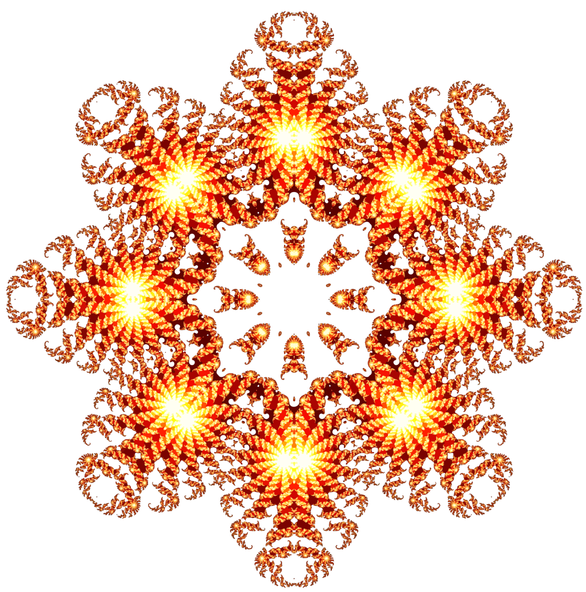
CAMPION COLLEGE
AT THE UNIVERSITY OF REGINA

LUTHER COLLEGE
UNIVERSITY OF REGINA

Math 111
Calculus II



by Robert G. Petry and Fotini Labropulu

published by Campion College and Luther College

Contents

* Sections denoted with an asterisk are optional and may be omitted in the course.

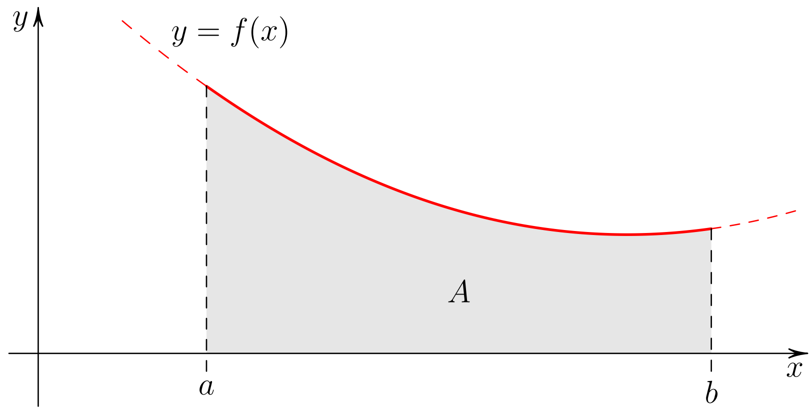
Unit 1: Integration Review

1.1 The Meaning of the Definite Integral

The **definite integral** of the function $f(x)$ between $x = a$ and $x = b$ is written:

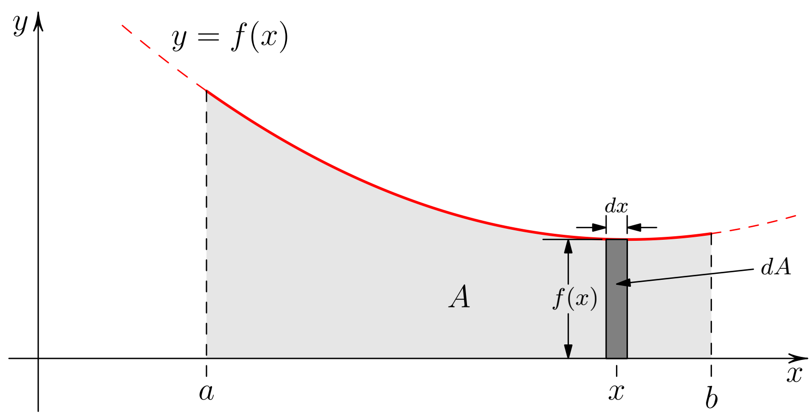
$$\int_a^b f(x)\,dx$$

Geometrically it equals the **area** A between the curve $y = f(x)$ and the x-axis between the vertical lines $x = a$ and $x = b$:



More precisely, assuming $a < b$, the definite integral is the net sum of the **signed** areas between the curve $y = f(x)$ and the x-axis where areas below the x-axis (i.e. where $f(x)$ dips below the x-axis) are counted **negatively**.

The notation used for the definite integral, $\int_a^b f(x)\,dx$, is elegant and intuitive. We are $\int$umming ($\int dA$) the (infinitesimally) small differential rectangular areas $dA = f(x) \cdot dx$ of height $f(x)$ and width dx at each value x between $x = a$ and $x = b$:



We will see soon how viewing integrals as sums of differentials can be used to come up with formulas for calculations aside from just area.

1.2 The Fundamental Theorem of Calculus

As seen in a previous calculus course, the definite integral can be written as a limiting sum (Riemann Sum) of N rectangles of finite width $\Delta x = (b-a)/N$ where we let the number of rectangles (N) go to infinity (and consequently the width $\Delta x \to 0$). This method of evaluating a definite integral is hard or impossible to compute exactly yfor most functions. An easy way to evaluate definite integrals is due to the **Fundamental Theorem of Calculus** which relates the calculation of a definite integral with the evaluation of the **antiderivative** $F(x)$ of $f(x)$:

Theorem: 1.1. The Fundamental Theorem of Calculus:

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

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

for any F an antiderivative of f, i.e. $F'(x) = f(x)$.

Notationally we write $F(b) - F(a)$ with the shorthand $F(x)\big|_a^b$, i.e.

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

where, unlike the integral sign, the bar is placed *on the right*.

1.3 Indefinite Integrals

Because of the intimate relationship between the antiderivative and the definite integral, we define the **indefinite integral** of $f(x)$ (with no limits a or b) to just be the antiderivative, i.e.

$$\int f(x)\ dx = F(x) + C$$

where $F(x)$ is an antiderivate of $f(x)$ (so $F'(x) = f(x)$) and C is an arbitrary constant. The latter is required since the antiderivative of a function is not unique as $\frac{d}{dx}C = 0$ implies we can always add a constant to an antiderivative to get another antiderivative of the same function.

Using our notation for indefinite integrals and our knowledge of derivatives gives the following.

Table of Indefinite Integrals

1. $\displaystyle\int x^n\, dx = \frac{1}{n+1}x^{n+1} + C \qquad (n \neq -1)$
2. $\displaystyle\int \cos x\, dx = \sin x + C$
3. $\displaystyle\int \sin x\, dx = -\cos x + C$
4. $\displaystyle\int \sec^2 x\, dx = \tan x + C$
5. $\displaystyle\int \sec x \tan x\, dx = \sec x + C$
6. $\displaystyle\int \csc^2 x\, dx = -\cot x + C$
7. $\displaystyle\int \csc x \cot x\, dx = -\csc x + C$
8. $\displaystyle\int cf(x)\, dx = c\int f(x)\, dx$
9. $\displaystyle\int [f(x) \pm g(x)]\ dx = \int f(x)\, dx \pm \int g(x)\, dx$

In the last two integration formulae $f(x)$ and $g(x)$ are functions while c is a constant. For indefinite integrals we say, for example, that $\frac{1}{n+1}x^{n+1} + C$ is the **(indefinite) integral** of x^n where x^n is the **integrand**. The **process** of finding the integral is called **integration**. Each of these indefinite integrals may be verified by differentiating the right hand side and verifying that the integrand is the result.

1.4 Integration by Substitution

The last two general integral results allow us to break up an integral of sums or differences into integrals of the individual pieces and to pull out any constant multipliers. Another useful way of solving an integral is to use the **Substitution Rule** which arises by working the differentiation Chain Rule in reverse.

Theorem: 1.2. Substitution Rule (Indefinite Integrals): Suppose $u = g(x)$ is a differentiable function whose range of values is an interval I upon which a further function f is continuous, then

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

where the right hand integral is to be evaluated at $u = g(x)$ after integration.

Here the du appearing on the right side is the differential:

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

which, recall, can be remembered by thinking $\frac{du}{dx} = g'(x)$ and multiplying both sides by dx.

When using the Substitution Rule with definite integrals we can avoid the final back-substitution of $u = g(x)$ of the indefinite case by instead just changing the limits of the integral appropriately to the u-values corresponding to the x-limits:

Theorem: 1.3. Substitution Rule (Definite Integrals): Suppose $u = g(x)$ is a differentiable function whose derivative g' is continuous on $[a, b]$ and a further function f is continuous on the range of $u = g(x)$ (evaluated on $[a, b]$), then

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

1.5 Integration Examples

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \left(t^2 + \sqrt{t} - \frac{2}{t^2}\right)\, dt$
2. $\displaystyle\int_0^1 \left(t^2 + 1\right)^2\, dt$
3. $\displaystyle\int x^2 \left(x^3 + 2\right)^{\frac{1}{3}}\, dx$
4. $\displaystyle\int_0^{\frac{\pi}{4}} (\sec x - \tan x) \sec x\, dx$
5. $\displaystyle\int \frac{\cos\sqrt{x}}{\sqrt{x}}\, dx$
6. $\displaystyle\int_1^2 x\sqrt{x-1}\, dx$
7. $\displaystyle\int \sin(5\theta)\, d\theta$
8. $\displaystyle\int_2^3 \frac{3x^2 - 1}{\left(x^3 - x\right)^2}\, dx$
9. $\displaystyle\int t^2 \sin\left(1 - t^3\right) dt$
10. $\displaystyle\int \frac{x - \sqrt{3x}}{\sqrt{2x}}\, dx$
11. $\displaystyle\int_0^4 (4x + 9)^{\frac{3}{2}}\, dx$
12. $\displaystyle\int (\cos\theta + \sin\theta)(\cos\theta - \sin\theta)^4\, d\theta$

Unit 2: Inverses and Other Functions

2.1 Inverse Functions

Example:
The inverse function of the function $f(x) = x^3$ is $g(x) = x^{\frac{1}{3}}$.

Intuitively $g(x) = x^{\frac{1}{3}}$ is the inverse of is $f(x) = x^3$ because g undoes the action of f. So if f acts on the value 2 so $f(2) = 2^3 = 8$ and we act g on the result, $g(8) = 8^{\frac{1}{3}} = 2$ we are returned to the original value.

One may wonder whether all functions have inverses. The answer is no. A necessary and sufficient condition for a function to have an inverse is that the function be *one-to-one*.

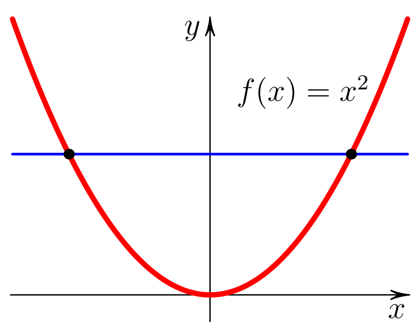
Definition: A function f with domain A and range B is said to be **one-to-one** if whenever $f(x_1) = f(x_2)$ (in B) one has that $x_1 = x_2$ (in A).

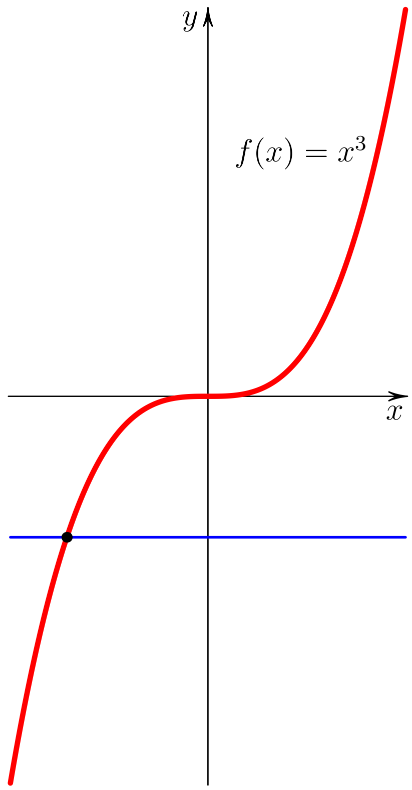
A logically equivalent condition is that if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$. In words, no two elements in the domain A have the same image in the range B.

2.1.1 Horizontal Line Test

The **Horizontal Line Test** says that a function $f(x)$ will be one-to-one if and only if every horizontal line intersects the graph of $y = f(x)$ at most once.

Example:
The horizontal line test shows that the function $y = x^2$ is not one-to-one while $y = x^3$ is one-to-one.





The following theorem is intuitively true when one considers the Horizontal Line Test.

Theorem: 2.1. Suppose a function f has a domain D consisting of an interval. If the function is increasing everywhere or decreasing everywhere on D then f is one-to-one.

Examples:
Determine whether the given functions are one-to-one:

1. $f(x) = 2x^3 + 5$
2. $f(x) = \dfrac{3-x}{x+1}$

Definition: Suppose f is a one-to-one function defined on domain A with range B. The **inverse function of $\boldsymbol{f}$** denoted by $\boldsymbol{f^{-1}}$ is defined on domain B with range A and satisfies

$$f^{-1}(y) = x \iff f(x) = y$$

for any y in B.

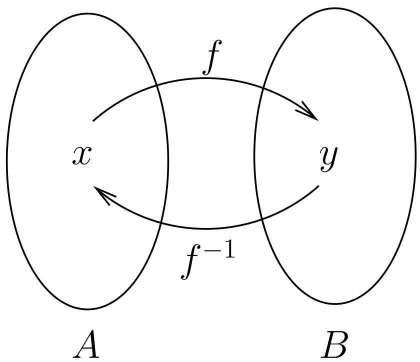
Here the symbol $\iff$ means "if and only if". *This* if and only if *that*" itself means that both the following hold

- "If *this* then *that*.
- "If *that* then *this*.

Also observe that the notation f^{-1} makes it clear that f is the function for which this is the inverse.

Notes:

1. $f^{-1} \neq \dfrac{1}{f}$ We call $\frac{1}{f}$ the *reciprocal* of f.
2. The definition says that if f maps x to y, then f^{-1} maps y back to x.



3. The domain of f^{-1} is the range of f while the range of f^{-1} is the domain of f.
4. Reversing the roles of x and y gives

$$f^{-1}(x) = y \iff f(y) = x$$

or equivalently

$$f(y) = x \iff f^{-1}(x) = y$$

This implies that f itself is the inverse function of f^{-1}.

5. The following hold (see last diagram)

$$\begin{aligned} f^{-1}(f(x)) &= x \quad \text{for every } x \text{ in } A \\ f\left(f^{-1}(y)\right) &= y \quad \text{for every } y \text{ in } B \end{aligned}$$

The first relationship highlights the utility of the inverse function in *solving equations* for if we have, say, $f(x) = 3$ for some one-to-one function f for which we know the inverse $f^{-1}(x)$, it follows, applying f^{-1} to both sides that $f^{-1}(f(x)) = x = f^{-1}(3)$. We are applying inverse functions all the time when we isolate variables in equations.

This explains why "cube-rooting both sides" of $x^3 = 64$ is a safe way to find the solution to this equation while "square-rooting both sides" of $x^2 = 64$ is not. The latter finds only one of the two solutions. (Applying a *function* in this way can only produce one number.)

2.1.2 Finding Inverse Functions

To find the inverse function of a one-to-one function f proceed with the following steps:

1. Write $y = f(x)$
2. Solve the equation for x in terms of y (if possible).
3. Interchange the roles of x and y. The resulting equation is $y = f^{-1}(x)$.

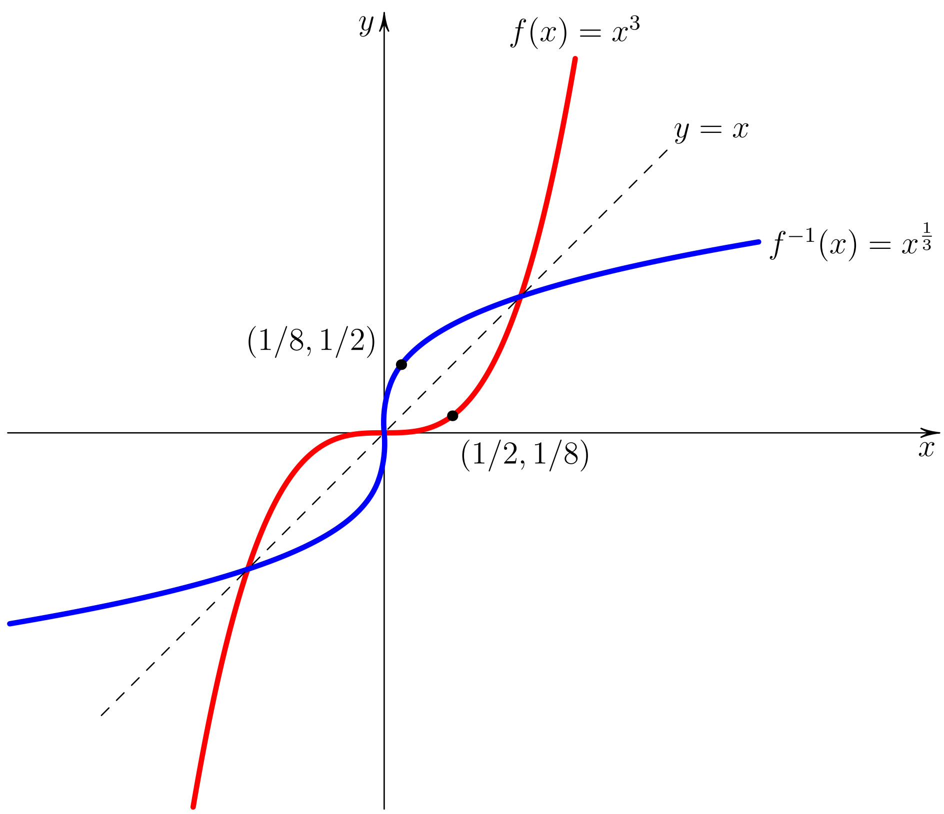
Examples:
Find the inverse function of the given function:

1. $f(x) = 2x^3 + 5$
2. $f(x) = \dfrac{3-x}{x+1}$
3. $f(x) = x^2 - 9$

Note that if you are able to solve your expression uniquely for x in terms of y in the second step it follows that the function is one-to-one since, given any y value in the range B there can only be a single value x in A which maps to it, namely the value which results from evaluating your solved expression with y.

2.1.3 Graphs of Inverse Functions

The definition of the inverse function implies that if (x, y) lies on the graph of $y = f(x)$ then (y, x) will lie on the graph of f^{-1}. Geometrically this means that the graph of f^{-1} may be obtained by reflecting the graph of f about the line $y = x$.



Graphically a discontinuity in f would imply a discontinuity in f^{-1} and vice versa. We have the following theorem.

Theorem: 2.2. Suppose f is a one-to-one continuous function defined on an interval then its inverse f^{-1} is also continuous.

2.1.4 Derivative of an Inverse Function

If we let $g(x)$ be the inverse of f then our earlier relationship $x = f\left(f^{-1}(x)\right) = f(g(x))$. Differentiating the left side with respect to x just gives 1. Differentiating the right side of the equation with respect to x can be done with the Chain Rule. Solving for the derivative $g'(a)$ gives the following result.

Theorem: 2.3. Suppose f is a one-to-one differentiable function with inverse $g = f^{-1}$. If $f'(g(a)) \neq 0$ then the inverse function is differentiable at a with

$$g'(a) = \frac{1}{f'(g(a))}$$

More generally the derivative of the inverse function is

$$g'(x) = \frac{1}{f'(g(x))} \ .$$

Example:
For the function $f(x) = \dfrac{1}{x-1}$:

1. Show that f is one-to-one.
2. Calculate $g = f^{-1}$ and find its domain and range.
3. Calculate $g'(2)$ using your result from part 2.
4. Find $g'(2)$ from the formula $g'(x) = \dfrac{1}{f'(g(x))}$.

Examples:
Find the following derivatives:

1. $\left(f^{-1}\right)'(1)$ if $f(x) = x^3 + x + 1$.
2. $g'(-1)$ if $f(x) = 3x - \cos x$ and $g = f^{-1}$.

2.1.5 Creating Invertible Functions

So far one-to-one (and hence invertible) functions seem uncommon. However this is only because we only considered functions defined on their natural domains, i.e. the set of numbers for which the function may be evaluated. We can choose to define a function with a smaller domain and by suitable restriction we can create a function that is one-to-one and hence invertible.

Example:
Define the function $f(x)$ to have the value $f(x) = x^2$ but only be defined on the domain $A = [0, \infty)$. Since f is increasing everywhere on this interval it is one-to-one and hence has an inverse, $f^{-1}(x) = x^{\frac{1}{2}} = \sqrt{x}$. If we restricted the domain to be $A = (-\infty, 0]$ the inverse would be $f^{-1}(x) = -\sqrt{|x|}$!

2.2 Exponential Functions

If we write the number 2^5, then this, recall, means $2 \times 2 \times 2 \times 2 \times 2$. We call 2 the *base* and 5 the *exponent*. We have already seen that one way to create a function is to replace the base with a variable. This produces *power functions* like

$$f(x) = x^2 \qquad y = x^{\frac{1}{2}} = \sqrt{x} \qquad y = x^{-1} = \tfrac{1}{x}$$

In general, a power function is of the form $y = x^r$ where r is any real constant.

If, on the other hand, we let the exponent be a variable and the base a constant, like:

$$f(x) = 2^x, \ y = (1/2)^x$$

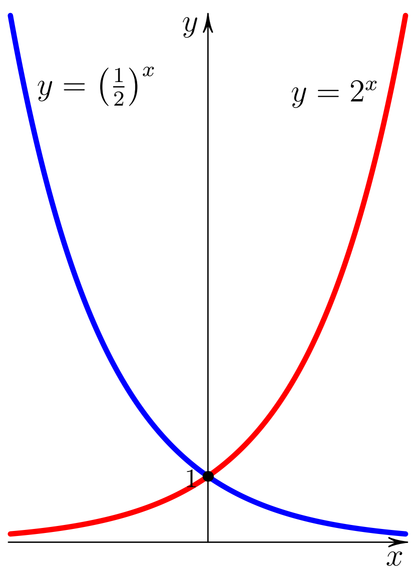
we get *exponential functions*.

Definition: Let $a > 0$. The function

$$f(x) = a^x$$

is an exponential function.

The graph of an exponential function has the following form depending on whether a is greater than or less than 1. Two typical values of a are shown.

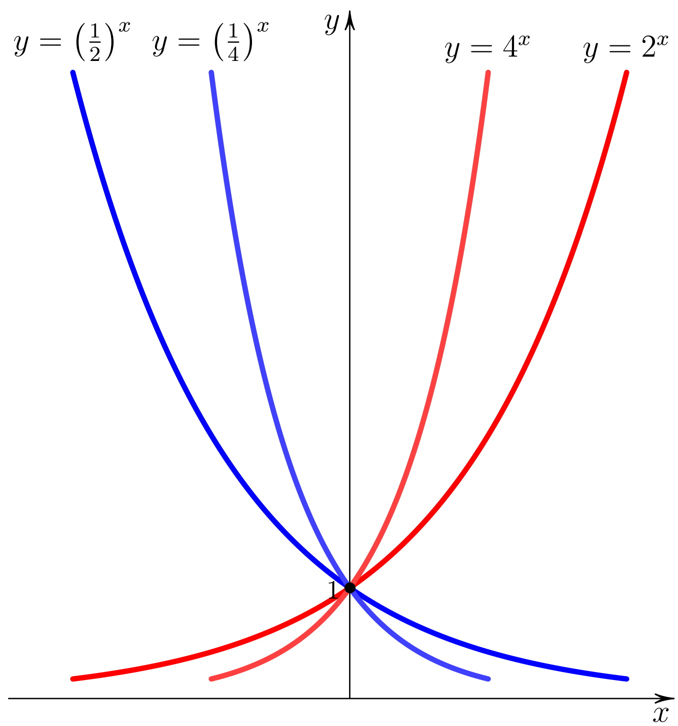


Notes:

1. If $x = 0$ then $a^x = a^0 = 1$. Therefore all exponential functions go through the point $(0, 1)$.
2. If $x = n$, a positive integer then $a^x = a^n = \underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ times}}$.
3. If $x = -n$, n a positive integer, then $a^x = a^{-n} = \frac{1}{a^n}$.
4. If $x = \frac{1}{n}$, n a positive integer, then $a^x = a^{\frac{1}{n}} = \sqrt[n]{a}$. (Hence $a < 0$ is excluded.)
5. If x is rational, $x = \frac{p}{q}$, then $a^x = a^{\frac{p}{q}} = (a^p)^{\frac{1}{q}} = \sqrt[q]{a^p}$.

6. If $a \neq 1$ (and $a > 0$) then $f(x) = a^x$ is a continuous function with domain $\mathbb{R}$ and range $(0, \infty)$.
7. If $0 < a < 1$ then $f(x) = a^x$ is a decreasing function.
8. If $a > 1$ then $f(x) = a^x$ is an increasing function.
9. If $a, b > 0$ and $x, y \in \mathbb{R}$ then (a) $a^x a^y = a^{x+y}$ (b) $\frac{a^x}{a^y} = a^{x-y}$ (c) $(a^x)^y = a^{xy}$ (d) $(ab)^x = a^x b^x$. These relations are readily apparent when one considers x and y as positive integers.

For two bases greater than one the base which is larger is the steeper curve while for two bases less than one the base which is smaller is steeper.



As depicted in the previous graphs, we have the following limits:

Theorem: 2.4. For exponential functions we have the following limits at infinity:

If $a > 1$, then $\lim\limits_{x \to -\infty} a^x = 0$ and $\lim\limits_{x \to \infty} a^x = \infty$.
If $0 < a < 1$, then $\lim\limits_{x \to -\infty} a^x = \infty$ and $\lim\limits_{x \to \infty} a^x = 0$.

So the x-axis is a horizontal asymptote for a^x provided $a > 0$, $a \neq 1$.

2.2.1 The Natural Exponential Function

Consider the derivative of $f(x) = a^x$:

$$f'(x) = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} = \lim_{h \to 0} \frac{a^x a^h - a^x}{h} = \left(\lim_{h \to 0} \frac{a^h - 1}{h} \right) a^x$$

where here we are able to pull out the a^x from the limit because a^x does not involve the limit variable h. The result shows the derivative is proportional to the function $f(x) = a^x$ itself with constant of proportionality c given by the evaluation of the limit:

$$c = \lim_{h \to 0} \frac{a^h - 1}{h}$$

Due to the presence of the constant a in the limit, one anticipates correctly that the constant c depends on the choice of base a. Interestingly, one can ask the question if there is some choice of base a for which the constant is $c = 1$. The answer is yes, the base is given by *Euler's Number*:

$$e = 2.71828\ldots$$

for which we have that $c = 1$ in the above limit:

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1 \ .$$

More constructively, as opposed to e being the solution of such a limit equation, it will be shown that e may be written as the following limit:

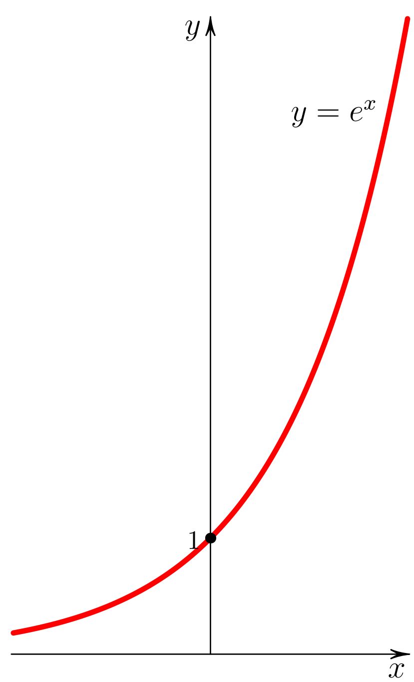
$$e = \lim_{h \to 0} (1 + h)^{\frac{1}{h}} \ ,$$

or, setting $h = 1/n$,

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n \ .$$

Definition: If $a = e = 2.71828\ldots$, then $f(x) = e^x$ is the *natural exponential function.*

Since $e = 2.71\ldots > 1$ the natural exponential function shares all the aforementioned properties of $f(x) = a^x$ where $a > 1$. (i.e. continuous, increasing function with domain $\mathbb{R}$, range $(0, \infty)$, limits, etc.)



You should identify the natural exponential key $\boxed{e^x}$ on your calculator.

2.2.2 Derivative of e^x

Furthermore from the preceding discussion we have the important result:

Theorem: 2.5. The derivative of the natural exponential function is:

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

Proof is, as above,

$$\frac{d}{dx}e^x = \lim_{h\to 0}\frac{e^{x+h}-e^x}{h} = \lim_{h\to 0}\frac{e^x e^h - e^x}{h} = \left(\lim_{h\to 0}\frac{e^h-1}{h}\right)e^x = (1)e^x = e^x \ .$$

A corollary of this theorem, applying the Chain Rule to the function e^u with $u = g(x)$ is:

Theorem: 2.6. $\quad \dfrac{d}{dx}e^u = e^u\dfrac{du}{dx} \quad$ or $\quad \dfrac{d}{dx}\left[e^{g(x)}\right] = e^{g(x)}g'(x)$

Examples:
Find $\frac{dy}{dx}$ for the following:

1. $y = e^{2x}$
2. $y = e^{x^3+x}$
3. $y = e^{\sec x} + \sec\left(e^x\right)$
4. $y = e^{x^4}\sin\left(x^2+1\right)$
5. $xy + e^x = 2xy^2$

2.2.3 Integral of e^x

Since $\frac{d}{dx}e^x = e^x$, we have the following:

Theorem: 2.7. The indefinite integral of e^x is

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

Examples:
Evaluate the following integrals:

1. $\displaystyle\int e^{2x}\,dx$
2. $\displaystyle\int x^3e^{x^4+1}\,dx$
3. $\displaystyle\int_0^1 e^{-x}\,dx$
4. $\displaystyle\int \frac{2e^x}{\left(1+e^x\right)^2}\,dx$
5. $\displaystyle\int \frac{2+3e^x}{e^x}\,dx$
6. $\displaystyle\int \frac{1-4e^{3x}}{e}\,dx$

2.2.4 Simplifying Exponential Expressions

Using the rules of exponents we are often able to consolidate expressions involving several exponents into an expression involving one exponent.

Example:
The expression $\dfrac{e^2\sqrt{e^x}}{(2e^x)^3}$ may be simplified as follows:

$$\begin{aligned}
\frac{e^2\sqrt{e^x}}{(2e^x)^3} &= \frac{e^2\left(e^x\right)^{\frac{1}{2}}}{2^3\left(e^x\right)^3} \quad \left(\text{since } \sqrt[n]{a} = a^{\frac{1}{n}},\ (ab)^x = a^x b^x \right) \\
&= \frac{e^2 e^{\frac{1}{2}x}}{2^3 e^{3x}} \quad (\text{ since } (a^x)^y = a^{xy}) \\
&= \frac{e^{2+\frac{1}{2}x}}{8e^{3x}} \quad (\text{ since } a^x a^y = a^{x+y}) \\
&= \frac{1}{8} e^{2+\frac{1}{2}x-3x} \quad \left(\text{since } \frac{a^x}{a^y} = a^{x-y} \right) \\
&= \frac{1}{8} e^{2-\frac{5}{2}x}
\end{aligned}$$

The usefulness in consolidating exponents in this manner is clear when solving equations.

Example:
Solving the equation

$$\frac{e^2\sqrt{e^x}}{(2e^x)^3} = \frac{1}{2}$$

Is equivalent, by using our previous result and multiplying both sides by 8, to

$$e^{2-\frac{5}{2}x} = 4$$

Now if we could apply an *inverse* to the natural exponential function on both sides we could solve for x.

2.3 Logarithmic Functions

We finished the last section by suggesting that inverses of exponential functions would be useful for, among other things, solving equations involving exponentials. Since the exponential function $f(x) = a^x$ with constant $a > 0$ and $a \neq 1$ is either everywhere decreasing ($0 < a < 1$) or increasing ($1 < a$) on open interval $\mathbb{R} = (-\infty, \infty)$, the exponential function is one-to-one and hence has an inverse function f^{-1}.

