Chapter 14 (14.1-14.8)

Partial Derivatives

Note: *This module is prepared from Chapter 14 (Thomas' Calculus, 13th edition) just to help the students. The study material is expected to be useful but not exhaustive. For detailed study, the students are advised to attend the lecture/tutorial classes regularly, and consult the text book.*

Appeal: Please do not print this e-module unless it is really necessary.

Dr. Suresh Kumar, Department of Mathematics, BITS Pilani, Pilani Campus

Contents

SECTION 14.1 (Functions of Several Variables)

Functions of one real variable

Let A and B be subsets of $\mathbb{R}$, the set of real numbers. Then a rule f that assigns to each member $x \in A$ a unique member $y \in B$, is called a function from A to B, written as $f : A \to B$. The element y is called image of x under f. It is also called as the value of f at x and is denoted by $f(x)$, that is, $y = f(x)$ while x is called pre-image of y under f. The set A (the set of all pre-images) is called domain of f (denoted by $D(f)$) while the set of images of all the elements of A under f, that is, $R(f) = \{f(x) : x \in A\}$ is called range of f. We call the set B as co-domain of f. The set $G(f) = \{(x, f(x)) : \ x \in A\}$ is called graph of f. Notice that the graph of f contains ordered pairs $(x, f(x))$ of real numbers, which can be shown geometrically as points $(x, f(x))$ in the XY-plane.

Ex. Consider the function $f : [-1, 1] \to \mathbb{R}$ defined by $f(x) = \sqrt{1 - x^2}$. Then, we have $D(f) = [-1, 1]$, $R(f) = \{\sqrt{1 - x^2} : x \in [-1, 1]\} = [0, 1]$ and $G(f) = \{(x, \sqrt{1 - x^2}) : x \in [-1, 1]\}$, which is displayed geometrically in XY-plane in the left panel of Figure 1. So domain of f in the interval $[-1, 1]$ on the X-axis, and geometrically its graph is the semicircle.

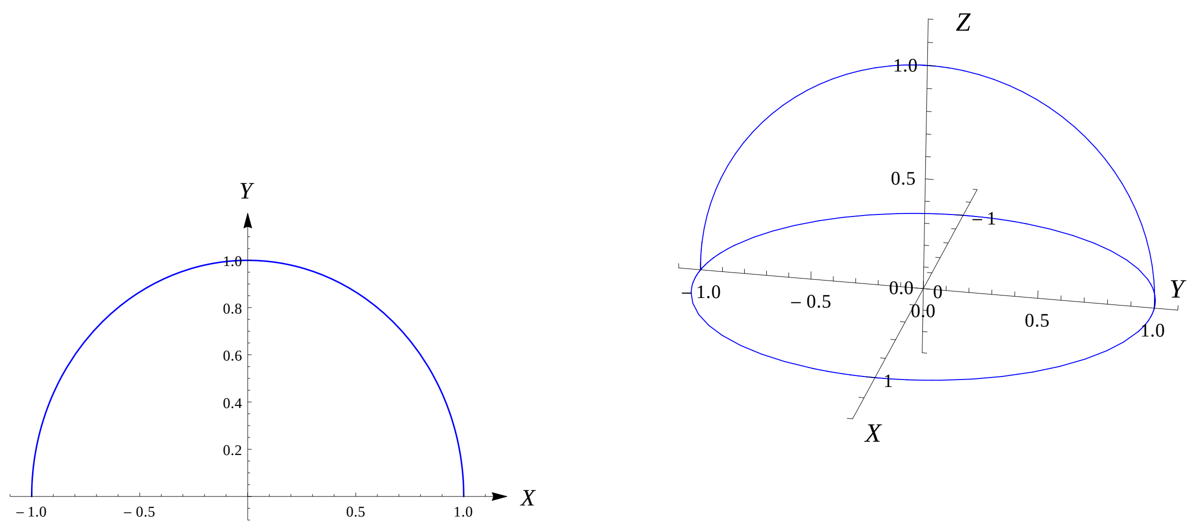


Figure 1: Left panel: $f(x) = \sqrt{1 - x^2}$. Right panel: $f(x, y) = \sqrt{1 - x^2 - y^2}$

Functions of two real variables

Let A, B and C be subsets of $\mathbb{R}$, the set of real numbers. Then a rule f that assigns to each member $(x, y) \in A \times B$ a unique member $z \in C$, is called a function from $A \times B$ to C, written as $f : A \times B \to C$.

The element z is called image of (x, y) under f. It is also called as the value of f at (x, y) and is denoted by $f(x, y)$, that is, $z = f(x, y)$ while (x, y) is called pre-image of z under f. The set $A \times B$ (the set of all pre-images) is called domain of f (denoted by $D(f)$) while the set of images of all the elements of $A \times B$ under f, that is, $R(f) = \{f(x, y) : (x, y) \in A \times B\}$ is called range of f. We call the set C as co-domain of f. The set $G(f) = \{(x, y, f(x, y)) : (x, y) \in A \times B\}$ is called graph of f. Notice that the graph of f contains the triplets $(x, y, f(x, y))$ of real numbers, which can be shown geometrically as points $(x, y, f(x, y))$ in the three dimensional XYZ-space.

Ex. Consider the function $f : \{(x, y) : x^2 + y^2 \leq 1\} \to \mathbb{R}$ defined by $f(x, y) = \sqrt{1 - x^2 - y^2}$. Then, we have $D(f) = \{(x, y) : x^2 + y^2 \leq 1\}$, $R(f) = \{\sqrt{1 - x^2 - y^2} : (x, y) \in D(f)\} = [0, 1]$ and $G(f) = \{(x, y, \sqrt{1 - x^2 - y^2}) : (x, y) \in D(f)\}$, which is displayed geometrically in the XYZ-space in the right panel of Figure 2. So domain of f is the region within and on the circle $x^2 + y^2 = 1$ in the XY-plane, and geometrically its graph is the hemispherical surface.

Remark: Likewise, we can define a function of three variables. However, functions of more than two variables can not be represented geometrically, and therefore are not of much interest. Consider the function $f : \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\} \to \mathbb{R}$ defined by $f(x) = \sqrt{1 - x^2 - y^2 - z^2}$. Then, we have $D(f) = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$, $R(f) = \{\sqrt{1 - x^2 - y^2 - z^2} : (x, y, z) \in D(f)\} = [0, 1]$ and $G(f) = \{(x, y, z, \sqrt{1 - x^2 - y^2 - z^2}) : (x, y, z) \in D(f)\}$. So domain of f is the region within and on the sphere $x^2 + y^2 + z^2 = 1$ in the XYZ-space, and of course its graph $G(f)$ can not be represented geometrically.

Table 1 shows some two and three variable functions with their respective domains and ranges.

Table 1: Some functions with domains and ranges.

Function	Domain	Range
$z = \sqrt{y - x^2}$	$y \geq x^2$	$[0, \infty)$
$z = \frac{1}{xy}$	$xy \neq 0$	$(-\infty, 0) \cup (0, \infty)$
$z = \sin xy$	Entire plane	$[-1, 1]$
$w = \sqrt{x^2 + y^2 + z^2}$	Entire space	$[0, \infty)$
$w = \frac{1}{x^2+y^2+z^2}$	$(x, y, z) \neq (0, 0, 0)$	$(0, \infty)$
$w = xy \ln z$	Half-space $z > 0$	$(-\infty, \infty)$

Neighbourhoods

Roughly speaking, neighbourhood of a point is any open region (the region carrying only its interior points and hence without the boundary points) containing the point.

Formally, neighbourhood of a point x_0 on X-axis is any open interval containing the point x_0. For any positive real number δ, the interval $(x_0 - \delta, x_0 + \delta)$ is a neighbourhood of x_0. We shall denote it by $N_\delta(x_0)$ and call it δ-neighbourhood of x_0. In fact, the interval $(x_0 - \delta, x_0 + \delta)$ is neighbourhood of its every point. A neighbourhood of x_0 without x_0 is called deleted neighbourhood of x_0.

Neighbourhood of a point (x_0, y_0) in the XY-plane is any open area or region containing the point (x_0, y_0). For any positive real number δ, the area $\{(x, y) : |x - x_0| < \delta,\ |y - y_0| < \delta\}$ is a neighbourhood of (x_0, y_0) and is a square neighbourhood of (x_0, y_0). Similarly, $\{(x, y) : \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta\}$ is circular neighbourhood of (x_0, y_0). We shall denote any δ-neighbourhood of (x_0, y_0) by $N_\delta(x_0, y_0)$.

Likewise, neighbourhood of a point (x_0, y_0, z_0) in the XYZ-space is any open region containing the point (x_0, y_0, z_0). For instance, $N_\delta(x_0, y_0, z_0) = \{(x, y, z) : |x-x_0| < \delta,\ |y-y_0| < \delta,\ |z-z_0| < \delta\}$ is cubical δ-neighbourhood of (x_0, y_0, z_0) while $N_\delta(x_0, y_0, z_0) = \{(x, y, z) : \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} < \delta\}$ is spherical δ-neighbourhood of (x_0, y_0, z_0).

Interior and boundary points in XY-plane

A point (x_0, y_0) in a region (set) R in XY-plane is an interior point of R if there exists a nbd of (x_0, y_0) entirely lying inside R as shown in the left panel of Figure 2. If every nbd of (x_0, y_0) contains points inside as well as outside of R, then it is called a boundary point of R (see right panel of Figure 2). The interior points of a region, as a set, make up the interior of the region. The region's boundary points make up its boundary. A region is open if it consists entirely of interior points. A region is closed if it contains all its boundary points.

