MTH 201 MATHEMATICAL METHODS 1: (3 Units) (L30: P 0: T 1)
Real-valued functions of a real variable. Real-valued functions of two or three variables. Review of differentiation and integration and their applications. Mean value theorem. Taylor series. Partial derivatives chain rule, extrema, langrange multipliers. Increments, differentials and linear approximations. Evaluation of line, integrals. Multiple integrals. Pre-requisite -MTH 103.

1. Real-valued functions

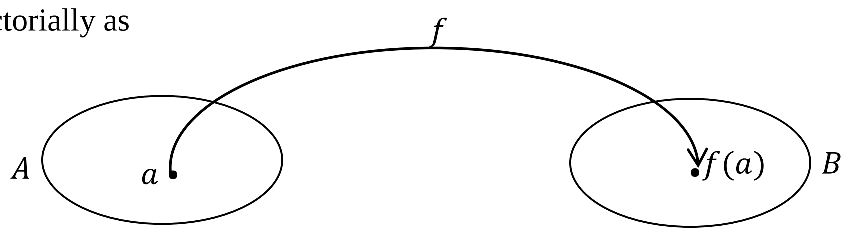
A function whose range lies within the real numbers i.e., no-root numbers and non-complex numbers, is said to be real function, also called a real-valued function. In mathematical analysis, and applications in geometry, applied mathematics, engineering, and natural sciences, a real-valued function is a function whose domain is the real variables $\mathbb{R}^n$. So, a real-valued function is a function whose domain is a subset $D \in \mathbb{R}$ of the set $\mathbb{R}$ of real numbers and the codomain is $\mathbb{R}$; such a function can be represented by a graph in the Cartesian plane.

In simplest terms the domain of a function is the set of all values that can be plugged into a function and have the function exist and have a real number for a value. However, for the domain, we need to avoid division by zero, square roots of negative numbers, logarithms of zero and logarithms of negative numbers. The range of a function is simply the set of all possible values that a function can take.

Most real functions that are considered and studied are differentiable in some interval. The most widely considered of such functions are the real functions, which are the real-valued functions of a real variable, that is, the functions of a real variable whose *codomain is the set of real numbers*.

1.1 Real-valued function of a real variable

Recall that a function is simply a rule for associating with each element of a set $A \in D$ in the domain to an element of a set B in the codomain. We often represent such an abstract function pictorially as



We can also represent the same situation with the notation;

$$f: A \longrightarrow B \qquad : \qquad x \longmapsto f(a)$$

which indicates that f is a function that takes elements of the set A to elements of the set B according to the rule $a \in A$ maps to $f(b) \in B$. The set A is referred to as the *domain* of the function f and the set B is referred to as the *target* space and the set

$$\{b \in B \mid b = f(a) \quad \text{for some} \quad a \in A\}$$

is called the image (or range) of f.

When the set A is actually a subset of $\mathbb{R}^n$ then we say that $f : \mathbb{R}^n \longrightarrow B$ is a function of several variables (even if the dimension n is huge). If the target space B is $\mathbb{R}$, the set of real numbers, then we say that $f : A \longrightarrow \mathbb{R}$ is a *real-valued function*. Example, a linear equation, $y = 2x$, and a quadratic equation, $y = x^2 - 4x + 3$ are real-valued functions of one independent variable x.

Examples

(a) Find the domain of each of the following real-valued functions of real variable

$$f(x) = \frac{x^2 + 2x + 1}{x^2 - 8x + 12}$$

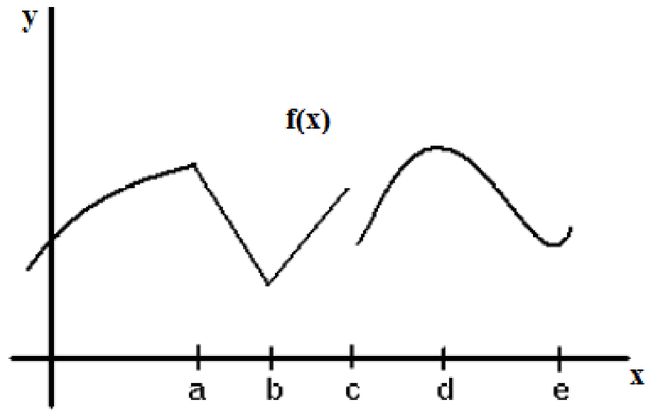
Solution:
$f(x) = \frac{x^2+2x+1}{x^2-8x+12} = \frac{(x+1)^2}{(x-6)(x-2)}$ is defined for all satisfying $(x-6)(x-2) \neq 0$, $x \neq 2,6$. $\therefore$ Domain $(f) = R - \{2,6\}$

(b) Find the domain and range of each of the following real-valued function: $f(x) = \sqrt{x-1}$.
Solution: $f(x) = \sqrt{x-1}$ is defined for all x satisfying $x - 1 \geq 0$, i.e., $x \geq 1$.
Now, let $y = \sqrt{x-1}$. Clearly, $y \geq 0$ for all $x \in [1, \infty]$. So, range $f(x) = [0, \infty]$

(c) Find the domain and range of each of the following real-valued function: $f(x) = |x - 1|$.
Solution: We have, $f(x) = |x-1|$. Clearly, $f(x)$ is defined for all $x \in R$. So, domain $(f) = R$. Also, $f(x) = |x-1| \geq 0$ for all $x \in R$. So, range $(f) = [0, \infty]$

1.1.1 Graphs of real-valued functions of a single variable

A real-valued function of real numbers can be represented in the Cartesian plane by a graph, and such graph can identify properties of that function: for example from the graph



we see that $f(x)$ has

- a discontinuity at $x = c$
- derivatives everywhere except at the points a, b, and c
- local maxima at the points a and d
- local minima at the points b and e

From the graph, the set of y-values taken on by f is the range of the function. The symbol y is the dependent variable of f, and f is said to be a function of the independent variables x. As we shall see later, graphical methods are also very useful in elucidating the behavior of real-valued functions of several variables.

1.2 Real-valued function of several variables

A real-valued function f defined on a subset D of $\mathbb{R}^2$ is a rule that assigns to each point $f(x,y)$ in D a real number $f(x,y)$. The largest possible set D in $\mathbb{R}^2$ on which f is defined is called the domain of f, and the range of f is the set of all real numbers $f(x,y)$ as (x,y) varies over the domain D. A similar definition holds for functions $f(x,y,z)$ defined on points (x,y,z) in $\mathbb{R}^3$. So, suppose D is a set of n-tuples of real numbers, $(x_1, x_2, x_3, \dots, x_n)$. A real-valued function f on D is a rule that assigns a single real number

$$y = f(x_1, x_2, x_3, \dots, x_n)$$

to each element in D. The set D is the function's domain. The set of y-values taken on by f is the range of the function. The symbol y is the dependent variable of f, and f is said to be a function of the n independent variables x_1 to x_n. We also call the x's the function's input variables and we call y the function's output variable.

Subsequently, a real-valued function of two variables is just a function whose domain is $\mathbb{R}^2$ and whose range is a subset of $\mathbb{R}^1$ or simply $\mathbb{R}$, the real numbers. If we view the domain D as column vectors in $\mathbb{R}^n$, we write the function as

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

where x_1, x_2, x_3 are the independent variables. Example is the volume of a right circular cylinder which is a function of the radius and height,

$$V = f(r, h) \quad \text{or} \quad V = \pi r^2 h$$

The volume of the cylinder is a function of its radius and height. Observe that the volume increases as both the radius and the height increase.

Examples

(a) The domain of the function: $f(x, y) = xy$, is all the $\mathbb{R}^2$, and the range of f is all of $\mathbb{R}$.

(b) The domain of the function: $f(x, y) = \frac{1}{x-y}$ is all of $\mathbb{R}^2$ except the points (x, y) for which $x = y$. That is, the domain is the set of $D = \{(x, y): x \neq y\}$. The range of f is all real numbers except 0.

(c) The domain of the function: $f(x, y) = \sqrt{1 - x^2 - y^2}$ is the set $D = \{(x, y): x^2 + y^2 \leq 1\}$,since the quantity inside the square root is non-negative if and only if $1 - (x^2 + y^2) \geq 0$. We see that D consists of all points on and inside the unit circle

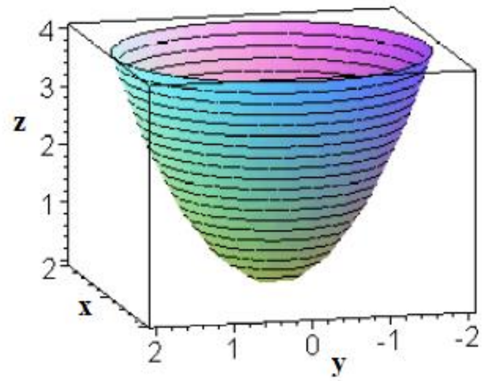
(d) The domain of the function: $f(x, y, z) = e^{x+y+z}$ is all of $\mathbb{R}^3$, and the range of f is all positive real numbers.

1.2.1 Graphs of Real-Valued Functions of Several Variables

Let $f: \mathbb{R}^n \longmapsto \mathbb{R}$ be a real-valued function of several variables. The graph of f is the set of points in $\mathbb{R}^{n+1}$ of the form

$$\{(x_1, x_2, x_3, \ldots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_{n+1} = f(x_1, x_2, x_3, \ldots, x_n)\}$$

Example, consider the function: $\mathbb{R}^n \longmapsto \mathbb{R}: (x, y) \longmapsto x^2 + y^2$. The graph of f is the set of points $\{(x, y, z) \in \mathbb{R}^3 \mid z = x^2 + y^2\}$



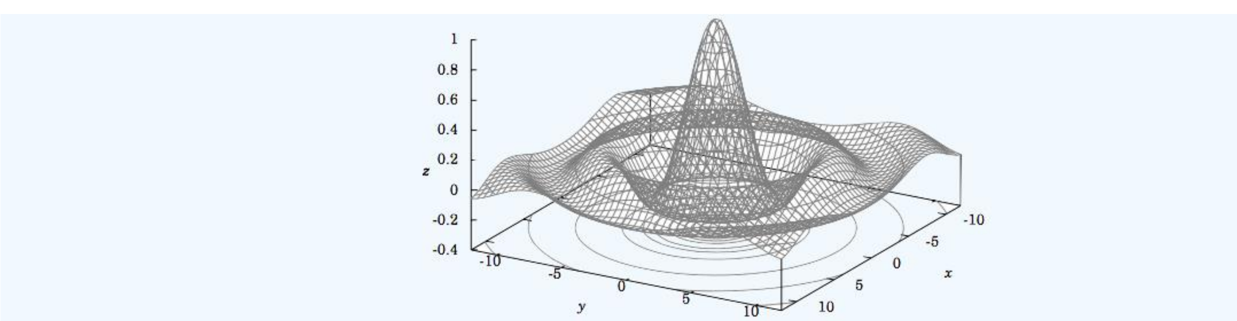
Unfortunately, the use of graphs to visualize functions is really only effective for functions with 1 or 2 variables. To draw the graph of a function of 3 variables requires a 4-dimensional space to draw the picture (and this is pretty hard to even imagine at your level).

A function $f(x,y)$ defined in $\mathbb{R}^2$ is often written as $z = f(x,y)$, so that the graph of $f(x,y)$ is the set $\{(x,y,z): z = f(x,y)\}$ in $\mathbb{R}^3$. So we see that this graph is a surface in $\mathbb{R}^3$, since it satisfies an equation of the form $F(x,y,z) = 0$ (namely, $F(x,y,z) = f(x,y) - z$). The traces of this surface in the planes $z = c$, where c varies over $\mathbb{R}$, are called the level curves of the function. Equivalently, the level curves are the solution sets of the equations $f(x,y) = c$, for c in $\mathbb{R}$. Level curves are often projected onto the xy -plane to give an idea of the various "elevation" levels of the surface (as is done in topography).