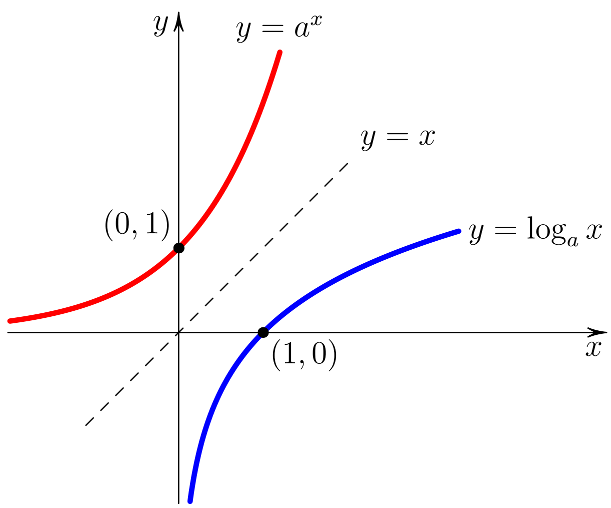
Definition: Given constant $a > 0$, $a \neq 0$, the **logarithmic function of base *a***, written $\log_a x$ is defined by

$$\log_a x = y \iff a^y = x$$

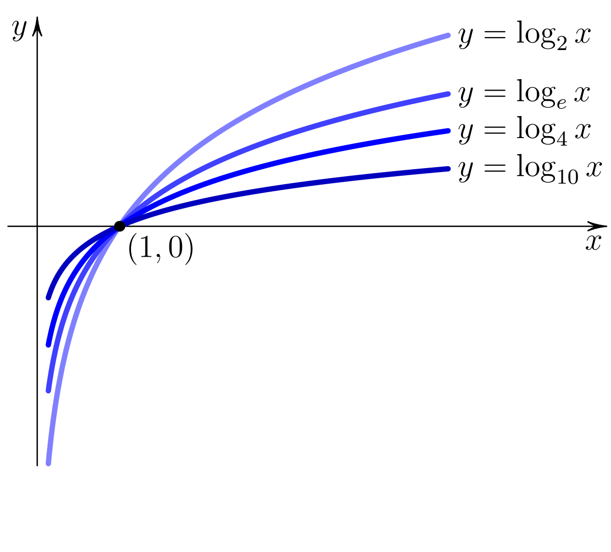
That is, it is the inverse of the exponential function $f(x) = a^x$.

In words, the logarithm of a value x to a base a is the *exponent* to which you must take a to get x.

For the case $\boxed{a > 1}$, which, as you recall, is the case for $a = e = 2.71\ldots$, a representative graph of $y = a^x$ and its inverse $y = \log_a x$ are as follows:



For base $a > 1$ we saw that larger values of a led to steeper $y = a^x$ curves, it follows that larger values of a will make the logarithmic curves more horizontal in this case:



2.3.1 Logarithmic Function Properties

Because of their relationship to exponentials as inverses the following are true for logarithmic functions:

1. $y = \log_a x$ has domain $(0, \infty)$ and range $\mathbb{R}$.
2. $y = \log_a x$ is continuous on its domain.
3. $y = f(x) = \log_a x$ is one-to-one with inverse function $f^{-1}(x) = a^x$.
4. $\log_a(1) = 0$
5. The following limits hold (see graph for $a > 1$ case):
   - If $0 < a < 1$ then $f(x) = \log_a(x)$ is a decreasing function with
     $$\lim_{x\to 0^+} \log_a x = +\infty \qquad \lim_{x\to\infty} \log_a x = -\infty$$
   - If $a > 1$, then $f(x) = \log_a(x)$ is an increasing function with
     $$\lim_{x\to 0^+} \log_a x = -\infty \qquad \lim_{x\to\infty} \log_a x = \infty$$

   Note the y-axis is a vertical asymptote in either case.
6. The following inverse relations hold:
   $$\begin{aligned} \log_a (a^x) &= x \quad \text{for any } x \text{ in } \mathbb{R} \\ a^{\log_a x} &= x \quad \text{for any } x > 0 \end{aligned}$$

The special multiplication, division, and power laws of exponents induce the following important logarithmic results.

Theorem: 2.8. For $x > 0$ and $y > 0$ and any real number r the following hold:

1. $\log_a(xy) = \log_a x + \log_a y$
2. $\log_a \left(\frac{x}{y}\right) = \log_a x - \log_a y$
3. $\log_a (x^r) = r \log_a x$

To prove the theorem note that if $x > 0$ and $y > 0$ then $m = \log_a x$ and $n = \log_a y$ exist and, exponentiating both sides, it follows that $x = a^m$ and $y = a^n$. Evaluating the first equation's left hand side we have:
$$\log_a(xy) = \log_a (a^m a^n) = \log_a \left(a^{m+n}\right) = m + n = \log_a x + \log_a y$$
The other conclusions are similarly proven.

These results can be used to simplify logarithmic expressions:

Examples:
Simplify the following:

1. $\log_2 4 + \log_2 10 - \log_2 5$
2. $\log_5 3 + \log_5 3^4 + \log_5 1$

2.3.2 The Natural Logarithmic Function

Definition: The logarithmic function with base a equal to $e = 2.71\ldots$ is called the **natural logarithmic function** and is denoted by $\boldsymbol{\ln x}$. In symbols:

$$\ln x = \log_e x$$

All the properties for a logarithm with base $a > 1$ apply to the natural logarithm. In terms of the notation for natural logarithms and exponentials we have the definition:

$$\ln x = y \iff e^y = x$$

and the properties:

$$\begin{aligned}
\ln e &= 1 \\
\ln e^x &= x \quad (x \in \mathbb{R}) \\
e^{\ln x} &= x \quad (x > 0) \\
\ln(xy) &= \ln x + \ln y \quad (x, y > 0) \\
\ln\left(\frac{x}{y}\right) &= \ln x - \ln y \quad (x, y > 0) \\
\ln\left(x^r\right) &= r \ln x \quad (x > 0, r \in \mathbb{R})
\end{aligned}$$

Note the following:

$$\log_a(x + y) \neq \log_a x + \log_a y$$

$$\log_a(x - y) \neq \log_a x - \log_a y$$

In the specific case of natural logarithms ($a = e$):

$$\ln(x + y) \neq \ln x + \ln y$$

$$\ln(x - y) \neq \ln x - \ln y$$

You should identify the natural logarithm key $\boxed{\ln}$ on your calculator. Note that the key $\boxed{\log}$ on the calculator means base 10 logarithm $\log_{10} x$.[1]

Examples:
Simplify the following:

1. $\ln 5 + 2 \ln 3 + \ln 1$
2. $\dfrac{1}{2} \ln(4t) - \ln(t^2 + 1)$
3. $e^{\ln\left(x^2+1\right)} + 3x^2 - 5$

[1]However in other areas (some computer applications) the symbol log will often refer to a natural logarithm so one needs to be careful.

2.3.3 Solving Exponential and Logarithmic Equations

Solving equations involving logarithmic or exponential functions typically involves using properties of these functions to simplify those expressions involving the variable and then applying the appropriate inverse function to undo the exponential or logarithm. Finally one may solve for the variable.[2]

Example:
We saw that the equation

$$\frac{e^2\sqrt{e^x}}{(2e^x)^3} = \frac{1}{2}$$

could be written, using properties of exponentials, as

$$e^{2-\frac{5}{2}x} = 4\ .$$

Applying ln, the inverse of the exponential e^x, to both sides of the equation, gives

$$2 - \frac{5}{2}x = \ln 4\ .$$

Solving for x gives

$$x = \frac{2}{5}\left(2 - \ln 4\right)\ .$$

Examples:
Solve the following equations for x:

1. $5e^{x-3} = 4$
2. $\ln\left(x^2 - 3\right) = 0$
3. $4e^x e^{-2x} = 6$
4. $\ln(2\ln x - 5) = 0$
5. $e^{x^2-5x+6} = 1$
6. $3e^{2x-4} = 10$
7. $\ln\left(\dfrac{x-2}{x-1}\right) = 1 + \ln\left(\dfrac{x-3}{x-1}\right)$

The following relates logarithms in other bases to the natural logarithm.

Theorem: 2.9. For $a > 0$, $a \neq 1$ we have:

$$\log_a x = \frac{\ln x}{\ln a}$$

[2]More complicated equations may allow themselves to be written as a product of factors equal to zero:

$$(\text{factor}_1)(\text{factor}_2)\cdot\ldots\cdot(\text{factor}_n) = 0$$

where the factors themselves involve logarithms or exponentials. Note that a strictly exponential factor equalling zero will provide no solution as $a^x \neq 0$ for all x. A strictly logarithmic factor equalling zero will be equivalent to the argument of the logarithm equalling 1.

Proof comes from observing that since $a^{\log_a x} = x$ we can take the natural logarithm of both sides and then use the power rule for the natural logarithm to get

$$(\log_a x)(\ln a) = \ln x \ .$$

Solving for $\log_a x$ gives our result.

This theorem is useful for evaluating an arbitrary base a logarithm on a calculator.

Examples:
Write in terms of the natural logarithm (ln):

1. $\log_5 7$
2. $\log_{20}\left(x^2+1\right)$
3. $\log_{10}\left(e^{2x}\right)$

2.3.4 Derivative of the Natural Logarithmic Function

Theorem: 2.10. The derivative of the natural logarithmic function is

$$\frac{d}{dx}(\ln x) = \frac{1}{x}$$

To prove the theorem note that if $y = \ln x$ then by definition of the logarithm as an inverse we have

$$e^y = x \ .$$

Differentiating this implicit equation with respect to x on both sides gives

$$e^y y' = 1 \ ,$$

and so

$$y' = \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x} \ .$$

A corollary of this result, applying the Chain Rule to the function $\ln u$ with $u = g(x)$ is

Theorem: 2.11. $\frac{d}{dx}\ln u = \frac{1}{u}\frac{du}{dx}$ or $\frac{d}{dx}\ln\left[g(x)\right] = \frac{g'(x)}{g(x)}$

Examples:
Differentiate the following functions:

1. $y = \ln\left(x^2 - 3x + 1\right)$
2. $y = \ln\left(x + \ln x\right)$
3. $y = \ln\left(\frac{x+1}{\sqrt{x+2}}\right)$
4. $y = e^{(2+x\ln x)}$

2.3.5 Derivatives Using Arbitrary Bases

Theorem: 2.12. The derivative of the logarithm function to base $a > 0$ $(a \neq 1)$ is

$$\frac{d}{dx}\left(\log_a x\right) = \frac{1}{x \ln a} \qquad\qquad \frac{d}{dx}\left[\log_a g(x)\right] = \frac{g'(x)}{g(x)\ln a}$$

Proof of the former derivative follows from the identity $\log_a x = \frac{\ln x}{\ln a}$:

$$\frac{d}{dx}\left(\log_a x\right) = \frac{d}{dx}\left(\frac{\ln x}{\ln a}\right) = \frac{d}{dx}\left(\frac{1}{\ln a}\cdot \ln x\right) = \frac{1}{\ln a}\cdot\frac{d}{dx}\left(\ln x\right) = \frac{1}{\ln a}\cdot\frac{1}{x} = \frac{1}{x\ln a}$$

Here note that we used that $\frac{1}{\ln a}$ is constant since a is constant. The latter derivative in the theorem follows from the Chain Rule applied to this former result.

Theorem: 2.13. The derivative of an exponential function with base $a > 0$, $a \neq 1$ is

$$\frac{d}{dx}a^x = a^x \ln a \qquad\qquad \frac{d}{dx}\left[a^{g(x)}\right] = a^{g(x)}g'(x)\ln a$$

Proof of the former derivative follows by the observation that by our inverse identies the base a may be written $a = e^{\ln a}$ and using the Chain Rule:

$$\frac{d}{dx}\left(a^x\right) = \frac{d}{dx}\left(e^{\ln a}\right)^x = \frac{d}{dx}e^{x\ln a} = e^{x\ln a}\frac{d}{dx}\left(x\ln a\right) = e^{x\ln a}\left(\ln a\right) = \left(e^{\ln a}\right)^x\left(\ln a\right) = a^x\ln a$$

Once again the latter derivative given in the theorem is just the result arising from using the Chain Rule with the former result.

Examples:
Differentiate the following functions:

1. $y = \log_{10}\left(3x^2 + e^x\right)$
2. $y = 5^{2e^x + 3x}$
3. $y = a^{3x}\log_4 x$
4. $y = 4^{\cos x}$

2.3.6 Logarithmic Differentiation

Using the properties of logarithms makes taking derivatives of logarithms of products, quotients, and powers easy.

Example:
To differentiate $y = \ln[x(x^2+1)(x-3)]$ is easily done if we expand the logarithm first and then differentiate:

$$\begin{aligned}\frac{dy}{dx} &= \frac{d}{dx}\ln[x(x^2+1)(x-3)] \\ &= \frac{d}{dx}\left[\ln x + \ln(x^2+1) + \ln(x-3)\right] \\ &= \frac{1}{x} + \frac{2x}{x^2+1} + \frac{1}{x-3}\end{aligned}$$

Wouldn't it be nice if when working with products, etc., we were always differentiating their logarithm? In *logarithmic differentiation* we take the *logarithm of both sides* of an equation before differentiating.

Example:
To differentiate $y = x\left(2x^3+1\right)^3 (x+5)^{\frac{1}{2}} \left(x^2+3x-1\right)^{\frac{1}{3}}$ one could use the (generalized) Product Rule. Instead, try taking the logarithm of both sides of the equation to get:

$$\ln y = \ln x + 3\ln\left(2x^3+1\right) + \frac{1}{2}\ln(x+5) + \frac{1}{3}\ln\left(x^2+3x-1\right)$$

Next differentiate both sides of the equation with respect to x to get:

$$\frac{1}{y}y' = \frac{1}{x} + 3\frac{6x^2}{2x^3+1} + \frac{1}{2}\frac{1}{x+5} + \frac{1}{3}\frac{2x+3}{x^2+3x-1}$$

Multiplying both sides by y and substituting in its value gives our derivative:

$$y' = \left[\frac{1}{x} + \frac{18x^2}{2x^3+1} + \frac{1}{2(x+5)} + \frac{2x+3}{3\left(x^2+3x-1\right)}\right] x\left(2x^3+1\right)^3 (x+5)^{\frac{1}{2}} \left(x^2+3x-1\right)^{\frac{1}{3}}$$

Note that implicit differentiation is used to differentiate the $\ln y$ that shows up on the left hand side of the equation with respect to x. This gives the $\frac{1}{y}y'$ which is why we need to multiply both sides by y (for which we have the function).

Steps in Logarithmic Differentiation

1. Take logarithms of both sides of the equation $y = f(x)$.
2. Differentiate with respect to x on both sides, remembering to use implicit differentiation on $\ln y$ to get $\frac{1}{y}y'$.
3. Solve for y' and substitute $f(x)$ for y.

Examples:
Find the following derivatives using logarithmic differentiation:

1. $y = \dfrac{\left(x^3+5\right)\left(x^2-3x\right)^4}{x-2}$
2. $y = \left(x^2+3\right)^{x^3}$
3. $f(x) = (e^x+1)^{\ln x}$
4. $y = (2x+1)^{\sqrt{x}}$

2.3.7 Integral of $\frac{1}{x}$ and a^x

The Power Rule for integration is

$$\int x^n\,dx = \frac{1}{n+1}x^{n+1} \quad (n \neq -1)$$

The answer for the indefinite integral clearly indicates that it cannot work for $n = -1$ as one would be dividing by zero. Since $\frac{d}{dx}\ln x = \frac{1}{x}$ however, we now have an antiderivative for $x^{-1} = \frac{1}{x}$, namely $\ln x$. This will only work for values of $x > 0$ since the domain of $\ln x$ is only positive numbers. However a second antiderivative of $\frac{1}{x}$ that will work when $x < 0$ is $\ln(-x)$, since $\frac{d}{dx}\ln(-x) = \frac{1}{-x}\cdot(-1) = \frac{1}{x}$ by the Chain Rule. We can combine the results using absolute value bars in the following theorem.

Theorem: 2.14. The indefinite integral of x^{-1} is

$$\int \frac{1}{x}\,dx = \ln|x| + C$$

A useful corollary of this result is that one can now integrate the tangent function. Using the substitution $u = \cos x$ (so $du = -\sin x\,dx$) one has

$$\int \tan x\,dx = \int \frac{\sin x}{\cos x}\,dx = -\int \frac{du}{u} = -\ln|u|+C = -\ln|\cos x|+C = \ln\left(|\cos x|^{-1}\right)+C = \ln|\sec x|+C$$

Since $\frac{d}{dx}a^x = a^x\ln a$ it follows that $\frac{d}{dx}\frac{a^x}{\ln a} = a^x$ and so we also have the result:

Theorem: 2.15. $\displaystyle\int a^x\,dx = \frac{a^x}{\ln a} + C$

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \frac{3}{2x}\,dx$
2. $\displaystyle\int \frac{x^2}{x^3+5}\,dx$
3. $\displaystyle\int \frac{\ln x}{x}\,dx$
4. $\displaystyle\int_1^4 \frac{10^{\sqrt{x}}}{\sqrt{x}}\,dx$
5. $\displaystyle\int e^x 5^{e^x}\,dx$
6. $\displaystyle\int \frac{\sec x\tan x}{3+5\sec x}\,dx$
7. $\displaystyle\int \cot x\,dx$

2.4 Exponential Growth and Decay

Many quantities, such as the number of cells being cultured in a lab dish or the number of radioactive nucleii of a particular isotope in a radioactive sample remaining undecayed, have a population $y(t)$ that satisfies the differential equation

$$\boxed{\frac{dy}{dt} = ky}\,.$$

Here the constant k, called the **relative growth rate**, characterizes the population under consideration. It will be positive ($k > 0$) if the population $y(t)$ is increasing in time and negative ($k < 0$) if it is decreasing. The preceding equation is called the **law of natural growth** or **law of natural decay** respectively. The constant k is called **relative** since if we solve for it, $k = \frac{1}{y}\frac{dy}{dt}$, we see the rate dy/dt is constant only relative to the population size y at an given time.

The differential $dy = \frac{dy}{dt}dt$ satisfies

$$dy = ky\,dt\,.$$

Over a fixed time interval Δt we have the analogous relation

$$\Delta y = ky\Delta t\,,$$

where y, by the Mean Value Theorem, is evaluated at some time t in the interval. Assuming Δt is small enough this can be effectively any time t in the interval as y will be approximately constant over such an interval. The change Δy in the population $y(t)$ over a fixed small time interval Δt is therefore proportional to the population itself

$$\Delta y \propto y\,.$$

which is expected for a population whose growth (or loss) depends on the current size of the population. Additionally the relation shows the change will also be approximately proportional to the length of the time interval Δt considered,

$$\Delta y \propto \Delta t\,,$$

assuming again that Δt is sufficiently small, a result that is also reasonable.

To understand how y changes in time we need to find the function $y(t)$ that satisfies (solves) the differential equation

$$\frac{dy}{dt} = ky\,.$$

If the right hand side of the equation just involved t explicitly, like $\frac{dy}{dt} = t^2$, the answer would just be the antiderivative $y(t) = \int t^2 = \frac{1}{3}t^3 + C$. Our differential equation is not of this form, however, as it has the dependent variable y on the right hand side. Solving such a differential equation such as ours can be done by the process of **separation of variables**. Inspired by the Leibniz notation, one formally proceeds by isolating, if possible, a function of the dependent variable y and its differential dy on one side of the equation and a function of the independent variable t and its differential dt on the other to get

$$\frac{dy}{y} = k\,dt\,.$$

One then integrates both sides:

$$\int \frac{dy}{y} = \int k\,dt$$
$$\Rightarrow\; \ln y = kt + D\,,$$

where we have combined the integration constants C_1 and C_2 arising from both sides of the integral into $D = C_2 - C_1$. Finally we can solve for y by taking the natural exponential of both sides:

$$\begin{aligned} e^{\ln y} &= e^{kt+D} \\ \Rightarrow\ y &= e^{kt}e^{D} \end{aligned}$$

Calling a new (positive) constant $C = e^D$ we have the final solution of the differential equation

$$y(t) = Ce^{kt}\,.$$

Despite the lack of rigour in our separation of variable approach, one may readily confirm that $y(t) = Ce^{kt}$ does satisfy the original differential equation as required.

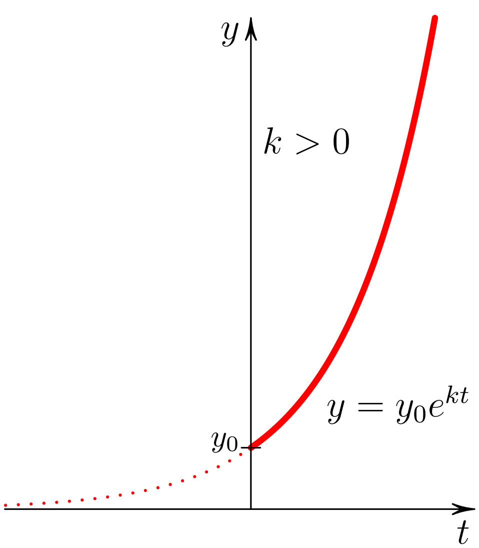
The constant of integration, C can be determined by providing an additional piece of information regarding the system. If, for instance, one knows the initial size of the population is $y(0) = y_0$, then the solution of the resulting initial value problem gives

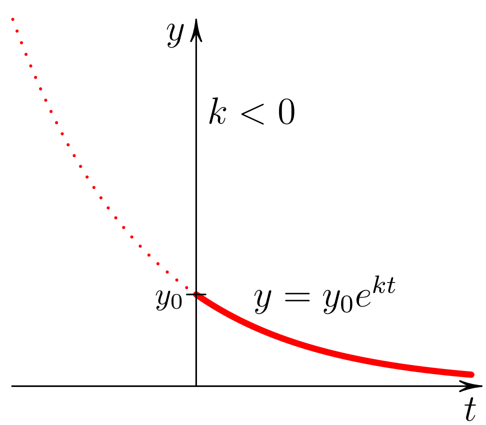
$$y_0 = y(0) = Ce^{k(0)} = Ce^0 = C(1) = C\,.$$

Placing this value for $C = y_0$ back in $y(t)$ gives

$$\boxed{y(t) = y_0 e^{kt}}\,.$$

As such the population at arbitrary time, assuming it is undergoing exponential growth or decay, is characterized completely by the growth constant k and its initial size y_0. The graphs of the cases where $k > 0$ (growth) and $k < 0$ (decay) are shown below.





Example:
Fox squirrels introduced into a city see their population increase from 50 to 12000 in 4 years. Assuming the growth was exponential over this time period,

1. Find the relative growth rate k.
2. When will the squirrel population exceed 1 million?
3. Is the latter likely? Explain.

When $k < 0$ we have a decay formula and the amount y decreases over time. Then the positive constant $\lambda = |k| = -k$, called the **decay constant**, may be introduced and our formula becomes

$$y(t) = y_0 e^{-\lambda t} .$$

Rather than the decay constant, one often uses the **half-life** constant T for a radioactive sample. It is defined to be the time required for half of the initial decaying substance to disappear (i.e. decay into a new form), and so $y(T) = \frac{1}{2}y_0$. This can be used to determine the decay constant λ.

Example:
Cobalt-60 is a radioactive isotope used in early radiotherapy and other applications. Sixty is the **mass number** of the nucleus, the number of nucleons (protons and neutrons) it contains. A sample of Cobalt-60 undergoes exponential decay with a half-life of 5.2714 years.

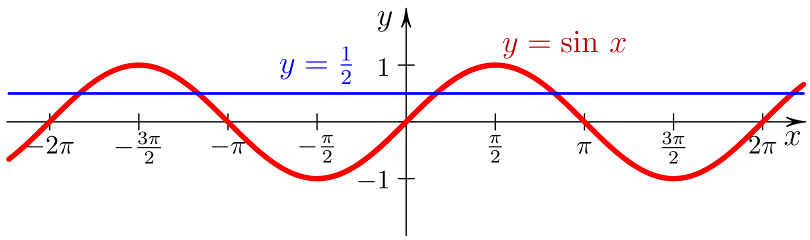
1. Find the decay constant $\lambda = -k$ of Cobalt-60.
2. How long would it take for a sample containing 40 grams of the isotope to decay to a sample containing only 10 grams of it?

Finally we note that there are many examples of quantities besides population counts which satisfy the differential equation $\frac{dy}{dt} = ky$ with solution $y = y_0 e^{kt}$. As an example, the voltage across a discharging capacitor in an electronic circuit containing only a resistor and a capacitor (an RC circuit) undergoes exponential decay from an initial voltage V_0.

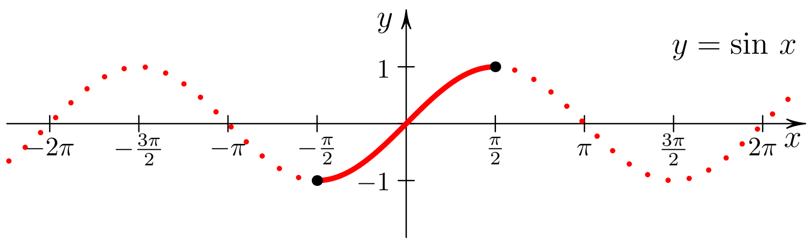
2.5 Inverse Trigonometric Functions

2.5.1 Inverse Sine

The sine function $y = \sin x$ on its natural domain $(-\infty, \infty)$ is not a one-to-one function. It clearly fails the horizontal line test as the intersection with the line $y = \frac{1}{2}$ clearly shows:

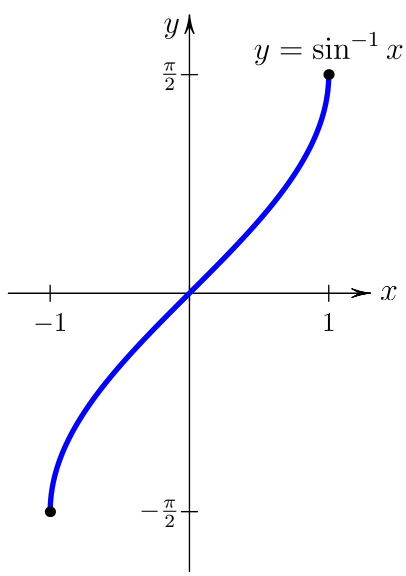


However the function $y = \sin x$ on domain $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is a one-to-one function:



Definition: The inverse function of $y = \sin x$; $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is called the **inverse sine function** or **arcsine function** and is denoted by $\boldsymbol{y = \sin^{-1} x}$ or $\boldsymbol{y = \arcsin x}$. It satisfies

$$y = \sin^{-1} x \iff x = \sin y \quad \left(-\frac{\pi}{2} \le y \le \frac{\pi}{2}\right)$$



The domain of inverse sine is $[-1, 1]$ and range is $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The usual inverse identities apply:

$$\begin{aligned} \sin^{-1}(\sin x) &= x \quad \text{for} \quad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ \sin\left(\sin^{-1} x\right) &= x \quad \text{for} \quad -1 \leq x \leq 1 \end{aligned}$$

Examples:
Evaluate the following:

1. $\sin^{-1}\left(\frac{\sqrt{3}}{2}\right)$
2. $\tan\left[\sin^{-1}\left(\frac{1}{2}\right)\right]$
3. $\sin\left(2\sin^{-1} x\right)$

Theorem: 2.16. The derivative of inverse sine is

$$\frac{d}{dx}\left(\sin^{-1} x\right) = \frac{1}{\sqrt{1-x^2}} \ ,$$

where $-1 < x < 1$.

To prove the theorem note that if $y = \sin^{-1} x$ then by definition of the inverse:

$$\sin y = x$$

Implicit differentiation of both sides with respect to x yields

$$(\cos y)\, y' = 1$$

and so $\frac{dy}{dx} = \frac{1}{\cos y}$. By the trigonometric identity $\cos^2 y + \sin^2 y = 1$ it follows that

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2}$$

where here the positive solution was taken since $-\frac{\pi}{2} < y < \frac{\pi}{2}$ implies $\cos y > 0$. Inserting this in the formula for $\frac{dy}{dx}$ gives the result.

Note that the Chain Rule result is, as expected:

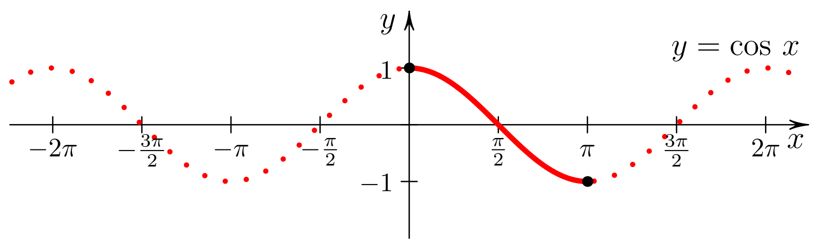
$$\frac{d}{dx}\left(\sin^{-1} g(x)\right) = \frac{g'(x)}{\sqrt{1 - g^2(x)}} \ .$$

Examples:
Differentiate the following:

1. $y = \sin^{-1}(\ln x + 3)$
2. $y = e^{\sin^{-1} x} + \sin^{-1}(e^x)$

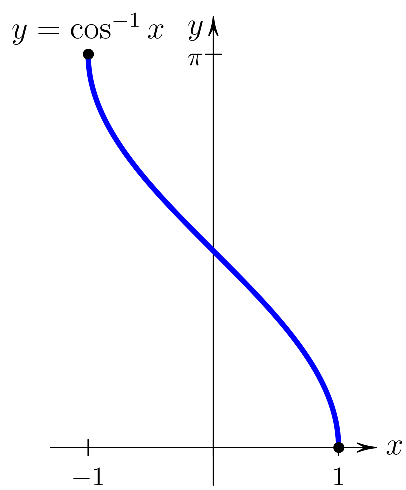
2.5.2 Inverse Cosine

The function $y = \cos x$ on domain $[0, \pi]$ is a one-to-one function:



Definition: The inverse function of $y = \cos x$; $[0, \pi]$ is called the **inverse cosine function** or **arccosine function** and is denoted by $\boldsymbol{y = \cos^{-1} x}$ or $\boldsymbol{y = \arccos x}$. It satisfies

$$y = \cos^{-1} x \iff x = \cos y \quad (0 \le y \le \pi)$$



The domain of inverse cosine is $[-1, 1]$ and range is $[0, \pi]$. The inverse identities are:

$$\begin{aligned} \cos^{-1}(\cos x) &= x \quad \text{for} \quad 0 \le x \le \pi \\ \cos\left(\cos^{-1} x\right) &= x \quad \text{for} \quad -1 \le x \le 1 \end{aligned}$$

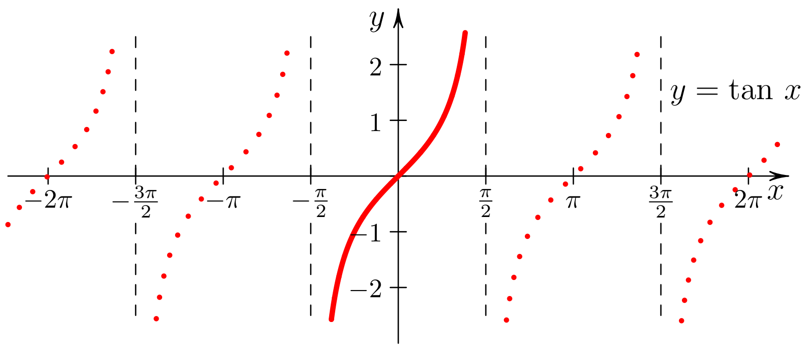
Theorem: 2.17. The derivative of inverse cosine is

$$\frac{d}{dx}\left(\cos^{-1} x\right) = -\frac{1}{\sqrt{1 - x^2}} \ ,$$

where $-1 < x < 1$.