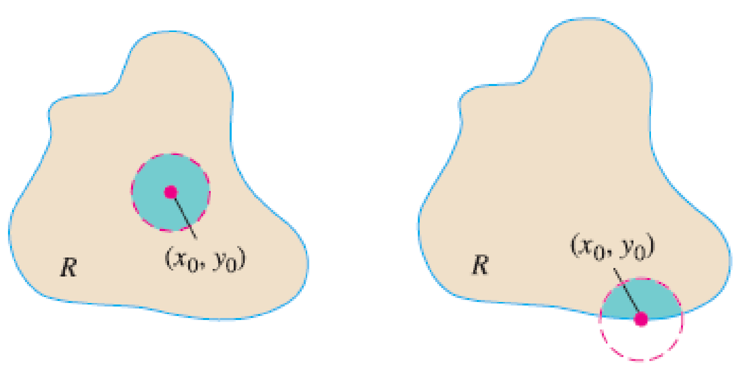
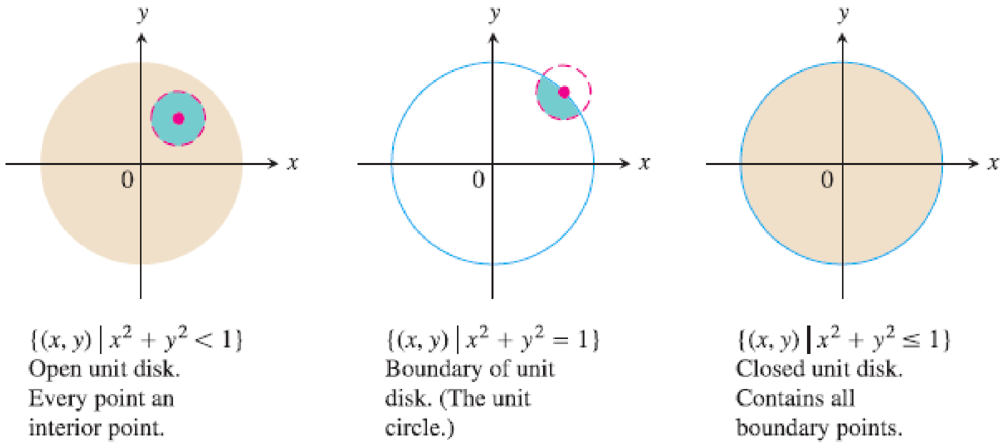


Figure 2: **Left:** Interior point (x_0, y_0) of R. **Right:** Boundary point (x_0, y_0) of R.



A region in the plane is bounded if it lies inside a disk of finite radius. A region is unbounded if it is not bounded. For example, domain of the function $f(x, y) = \sqrt{y - x^2}$ is given by $y \geq x^2$, which is a closed and unbounded region as shown in Figure 3.

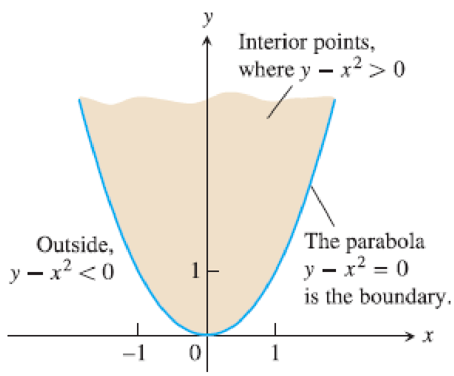


Figure 3:

Level Curves, and Contours of Functions of Two Variables

The set of points in the domain of the function $z = f(x, y)$ where it has a constant value $f(x, y) = c$ is called a level curve of f. A contour curve is a curve $f(x, y) = c$ in space in which the plane $z = c$ cuts the surface $z = f(x, y)$. For example, consider the function $f(x, y) = 100 - x^2 - y^2$. The level curves $f(x, y) = 0$ $f(x, y) = 51$, and $f(x, y) = 75$ in the domain of f are shown in left panel of Figure 4.

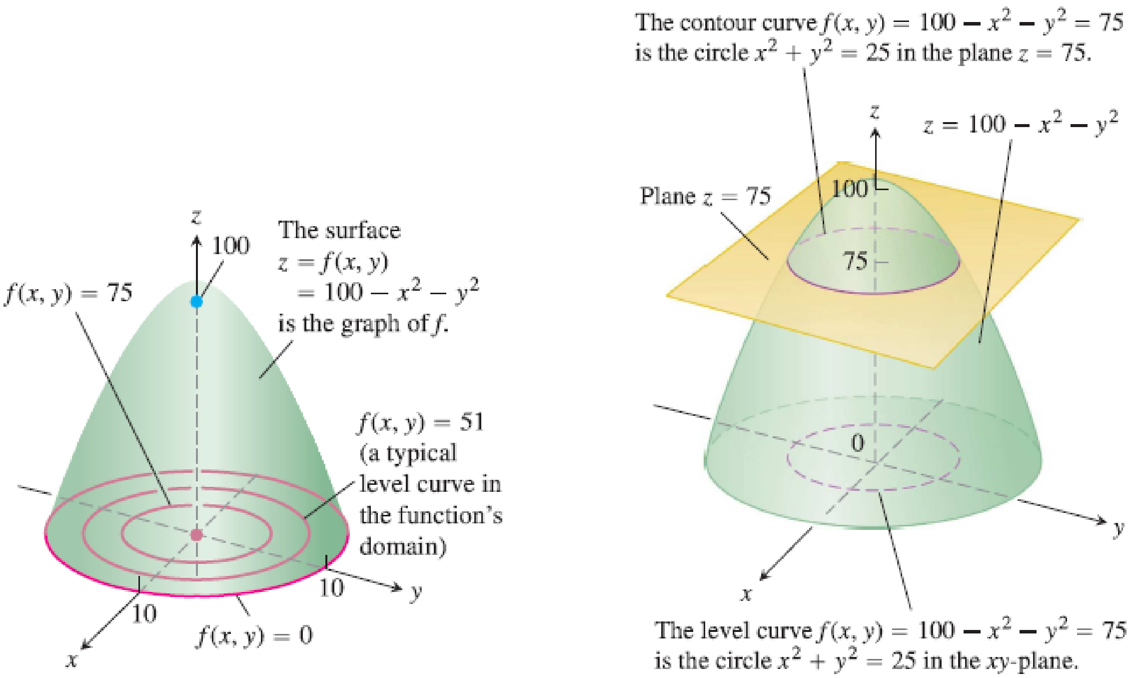


Figure 4:

Interior and boundary points in XYZ-space

A point (x_0, y_0, z_0) in a region (set) R in XYZ-space is an interior point of R if there exists a nbd of (x_0, y_0, z_0) entirely lying inside R as shown in the left panel of Figure 5. If every nbd of (x_0, y_0) contains points inside as well as outside of R, then it is called a boundary point of R (see right panel of Figure 5). The interior points of a region, as a set, make up the interior of the region. The region's boundary points make up its boundary. A region is open if it consists entirely of interior points. A region is closed if it contains all its boundary points.

The set of points (x, y, z) in space where a function of three independent variables has a constant value $f(x, y, z) = c$ is called a level surface of f. For example, the level surfaces of the function $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ are concentric spheres as shown in Figure 6.

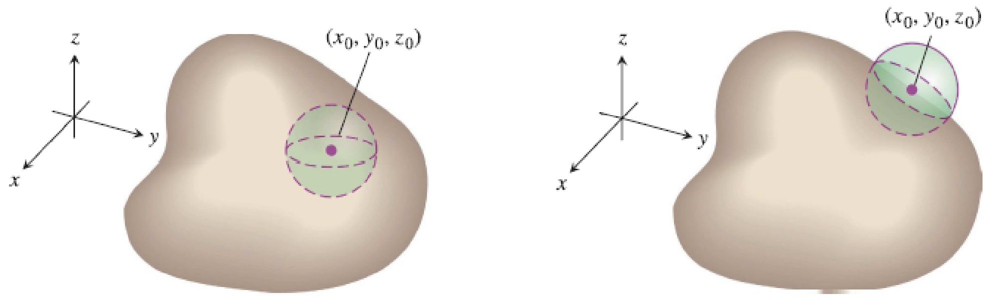


Figure 5: **Left:** Interior point (x_0, y_0, z_0) of R. **Right:** Boundary point (x_0, y_0, z_0) of R.

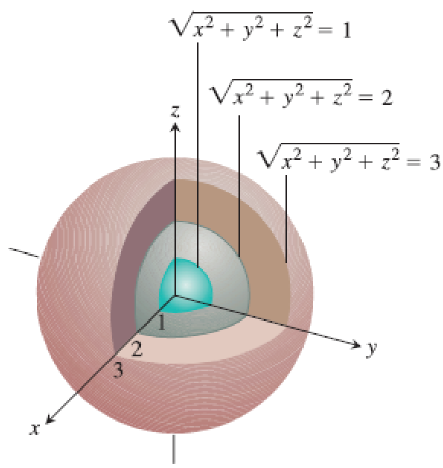


Figure 6: The level surfaces of $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ are concentric spheres.

SECTION 14.2 (Limits and Continuity in Higher Dimensions)

Limit of a function of one real variable

Meaning of $x \to x_0$

If a real variable $x \neq x_0$ takes infinitely many values in each neighbourhood of the point x_0, then we say that x approaches or tends to x_0, and we write $x \to x_0$. It implies that each δ-neighbourhood $(x_0 - \delta, x_0 - \delta)$ of x_0 contains infinitely many values of the variable x. We may choose positive values of δ very close to 0. Still the δ-neighbourhood will carry infinitely many values of the variable x. Thus, x takes values arbitrarily close to x_0, and it makes sense to say that x approaches x_0 or $x \to x_0$. If x approaches x_0 by taking values less than x_0, then x_0 is called as the left hand limit of x, and we write $x \to x_0^-$. If x approaches x_0 by taking values greater than x_0, then x_0 is called as the right hand limit of x, and we write $x \to x_0^+$.

Ex. Suppose x takes values 0.9, 0.99, 0.999, ... Then $x \neq 1$ and takes infinitely many values in each neighbourhood of 1. So x approaches 1. Also, all the values of x are less than 1. So $x \to 1^-$.

Ex. Suppose x takes values 1.1, 1.01, 1.001, ... Then $x \neq 1$ and takes infinitely many values in each neighbourhood of 1. So x approaches 1. Also, all the values of x are greater than 1. So $x \to 1^+$.