Example

The graph of the function: $f(x,y) = \frac{sin\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}}$

is shown below. Note that the level curves (shown both on the surface and projected onto the xy -plane) are groups of concentric circles.



Observe what happens to the function in the example above, at the point $(x, y) = (0,0)$, since both the numerator and denominator are 0 at that point. The function is not defined at (0,0), but the limit of the function exists (and equals 1) as (x, y) approaches (0,0).

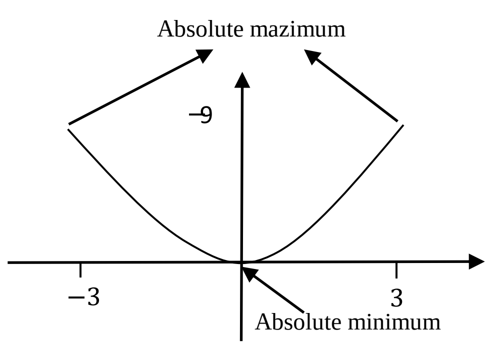
Applications of Differentiation

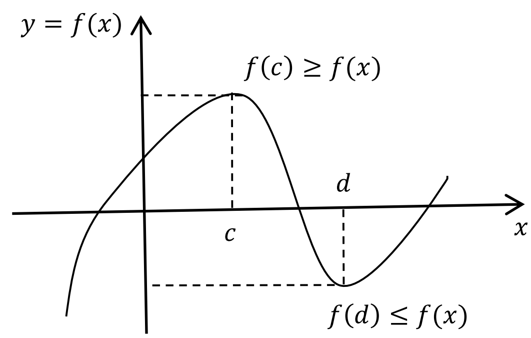
Maximum and Minimum Values

Definition: Let c be a number in the domain D of a function . Then $f(c)$ is the

- Absolute maximum value of f on D if $f(c) \geq f(x)$ for any $x \in D$
- Absolute minimum value of f on D if $f(d) \leq f(x)$ for any $x \in D$

Example: $f(x) = x^2, x \in [-3,3]$

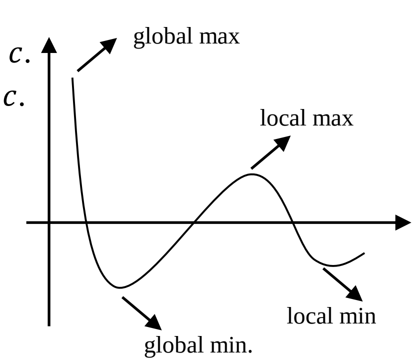




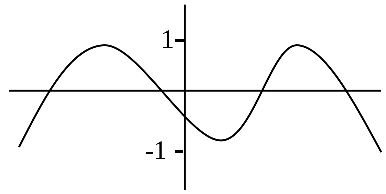
$c = 0$ - absolute minimum
$c = \pm 3$ - absolute maximum
$y = 0, 9$ are extreme values

Other names: an absolute maximum or minimum is sometimes called a global maximum or minimum. The maximum and minimum values of are called extreme values of f.

Definition: The number $f(c)$ is a
- Local maximum value of f if $f(c) \geq f(x)$ for any x near c.
- Local minimum value of f if $f(c) \leq f(x)$ for any x near c.

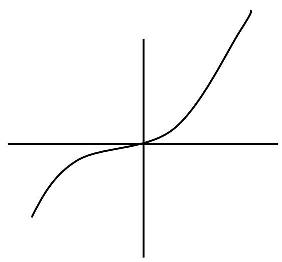


Example: $f(x) = \sin x$



Maximum: $x = \frac{\pi}{2} + 2k\pi$
Minimum: $x = -\frac{\pi}{2} + 2k\pi$

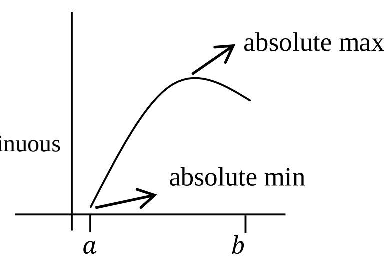
Example: $y = x^3$

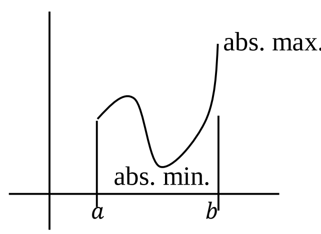


$f(x)$ is everywhere, no min/max.

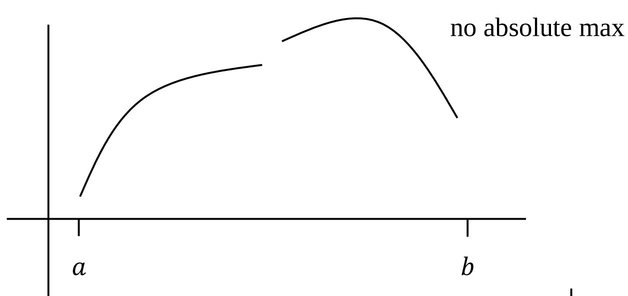
When does a function have extreme values?
Theorem: If f is differentiable (continuous) on a closed interval $[a, b]$, then f attains an absolute maximum value $f(c)$ and absolute minimum value $f(d)$ at some numbers $c, d \in [a, b]$.

f is continuous

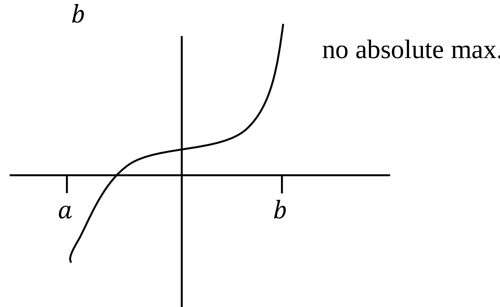




f is discontinuous:



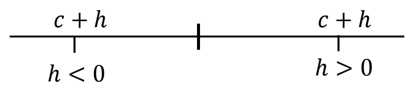
f is continuous but on an open interval (a, b):



Fermat's Theorem: If f has a local maximum or minimum at c, and if $f'(c)$ exists, then $f'(c) = 0$.

Proof:
Assume f has local max. at c, then
$f(c) \geq f(x)$ for x close to c and $f(c) \geq f(c+h)$, where h is close to 0
Thus,

$$f(c+h) - f(c) \leq 0$$



Now, let $h > 0$, then

$$\frac{f(c+h) - f(c)}{h} \leq 0$$

Hence,

$$f'(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \leq 0$$

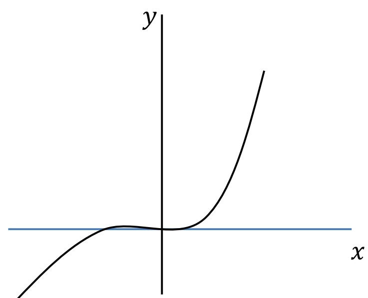
Let $h < 0$,

$$\frac{f(c+h)-f(c)}{h} \geq 0 \text{ and } f'(c) = \lim_{h \to 0} \frac{f(c+h)-f(c)}{h} \geq 0. \quad \text{Thus, } f'(c) = 0.$$

Caution: The converse of Fermat's Theorem is false in general.

Example: $f(x) = x^3$.

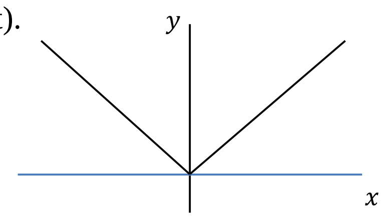
$f'(x) = 3x^2$ and $f'(0) = 0$,
but no local min/max.



$y = x^3$. The cubic function.

Example: $f(x) = |x|$

0 is a local min, but $f'(0)$ dne (does not exist).



Note: However, Fermat's Theorem suggests that we should start looking for extreme values at numbers where the derivative is zero or does not exist.

Definition: A critical number of a function f is a number $c \in D$ such that either $f'(c) = 0$ or $f'(c)$ does not exist.

Example: Find the critical numbers of $f(x) = \frac{x}{1+x^2}$.

$$f'(x) = \frac{(1+x^2) - x(2x)}{(1+x^2)^2} = \frac{1-x^2}{(1+x^2)^2} = 0$$

Thus,

$1 - x^2 = 0 \implies x = \pm 1$, which is the critical numbers.

Rephrase Fermat's Theorem: If f has a local maximum or minimum at c, then c is a critical number of f.

Closed Interval Method: To find the absolute maximum and minimum values of a continuous function f on a closed interval $[a, b]$:

- Find the values of f at the critical numbers of f in (a, b).
- Find the values of f at the endpoints of the interval.
- The largest of the values is the absolute maximum; the smallest value is the absolute minimum.

Example: Find the absolute maximum and minimum values of

$f(x) = x^3 - 6x^2 + 5$ on $[-3,5]$.

$f'(x) = 3x^2 - 12x = 0$ or $x(x-4) = 0$

Thus,

$x = 0,4 \in [-3,5]$

$f(0) = 5$ and $f(4) = -27 \implies (0,5)$ is the abs. max.

$f(-3) = -76$ and $f(5) = -20$, so $(-3,-76)$ is the abs. min.

Example: $f(x) = e^x + e^{-2x}$, $0 \le x \le 1$.

$$f'(x) = e^x + e^{-2x}(-2) = e^x - 2e^{-2x}$$
$$= e^{-2x}(e^{3x} - 2) = 0$$

but,,

$e^{-2x} \neq 0$, so, $e^{3x} = 2$ or $3x = \ln 2 \implies x = \frac{1}{3}\ln 2$

Thus,

$f\left(\frac{1}{3}\ln 2\right) = 1.89$, so, $\left(\frac{1}{3}\ln 2, 1.89\right)$ - abs. min.

and

$f(0) = 2$, and $f(1) = 2.84 \implies (1, 2.84)$ - abs. max.

Rolle's Theorem: Let f be a function such that

- f is continuous on $[a, b]$
- f is differentiable on (a, b)
- $f(a) = f(b)$

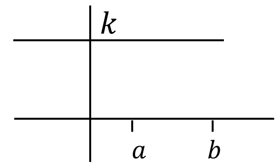
Then, there is a number c in (a, b) such that $f'(c) = 0$.

Proof:

Case I: $f(x) = \text{const.} = k,\ f(a) = f(b) = k$

$f'(x) = 0$ for any $x \in (a, b)$

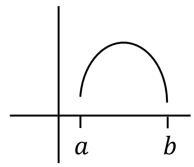
Take any c from (a, b)

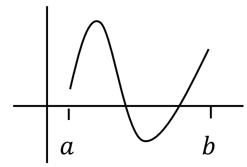


Case II: $f(x) > f(a)$ for some $x \in (a, b)$

$f(x)$ must have max. value for some $x \in (a, b)$

$\Longrightarrow$ by Fermat's theorem, $f'(c)$

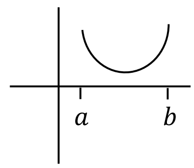


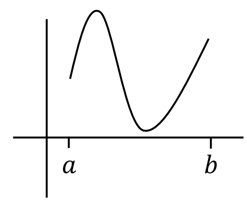


Case III: $f(x) < f(a)$ for some $x \in (a, b)$

$f(x)$ must have min value for $c \in (a, b)$

Thus, by Fermat's theorem, $f'(c) = 0$





Example: Prove that the equation $x^{13} + 7x - 5 = 0$ has exactly one (real) root.

$f(x) = x^{13} + 7x - 5$

So,

$f(0) = -5$ and $f(1) = 3$, $(-5 \leq 0 \leq 3)$

By IVT (Intermediate Value Theorem), there is $c \in (0,1)$ such that $f'(c) = 0$.