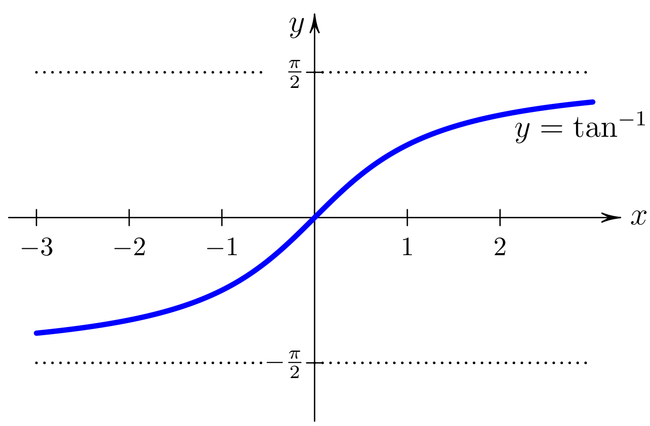
2.5.3 Inverse Tangent

The function $y = \tan x$ on domain $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ is a one-to-one function:



Definition: The inverse function of $y = \tan x$; $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ is called the **inverse tangent function** and is denoted by $\boldsymbol{y = \tan^{-1} x}$ or $\boldsymbol{y = \arctan x}$. It satisfies

$$y = \tan^{-1} x \iff x = \tan y \quad \left(-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}\right)$$



The domain of inverse tangent is $(-\infty, \infty)$ and range is $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. The usual inverse identities apply:

$$\begin{aligned} \tan^{-1}(\tan x) &= x \quad \text{for} \quad -\frac{\pi}{2} < x < \frac{\pi}{2} \\ \tan\left(\tan^{-1} x\right) &= x \quad \text{for} \quad x \in \mathbb{R} \end{aligned}$$

as well as the limits:

$$\lim_{x \to -\infty} \tan^{-1} x = -\frac{\pi}{2} \qquad\qquad \lim_{x \to \infty} \tan^{-1} x = \frac{\pi}{2}$$

So $y = \pm\frac{\pi}{2}$ are horizontal asymptotes of the function.

Theorem: 2.18. The derivative of inverse tangent is

$$\frac{d}{dx}\left(\tan^{-1} x\right) = \frac{1}{1 + x^2} \ .$$

2.5.4 Other Trigonometric Inverses

Similarly one defines $y = \csc^{-1} x$, $y = \sec^{-1} x$, and $y = \cot^{-1} x$.[3]

Notes:

- Since trigonometric functions are functions of angles, inverse trigonometric functions return angles. All our angles above are in radians. On your calculator you must have it set to radian mode to get these inverse trigonometric function results. If you have your calculator set to degree mode you will get your answers in degrees.
- The -1 in $\sin^{-1} x$ means inverse *not* taking to the power of -1 (reciprocal) like the 2 in $\sin^2 x$ means. If you mean take to the power of -1, i.e. $\frac{1}{\sin x}$ then you must write $(\sin x)^{-1}$ or simply use the reciprocal trigonometric function $\csc x$.
- It is because none of the trig functions are one-to-one and hence not invertible on their natural domains that solving a trigonometric equation $\text{trig}(x) = \#$ requires more than just "applying the inverse" to both sides (unlike, say comparable logarithmic or exponential equations). So to solve $\sin x = \frac{1}{2}$ the result $x = \sin^{-1}(1/2) = \pi/6$ is only one of many solutions. (See the intersections between $y = \sin x$ and $y = 1/2$ in our initial graph in this section.)

A complete table of the inverse trigonometric derivatives is as follows:

$$\frac{d}{dx}\left(\sin^{-1} x\right) = \frac{1}{\sqrt{1-x^2}} \qquad \frac{d}{dx}\left(\cos^{-1} x\right) = -\frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}\left(\tan^{-1} x\right) = \frac{1}{1+x^2} \qquad \frac{d}{dx}\left(\cot^{-1} x\right) = -\frac{1}{1+x^2}$$

$$\frac{d}{dx}\left(\sec^{-1} x\right) = \frac{1}{x\sqrt{x^2-1}} \qquad \frac{d}{dx}\left(\csc^{-1} x\right) = -\frac{1}{x\sqrt{x^2-1}}$$

Here the derivatives exist on the domains of the inverse trigonometric function except at those values where the expression is undefined. For the chain rule formulae simply replace x by $g(x)$ in each formula and multiply the result by $g'(x)$.

Examples:
Differentiate the following functions:

1. $y = \sin^{-1}(2x-1)$
2. $y = \tan^{-1}\left(\frac{x}{3}\right) + \ln\sqrt{\frac{x-3}{x+3}}$
3. $y = x\cos^{-1} x - \sqrt{1-x^2}$
4. $y = \sin^{-1}\left(\tan^{-1} x\right)$
5. $y = \tan^{-1}(\ln x)\, e^{x^2+3}$
6. $y = \cos^{-1}\left(e^{2x} - 5\right)$
7. $y = \tan^{-1}\left(x^2+3\right) - \tan\left(\cos^{-1} x + 1\right)$

[3] Note that the convention for the inverse secant and inverse cosecant functions used here is that the domain of secant (cosecant), and hence the range of inverse secant (cosecant) is $[0, \pi/2) \cup [\pi, 3\pi/2)$ ($(0, \pi/2] \cup (\pi, 3\pi/2]$). One can also choose the more intuitive interval $[0, \pi/2) \cup (\pi/2, \pi]$ for secant and $[-\pi/2, 0) \cup (0, \pi/2]$ for cosecant but then absolute value bars are required about the x outside the radical in the derivative formulae.

8. $f(t) = \sec^{-1}\left(e^t + \ln t\right)$
9. $\sin^{-1} y = x^2 + y^2 + e^y$

The derivatives of the inverse trigonometric functions give the following results:

Theorem: 2.19. $\displaystyle\int \frac{1}{\sqrt{1-x^2}}\,dx = \sin^{-1} x + C \qquad \int \frac{1}{1+x^2}\,dx = \tan^{-1} x + C$

If one considers the more general integral

$$\int \frac{1}{\sqrt{a^2-x^2}}\,dx\ ,$$

where $a > 0$ is constant this can be solved by first noting that

$$\sqrt{a^2-x^2} = \sqrt{a^2\left(1-\frac{x^2}{a^2}\right)} = \sqrt{a^2}\sqrt{1-\frac{x^2}{a^2}} = |a|\sqrt{1-\frac{x^2}{a^2}} = a\sqrt{1-\frac{x^2}{a^2}}$$

and then using the substitution $u = \frac{x}{a}$ (so $du = \frac{dx}{a}$) to get:

$$\int \frac{1}{\sqrt{a^2-x^2}}\,dx = \int \frac{1}{a\sqrt{1-\frac{x^2}{a^2}}}\,dx = \int \frac{1}{\sqrt{1-u^2}}\,du = \sin^{-1} u + C = \sin^{-1}\frac{x}{a} + C\ .$$

Similar generalization can be done to the inverse tangent integral. We thus have:

Theorem: 2.20. For constant $a > 0$,

$$\int \frac{1}{\sqrt{a^2-x^2}}\,dx = \sin^{-1}\frac{x}{a} + C \qquad\qquad \int \frac{1}{x^2+a^2}\,dx = \frac{1}{a}\tan^{-1}\frac{x}{a} + C$$

Integrals with literal constants like this (i.e. a) are what one includes typically in integral tables.

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \frac{3}{\sqrt{4-2x^2}}\,dx$
2. $\displaystyle\int \frac{\tan^{-1} x}{1+x^2}\,dx$
3. $\displaystyle\int \frac{4}{t\left[9+(\ln t)^2\right]}\,dt$
4. $\displaystyle\int \frac{e^x}{\sqrt{1-8e^{2x}}}\,dx$
5. $\displaystyle\int \frac{3x+4}{2x^2+3}\,dx$
6. $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\cos x}{1+\sin^2 x}\,dx$

2.6 L'Hôpital's Rule

We have already evaluated limits that are indeterminate forms of the type $\frac{0}{0}$ and $\frac{\infty}{\infty}$.

Example:
Evaluate the following limits:

1. $\displaystyle\lim_{x\to 1} \frac{x^2-1}{x^2-3x+2}$
2. $\displaystyle\lim_{x\to\infty} \frac{3x^2-5x+2}{4x^2+3x-10}$

In general:

- If $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, then $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)}$ is called the **indeterminate form of type $\frac{0}{0}$**.
- If $\lim_{x\to a} f(x) = \pm\infty$ and $\lim_{x\to a} g(x) = \pm\infty$, then $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)}$ is called the **indeterminate form of type $\frac{\infty}{\infty}$**.

Our techniques used above will not work for evaluating all limits of this type:

Example:
The limit $\displaystyle\lim_{x\to 0} \frac{2^x-1}{x}$ is an indeterminate form of type $\frac{0}{0}$ while $\displaystyle\lim_{x\to\infty} \frac{\ln x}{x}$ is of type $\frac{\infty}{\infty}$. Neither limit may be resolved using the methods of the previous example.

Theorem: 2.21. If f and g are differentiable functions with $g'(x) \neq 0$ on an open interval containing the value a and $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)}$ is an indeterminate form of type $\frac{0}{0}$ or $\frac{\infty}{\infty}$, (i.e. $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$ or $\lim_{x\to a} f(x) = \pm\infty$ and $\lim_{x\to a} g(x) = \pm\infty$) then

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

provided the limit on the right hand side either exists or is $\pm\infty$. This is **L'Hôpital's Rule**.

Example:
Evaluate the previous limits using L'Hôpital's Rule:

1. $\displaystyle\lim_{x\to 0} \frac{2^x-1}{x} = \lim_{x\to 0} \frac{2^x \ln 2 - 0}{1} = 2^0 \ln 2 = (1)(\ln 2) = \ln 2$
2. $\displaystyle\lim_{x\to\infty} \frac{\ln x}{x} = \lim_{x\to\infty} \frac{\frac{1}{x}}{1} = \lim_{x\to\infty} \frac{1}{x} = 0$

Note that when applying L'Hôpital's Rule one is **not** using the Quotient Rule! The derivatives in the numerator and denominator are taken separately.

Examples:
Evaluate the following limits:

1. $\lim\limits_{x\to 0} \dfrac{e^x - 1}{x}$
2. $\lim\limits_{x\to 0} \dfrac{\sin x}{x}$
3. $\lim\limits_{x\to 0} \dfrac{\cos x + 2x - 1}{3x}$
4. $\lim\limits_{x\to 2} \dfrac{x^2 + 3x + 5}{x^2 - 4}$
5. $\lim\limits_{x\to \infty} \dfrac{\ln x}{\sqrt{x}}$
6. $\lim\limits_{x\to 0} \dfrac{x - \sin x}{x^3}$
7. $\lim\limits_{x\to \infty} \dfrac{e^x}{\ln x}$
8. $\lim\limits_{x\to 0} \dfrac{\cos x}{x^2 - 1}$
9. $\lim\limits_{x\to 5} \dfrac{\sqrt{x-1} - 2}{x^2 - 25}$
10. $\lim\limits_{x\to 0} \dfrac{\sin x}{x - \tan x}$
11. $\lim\limits_{x\to 0} \dfrac{3^x - 1}{x}$
12. $\lim\limits_{x\to 0} \dfrac{4e^{2x} - 4}{e^x - 1}$

2.6.1 Indeterminate Forms of type $0 \cdot \infty$ and $\infty - \infty$

Consider the following indeterminate forms:

- If $\lim\limits_{x\to a} f(x) = 0$ and $\lim\limits_{x\to a} g(x) = \pm\infty$ then $\lim\limits_{x\to a} f(x)g(x)$ is called the **indeterminate form of type $0 \cdot \infty$**.
- If $\lim\limits_{x\to a} f(x) = \infty$ and $\lim\limits_{x\to a} g(x) = \infty$ then $\lim\limits_{x\to a} [f(x) - g(x)]$ is called the **indeterminate form of type $\infty - \infty$**.

To solve an indeterminate form of type $0 \cdot \infty$, write the product $f \cdot g$ as either

$$f \cdot g = \frac{f}{1/g} \quad \text{or} \quad f \cdot g = \frac{g}{1/f}$$

This will convert the indeterminate form into a form of type $\frac{0}{0}$ or $\frac{\infty}{\infty}$ which can then potentially evaluated using L'Hôpital's Rule.

For indeterminate forms of type $\infty - \infty$ try to convert the difference into a quotient (by using a common denominator or factoring out common terms or rationalization) to once again reduce the limit to type $\frac{0}{0}$ or $\frac{\infty}{\infty}$.

Examples:
Evaluate the following limits:

1. $\lim\limits_{x\to 0^+} x^2 \ln x$
2. $\lim\limits_{x\to \frac{\pi}{2}} (2x-\pi)\sec x$
3. $\lim\limits_{x\to 0} \left(\dfrac{1}{e^x-1} - \dfrac{1}{x}\right)$
4. $\lim\limits_{x\to 1} \left(\dfrac{1}{x^2-1} - \dfrac{1}{x-1}\right)$
5. $\lim\limits_{x\to \infty} \left(\sqrt{x^2+x} - x\right)$
6. $\lim\limits_{x\to \infty} (x - \ln x)$

2.6.2 Exponential Indeterminate Forms

Several indeterminate forms arise from $\lim\limits_{x\to a} [f(x)]^{g(x)}$.

- If $\lim\limits_{x\to a} f(x) = 0$ and $\lim\limits_{x\to a} g(x) = 0$ then **indeterminate form of type $\mathbf{0^0}$**.
- If $\lim\limits_{x\to a} f(x) = \infty$ and $\lim\limits_{x\to a} g(x) = 0$ then **indeterminate form of type $\mathbf{\infty^0}$**.
- If $\lim\limits_{x\to a} f(x) = 1$ and $\lim\limits_{x\to a} g(x) = \pm\infty$ then **indeterminate form of type $\mathbf{1^\infty}$**.

Each of these can be evaluated either by taking the natural logarithm:

$$y = [f(x)]^{g(x)} \;\Rightarrow\; \ln y = g(x)\ln[f(x)] \;,$$

or by writing the function as an exponential:

$$[f(x)]^{g(x)} = e^{g(x)\ln[f(x)]}$$

In either case an indeterminate form of type $0\cdot\infty$ will result.

Example:
As a practical example prove our limit formula for e by evaluating $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}}$, a limit of indeterminate form 1^∞.

Let $y = (1+x)^{\frac{1}{x}}$. Taking the natural logarithm of both sides results in

$$\ln y = \frac{1}{x}\ln(1+x)$$

Taking the limit as $x \to 0$ of the righthand side gives an indeterminate form of type $\frac{0}{0}$ readily evaluated using L'Hôpital's Rule:

$$\lim_{x\to 0}\ln y = \lim_{x\to 0}\frac{\ln(1+x)}{x} = \lim_{x\to 0}\frac{\frac{1}{1+x}}{1} = \lim_{x\to 0}\frac{1}{1+x} = \frac{1}{1+0} = 1$$

So

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}} = \lim_{x\to 0} y = \lim_{x\to 0} e^{\ln y} = e^{\left(\lim\limits_{x\to 0} \ln y\right)} = e^1 = e$$

This was the limit stated for e given before.

Examples:
Evaluate the following limits:

1. $\lim\limits_{x\to\infty} (1+e^x)^{e^{-x}}$
2. $\lim\limits_{x\to 0^+} (e^x-1)^x$
3. $\lim\limits_{x\to\infty} \left(1+\dfrac{3}{x}+\dfrac{5}{x^2}\right)^x$
4. $\lim\limits_{x\to 1^-} (1-x)^{\ln x}$
5. $\lim\limits_{x\to \frac{\pi}{2}^-} (\tan x)^{\cos x}$

Unit 3: Integration Methods

3.1 Integration by Parts

Just as the Chain Rule for differentiation leads to the useful Method of Substitution for solving integrals, so too does the Product Rule result in a useful method for solving integrals. Starting with the Product Rule

$$\frac{d}{dx}\left[f(x)g(x)\right] = f'(x)g(x) + f(x)g'(x) \ ,$$

one can integrate both sides of the equation to get:

$$\int \frac{d}{dx}\left[f(x)g(x)\right]dx = \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx$$

An antiderivative of the derivative of a function is just the function itself so the left-hand side becomes $f(x)g(x) + C$. The constant C may be absorbed into the indefinite integrals on the right and so one has:

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

Reordering the terms gives

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

The formula suggests that a useful strategy for evaluating an integral is to consider an integrand as a product of two terms, one of which may be differentiated ($f(x)$) and one which may be integrated ($g'(x)$) to produce a new integral that is perhaps more easy to evaluate than the original. It is customary to define $u = f(x)$ and $v = g(x)$. One then has the corresponding differentials $du = f'(x)dx$ and $dv = g'(x)dx$. The formula becomes:

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

This is the **Integration by Parts** formula.

Example:
Integrate $\int x^2 \ln x\,dx$.

The fact that $\ln x$ is easily differentiated and x^2 easily integrated suggests we reorder the terms and identify u and dv as follows:

$$\int \underbrace{\ln x}_{=u}\,\underbrace{x^2\,dx}_{=dv}$$

Then $u = \ln x$ implies (differentiating) that $du = \frac{1}{x}\,dx$. The differential $dv = x^2\,dx$ is integrated to give $v = \frac{1}{3}x^3$. The Integration by Parts formula $\int u\,dv = uv - \int v\,du$ implies:

$$\begin{aligned}
\int (\ln x)\cdot\left(x^2\,dx\right) &= (\ln x)\cdot\left(\frac{1}{3}x^3\right) - \int\left(\frac{1}{3}x^3\right)\cdot\left(\frac{1}{x}\,dx\right) \\
&= \frac{1}{3}x^3\ln x - \frac{1}{3}\int x^2\,dx \\
&= \frac{1}{3}x^3\ln x - \frac{1}{3}\cdot\frac{1}{3}x^3 + C \\
&= \frac{1}{3}x^3\ln x - \frac{1}{9}x^3 + C
\end{aligned}$$

In general, to apply Integration by Parts select u and dv so that

1. The product $u\,dv$ is equal to the original integrand.
2. dv can be integrated.
3. The new integral $\int v\,du$ is easier than the original integral.
4. For integrals involving $x^p e^{ax}$ try $u = x^p$, $dv = e^{ax}\,dx$.
5. For integrals involving $x^p(\ln x)^q$ try $u = (\ln x)^q$, $dv = x^p\,dx$.

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \ln x\,dx$
2. $\displaystyle\int xe^x\,dx$
3. $\displaystyle\int x^2e^{-x}\,dx$
4. $\displaystyle\int e^x\cos x\,dx$
5. $\displaystyle\int_0^1 \tan^{-1} x\,dx$
6. $\displaystyle\int x^3(\ln x)^2\,dx$
7. $\displaystyle\int \sin(\ln x)\,dx$
8. $\displaystyle\int \theta\sec^2\theta\,d\theta$
9. $\displaystyle\int x^5e^{-x^3}\,dx$
10. $\displaystyle\int x\sin(x^2)\,dx$
11. $\displaystyle\int \cos^2 x\,dx$

3.2 Trigonometric Integrals

Strategy for Evaluating $\int \sin^m x \cos^n x \, dx$

1. For an odd power of sine ($m = 2k+1$), save one sine factor and express the remaining sine factors in terms of cosine using the identity $\sin^2 x = 1 - \cos^2 x$:

$$\int \sin^{2k+1} x \cos^n x \, dx = \int \left(\sin^2 x\right)^k \cos^n x \sin x \, dx = \int \left(1 - \cos^2 x\right)^k \cos^n x \sin x \, dx$$

   Then substitute $u = \cos x$.

2. For an odd power of cosine ($n = 2k+1$), save one cosine factor and expressing the remaining cosine factors in terms of sine using the identity $\cos^2 x = 1 - \sin^2 x$:

$$\int \sin^m x \cos^{2k+1} x \, dx = \int \sin^m x \left(\cos^2 x\right)^k \cos x \, dx = \int \sin^m x \left(1 - \sin^2 x\right)^k \cos x \, dx$$

   Then substitute $u = \sin x$.

3. If the powers of both sine and cosine are even, use the trigonometric identities:

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

   These may need to be used repeatedly. The identity $\sin x \cos x = \frac{1}{2} \sin 2x$ may also be useful.

Either 1 or 2 can be used if the powers of sine and cosine are both odd.

Examples:
Evaluate the following integrals:

1. $\int \sin^4 x \cos^3 x \, dx$
2. $\int \sin^3 x \, dx$
3. $\int \cot^5 x \sin^2 x \, dx$
4. $\int \sin^2 x \, dx$
5. $\int \sin^4 x \, dx$
6. $\int \cos^2 x \sin^2 x \, dx$

Strategy for Evaluating $\int \tan^m x \sec^n x \, dx$

1. For an odd power of tangent ($m = 2k+1$), save a factor of $\sec x \tan x$ and express the remaining factors of tangent in terms of $\sec x$ using the identity $\tan^2 x = \sec^2 x - 1$:

$$\int \tan^{2k+1} x \sec^n x \, dx = \int \left(\tan^2 x\right)^k \sec^{n-1} x \sec x \tan x \, dx = \int \left(\sec^2 x - 1\right)^k \sec^{n-1} x \sec x \tan x \, dx$$

   Then substitute $u = \sec x$.

2. For an even power of secant ($n = 2k$), save a factor of $\sec^2 x$ and express the remaining secant factors in terms of $\tan x$ using the identity $\sec^2 x = 1 + \tan^2 x$:

$$\int \tan^m x \sec^{2k} x \, dx = \int \tan^m x \left(\sec^2 x\right)^{k-1} \sec^2 x \, dx = \int \tan^m x \left(1 + \tan^2 x\right)^{k-1} \sec^2 x \, dx$$

Then substitute $u = \tan x$.

3. If m is even and $n = 0$ (i.e. no factors of secant), convert a single factor of $\tan^2 x$ using $\tan^2 x = \sec^2 x - 1$. The first term will then be integrable and the procedure may be repeated on the second integral of now lower power.

Strategy for Evaluating $\int \cot^m x \csc^n x \, dx$

1. For an odd power of cotangent ($m = 2k+1$), save a factor of $\csc x \cot x$ and express the remaining factors of cotangent in terms of $\csc x$ using the identity $\cot^2 x = \csc^2 x - 1$:

$$\int \cot^{2k+1} x \csc^n x \, dx = \int \left(\cot^2 x\right)^k \csc^{n-1} x \csc x \cot x \, dx = \int \left(\csc^2 x - 1\right)^k \csc^{n-1} x \csc x \cot x \, dx$$

Then substitute $u = \csc x$.

2. For an even power of cosecant ($n = 2k$), save a factor of $\csc^2 x$ and express the remaining factors of cosecant in terms of $\cot x$ using the identity $\csc^2 x = 1 + \cot^2 x$:

$$\int \cot^m x \csc^{2k} x \, dx = \int \cot^m x \left(\csc^2 x\right)^{k-1} \csc^2 x \, dx = \int \cot^m x \left(1 + \cot^2 x\right)^{k-1} \csc^2 x \, dx$$

Then substitute $u = \cot x$.

3. If m is even and $n = 0$ (i.e. no factors of cosecant), convert a single factor of $\cot^2 x$ using $\cot^2 x = \csc^2 x - 1$. The first term will then be integrable and the procedure may be repeated on the second integral of now lower power.

Note: This strategy is identical for that of tangents and secants with the identification $\tan \Rightarrow \cot$ and $\sec \Rightarrow \csc$.

Examples:
Evaluate the following integrals:

1. $\int \tan^3 x \sec^3 x \, dx$
2. $\int \tan^2 x \sec^4 x \, dx$
3. $\int \tan^3 x \, dx$
4. $\int \sec x \, dx$
5. $\int \sec^3 x \, dx$
6. $\int \tan^4 x \, dx$
7. $\int \cot^3 x \csc^4 x \, dx$
8. $\int \cot^3 x \csc^3 x \, dx$
9. $\int \csc x \, dx$
10. $\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} \csc^4 x \, dx$

Strategy for Evaluating $\int \sin mx \cos nx\,dx$, $\int \sin mx \sin nx\,dx$, $\int \cos mx \cos nx\,dx$

Apply the corresponding trigonometric identity:

- $\sin a \cos b = \dfrac{1}{2}\left[\sin(a-b) + \sin(a+b)\right]$
- $\sin a \sin b = \dfrac{1}{2}\left[\cos(a-b) - \cos(a+b)\right]$
- $\cos a \cos b = \dfrac{1}{2}\left[\cos(a-b) + \cos(a+b)\right]$

with $a = mx$ and $b = nx$.

> **Examples:**
> Evaluate the following integrals:
>
> 1. $\displaystyle\int \sin 4x \cos 5x\,dx$
> 2. $\displaystyle\int \sin 2x \sin 6x\,dx$
> 3. $\displaystyle\int_0^{\frac{\pi}{4}} \cos 2x \cos 4x\,dx$
> 4. $\displaystyle\int \sin 2x \sin 6x \cos 2x\,dx$

A Note on the Identities

Note that the various trigonometric identities on this handout follow readily from the three basic identities:

a) $\sin^2 x + \cos^2 x = 1$
b) $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$
c) $\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$

The half-angle identities follow from c) setting $y = x$ and then replacing alternately $\sin^2 x$ or $\cos^2 x$ using a). Dividing a) by $\cos^2 x$ gives the identity involving tangent and secant, while dividing a) by $\sin^2 x$ gives the identity involving cotangent and cosecant. The last three identities on this page follow by solving for the various products using the $+$ and $-$ equations from the appropriate angle addition formula b) or c).

3.3 Trigonometric Substitution

Some integrals, typically involving roots, may be resolved by using the Substitution Method where the old variable is defined in terms of a new variable via a trigonometric function.

Example:
Find the indefinite integral $\int \sqrt{4-x^2}\,dx$. We consider the substitution $\theta(x)$ defined via

$$x = 2\sin\theta$$

(and so $dx = 2\cos\theta\,d\theta$). Unlike our usual application of the substitution method here we have defined θ implicitly. To make $\theta(x)$ unique as required we add the additional constraint $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$. (Equivalently we recognize that the explicit substitution which has been done is just $\theta = \sin^{-1}\left(\frac{x}{2}\right)$ which, recall, is defined with this range.) The integral becomes

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

Note that when we solved the identity $1-\sin^2\theta = \cos^2\theta$ for $\cos\theta$ we used that $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$ to get $\cos\theta = \sqrt{1-\sin^2\theta}$ since $\cos\theta$ is indeed positive on the interval. This choice of positive sign was also used in our final step where again $\cos\theta$ was represented by a positive value:

$$\cos\theta = \sqrt{1-\sin^2\theta} = \sqrt{1-x^2/4} = \sqrt{(4-x^2)/4} = \sqrt{4-x^2}\sqrt{1/4} = \frac{1}{2}\sqrt{4-x^2}$$

Here we could have also drawn a right triangle with angle θ and length x opposite and hypotenuse of 2 to work out $\cos\theta$.

This method is called **Trigonometric Substitution**. More generally if an integrand contains one of $\sqrt{a^2-x^2}$, $\sqrt{a^2+x^2}$, or $\sqrt{x^2-a^2}$ (where $a>0$ is constant) then the radical sign can be removed via the appropriate substitution:

Expression	Substitution	Identity
$\sqrt{a^2-x^2}$	$x = a\sin\theta \quad \left(-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}\right)$	$1-\sin^2\theta = \cos^2\theta$
$\sqrt{a^2+x^2}$	$x = a\tan\theta \quad \left(-\frac{\pi}{2} < \theta < \frac{\pi}{2}\right)$	$1+\tan^2\theta = \sec^2\theta$
$\sqrt{x^2-a^2}$	$x = a\sec\theta \quad \left(0 \le \theta < \frac{\pi}{2} \text{ or } \pi \le \theta < \frac{3\pi}{2}\right)$	$\sec^2\theta - 1 = \tan^2\theta$

Note that all the ranges of θ have been chosen so that θ will equal the relevant inverse trigonometric function with argument x/a.

Example:
Prove the Archimedian result that the area of a circle of radius R is $A = \pi R^2$.

The area of a semi-circle of radius R is the area under the curve $y = \sqrt{R^2 - x^2}$ between $x = -R$ and $x = R$ and so the area of a circle is

$$A = 2\int_{-R}^{R} \sqrt{R^2 - x^2}\, dx$$

Using substitution $x = R\sin\theta$, with $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$ gives $dx = R\cos\theta\, d\theta$. So $\theta = \sin^{-1}(x/R)$ and the limits become, for $x = R$, $\theta = \sin^{-1}(R/R) = \sin^{-1}(1) = \pi/2$ and for $x = -R$,
$\theta = \sin^{-1}(-R/R) = \sin^{-1}(-1) = -\pi/2$. The solution of the integral follows, similar to the last example,

$$\begin{aligned}
A = 2\int_{-R}^{R} \sqrt{R^2 - x^2}\, dx &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{R^2 - R^2\sin^2\theta} \cdot R\cos\theta\, d\theta \\
&= 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{R^2}\sqrt{1 - \sin^2\theta} \cdot R\cos\theta\, d\theta \\
&= 2R^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\theta\cos\theta\, d\theta = 2R^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta\, d\theta \\
&= 2R^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2}(1 + \cos 2\theta)\, d\theta = R^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1 + \cos 2\theta)\, d\theta \\
&= R^2 \left[\theta + \frac{1}{2}\sin 2\theta\right|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= R^2 \left\{ \left[\frac{\pi}{2} + \frac{1}{2}\sin\pi\right] - \left[-\frac{\pi}{2} + \frac{1}{2}\sin(-\pi)\right] \right\} \\
&= \pi R^2
\end{aligned}$$

Examples:
Evaluate the following integrals:

1. $\displaystyle\int_1^2 \frac{1}{x^2\sqrt{16 - x^2}}\, dx$
2. $\displaystyle\int \frac{\sqrt{x^2 - 9}}{x^4}\, dx$
3. $\displaystyle\int \frac{1}{(x^2 + 2x + 2)^2}\, dx$
4. $\displaystyle\int \frac{2x - 3}{x^2 - 4x + 8}\, dx$
5. $\displaystyle\int \frac{1}{x^3\sqrt{x^2 - 25}}\, dx$
6. $\displaystyle\int \frac{x^2}{(2 - 9x^2)^{\frac{3}{2}}}\, dx$
7. $\displaystyle\int \frac{1}{(5 - 4x - x^2)^{\frac{5}{2}}}\, dx$
8. $\displaystyle\int \frac{\sqrt{x - 4}}{x}\, dx$

3.4 Partial Fraction Decomposition

The rational function $\dfrac{x+2}{x^3-x^2}$ can be shown to be equal to $-\dfrac{3}{x}-\dfrac{2}{x^2}+\dfrac{3}{x-1}$. Therefore the integral of the former rational function is:

$$\int \frac{x+2}{x^3-x^2}\,dx = \int\left(-\frac{3}{x}-\frac{2}{x^2}+\frac{3}{x-1}\right)dx = -3\ln|x|+\frac{2}{x}+3\ln|x-1|+C$$

This example suggests that determining a technique to decompose a rational function in this way would provide a method for its integration.

A polynomial $P(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, $a_n \neq 0$ is said to have degree n. A function $f(x) = \frac{P(x)}{Q(x)}$ where $P(x)$ and $Q(x)$ are polynomials is a rational function.

The rational number $\frac{7}{4}$ is called improper because the numerator is larger than the denominator. Through division of 4 into 7 one can write $\frac{7}{4}$ as $1\frac{3}{4}$ where the fractional part, $\frac{3}{4}$ is a proper fraction. Analagous definitions are made for rational functions.

Definition: A rational function $f(x) = \frac{P(x)}{Q(x)}$ is called **proper** if the degree of P is less than the degree of Q. Otherwise $f(x)$ is called **improper** if $\deg(P) \geq \deg(Q)$.