Ex. Suppose x takes values 1, 1/2, 1/3, ... Then $x \neq 0$ and takes infinitely many values in each neighbourhood of 0. So x approaches 0. Also, all the values of x are greater than 0. So $x \to 0^+$.

Limit of $f(x)$ as $x \to x_0$

Let f be a function defined in some neighbourhood of x_0 except possibly at x_0. Then a real number l is said to be limit of $f(x)$, symbolically written as $\lim_{x \to x_0} f(x) = l$, if given any positive real number ϵ (however small), there exists $\delta > 0$ (depending on ϵ) such that

$0 < |x - x_0| < \delta \quad \Longrightarrow \quad |f(x) - l| < \epsilon.$

or $x \in (x_0 - \delta, x_0 + \delta) - \{x_0\} \quad \Longrightarrow \quad f(x) \in (l - \epsilon, l + \epsilon).$

Thus, $\lim_{x \to x_0} f(x) = l$ if corresponding to each ϵ-neighbourhood of l, there exists a deleted δ-neighbourhood of x_0 such that the values of $f(x)$ corresponding to the deleted δ-neighbourhood of x_0 lie in the ϵ-neighbourhood of l. Geometrically, it implies that the portion of the curve $f(x)$ corresponding to the deleted δ-neighbourhood of x_0 lies inside the horizontal strip created by the lines $f(x) = l - \epsilon$ and $f(x) = l + \epsilon$ parallel to X-axis.

Further, $\lim_{x \to x_0^-} f(x) = l$ (left hand limit of $f(x)$) if $x \in (x_0 - \delta, x_0) \quad \Longrightarrow \quad f(x) \in (l - \epsilon, l + \epsilon)$, and $\lim_{x \to x_0^+} f(x) = l$ (right hand limit of $f(x)$) if $x \in (x_0, x_0 + \delta) \quad \Longrightarrow \quad f(x) \in (l - \epsilon, l + \epsilon)$.

Obviously, $\lim_{x \to x_0} f(x) = l$ if and only if $\lim_{x \to x_0^-} f(x) = l = \lim_{x \to x_0^+} f(x)$.

Ex. Show that $\lim_{x \to 1} (2x + 1) = 3$.

Sol. Let $\epsilon > 0$ be given. Then

$|2x + 1 - 3| = |2x - 2| = 2|x - 1| < \epsilon$ provided $|x - 1| < \epsilon/2$.

Choosing $\delta = \epsilon/2$, we have

$|x - 1| < \delta \quad \Longrightarrow \quad |2x + 1 - 3| < \epsilon.$

Thus, $\lim_{x \to 1} (2x + 1) = 3$.

For the sake of illustration, we plot $f(x) = 2x + 1$ in left panel of Figure 7. We choose $\epsilon = 0.4$ so that $\delta = \epsilon/2 = 0.2$. The horizontal dotted blue lines are $f(x) = l - \epsilon = 3 - 0.4 = 2.6$ and $f(x) = l + \epsilon = 3 + 0.4 = 3.4$ while the vertical red lines are $x = x_0 - \delta = 1 - 0.2 = 0.8$ and $x = x_0 + \delta = 1 + 0.2 = 1.2$. We can see that the part of the curve in the vertical red strip completely lies inside the horizontal blue strip, as expected.

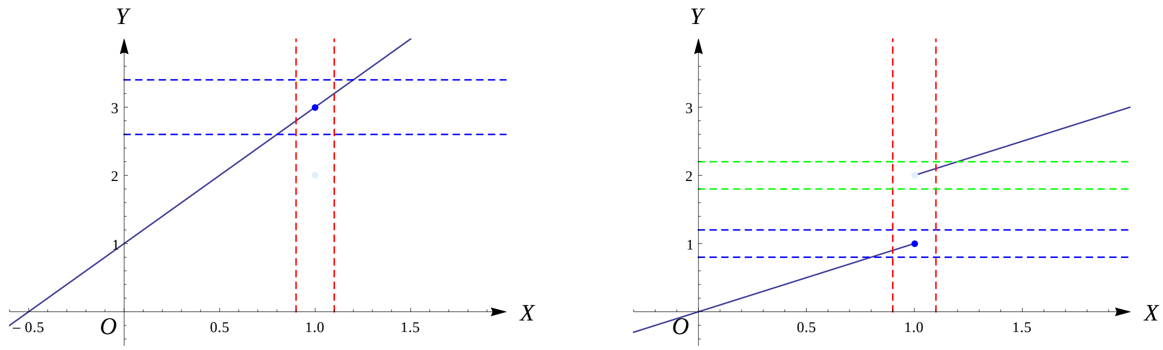


Figure 7: Left panel: $f(x) = 2x + 1$. Right panel: $f(x) = x$ for $x \leq 1$ and $f(x) = x + 1$ for $x > 1$.

Ex. Show that $\lim_{x\to 1} f(x)$, where $f(x) = \begin{cases} x & \text{if } x \leq 1 \\ x+1 & \text{if } x > 1 \end{cases}$, does not exist.

Sol. We shall show that $\lim_{x\to 1^-} f(x) = 1$ and $\lim_{x\to 1^+} f(x) = 2$.

Let $\epsilon > 0$ be given. Then for $x < 1$, we have $|f(x) - 1| = |x - 1| < \epsilon$ provided $x \in (1 - \epsilon, 1)$.

Choosing $\delta = \epsilon$, we have

$x \in (1 - \delta, 1) \quad \Longrightarrow \quad |f(x) - 1| < \epsilon.$

Thus, $\lim_{x\to 1^-} f(x) = 1$.

Next for $x > 1$, we have

$|f(x) - 2| = |x + 1 - 2| = |x - 1| < \epsilon$ provided $x \in (1, 1 + \epsilon)$.

Choosing $\delta = \epsilon$, we have

$x \in (1, 1 + \delta) \quad \Longrightarrow \quad |f(x) - 2| < \epsilon.$

Thus, $\lim_{x\to 1^+} f(x) = 2$.

Since $\lim_{x\to 1^-} f(x) \neq \lim_{x\to 1^+} f(x)$, so $\lim_{x\to 1} f(x)$ does not exist.

For the sake of illustration, we plot $f(x) = \begin{cases} x & \text{if } x \leq 1 \\ x+1 & \text{if } x > 1 \end{cases}$ in right panel of Figure 7. We choose $\epsilon = 0.2$ so that $\delta = \epsilon = 0.2$. The horizontal dotted blue lines are $f(x) = l - \epsilon = 1 - 0.2 = 0.8$ and $f(x) = l + \epsilon = 1 + 0.2 = 1.2$ while the horizontal dotted blue lines are $f(x) = l - \epsilon = 2 - 0.2 = 1.8$ and $f(x) = 2 + \epsilon = 2 + 0.2 = 2.2$. The vertical red lines are $x = x_0 - \delta = 1 - 0.2 = 0.8$ and $x = x_0 + \delta = 1 + 0.2 = 1.2$. We can see that the part of the curve corresponding to $x \in (0.8, 1)$ in the vertical Red strip completely lies inside the horizontal blue strip while the part of the curve corresponding to $x \in (1, 1.2)$ in the vertical Red strip completely lies inside the horizontal Green strip, as expected.

Ex. Discuss the geometry of $\lim_{x\to 0} \sqrt{1 - x^2} = 1$.

Sol. We know that $\sqrt{1 - x^2}$ is semicircular curve with the domain $[-1, 1]$ on X-axis (See left panel of Figure 1). It is easy to see that when $x \to 0$ either from left or right in the neighbourhood of 0, the corresponding part of the curve converges to the point $(0,1)$ on Y-axis. It implies that $\sqrt{1-x^2}$ tends to 1 for both the paths along which $x \to 0$. That is why, $\lim_{x\to 0}\sqrt{1-x^2} = 1$.

Remark: If $\lim_{x\to x_0} f(x)$ exists, then the graph of the function $f(x)$, geometrically, converges or strikes at the same point (whose y-coordinate is the limit of $f(x)$) from left as well as right in the neighbourhood of x_0. For instance, see left panel of Figure 2. If $\lim_{x\to x_0} f(x)$ does not exist but $\lim_{x\to x_0^-} f(x)$ and $\lim_{x\to x_0^+} f(x)$ exist finitely, then the graph of the function $f(x)$, geometrically, from the left strikes at a point different from the point where it strikes from the right in the neighbourhood of x_0. For an example, see right panel of Figure 2.

Remark: The limit of a function at a point may exist even if the function is not defined there at. For example, consider the function $f(x) = 2x+1$, $x \neq 1$, which is not defined at $x = 1$ as per its given definition. But $\lim_{x\to 1}(2x+1) = 3$.

For, let $\epsilon > 0$ be given. Then

$|2x+1-3| = |2x-2| = 2|x-1| < \epsilon$ provided $0 < |x-1| < \epsilon/2$.

Choosing $\delta = \epsilon/2$, we have

$0 < |x-1| < \delta \quad \Longrightarrow \quad |2x+1-3| < \epsilon$.

Thus, $\lim_{x\to 1}(2x+1) = 3$.

Limit of a function of two real variables

Meaning of $(x,y) \to (x_0,y_0)$

If the ordered pair $(x,y) \neq (x_0,y_0)$ of real variables x and y takes infinitely many values in each neighbourhood of the point (x_0,y_0), then we say that (x,y) tends to (x_0,y_0), and we write $(x,y) \to (x_0,y_0)$. It implies that each δ-neighbourhood $N_\delta(x_0,y_0)$ (square, circular or whatever sahpe) of (x_0,y_0) contains infinitely many values of (x,y). We may choose positive values of δ very close to 0. Still the δ-neighbourhood will carry infinitely many values of (x,y). Thus, (x,y) takes values arbitrarily close to (x_0,y_0), and it makes sense to say that (x,y) approaches (x_0,y_0) or $(x,y) \to (x_0,y_0)$.