Now, suppose there are two roots a, b such that $f(a) = 0 = f(b)$,

$\Longrightarrow$ by Rolle's theorem,

$\exists\ d \in (a, b)$, s.t., $f'(d) = 0$

But,

$f'(x) = 13x^2 + 7 \geq 7 > 0$, which is a contradiction.

Mean Value Theorem (MVT): Let f be a function such that

- f is continuous on $[a, b]$
- f is differentiable on (a, b)

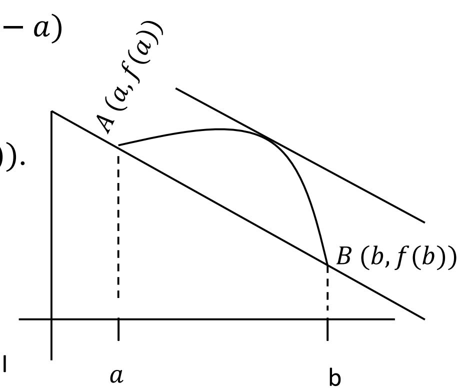
Then there is a number $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

or

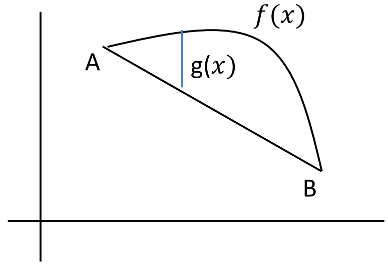
$$f(b) - f(a) = f'(c)(b - a)$$

Geometric interpretation: There is a point in (a, b) such that the tangent line at that point is parallel to the secant line going through points $(a, f(a))$ and $(b, f(b))$.



Proof:

Equation of a line AB,



$$y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$$

Let $g(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$

Recall that by Rolle's theorem, we have

1) $g(x)$ is continuous on $[a,b]$
2) $g(x)$ is differentiable on (a,b) such that $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$
3) $g(a) = f(a) - f(a) - \frac{f(b)-f(a)}{b-a}(a-a) = 0$

$$g(b) = f(b) - f(a) - \frac{f(b)-f(a)}{b-a}(b-a) = 0$$

So, $g(a) = g(b)$

By Rolle's theorem, there $\exists$ $c \in (a,b)$ s.t. $g'(c) = 0$

So,

$$0 = g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$$

Hence,

$$f'(c) = \frac{f(b)-f(a)}{b-a}$$

Example: Suppose we know that $f(x)$ is continuous and differentiable on $[-7,0]$, $f(-7) = -3$, and $f'(x) \leq 2$. What is the largest possible value for $f(0)$?

By MVT, there is $c \in [-7,0]$, such that,

$$f'(c) = \frac{f(0)-f(-7)}{0-(-7)} = \frac{f(0)-(-3)}{7}$$

i.e.,

$$(7)f'(c) = f(0) + 3$$
$$f(0) = (7)f'(c) - 3 \leq 7(2) - 3 = 11, \text{ (given } f'(c) \leq 2)$$

Remark: If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant on (a,b).

Proof:

Let $x_1, x_2 \in (a,b)$ s.t. $x_1 < x_2$

$f(x)$ is differentiable on $(a,b) \Longrightarrow$ it is diff. on (x_1, x_2) and, thus, continuous on $[x_1, x_2]$

By MVT, there $\exists$ $c \in (x_1, x_2)$ s.t.

$$f'(c) = \frac{f(x_2)-f(x_1)}{x_1, x_2} = 0 \Longrightarrow f(x_1) = f(x_2)$$

Since x_1, x_2 are any numbers from (a,b), f is constant on (a,b).

Example: We illustrate The Mean Value Theorem by considering $f(x) = x^3$ on the interval [1; 3].

Observe that f is a polynomial and so continuous everywhere. For any x we see that

$$f'(x) = 3x^2$$

So f is continuous on [1; 3] and differentiable on (1, 3). So the Mean Value theorem applies to f and [1; 3].
Thus,

$$\frac{f(b) - f(a)}{b - a} = \frac{f(3) - f(1)}{3 - 1} = \frac{27 - 1}{2} = 13$$

Hence,

$$f'(c) = 3c^2$$

So we seek c in [1,3] with

$$3c^2 = 13$$

iff

$$c^2 = \frac{13}{3} \quad \text{or} \quad c = \pm\sqrt{\frac{13}{3}}$$

However, $-\sqrt{\frac{13}{3}}$ is not in the interval (1,3), but $\sqrt{\frac{13}{3}}$ is a little bigger than $\sqrt{\frac{12}{3}} = \sqrt{4} = 2$.

So, $c = \sqrt{\frac{13}{3}}$ is in the interval (1,3), and

$$f'(c) = f'^{\left(\sqrt{\frac{13}{3}}\right)} = 13 = \frac{f(3) - f(1)}{3 - 1} = \frac{f(b) - f(a)}{b - a}$$

Let's look again at the two theorems together.

Rolle's Theorem: Let $a < b$. If f is continuous on $[a, b]$ and differentiable on (a, b) and $f(a) < f(b)$, then there is a c in (a, b) with $f'(c) = 0$.

Mean Value Theorem: Let $a < b$. If f is continuous on $[a, b]$ and differentiable on (a, b) then there is a c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Corollary: If $f'(x) = g'(x)$ for all $x \in (a, b)$, then $f - g$ is constant on (a, b).
Proof:

$$h(x) = f(x) - g(x)$$

Thus,

$h'(x) = f'(x) - g'(x) = 0$ for all $x \in (a, b)$

By previous fact, $h(x) =$ const.

$\Longrightarrow f - g =$ const.

How Derivatives Affect the Shape of a Graph

Increasing/Decreasing Test:

- If $f'(x) > 0$ on an interval, then f is increasing on that interval.
- If $f'(x) < 0$ on an interval, then f is decreasing on that interval.

Proof:

Let x_1, x_2 and if $f'(x) > 0$, then $f(x_1) < f(x_2)$

Thus, $f(x)$ is differentiable on (x_1, x_2)

By MVT, there is $c \in (x_1, x_2)$, such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_1 - x_2}$$

or,

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0$$

$$\Longrightarrow \qquad f(x_2) > f(x_1)$$

Thus, $f(x)$ is increasing.

Example: Find the intervals where $f(c) = 4x^3 + 3x^2 - 6x + 1$ is increasing or decreasing.

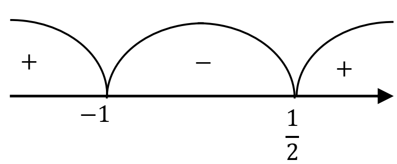
$f'(x) = 12x^2 - 6x - 6 = 0$ or $2x^2 - x - 1 = 0$

i.e.,

$(2x - 1)(x + 1) = 0$, which gives, $x = \frac{1}{2}, -1$ are critical points.

Hence, $f(x)$ is increasing

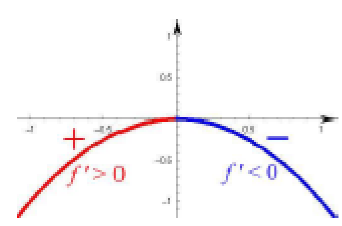
on $(-\infty, -1) \cup \left(\frac{1}{2}, \infty\right)$ or $\left(-1, \frac{1}{2}\right)$



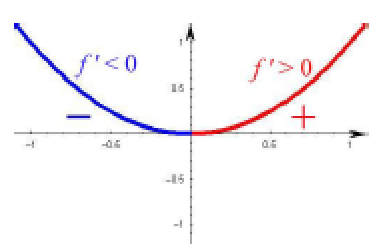
First Derivative Test:

Let c be a critical number of a continuous function f. Then,

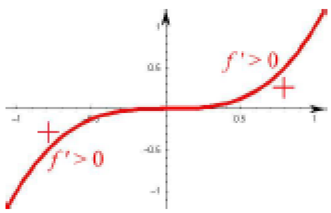
- If f' changes from positive to negative at c, then has a local max at c.

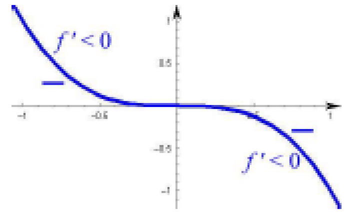


- If f' changes from negative to positive at c, then f has a local min at c.



- If f does not change sign at c, then f has no local max or min at c.



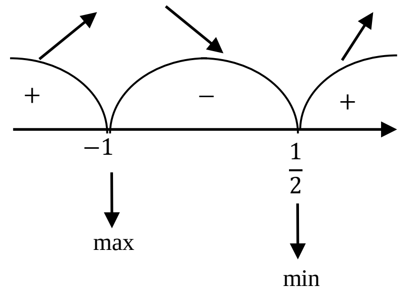


Example: Given $f(x) = 4x^3 + 3x^2 - 6x + 1$. Find the local minimum/maximum values.

Recall from the example above that the critical points are $\left(-1, \frac{1}{2}\right)$.

Hence,

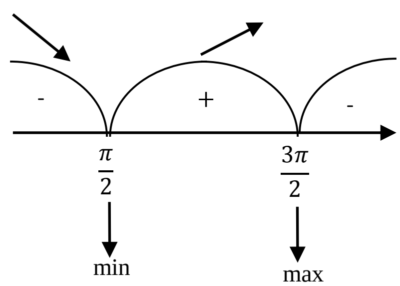
$$f(-1) = 6 \quad \text{and} \quad f\left(\frac{1}{2}\right) = -0.75$$



Therefore, $(-1,6)$ is local max.

$\left(\frac{1}{2}, -0.75\right)$ is local min.

Example: Given that $(x) = \cos^2 x - 2\sin x, \ \ 0 \le x \le 2\pi$, find the local minimum/maximum values of f.



$$f'(x) = 2\cos x(-\sin x) - 2\cos x = 0$$

i.e.,

$$-2\cos x(1 + \sin x) = 0$$

which gives,

$$\cos x = 0 \quad \text{or} \quad \sin x = -1$$

Thus,

$$x = \frac{\pi}{2}, \frac{3\pi}{2}$$

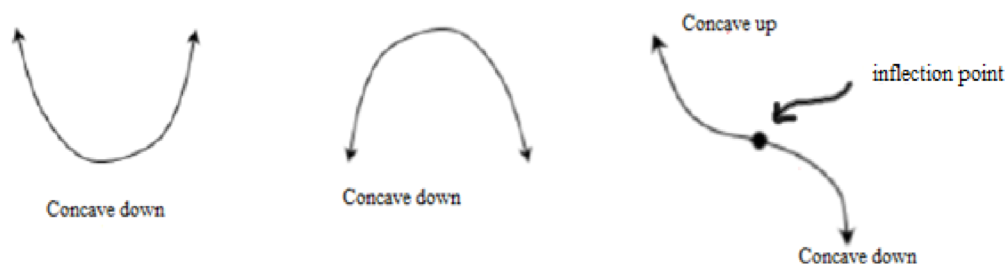
The critical points are: $f\left(\frac{\pi}{2}\right) = -2$, and $f\left(\frac{3\pi}{2}\right) = 2$.

Therefore,

$\left(\frac{\pi}{2}, -2\right)$ – local min. and $\left(\frac{3\pi}{2}, 2\right)$ – local max.

Concavity and the Second Derivative Test

The first derivative describes the direction of the function. The second derivative describes the concavity of the original function. Concavity describes the direction of the curve, how it bends...



Just like direction, concavity of a curve can change, too. The points of change are called *inflection points*.

Concavity Test:

- If $f''(x) > 0$ on an interval, then the graph of f is concave up (CU) on that interval.
- If $f''(x) < 0$ on an interval, then the graph of f is concave down (CD) on that interval.

Definition: A point P on a curve $y = f(x)$ is called an inflection point (IP) if f is continuous there and the curve changes from CU to CD or vise versa at P.