Note:

1. If $f(x) = P(x)/Q(x)$ is proper then it is possible to express it as a sum of simpler fractional functions called **partial fractions** which are integrable.
2. If $f(x)$ is improper, then use long division to divide P by Q until a remainder $R(x)$ is obtained such that $\deg R < \deg Q$. Then
$$f(x) = \frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)}$$
where $S(x)$ and $R(x)$ are polynomials. $S(x)$ is then integrable as it is a polynomial while the proper rational function $R(x)/Q(x)$ can in turn be integrated by the method of partial fractions thereby making $f(x)$ integrable.

Examples:
For the following rational functions determine if they are proper or improper. For those that are improper write them as a polynomial plus a proper rational function.

1. $f(x) = \dfrac{x+1}{x^3-3x^2+2}$
2. $f(x) = \dfrac{x^2+1}{x^2+3x}$
3. $f(x) = \dfrac{x^4+5x^2+1}{x^2+2}$

Definition: Let $g(x) = ax^2+bx+c$ be a quadratic function with real coefficients. If $b^2-4ac \geq 0$ then $g(x)$ is called **reducible** because it can be written as a product of linear factors with real coefficients. If $b^2-4ac < 0$ then $g(x)$ is called **irreducible** because it cannot be written as a product of linear factors with real coefficients.

Examples:

1. The function $g(x) = x^2 + 5x + 6$ has $b^2 - 4ac = 25 - 24 = 1 > 0$ and so is reducible. It clearly factors as $g(x) = (x + 2)(x + 3)$.

2. The function $g(x) = 2x^2 + 4x + 5$ has $b^2 - 4ac = 16 - 40 = -24 < 0$ and is irreducible.

Note: It can be shown, as a consequence of the *Fundamental Theorem of Algebra*, that any polynomial $Q(x)$ with real coefficients can be factored as a product of linear factors of the form $(ax + b)$ and/or quadratic irreducible factors of the form $ax^2 + bx + c$, where a, b, and c are real numbers.

Theorem: 3.1. If $P(x)$ and $Q(x)$ are polynomials and $\deg P < \deg Q$ the it follows that

$$\frac{P(x)}{Q(x)} = F_1 + F_2 + \ldots + F_n$$

where each F_i has one of the forms

$$\frac{A}{(ax+b)^i} \quad \text{or} \quad \frac{Ax+B}{(ax^2+bx+c)^j}$$

for some nonnegative integers i and j. The sum $F_1 + F_2 + \ldots + F_n$ is called the **partial fraction decomposition** of $\frac{P(x)}{Q(x)}$ and each F_i is called a **partial fraction**. The denominator polynomials are real linear functions and irreducible quadratics respectively.

Steps for finding Partial Fraction Decomposition

To decompose $f(x) = \frac{P(x)}{Q(x)}$ into partial fractions do the following:

1. If $\deg P \geq \deg Q$ then use long division to get
$$\frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)}$$

2. Express $Q(x)$ as a product of linear and/or quadratic irreducible factors.

3. Express the proper rational function ($P(x)/Q(x)$ or $R(x)/Q(x)$) as a sum of partial fractions of the form
$$\frac{A}{(ax+b)^i} \quad \text{and/or} \quad \frac{Ax+B}{(ax^2+bx+c)^j}$$

4. Evaluate the constants.

Once the partial fraction decomposition has been accomplished the necessary integration may be completed.

Upon factoring $Q(x)$ there are four cases that are logically possible.

Case I: $Q(x)$ is a product of distinct linear factors.

Suppose that

$$Q(x) = (a_1x + b_1)(a_2x + b_2)\dots(a_kx + b_k)$$

where no factor is repeated. Then there exist constants A_1, A_2, $\dots A_k$ such that

$$\frac{P(x)}{Q(x)}\left(\text{or } \frac{R(x)}{Q(x)}\right) = \frac{A_1}{a_1x + b_1} + \frac{A_2}{a_2x + b_2} + \dots + \frac{A_k}{a_kx + b_k}$$

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \frac{1}{x^2 + 2x - 3}\,dx$
2. $\displaystyle\int \frac{4x^2 + 13x - 9}{x^3 + 2x^2 - 3x}\,dx$
3. $\displaystyle\int \frac{4x^2 + 3x + 1}{x^2 - 1}\,dx$

Case II: $Q(x)$ is a product of linear factors some of which are repeated.

If $Q(x)$ has a factor $(ax + b)^r$ then the partial fraction decomposition will have the following terms due to that factor:

$$\frac{A_1}{ax + b} + \frac{A_2}{(ax + b)^2} + \dots + \frac{A_r}{(ax + b)^r}$$

where A_1, A_2, $\dots A_r$ are constants. Use distinct constants (i.e. A, B, C, etc.) for each factor.

Example:

The (proper) rational function $\dfrac{x^4 + 1}{x(3x + 2)^3(2x - 1)^2}$ decomposes into

$$\frac{x^4 + 1}{x(3x + 2)^3(2x - 1)^2} = \frac{A}{x} + \frac{B_1}{3x + 2} + \frac{B_2}{(3x + 2)^2} + \frac{B_3}{(3x + 2)^3} + \frac{C_1}{2x - 1} + \frac{C_2}{(2x - 1)^2}$$

where the constants A, B_1, B_2, B_3, C_1, and C_2 would then have to be determined.

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \frac{x^3 - 4x - 1}{x(x - 1)^3}\,dx$
2. $\displaystyle\int \frac{3x^2 + 5x - 10}{x^2(3x - 5)}\,dx$

Case III: $Q(x)$ contains a nonrepeated irreducible quadratic factor.

If $Q(x)$ has a nonrepeated irreducible factor $ax^2 + bx + c$ (so $b^2 - 4ac < 0$), then the partial fraction decomposition will have the following term due to that factor:

$$\frac{Ax + B}{ax^2 + bx + c}$$

where A and B are constants.

Example:
Evaluate the integral $\displaystyle\int \frac{x^3 - 4x^2 + 2}{(x^2+1)(x^2+2)}\,dx$

Case IV: $Q(x)$ contains a repeated irreducible quadratic factor.

If $Q(x)$ has an irreducible factor $\left(ax^2 + bx + c\right)^r$ then the partial fraction decomposition will have the following terms due to that factor:

$$\frac{A_1x + B_1}{ax^2 + bx + c} + \frac{A_2x + B_2}{(ax^2 + bx + c)^2} + \ldots + \frac{A_rx + B_r}{(ax^2 + bx + c)^r}$$

where A_1, A_2, $\ldots$, A_r, and B_1, B_2, $\ldots$, B_r are constants.

Example:
Evaluate the following integral: $\displaystyle\int \frac{2x^6 + 5x^4 + 2x^2 + 1}{x\,(x^2+1)^2}\,dx$

Having looked at all the cases we can write down the partial fraction decomposition of arbitrary rational functions.

Examples:
Write down the form of the partial fraction decomposition of the following rational functions. Do not evaluate the constants.

1. $\displaystyle\frac{x^2 - x - 21}{2x^3 - x^2 + 8x - 4}$
2. $\displaystyle\frac{x^3 + 2}{(x^2+4)^2}$
3. $\displaystyle\frac{x^6 + 5x^3 + x - 1}{x^4 + 5x^2 + 4}$
4. $\displaystyle\frac{x + 1}{(x^2-4)^2(x^2+3)}$

Rationalizing Substitutions

Some nonrational functions can be changed into rational functions by means of a substitution.

Example:
Evaluate $\displaystyle\int \frac{4\sqrt{x}}{x-2}\,dx$

3.5 General Strategies for Integration

Unlike differentiation which is largely a deterministic application of rules, integration is an art, with many indefinite integrals not even having an antiderivative that may be written in terms of known functions.

The following basic strategies have been seen

1. Basic Formulas of Integration
2. Substitution
3. Integration by Parts
4. Trigonometric Integrals
5. Trigonometric Substitution
6. Partial Fraction Decomposition
7. Rationalizing Substitution

One or more of these strategies along with using functional identities to rewrite the integrand may need to be applied to evaluate an integral.

Examples:
Evaluate the following integrals:

1. $\displaystyle\int \frac{e^{3t}}{1+e^{6t}}\,dt$
2. $\displaystyle\int e^{x+e^x}\,dx$
3. $\displaystyle\int \frac{1+e^x}{1-e^x}\,dx$
4. $\displaystyle\int x^2\ln(1+x)\,dx$
5. $\displaystyle\int \tan x\sec^6 x\,dx$
6. $\displaystyle\int \frac{e^{3x}}{1+e^x}\,dx$
7. $\displaystyle\int \frac{\cos^3 x}{\sqrt{1+\sin x}}\,dx$
8. $\displaystyle\int \frac{x}{\csc\left(5x^2\right)}\,dx$
9. $\displaystyle\int \left(2x+2^x+2^\pi\right)\,dx$
10. $\displaystyle\int \frac{7x^2+20x+65}{x^4+4x^3+13x^2}\,dx$

3.6 Improper Integrals

The definite integral due to Riemann which we use involves functions integrated over a closed interval $[a, b]$. Functions which are piecewise continuous where there are only a finite number of jump discontinuities are integrable. We now consider *improper integrals* where these restrictions do not hold. We consider two cases:

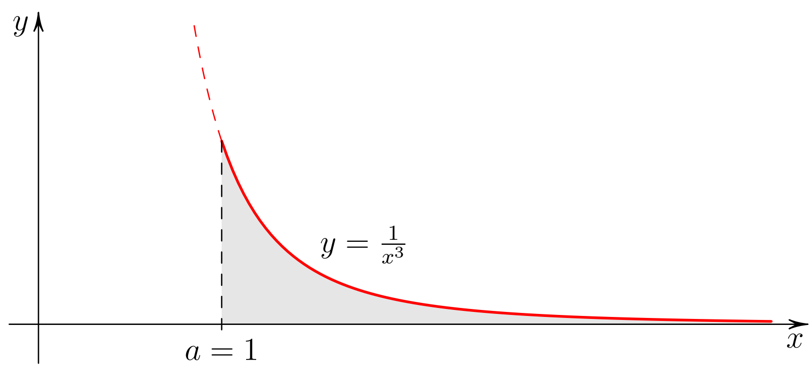
Improper Integrals of the First Kind : The interval of integration is infinite.

Improper Integrals of the Second Kind : The interval of integration contains an infinite discontinuity.

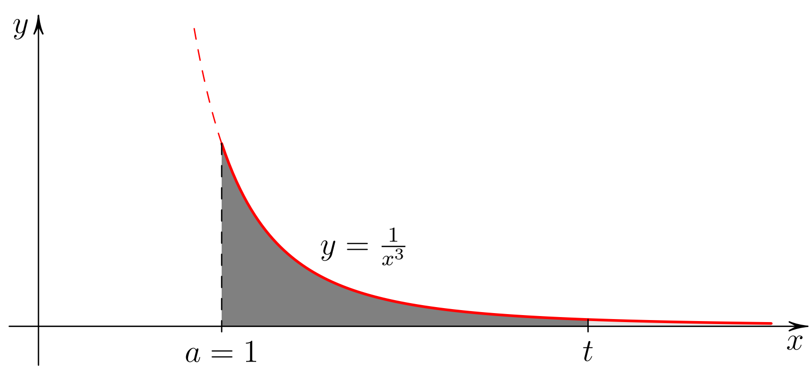
We can define definite integrals under these circumstances by considering suitable limits of integrals over closed intervals.

3.6.1 Improper Integrals of the First Kind

Suppose we wish to find the area under the curve $y = \frac{1}{x^3}$ over the interval $[1, \infty)$ shaded in the following diagram.



Intuitively one would find the area by evaluating the area under the curve (the definite integral) over the closed interval $[1, t]$, and then consider the limit of that as $t \to \infty$:



Should such a (finite) limit exist we would define that to be the area under the curve over the open interval $[1, \infty)$.

The previous discussion prompts the following definition for the improper integral over an infinite interval $[a, \infty)$ and, similarly, over intervals $(-\infty, b]$, and $(-\infty, \infty)$.

Definition: Define the following improper integrals of the first kind:

a) $\displaystyle\int_a^\infty f(x)\,dx = \lim_{t\to\infty}\int_a^t f(x)\,dx$ (Where the latter integrals must exist for every $t \geq a$.)

b) $\displaystyle\int_{-\infty}^b f(x)\,dx = \lim_{t\to-\infty}\int_t^b f(x)\,dx$ (Where the latter integrals must exist for every $t \leq b$.)

c) $\displaystyle\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^a f(x)\,dx + \int_a^\infty f(x)\,dx$ (Where a is any real number.)

The improper integrals in a) and b) are **convergent** if the limit exists (i.e. is finite) and **divergent** otherwise. For c) the integral is convergent if and only if both integrals on the right side are convergent.

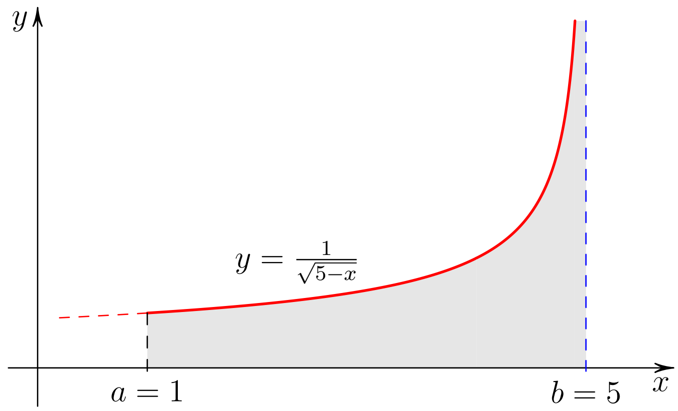
Note that $\displaystyle\int_{-\infty}^\infty f(x)\,dx$ is **not** defined to be $\displaystyle\lim_{t\to\infty}\int_{-t}^t f(x)\,dx$. The integral over $(-\infty,\infty)$ by definition must be broken into two pieces for which independent limits must be taken.

Examples:
Determine whether the following integrals converge or diverge. Find the value of any convergent integral.

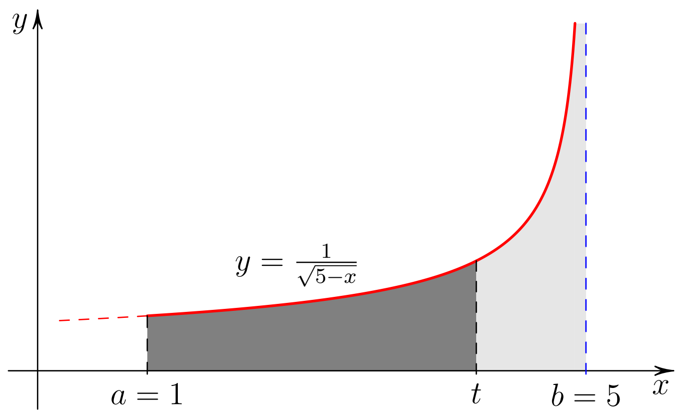
1. $\displaystyle\int_1^\infty \frac{1}{x^3}\,dx$
2. $\displaystyle\int_2^\infty \frac{1}{x-1}\,dx$
3. $\displaystyle\int_{-\infty}^0 xe^x\,dx$
4. $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx$
5. $\displaystyle\int_{-\infty}^0 xe^{-x^2}\,dx$
6. $\displaystyle\int_1^\infty \frac{\ln x}{x}\,dx$

3.6.2 Improper Integrals of the Second Kind

In the second case we consider those situations where the function being integrated has an infinite discontinuity at some point over which we want to integrate. Consider the area under the curve $y = \dfrac{1}{\sqrt{5-x}}$ between $x = 1$ and $x = 5$. The situation is shown in the following diagram.



The function has an infinite discontinuity at the right endpoint ($b = 5$). Intuitively we can imagine finding the area under the curve over the closed interval $[1, t]$ with $t < b$ and then consider the limit as $t \to b$:



This discussion suggests the following definition for improper integrals involving infinite integrands. Our example illustrated an integral where the right endpoint had the discontinuity. Similarly integrals with a discontinuity at the left endpoint or within the interval are defined.

Definition: Define the following improper integrals of the second kind:

a) Suppose $f(x)$ is continuous on $[a, b)$ but discontinuous at $x = b$ then:
$$\int_a^b f(x)\,dx = \lim_{t \to b^-} \int_a^t f(x)\,dx$$

b) Suppose $f(x)$ is continuous on $(a, b]$ but discontinuous at $x = a$ then:
$$\int_a^b f(x)\,dx = \lim_{t \to a^+} \int_t^b f(x)\,dx$$

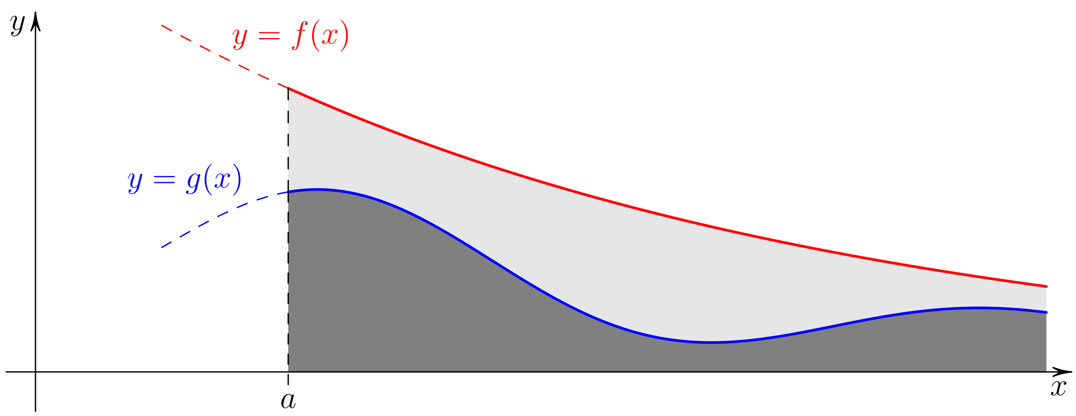
c) Suppose $f(x)$ is continuous on $[a, b]$ except at a value c in (a, b) then:
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

The improper integrals in a) and b) are **convergent** if the limit exists (i.e. is finite) and **divergent** otherwise. For c) the integral is convergent if and only if both integrals on the right side are convergent.

Examples:
Determine whether the following integrals converge or diverge. Find the value of any convergent integral.

1. $\displaystyle\int_1^5 \frac{1}{\sqrt{5-x}}\,dx$
2. $\displaystyle\int_0^1 x\ln x\,dx$
3. $\displaystyle\int_{-2}^7 \frac{1}{(x+1)^{\frac{2}{3}}}\,dx$
4. $\displaystyle\int_0^2 \frac{1}{x^2-4x+3}\,dx$

A consideration of the areas represented by improper integrals in the following diagram makes the following theorem plausible:



Theorem: 3.2. Let f and g be continuous functions satisfying $f(x) \geq g(x) \geq 0$ for all $x \geq a$. If $\int_a^\infty f(x)\,dx$ is convergent then $\int_a^\infty g(x)\,dx$ is convergent. If $\int_a^\infty g(x)\,dx$ is divergent then $\int_a^\infty f(x)\,dx$ is divergent.

Analagous theorems for the infinite intervals $(-\infty, b]$ and $(-\infty, \infty)$ as well as for improper integrals of the second kind may also be written. The theorems are useful for determining convergence or divergence of functions that are difficult to integrate.

Examples:
Determine whether the following integrals are convergent or divergent.

1. $\displaystyle\int_1^\infty \frac{1}{\sqrt{x^3+1}}\,dx$
2. $\displaystyle\int_0^1 \frac{e^{-x}}{x^{\frac{2}{3}}}\,dx$

Unit 4: Sequences and Series

4.1 Sequences

Definition: An ordered list of numbers:

$$\{a_1, a_2, a_3, \ldots, a_n, \ldots\}$$

is called a **sequence**. The numbers are called **terms** with a_1 here being the *first term*, and, more generally, a_n being the n^{th} term in the sequence.

The above sequence may represented by the compact notation $\boldsymbol{\{a_n\}}$ or sometimes with the index limits made explicit as $\boldsymbol{\{a_n\}_{n=1}^{\infty}}$. An explicit index is useful if we start enumerating the sequence from a value other than 1.

Some texts will distinguish **finite** and **infinite** sequences depending on whether the sequence terminates or not. For our purposes we will be assuming infinite sequences unless otherwise noted.

An equivalent way of thinking of a sequence is as a function f whose domain is the positive integers. In this case $a_n = f(n)$. Writing a_n as just such a function of the index is a convenient way of representing a sequence.

Examples:
The following are several ways to represent the same sequences.

1. $\left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right\} = \left\{\frac{1}{n}\right\} = \left\{\frac{1}{n}\right\}_{n=1}^{\infty}$, $a_n = \frac{1}{n}$
   Note this sequence could also have been represented by $\left\{\frac{1}{n+1}\right\}_{n=0}^{\infty}$
2. $\left\{\frac{1}{2}, \frac{-4}{5}, \frac{9}{8}, \ldots, \frac{(-1)^{n+1}n^2}{3n-1}, \ldots\right\} = \left\{\frac{(-1)^{n+1}n^2}{3n-1}\right\}_{n=1}^{\infty}$, $a_n = \frac{(-1)^{n+1}n^2}{3n-1}$
3. $\{4, 4, 4, \ldots, 4, \ldots\} = \{4\}_{n=1}^{\infty}$, $a_n = 4$

A sequence may not have a simple defining function in terms of the index n.

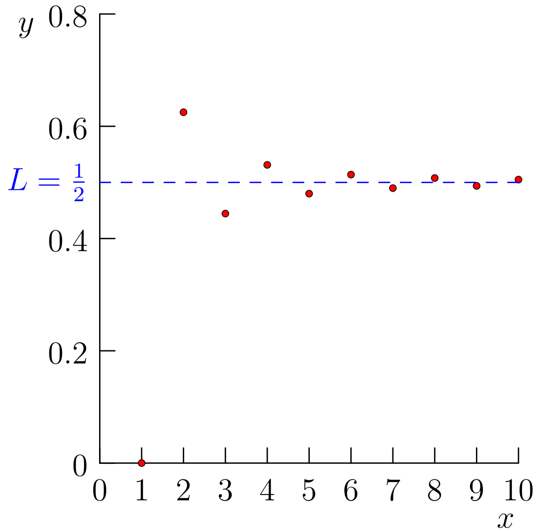
Example:
The sequence generated by the digits of $\pi = 3.14159\ldots$

$$\{3, 1, 4, 1, 5, 9, \ldots\}$$

is not representable by a simple function $f(n)$.

Since theoretically any sequence is a function $a_n = f(n)$ on the set of positive integers we can graphically represent it by plotting the coordinate points (n, a_n).

Example:
A graph of the sequence $\left\{\dfrac{\cos(n\pi) + n^2}{2n^2}\right\}$ is as follows:



The above graph clearly approaches the value $1/2$ as n gets large. In symbols we would write

$$\lim_{n\to\infty} \frac{\cos(n\pi) + n^2}{2n^2} = \frac{1}{2}$$

This limit is analagous to the limit of a function $f(x)$ as $x \to \infty$ with the only difference being that n is restricted to positive integers. This discussion motivates the following definition for the limit of a sequence.[1]

Definition: If the terms a_n of sequence $\{a_n\}$ get arbitrarily close to the value L for sufficiently large n then we say the sequence **converges** to **limit** $\boldsymbol{L}$ or **is convergent** with **limit** $\boldsymbol{L}$. Symbolically $a_n \to L$ as $n \to \infty$ or

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

If a sequence is not convergent (i.e. it has no limit) then the sequence **diverges** or **is divergent**.

A divergent sequence may have a trend to infinity.[2]

Definition: If the terms a_n of sequence $\{a_n\}$ get arbitrarily large (positively) for sufficiently large n we say that the sequence $\{a_n\}$ **diverges to infinity** and we write

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

An analagous definition holds for a sequence to diverge to $-\infty$.

Example:
The *Fibonacci Sequence* satisfies $a_1 = 1$, $a_2 = 1$ and $a_n = a_{n-2} + a_{n-1}$ for $n > 2$, i.e.

$$\{1, 1, 2, 3, 5, 8, 13, 21, \ldots\}$$

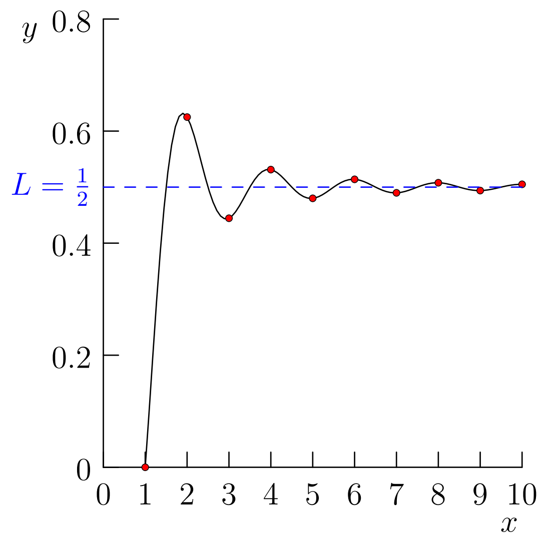
The sequence diverges to ∞ ($\lim_{n\to\infty} a_n = \infty$).

[1] A more rigorous definition of the limit of a sequence is that $a_n \to L$ as $n \to \infty$ if and only if for any $\epsilon > 0$ there exists $m > 0$ such that $n > m$ implies $|a_n - L| < \epsilon$.

[2] A more rigorous definition for $\lim\limits_{n\to\infty} a_n = \infty$ is that for any $M > 0$ there exists an index $m > 0$ such that $n > m$ implies $a_n > M$.

The limit of a sequence with $a_n = f(n)$ is essentially the limit of $f(x)$ as $x \to \infty$ with x restricted to the positive integers (instead of the continuous real axis).

Example:
If we plot $f(x) = \frac{\cos(x\pi)+x^2}{2x^2}$ over our earlier sequence we have:



The limit of $f(n)$, with n an integer, clearly cannot differ from that of $f(x)$ if the latter exists, thereby leading to the following theorem.

Theorem: 4.1. If $\lim\limits_{x\to\infty} f(x) = L$ then the limit of sequence $\{a_n\}$ with $a_n = f(n)$ is also L,

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

(Note the converse of this theorem is not true, $\lim_{n\to\infty} a_n = L \not\Rightarrow \lim_{x\to\infty} f(x) = L\ .$)

For a sequence which converges to $L = 0$ we have the following result:

Theorem: 4.2. $\lim\limits_{n\to\infty} a_n = 0$ if and only if $\lim\limits_{n\to\infty} |a_n| = 0$.

These theorems are convenient for evaluating the limits of certain sequences.

Example:
Find the limits of the following sequences:

1. $\left\{\dfrac{2n}{5n-3}\right\}$
2. $\left\{\dfrac{5n}{e^{2n}}\right\}$
3. $\left\{\dfrac{(-1)^n(3n+1)}{n^2+5}\right\}$

Theorem: 4.3. Given sequences $\{a_n\}$ and $\{b_n\}$ are convergent, c is a constant, and f a function defined at a_n and continuous at $L = \lim\limits_{n\to\infty} a_n$, then the following hold:

1. $\lim_{n\to\infty} c = c$
2. $\lim_{n\to\infty} (a_n \pm b_n) = \lim_{n\to\infty} a_n \pm \lim_{n\to\infty} b_n$
3. $\lim_{n\to\infty} ca_n = c \lim_{n\to\infty} a_n$
4. $\lim_{n\to\infty} (a_n \cdot b_n) = \left(\lim_{n\to\infty} a_n\right) \cdot \left(\lim_{n\to\infty} b_n\right)$
5. $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}$ (Here we require $\lim_{n\to\infty} b_n \neq 0$.)
6. $\lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right)$

By the last item of the previous theorem applied to $f(x) = x^k$ one has the corollary

Theorem: 4.4. Given non-negative sequence $\{a_n\}$ (i.e. $a_n \geq 0$) and power $k > 0$ one has

$$\lim_{n\to\infty} (a_n)^k = \left(\lim_{n\to\infty} a_n\right)^k .$$

Here the sequence must have $a_n \geq 0$ for $(a_n)^k$ to be defined and the power k cannot be negative to accommodate sequences for which $\lim_{n\to\infty} a_n = 0$.

The following theorem results from consideration of the behaviour of the limit of the exponential function $\lim_{x\to\infty} r^x$ when $r \geq 0$ in Theorem 4.1 and use of Theorem 4.2 noting that $\lim_{n\to\infty} |r^n| = \lim_{n\to\infty} |r|^n$ when $-1 < r < 0$.

Theorem: 4.5. The geometric sequence $\{ar^n\} = \left\{ar, ar^2, ar^3, \ldots, ar^n, \ldots\right\}$ $(a \neq 0)$ is convergent when $-1 < r \leq 1$ with

$$\lim_{n\to\infty} ar^n = \begin{cases} 0 & \text{if } -1 < r < 1 \\ a & \text{if } r = 1 \end{cases} .$$

It is divergent for all other values of r, diverging to infinity (∞) for $r > 1$.

Examples:
Determine whether the following sequences are divergent or convergent. For convergent sequences determine the limit.

1. $\left\{\left(\frac{n+1}{8n}\right)^{\frac{1}{3}}\right\}$
2. $\left\{5\left(\frac{1}{2}\right)^n\right\}$
3. $\{2^n\}$
4. $\left\{\left(-\frac{1}{3}\right)^n\right\}$
5. $\{(-3)^n\}$

A special class of sequences are those that are **monotonic**.

Definition: A sequence is **monotonic** if it is either

> **increasing** : $a_n < a_{n+1}$ for all n, or
>
> **decreasing** : $a_n > a_{n+1}$ for all n, or
>
> **nondecreasing** : $a_n \leq a_{n+1}$ for all n, or
>
> **nonincreasing** : $a_n \geq a_{n+1}$ for all n.

Note that increasing sequences are, by definition, nondecreasing as are decreasing sequences nonincreasing. An example of a nondecreasing sequence that is not an increasing sequence is

$$\{1, 1, 2, 2, 3, 3, 4, 4, \ldots\}$$

Examples:
Classify the monotonicity of the following sequences.

1. $\left\{\dfrac{2}{n+3}\right\}$
2. $\left\{\dfrac{2n+3}{3n+5}\right\}$
3. $\left\{\dfrac{3n+2}{2n+1}\right\}$

Definition: Sequence $\{a_n\}$ is **bounded below** if there exists some N such that $N \leq a_n$ for all n. The sequence $\{a_n\}$ is **bounded above** if there exists some M such that $a_n \leq M$ for all n. The sequence $\{a_n\}$ is **bounded** if it is bounded both below and above.

Examples:

1. The sequence $\left\{n^2\right\}$ is bounded below since $a_n > 0$ for all $n \geq 1$.
2. The sequence $\left\{\frac{n}{n+1}\right\}$ is bounded since $0 < a_n < 1$.

The following theorem can be used to determine whether a monotonic sequence converges or not.

Theorem: 4.6. A monotonic sequence converges if and only if it is bounded.

4.2 Series

Definition: Given a sequence $\{a_k\}$, the summation of its terms,

$$a_1 + a_2 + a_3 + \cdots + a_k + \cdots$$

is called an **(infinite) series**. A series is abbreviated $\sum_{k=1}^{\infty} a_k$ or sometimes without explicit index limits as $\sum a_k$.

Due to the sum being over an infinite number of terms it need not exist.

Example:
The series

$$\sum_{k=1}^{\infty} 1 = 1 + 1 + \cdots + 1 + \cdots$$

clearly cannot approach a number when added.

To rigorously define what we mean by the value of the sum of a series we introduce the following.

Definition: Given the series $\sum_{k=1}^{\infty} a_k$ define the sum of the first n terms of the series to be the $\boldsymbol{n}^{\text{th}}$ **partial sum $\boldsymbol{S_n}$**:

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

Example:
For the series $1 + 1 + 1 + \cdots + 1 + \cdots$ the n^{th} partial sums are

$$\begin{aligned} S_1 &= 1 \\ S_2 &= 1 + 1 = 2 \\ S_3 &= 1 + 1 + 1 = 3 \\ &\vdots \\ S_n &= \underbrace{1 + 1 + 1 + \cdots + 1}_{n \text{ times}} = n \end{aligned}$$

The partial sums S_n for a series $\sum a_k$ themselves form a sequence $\{S_n\}$ the limit of which we will consider the sum of the series.