Now, see the critical difference between $x \to x_0$ and $(x,y) \to (x_0,y_0)$. In the case, $x \to x_0$, the variable x can take values only on the X-axis (one dimension), and thus can approach x_0 from left or right directions along X-axis. That is why, we talk about left and right hand limits when $x \to x_0$. On the other hand, in the case $(x,y) \to (x_0,y_0)$, the ordered pair (x,y) can take values in the XY-plane (two dimensions), and thus can approach (x_0,y_0) along infinitely many paths in the XY-plane.

Ex. Suppose (x,y) takes values $(0.9,0.9)$, $(0.99,0.99)$, $(0.999,0.999)$, ... Then $(x,y) \neq (1,1)$ and takes infinitely many values in each neighbourhood of $(1,1)$. So (x,y) approaches $(1,1)$. Also, all the values of

(x, y) lie on the line $y = x$. So $(x, y) \to (1, 1)$ by taking values on the straight line path $y = x$.

Ex. Suppose (x, y) takes values $(1.1, (1.1)^2)$, $(1.01, (1.01)^2)$, $(1.001, (1.001)^3)$, ... Then $(x, y) \neq (1, 1)$ and takes infinitely many values in each neighbourhood of $(1, 1)$. So (x, y) approaches $(1, 1)$. Also, all the values of (x, y) lie on the parabola $y = x^2$. So $(x, y) \to (1, 1)$ by taking values on the parabolic path $y = x^2$.

Limit of $f(x, y)$ as $(x, y) \to (x_0, y_0)$

Let f be a function defined in some neighbourhood of (x_0, y_0) except possibly at (x_0, y_0). Then a real number l is said to be limit of $f(x, y)$, symbolically written as $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y) = l$, if given any positive real number ϵ (however small), there exists $\delta > 0$ (depending on ϵ) such that

$0 < |x - x_0| < \delta, 0 < |y - y_0| < \delta \quad \Longrightarrow \quad |f(x) - l| < \epsilon.$

or $x \in (x_0 - \delta, x_0 + \delta) - \{x_0\}, y \in (y_0 - \delta, y_0 + \delta) - \{y_0\} \quad \Longrightarrow \quad f(x) \in (l - \epsilon, l + \epsilon).$

Thus, $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y) = l$ if corresponding to each ϵ-neighbourhood of l, there exists a deleted δ-neighbourhood of (x_0, y_0) such that the values of $f(x, y)$ corresponding to the deleted δ-neighbourhood of (x_0, y_0) lie in the ϵ-neighbourhood of l. Geometrically, it implies that the portion of the surface $f(x, y)$ corresponding to the deleted δ-neighbourhood of (x_0, y_0) lies inside the sandwiched space between the planes $f(x) = l - \epsilon$ and $f(x) = l + \epsilon$ parallel to XY-plane.

Further, if $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y)$ exists and is equal to l (say), then $f(x, y) \to l$ for every possible path along which (x, y) approaches (x_0, y_0), that is, the limit is independent of path. In case, $f(x, y)$ approaches to two different values for two different paths along which (x, y) approaches (x_0, y_0), then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y)$ does not exist.

Ex. Show that $\lim\limits_{(x,y)\to(1,1)} (2x + 2y + 1) = 5$.

Sol. Let $\epsilon > 0$ be given. Then

$|2x + 2y + 1 - 5| = |2x + 2y - 4| = |2(x - 1) + 2(y - 1)| \leq 2|x - 1| + 2|y - 1| < \epsilon$ provided $|x - 1| < \epsilon/2$, $|y - 1| < \epsilon/2$.

Choosing $\delta = \epsilon/2$, we have

$|x - 1| < \delta, |y - 1| < \delta \quad \Longrightarrow \quad |2x + 2y + 1 - 5| < \epsilon.$

Thus, $\lim\limits_{x\to 1} (2x + 2y + 1) = 5$.

Geometry: For the sake of illustration, we plot the plane $f(x) = 2x + 2y + 1$ (Red color plane) in the left panel of Figure 8 in the domain $\{(x, y) : 0 \leq x \leq 2, 0 \leq y \leq 3\}$. Let us choose $\epsilon = 0.4$ so that $\delta = \epsilon/2 = 0.2$. Then $f(x) = l - \epsilon = 5 - 0.4 = 4.6$ (Blue color plane) and $f(x) = l + \epsilon = 5 + 0.4 = 5.4$ (Green color plane) are planes parallel to the XY-plane. Next, the square-neighbourhood of $(1, 1)$ in the XY-plane is $\{(x, y) : |x - 1| < 0.2, |y - 1| < 0.2\}$, which is not shown in Figure 8. We can imagine that the part of the Red plane corresponding to the square-neighbourhood $\{(x, y) : |x - 1| < 0.2, |y - 1| < 0.2\}$ of $(1, 1)$ lies inside the sandwiched space between the Green ($f(x) = 4.6$) and Blue ($f(x) = 5.4$) colored planes, as expected.

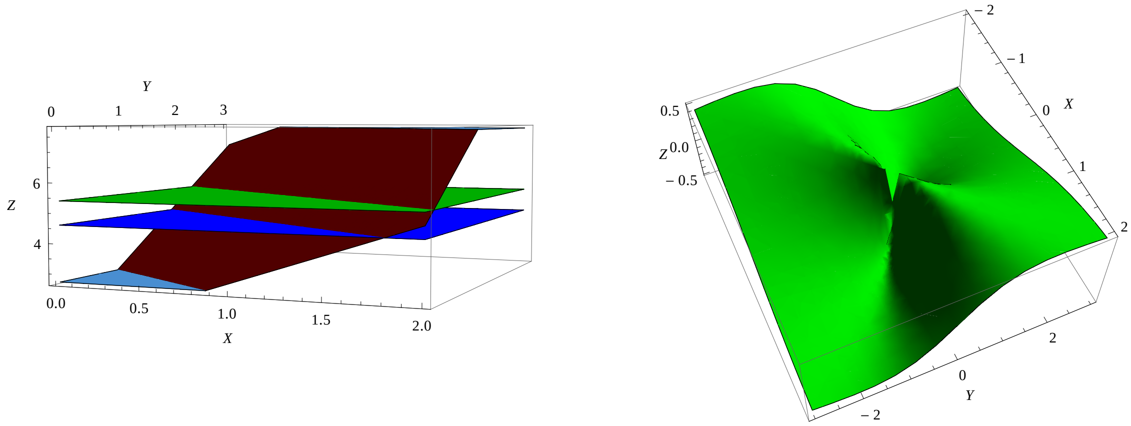


Figure 8: Left panel: $f(x) = 2x + 2y + 1$ (Red plane). Right panel: $f(x, y) = \frac{xy}{x^2+y^2}$, $(x, y) \neq (0, 0)$.

Ex. Show that $\lim\limits_{(x,y)\to(0,0)} f(x, y)$, where $f(x, y) = \frac{xy}{x^2+y^2}$, $(x, y) \neq (0, 0)$ does not exist.

Sol. We shall show that the given limit is path dependent.

Suppose $(x, y) \to (0, 0)$ along the straight line path $y = mx$. Then

$$\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{(x,y)\to(0,0)} \frac{xy}{x^2 + y^2} = \lim_{x\to 0} \frac{x.mx}{x^2 + m^2x^2} = \frac{m}{1 + m^2}.$$

So the given limit depends on m, the slope of the path line. Hence, it get different values for different values of m, and consequently it does not exist.

Geometry: The right panel of Figure 8 shows the the surface $f(x, y) = \frac{xy}{x^2+y^2}$, $(x, y) \neq (0, 0)$ in Green color. We can see that the part of the surface corresponding to a neighbourhood of the point $(0, 0)$ is not smooth rather it looks ruptured. Any path in the neighbourhood of $(0, 0)$ in the XY-plane corresponds to a curve on the Green surface. When $(x, y) \to (0, 0)$ along different paths $y = mx$, the corresponding curves end up or converge to points near the ruptured part at different heights, which is evident from the figure. That is why, the limit get different values along different paths, and consequently it does not exist.

Remark: Note that the two path approach or path dependent test is used when the limit does not exist uniquely. In case, the limit exists uniquely, you will get same answer for the limit along every chosen path. However, this way you can not ensure the existence of limit because you can not exhaust infinitely many paths. So in order to prove the existence of limit, use $\epsilon - \delta$ approach. Also, in some problems, changing the variables to polar coordinate system is quite useful for applying $\epsilon - \delta$ approach.

Ex. Show that $\lim_{(x,y)\to(0,0)} \frac{4x^2y}{x^2+y^2} = 0$.

Sol. Let $\epsilon > 0$ be given. Then, we have

$\left|\frac{4x^2y}{x^2+y^2} - 0\right| = \left|\frac{4x^2y}{x^2+y^2}\right|$.

Using $x = r\cos\theta$ and $y = r\sin\theta$, we get

$\left|\frac{4x^2y}{x^2+y^2} - 0\right| = |r\cos^2\theta\sin\theta| \leq r = \sqrt{x^2+y^2} < \epsilon$ for $\sqrt{x^2+y^2} < \epsilon$.

Choosing $\epsilon = \delta$, we have

$\left|\frac{4x^2y}{x^2+y^2} - 0\right| < \epsilon$ for $\sqrt{x^2+y^2} < \delta$.

Hence, $\lim_{(x,y)\to(0,0)} \frac{4x^2y}{x^2+y^2} = 0$.

Ex. Discuss the geometry of $\lim_{(x,y)\to(0,0)} \sqrt{1-x^2-y^2} = 1$.