Example: Given that $f(x) = x^2 \ln x$, find the intervals of concavity and the inflection points.

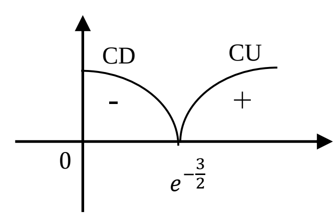
$$f'(x) = x^2\left(\frac{1}{x}\right) + 2x(\ln x) = x + 2x(\ln x)$$

and

$$f''(x) = 1 + 2x\left(\frac{1}{x}\right) + 2(\ln x) = 3 + 2\ln x = 0$$

i.e.,

$$\ln x = -\frac{3}{2} \text{ or } x = e^{-\frac{3}{2}}$$



So, $f(x)$ is CD on $\left(0, e^{-\frac{3}{2}}\right)$ and CU on $\left(e^{-\frac{3}{2}}, \infty\right)$

Inflection point (IP) is $\left(e^{-\frac{3}{2}}, f\left(e^{-\frac{3}{2}}\right)\right) = (0.223, f(0.223))$

Second Derivative Test: Let f'' be continuous near c:

- If $f'(c) = 0$ and $f''(c) > 0$, then f has a local minimum at c, (CU).
- If $f'(c) = 0$ and $f''(c) < 0$, then f has a local maximum at c, (CD).

Example: $f(x) = \frac{x^2}{x-1}$, find the local minimum and maximum values.

$$f'(x) = \frac{2x(x-1) - x^2}{(x-1)^2} = \frac{x^2 - 2x}{(x-1)^2} = \frac{x(x-2)}{(x-1)^2} = 0$$

Thus,

$x = 0,2$

and

$$f''(x) = \frac{(2x-1)(x-1)^2-(x^2-2x)[2(x-1)]}{(x-1)^4} = \frac{2[(x-1)^2-(x^2-2x)]}{(x-1)^3} = \frac{2}{(x-1)^3}$$

It follows that,

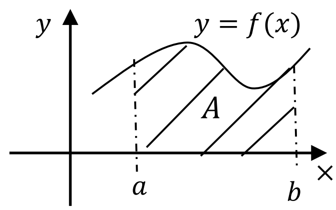
$$f''(0) = -2 < 0 \rightarrow \text{max.}$$
$$f''(2) = 2 > 0 \rightarrow \text{ min.}$$

Application of Definite Integral (Area between Curves)

<u>Area of Region between Curves</u>

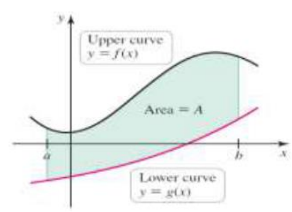
1. If $f(x) \geq 0$ on $[a,b]$, then the area under the curve $y = f(x)$ over $[a,b]$ is,

$$A = \int_a^b f(a)dx$$



2. If $f(x)$ and $g(x)$ are continuous with $f(x) \geq g(x)$ on $[a,b]$, then the area of the region between the curves $y = f(x)$ and $y = g(x)$ from a to b is

$$A = \int_a^b [f(x) - g(x)]dx$$

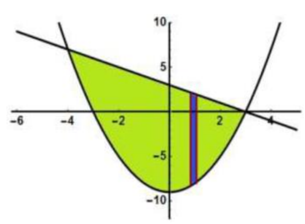


Example: Find the area bounded by the graph of $y = 3 - x$ and $y = x^2 - 9$.

To get points of intersection, we solve

$3 - x = x^2 - 9 \implies (x+4)(x-3) = 0$. Hence, $x = -4, 3$.

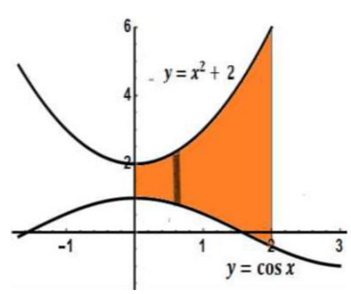
Thus, $\quad A = \int_{-4}^{3} [(3-x) - (x^2 - 9)]dx = \frac{343}{6}$



Example: Find the area of the region bounded by the graphs of $y = \cos x$, $y = x^2 + 2$ for $0 \leq x \leq 2$.

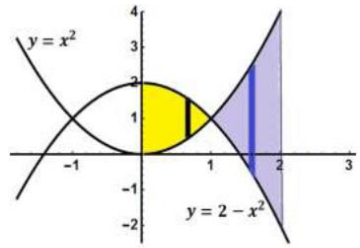
$$A = \int_0^2 [y_{top} - y_{bottom}]dx = \int_0^2 [(x^2+2) - \cos x]dx$$

$$= \frac{20}{3} - \sin 2$$



Note: Sometimes, the upper and lower boundary is not defined by a single rule as in the following example:

Example: Find the area bounded by the graphs of $y = x^2$ and $y = 2 - x^2$ for $0 \le x \le 2$.



The point of intersection is, $x^2 = 2 - x^2$ or $x = \pm 1$

Thus,

$$A = \int_0^1 [(2 - x^2) - x^2]dx + \int_1^2 [x^2 - (2 - x^2)]dx = 4$$

Note: Some regions are best treated by regarding $\boldsymbol{x}$ as a function of $\boldsymbol{y}$. If a region is bounded by curves with equations

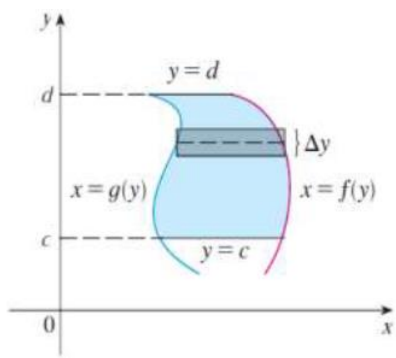
$$x = f(y), \quad x = g(y), \quad y = c \text{ and } y = d$$

where

$$f(y) \ge g \ge (y) \quad \text{on} \quad c \le y \le d$$

Then this area is given by;

$$A = \int_c^d [f(y) - g(y)]dy$$



Example: Find the area enclosed by the line $y = x - 1$ and the parabola $y^2 = 2x + 6$.

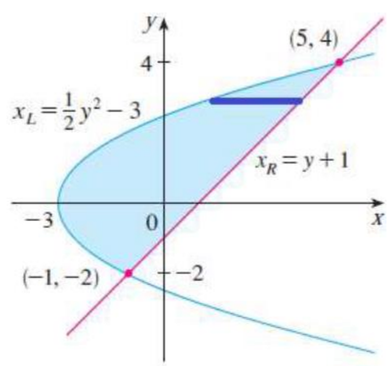
To get the intersection, we have;

$x = y + 1$ and $x = \frac{y^2}{2} - 3$

i.e, $y^2 - 2y - 8 = 0$ or $y = -2, 4$.

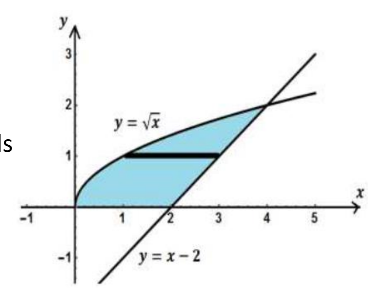
Thus,

$$A = \int_{-2}^4 [x_{right} - x_{left}]dy$$
$$= \int_{-2}^4 \left[(y + 1) - \left(\frac{y^2}{2} - 3\right)\right] dy = 18$$



Example: Find the area of the region in the 1st quadrant that is bounded above by $y = \sqrt{x}$ and below by the $\boldsymbol{x - axis}$ and the line $\boldsymbol{y = x - 2}$.

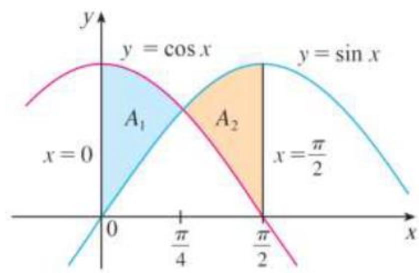
The intersection points are:



$x = y^2$ and $x = y + 2$
i.e.,
$y^2 - y - 2 = 0$ or $y = -1, 2$
Thus,

$$A = \int_0^2 [(y+2) - y^2]dy = \frac{10}{3}$$

Example: Find the area of the region bounded by the curves $y = \sin x$, $y = \cos x, x = 0$ and $x = \frac{\pi}{2}$.



The intersection points are:
$\sin x = \cos x$, $0 \le x \le \frac{\pi}{2}$
Thus,
$\tan x = 1$ or $x = \frac{\pi}{4}$
Observe that $\cos x \ge \sin x$ when $0 \le x \le \frac{\pi}{4}$,
but $\sin x \ge \cos x$ when $\frac{\pi}{4} \le x \le \frac{\pi}{2}$
Therefore the required area is:

$$A = A_1 + A_2$$

$$A_1 = \int_0^{\frac{\pi}{4}} (\cos x - \sin x)dx = \frac{1}{\sqrt{2}} - 1$$

and

$$A_2 = \int_{\frac{\pi}{4}}^{\pi_2} (\sin x - \cos x)dx = -1 + \frac{2}{\sqrt{2}}$$

Thus,

$$A = A_1 + A_2 = 2\sqrt{2} - 2$$

Exercises

Sketch and find the area bounded by the graphs:
1. $y = 4 - x^2$, $y = -x + 2$
2. $y = e^x$, $y = 2e^{-x} + 1$, $x = 0$

3. $y = x^2$, $y = x$, $x = 0$, $x = 2$.

4. $y = \cos x$, $y = 2 - \cos x$, $x = 0$, $x = 2\pi$.

5. $x = y^2$, $y = x - 2$.

6. $y = \frac{4}{2-x}$, $y = 4$, $x = 0$.

7. $y = \sin x$, $y = \cos x$, $x = 0$, $x = \frac{\pi}{2}$.

8. $y = x^2$, $y = x^3$, $x = 0$, $x = 2$.

9. $y = x^2 + 2x$, $y = -x + 4$, $x = -2$, $x = -4$.

10. $y = \tan x$, $y = 0$, $x = -\frac{\pi}{4}$, $x = \frac{\pi}{3}$.

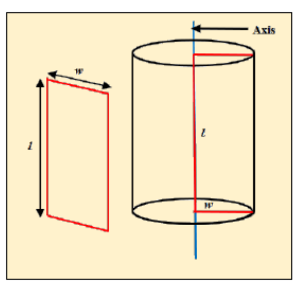
11. $y = \sqrt{x} - 1$, $x = 0$, $x = 4$.

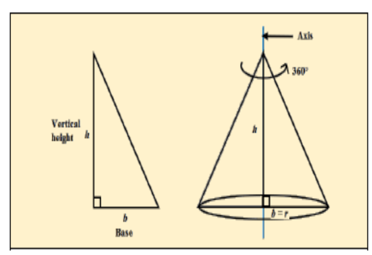
12. $y = -x^2 - 2x$, $y = x^2 - 4$.

Applications of The Definite Integral - The volume of a solid of revolution

If we rotate a plane figure about a straight line (called the axis) through a complete revolution or 360°, it sweeps out a three dimensional (3D) region. The shape of the 3D region depends on the shape of the 2D region. The solids obtained by this process are called solids of revolution. The volume of such a solid obtained by rotation is the volume of a solid of revolution.

If a rectangle is rotated through one complete turn about its length, the solid of revolution will be a cylinder. We can visualize a cylinder as the shape swept out by the rectangle as it rotates a full turn of 360° or one complete revolution or through 2π radians.





Similarly, if a triangle is rotated through one complete revolution about its vertical height, a cone is formed.

Rotation of regions on the Cartesian Plane

We can form solids of revolution by the rotation of regions about the vertical or horizontal axes on the Cartesian Plane. If the plane region has a definite shape, then the solid will have a definite shape as well.