Definition: Let series $\sum a_k$ have n^{th} partial sums S_n. If the sequence $\{S_n\}$ is convergent, so

$$\lim_{n\to\infty} S_n = S\ ,$$

then we say that the series $\sum a_k$ is **convergent** and call S the **sum** of the series,

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots + a_k + \cdots = S\ .$$

If a series is not convergent then it is **divergent**.

Example:

The series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{(k+1)(k+2)}$ can be written using partial fraction decomposition as $\displaystyle\sum_{k=1}^{\infty} \left(\frac{1}{k+1} - \frac{1}{k+2}\right)$. The n^{th} partial sum is therefore

$$S_n = \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n+1} - \frac{1}{n+2}\right) = \frac{1}{2} - \frac{1}{n+2}$$

Since $\displaystyle\lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(\frac{1}{2} - \frac{1}{n+2}\right) = \frac{1}{2}$ the series is convergent with sum $1/2$, i.e.

$$\sum_{k=1}^{\infty} \frac{1}{(k+1)(k+2)} = \frac{1}{2}\,.$$

Because of the cancellation arising in the partial sum the series is called a **telescoping series**.

Example:

We saw the series $\displaystyle\sum_{k=1}^{\infty} 1$ has partial sum $S_n = n$. Therefore $\displaystyle\lim_{n\to\infty} S_n = \lim_{n\to\infty} n = \infty$ and so the sequence $\{S_n\}$ and hence the series $\displaystyle\sum_{k=1}^{\infty} 1$ are divergent (as expected).

Theorem: 4.7. The **geometric series**

$$\sum_{k=1}^{\infty} ar^{k-1} = a + ar + ar^2 + \cdots + ar^k + \cdots = a\left(1 + r + r^2 + \cdots + r^k + \cdots\right)$$

is convergent if $-1 < r < 1$ with sum $\dfrac{a}{1-r}$ and is otherwise divergent.[3]

Examples:

Determine whether the following series are convergent and if so, find the sum.

1. $2 + \dfrac{2}{2} + \dfrac{2}{4} + \cdots + 2\left(\dfrac{1}{2}\right)^{n-1} + \cdots$
2. $\displaystyle\sum_{n=1}^{\infty} (3)^{n-1}$
3. $1 + x + x^2 + x^3 + \cdots \quad (\text{for } |x| < 1)$

A necessary (but not sufficient) requirement for a series to converge is that its terms must approach zero as $n \to \infty$.

Theorem: 4.8. If series $\displaystyle\sum_{k=1}^{\infty} a_k$ is convergent then $\displaystyle\lim_{k\to\infty} a_k = 0\,.$

[3]Proof follows by noting that

$$(1-r)S_n = a + ar + ar^2 + \cdots + ar^{n-1} - \left(ar + ar^2 + ar^3 + \cdots + ar^n\right) = a - ar^n,$$

and so $S_n = \frac{a(1-r^n)}{1-r}$. Then $S = \displaystyle\lim_{n\to\infty} S_n = \frac{a}{1-r}$ since $\displaystyle\lim_{n\to\infty} r^n = 0$ for $-1 < r < 1\,.$

Note the converse of the last theorem, namely that if $\lim\limits_{k\to\infty} a_k = 0$ then $\sum\limits_{k=1}^{\infty} a_k$ is convergent, is **not** true, as demonstrated in the following example.

> **Example:**
> The **harmonic series**,
> $$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} + \cdots$$
> is divergent. To see this note that the harmonic series
> $$1 \;+\; \frac{1}{2} \;+\; \frac{1}{3} + \frac{1}{4} \;+\; \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \;+\; \frac{1}{9} + \cdots$$
> is strictly greater than
> $$\underbrace{\frac{1}{1}}_{>\frac{1}{2}} + \underbrace{\frac{1}{2}}_{=\frac{1}{2}} + \underbrace{\frac{1}{4} + \frac{1}{4}}_{=\frac{1}{2}} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{=\frac{1}{2}} + \frac{1}{16} + \cdots\,,$$
> which diverges to infinity as one can exceed any multiple of 1/2 one wants by taking enough terms. Specifically, the partial sums of the harmonic series form an increasing sequence in which $S_{2^n} > \frac{1}{2}(n+1)$ and therefore the sequence $\{S_n\}$ is unbounded (and hence divergent).

The contrapositive of Theorem 4.8 (which logically must be true) provides a useful method to test if some series are divergent:

Theorem: 4.9. Term Test for Divergence:
If the terms a_k of series $\sum a_k$ approach a non-zero limit $\left(\lim\limits_{k\to\infty} a_k = L \neq 0\right)$ or $\lim\limits_{k\to\infty} a_k$ does not exist, then $\sum a_k$ is divergent.

> **Example:**
> Determine whether the series
> $$\frac{1}{3} + \frac{2}{5} + \frac{3}{7} + \cdots + \frac{k}{2k+1} + \cdots$$
> converges or diverges.

Note that if we find $\lim\limits_{n\to\infty} a_k = 0$ we know nothing about the convergence or divergence of series $\sum a_k$.

Theorem: 4.10. If c is any constant and $\sum\limits_{k=1}^{\infty} a_k$, $\sum\limits_{k=1}^{\infty} b_k$ are convergent series then the following series are convergent with the given results:

1. $\sum\limits_{k=1}^{\infty} c a_k = c \sum\limits_{k=1}^{\infty} a_k$
2. $\sum\limits_{k=1}^{\infty} (a_k \pm b_k) = \sum\limits_{k=1}^{\infty} a_k \pm \sum\limits_{k=1}^{\infty} b_k$

If $\sum_{k=1}^{\infty} a_k$ is divergent and $c \neq 0$ then $\sum_{k=1}^{\infty} ca_k$ is divergent. If one of $\sum_{k=1}^{\infty} a_k$, $\sum_{k=1}^{\infty} b_k$ is convergent and one is divergent then $\sum_{k=1}^{\infty}(a_k \pm b_k)$ is divergent.

Example:
Prove that the following series converges and find its sum.

$$\sum_{k=1}^{\infty}\left[\frac{2}{3^{k-1}} + \frac{7}{k(k+1)}\right]$$

Examples:
Determine whether each series is convergent or divergent. For convergent series, find the sum.

1. $\displaystyle\sum_{k=1}^{\infty} \frac{3k}{5k-1}$
2. $\displaystyle\sum_{k=1}^{\infty} k!$ Here the factorial $k! = k \cdot (k-1) \cdots (1)$ for $k \geq 1$ (and $0!$ is defined to be 1).
3. $\displaystyle\sum_{k=1}^{\infty}\left(\frac{1}{3^k} - \frac{1}{4^k}\right)$
4. $\displaystyle\frac{1}{2} + \frac{2}{3} + \cdots + \frac{n}{n+1} + \cdots$
5. $\displaystyle\sum_{k=1}^{\infty}\left[\left(\frac{3}{2}\right)^k + \left(\frac{2}{3}\right)^k\right]$
6. $\displaystyle\sum_{k=0}^{\infty} \frac{6^k}{7^{k+1}}$
7. $\displaystyle\sum_{n=1}^{\infty} \ln\left(\frac{2n}{3n-7}\right)$

4.3 Testing Series with Positive Terms

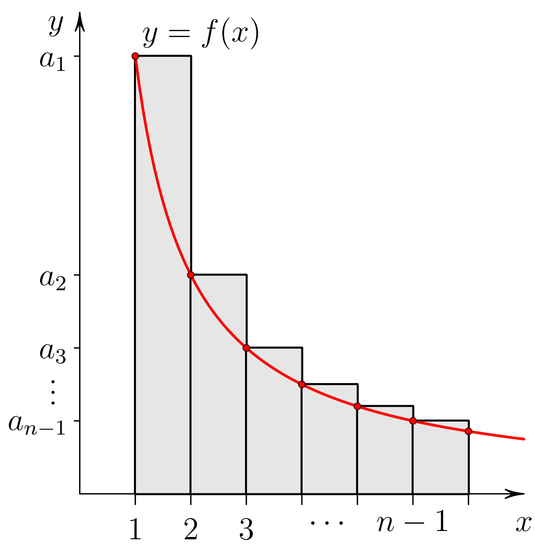
It is often difficult to find an exact sum of a series. In most cases a simple formula for the partial sum S_n cannot be found. It is therefore of interest to develop techniques to test whether a series is convergent or divergent. We start by considering series $\sum a_k$ with positive ($a_k > 0$) terms. Because the terms are positive the sequence of partial sums, $\{S_n\}$, is increasing.

4.3.1 The Integral Test

Suppose a series $\sum_{k=1}^{\infty} a_k$ has terms $a_k = f(k)$ written in terms of a function $f(x)$ that is continuous, positive, and decreasing for $x \geq 1$. The integral $\int_1^n f(x)\,dx$ will be smaller than the partial sum S_{n-1},

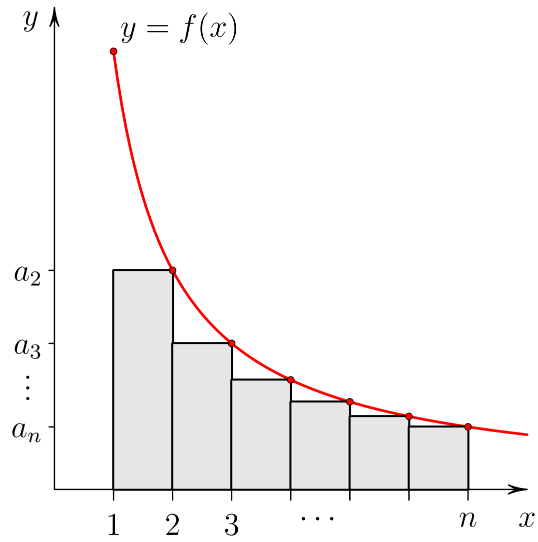
$$S_{n-1} = \sum_{k=1}^{n-1} a_k = a_1 + a_2 + \cdots + a_{n-1} = (a_1)(1) + (a_2)(1) + \cdots + (a_{n-1})(1)\,,$$

since the latter can be considered the total area of rectangles of height a_k and width 1 for $k = 1$ to $k = n-1$ as shown in the following diagram:



Here we are considering the a_k to be the height on the left side of the rectangles. Consider the case that $\int_1^{\infty} f(x)\,dx$ is divergent. Since $f(x)$ is positive, $\int_1^t f(x)\,dx$ is an increasing function of t and $\int_1^{\infty} f(x)\,dx = +\infty$. Suppose that increasing sequence $\{S_n\}$ were bounded with upper bound M. Then the relationship $\int_1^n f(x)\,dx < S_{n-1}$ implies that $\int_1^n f(x)\,dx < M$ for any integer n and therefore $\int_1^t f(x)\,dx < M$ for any real $t \geq 1$, a contradiction to the divergence of $\int_1^{\infty} f(x)\,dx$. Hence $\{S_n\}$ must be an unbounded monotonic sequence and therefore is divergent. Thus if $\int_1^{\infty} f(x)\,dx$ is divergent then $\sum a_k$ is divergent.

Alternatively if we consider rectangles with height being a_k on the right (so $k = 2$ to $k = n$) we have the following diagram:



In this case it follows that the integral $\int_1^n f(x)\,dx$ must be greater than:

$$(a_2)(1) + (a_3)(1) + \cdots + (a_n)(1) = a_2 + a_3 + \cdots + a_n = S_n - a_1$$

Consider the case where $\int_1^\infty f(x)\,dx$ is convergent. Let M be the value of the integral. Since $f(x)$ is positive, $\int_1^n f(x)\,dx < M$. From $S_n - a_1 < \int_1^n f(x)\,dx$ it follows that for any n, $S_n < M + a_1$ and thus monotonic sequence $\{S_n\}$ is bounded and therefore convergent. Thus if $\int_1^\infty f(x)$ is convergent, then $\sum a_k$ is convergent.

We summarize our result in the following theorem.

Theorem: 4.11. The Integral Test:

Let $f(x)$ be a continuous positive decreasing function for $x \geq 1$ and $\displaystyle\sum_{k=1}^{\infty} a_k$ be a series with $a_k = f(k)$.

1. If $\int_1^\infty f(x)\,dx$ is convergent then $\displaystyle\sum_{k=1}^{\infty} a_k$ is also convergent.
2. If $\int_1^\infty f(x)\,dx$ is divergent then $\displaystyle\sum_{k=1}^{\infty} a_k$ is also divergent.

Examples:
Determine whether each of the following series is convergent or divergent.

1. $\displaystyle\sum_{n=1}^{\infty} ne^{-n^2}$
2. $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ (The harmonic series)
3. $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^p}$ (The hyperharmonic or p-series)
4. $1 + \dfrac{1}{2^2} + \dfrac{1}{3^2} + \cdots + \dfrac{1}{n^2} + \cdots$

5. $\displaystyle\sum_{n=1}^{\infty} \frac{5}{\sqrt{n}}$

6. $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n^2}$

We summarize our results for the p-series in the following theorem.

Theorem: 4.12. The ***p*-series** $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^p}$ is convergent for $p > 1$ and divergent otherwise.

Notes on the Integral Test:

1. The Integral Test can be relaxed to consider a function $f(x)$ that is continuous positive and decreasing only on $x \geq n$ with the corresponding integral $\int_n^\infty f(x)\,dx$. This determines convergence or not of the series $\sum_{k=n}^{\infty} a_k$ but this in turn determines convergence of the entire series $\sum_{k=1}^{\infty} a_k$ since these two series differ only by a finite number of terms having a finite sum.

2. It follows from the Integral Test that the improper integral and the series either are both convergent or both divergent. This means that determining the convergence or not of a series can be used to determine the convergence properties of the improper integral of a continuous positive decreasing function $f(x)$ if that were desired.

Estimating the Series Sum

Even if a series is convergent it may be impossible to sum due to the impossibility of finding a closed form for the partial sum S_n for which we can take the limit. In that case one may resort to numerically calculating the partial sum S_n itself, for "large" n as an approximation for the sum of the series, $S \approx S_n$. The error in the approximation is the **remainder** $R_n = S - S_n$ which is the sum of the terms that were not included:

$$S = \sum_{k=1}^{\infty} a_k = \underbrace{a_1 + a_2 + \cdots + a_n}_{S_n = \sum_{k=1}^{n} a_k} + \underbrace{a_{n+1} + a_{n+2} + \cdots}_{R_n = \sum_{k=n+1}^{\infty} a_k} = S_n + R_n$$

For a convergent series $\sum a_k$ with $a_k = f(k)$ where $f(x)$ is a continuous positive decreasing function one can place bounds on the size of the remainder, thereby estimating the error in the numerical approximation. In our previous discussion we found that the n^{th} partial sum satisfied

$$S_n - a_1 < \int_1^n f(x)\,dx < S_{n-1}$$

In the case of convergence these inequalities imply

$$S - a_1 \leq \int_1^\infty f(x)\,dx \leq S\ .$$

If we start summing at the n^{th} term instead of the first this generalizes to

$$(a_n + a_{n+1} + a_{n+2} + \cdots) - a_n \leq \int_n^\infty f(x)\,dx \leq a_n + a_{n+1} + a_{n+2} + \cdots\ ,$$

from which it follows that

$$R_n \le \int_n^\infty f(x)\,dx \le R_{n-1}\ .$$

Thus $\int_n^\infty f(x)\,dx$ is an upper bound for the error R_n. The substitution $n-1 \to n$ for the inequality on the right implies $\int_{n+1}^\infty f(x)\,dx \le R_n$, thereby providing a lower bound on the remainder (error) as well. We summarize the result in the following theorem:

Theorem: 4.13. For convergent series $\displaystyle\sum_{k=1}^\infty a_k$ with $a_k = f(x)$ where $f(x)$ is a continuous positive decreasing function, the remainder $R_n = \displaystyle\sum_{k=n+1}^\infty a_k = S - S_n$ satisfies:

$$\int_{n+1}^\infty f(x)\,dx \le R_n \le \int_n^\infty f(x)\,dx\ .$$

Example:
Leonhard Euler was able to show the sum of the p-series with $p = 2$ is

$$\sum_{k=1}^\infty \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{25} + \cdots = \frac{\pi^2}{6} = 1.644934\ldots$$

1. Find S_4 .
2. Find the remainder R_4 .
3. Show R_4 falls within the bounds of the last theorem.
4. What partial sum S_n is required to be in error less than 0.01?

4.3.2 The Basic Comparison Test

We have seen several examples of series with their associated convergence properties:

geometric: $\displaystyle\sum_{k=1}^\infty ar^{k-1}$ is convergent for $|r| < 1$, divergent for $|r| \ge 1$

telescopic: $\displaystyle\sum_{k=1}^\infty \frac{1}{k(k+1)}$ (for example) is convergent

harmonic: $\displaystyle\sum_{k=1}^\infty \frac{1}{k}$ is divergent

***p*-series:** $\displaystyle\sum_{k=1}^\infty \frac{1}{k^p}$ is convergent for $p > 1$, divergent for $p \le 1$

We now develop some series convergence tests that use the known convergence properties of one series to determine that of another. The first test is a discrete analogue to our improper integral test found in Theorem 3.2.

Theorem: 4.14. Basic Comparison Test:
Let $\sum a_k$ and $\sum b_k$ be series with positive terms satisfying $a_k \leq b_k$ for all k.

1. If $\sum b_k$ is convergent then $\sum a_k$ is convergent.
2. If $\sum a_k$ is divergent then $\sum b_k$ is divergent.

Proof: Let $S_n = \sum_{k=1}^{n} a_k$ and $T_n = \sum_{k=1}^{n} b_k$ denote the partial sums of the series $\sum a_k$ and $\sum b_k$ respectively. Since $a_k \leq b_k$ it follows that $S_n \leq T_n$ for any n. Furthermore sequences $\{S_n\}$ and $\{T_n\}$ are both increasing (and hence monotonic) since terms a_k and b_k are positive.

Part 1 of the theorem follows from noting that monotonic sequence $\{T_n\}$ has upper bound $T = \sum_{k=1}^{\infty} b_k$ since the series $\sum b_k$, and hence the sequence $\{T_k\}$ converges. The sequence $\{S_n\}$ must also then have this upper bound since $S_n \leq T_n \leq T$. Thus $\{S_n\}$ is a monotonic bounded sequence and hence converges to S. Therefore $\sum a_k$ is convergent.

Part 2 follows from noting that if $\sum a_k$ is divergent then monotonic sequence $\{S_n\}$ is unbounded, which implies it has no upper bound as it is bounded below by 0. Now $S_n \leq T_n$ implies that monotonic sequence $\{T_n\}$ has no upper bound and hence does not converge, thereby proving $\sum b_k$ is divergent.

Example:
Determine whether the series converges or diverges.

1. $\displaystyle\sum_{k=1}^{\infty} \frac{1}{2+5^k}$
2. $\displaystyle\sum_{n=2}^{\infty} \frac{3}{\sqrt{n}-1}$
3. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n3^n}$

Remainder Estimate

Note that if one uses convergent series $\sum b_k$ to show $\sum a_k$ is convergent by the Basic Comparison Test, it follows, since $a_k \leq b_k$, that the remainder $R_n = \sum_{k=n+1}^{\infty} a_k$ is less than or equal to the remainder $\tilde{R}_n = \sum_{k=n+1}^{\infty} b_k$. Hence if we have an estimate for the size of the error $\tilde{R}_n \leq \epsilon$ for series $\sum b_k$, this implies $R_n \leq \epsilon$ for series $\sum a_k$.

4.3.3 The Limit Comparison Test

Our next test involves taking the limit of the ratio of terms of two series, one of whose convergence properties are presumed known.

Theorem: 4.15. The Limit Comparison Test:
Let $\sum a_k$ and $\sum b_k$ be series with positive terms.

1. If $\displaystyle\lim_{k\to\infty} \frac{a_k}{b_k} = L > 0$ then both series are convergent or both divergent.
2. If $\displaystyle\lim_{k\to\infty} \frac{a_k}{b_k} = 0$ and $\sum b_k$ is convergent then $\sum a_k$ is convergent.

3. If $\lim_{k\to\infty} \frac{a_k}{b_k} = \infty$ and $\sum b_k$ is divergent then $\sum a_k$ is divergent.

The convergence conclusions of the Limit Comparison Test can be remembered by noting that the limit condition effectively suggests that the tail of the series satisfies $a_k = Lb_k$, in other words the series tail is effectively a multiple of that of the other series by a constant $c = L$. For $c = L \neq 0$ we saw in Theorem 4.10 that the new series has the same convergence properties as the original.

Proof: Consider the case where $\lim_{k\to\infty} \frac{a_k}{b_k} = L > 0$. Then there exists $\tilde{M} > 0$ and $\tilde{N} > 0$ such that

$$\tilde{M} < \frac{a_k}{b_k} < \tilde{N} \quad \text{for } k > n$$

Let M be the minimum of the finite set of numbers $\left\{ \frac{a_k}{b_k} | k \leq n \right\}$ and the number $\tilde{M}$. Similarly let N be the maximum of the finite set of numbers $\left\{ \frac{a_k}{b_k} | k \leq n \right\}$ and the number $\tilde{N}$. It follows that for all k:

$$M < \frac{a_k}{b_k} < N$$

Since $b_k > 0$ we have, for all k,

$$Mb_k < a_k < Nb_k \ .$$

If $\sum b_k$ converges then $\sum Nb_k$ converges and $a_k < Nb_k$ implies $\sum a_k$ converges by the Basic Comparison Test. Similarly if $\sum b_k$ diverges then $\sum Mb_k$ diverges and $Mb_k < a_k$ implies $\sum a_k$ diverges by the Basic Comparison Test.

In the case $\lim_{k\to\infty} \frac{a_k}{b_k} = 0$ we can only argue that $a_k/b_k < N$ and so only $\sum b_k$ convergent implies $\sum a_k$ convergent. In the case $\lim_{k\to\infty} \frac{a_k}{b_k} = \infty$ we can only argue that $M < a_k/b_k$ and so only $\sum b_k$ divergent implies $\sum a_k$ divergent.

Example:
Determine whether the following series are convergent or divergent.

1. $\displaystyle\sum_{k=1}^{\infty} \frac{1}{\sqrt[3]{k^2+1}}$
2. $\displaystyle\sum_{n=1}^{\infty} \frac{3n^2+5n}{2^n\,(n^2+1)}$
3. $\displaystyle\sum_{k=1}^{\infty} \frac{1+2^k}{1+3^k}$
4. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2 \ln n}$

Note that since convergence is entirely determined by the infinite tail of a series, we can relax the Basic and Limit Comparison Tests to require that they only have positive terms for $k > n$ for some fixed n and, in the case of the Basic Comparison Test, that additionally $a_k \leq b_k$ for $k > n$.

4.4 The Alternating Series Test

We now consider series where all the terms are not positive. An alternating series is a special case of such a series.

Definition: An **alternating series** is a series of either of the forms

$$a_1 - a_2 + a_3 - \cdots + (-1)^{k-1} a_k + \cdots = \sum_{k=1}^{\infty} (-1)^{k-1} a_k \ ,$$

$$-a_1 + a_2 - a_3 + \cdots + (-1)^{k} a_k + \cdots = \sum_{k=1}^{\infty} (-1)^{k} a_k \ ,$$

where a_k is positive for all k.

Theorem: 4.16. Alternating Series Test:
If an alternating series of the form $\displaystyle\sum_{k=1}^{\infty}(-1)^{k-1}a_k$ or $\displaystyle\sum_{k=1}^{\infty}(-1)^{k}a_k$ with $a_k > 0$ satisfies

1. $a_{k+1} \le a_k$ for all k ,
2. $\displaystyle\lim_{k\to\infty} a_k = 0$,

then the alternating series is convergent.

Note that $a_{k+1} \le a_k$ is equivalent to $a_{k+1} - a_k \le 0$ and $\frac{a_{k+1}}{a_k} \le 1$.

Proof: Suppose we have an alternating series of the form $\sum_{k=1}^{\infty}(-1)^{k-1}a_k$. Consider the even partial sums S_2, S_4,..., where, in general the $(2n)^{\text{th}}$ partial sum is S_{2n} for n a positive integer given by:

$$S_{2n} = a_1 - a_2 + a_3 - a_4 + a_5 - a_6 \cdots + a_{2n-1} - a_{2n} \ .$$

Then grouping the terms in pairs one has

$$S_{2n} = (a_1 - a_2) + (a_3 - a_4) + (a_5 - a_6) \cdots + (a_{2n-1} - a_{2n}) \ ,$$

where each term in parentheses is nonnegative since $a_k \ge a_{k+1}$. This implies $\{S_{2n}\}$ is a nondecreasing sequence ($S_2 \le S_4 \le S_6 \le \ldots \le S_{2n} \le \ldots$). The terms of the even partial sum S_{2n} may be regrouped as

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n} \ ,$$

where, once again, the terms in parentheses are positive. This shows $S_{2n} < a_1$ for all n. The monotonic sequence of even partial sums $\{S_n\}$ is bounded and hence has limit S. The odd partial sums are S_{2n+1} for n a positive integer and may be written

$$S_{2n+1} = S_{2n} + a_{2n+1} \ .$$

Taking the limit of the odd partial sum sequence $\{S_{2n+1}\}$ gives

$$\lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} (S_{2n} + a_{2n+1}) = \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} a_{2n+1} = S + 0 = S$$

Since both even and odd partial sum sequences approach the same limit S, the sequence $\{S_n\}$ approaches S as well and alternating sequence $\sum_{k=1}^{\infty}(-1)^{k-1}a_k$ is convergent. Since alternating sequence $\sum_{k=1}^{\infty}(-1)^{k}a_k = (-1)\sum_{k=1}^{\infty}(-1)^{k-1}a_k$ this completes the proof for the other possible case.

Example:
The alternating harmonic series

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

is convergent since $a_k = \frac{1}{k}$ satisfies $a_{k+1} = \frac{1}{k+1} \leq \frac{1}{k} = a_k$ and $\lim_{k\to\infty} a_k = \lim_{k\to\infty} \frac{1}{k} = 0$.

Example:
Determine whether the following alternating series converge or diverge.

1. $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{2n}{4n^2 - 3}$
2. $\displaystyle\sum_{k=1}^{\infty} (-1)^{k} \frac{2k}{4k - 3}$
3. $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{\ln n}{n}$
4. $\displaystyle\sum_{k=1}^{\infty} \cos(k\pi) \frac{3k^2 + 2}{2k^2 + 1}$

Remainder Estimate

An estimate of the error made when approximating the sum S of an alternating series with the n^{th} partial sum S_n is given by the following theorem.

Theorem: 4.17. For an alternating series with terms of absolute value $a_k > 0$ the remainder $R_n = S - S_n$ satisfies $|R_n| < a_{n+1}$.

4.5 Tests of Absolute Convergence

For series $\sum a_k$ whose terms have mixed sign, one can consider the convergence properties of the series $\sum |a_k|$ with nonnegative terms generated by taking the absolute value of the terms of the original series.

4.5.1 Absolute Convergence

Definition: A series $\sum a_k$ is **absolutely convergent** if the series $\sum |a_k|$ is convergent.

A series may be convergent that is not absolutely convergent, prompting the following definition.

Definition: A series $\sum a_k$ that is convergent but not absolutely convergent is called **conditionally convergent**.

> **Example:**
> The alternating harmonic series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ is conditionally convergent because it converges but the harmonic series, which is the sum of the absolute values of its terms, does not.

The following theorem shows a convergent series can only be absolutely or conditionally convergent.

Theorem: 4.18. If a series $\sum a_k$ is absolutely convergent then it is convergent.

The theorem also shows that convergence of some series $\sum a_k$ may be determined by considering convergence of $\sum |a_k|$.

> **Example:**
> Determine whether the following series are absolutely convergent or conditionally convergent.
>
> 1. $\displaystyle\sum_{k=1}^{\infty}(-1)^{k-1}\frac{1}{k^2}$
> 2. $\displaystyle\sum_{n=1}^{\infty}\frac{\sin n}{n^2}$
> 3. $\displaystyle\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{\sqrt{k}}$

4.5.2 The Ratio Test

The following convergence test considers the limit of the ratio of terms within a series.

Theorem: 4.19. The Ratio Test:

Suppose the ratio of consecutive terms of series $\displaystyle\sum_{k=1}^{\infty} a_k$ has limit

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = L\ ,$$

then

1. If $L < 1$ the series $\sum a_k$ is absolutely convergent (and hence convergent).
2. If $L = 1$ the test is inconclusive.
3. If $L > 1$ the series $\sum a_k$ is divergent.

If $\lim\limits_{k\to\infty} \left|\dfrac{a_{k+1}}{a_k}\right| = \infty$ then $\sum a_k$ is also divergent.

Note that if the Ratio Test is inconclusive ($L = 1$) this means that $\sum a_k$ is potentially absolutely convergent, conditionally convergent, or divergent.

The convergence conclusions of the Ratio Test can be remembered by noting that the limit condition effectively suggests that the tail of the series satisfies $a_{k+1} = La_k$, in other words it behaves like a geometric series with $r = L$. From this it follows $r = L < 1$ should converge and $r = L > 1$ should diverge.

Example:
Determine whether the following series are absolutely convergent, conditionally convergent, or divergent.

1. $\displaystyle\sum_{k=1}^{\infty}(-1)^k\frac{3^k}{k!}$
2. $\displaystyle\sum_{n=1}^{\infty}\frac{4^n}{n^2}$
3. $\displaystyle\sum_{k=1}^{\infty}e^{-k}k!$
4. $\displaystyle\sum_{n=1}^{\infty}\frac{2^n}{2n^2+1}$

4.5.3 The Root Test

The next convergence test considers the limit of the k^{th} root of $|a_k|$.

Theorem: 4.20. The Root Test:
Suppose the terms of series $\displaystyle\sum_{k=1}^{\infty} a_k$ satisfy

$$\lim_{k\to\infty} \sqrt[k]{|a_k|} = L \ ,$$

then

1. If $L < 1$ the series $\sum a_k$ is absolutely convergent (and hence convergent).
2. If $L = 1$ the test is inconclusive.
3. If $L > 1$ the series $\sum a_k$ is divergent.

If $\lim\limits_{k\to\infty} \sqrt[k]{|a_k|} = \infty$ then $\sum a_k$ is also divergent.

The convergence conclusions of The Root Test can be remembered by noting that the limit condition effectively suggests that the tail of the series behaves like $a_k = L^k$, in other words like a geometric series with $r = L$. From this it follows that $r = L < 1$ should converge and $r = L > 1$ should diverge.

Example:
Determine the convergence or divergence of the following series.

1. $\displaystyle\sum_{k=1}^{\infty} \frac{2^{3k+1}}{k^k}$
2. $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n}{(\ln n)^n}$

4.5.4 Rearrangement of Series

For a finite summation of terms the order in which we add the numbers does not matter, i.e. $1+4+2 = 4+1+2$, or, in symbols $a_1+a_2+a_3 = a_2+a_1+a_3$. The second summation is a **rearrangement** of the first. To make the idea precise we note that the indices on the second summation, $(2, 1, 3)$ are a **permutation** of those on the first $(1, 2, 3)$. A permutation on the infinite set of positive indices $(1, 2, 3, \ldots)$ of a series can similarly be defined thereby making the intuitive definition of a **rearrangement of a series** precise.

If a series is absolutely convergent then we get the same sum regardless of the order in which the terms are added (as expected), as summarized in the following theorem.

Theorem: 4.21. Any rearrangment of absolutely convergent series $\sum a_k$ has the same sum as the original series.