Sol. We know that $\sqrt{1-x^2-y^2}$ is hemispherical surface with the circular domain $x^2+y^2 \leq 1$ centred at $(0,0)$ in the XY-plane (See right panel of Figure 1). Any path in the neighbourhood of $(0,0)$ inside the circular domain in the XY-plane corresponds to a curve on the hemispherical surface. It is easy to imagine/visualize that when $(x,y) \to (0,0)$ along any path in the neighbourhood of $(0,0)$, the corresponding curve on the hemispherical surface converges to the point $(0,0,1)$ on Z-axis. It implies that $\sqrt{1-x^2-y^2}$ tends to 1 for every path along which $(x,y) \to (0,0)$. That is why, $\lim_{(x,y)\to(0,0)} \sqrt{1-x^2-y^2} = 1$.

Remark: If $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ exists, then the graph (surface) of the function $f(x,y)$, geometrically, converges or strikes at the same point (whose z-coordinate is the limit of $f(x,y)$) from all directions in the neighbourhood of (x_0,y_0).

Remark: The limit of a function at a point may exist even if the function is not defined there at. For example, consider the function $f(x,y) = \sqrt{1-x^2-y^2}$, $(x,y) \neq (0,0)$, which is not defined at $x = (0,0)$ as per its given definition. Its graph is the hemispherical surface punctured at the point $(0,0,1)$ with the circular domain $x^2+y^2 \leq 1$ centred at $(0,0)$ in the XY-plane. We find $\lim_{(x,y)\to(0,0)} \sqrt{1-x^2-y^2} = 1$ since when $(x,y) \to (0,0)$ along any path in the neighbourhood of $(0,0)$, the corresponding curve on the hemispherical surface converges to the point $(0,0,1)$ on Z-axis.

Continuity of a function of one real variable

Let f be a function defined in some neighbourhood of x_0. Then f is said to be continuous at x_0 if $\lim_{x\to x_0} f(x) = f(x_0)$. Further, f is left continuous at x_0 if $\lim_{x\to x_0^-} f(x) = f(x_0)$, and right continuous if $\lim_{x\to x_0^+} f(x) = f(x_0)$. We say that f is continuous on an interval if it is continuous at each point of the interval.

Geometry: The existence of limit at x_0 ensures that the graph of $f(x)$ strikes at the same point from both sides. Also, the limit is equal to $f(x_0)$. So the point of strike is $(x_0, f(x_0))$. Therefore, continuity of

f at x_0 implies that there exists at least one neighbourhood of x_0 in which the $f(x)$ has continuous graph, that is, without any break point. Consequently, continuity over an interval implies that the function has continuous graph in the interval.

Ex. The function $f(x) = \sqrt{1-x^2}$ is continuous at $x = 0$ since $\lim\limits_{x\to 0} f(x) = 1 = f(0)$. In fact, $f(x) = \sqrt{1-x^2}$ is continuous at every point in its domain interval $[-1,1]$. That is why, graph of $f(x) = \sqrt{1-x^2}$ is the continuous semicircular curve from $(-1,0)$ to $(1,0)$ as shown in left panel of the Figure 1.

Ex. The function $f(x) = \begin{cases} x & \text{if } x \leq 1 \\ x+1 & \text{if } x > 1 \end{cases}$ is not continuous at $x = 1$ since $\lim\limits_{x\to 1} f(x)$ does not exist. The graph of $f(x)$ breaks at the point $x = 1$ as may be seen in the right panel of Figure 7. Notice that the function is continuous at every real number except $x = 1$.

Continuity of a function of two real variables

Let f be a function defined in some neighbourhood of (x_0, y_0). Then f is said to be continuous at (x_0, y_0) if $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0, y_0)$. We say that f is continuous in an open region in the XY-plane if it is continuous at each point of the region.

Geometry: The existence of limit at (x_0, y_0) ensures that the graph (surface) of $f(x,y)$ points towards the same point from all directions in each neighbourhood of (x_0, y_0). Also, the limit is equal to $f(x_0, y_0)$. So the point of convergence of the surface is $(x_0, y_0, f(x_0, y_0))$. Therefore, continuity of f at (x_0, y_0) implies that there exists at least one neighbourhood of (x_0, y_0) in which the $f(x,y)$ has continuous surface, that is, without any break point or point hole. Consequently, continuity over a region implies that the function has continuous surface in the region.

Ex. The function $f(x,y) = \sqrt{1-x^2-y^2}$ is continuous at $(0,0)$ since $\lim\limits_{x\to 0} f(x,y) = 1 = f(0,0)$. In fact, $f(x,y) = \sqrt{1-x^2-y^2}$ is continuous at every point in its domain circular region $x^2+y^2 \leq 1$. That is why, graph of $f(x) = \sqrt{1-x^2-y^2}$ is the continuous hemispherical surface as shown in right panel of the Figure 1.

Ex. The function $f(x,y) = \frac{xy}{x^2+y^2}$, $(x,y) \neq (0,0)$, is not continuous at $(0,0)$ since the function is not defined at $(0,0)$. Even if the function is defined to have some value at $(0,0)$, it can not be continuous at (0,0) since $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist. The graph surface of $f(x,y)$ breaks at the point $(0,0)$ as may be seen in the right panel of Figure 8. Notice that the function is continuous everywhere in the XY-plane except $(0,0)$.

SECTION 14.3 (Partial Derivatives)

Differentiability of a function of one real variable

Let f be a function defined in some neighbourhood of x, and $\Delta f = f(x + \Delta x) - f(x)$ be change in f corresponding to an infinitesimal change Δx in x. Then f is said to be differentiable at x if $\Delta f = f(x + \Delta x) - f(x)$ can be expressed in the form

$$\Delta f = f(x + \Delta x) - f(x) = A\Delta x + \Delta x\ \phi(\Delta x), \tag{1}$$

where A is independent of Δx, and ϕ is a function of Δx such that $\phi(\Delta x) \to 0$ as $\Delta x \to 0$.

Dividing (1) throughout by Δx, and then taking limit as $\Delta x \to 0$, we get

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

We define the limit $\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$ or $\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$ to be the derivative of f with respect to x, and denote it by $\frac{df}{dx}$ or $f'(x)$. So $\frac{df}{dx} = f'(x) = A$. Obviously, the existence of $f'(x)$ is necessary for the function $f(x)$ to be differentiable at x. Now, (1) can be rewritten as

$$\Delta f = f(x + \Delta x) - f(x) = f'(x)\Delta x + \Delta x\ \phi(\Delta x), \tag{2}$$

Next, we define the expression $f'(x)\Delta x$ to be the differential of f and denote it by df. So $df = f'(x)\Delta x$. If we choose $f(x) = x$, then $dx = 1.\Delta x = \Delta x$. Therefore, the differential of f can be written as $df = f'(x)dx$.

Remark1: Notice that by definition of differential, $dx = \Delta x$. Likewise, we have $df = \Delta f$. It means, differentials can be replaced by the corresponding infinitesimals and vice versa. That is why, the expression $\frac{df}{dx}$ is interpreted in two ways. First, simply as the ratio of the differentials df and dx or the infinitesimals Δf and Δx. Second, it is the limit of the ratio of the infinitesimals Δf and Δx, that is, $\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$. In this case, it is derivative of f with respect to x. Thus, the differential of f is df while the derivative of f, denoted by $f'(x)$ or $\frac{df}{dx}$, is the limit of the ratio of the infinitesimals Δf and Δx. In fact, the derivative measures a rate of change, while the differential measures the change itself. In practical situations, sometimes we are interested in the rate of change of some quantity and sometimes we want to know the change itself. Therefore, derivative and differential are not same but both are useful.

Remark2: If y is a differentiable function of x, and x is a differentiable function of t, then y is a differentiable function of t, and

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

known as the chain rule of derivatives. In the chain rule, it appears like we have multiplied and divided by dx. But it is not true because the expressions $\frac{dy}{dt}$, $\frac{dy}{dx}$ and $\frac{dx}{dt}$ are derivatives and not the ratios of the differentials dy, dx and dt. For, let Δt be an infinitesimal change in t, and Δx and Δy be corresponding changes in x and y, respectively. Then we can write,

$$\frac{\Delta y}{\Delta t} = \frac{\Delta y}{\Delta x}\frac{\Delta x}{\Delta t}$$

Here, we have really multiplied and divided by Δx. In the limit $\Delta t \to 0$, we get

$$\lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \left(\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}\right)\left(\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}\right).$$

which implies the chain rule

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

Remark3: Consider the integral $\int x \sin(x^2)dx$. Here, dx is the differential of x. It has no role at all in the integration, and is optional to write with the integrand function. But it is useful to write it in the integral for two reasons. First, it tells us explicitly that the variable of integration is x. Second, it helps us in using the change of variable conveniently. For instance, to solve the given integral, we use $x^2 = t$. This implies that $d(x^2) = dt$ or $\frac{d}{dx}(x^2)dx = dt$ or $2xdx = dt$. Consequently,

$$\int x \sin(x^2)dx = \frac{1}{2}\int \sin t dt = -\frac{1}{2}\cos t + C = -\frac{1}{2}\cos(x^2) + C,$$

where C is constant of integration.

Geometry of $f'(x)$: The ratio $\dfrac{f(x+\Delta x) - f(x)}{\Delta x}$ is slope of the secant passing through the points $P(x, f(x))$ and $Q(x+\Delta x, f(x+\Delta x))$ of the curve $f(x)$. Since , by definition, tangent to the curve at the point $P(x, f(x))$ is the limiting position of secant PQ as $Q \to P$ or $\Delta x \to 0$, so the slope of the secant PQ becomes the slope of the tangent at P in the limit $\Delta x \to 0$. Therefore, $f'(x) = \lim\limits_{\Delta x \to 0} \dfrac{f(x+\Delta x) - f(x)}{\Delta x}$ is the slope of the tangent to the curve $f(x)$ at the point $P(x, f(x))$.

Ex. Show that the function $f(x) = x^2$ is differentiable.

Sol. We have

$f(x+\Delta x) - f(x) = (x+\Delta x)^2 - x^2 = 2x\Delta x + (\Delta x)^2 = A\Delta x + \Delta x\ \phi(\Delta x)$,

where $A = 2x$ is independent of Δx, and $\phi(\Delta x) = \Delta x \to 0$ as $\Delta x \to 0$. Therefore, $f(x)$ is differentiable and $f'(x) = \lim\limits_{\Delta x \to 0} \dfrac{f(x+\Delta x) - f(x)}{\Delta x} = A = 2x$.