In the computation of the volume of a solid shape, as long as the rotational solid resulting from your graph has no hollow space in it, we can use the disk method which we shall however restrict ourselves in our discussion.

The Disk Method

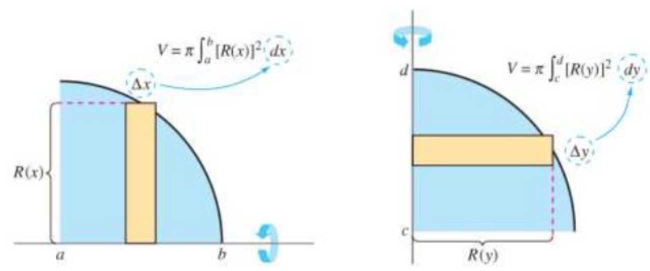
To find the volume of a solid of revolution with the disk method, we use one of the following:

- Horizontal Axis of Revolution

Volume, $V = \pi \int_a^b [R(x)]^2 dx$

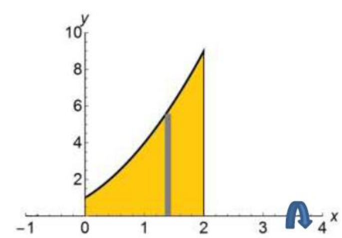
- Vertical Axis of Revolution

Volume, $V = \pi \int_c^d [R(y)]^2 dy$



Horizontal axis of revolution   Vertical axis of revolution

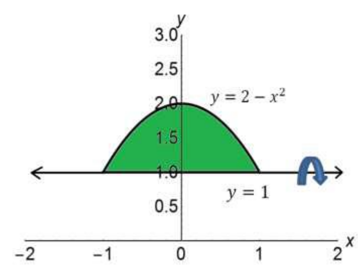
Example: Let R be the region bounded by the curve $y = (x+1)^2$, the $x-$ axis, and the lines $x = 0$ and $x = 2$. Find the volume of the solid of revolution obtained by revolving R about the $x-$ axis.



Volume, $V = \pi \int_0^2 [(x+1)^2]^2 dx = \pi \int_0^2 (x+4)^2 dx = \frac{242\pi}{5}$

Note: Revolving about a line that is not a coordinate axis.

Example: Find the volume of the solid formed by revolving the region bounded by $f(x) = 2 - x^2$ and $g(x) = 1$ about the line $y = 1$.



Solve $2 - x^2 = 1$ to determine that the limits of integration are ± 1, and

$$R(x) = (2 - x^2) - 1 = 1 - x^2$$

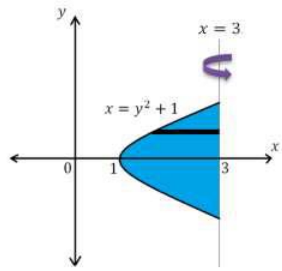
Thus, the volume is given by,

$$V = \pi \int_{-1}^{1} (1 - x^2)^2 dx = \frac{16\pi}{5}$$

Example: Find the volume of the solid generated by revolving the region between the parabola $x = y^2 + 1$ and the line $x = 3$ about the line.

$$R(y) = 3 - (y^2 + 1) = 2 - y^2$$

$$V = \pi \int_{-\sqrt{2}}^{\sqrt{2}} (2 - y^2)^2 dy = \frac{64\sqrt{2}\,\pi}{15}$$



Line and Multiple Integrals

We already know how to perform integrals like

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

This integral of a single variable is the simplest example of a 'line integral'. A line integral is just an integral of a function along a path or curve – which we already discussed under the applications of integration in our earlier lecture. In this case, the curve is a straight line – a segment of the x-axis that starts at $x = a$ and ends at $x = b$. Just so we're clear on notation, I'll write the indefinite integral of $f(x)$ with respect to x as

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

where c is a constant, and the definite integral of $f(x)$ from $x = a$ to $x = b$ as

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

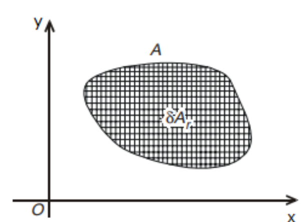
Multiple (Double) Integrals

The double integral is the analogue of the single integral and it shows the volume V bounded by the plane region D and the surface, $z = f(x, y)$, and it corresponds to the double integral

$$V = \iint_D f(x, y)dxdy$$

Here, instead of closed intervals $[a, b]$ in the line, we deal with closed regions D in the plane.

Now let $f(x, y)$ be a single valued and bounded function of two independent variables x and y defined in a closed region A in xy-plane. Let A be divided into n elementary areas, $\delta A_1, \delta A_2, \dots, \delta A_n$; and let (x_r, y_r) be any point inside the rth elementary area δA_r.



Consider the sum

$$f(x_1, y_1)\delta A_1 + f(x_2, y_2)\delta A_2 + \cdots + f(x_n, y_n)\delta A_n = \sum_{r=1}^{n} f(x_r, y_r)\delta A_r$$

Then the limit of the sum, if it exists, as $n \to \infty$ and each sub-elementary area approaches to zero, is termed as a double integral of $f(x, y)$ over the region A, and expressed as

$$\iint_A f(x, y)dA$$

Thus,

$$\iint_A f(x, y)dA = \lim_{\substack{n\to\infty \\ \delta A_r \to 0}} \sum_{r=1}^{n} f(x_r, y_r)\delta A_r$$

Evaluation of Double Integral

Evaluation of double integral

$$\iint_R f(x, y)dxdy$$

can be discussed under following three possible cases:

Case I: When the region R is bounded by two continuous curves $y = \psi\,(x)$ and $y = \varphi(x)$ and the two lines (ordinates) $x\ = a$ and $x = b$.

In such a case, integration is first performed with respect to y keeping x as a constant and then the resulting integral is integrated within the limits $x = a$ and $x = b$.

Mathematically expressed as:

$$\iint_R f(x, y)dxdy = \int_{x=a}^{x=b} \left(\int_{y=\phi(x)}^{y=\psi(x)} f(x, y)dy \right) dx$$

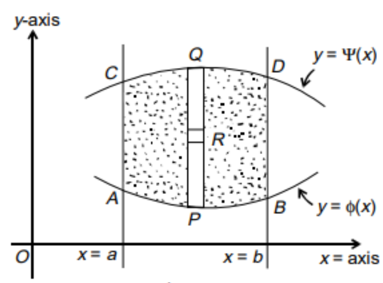


Fig. 2

Geometrically, the process is shown in Fig. 2, where integration is carried out from inner rectangle (i.e., along the one edge of the ‘vertical strip PQ’ from P to Q) to the outer rectangle.

Case 2: When the region R is bounded by two continuous curves $x = \varphi(y)$ and $x = \psi(y)$ and the two lines (abscissa) $y = a$ and $y = b$.

In such a case, integration is first performed with respect to x keeping y as a constant and then the resulting integral is integrated between the two limits $y = a$ and $y = b$.

Mathematically expressed as:

$$\iint_R f(x,y)dxdy = \int_{y=a}^{y=b} \left(\int_{x=\phi(y)}^{x=\psi(y)} f(x,y)dx\right) dy$$

Geometrically, the process is shown in Fig. 3, where integration is carried out from inner rectangle (i.e., along the one edge of the horizontal strip PQ from P to Q) to the outer rectangle.

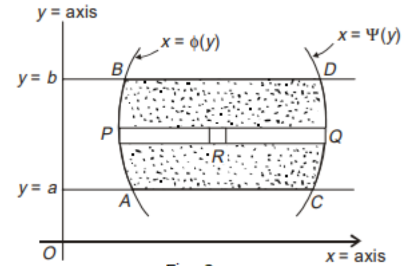


Fig. 3

Case 3: When both pairs of limits are constants, the region of integration is the rectangle ABCD (say). In this case, it is immaterial whether $f(x,y)$ is integrated first with respect to x or y, the result is unaltered in both the cases (Fig. 4).

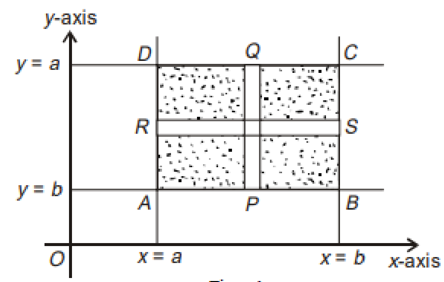
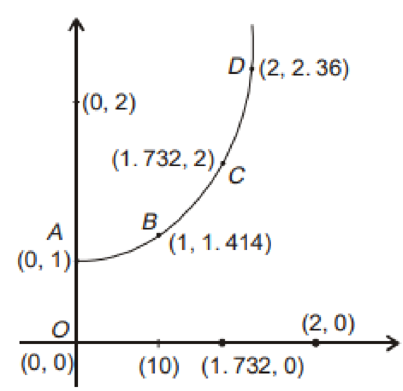


Fig. 4

Observations: While calculating double integral, in either case, we proceed outwards from the innermost integration and this concept can be generalized to repeated integrals with three or more variable also.

Example 1: Evaluate $\int_0^1 \int_0^{\sqrt{1+x^2}} \frac{1}{(1+x^2+y^2)} dydx$

Solution: Clearly, here $y = f(x)$ varies from 0 to $\sqrt{1+x^2}$ and finally x (as an independent variable) goes between 0 to 1.



$$I = \int_0^1 \left(\int_0^{\sqrt{1+x^2}} \frac{1}{(1+x^2)+y^2} dy \right) dx$$

$$= \int_0^1 \left(\int_0^{\sqrt{1+x^2}} \frac{1}{a^2+y^2} dy \right) dx, \; a^2 = (1+x^2)$$

$$= \int_0^1 \left(\frac{1}{a} \tan^{-1} \frac{y}{a} \right)_0^{\sqrt{1+x^2}} dx$$

$$= \int_0^1 \frac{1}{\sqrt{1+x^2}} \left(\tan^{-1} \frac{\sqrt{1+x^2}}{\sqrt{1+x^2}} - \tan^{-1} 0 \right) dx$$

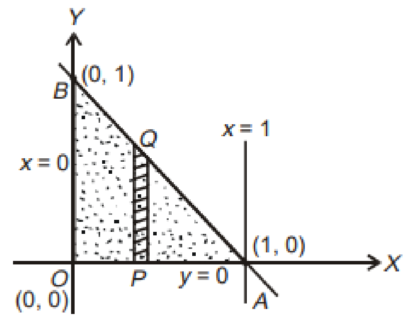
$$= \int_0^1 \frac{1}{\sqrt{1+x^2}} \left(\frac{\pi}{4} - 0 \right) dx = \frac{\pi}{4} \left[\log\left\{x + \sqrt{1+x^2}\right\} \right]_0^1$$

$$= \frac{\pi}{4} \log\left(1+\sqrt{2}\right)$$

Example 2: **Evaluate $\iint e^{2x+3y}\,dxdy$ over the triangle bounded by the lines $x = 0$, $y = 0$ and $x + y = 1$.**

Solution: Here the region of integration is the triangle $OABO$ as the line $x + y = 1$ intersects the axes at points (1, 0) and (0, 1). Thus, precisely the region R (say) can be expressed as:

$$0 \le x \le 1,\ 0 \le y \le 1 - x$$



$$\therefore \qquad I = \iint_R e^{2x+3y}dxdy$$

$$= \int_0^1\left(\int_0^{1-x} e^{2x+3y}dy\right)dx$$

$$= \int_0^1\left[\frac{1}{3}e^{2x+3y}\right]_0^{1-x} dx$$

$$= \frac{1}{3}\int_0^1 \left(e^{3-x} - e^{2x}\right)dx$$

$$= \frac{1}{3}\left[\frac{e^{3-x}}{-1} - \frac{e^{2x}}{2}\right]_0^1$$

$$= \frac{-1}{3}\left[\left(e^2 + \frac{e^2}{2}\right) - \left(e^3 + \frac{1}{2}\right)\right]$$

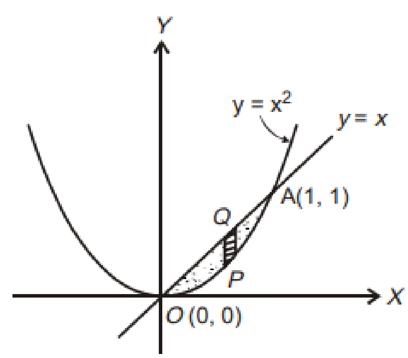
$$= \frac{1}{6}\left[2e^3 - 3e^2 + 1\right] = \frac{1}{6}\left[(2e+1)(e-1)^2\right].$$

Example 3: **Evaluate the integral $\iint_R xy(x+y)dxdy$ over the area between the curves $y = x^2$ and $y = x$.**

Solution: We have $y = x^2$ and $y = x$ which implies $x^2 - x = 0$ i.e. either $x = 0$ or $x = 1$

Further, if $x = 0$ then $y = 0$; if $x = 1$ then $y = 1$. Means the two curves intersect at points (0, 0), (1, 1).