However for a series that is only conditionally convergent the order in which we add the terms does matter. Indeed we get the following remarkable result

Theorem: 4.22. Riemann Rearrangement Theorem: Let $\sum a_k$ be a conditionally convergent series with sum S. Then for any real number R there exists a rearrangement of series $\sum a_k$ having sum R. Additionally there exist rearrangements of series $\sum a_k$ which diverge to $+\infty$, $-\infty$, and which fail to approach any limit, finite or infinite.

Example:
By the latter theorem it follows that the (conditionally convergent) alternating harmonic series can be rearranged to sum to any number or to diverge.

4.6 Procedure for Testing Series

We have seen several methods for testing for the convergence and divergence of series. The form of the series should suggest the type of test to be used. The following steps will be helpful for determining convergence and divergence of series.

1. Recognize known series with associated convergence and divergence properties:

   geometric series: $\displaystyle\sum_{k=1}^{\infty} ar^{k-1} = \sum_{k=0}^{\infty} ar^k$ is convergent for $|r| < 1$ and otherwise divergent.

   ***p*-series:** $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^p}$ is convergent for $p > 1$ and otherwise divergent.
   (Note that $p = 1$ is $\sum 1/k$ the (divergent) **harmonic** series.)

2. If $\lim_{k\to\infty} a_k \neq 0$ or that limit does not exist then the series is divergent by the **Term Test for Divergence**.

3. If $\lim_{k\to\infty} a_k = 0$ then proceed as follows:

   (a) If the terms of the series are **positive**, use one of the following tests.

   Basic Comparison Test: Useful when a_k is a rational or algebraic function of k (i.e. involving roots of polynomials). Consider a suitable geometric or p-series for comparison. Remember any comparison series must be positive.

   Limit Comparison Test: Same criteria as the Basic Comparison Test. Choose this one if evaluating the limit of the ratio of comparing terms is easier than proving an inequality between them as required in the Basic Comparison Test.

   Ratio Test: Useful for series involving factorials or other products (including a constant raised to power k). Do not use this test for rational or algebraic functions of k as these result in inconclusive ($L = 1$) results.

   Root Test: Useful if a_k may be written $a_k = (b_k)^k$.

   Integral Test: Useful if $a_k = f(k)$ for positive, continuous, decreasing $f(x)$ and $\int_1^\infty f(x)\,dx$ is easily evaluated.

   (b) If the series is **alternating** (either $\sum(-1)^k a_k$ or $\sum(-1)^{k-1} a_k$ for $a_k > 0$) either:

   i. Use the **Alternating Series Test**.

   ii. Apply a positive series test from 3(a) above to the absolute value of the alternating series (either $\sum \left|(-1)^k a_k\right| = \sum a_k$ or $\left|\sum(-1)^{k-1} a_k\right| = \sum a_k$) since the convergence of the latter implies the convergence of the alternating series.

   (c) If the terms of the series $\sum a_k$ are neither positive nor alternating, apply a test from 3(a) above to the absolute value of the series, $\sum |a_k|$. If $\sum |a_k|$ is convergent then $\sum a_k$ is also convergent.

   (d) If the series only satisfies theorem criteria (positivity, decreasing, alternating, etc.) after a certain point (i.e. for $k \geq n$ for some n) apply the above steps to the tail series $\sum_{k=n}^{\infty} a_k$. The convergence or divergence of the entire series will be the same as that of the tail series since they only differ by a finite number of terms of finite sum.

Examples:
Determine the convergence (absolute or conditional) or divergence of the following series.

1. $\displaystyle\sum_{k=1}^{\infty} \frac{2k^2}{k^2+1}$

2. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{n}{n^2+1}$

3. $\displaystyle\sum_{n=1}^{\infty} \frac{1000-n}{n!}$

4. $\displaystyle\sum_{k=1}^{\infty} e^{-2k}$

5. $\displaystyle\sum_{n=1}^{\infty} \frac{5^n}{n^6\, 3^{n+1}}$

6. $\displaystyle\sum_{k=1}^{\infty} \sin\left(\frac{\pi}{2}+\frac{1}{k}\right)$

7. $\displaystyle\sum_{k=1}^{\infty} \frac{k^k}{10^k}$

8. $\displaystyle\sum_{n=1}^{\infty} (-1)^{n-1} \frac{\sqrt{n}}{n^2+1}$

9. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{3^n}+\frac{5}{\sqrt{n}}\right)$

4.7 Power Series

We have seen that in many cases the terms of the series $\sum a_k$ may be written as a function of the summation index, namely $a_k = f(k)$, for some function f. However if the series $\sum a_k$ is convergent its value is a number, namely its sum. The final sum does not depend on the index k in the same way that a definite integral $\int_a^b f(t)\,dt$ results in a number independent of the dummy variable t.

Consider, however, the situation where the terms a_k depend additionally on an actual variable, say x, different from the summation index, present in the series (i.e. $a_k = a_k(x)$). In this case the sum of the series (and indeed its convergence) depends on the value of x.

Example:
The geometric series with $a = 1$ and $r = x$ is given by

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots + x^k + \cdots$$

Here $a_k(x) = x^k$. The sum is now a function of x, namely $\dfrac{1}{1-x}$, and is valid for $|x| < 1$ for which the series is convergent.

We could introduce the variable x into the terms of a series in many ways, for instance

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k!}\,.$$

If we choose to introduce it as in our geometric series above, namely so that the terms of the series look like terms in a polynomial, $c_k x^k$, we have a **power series**.

Definition: Let x be a variable. A series of the form

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

where c_k are real constants (for $k = 0, 1, 2 \ldots$) is a **power series in $\boldsymbol{x}$**. The constants c_k are called the **coefficients** of the series.

Example:
The geometric series with $r = x$ above, $1 + x + x^2 + \cdots + x^k + \cdots$, is a power series in x with coefficients $c_k = 1$ for all k.

If we choose to make the k^{th} term of the series have the more general form $c_k(x-a)^k$ we get the following.

Definition: Given real constant coefficients c_k and real constant a the **power series in $\boldsymbol{(x-a)}$** is

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

The series is also known as the **power series about $\boldsymbol{a}$** or **centred on $\boldsymbol{a}$**.

A power series in x is just a special case of this last definition with $a = 0$.

Since the convergence of a power series (and indeed its sum should it converge) will, in general, depend on the value of the variable x, an obvious question is to find the values of x for which the power series is convergent.

> **Example:**
> Find the values of x for which the following power series are convergent.
>
> 1. $\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$
> 2. $\displaystyle\sum_{n=0}^{\infty} \frac{n^2}{2^n} x^n$
> 3. $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{e^n} (x-e)^n$

As suggested by the previous example a power series about a will converge on an interval centred on a as detailed in the following theorem.

Theorem: 4.23. The power series $\displaystyle\sum_{k=0}^{\infty} c_k (x-a)^k$ will either:

1. Converge only at $x = a$.
2. Converge for $|x-a| < R$ and diverge for $|x-a| > R$ for some positive real number R.
3. Converge for all x.

Definition: The **radius of convergence R** for a power series $\sum_{k=0}^{\infty} c_k (x-a)^k$ is the value R if Part 2 of Theorem 4.23 applies. For Part 1 the radius of convergence is defined to be $R = 0$ and for Part 3 the radius of convergence is defined to be $R = \infty$.

Definition: The **interval of convergence I** of a power series $\sum_{k=0}^{\infty} c_k (x-a)^k$ is the set of values of x for which the series converges. For the three possibilities of convergence one has:

1. The set containing the single value $x = a$ (i.e. $\{a\}$) for $R = 0$.
2. One of $[a-R, a+R]$, $[a-R, a+R)$, $(a-R, a+R]$, or $(a-R, a+R)$ for $R > 0$ finite.
3. $(-\infty, \infty)$ for $R = \infty$.

The choice of interval in the second case depends upon the convergence or not of the series at the interval endpoint values $x = a \pm R$.

> **Example:**
> For the previous example one has the following radii and intervals of convergence:
>
> 1. $R = \infty$, $I = (-\infty, \infty)$
> 2. $R = 2$, $I = (-2, 2)$
> 3. $R = e$, $I = (0, 2e)$

Example:
Find the radius and interval of convergence of each of the following series.

1. $\displaystyle\sum_{k=0}^{\infty}(-1)^k \frac{1}{k+1}(x-3)^k$

2. $\displaystyle\sum_{n=0}^{\infty} n^3(x-5)^n$

3. $\displaystyle\sum_{k=0}^{\infty} k!(2x-1)^k$

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

4.8 Representing Functions with Power Series

The sum of a power series for a given x from its interval of convergence I results in a number. As such it is natural to consider the power series as defining a function $f(x)$ of x on I with the value of the function being the sum. We write

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \quad (x \in I)$$

If the sum of the power series can be written in a closed form, then the power series can be considered a representation of that function valid on I.

Example:
The power series $\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots$ converges for $|x| < 1$, and therefore is a function of x on $I = (-1, 1)$. On that interval the sum for given x is $\frac{1}{1-x}$. The power series thus is a representation of the function $\frac{1}{1-x}$ on this restricted domain:

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^k + \cdots = \sum_{k=0}^{\infty} x^k \quad \text{for } |x| < 1$$

One can find power series representations of other functions using known power series.

Example:
Find representations of the following functions with power series. Indicate the values of x for which the representations are valid.

1. $\dfrac{1}{1+x}$
2. $\dfrac{1}{1-x^3}$
3. $\dfrac{x}{2-3x}$

As functions power series behave like polynomials. They are continuous and can be termwise differentiated and integrated to produce new functions of x as detailed in the following theorem.

Theorem: 4.24. Suppose $f(x)$ is defined by the power series

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

with radius of convergence $R > 0$ (finite or infinite). Then on the interval $(a - R, a + R)$:

1. $f(x)$ is continuous.
2. $f(x)$ is differentiable with derivative

$$f'(x) = \sum_{k=0}^{\infty} k c_k (x-a)^{k-1} = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots + k c_k (x-a)^{k-1} + \cdots$$

3. $f(x)$ is integrable with integral

$$\int f(x)\,dx = C+\sum_{k=0}^{\infty}\frac{c_k}{k+1}(x-a)^{k+1} = C+c_0(x-a)+\frac{c_1}{2}(x-a)^2+\frac{c_2}{3}(x-a)^3+\cdots+\frac{c_k}{k+1}(x-a)^{k+1}+\cdots$$

The series resulting from differentiation and integration both have radius of convergence R.

Note that the above theorem shows that for power series the calculus operations of differentiation and integration can be exchanged with summation, just as occurs with finite sums.

$$\frac{d}{dx}\sum c_k(x-a)^k = \sum \frac{d}{dx}\left[c_k(x-a)^k\right]$$

$$\int\left[\sum c_k(x-a)^k\right]\,dx = \sum\left[\int c_k(x-a)^k\,dx\right]$$

If the power series is a representation of a function with a closed form, differentiation and integration can be used to find power series representations of other functions.

Example:
Find a power series representation for each of the following functions. Indicate the interval for which the representation is valid.

1. $\dfrac{1}{(1+x)^2}$
2. $\ln(1+x)$
3. $\tan^{-1}x$

Because power series representations are easy to integrate, they may be used as a means to integrate difficult functions. The integral will only be valid for x lying within $(a-R, a+R)$.

Example:
Integrate each of the following using a power series.

1. $\displaystyle\int \frac{1}{1+x^3}\,dx$
2. $\displaystyle\int \frac{x}{1-x^4}\,dx$

4.9 Maclaurin Series

We expect some functions $f(x)$ can be represented by an infinite series $\sum_{k=0}^{\infty} c_k x^k$ for a certain, potentially restricted, domain. While we verified that $\frac{1}{1-x}$ could be represented by the geometric series, we would like a general mechanism for determining the coefficients of the power series that correspond to an arbitrary $f(x)$. Suppose $f(x)$ has a power series expansion:

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + \cdots .$$

One observes that if we set $x = 0$ that $f(0) = c_0$. In other words, we can determine c_0 by evaluating f at $x = 0$. We can differentiate the power series above term by term to get:

$$f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + 4c_4 x^3 + \cdots$$

If we evaluate this at $x = 0$ we get $c_1 = f'(0)$. The next derivative is:

$$f''(x) = 2c_2 + 6c_3 x + 12c_4 x^2 + \cdots$$

and so $c_2 = \frac{f''(0)}{2}$. Repeated differentiation and evaluation relates the k^{th} coefficient c_k to the derivative $f^{(k)}$ evaluated at $x = 0$ as follows:

$$c_k = \frac{f^{(k)}(0)}{k!}$$

where recall the **factorial** is defined by $\boldsymbol{k! = k \cdot (k-1) \cdots 2 \cdot 1}$. The formula is true for $k = 0$ as well with the convention $f^{(0)}(x) = f(x)$ and noting that $\mathbf{0! = 1}$ by definition. Plugging our c_k into the original power series we get the following definition.

Definition: Given function $f(x)$ differentiable to all orders at $x = 0$ the power series in x given by

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

is the **Maclaurin series** for $f(x)$.

As a power series, the Maclaurin series must converge for $|x| < R$ for some radius of convergence R dependent upon the function.[4]

Examples:
Find the Maclaurin series and its interval of convergence for each of the following functions.

1. $f(x) = \dfrac{1}{1-x}$
2. $f(x) = e^x$
3. $f(x) = \sin x$
4. $f(x) = \ln(1+x)$

[4] We will discuss shortly whether the Maclaurin series of $f(x)$ actually is a valid representation of the function on this interval.

Some important Maclaurin series and their domains of validity are:

$$\begin{aligned}
\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + \cdots \quad ; \ (-1,1) \\
e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad ; \ (-\infty,\infty) \\
\sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \quad ; \ (-\infty,\infty) \\
\cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \quad ; \ (-\infty,\infty) \\
\tan^{-1} x &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \quad ; \ [-1,1] \\
\ln(1+x) &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \quad ; \ (-1,1]
\end{aligned}$$

As we saw with the more general power series, we can derive series of new functions from known Maclaurin series.

Example:
For the following functions we have the power series:

1. $e^{x^3} = \sum_{k=0}^{\infty} \frac{\left(x^3\right)^k}{k!} = \sum_{k=0}^{\infty} \frac{x^{3k}}{k!}$

   Series is valid for x^3 in $(-\infty,\infty)$ which implies for x in $(-\infty,\infty)$.

2. $\sin^2 x = \frac{1}{2}(1-\cos 2x) = \frac{1}{2} - \frac{1}{2}\sum_{k=0}^{\infty}(-1)^k \frac{(2x)^{2k}}{(2k)!} = \frac{1}{2} + \sum_{k=0}^{\infty}(-1)^{k+1}\frac{2^{2k-1}x^{2k}}{(2k)!}$

   Series is valid for $2x$ in $(-\infty,\infty)$ and so for x in $(-\infty,\infty)$.

These series can be confirmed to be the Maclaurin series of the given functions by direct computation.

4.10 Taylor Series

Maclaurin series can be generalized by expanding in powers of $(x-a)^k$ rather than x^k for some constant a. A similar argument to that before gives the following definition.

Definition: Given function $f(x)$ differentiable to all orders at $x = a$ the power series in $(x-a)$ given by

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

is the **Taylor Series** for the function $f(x)$ **at a** (or about a or centred on a).

The Taylor series expansion at a will converge within some radius R about a, i.e. for $|x-a| < R$.

We note that Maclaurin series is just a special case of the Taylor series when $a = 0$.

Examples:
Find the Taylor Series of the given function at the specified value and determine the interval of convergence.

1. $f(x) = \sin x$ at $a = \dfrac{\pi}{2}$
2. $f(x) = \ln x$ at $a = 1$
3. $f(x) = e^x$ at $a = 2$

Assuming convergence to $f(x)$, Taylor series gives us a mechanism for calculating trigonometric and other functions, namely by evaluating the first n terms of the series at x. How many terms of the series are required for a good approximation will depend on the function, the value a about which it is expanded, and x.

Tayor series allows expansion of the function f about values other than $a = 0$ which is useful for functions that are not defined at 0. Also in general fewer terms of the expansion will be required for a good approximation if the Taylor series is generated about a value a near the x of interest. Indeed by truncating the Taylor series at the $k = 1$ term

$$f(x) \approx f(a) + f'(a)(x-a) \ ,$$

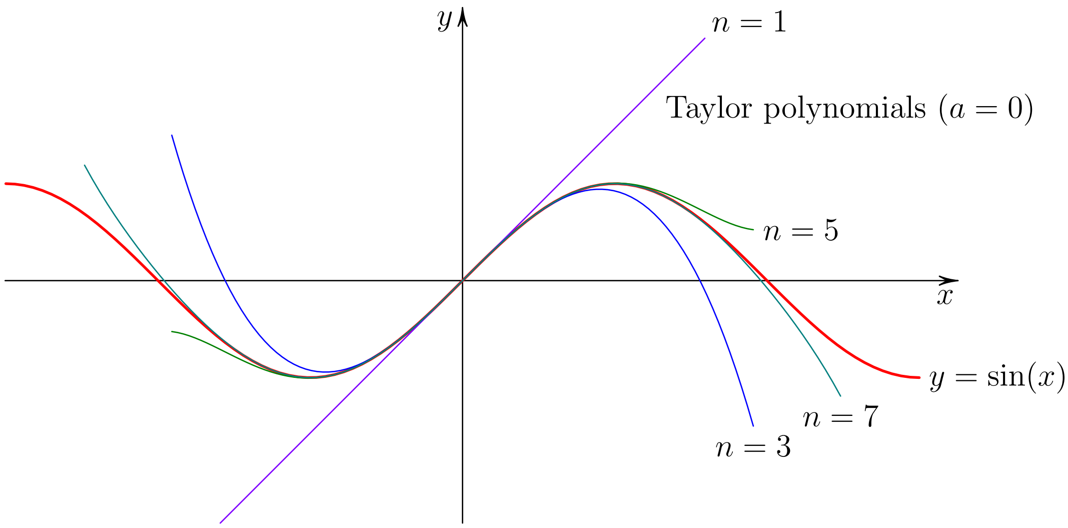
or at the $k = 2$ term,

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 \ ,$$

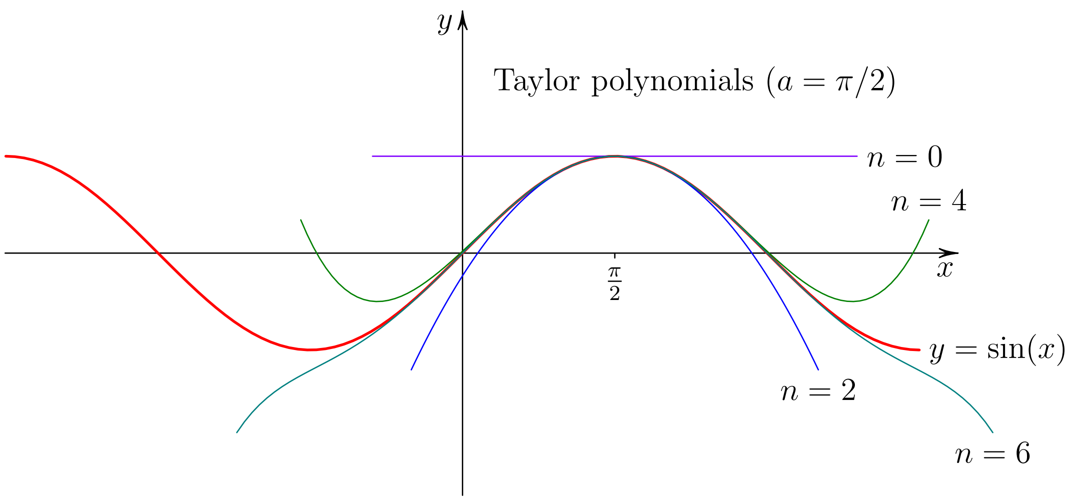
we are just reproducing the linear and quadratic approximations of a function at $x = a$ arrived at our previous course. In general truncating at the $k = n^{th}$ term,

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

yields increasingly better approximations. The right hand side is called the n^{th} **Taylor polynomial** of the function $f(x)$ at a. The following diagram shows the first four Taylor polynomials of $\sin x$ at $a = 0$ (i.e. the Maclaurin series):



Here are the first four polynomials for the Taylor series of $\sin x$ at $a = \pi/2$:



One observes both the greater accuracy of the higher order approximations as well as the utility of expanding the Taylor series at a value a near the x at which you wish to approximate the function.

One useful result of Theorem 4.24 is that it justifies our original proof for the coefficients for the Maclaurin series where (recall) we required the power series to be differentiable. Generalizing this result to Taylor series we have the following result:

Theorem: 4.25. If function $f(x)$ is represented by the power series

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

on an open interval containing a then the coefficients are the Taylor series coefficients $c_k = f^{(k)}(a)/k!$, (i.e. $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$).

In other words, if a function $f(x)$ has a power series that converges to $f(x)$ it will be the Taylor series. The theorem does not say, however, that the Taylor series at a for a given function f will necessarily converge to f. For instance define the continuous piecewise function:

$$f(x) = \begin{cases} 0 & \text{if } x < -\pi/2 \\ \cos x & \text{if } -\pi/2 \leq x \leq \pi/2 \\ 0 & \text{if } x > \pi/2 \end{cases}$$

Then f will have the same Taylor series at $a = 0$ (i.e. Maclaurin series) as the function $\cos x$. However that series clearly cannot represent both functions. One must determine that the Taylor series for f at a really does converge to the function.

Theorem: 4.26. If a function f has derivatives to all orders in an interval centred on a, then

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

will hold on the interval if and only if

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

for all x in the interval, where

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

Unit 5: Integration Applications

5.1 Areas Between Curves

We have already seen this topic in a previous course. We review it here for the purpose of reminding ourselves how to interpret differentials in integration formulae.

We have already calculated an area between curves. Namely, when $y = f(x) \geq 0$ on $[a, b]$ then

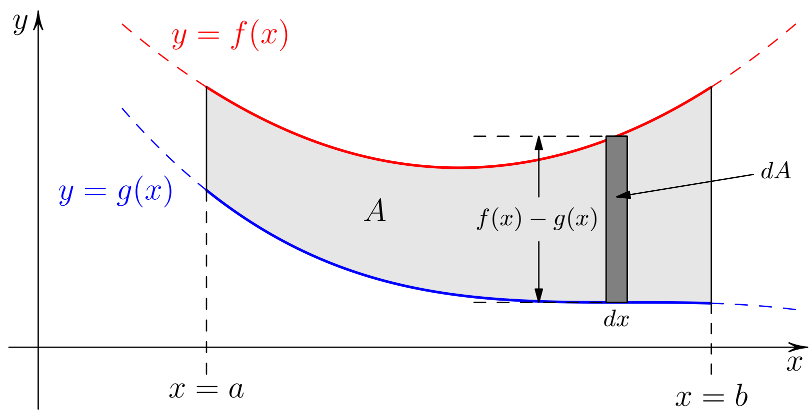
$$A = \int_a^b f(x)\,dx$$

is the area below the curve $y = f(x)$ and above the x-axis. But since the y-axis is just the curve $y = 0$, the above integral is also the area between the curve $y = f(x)$ and the curve $y = 0$. If the lower curve is $y = g(x)$ instead of $y = 0$ we get the following result:

Theorem: 5.1. If $y = f(x)$ and $y = g(x)$ are integrable on the x-interval $[a, b]$ with $f(x) \geq g(x)$ on the interval then the **area between the curves** $y = f(x)$ and $y = g(x)$ and the lines $x = a$ and $x = b$ is

$$A = \int_a^b [f(x) - g(x)]\ dx\ .$$

The situation is illustrated in the following diagram:



A simple way to remember the theorem is the differential notation. We are finding the sum $A = \int dA$ of the infinitesimal rectangle areas each of area dA where

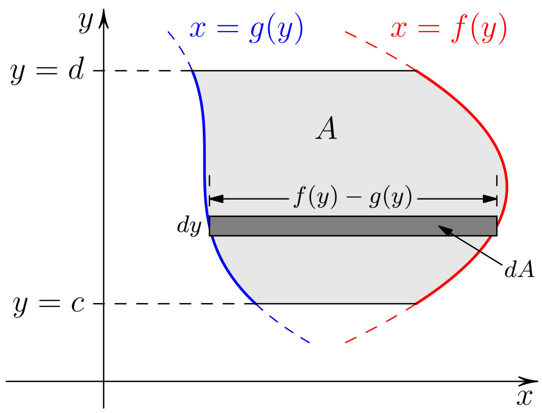
$$dA = \underbrace{[f(x) - g(x)]}_{\text{height}} \cdot \underbrace{dx}_{\text{width}}\ .$$

Sometimes the region for which one wants an area is better described by functions $x = f(y)$ and $x = g(y)$ denoting the right and left boundaries of the region respectively. One then has the result:

Theorem: 5.2. If $x = f(y)$ and $x = g(y)$ are integrable on the y-interval $[c, d]$ with $f(y) \geq g(y)$ on the interval then the **area between the curves** $x = f(y)$ and $x = g(y)$ and the lines $y = c$ and $y = d$ is

$$A = \int_c^d [f(y) - g(y)]\ dy\ .$$

The situation is depicted below:



In this case the integral represents the addition of horizontal area elements dA of area

$$dA = \underbrace{[f(y) - g(y)]}_{\text{width}} \cdot \underbrace{dy}_{\text{height}} \ .$$

Sometimes no interval $[a, b]$ or $[c, d]$ is given. If one is asked to find the region bounded by $y = h_1(x)$ and $y = h_2(x)$ one uses vertical area elements (first theorem) to find the area. One must solve $h_1(x) = h_2(x)$ to find x-values x_i which enclose the bounded region(s). There may be multiple enclosed regions and these values determine the endpoints of the integration. One must ascertain which function is the higher one on each region.

If one is asked to find the region bounded by $x = h_1(y)$ and $x = h_2(y)$ then one uses horizontal area elements (second theorem) to find the area. One must solve $h_1(y) = h_2(y)$ to find y-values y_i which enclose the bounded region(s). Once again these determine the endpoints of integration. One must determine which function is greater on each region.

Examples:
Find the area of the region bounded by the following curves:

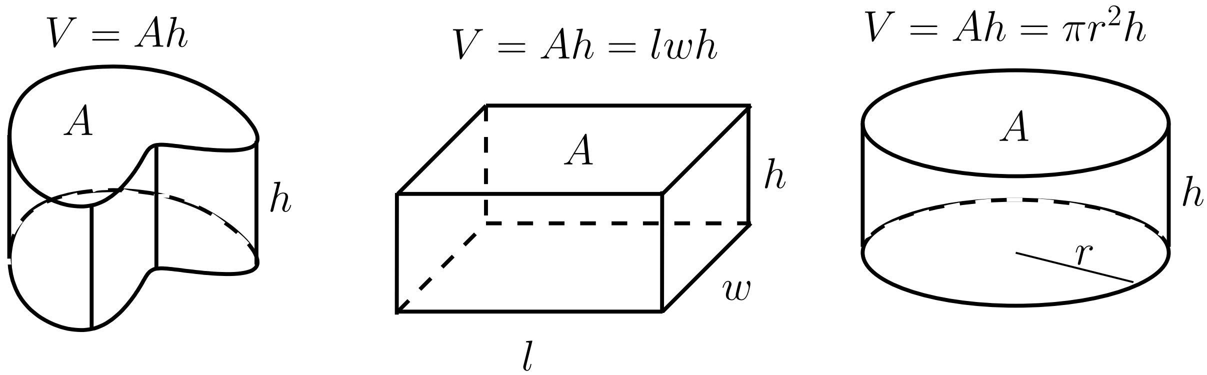
1. $y = 6 - x^2$ and $y = 3 - 2x$
2. $x = y^2$ and $x - y - 2 = 0$

5.2 Calculation of Volume

5.2.1 Volume as a Calculus Problem

The **volume** of a three-dimensional object is the amount of space it occupies (expressed in cubic units). If we had an odd-shaped jar and had a bag of caramels that were a cm on each side we could estimate the jar's volume by counting the number of caramels that would fit in the jar. This would be the volume in cubic centimetres (i.e. millilitres). In a more accurate estimate we would fill the jar with water and then measure the number of millilitres with a measuring cup.

We are interested in how we can use calculus to find such volumes. An object for which we have a formula for volume is a **right cylinder**. A cylinder is a three-dimensional geometric figure that has two congruent and parallel bases. When the bases are aligned one directly above the other one has a right cylinder. The following are three examples of right cylinders:



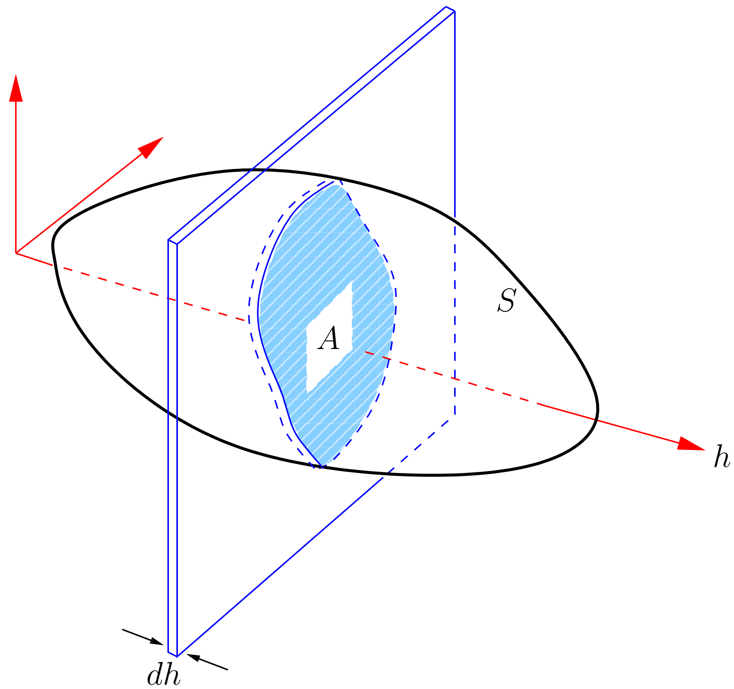
The first shows the most general case, while the second and third are special cases, a rectangular parallelepiped and a right circular cylinder, respectively. In general the area of a right cylinder is:

$$V = Ah$$

In the special case where the area is expressible in terms of other lengths we can further specialize the formula. So a right circular cylinder has volume:

$$V = \pi r^2 h$$

Returning to the general case of finding a volume of an arbitrary 3D object we may follow the same route we did for solving general areas. In the area case we sliced the figure into increasingly thinner rectangles ($\Delta x \to 0$) and then added those up. For a 3D object we can imagine slicing it into a large number of increasingly thinner (shorter) right cylinders (of height $\Delta h \to 0$) and adding them up. To do this we take an abitrary axis through our solid S and partition the axis into equal widths Δh. At each point h along the axis we will have a **cross-sectional area** A in the plane perpendicular to our axis h.