Ex. Show that the function $f(x) = |x|$ is not differentiable at $x = 0$.

Sol. We have

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

But $\displaystyle\lim_{\Delta x \to 0^-} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to 0^-} \frac{-\Delta x}{\Delta x} = -1$ and $\displaystyle\lim_{\Delta x \to 0^+} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to 0^+} \frac{\Delta x}{\Delta x} = 1$.

So $\displaystyle\lim_{\Delta x \to 0} \frac{|\Delta x|}{\Delta x}$ does not exist and hence $f'(0)$ does not exist. Consequently, $f(x) = |x|$ is not differentiable at 0.

Geometry: The plot of $f(x) = |x|$ is shown in the Figure 4. We see that tangent at $(0, 0)$ from the left is the line $f(x) = -x$ with slope $\displaystyle\lim_{\Delta x \to 0^-} \frac{|\Delta x|}{\Delta x} = -1$ (left hand derivative) while the tangent at $(0, 0)$ from the right is the line $f(x) = x$ with slope $\displaystyle\lim_{\Delta x \to 0^+} \frac{|\Delta x|}{\Delta x} = 1$ (right hand derivative). So we do not have a unique tangent at $(0, 0)$, and consequently $f'(0)$ does not exist.

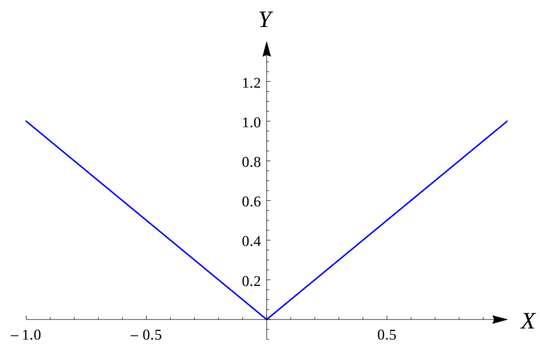


Figure 9: $f(x) = |x|$.

Differentiability implies continuity: Let a function $f(x)$ be differentiable at $x = x_0$. Then $f'(x_0)$ exists and $f(x_0 + \Delta x) - f(x_0) = f'(x_0)\Delta x + \Delta x\ \phi(\Delta x)$, where $\phi(\Delta x) \to 0$ as $\Delta x \to 0$. So in the limit $\Delta x \to 0$, we get $\displaystyle\lim_{\Delta x \to 0} f(x_0 + \Delta x) = f(x_0)$. This shows that $f(x)$ is continuous at $x = x_0$.

However, a continuous function need not be differentiable. For example, $f(x) = |x|$ is continuous at $x = 0$ but not differentiable at $x = 0$.

Why is differentiability considered on open intervals in mean value theorems?

We do have a notion of a function being differentiable on a closed interval.

The reason that mean value theorems, for instance, the Rolle's theorem talks about differentiability on the open interval (a, b) is that it is a weaker assumption than requiring differentiability on $[a, b]$.

Normally, theorems might try to make the assumptions as weak as possible, to be more generally applicable. For instance, the function:

$f(x) = x \sin(1/x)$, $x > 0$ and $f(0) = 0$

is continuous at 0, and differentiable everywhere except at 0.

We can still apply Rolle's theorem to this function on say the interval $(0, 1/\pi)$. If the statement of Rolle's theorem required the use of the closed interval, then we could not apply it to this function.

Differentiability of a function of two real variables

Let f be a function defined in some neighbourhood of (x, y), and $\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y)$ be change in f corresponding to an infinitesimal change Δx in x and an infinitesimal change Δy in y . Then f is said to be differentiable at (x, y) if $\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y)$ can be expressed in the form

$$\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y) = A\Delta x + B\Delta y + \Delta x\ \phi(\Delta x, \Delta y) + \Delta y\ \psi(\Delta x, \Delta y), \tag{3}$$

where A and B both are independent of Δx and Δy; and the functions $\phi(\Delta x, \Delta y)$ and $\psi(\Delta x, \Delta y)$ both tend to 0 as $(\Delta x, \Delta y) \to (0, 0)$.

Choosing $\Delta y = 0$ and dividing (3) throughout by Δx, and then taking limit as $\Delta x \to 0$, we get

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

We define the limit $\displaystyle\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$ or $\displaystyle\lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$ to be the partial derivative of f with respect to x, and denote it by $\dfrac{\partial f}{\partial x}$ or $f_x(x, y)$. So $\dfrac{\partial f}{\partial x} = f_x(x, y) = A$.

Similarly, choosing $\Delta x = 0$ and dividing (3) throughout by Δy, and then taking limit as $\Delta y \to 0$, we get

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

We define the limit $\displaystyle\lim_{\Delta y \to 0} \frac{\Delta f}{\Delta y}$ or $\displaystyle\lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$ to be the partial derivative of f with respect to y, and denote it by $\dfrac{\partial f}{\partial y}$ or $f_y(x, y)$. So $\dfrac{\partial f}{\partial y} = f_y(x, y) = B$.

We notice that the existence of the partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ is essential for the function $f(x, y)$ to be differentiable at (x, y). Now, (3) can be rewritten as

$$\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y) = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \Delta x\ \phi(\Delta x, \Delta y) + \Delta y\ \psi(\Delta x, \Delta y), \tag{4}$$

Next, we define the expression $\dfrac{\partial f}{\partial x}\Delta x + \dfrac{\partial f}{\partial y}\Delta y$ to be the total differential of f and denote it by df. So $df = \dfrac{\partial f}{\partial x}\Delta x + \dfrac{\partial f}{\partial y}\Delta y$. If we choose $f(x, y) = x$, then $dx = 1.\Delta x = \Delta x$. Similarly, the choice $f(x, y) = y$ leads to $dy = 1.\Delta y = \Delta y$. Therefore, the total differential of f can be written as

$$\boxed{df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.}$$

Ex. Show that the function $f(x) = x^2 + y^2$ is differentiable.

Sol. We have

$f(x + \Delta x) - f(x) = (x + \Delta x)^2 + (y + \Delta y)^2 - x^2 - y^2 = 2x\Delta x + 2y\Delta y + (\Delta x)^2 + (\Delta y)^2$

$= A\Delta x + B\Delta y + \Delta x\ \phi(\Delta x, \Delta y) + \Delta y\ \phi(\Delta x, \Delta y)$,

where $A = 2x$ and $B = 2y$ both are independent of Δx and Δy, and $\phi(\Delta x, \Delta y) = \Delta x \to 0$ and $\psi(\Delta x, \Delta y) = \Delta y \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. Therefore, $f(x, y)$ is differentiable at (x, y).

Ex. Show that the function $f(x) = \begin{cases} \frac{x^2+y^2}{|x|+|y|} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is continuous at $(0, 0)$ but not differentiable.

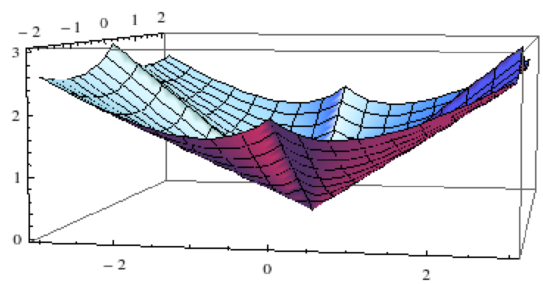


Figure 10: Graph of $f(x, y)$ shows that the surface is continuous but with a sharp edge at $(0, 0, 0)$.

Sol. Let $\epsilon > 0$ be given. Then

$|f(x, y) - f(0, 0)| = \dfrac{x^2 + y^2}{|x| + |y|} \leq \dfrac{[|x| + |y|]^2}{|x| + |y|} = |x| + |y| < \epsilon$ if $|x| < \epsilon/2$ and $|y| < \epsilon/2$.

Choosing $\delta = \epsilon/2$, we have

$|x| < \delta,\ |y| < \delta \implies |f(x, y) - f(0, 0)| < \epsilon.$

This shows that $f(x, y)$ is continuous at $(0, 0)$.

Now, $f_x(0, 0) = \lim\limits_{\Delta x \to 0} \dfrac{f(0 + \Delta x, 0) - f(0, 0)}{\Delta x} = \lim\limits_{\Delta x \to 0} \dfrac{\Delta x}{|\Delta x|}$, which does not exist.

Similarly, $f_y(0, 0) = \lim\limits_{\Delta y \to 0} \dfrac{f(0, 0 + \Delta y) - f(0, 0)}{\Delta y} = \lim\limits_{\Delta y \to 0} \dfrac{\Delta y}{|\Delta y|}$ does not exist.

So $f(x, y)$ is not differentiable at $(0, 0)$ since existence of partial derivatives is necessary for differentiability.

Differentiability implies continuity: Let a function $f(x, y)$ be differentiable at (x_0, y_0). Then $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ both exists and

$f(x_0 + \Delta x, y_0 + \Delta y) - f(x, y) = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \Delta x\ \phi(\Delta x, \Delta y) + \Delta y\ \psi(\Delta x, \Delta y)$,

where $\phi(\Delta x, \Delta y) \to 0$ and $\psi(\Delta x, \Delta y) \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. So in the limit $(\Delta x, \Delta y) \to (0, 0)$, we get $\lim\limits_{(\Delta x,\Delta y)\to(0,0)} f(x_0+\Delta x, y_0+\Delta y) = f(x_0,y_0)$. This shows that $f(x,y)$ is continuous at (x_0,y_0).

However, a continuous function need not be differentiable as we have seen in the previous example.

Ex. Show that the function $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y)\neq(0,0) \\ 0 & \text{if } (x,y)=(0,0) \end{cases}$ is not continuous at $(0,0)$ but partial derivatives exist.