$\therefore$ The region R of integration is doted and can be expressed as: $0 \le x \le 1,\ x^2 \le y \le x$.



$$\therefore \qquad \iint_R xy(x+y)dxdy = \int_0^1\left(\int_{x^2}^{x} xy(x+y)dy\right)dx$$

$$= \int_0^1\left\{\left(x^2\frac{y^2}{2} + x\frac{y^3}{3}\right)_{x^2}^{x}\right\}dx$$

$$= \int_0^1\left\{\left(\frac{x^4}{2} + \frac{x^4}{3}\right) - \left(\frac{x^6}{2} + \frac{x^7}{3}\right)\right\}dx$$

$$= \int_0^1 \left(\frac{5}{6}x^4 - \frac{1}{2}x^6 - \frac{1}{3}x^7 \right) dx$$

$$= \left[\frac{5}{6} \times \frac{x^5}{5} - \frac{1}{2}\frac{x^7}{7} - \frac{1}{3}\frac{x^8}{8} \right]_0^1 = \frac{1}{6} - \frac{1}{14} - \frac{1}{24} = \frac{3}{56}$$

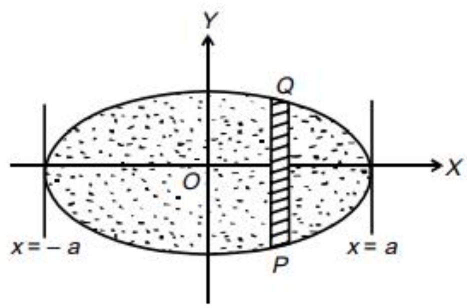
Example 4: Evaluate $\iint (x+y)^2\, dxdy$ over the area bounded by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Solution: For the given ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, the region of integration can be considered as bounded by the curves $y = -b\sqrt{1 - \frac{x^2}{a^2}}$, $y = b\sqrt{1 - \frac{x^2}{a^2}}$ and finally x goes from $-a$ to a

$\therefore$
$$I = \iint (x+y)^2\, dxdy = \int_{-a}^{a} \left(\int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} (x^2 + y^2 + 2xy)\, dy \right) dx$$

$$I = \int_{-a}^{a} \left(\int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} (x^2 + y^2)\, dy \right) dx$$

[Here $\int 2xy\, dy = 0$ as it has the same integral value for both limits i.e., the term xy, which is an odd function of y, on integration gives a zero value.]



$$I = 4\int_0^a \left(\int_0^{b\sqrt{1-x^2/a^2}} (x^2 + y^2)\, dy \right) dx$$

$$I = 4\int_0^a \left[x^2 y + \frac{y^3}{3} \right]_0^{b\sqrt{1-x^2/a^2}} dx$$

$\Rightarrow$
$$I = 4\int_0^a \left[x^2 b \left(1 - \frac{x^2}{a^2}\right)^{1/2} + \frac{b^3}{3}\left(1 - \frac{x^2}{a^2}\right)^{3/2} \right] dx$$

On putting $x = a\sin\theta$, $dx = a\cos\theta\, d\theta$; we get

$$I = 4b\int_0^{\pi/2} \left(\left(a^2 \sin^2\theta \cos\theta\right) + \frac{b^3}{3}\cos^3\theta \right) a\cos\theta d\theta$$

$$= 4ab\int_0^{\pi/2} \left(a^2 \sin^2\theta \cos^2\theta + \frac{b^3}{3}\cos^4\theta \right) d\theta$$

Now using formula $\int_0^{\pi/2} \sin^p x \cos^q x \, dx = \dfrac{\frac{1}{2}\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{q+1}{2}\right)}{\Gamma\left(\frac{p+q+2}{2}\right)}$

and $\quad \int_0^{\pi/2} \cos^n x \, dx = \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} \dfrac{\sqrt{\pi}}{2}$, $\qquad$ (in particular when $p = 0$, $q = n$)

$$\iint (x+y)^2 \, dx dy = 4ab\left\{a^2 \frac{\Gamma\frac{3}{2}\Gamma\frac{3}{2}}{2\Gamma 3} + \frac{b^2}{3}\frac{\Gamma\frac{5}{2}\Gamma\frac{1}{2}}{2\Gamma 3}\right\}$$

$$= 4ab\left\{a^2 \frac{\frac{\sqrt{\pi}}{2}\frac{\sqrt{\pi}}{2}}{2.2.1} + \frac{b^2}{3}\frac{\frac{3}{2}\frac{\sqrt{\pi}}{2}\sqrt{\pi}}{2.2.1}\right\} \qquad \left| \because \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\right.$$

$$= 4ab\left\{\frac{\pi a^2}{16} + \frac{\pi b^2}{16}\right\} = \frac{\pi ab\left(a^2+b^2\right)}{4}$$

ASSIGNMENT

1. Evaluate $\int_0^1 \int_0^1 \dfrac{dx\,dy}{\sqrt{\left(1-x^2\right)\left(1-y^2\right)}}$

2. Evaluate $\iint_R xy \, dx \, dy$, where A is the domain bounded by the x-axis, ordinate $x = 2a$ and the curve $x^2 = 4ay$.

3. Evaluate $\iint e^{ax+by} dy \, dx$, where R is the area of the triangle $x = 0$, $y = 0$, $ax + by = 1$ ($a > 0$, $b > 0$). [Hint: See example 2]

4. Prove that $\int_1^2 \int_3^1 \left(xy + e^y\right) dy \, dx = \int_3^1 \int_1^2 \left(xy + e^y\right) dx \, dy$.

5. Show that $\int_0^1 dx \int_0^1 \dfrac{x-y}{(x+y)^3} dy \neq \int_0^1 dy \int_0^1 \dfrac{x-y}{(x+y)^3} dx$.

6. Evaluate $\int_0^\infty \int_0^\infty e^{-x^2\left(1+y^2\right)} x \, dx \, dy$ $\qquad$ [Hint: Put $x^2(1 + y^2) = t$, taking y as const.]

Taylor Series and Maclaurin Series

Let's start our discussion with a function that can be represented by a power series. A series

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

is called a *power series* . The domain of the power series function is the set of all x values for which the series converges.

Here is a simple example to demonstrate that in the typical power series you will have convergence for some values of x and divergence for others.

$$\sum_{n=0}^{\infty} x^n$$

This series is quite clearly a geometric series, and converges for $|x| < 1$.

A basic variant of the power series is the *power series centered at* $x = a$:

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_2(x-a)^3 + \cdots$$

We first notice that

$$f(a) = c_0$$

We can find the derivative of $f(x)$ by differentiating the individual terms of the power series,

$$f'(x) = \sum_{n=1}^{\infty} n c_n (x-a)^{n-1}$$

Observe that the derivative is also a power series, so we can proceed to compute all of its higher derivatives.

$$f''(x) = \sum_{n=2}^{\infty} n(n-1) c_n (x-a)^{n-2}$$

$$f'''(x) = \sum_{n=3}^{\infty} n(n-1)(n-2) c_n (x-a)^{n-3}$$

$$\cdots$$

$$f^{(k)}(x) = \sum_{n=3}^{\infty} n(n-1)(n-2)\ldots(n-k+1) c_n (x-a)^{n-k}$$

When we evaluate the derivatives at a, we get the constant term in each power series,

$$f'(a) = 1.c_1; \quad f''(a) = 2.1.c_2; \quad f'''(a) = 3.2.1.c_3; \quad \ldots \quad f^{(k)}(a) = k!.c_k$$

Solving the equation for the k-th coefficient c_k, we get

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

We have proved the following theorem.

Theorem: If f has a power series expansion at a with radius of convergence $R > 0$, that is

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \text{ for all } |x-a| < R,$$

then its coefficients are given by the formula,

$$c_n = \frac{f^{(n)}(a)}{n!}$$

Remark: Substituting this formula back into the series, we see that if f has a power series expansion at a, then it must be of the form:

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

$$= f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^n(a)}{n!}(x-a)^n + R_n(x)$$

Where $R_n(x)$ is the remainder or error, and

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

Definition: Let f be a function with all derivatives in the open interval $(a-r, a+r)$, then the series

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

is called the Taylor series of the function f at a on $(a-r, a+r)$ if and only if

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

with

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}, \ \ c \in (a-r, a+r)$$

When $a = 0$, the series becomes

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

and is given the special name Maclaurin series.

Example: Let us look at the power series expansion of the exponential function $f(x) = e^x$ centred at 0,

$$f(x) = e^x = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^n(0)}{n!}x^n$$

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

which is a power series expansion of the exponential function $f(x) = e^x$. The power series is centered at 0. The derivatives $f^k(x) = e^x$, so $f^k(0) = e^0 = 1$. So the Taylor series of the function f at 0, or the Maclaurin series of f, is

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

which agrees with the power series definition of the exponential function.

Definition: If $f(x)$ is the sum of its Taylor series expansion, it is the limit of the sequence of partial sums,

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

We call the n-th partial sum the n-th-degree Taylor polynomial of f at a.

One important application of Taylor series is to approximate a function by its Taylor polynomials. This is very useful in physics and engineering, where people only need a good approximation for most scenarios, and polynomials are usually much easier to deal with than a transcendental function. The following theorem justifies the use of Taylor polynomials for function approximation.

Theorem (Taylor's Theorem): Let $n > 1$ be an integer, and let $a \in \mathbb{R}$ be a point. If $f(x)$ is a function that is n times differentiable at the point a, then there exists a function $h_n(x)$ such that,

$$f(x) = T_n(x) + h_n(x)(x-a)^n$$

where

$$\lim_{x \to a} h_n(x) = 0$$

The term

$$R_n(x) = f(x) - T_n(x) = h_n(x)(x-a)^n$$

is called the *Peano form of the remainder*.

Sometimes we would like a better estimate on the remainder term, so that we could have a better understanding of how good the Taylor polynomials approximate the original functions. However, we can only do this under stronger regularity assumptions on $f(x)$.

Theorem (Lagrange Form of the Remainder). Let $n \geq 1$ be an integer, and let $a \in \mathbb{R}$ be a point. If $f(x)$ is a function that is $n+1$ times differentiable on an open interval I containing a, then for all $x \in I$, there exists a number z strictly between a and x such that,

$$R_n(x) = \frac{f^{(n+1)}(z)}{(n+1)!}(x-a)^{n+1}$$

This is the *Lagrange form of the remainder*.

Example: Find the Maclaurin series for $f(x) = \sin x$, and show that its sum equals $\sin x$.