The area A will depend upon the position h along the axis and so it is a function $A(h)$ of h. The volume of our object will then be the limit of the Riemann Sum of the n cylinders with volumes V_i:

$$V = \lim_{n\to\infty} \sum_{i=1}^{n} \underbrace{A(h_i^*)\Delta h}_{V_i}$$

The position h_i^* is taken to be a point in the corresponding interval along h of the i^{th} cylinder. This limit is the definite integral:

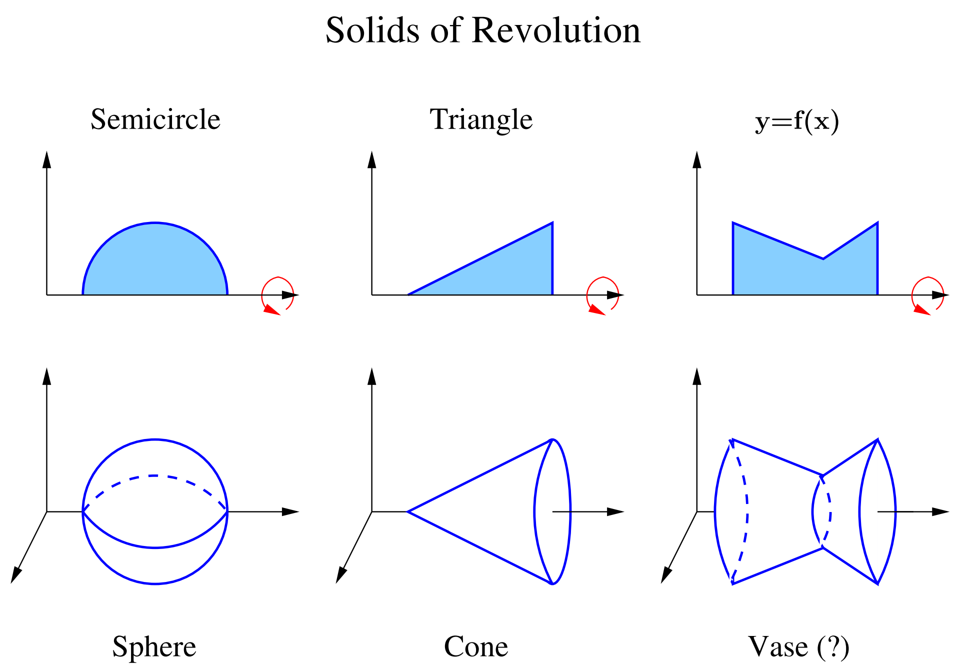
$$V = \int_a^b A(h)dh \; ,$$

where a and b denote the position of the first and last cylinders needed. Thinking in differential notation we are summing ($\int dV$) the volumes of infinitessimally thin cylinders given by:

$$dV = A\, dh$$

5.2.2 Solids of Revolution

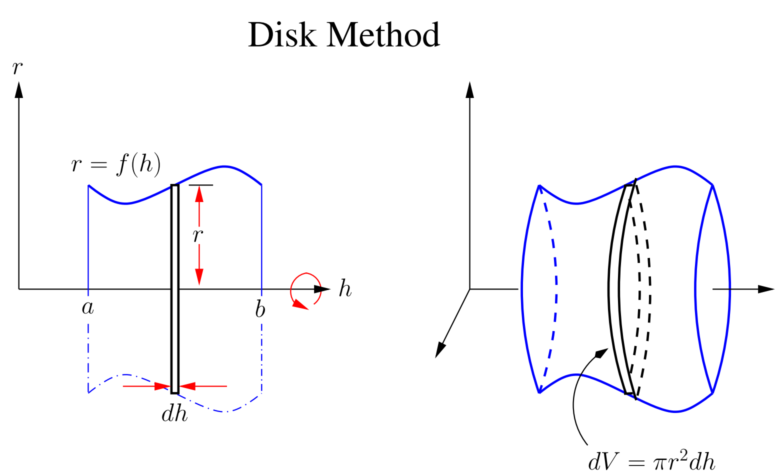
In general the function $A(h)$ for a particular solid S will be complicated. However if the solid has a symmetry and the axis h is directly along that axis of symmetry the function $A(h)$ may be calculated. A certain class of solids, called **Solids of Revolution** have such a symmetry because they are generated by sweeping out a two dimensional area around just such an axis:



As can be seen in the diagram some common shapes (spheres and cones) can be thought of as solids of revolution of simple areas (semicircles and triangles). Other interesting shapes may be constructed by choosing other functions for the upper boundary.

5.2.3 The Disk Method

For a solid of revolution for which the planar area of the revolution extends from the axis of symmetry to a boundary curve describable by a function, the volume is now calculable because the area function $A(h)$ can be found:



As shown in the diagram, the general right cylinders are now right **circular** cylinders with area $A = \pi r^2$. Our volume element is the **disk** of differential volume:

$$\boxed{dV = \pi r^2 dh}$$

where the radius r for the given disk is $r = f(h)$. The volume is then the sum ($V = \int dV$) given by

$$V = \int_a^b \pi r^2 \, dh = \int_a^b \pi \left[f(h)\right]^2 dh$$

If the symmetry axis is the x-axis (so $h = x$ and $r = y = f(x)$) then the formula for a solid of revolution about the x-axis becomes

$$V = \int_a^b \pi \left[f(x)\right]^2 dx$$

This method of calculating volumes is called the **Disk Method** because we are approximating the volume with disks (i.e. short right cylinders).

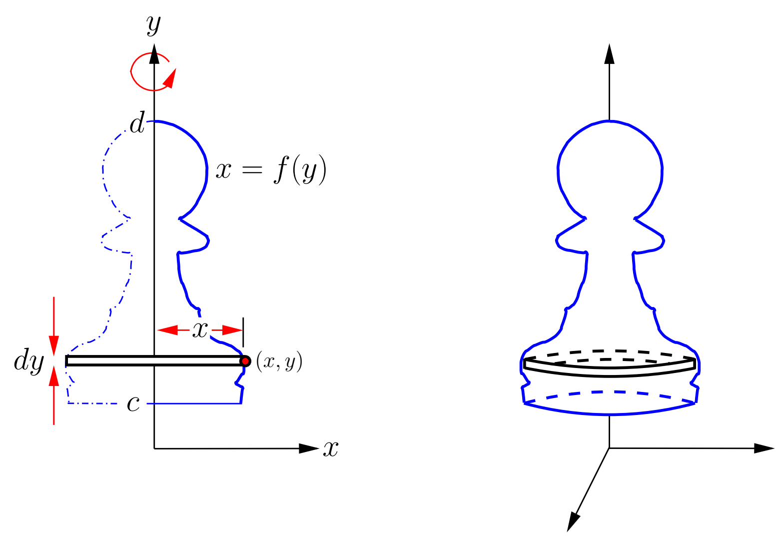
Example:
Find the volume of the solid obtained by revolving the following regions about the x-axis:

- The region bounded by the curves $y = x^{\frac{3}{2}}$, $x = 4$ and $y = 0$.
- The region bounded by the curve $y = 2x - x^2$ and the line $y = 0$.

The constant π is defined to be the ratio of the circumferance to the diameter of any circle, $\pi = \frac{C}{d}$. Archimedes (c. 287-212 BCE) discovered (long before the invention of calculus) that the constant π surprisingly shows up in the formulas for the area of a circle ($A = \pi R^2$) as well as for a sphere's volume ($V = \frac{4}{3}\pi R^3$) and surface area ($S = 4\pi R^2$).

Example:
Use calculus to show the volume of a sphere of radius R is $V = \frac{4}{3}\pi R^3$.

We sometimes wish to revolve a region about the y-axis. For a region bounded by the curve $x = f(y)$ and the lines $x = 0$ (i.e. the y-axis), $y = c$, and $y = d$ we can use horizontal disks. In this case $dh = dy$ (since the symmetry axis h is now y) and $r = x = f(y)$ as shown in the following diagram.



The formula for the volume of such a region revolved about the y-axis is therefore:

$$V = \int_c^d \pi \left[f(y)\right]^2 dy$$

Example:
Find the volume of revolution of the solid obtained by revolving the region bounded by $y = x^2$, $x = 0$, and $y = 4$ about the y-axis.

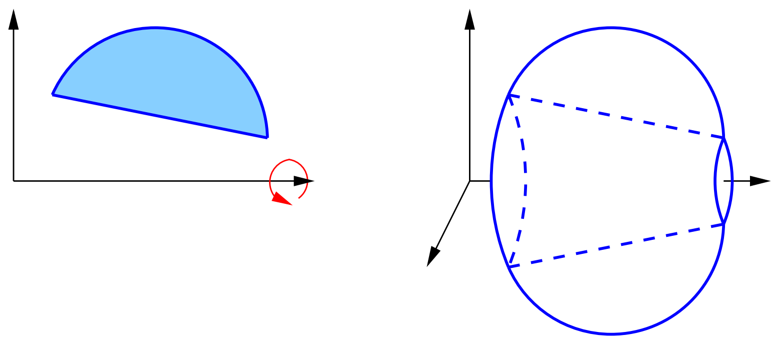
In general we may be interested in revolving an area about a horizontal or vertical axis which is neither the x nor the y-axis. In this case we must work from the differential volume dV to determine the volume integral. Follow the following steps which will work for any volume problem:

1. Draw the region to be revolved labelling the curves appropriately as well as points of intersection. Draw the symmetry line about which the curve is to be revolved.
2. Decide on the differential volume element you will use and indicate a typical cross-section rectangle on your region.
3. Label the endpoints of rectangle with coordinates (x, y), etc. Note that for vertical elements the endpoint coordinates will have the same endpoint x values while you will need to distinguish the y values (i.e. y_1, y_2) as these values will differ. If a value is constant (e.g. 0) for all potential cross-section rectangles label it as such instead of using a variable. The opposite holds if you are using a horizontal element (i.e. same y values at endpoints, different x's.)
4. Label the differential width of the cross-section rectangle as dx or dy according to whether you are using a vertical or horizontal elements respectively.
5. Write down your integral, $V = \int dV$, and substitute your volume element for dV.
6. Next replace each variable in the integral in terms of the variables of your cross-section rectangle endpoint coordinate variables. Some calculation may be required here, especially if you are not revolving about a coordinate axis.
7. From your diagram determine the limits of integration. These must be along the same axis as found in your differential, i.e. x for dx or y for dy. Note the limits are for the rectangles only within the region being revolved.
8. All variables of the integrand must be in terms of the integration variable. Replace any variables in your integral that are not the same as the differential variable in terms of the differential variable by considering the curve that the endpoint involving the variable lies upon. You may need to solve for the variable using the curve equation to do so.
9. Calculate the integral.

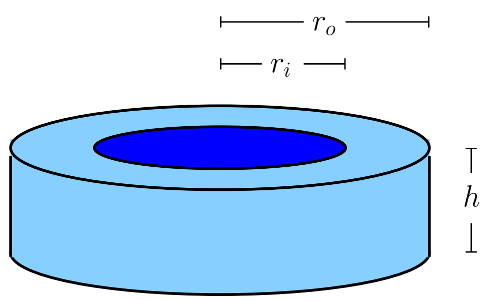
Example:
Find the volume of the solid generated by revolving the region bounded by $y = x^2$ and $y = 4$ about the line $y = 4$.

5.2.4 The Washer Method

Suppose we generate a solid of revolution in which the lower boundary is not the axis of revolution but, more generally the curve $g(h)$ above it. These regions are what we considered when we looked at areas between curves having an upper boundary determined by a function f and a lower boundary determined by a function g. If such a region is revolved about a chosen axis of symmetry h then the solid of revolution generated will have a hole in it and the disk method will not, at first glance, work.



The disk method can be used to calculate the volume of the outer solid and then again the inner solid with the actual volume being the difference of the two, but instead we introduce a new method which solves the problem with a single integral. Consider a right circular cylinder of height h and outer radius r_o with an interior cylinder of inner radius r_i removed:



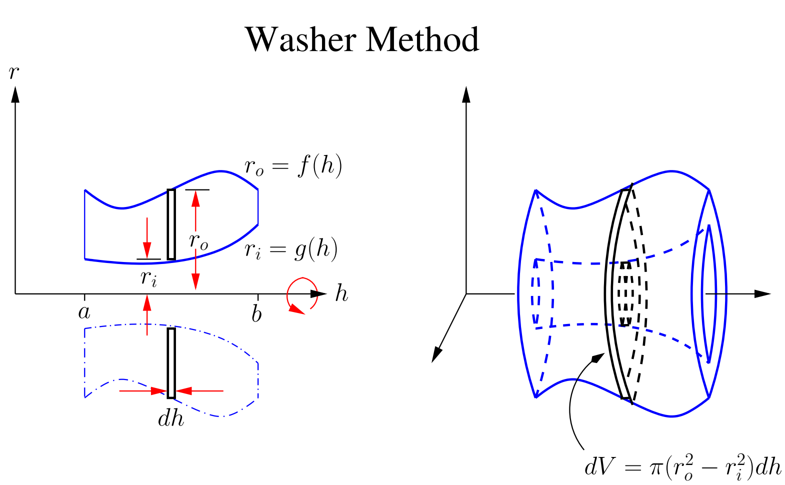
The volume of the solid is just the difference of the volumes of the cylinders:

$$V = V_o - V_i = \pi r_o^2 h - \pi r_i^2 h$$

Factoring out the π and h gives:

$$V = \pi(r_o^2 - r_i^2)h$$

If we imagine shrinking the height h to a small height Δh the shape looks like a metal *washer*. In the **Washer Method** we will slice solids of revolution with symmetrical holes in them into such washer-shaped volumes and sum them up to find the total volume of the solid.



In the Washer Method the height of the washer Δh is replaced with the differential dh so the differential element of volume is:

$$\boxed{dV = \pi(r_o^2 - r_i^2)dh}$$

The total volume is just the sum ($V = \int dV$) and we have:

$$V = \int_a^b \pi(r_o^2 - r_i^2)dh = \int_a^b \pi\left([f(h)]^2 - [g(h)]^2\right) dh$$

where we have replaced the outer and inner radii by their functional values at h (see diagram). If the symmetry axis is the x-axis (so $h = x$, $r_o = y_2 = f(x)$, and $r_i = y_1 = g(x)$) then the formula for a solid of revolution about the x-axis becomes

$$V = \int_a^b \pi\left([f(x)]^2 - [g(x)]^2\right) dx$$

This may always be found starting with the differential volume element $dV = \pi(r_o^2 - r_i^2)dh$ and following the general steps.

> **Example:**
> Find the volume of the solid obtained by revolving the region bounded by the curves $y = x^2 + 1$ and $y = 3 - x^2$ about the x-axis

If the symmetry axis is the y-axis then $h = y$ and we can write the outer radii as $r_o = x_2 = f(y)$ and the inner radii as $r_i = x_1 = g(y)$. The general formula for a solid of revolution about the y-axis becomes:

$$V = \int_c^d \pi\left([f(y)]^2 - [g(y)]^2\right) dy$$

Once again we use our convention of using constants c and d for limits along the y-axis.

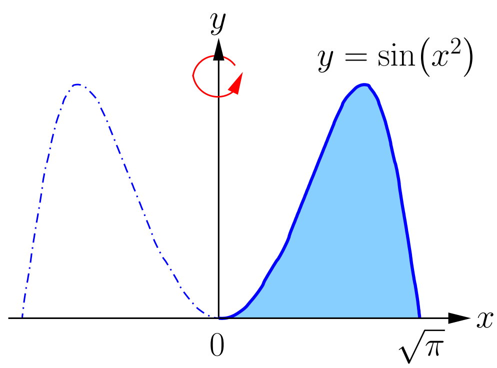
> **Example:**
> Find the volume of the solid obtained by revolving the region bounded by $y = 2x$ and $y = x^2$ about the y-axis.

If one is revolving about an axis that is not a coordinate axis there is no immediate formula since the curve functions no longer represent distances from the axes of revolution. Follow the steps outlined previously starting with the volume element and identifying the variables from your diagram. In this case the radii will need to be calculated in terms of the variables.

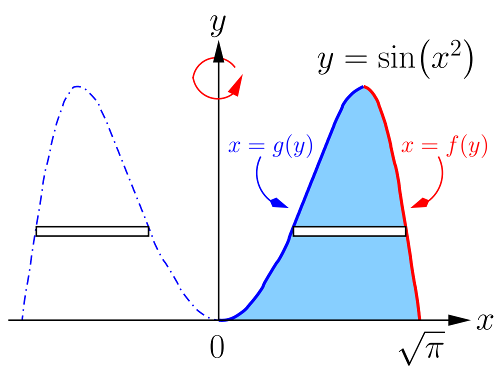
> **Example:**
> Find the volume of the solid obtained by revolving the region bounded by the curves $y = \sqrt[3]{x}$ and $y = \frac{1}{4}x$ and lying in the first quadrant ($x \geq 0$ and $y \geq 0$) about the line $x = 10$.

5.3 The Shell Method

Suppose one wished to calculate the volume of the solid generated by revolving the region bounded by the curve $y = \sin(x^2)$ and the line $y = 0$ about the y-axis as shown.

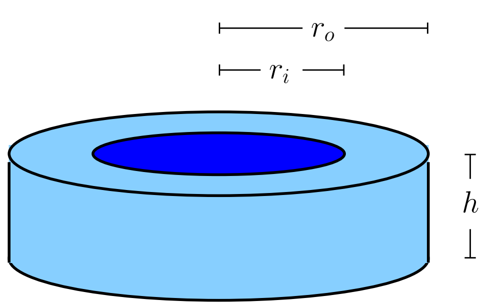


While one can imagine dividing the solid into washer-shaped regiond perpendicular to the y-axis, there would be difficulties.



Ultimately we would need to solve for x in terms of y to find the inner and outer radii of the washer in terms of the position y of the disk. This could be tricky in part because of the trigonometric function. Moreover in inverting the function we would somehow need to generate both an upper and lower function $x(y)$ corresponding to the two radii of the washer!

We now consider a different approach by the introduction of a new volume element. Recall the cylinder with the inner cylinder removed.



We found before the formula for the volume of this solid was just the difference of the volume of the cylinders:

$$V = \pi \left(r_o^2 - r_i^2\right) h$$

Notice that we can factor the difference of squares:

$$V = \pi \left(r_o + r_i\right) \left(r_o - r_i\right) h$$

Defining the average radius r to be:

$$r = \frac{r_o + r_i}{2} \ ,$$

and the difference of the radii to be

$$\Delta r = r_o - r_i \ ,$$

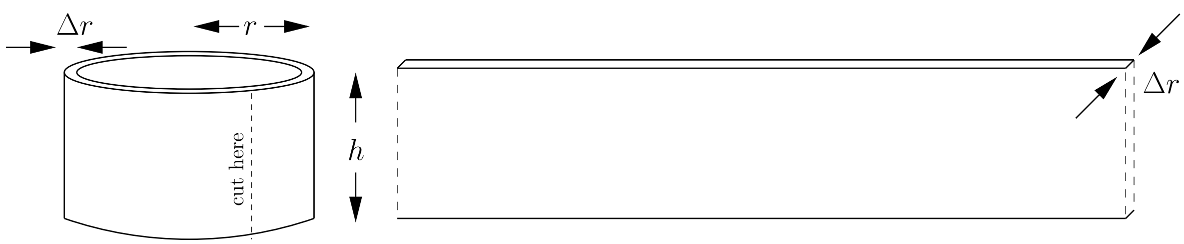
we can rewrite the volume as

$$V = \pi(2r)(\Delta r)h \ ,$$

or equivalently

$$V = 2\pi r h \, \Delta r \ .$$

Now suppose we make the two radii approximately equal. Then the average radius r approximately equals either of them with little error, while their difference Δr will be small. The solid will be like a cylindrical tin can (with the ends cut off) which we will refer to as a "shell". The volume formula for the shell now becomes quite intuitive. If we imagine cutting the shell along its length and unrolling it we would effectively have a rectangular sheet with top length equalling the circumference of the circle $(2\pi r)$ and the other side length equalling the height. The thin thickness will be Δr.

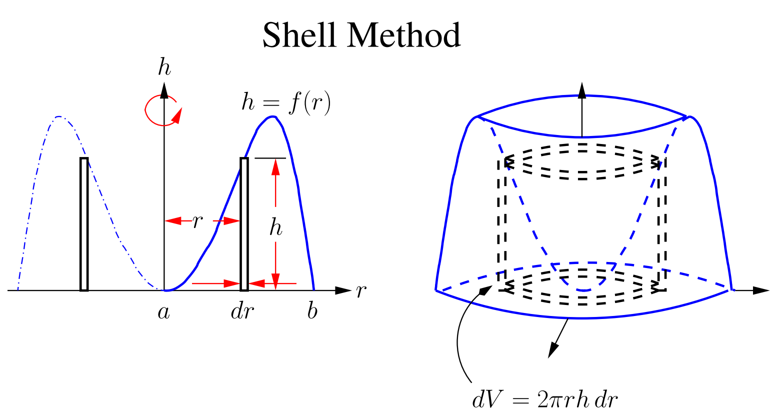


So the volume of the rectangular sheet (the shell) is then:

$$V = (\text{circumference})(\text{height})(\text{depth}) = (2\pi r)(h)(\Delta r) = 2\pi r h \, \Delta r$$

as found before.

Now in the Shell Method we imagine slicing the solid of revolution into concentric cylindrical shells and adding up the volumes.



In the limit that the number of shells goes to infinity, so $\Delta r \to 0$, the limit of the Riemann Sum is an integral ($\int dV$) of differential volume

$$\boxed{dV = 2\pi r h\, dr}$$

with, as usual the finite V and Δr replaced with dV and dr in our shell volume formula. Specifically, for the region depicted above we have for the volume of the solid of revolution

$$V = \int_a^b 2\pi r h\, dr = \int_a^b 2\pi r f(r)\, dr$$

Notice that in the Shell Method the differential dr is now *perpendicular* to the axis of symmetry.

If the axis of symmetry is the y-axis then, setting $h = y$ and $r = x$ in the above formula we get the following result for an a solid of revolution generated by revolving the region bounded by the curve $y = f(x)$, $y = 0$, $x = a$, and $x = b$ about the y-axis:

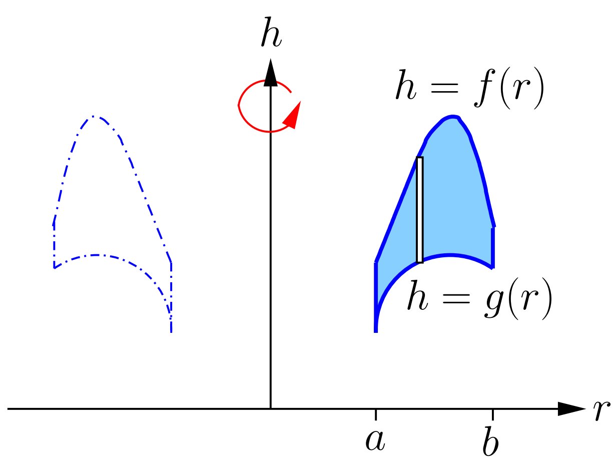
$$V = \int_a^b 2\pi x f(x)\, dx$$

Example:
Find the volume of the solid generated by revolving the region bounded by the curves $y = \sin(x^2)$ and $y = 0$ between $x = 0$ and $x = \sqrt{\pi}$ about the y-axis.

If the axis of symmetry is the x-axis then, setting $h = x$ and $r = y$ in the above formula we get the following formula for a solid of revolution generated by the region bounded by the curves $x = f(y)$, $x = 0$, $y = c$, and $y = d$.

$$V = \int_c^d 2\pi y f(y)\, dy$$

More generally the lower bound of the region need not be along the radial axis:



If the upper bound of the figure remains described by the function $h = f(r)$ while the lower bound is now the function $h = g(r)$ the the height of the shell is $h = f(r) - g(r)$. The volume for the solid generated by revolving about the h-axis is then

$$V = \int_a^b 2\pi r h\, dr = \int_a^b 2\pi r\, [f(r) - g(r)]\, dr$$

One observes that, in the example depicted, the washer method would now be impossible in a single integral. For even should we be able to solve for $r(h)$ on the curves, there is no single washer generated toward the bottom of the region. One also notes in this example that the region is generally bounded by four curves, two are vertical lines. In general the shell method works well for solids with straight cylindrical boundaries parallel to the axis of revolution.

In the event we are revolving about the y-axis, so $h = y$ and $r = x$ in the above diagram, the solid generated by revolving the region bounded by $y = f(x)$, $y = g(x)$, $x = a$ and $x = b$ is

$$V = \int_a^b 2\pi x \, [f(x) - g(x)] \, dx$$

On the other hand the volume of the solid generated by revolving about the x-axis the region bounded by $x = f(y)$, $x = g(y)$, and the horizontal lines $y = c$ and $y = d$ is, setting $h = x$, and $r = y$,

$$V = \int_c^d 2\pi y \, [f(y) - g(y)] \, dy$$

These formulae may all, as usual be derived following the general steps starting with the differential volume element $dV = 2\pi r h \, dr$.

Example:
Find the volume of the solid generated by revolving the region bounded by the curves $x^2 = 4y$ and $y = 4$ about the x-axis.

If one wishes to use the shell method to evaluate a solid generated by revolution of a region about a non-coordinate axis the differential volume element $dV = 2\pi r h \, dh$ and the general steps must be used.

Example:
Find the volume of the solid generated by revolving the region bounded by the curves $y = x^2$ and $x + y = 2$ about the line $x = 3$.

5.4 Determining the Volume Method

We have considered several methods for calculating volumes of solids of revolution. Which to use in a particular instance can be determined by the following set of steps:

1. Draw the region to be revolved and identify the axis of revolution.

2. Next imagine a line *perpendicular* to the axis of revolution. Where this intersects the region represents (the cross-section of) a particular volume element. Slide the line through the region keeping it perpendicular to the axis of revolution. If all such lines intersect the boundary of the region two or fewer times then the *Disk/Washer Method* will potentially work. An exception to the intersection rule is when you first enter and leave the region; here one can intersect the boundary in a straight line.

   As you pass the line through the region consider the endpoints of the intersection with the region. If these do not always lie on the same curves then the region will need to be partitioned into two (or more) regions each requiring a separate integral to use this method.

3. Repeat the previous step but now with lines *parallel* to the axis of revolution. If the region passes this test then it is a candidate for the *Shell Method*.

4. For any region that passes the *Disk/Washer* test, if one of the boundary curves of the region is the axis of revolution then use the *Disk Method* for that interval, otherwise the *Washer Method* will be used.

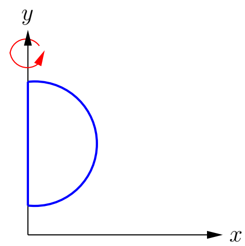
5. To distinguish multiple candidate methods, next look at the curve equations. Identify the variable of integration (dx or dy). It will be the variable of the coordinate axis *parallel* to the axis of revolution in the case of the *Disk/Washer Method* and *perpendicular* to the axis of revolution for the *Shell Method*. The curves in the boundary region (aside from any straight lines at the beginning or end of the region for the method in question) will have to have the non-differential variable written as a function of the differential variable. If these curves cannot be solved for the non-differential variable, reject the method.

6. If several candidate methods still remain, set up the volume integral for one candidate. If it is not easily solved, try the other candidate.

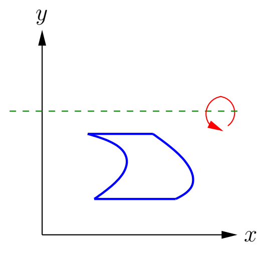
7. If no candidates lead to a solution, then the solid cannot be calculated with a single volume method. Try partitioning the region into several areas and calculate the volume for each region revolved separately. If these can be calculated the volume of the entire solid is their sum.

Example:
For each of the following regions identify the simplest volume method(s) to use to calculate the solid of revolution found by rotating the region about the indicated axis. For the method state whether to use vertical or horizontal differential volume elements.

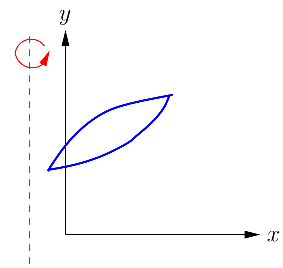
a)



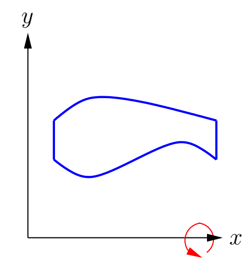
b)



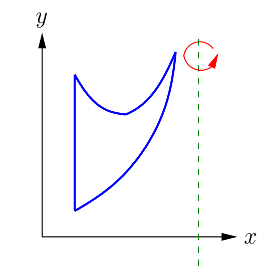
c)



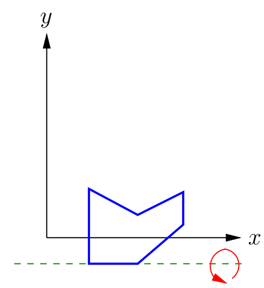
d)



e)

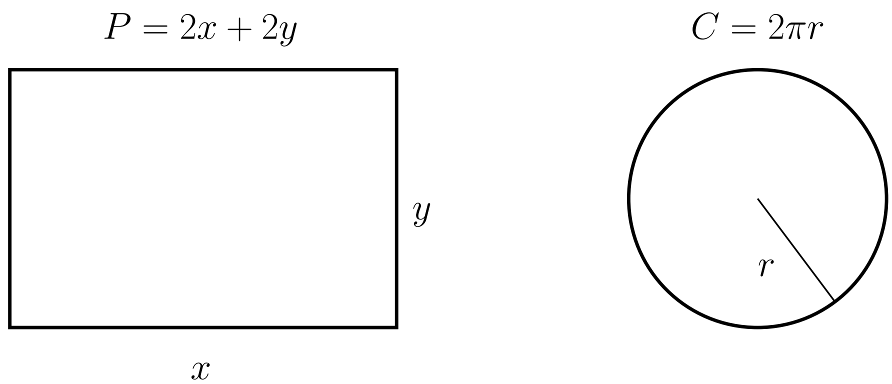


f)

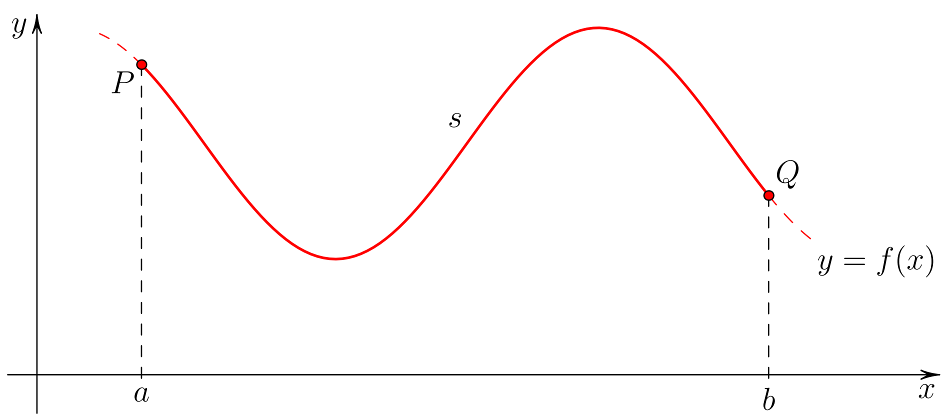


5.5 Arc Length

From basic geometry the perimeter P of a polygon or the circumference C of a circle are the distance around the given object. They are the length one would measure if the boundary of the object were cut, straightened out, and measured with a ruler.



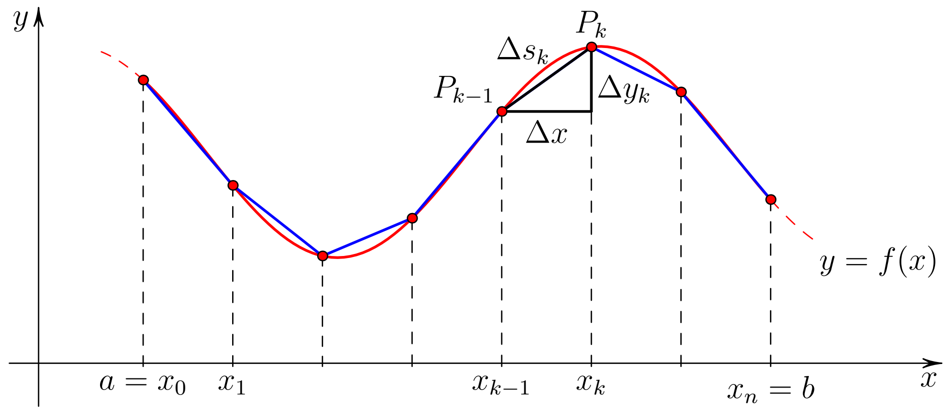
Intuitively the **arc length s** of a curve between points P and Q is similarly the "length of a piece of string" that follows the curve joining those two points.