Sol. We shall show that the limit $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ is path dependent.
Suppose $(x,y)\to(0,0)$ along the straight line path $y=mx$. Then
$\lim\limits_{(x,y)\to(0,0)} f(x,y) = \lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2} = \lim\limits_{x\to0} \dfrac{x.mx}{x^2+m^2x^2} = \dfrac{m}{1+m^2}$.
So the given limit depends on m, the slope of the path line. Hence, it get different values for different values of m, and consequently it does not exist. So $f(x,y)$ is not continuous at $(0,0)$.
Now, $f_x(0,0) = \lim\limits_{\Delta x\to0} \dfrac{f(0+\Delta x,0)-f(0,0)}{\Delta x} = \lim\limits_{\Delta x\to0} \dfrac{0}{\Delta x} = 0$.
Similarly, $f_y(0,0) = \lim\limits_{\Delta y\to0} \dfrac{f(0,0+\Delta y)-f(0,0)}{\Delta y} = \lim\limits_{\Delta y\to0} \dfrac{0}{\Delta y} = 0$.
Note: From the previous two examples, we may notice that continuity of $f(x,y)$ has nothing to do with its partial derivatives.

Sufficient condition for differentiability: If the partial derivatives f_x and f_y are continuous at (x_0,y_0), then $f(x,y)$ is differentiable at (x_0,y_0).

Geometry of partial derivative: Consider the partial derivative $f_x(x_0,y_0)$ of $f(x,y)$ at (x_0,y_0). The partial derivative $f_x(x_0,y_0)$, in fact, is the ordinary derivative of $f(x,y_0)$, which is the curve in XYZ-space determined by the section of the plane $y=y_0$ with the surface $z=f(x,y)$. Therefore, $f_x(x_0,y_0)$ is the slope of the tangent to the curve $z=f(x,y_0)$ at the point $(x_0,y_0,f(x_0,y_0))$ as shown in Figure 10.

Ex. Find f_x and f_y given that $f(x,y)=\tan^{-1}\left(\frac{x}{y}\right)$.

Sol. We have

$$f_x = \frac{1}{1+\frac{x^2}{y^2}}\frac{\partial}{\partial x}\left(\frac{x}{y}\right) = \frac{y^2}{x^2+y^2}\frac{1}{y} = \frac{y}{x^2+y^2}.$$

$$f_y = \frac{1}{1+\frac{x^2}{y^2}}\frac{\partial}{\partial y}\left(\frac{x}{y}\right) = \frac{y^2}{x^2+y^2}\frac{-x}{y^2} = \frac{-x}{x^2+y^2}.$$

SECTION 14.4 (The Chain Rule)

In the following, we state some chain rules of differentiation without proof.

- If f is differentiable function of x, and x is differentiable function of t, then f is differentiable function of t, and

$$\frac{df}{dt} = \frac{df}{dx}\frac{dx}{dt}.$$

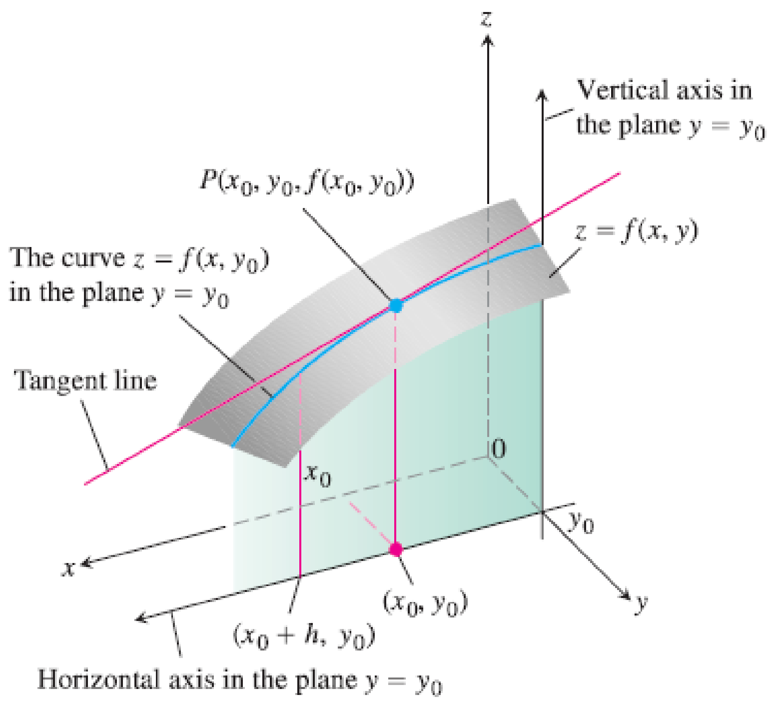


Figure 11:

Here, to reach from f to t, there is one path, namely, f to x and then x to t.

- If f is differentiable function of x, y, and x, y are differentiable function of t, then f is differentiable function of t, and

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

In this formula, two terms are appearing in addition because there are two possible paths to reach from f to t. One is from f to x and x to t. Second is from f to y and y to t.

- If f is differentiable function of x, y, and x, y are differentiable function of u, v, then f is differentiable function of u, v and

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u},$$

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}.$$

Here, we have two paths from f to u, namely, $f \to x \to u$ and $f \to y \to u$. Similarly, the two paths from f to v are $f \to x \to v$ and $f \to y \to v$.

Likewise, we can write the chain rule for a composite function f with any number of variables.

Ex. Use the Chain Rule to find the derivative of $f(x, y) = xy$ with respect to t along the path $x = \cos t$, $y = \sin t$.

Sol. We have
$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} == y(-\sin t) + x\cos t = -\sin^2 t + \cos^2 t = \cos 2t.$$

Ex. Given a function $f(x-y, y-z, z-x)$, show that $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z} = 0$.

Sol. Let $p = x - y$, $q = y - z$ and $r = z - x$ so that f is a function of p, q, r, and p, q, r are functions of x, y, z. So
$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial p}\frac{\partial p}{\partial x} + \frac{\partial f}{\partial q}\frac{\partial q}{\partial x} + \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} = \frac{\partial f}{\partial p}.1 + \frac{\partial f}{\partial q}.0 + \frac{\partial f}{\partial r}(-1) = \frac{\partial f}{\partial p} - \frac{\partial f}{\partial r}.$$

Similarly,
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial p}\frac{\partial p}{\partial y} + \frac{\partial f}{\partial q}\frac{\partial q}{\partial y} + \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} = \frac{\partial f}{\partial q} - \frac{\partial f}{\partial p}$$

and
$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial p}\frac{\partial p}{\partial z} + \frac{\partial f}{\partial q}\frac{\partial q}{\partial z} + \frac{\partial f}{\partial r}\frac{\partial r}{\partial z} = \frac{\partial f}{\partial r} - \frac{\partial f}{\partial q}.$$

Adding $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$, we get the required result.

Implicit Differentiation

Let $f(x, y) = 0$ be an implicit relation connecting x and y. Then y may be regarded as function of x via the relation $f(x, y) = 0$. Thus, f is function of x, y, and x, y both may be regarded as functions of x. So
$$\frac{df}{dx} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}.$$

Also, $f(x, y) = 0$ implies $\frac{df}{dx} = 0$. Thus, we obtain
$$\boxed{\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} = -\frac{f_x}{f_y}}.$$

This formula expresses $\frac{dy}{dx}$ in terms of the partial derivatives of $f(x, y)$. It eases the life when we calculate $\frac{dy}{dx}$ from an implicit relation in x and y.

Ex. Find $\frac{dy}{dx}$ from the relation $x^y + y^x = 1$.

Sol. Assuming $f(x, y) = x^y + y^x - 1$, we get
$$\frac{dy}{dx} = -\frac{f_x}{f_y} = \frac{yx^{y-1} + y^x \ln y}{x^y \ln x + xy^{x-1}}.$$

Ex. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ from the relation $x^3 + z^2 + ye^{xz} + z\cos y = 0$.

Sol. Assuming $f(x, y) = x^3 + z^2 + ye^{xz} + z\cos y$, we get

$$\frac{\partial z}{\partial x} = -\frac{f_x}{f_z} = -\frac{3x^2 + yze^{xz}}{2z + xye^{xz} + \cos y},$$

$$\frac{\partial z}{\partial y} = -\frac{f_y}{f_z} = -\frac{e^{xz} - z\sin y}{2z + xye^{xz} + \cos y}.$$

Higher Order Partial Derivatives

Let $f(x, y)$ be a function of x and y. Then, as we have seen earlier, its first order derivatives are

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

and

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

The second order partial derivatives of $f(x, y)$ are

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

$$f_{xy}(x, y) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \lim_{k \to 0} \frac{f_x(x, y+k) - f_x(x, y)}{k},$$

$$f_{yx}(x, y) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \lim_{h \to 0} \frac{f_y(x+h, y) - f_y(x, y)}{h},$$

$$f_{yy}(x, y) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \lim_{k \to 0} \frac{f_y(x, y+k) - f_y(x, y)}{h}.$$

Likewise we can define third order and higher order partial derivatives.

Ex. Find 2nd order partial derivatives of $f(x, y) = x^3 + x^2y^2 + y^3$.

Sol. We have $f_x = 3x^2 + 2xy^2$, $f_y = 2x^2y + 3y^2$.

$$f_{xx} = 6x + 2y^2, \quad f_{xy} = 4xy, \quad f_{yx} = 4xy, \quad f_{yy} = 2x^2 + 6y.$$

Caution: While dealing with problems in your text book, you will notice that $f_{xy} = f_{yx}$. In other words, it does not matter in which order we calculate the partial derivatives. However, you should keep in mind that the equality $f_{xy} = f_{yx}$ is ensured only when f_{xy} and f_{yx} are continuous functions, otherwise f_{xy} and f_{yx} need not be equal.

For example, consider the function

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

Then it can be shown that $f_{yx}(0,0) = 1$ and $f_{xy}(0,0) = -1$. So $f_{xy}(0,0) \neq f_{yx}(0,0)$.