First, we need to find the derivatives of $f(x)$ at 0:

$$\begin{aligned} f(x) &= \sin x, & f(0) &= 0 \\ f'(x) &= \cos x, & f'(0) &= 1 \\ f''(x) &= -\sin x, & f''(0) &= 0 \\ f'''(x) &= -\cos x, & f'''(0) &= -1 \\ f^{(4)}(x) &= \sin x, & f^{(4)}(0) &= 0 \\ &\dots & &\dots \end{aligned}$$

The derivatives repeat in a 4-cycle, so we can write the Maclaurin series as,

$$f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdots$$

$$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

To show that the sum of the Maclaurin series equals to the function $f(x) = \sin x$, we consider the n-th remainder term in Lagrange form:

$$R_n(x) = \frac{f^{(n+1)}(z)}{(n+1)!} x^{n+1}$$

where z is a number strictly between 0 and x. Notice that $f^{(n+1)}(z)$ is a sine function or a cosine function, so $|f^{(n+1)}(z)| \le 1$. Then we have,

$$-\frac{x^{n+1}}{(n+1)!} \le R_n(x) \le \frac{x^{n+1}}{(n+1)!}$$

However, we know that,

$$\lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} = 0$$

For all $x \in \mathbb{R}$, so that,

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

for all $x \in \mathbb{R}$. But since $R_n(x) = f(x) - T_n(x)$, this implies that the Taylor polynomials converges to $f(x)$ for all $x \in \mathbb{R}$, i.e., the sum of the Maclaurin series equals $f(x) = \sin x$.

Example: Find the Taylor series for $f(x) = e^x$ at $a = 1$.

All derivatives of $f(x)$ are e^x, so $f^{(n)}(1) = e$ for all $n \ge 0$. Thus its Taylor series at 1 is

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

with radius of convergence $R = \infty$. The following transformation verifies that we found the right expression for the Taylor series:

$$e^x = e \cdot e^{x-1} = e \sum_{n=0}^{\infty} \frac{(x-1)^n}{n!} = \sum_{n=0}^{\infty} \frac{e}{n!}(x-1)^n.$$

Exercise: Find the Maclaurin series of $f(x) = \cos(x)$.

First, we find the derivatives of $f(x)$ at 0:

$$\begin{aligned}
f(x) &= \cos(x), & f(0) &= 1,\\
f'(x) &= -\sin(x), & f'(0) &= 0,\\
f''(x) &= -\cos(x), & f''(0) &= -1,\\
f'''(x) &= \sin(x), & f'''(0) &= 0,\\
f^{(4)}(x) &= \cos(x), & f^{(4)}(0) &= 1,\\
&\cdots\cdots & &\cdots\cdots
\end{aligned}$$

The derivatives repeat in a 4-cycle, so we can write the Maclaurin series as

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

Example: Find the Maclaurin series for $f(x) = x\cos(x)$.

We know that the Maclaurin series for $\cos(x)$ is

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

Thus

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

Example: Find the Maclaurin series for $f(x) = e^{-x^2}$.

We know that the Maclaurin series for the exponential function e^u is

$$e^u = \sum_{n=0}^{\infty} \frac{u^n}{n!}$$

Substitute $u = -x^2$ in the expression above, we get

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

Remark: The Taylor series/Maclaurin series of a infinitely differentiable function does not necessarily equal to the original function. A proof is required to show that they are equal (or not equal) for a function under consideration. We used the Lagrange form of the remainder to prove it for $\sin x$ and used the differential equation method to prove it for e^x.

We collect the following table of important Maclaurin series for reference:

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

Example: Find the limit $\lim_{x\to 0} \dfrac{\sin(x)}{x}$.

We could evaluate the limit with L'Hôpital's Rule, but let's use the Maclaurin series instead.

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots}{x} = \lim_{x\to 0} \left(1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \cdots\right) = 1.$$

The result agrees with the answer we get from L'Hôpital's Rule:

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{\cos(x)}{1} = \frac{\cos(0)}{1} = 1.$$

Another important application of Taylor series is that they enable us to integrate functions that we previous could not handle.

Example : Evaluate $\int e^{-x^2}\, dx$ as an infinite series.

We know the Maclaurin series for the integrand is

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

with radius of convergence $R = \infty$. Using term-by-term integration, we get

$$\int e^{-x^2}\,dx = C + \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(2n+1)} x^{2n+1}$$

with radius of convergence $R = \infty$.

Exercises:

1. Write the Taylor series for $f(x) = \frac{1}{x}$ centered at $a = 1$.
2. Find the Taylor series for $f(x) = \sin x$ in $(x - \pi/4)$.
3. Compute and write a Maclaurin series (5 terms) for $f(x) = \frac{1}{1-\sin x}$.
4. Find the Taylor series for the function $x^4 + x - 2$ centered at $a = 1$.
5. Find the first 4 terms in the Taylor series for $(x-1)e^x$ near $x = 1$.
6. Find the first 3 terms in the Maclaurin series for (a) $\sin x^2$, (b) $\frac{x}{\sqrt{1-x^2}}$, (c) xe^{-x}, (d) $\frac{x}{1+x^2}$.
7. Using $\frac{1}{1+x} \equiv 1 - x + x^2 - x^3 + \cdots$, find the Maclaurin series for the function $\frac{1}{2+x}$.

Partial Differentiation

Introduction

In the first part of this course, you have met the idea of a derivative. To recap using first principle, recall that if you have a function, f (say), then the slope of the curve of f at a point x is said to be the number,

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

provided that this limit exists. If this limit does exist for every value of x, then the function f is said to be differentiable, or smooth.

If f is given by a power series in x, let's say

$$f(x) = a_0 + f_1 x + f_2 x^2 + \cdots$$

then f is differentiable and its derivative at $x = 0$ is equal to f_1. This is easy to see and we know from Taylor's theorem that a partial converse to this result exists. Namely, if a function is smooth and we can differentiate it several times, then we can approximate the function locally in terms of a polynomial. This idea is very important during this part of the course.

Let us introduce some terminology that is useful when dealing with derivatives.

Definition: *(The absolute value or modulus function)* The absolute value function $x \mapsto |x|$, which is read "x maps to mod x", is given by

$$|x| = \begin{cases} x; & x \geq 0 \\ x; & x \leq 0 \end{cases}$$

or $|x| = +\sqrt{x^2}$.

Definition: *(Big O notation)* We use the object ($\Delta x = h$)

$$O(h^2)$$

to mean any function of h that contains terms which are as small as h^2, and possibly smaller, when h is itself small. For instance, $2h^2 + h^{100}$ is $O(h^2)$, as is $99h^2 - h^{10}$; it's just a shorthand where we don't care too much what the higher terms are because they're very small if h is small. In fact, a function $f(h)$ is said to be $O(h)$ as $h \to 0$ precisely when the limit,

$$\lim_{h \to 0} \frac{f(h)}{h}$$

exists.

Now, *we can think of the derivative as a way of approximating the function f using a linear function*. For instance, one can also define the derivative of $f(x)$ to be the function $f'(x)$ such that the statement,

$$|f(x + h) - f(x) - hf'(x)| = O(h^2)$$

holds as $h \to 0$ for all values of x.

This simply means that if h is small, then for a fixed x, the function

$$h \mapsto f(x + h)$$

looks (locally at least) a lot like the linear function of h,

$$h \longmapsto f(x) + hf'(x)$$

To clarify these ideas let us consider a familiar example.

Example: Let us consider the function $f(x) = x^2$. Now, we know that $f'(x) = 2x$. But we also have,

$$|f(x+h) - f(x) - hf'(x)| = |(x+h)^2 - x^2 - 2xh| = h^2$$

Let us consider the function $f(x) = e^x$. Now, we know that $f'(x) = e^x$, but for a fixed x we also have,

$$|f(x+h) - f(x) - hf'(x)| = |e^{x+h} - e^x - he^x| = |e^x(e^h - 1 - h)|$$

Since

$$e^h = 1 + h + \frac{1}{2}h^2 + \frac{1}{6}h^3 + \cdots$$

we can justifiably write

$$|e^x(e^h - 1 - h)| = O(h^2) \text{ as } h \to 0 \text{ with } x \text{ fixed.}$$

So, to sum up this preamble to this part of the course, a derivative of a function $f(x)$ is computed using the familiar idea of a limiting process or first principle which provides a linear approximation to $f(x)$. In fact, one could say that the function (of h); $L(h) = f(x) + hf'(x)$ is the best linear approximation to the function $f(x)$ at x.

We could also define the quadratic function of h,

$$Q(h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2$$

and then we would find that $|f(x) - Q(x)| = O(h^3)$. For small h, something that is $O(h^3)$ is smaller in size than something that is $O(h^2)$, and *this is the mathematical way in which one writes that Q is a better approximation of f near x than L is.*

Example: Find the linear and quadratic approximations to the function $g(t) = cos(t)$ at the point, $t = 0$.

In this case,

$$g(0) = 1,\ g'(0) = 0, \text{ and } g''(0) = -1$$

From where,

$$L(h) = g(0) + hg'(0) = 1 \text{ and } Q(h) = g(0) + g'(0)h + \tfrac{1}{2}g''(0)h^2 = 1 - \tfrac{1}{2}h^2$$

Partial Derivatives

Often functions depend on more than one variable. For instance, the volume of a box is given by

$$V(x, y, z) = xyz$$

where x, y and z are the lengths of the sides of the box. Since calculus is so useful when studying problems in one variable, such as when maximizing, curve sketching, or deriving differential equations in physics, we would like to see if there is a calculus for functions which depend on two variables.

We shall write this as;

$$(x, y) \longmapsto f(x, y)$$

or just $f(x, y)$ for short. The symbol $\longmapsto$ is read “maps to” and indicates that f is a black box, with (x, y) as input, and some value $f(x, y)$ as output.

Now, a function such as

$$f(x, y) = x^2 + y^3$$

can have derivatives too. For instance, we can forget about y for a second and just think about the limit

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

Or we could try to evaluate the limit

$$\lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

We can perform both of these operations, and you should verify that the first one gives us $2x$ and the second, $3y^2$.

But the first answer is exactly what you get when you take f, hold y as a constant, and just differentiate the function, thinking of it as a function only of x. Indeed, we could write

$$\frac{dy}{dx}(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

Indeed we shall perform this process to differentiate functions of two variables, but *we shall use a slightly different notation instead to remind us of the fact that there are several variables in our function*. We actually use a ‘curly d’ and write

$$\frac{\partial f}{\partial x}(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

and also,

$$\frac{\partial f}{\partial y}(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

We do, however, read the symbol ‘∂’ as a ‘d’. A more compact notation is also used for partial derivative and the symbols

$$f_x(x, y) \text{ and } f_y(x, y)$$

are used in place of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

Example: Given $f(x, y) = xy$, find $f_x(x, y)$ and $f_y(x, y)$.

Using the definition, we find

$$f_x(x, y) = \lim_{h \to 0} \frac{(x + h)y - xy}{h} = y$$

and

$$f_y(x, y) = \lim_{k \to 0} \frac{x(y + k) - xy}{k} = x$$

Example: Find the partial derivative of the function, $f(x,y) = 3x^2y + 5xy^3$ for f_x and f_y.

$$f_x = \frac{\partial f(x,y)}{\partial x} = \frac{\partial}{\partial x}(3x^2y + 5xy^3) = 6xy + 5y^3$$

and

$$f_y = \frac{\partial f(x,y)}{\partial y} = \frac{\partial}{\partial y}(3x^2y + 5xy^3) = 3x^2 + 15xy^2$$

The more general case can be illustrated by considering a function $f(x,y,z)$ of three variables x, y and z. If y and z are held constant and only x is allowed to vary, the partial derivative of f with respect to x is denoted by $\frac{\partial f}{\partial x}$ and defined by

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z) - f(x,y,z)}{\Delta x}$$

and

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y, z) - f(x,y,z)}{\Delta y}$$

also

$$\frac{\partial f}{\partial z} = \lim_{\Delta z \to 0} \frac{f(x, y, z + \Delta z) - f(x,y,z)}{\Delta z}$$

Example: Let $f(x,y,z) = x^2yz + ye^z$, then

$$\frac{\partial f}{\partial x} = 2xyz, \quad \frac{\partial f}{\partial y} = x^2z + e^z, \text{ and } \frac{\partial f}{\partial z} = x^2y + ye^z$$

Given the fact that differentiation of a function of one variable provides a way of constructing linear and quadratic approximations to a function; does partial differentiation provide analogies of such approximations for functions of more than one variable? The answer to this question will be affirmative, but we must first examine the notation of higher-derivatives for functions of more than one variable.