Consider a curve defined by a function $y = f(x)$ with a continuous first derivative $f'(x)$ defined on interval $[a, b]$. Such a function is called a **smooth function** and its graph $y = f(x)$ is said to be **smooth**. An approximation to the arc length can be found by breaking the interval into n subintervals of equal length $\Delta x = \frac{b-a}{n}$ with endpoints

$$a = x_0, x_1, x_2, \ldots, x_{n-1}, x_n = b$$

The points $P_k = (x_k, f(x_k))$, $k = 0, 1, \ldots, n$, lie along the curve. The length of the chord (straight line segment) joining P_{k-1} to its neighbour P_k will approximate the arc length between those two points. Call that chord length Δs_k. It equals, by the Pythagorean Theorem, $\Delta s_k = \sqrt{(\Delta x)^2 + (\Delta y_k)^2}$ where we define $\Delta y_k = y_k - y_{k-1} = f(x_k) - f(x_{k-1})$.



By the Mean Value Theorem applied to the interval $[x_{k-1}, x_k]$ there exists an x_k^* in (x_{k-1}, x_k) satisfying

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} = f'(x_k^*)$$

Equivalently $\Delta y_k = f'(x_k^*)\Delta x$. From this it follows that we can write Δs_k as:

$$\begin{aligned}
\Delta s_k &= \sqrt{(\Delta x)^2 + (\Delta y_k)^2} \\
&= \sqrt{(\Delta x)^2 + \left[f'(x_k^*)\Delta x\right]^2} \\
&= \sqrt{(\Delta x)^2 + \left[f'(x_k^*)\right]^2 (\Delta x)^2} \\
&= \sqrt{\left(1 + \left[f'(x_k^*)\right]^2\right) \cdot (\Delta x)^2} \\
&= \sqrt{1 + \left[f'(x_k^*)\right]^2} \cdot \sqrt{(\Delta x)^2} \\
&= \sqrt{1 + \left[f'(x_k^*)\right]^2}\,\Delta x
\end{aligned}$$

An approximation to the arc length between the points $P(a, f(a))$ and $Q(b, f(b))$ is therefore the length of the polygonal curve we have constructed, namely the Riemann sum

$$\sum_{k=1}^{n} \Delta s_k = \sum_{k=1}^{n} \sqrt{1 + \left[f'(x_k^*)\right]^2}\,\Delta x$$

Taking the limit as $n \to \infty$ (so $\Delta x \to 0$) we have the following:

Definition: If function f has a continuous derivative f' over the x-interval $[a, b]$ then the length of the graph over the interval (the **arc length**) is

$$s = \int_a^b \sqrt{1 + [f'(x)]^2}\,dx \quad \text{or} \quad s = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx$$

Note that if a curve is described by the equation $x = g(y)$ with continuous derivative $g'(y)$ over the y-interval $[c, d]$ then the arc length between the endpoints equals

$$s = \int_c^d \sqrt{1 + [g'(y)]^2}\,dy \quad \text{or} \quad s = \int_c^d \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\,dy$$

Examples:
Find the lengths of the following curves over the given intervals.

1. $f(x) = 3x^{\frac{2}{3}} - 10$ over $8 \leq x \leq 27$
2. $y = \dfrac{x^4}{16} + \dfrac{1}{2x^2}$ over $1 \leq x \leq 2$

If we have a smooth function $f(x)$ on interval $[a, b]$ we can consider the arc length from the starting point $(a, f(a))$ up to the point $(x, f(x))$ for some value x in $[a, b]$. This length is clearly dependent on the choice of x and hence is a function $s(x)$.

Definition: Given a smooth curve described by the function $y = f(x)$ over the x-interval $[a, b]$, the arc length between the initial point $(a, f(a))$ and the point $(x, f(x))$ for x in $[a, b]$ is given by the **arc length function**

$$s(x) = \int_a^x \sqrt{1 + [f'(t)]^2}\, dt$$

Notes:

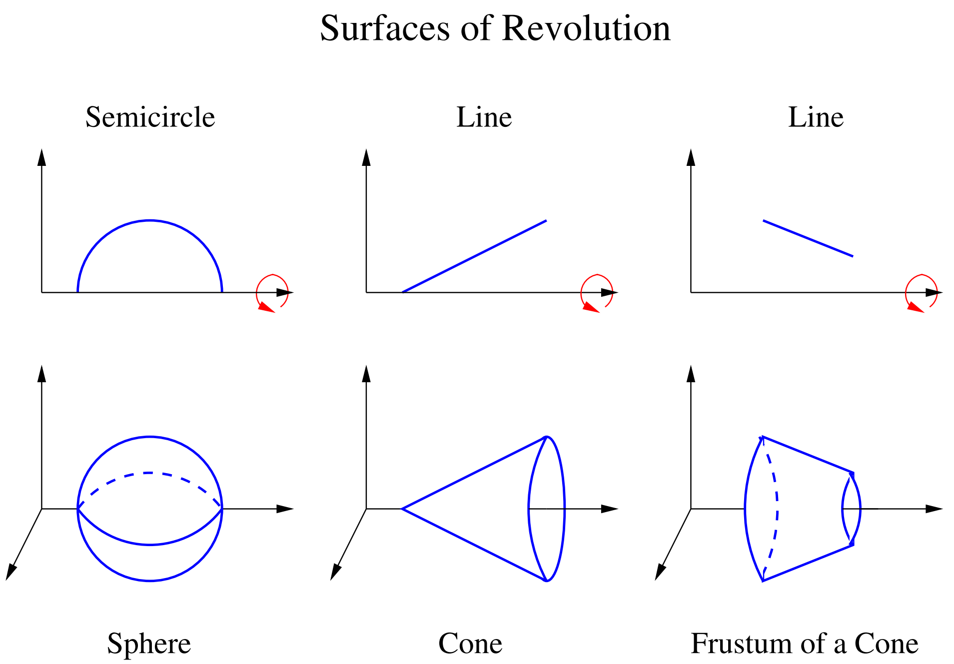
1. We changed the variable of integration to t to distinguish it from the limit of integration x.
2. The derivative of the arc length function $s(x)$ is, by the Fundamental Theorem of Calculus, $\frac{ds}{dx} = \sqrt{1 + [f'(x)]^2}$. This in turn implies, since $ds = \frac{ds}{dx}\, dx$, that the **differential of the arc length** is
$$ds = \sqrt{1 + [f'(x)]^2} dx \ .$$
Noting that $dy = f'(x)dx$ the differential of arc length is sometimes written
$$ds = \sqrt{dx^2 + dy^2}\ ,$$
reflecting its Pythagorean origins. The latter differential formula, along with the relation $dy = f'(x)dx$ for x-intervals $[a, b]$ or $dx = g'(y)dy$ for y-intervals $[c, d]$, and the idea of the integral as a sum ($s = \int ds$) make it easy to remember/derive the arc length formulae.

Examples:
Find the arc length function $s(x)$ of

1. $y = \sqrt{x - x^2} - \cos^{-1}\left(\sqrt{x}\right)$ with starting point $P(0, -\pi/2)$.
2. $y = e^x$ with starting point $P(0, 1)$.

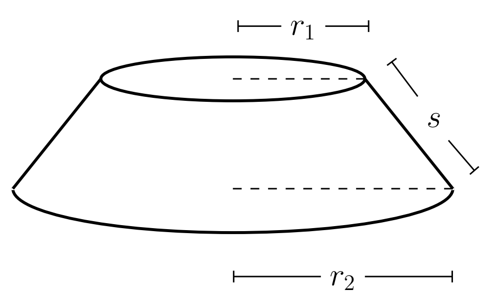
5.6 Areas of Surfaces of Revolution

Just as volumes of revolution are constructed by revolving an area about a line, a **surface of revolution** is generated by revolving the graph of a continuous function about a line.



Note that the surface generated in this manner does not produce any end surfaces such as the circular bottom of the cone or ends of the frustum of the cone. (The spherical surface however is fully generated.)

We turn ourselves to the problem of calculating the area of such a surface of revolution. Consider a simple object like the frustum of the cone above, with end radii of r_1 and r_2 and **slant height** s.

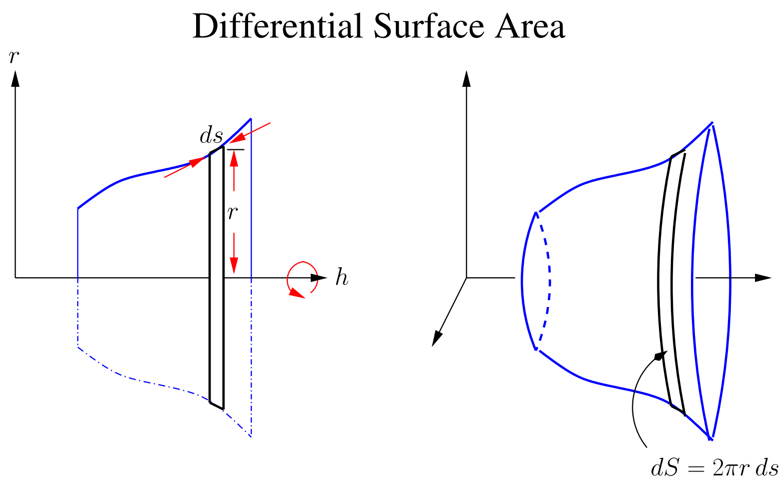


The **lateral** surface area of the frustum (i.e. the area of the side but not including the circular ends), can be shown to be

$$S = 2\pi r s \ ,$$

where $r = \frac{1}{2}(r_1 + r_2)$ is the average of the two radii.

The approximate area of a surface of revolution may be found by slicing it perpendicular to its axis of revolution (which we will call h). The intersection of these planes with the surface form the circular ends of conical frusta the net lateral surface area of which will approximate the area of the surface of revolution.



A Riemann sum for the area is thereby formed:

$$S \approx \sum_{k=1}^{n} \Delta S_k \ ,$$

where ΔS_k is the lateral surface area of the k^{th} conical frustum. Upon taking the limit of an infinite number ($n \to \infty$) of conical frusta one arrives at an exact expression for the surface area S in terms of an integral $S = \int dS$ where the differential area of the conical frustum is given, analagously to the finite surface area:

$$dS = 2\pi r\, ds$$

where here the finite slant height s is replaced by the **differential arc length** $ds = \sqrt{dx^2 + dy^2}$ of the curve generating the surface. Four different formulae are possible depending on whether the axis of symmetry h is the x or the y-axis and whether the curve-generating function is written as $y = f(x)$ over the x-interval $[a, b]$ or $x = g(y)$ over the y-interval $[c, d]$. In all cases we assume the function is positive and has continuous first derivative (i.e. is smooth).

Considering first the case where the axis of symmetry h is the x-axis and the curve revolved is described by the function $y = f(x)$, $a \le x \le b$, then our radius is $r = y = f(x)$ and we write the differential arc length in terms of the independent variable x as $ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$ so that $S = \int dS = \int 2\pi r\, ds$ is given by the formula:

$$S = \int_a^b 2\pi f(x)\sqrt{1 + [f'(x)]^2}\, dx \quad \text{or} \quad S = \int_a^b 2\pi f(x)\sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

If we still consider the case where we are rotating about the x-axis but the curve is instead described by the function $x = g(y)$, $c \le y \le d$ then the differential arc length must be written in terms of the independent variable y as $ds = \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy$. The radius $r = y$ still but as that is already expressed in terms of the variable of the differential (dy) variable our formula becomes (still using $S = \int dS = \int 2\pi r\, ds$):

$$S = \int_c^d 2\pi y\sqrt{1 + [g'(y)]^2}\, dy \quad \text{or} \quad S = \int_c^d 2\pi y\sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy$$

Consider next the cases where the surface is generated by revolving the curve around the y-axis. In this case the radius $r = x$. If the curve is described by the function $y = f(x)$, $a \le x \le b$ then the independent variable is x so we use the differential ds written in terms of dx to get, since $S = \int dS = \int 2\pi r\, ds$,

$$S = \int_a^b 2\pi x \sqrt{1 + [f'(x)]^2}\, dx \quad \text{or} \quad S = \int_a^b 2\pi x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

If instead the curve is written $x = g(y)$, $c \le y \le d$, then we must write the radius x in terms of the differential (dy) as $g(y)$ and also select the appropriate form of the differential arc length in terms of dy:

$$S = \int_c^d 2\pi g(y) \sqrt{1 + [g'(y)]^2}\, dy \quad \text{or} \quad S = \int_c^d 2\pi g(y) \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy$$

None of these formulae need to be memorized. Just remember that we are integrating ($S = \int dS$) the differential surface area $dS = 2\pi r\, ds$. Then look at the curve's function; the independent variable determines the form of ds and the limits on the integral. The radius r is the variable perpendicular to the axis of revolution. If that is not the independent variable then use the function for the curve to express it in terms of the independent (differential) variable.

Examples:
Find the areas of the following surfaces of revolution generated by the curves:

1. $y = x^3$ from $x = 1$ to $x = 2$ revolved about the x-axis.
2. $x = \frac{1}{8}y^3$ from $x = 1$ to $x = 8$ revolved about the y-axis.
3. $x = 1 + 2y^2$, $1 \le y \le 2$ revolved about the x-axis.
4. $y = \sqrt[3]{x}$, $1 \le x \le 8$ revolved about the y-axis.
5. $y = e^{-x}$ from $x = 0$ to $x = 1$ revolved about the x-axis.

We are now able to calculate the following result discovered by Archimedes:

Example:
Use calculus to show that the surface area of a sphere of radius R is $S = 4\pi R^2$.

Unit 6: Parametric Equations

6.1 Parametric Equations

We have seen how a function $F(x)$ can define a curve via the equation $y = F(x)$. Similarly a function $G(y)$ of the y-coordinate can define a curve via $x = G(y)$. Finally a relation $H(x, y) = 0$ such as $x^2 + y^2 - 25 = 0$ can define a curve in the coordinate plane.

Another way in which we can define a planar curve is to consider writing both x and y as functions of a further variable.

Definition: Let $f(t)$ and $g(t)$ be continuous functions of a real variable t with common domain D. Then

$$\begin{aligned} x &= f(t) \\ y &= g(t) \end{aligned}$$

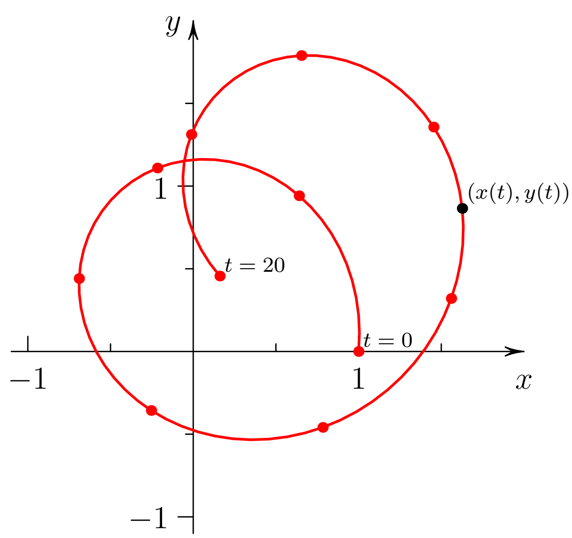
are called **parametric equations**. The variable $\boldsymbol{t}$ is called the **parameter**.

Geometrically such a parametric equation represents a curve. Given a value for t the point $(x, y) = (f(t), g(t))$ is defined. As t ranges over the domain D a curve $\mathcal{C}$ is defined.

Example:
The following shows the curve $\mathcal{C}$ generated by the parametric equations

$$\begin{aligned} x(t) &= \cos(t/2) + \frac{t}{20} \\ y(t) &= \sin(t/2) + \frac{t}{20} \end{aligned}$$

from $t = 0$ to $t = 20$. The red dots highlight the points at $t = 0, 2, 4, \ldots, 20$.



The parameter t can, as the choice of variable name suggests, represent a time. In this case $(x, y) = (f(t), g(t))$ can represent the position of an object in the plane at a given time. The curve, in this case, is the path that the object takes in time. However the parameter need not be time. The parameter could, for instance, represent a physical variable such as an angle θ. Another common physical parameter is

s, the arc length from some fixed point on the curve to the point (x, y). Finally the parameter may have no physical interpretation whatsoever.

The curve $\mathcal{C}$ in the last example could not be represented either by a function $y = F(x)$ or $x = G(y)$ as vertical and horizontal lines intersect the curve at more than one place. However, as a parametric curve the curve is represented entirely by functions, one just requires two of them.

Example:
Sketch and identify the curve given by $x = 3\cos\theta$, $y = 2\sin\theta$, for $0 \leq \theta \leq 2\pi$.

A further way to identify a curve from its parametric definition is to eliminate the parameter between the two equations to get a single equation involving only x and y.

Example:
In the previous example of the ellipse one finds $\cos\theta = \frac{x}{3}$ and $\sin\theta = \frac{y}{2}$. From this it follows that

$$\begin{aligned} \cos^2\theta + \sin^2\theta &= 1 \\ \Rightarrow \left(\frac{x}{3}\right)^2 + \left(\frac{y}{2}\right)^2 &= 1 \\ \Rightarrow \frac{x^2}{9} + \frac{y^2}{4} &= 1 \end{aligned}$$

We recognize the relation in the last line as defining a horizontal ellipse centred on the origin with semimajor axis of length $\sqrt{9} = 3$ and semiminor axis of $\sqrt{4} = 2$.

Example:
Eliminate the parameter t from the parametric equations $x = e^{2t}$, $y = t + 1$ to find an equation for the curve in terms of x and y only.

Note that a function $y = F(x)$ can always be written in parametric form by associating the independent variable with the parameter:

$$\begin{aligned} x(t) &= t \\ y(t) &= F(t) \end{aligned}$$

or, equivalently, $\mathcal{C}(t) = (t, F(t))$. Similarly $x = G(y)$ can be generated by the parametrization $\mathcal{C}(t) = (G(t), t)$.

Examples:

1. The curve $y = 1 + x^2$ can be written parametrically as

$$\begin{aligned} x(t) &= t \\ y(t) &= 1 + t^2 \end{aligned}$$

2. The curve $x = e^y$ can be written parametrically as

$$\begin{aligned} x(t) &= e^t \\ y(t) &= t \end{aligned}$$

6.2 Calculus of Parametric Curves

6.2.1 Tangent Slope and Concavity

For a curve defined by $y = F(x)$ we know the slope of the tangent at the point $(x, y) = (x, F(x))$ is given by the derivative $F'(x)$. What then is the slope of the tangent to a curve defined parametrically by $x = f(t)$, $y = g(t)$? Suppose we can eliminate the parameter t as we have done in the previous section to obtain $y = F(x)$ for some function F.[1] Then it follows that

$$g(t) = F(f(t)).$$

Differentiating both sides of the latter equation with respect to t one obtains, by the Chain Rule,

$$g'(t) = F'(f(t)) \cdot f'(t).$$

If $f'(t) \neq 0$ then $F'(f(t)) = \frac{g'(t)}{f'(t)}$ and hence the tangent slope at the point $(x(t), y(t))$ is given by

$$F'(x) = \frac{g'(t)}{f'(t)} \qquad (f'(t) \neq 0).$$

In Leibniz derivative notation this is easily remembered; the slope of the tangent line $x = f(t)$, $y = g(t)$ is given by

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} \qquad \left(\frac{dx}{dt} \neq 0\right).$$

Note that:

1. The curve has a horizontal tangent line when $\frac{dy}{dt} = 0$ (provided $\frac{dx}{dt} \neq 0$).
2. The curve has a vertical tangent line when $\frac{dx}{dt} = 0$ (provided $\frac{dy}{dt} \neq 0$).

To find the second derivative $\frac{d^2y}{dx^2}$, one may continue as before noting that this would be $F''(x)$. Differentiating $F'(f(t)) = \frac{g'(t)}{f'(t)}$ with respect to t on both sides gives, using the Chain Rule,

$$F''(f(t)) \cdot f'(t) = \frac{d}{dt}\left[\frac{g'(t)}{f'(t)}\right].$$

Thus, for the second derivative with respect to x we have

$$F''(x) = \frac{\frac{d}{dt}\left[\frac{g'(t)}{f'(t)}\right]}{f'(t)} \qquad (f'(t) \neq 0)\ ,$$

or in Leibniz notation,

$$\frac{d^2y}{dx^2} = \frac{\frac{d}{dt}\left(\frac{dy}{dx}\right)}{\frac{dx}{dt}} \qquad \left(\frac{dx}{dt} \neq 0\right).$$

Here we must find $\frac{dy}{dx}$ as a function of t as above and then differentiate.[2]

[1] While this may not be possible for the entire curve, this may plausibly be done for a portion of the curve near the point at which we desire the tangent slope.

[2] This latter formula may be remembered by imagining $\frac{dy}{dx}$ as a function of t and then differentiating this with respect to x. Assuming we can find an inverse of f (at least locally around the point of interest), so that $t = f^{-1}(x)$ and $\frac{dt}{dx}$ is defined, then, by the Chain Rule,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dt}\left(\frac{dy}{dx}\right) \cdot \frac{dt}{dx}$$

Noting that our inverse functions satisfy $\frac{dt}{dx} \cdot \frac{dx}{dt} = 1$ we can bring $\frac{dt}{dx}$ into the denominator as $\frac{dx}{dt}$ to recover our previously derived formula.

Note that $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ are of interest as these are, within the given the coordinate system, quantities that will be independent of the particular parameterization used to generate the curve.

Example:
Find the equation of the tangent line to the curve $x(t) = 1 - t^3$, $y(t) = t^2 - 3t + 1$ at the point corresponding to $t = 1$.

Example:
Find the equation(s) of the tangent(s) to the curve $x(t) = 1 - 2\cos^2 t$, $y(t) = (\tan t)(1 - 2\cos^2 t)$, t in $(-\pi/2, \pi/2)$ at the point $(0, 0)$.

Example:
For the curve given by $x(t) = 3t^2 - 6t$, $y(t) = \sqrt{t}$, with $t \geq 0$,

1. Find the points at which the tangent line is horizontal or vertical.
2. Find $\frac{d^2y}{dx^2}$ and discuss the concavity.

6.2.2 Area Under the Curve

For non-negative function $F(x)$, the area under the curve $y = F(x)$ and above the x-axis for $a \leq x \leq b$ is found using vertical area elements and is given by $A = \int_a^b F(x)\,dx = \int_a^b y\,dx$. If the curve is parametrically defined by $x = f(t)$, $y = g(t)$, for t in $[t_a, t_b]$, so that $a = f(t_a)$ and $b = f(t_b)$, the area becomes, upon change of variables via $x = f(t)$ and noting $dx = \frac{dx}{dt}\,dt$,

$$A = \int_a^b y\,dx = \int_{t_a}^{t_b} y\frac{dx}{dt}\,dt = \int_{t_a}^{t_b} g(t)f'(t)\,dt$$

Example:
Find the area enclosed by the x-axis and the curve $x = 1 + \sin t$, $y = \frac{\pi}{2}t - t^2$.

Similarly if a curve is defined by the non-negative function $x = G(y)$ for $c \leq y \leq d$ then the area between that curve and the y-axis is found using horizontal area elements and is given by $A = \int_c^d G(y)\,dy = \int_c^d x\,dy$. If the curve is parametrically defined by $x = f(t)$, $y = g(t)$, for t in $[t_c, t_d]$, such that $c = g(t_c)$ and $d = g(t_d)$, the area becomes, upon changinge variable via $y = g(t)$ and noting $dy = \frac{dy}{dt}\,dt$,

$$A = \int_c^d x\,dy = \int_{t_c}^{t_d} x\frac{dy}{dt}\,dt = \int_{t_c}^{t_d} f(t)g'(t)\,dt$$

These formulae can be generalized for areas between parametric curves by consideration of the coordinates of the endpoints of the area elements.

Example:
Find the area inside the **deltoid curve** described by the parametric equations
$x(t) = 2a\cos t + a\cos 2t$, $y(t) = 2a\sin t - a\sin 2t$, $0 \leq t \leq 2\pi$. Here a is a positive constant.

6.2.3 Arc Length

We have seen that the arc length of a curve $\mathcal{C}$ described by $y = F(x)$ for $a \leq x \leq b$ is given by

$$s = \int_a^b \sqrt{1 + [F'(x)]^2}\, dx = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

If curve $\mathcal{C}$ is also parameterized by $x = f(t)$, $y = g(t)$, for $t_a \leq t \leq t_b$ where $t_a = f(a)$ and $t_b = f(b)$ one has, by the change of variable $x = f(t)$, that

$$s = \int_{t_a}^{t_b} \sqrt{1 + \left(\frac{dy/dt}{dx/dt}\right)^2}\, \frac{dx}{dt}\, dt\ ,$$

where here we used our expression for the derivative $\frac{dy}{dx}$ in terms of t found before. If we further assume that our function $x = f(t)$ is an increasing function of x, so $dx/dt > 0$ and $\sqrt{(dx/dt)^2} = dx/dt$, we may rewrite this:

$$s = \int_{t_a}^{t_b} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt\ .$$

The previous result can be shown to be be true quite generally for any parameterized curve $\mathcal{C}$, not simply for curves that may be described by a function of x.

Theorem: 6.1. Suppose curve $\mathcal{C}$ is traced exactly once by the parameterization $x = f(t)$, $y = g(t)$ as t increases from t_i to t_f. If derivatives f' and g' exist and are continuous on $[t_i, t_f]$ then the arc length of $\mathcal{C}$ is

$$s = \int_{t_i}^{t_f} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt\ .$$

One can recover special case formulae from Chapter **??** for $y = F(x)$ using parameterization $x = t$, $y = F(t)$ and for $x = G(y)$ using parameterization $x = G(t)$, $y = t$.

Note that the formula is easily remembered by $s = \int ds$, where the differential arc length element $ds = \sqrt{dx^2 + dy^2}$. This formula requires some form of parameterization of the curve for its application. However, the differential $ds = \sqrt{dx^2 + dy^2}$ reflects the fact that the arc length of a curve is independent of the parameterization of the curve chosen to calculate it.

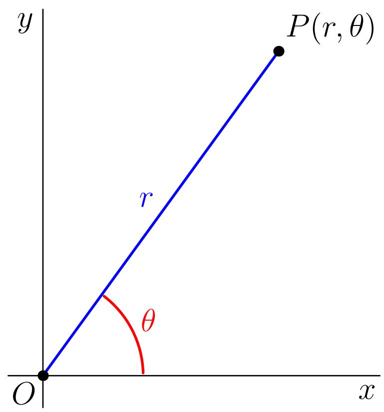
Example:
Find the length s of the arc of the circle $x = R\cos\theta$, $y = R\sin\theta$ for $0 \leq \theta \leq \frac{2\pi}{3}$ where R is the constant radius.

Unit 7: Polar Coordinates

7.1 Polar Coordinates

A coordinate system identifies a point P in space by an ordered list of one or more numbers called coordinates. We are familiar with the **Cartesian Coordinate System** in which we identify a point P in two-dimensional space by the ordered pair (x, y) where x is the x-coordinate of the point and y is the y-coordinate of the point.

In polar coordinates one selects a point to O to be the origin or pole of the coordinate system. The first coordinate of an arbitrary point P is the distance r from O to P. Next one draws a **polar axis** as a ray emanating from O which, for planar coordinates, is aligned with the positive x-axis. The second coordinate of P is the angle θ made between this axis and the ray OP. The point P is then represented by the ordered pair (r, θ) which are the **polar coordinates** of P.



Notes:

1. Angles θ are positive if measured in the usual counter-clockwise sense from the polar axis and are negative if measured clockwise.
2. By convention polar coordinates (r, θ) with a negative r represent the point P at $(|r|, \theta + \pi)$. In other words, they are at the positive distance $|r|$ directed along the ray opposite the ray subtending the angle θ.

Example:
Plot the point with each of the following polar coordinates.

1. $(2, 3\pi/4)$
2. $(2, -2\pi/3)$
3. $(-2, 3\pi/4)$

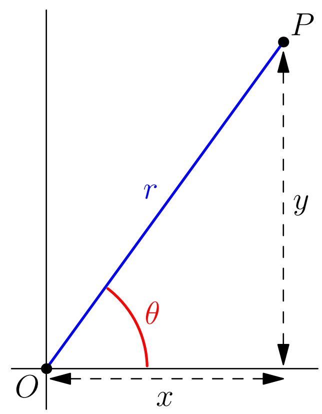
Unlike in Cartesian Coordinates, the polar coordinates for point P are not unique since the addition of a multiple of 2π to the angle returns one to the same physical point.

Example:
The following polar coordinates all represent the same point P:

$$(2, \pi/4), \qquad (2, 9\pi/4), \qquad (2, -7\pi/4), \qquad (-2, 5\pi/4)$$

7.1.1 Converting Between Polar and Cartesian Coordinates

Suppose P has Cartesian coordinates (x, y) and polar coordinates (r, θ).



As suggested by the diagram, one has

$$x = r\cos\theta \qquad\qquad y = r\sin\theta \ .$$

These formulae allow conversion from polar to Cartesian coordinates.

Example:
Find the Cartesian coordinates of $(r, \theta) = (3, \pi/4)$.

To convert from Cartesian to polar coordinates we solve for r and θ. Squaring and adding both sides of the two equations gives:

$$x^2 + y^2 = r^2\cos^2\theta + r^2\sin^2\theta = r^2$$

Dividing both sides of the second equation by the corresponding sides of the first yields:

$$\frac{y}{x} = \frac{r\sin\theta}{r\cos\theta} = \tan\theta$$

In summary we have, to convert from Cartesian to polar coordinates,

$$r = \sqrt{x^2 + y^2} \qquad\qquad \tan\theta = \frac{y}{x}$$

Note that when solving $\tan\theta = y/x$ for θ:

1. There will typically be two solutions in $[0, 2\pi)$. The selection of the correct value is made by looking at the quadrant determined by the original x and y values.
2. If $x = 0$ (so y/x is undefined) consider the y-value to determine if $\theta = \pi/2$ or $\theta = 3\pi/2$.

Example:
Find the polar coordinates of $(x, y) = \left(-1, -\sqrt{3}\right)$

7.1.2 Curves in Polar Coordinates

If we consider the coordinate θ as an independent variable and r a dependent variable determined by function $f(\theta)$ then $r = f(\theta)$ describes a curve in polar coordinates. More generally the solutions (r, θ) of the relation $g(r, \theta) = 0$ describe a curve.

Example:

1. Sketch the curve represented by the polar equation $r = 3\sin\theta$.
2. Find a Cartesian equation for this curve.

7.1.3 Symmetry in Polar Curves

Suppose a polar curve is defined either explictly by $r = f(\theta)$ or implicitly with the relation $g(r, \theta) = 0$. Then, just as we saw with Cartesian Coordinates, the curves may have certain symmetries arising from the following transformations:

$\theta \to -\theta$: If $f(-\theta) = f(\theta)$ or $g(r, -\theta) = g(r, \theta)$, then the curve is symmetric about the polar axis.

$r \to -r$: If $g(-r, \theta) = g(r, \theta)$ then the curve is symmetric about the origin O.

$\theta \to \pi - \theta$: If $f(\pi - \theta) = f(\theta)$ or $g(r, \pi - \theta) = g(r, \theta)$, then the curve is symmetric about the vertical $\theta = \pi/2$ line.

Example:
Sketch the curve $r = 1 + \cos\theta$.

7.1.4 Tangents

The slope of the tangent to a polar curve described by $r = f(\theta)$ may be found by converting to Cartesian Coordinates. Since $x = r\cos\theta$ and $y = r\sin\theta$ we have

$$x = f(\theta)\cos\theta \qquad\qquad y = f(\theta)\sin\theta$$

The curve is now represented by a parametric equation with parameter θ. Applying our rules for finding the derivative dy/dx of such equations from Chapter 6 gives $dy/dx = (dy/d\theta)/(dx/d\theta)$ and so:

$$\frac{dy}{dx} = \frac{f'(\theta)\sin\theta + f(\theta)\cos\theta}{f'(\theta)\cos\theta - f(\theta)\sin\theta}$$

Example:
Consider the curve $r = 1 + \cos\theta$.

1. Find the slope of the tangent line to the curve where $\theta = \pi/2$.
2. Find the points on the curve where the tangent line is horizontal or vertical.