Higher and Mixed Partial Derivatives

Given that we can differentiate a function $f(x,y)$ with respect to one of the variables, x or y, to form a partial derivative, can we evaluate derivatives of derivatives, that is, how do we obtain higher derivatives? We shall write the 'second partial derivative of f with respect to x' as

$$\frac{\partial^2 f}{\partial x^2}(x,y) = f_{xx}(x,y) = \lim_{\Delta x \to 0} \frac{f_x(x + \Delta x, y) - f_x(x,y)}{\Delta x}$$

and then do the same for y:

$$\frac{\partial^2 f}{\partial y^2}(x,y) = f_{yy}(x,y) = \lim_{\Delta y \to 0} \frac{f_y(x, y + \Delta y) - f_y(x,y)}{\Delta y}$$

We also have the 'mixed' second partial derivatives

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = f_{xy}(x,y) = \lim_{\Delta x \to 0} \frac{f_y(x+\Delta x, y) - f_y(x,y)}{\Delta x}$$

and

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = f_{yx}(x,y) = \lim_{\Delta y \to 0} \frac{f_x(x, y+\Delta y) - f_x(x,y)}{\Delta y}$$

You may think that the order in which we perform each subsequent differentiation when finding partial derivatives depends on the order in which each is performed. Put another way: is f_{xy} different from f_{yx}?

We shall label the answer to this question as a theorem.

Theorem: For smooth function $f(x,y)$ the mixed partials f_{xy} and f_{yx} are the same.

Notice that a natural corollary of this theorem is that the higher derivative $f_{xxyxxyxyx}$ is the same as $f_{xxxxxxyyy}$ as the order of differentiation is of no importance!

Example: Consider the function $f(x,y) = x^2 y + y e^x$, we then have the following derivatives:

$$f_x = 2xy + ye^x, \quad f_y = x^2 + e^x$$
$$f_{xx} = 2y + ye^x, \quad f_{yy} = 0, \quad f_{xy} = 2x + e^x, \quad f_{yx} = 2x + e^x$$
$$f_{xxx} = ye^x, \quad f_{xxy} = 2 + e^x, \quad f_{xyy} = 0, \quad f_{yyy} = 0$$

and so on!

Chain Rule

Suppose that we have the equation of a surface

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

and we want to obtain a calculus for a particle moving along that surface. We can then think of (x,y) as a function of time, t, and form the height function

$$z(t) = f(x(t), y(t))$$

A natural question to ask is 'what is the derivative of z with respect to time?'. This derivative measures the rate at which the height is changing as time changes. We can find this expression by thinking about the difference

$$z(t+h) - z(t)$$

which is

$$f\big(x(t+h), y(t+h)\big) - f(x(t), y(t))$$

With x and y smooth functions, we can write

$$x(t+h) = x(t) + x'(t)h + O(h^2)$$

and

$$y(t+h) = y(t) + y'(t)h + O(h^2)$$

Therefore

$$z'(t) = \lim_{h \to 0} \frac{z(t+h) - z(t)}{h} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Now, we can extend this idea by supposing that x and y are now functions of (t, s), and we now form

$$z(t,s) = f(x(t,s), y(t,s))$$

We can think of s being fixed and repeat the above process, replacing derivatives of x and y with partial derivatives. This would give us the result that

$$\frac{\partial z}{\partial t} = \lim_{h\to 0}\frac{z(t+h,s) - z(t,s)}{h} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

and, reasoning in an analogous fashion,

$$\frac{\partial z}{\partial s} = \lim_{k\to 0}\frac{z(t,s+k) - z(t,s)}{k} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

These expressions are known as the chain rule for functions of two variables. If we recall the chain rule from one-dimensional calculus, that

$$\frac{d}{dt}f\big(x(t)\big) = \frac{df}{dx}\frac{dx}{dt}$$

then we see that an extra term is required when we move to two dimensions. Explicitly,

$$\frac{d}{dt}f(x(t), y(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

And again, just for clarity

$$\frac{\partial}{dt}f(x(t,s), y(t,s)) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

Similarly, if x, y and z are all functions of a single variable t, then f can be considered as a function of t and

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}$$

Example: Let $f(x,y) = \cos(xy)$ and let $x(t) = 2t,\ y(t) = t^2$; let us find $\frac{d}{dt}f(x(t),y(t))$ using the chain rule.

Now

$$\frac{d}{dt}f\big(x(t), y(t)\big) = f_x\big(x(t), y(t)\big)x'(t) + f_y\big(x(t), y(t)\big)y'(t)$$

which equals

$$-\sin(xy)yx'(t) - \sin(xy)xy'(t) = -\sin(2t^3)(2t^2 + 4t^2) = -\sin(2t^3)6t^2$$

Also, substituting the expressions for x, y, into f, we have

$$\frac{d}{dt}\cos(2t^3) = -\sin(2t^3)6t^2$$

Example: Let $f(x,y,z) = x^2yz$; $x = e^t, y = t,$ and $z = 1 + t$.

Using the chain rule gives,

$$\frac{df}{dt} = (2xyz)e^t + x^2z + x^2y = 2te^{2t}(1+t) + e^{2t}(1+2t)$$

However, using direct substitution, we have

$$f = e^{2t}t(1+t)$$

Differentiating, gives

$$\frac{df}{dt} = e^{2t}(2t+1) + 2e^{2t}t(1+t) = e^{2t}(2t+1) + 2te^{2t}(1+t)$$

Lagrange Multipliers

Lagrange multipliers, named after Joseph Louis Lagrange, is a method for finding the extrema (local minima or local maxima) of a function subject to one or more (equality or inequality) constraints. This method reduces a problem in n variable with k constraints to a problem in $n + k$ variables with no constraint.

The method introduces a scalar variable, the Lagrange multiplier, for each constraint and forms a linear combination involving the multipliers as coefficients. The Lagrange multipliers are also called *undetermined multipliers*.

Many well-known machine learning algorithms make use of the method of Lagrange multipliers. For example, the theoretical foundations of principal components analysis (PCA) are built using the method of Lagrange multipliers with equality constraints. Similarly, the optimization problem in support vector machines SVMs is also solved using this method.

The method of Lagrange Multipliers

Suppose we have the following optimization problem:

Minimize $f(x)$

Subject to:

$$g_i(x) = 0, \quad (i = 1.2, \dots, n)$$

The method of Lagrange multipliers first constructs a function called the *Lagrange function* as given by the following expression:

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \cdots + \lambda_n g_n(x)$$

where λ represents a vector of Lagrange multipliers, i.e.,

$$\lambda = [\lambda_1, \lambda_2, \dots, \lambda_n]^T$$

Now, to find the points of local minimum of $f(x)$ subject to the equality constraints, we find the stationary points of the Lagrange function $L(x, \lambda)$, i.e., we solve the following equations:

$$\nabla L = 0 \quad \text{or} \quad \nabla[f(x) + \lambda(g(x) - c)] = 0$$

Once values for λ are determined, we can go on to find extrema of the unconstrained function,

$$F(x) = f(x) + \lambda(g(x) - c)$$

Note that the critical points of $F(x)$ are all on the curve $g(x) = c$ (As can be seen by setting $\nabla F = 0$. Therefore, the extrema of $F(x)$ are equal to the extrema of $f(x)$.

Example: Let's solve the following minimization problem:

Minimize: $f(x) = x^2 + y^2$

Subject to :

$$x + 2y - 1 = 0$$

The first step is to construct the Lagrange function:

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(x + 2y - 1)$$

Which yields the following equations:

$$\frac{\partial L}{\partial x} = 0 \quad \Longrightarrow \quad 2x + \lambda = 0 \qquad (1)$$

$$\frac{\partial L}{\partial y} = 0 \quad \Longrightarrow \quad 2y + 2\lambda = 0 \qquad (2)$$

$$\frac{\partial L}{\partial \lambda} = 0 \quad \Longrightarrow \quad x + 2y - 1 = 0 \qquad (3)$$

Solving (1) to (3) we have,

$$x = \frac{1}{5}, \quad y = \frac{2}{5}$$

Hence, the local minimum point lies at $(\frac{1}{5}, \frac{2}{5})$.

Example: Find the minimum of the following function subject to the given constraints:

Minimize: $g(x, y) = x^2 + 4y^2$

Subject to:

$$x + y = 0$$
$$x^2 + y^2 = 1$$

Constructing the Lagrange function, we have,

$$L(x, y, \lambda_1, \lambda_2) = x^2 + 4y^2 + \lambda_1(x + y) + \lambda_2(x^2 + y^2 - 1)$$

Thus, we have four equations:

$$\frac{\partial L}{\partial x} = 0 \quad \Longrightarrow \quad 2x + \lambda_1 + 2x\lambda_2 = 0 \qquad (1)$$

$$\frac{\partial L}{\partial y} = 0 \quad \Longrightarrow \quad 8y + \lambda_1 + 2y\lambda_2 = 0 \qquad (2)$$

$$\frac{\partial L}{\partial \lambda_1} = 0 \quad \Longrightarrow \quad x + y = 0 \qquad (3)$$

$$\frac{\partial L}{\partial \lambda_2} = 0 \quad \Longrightarrow \quad x^2 + y^2 - 1 = 0 \qquad (4)$$

Solving the above system of equations, gives us two solutions for (x, y), i.e., we get the two points:

$$\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \quad \text{and} \quad -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}$$

Remark: If you have a function to maximize, you can solve it in a similar manner, keeping in mind that maximization and minimization are equivalent problems, i.e.,

Maximize: $f(x)$ is equivalent to minimize: $-f(x)$

Example: Find the maximum for $f(x, y) = x^2 y$ with the condition that (x, y) lies on the circle around the origin with radius $\sqrt{3}$, i.e.,

$$x^2 + y^2 = 3$$

Constructing the Lagrange function, we have,

$$L(x, y, \lambda) = x^2 y + \lambda(x^2 + y^2 - 3)$$

Thus, we have three equations:

$$\frac{\partial L}{\partial x} = 0 \quad \Longrightarrow \quad 2xy + 2\lambda x = 0 \qquad (1)$$

$$\frac{\partial L}{\partial y} = 0 \quad \Longrightarrow \quad x^2 + 2\lambda x = 0 \qquad (2)$$

$$\frac{\partial L}{\partial \lambda} = 0 \quad \Longrightarrow \quad x^2 + y^2 - 3 = 0 \qquad (3)$$

Solving the equations, clearly, there are 4 critical points:

$$\left(\sqrt{2}, 1\right), \ \left(-\sqrt{2}, 1\right); \ \left(\sqrt{2}, -1\right), (-\sqrt{2}, -1)$$

By evaluating the Lagrangian at these points, we find

$$f\left(\sqrt{2}, 1\right) = 2, \quad f(-\sqrt{2}, 1) = 2$$

and

$$f\left(\sqrt{2}, -1\right) = -2, \quad f\left(-\sqrt{2}, -1\right) = -2$$

Therefore, the criterion function attains a maximum at $\left(\sqrt{2}, 1\right)$ and $(-\sqrt{2}, 1)$ and a minimum at the other two critical points.