For,

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h}.$$

Next,

$$f_y(h,0) = \lim_{k\to 0} \frac{f(h,k) - f(h,0)}{k} = \lim_{k\to 0} \frac{\frac{hk(h^2-k^2)}{h^2+k^2} - 0}{k} = h$$

and

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = \lim_{k\to 0} \frac{0-0}{k} = 0.$$

It follows that

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = \lim_{h\to 0} \frac{h-0}{h} = 1.$$

Likewise, we get

$$f_{xy}(0,0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = \lim_{k\to 0} \frac{-k-0}{k} = -1.$$

SECTION 14.5 (Directional Derivatives and Gradient Vectors)

The directional derivative of a function $f(x, y, z)$ at a point (x, y, z) in the direction of a unit vector $\hat{n} = l\hat{i} + m\hat{j} + n\hat{k}$, denoted by $\left(\frac{df}{dt}\right)_{\hat{n}}$ or $D_{\hat{n}}f$, is defined as the limit

$$\left(\frac{df}{dt}\right)_{\hat{n}} = \lim_{t\to 0} \frac{f(x + lt, y + mt, z + nt) - f(x, y, z)}{t},$$

provided it exists. Note that $(x + lt, y + mt, z + nt)$ is a neighbouring point of (x, y, z) on a line through (x, y, z) in the direction of $\hat{n} = l\hat{i} + m\hat{j} + n\hat{k}$. In particular, the directional derivative of $f(x, y, z)$ in a direction parallel to x-axis, where $\hat{n} = \hat{i}$, reads as

$$\left(\frac{df}{dt}\right)_{\hat{i}} = \lim_{t\to 0} \frac{f(x + t, y, z) - f(x, y, z)}{t},$$

which is nothing but the partial derivative of f with respect to x, that is, f_x. Similarly, directional derivatives of $f(x, y, z)$ in directions parallel to y-axis and z-axis are f_y and f_z, respectively. So we see that the partial derivatives f_x, f_y and f_z are special cases of the directional derivative $\left(\frac{df}{dt}\right)_{\hat{n}}$ of $f(x, y, z)$.

The parametric equations of the line through (x, y, z) in the direction $\hat{n} = l\hat{i}+m\hat{j}+n\hat{k}$ are $x(t) = x+lt$, $y(t) = y + mt$ and $z(t) = z + nt$. So by chain rule, we have

$$\frac{df}{dt} = \frac{\partial f}{\partial x}l + \frac{\partial f}{\partial y}m + \frac{\partial f}{\partial z}n = \left(\frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}\right).(l\hat{i} + m\hat{j} + n\hat{k}).$$

So the directional derivative $\left(\frac{df}{dt}\right)_{\hat{n}}$ may be written as

$$\left(\frac{df}{dt}\right)_{\hat{n}} = \nabla f.\hat{n} \ ,$$

where $\nabla f = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}$ is known as gradient vector of $f(x, y, z)$.

This vector representation of directional derivative is useful in view of the fact that $\left(\frac{df}{dt}\right)_{\hat{n}}$ represents the rate of change of $f(x, y, z)$ in the direction $\hat{n}$. If θ is angle between ∇f and $\hat{n}$, then we have

$$\left(\frac{df}{dt}\right)_{\hat{n}} = |\nabla f| \cos\theta.$$

Obviously, maximum value of directional derivative is $|\nabla f|$, and is obtained in the direction of ∇f. Similarly, minimum value of directional derivative is $-|\nabla f|$, in the direction opposite to ∇f. It is 0 in the direction perpendicular to ∇f.

Let $P(x, y, z)$ be a point on a level surface $f(x, y, z) = c$, where c is a constant. Suppose $\overrightarrow{r}(t) = x(t)\hat{i} + y(t)\hat{j} + z(t)\hat{k}$ be any curve on the level surface $f(x, y, z) = c$ passing through the point $P(x, y, z)$ so that $f(x(t), y(t), z(t)) = c$. Therefore, we have

$$0 = \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = \nabla f.\frac{d\overrightarrow{r}}{dt}.$$

Since $\frac{d\overrightarrow{r}}{dt}$ is tangent vector to each curve $\overrightarrow{r}(t) = x(t)\hat{i} + y(t)\hat{j} + z(t)\hat{k}$ on the level surface $f(x, y, z) = c$ passing through the point $P(x, y, z)$, so it follows that the gradient vector ∇f is normal to the level surface at $P(x, y, z)$. Further, it implies that the directional derivative at any point of a level surface is maximum along the direction of normal to the surface at that point.

Ex. Find gradϕ at the point $(1, -2, -1)$, given that $\phi = 3x^2y - y^3z^2$.

Ans. $-12\hat{i} - 9\hat{j} - 16\hat{k}$.

Ex. Find the directional derivative of $f = x^2 - y^2 + 2z^2$ at the point $P(1, 2, 3)$ in the direction of line PQ, where Q is the point (5,0,4). Also find the direction and magnitude of maximum directional derivetive at P.

Ans. $4\sqrt{21}/3$, $2\hat{i} - 4\hat{j} + 12\hat{k}$, $2\sqrt{41}$.

Ex. If $\overrightarrow{r} = x\hat{i} + y\hat{j} + z\hat{k}$, then show that $\nabla r^n = nr^{n-2}\overrightarrow{r}$.

Ex. Find the angle between the normals to the surfaces $x^2 + y^2 + z^2 = 9$ and $z = x^2 + y^2 - 3$ at the point $(2, -1, 2)$.

Ans. $\cos^{-1}(8\sqrt{21}/63)$.

Ex. Find the values of constants a and b so that the surfaces $ax^2 - byz = (a + 2)x$ and $4x^2y + z^3 = 4$ may intersect orthogonally at the point $(1, -1, 2)$.

Ans. $a = 2.5$, $b = 1$.

Ex. Let $f(x, y) = \begin{cases} 0, & y = x \\ 1, & y \neq x \end{cases}$.

Show that the directional derivative exists at $(0, 0)$ along the line $y = x$ though ∇f does not exists at $(0, 0)$.

Ex. Let $f(x, y) = \begin{cases} 0, & xy \neq 0 \\ 1, & xy = 0 \end{cases}$.

Show that the directional derivative does not exist at $(0, 0)$ along the line $y = x$ though ∇f exists at $(0, 0)$.

SECTION 14.6 (Tangent Planes and Differentials)

The tangent plane at the point $P_0(x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$ of a differentiable function f is the plane through P_0 normal to $\nabla f|_{P_0}$. So direction ratios of the normal to the tangent plane at $P_0(x_0, y_0, z_0)$ are $f_x(P_0)$, $f_y(P_0)$ and $f_z(P_0)$. So equation of the tangent plane is

$$f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0) = 0,$$

and the equation of the normal line at P_0 is given by

$$\frac{x - x_0}{f_x(P_0)} = \frac{y - y_0}{f_y(P_0)} = \frac{z - z_0}{f_z(P_0)}.$$

Ex. Find the equation of tangent plane and normal to the sphere $x^2 + y^2 + z^2 = 1$ at the point $(0, 0, 1)$.
Sol. Normal vector to the given sphere is
$\overrightarrow{n} = \nabla(x^2 + y^2 + z^2) = 2x\hat{i} + 2y\hat{j} + 2z\hat{k} = 2\hat{k}$, at the point $(0, 0, 1)$. So direction ratios of the normal to the tangent plane at $(0, 0, 1)$ are 0, 0 and 2. So equation of the tangent plane at $(0, 0, 1)$ is

$$0(x - 0) + 0(y - 0) + 2(z - 1) = 0 \quad \text{or} \quad z = 1.$$

Equation of normal at $(0, 0, 1)$ is given by

$$\frac{x - 0}{0} = \frac{y - 0}{0} = \frac{z - 0}{2},$$

which is nothing but the equation of z-axis, as expected.

Estimating the change in f in a given direction

To estimate the change in the value of a differentiable function f when we move a small distance ds from a point P_0 in a particular direction $\hat{n}$, we use the formula

$$df = (\nabla f|_{P_0}).\hat{n}(ds).$$

Ex. Estimate how much the value of $f(x, y, z) = y \sin x + 2yz$ will change if the point $P(x, y, z)$ moves 0.1 unit from $P_0(0, 1, 0)$ straight towards $P_1(2, 2, -2)$.
Sol. We find $\hat{n} = \frac{\overrightarrow{P_0P_1}}{|\overrightarrow{P_0P_1}|} = \frac{2}{3}\hat{i} + \frac{1}{3}\hat{j} - \frac{2}{3}\hat{k}$.
Next, $\nabla f|_{P_0} = y \cos x\hat{i} + (\sin x + 2z)\hat{j} + 2y\hat{k}|_{P_0(0,1,0)} = \hat{i} + 2\hat{k}$.
So with $ds = 0.1$, the resulting change in f is

$$df = (\nabla f|_{P_0}).\hat{n}(ds) = (-2/3)(0.1) \approx -0.067$$

Linearization of $f(x, y)$

The linearization of a function $f(x, y)$ at a point (x_0, y_0), where f is differentiable, is the function $L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$.
The approximation $L(x, y) \approx f(x, y)$ is the standard linear approximation of f at (x_0, y_0).
Note that $z = L(x, y)$ is tangent plane to the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$. Thus, the linearization of a function of two variables is a tangent-plane approximation.

In this approximation, we can estimate the error occurred as well by using the following result:

If f has continuous first and second partial derivatives throughout an open set containing a rectangle R centered at (x_0, y_0) and if M is any upper bound for the values of $|f_{xx}|$, $|f_{xy}|$ and $|f_{yy}|$ on R, then the error $E(x, y)$ occurred in replacing $f(x, y)$ on R by its linearization

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

satisfies the inequality

$$|E(x, y)| \leq \frac{1}{2}M(|x - x_0| + |y - y_0|)^2.$$

Ex. Find the linearization of $f(x, y) = x^2 - xy + \frac{1}{2}y^2 + 3$ at the point (3, 2). Also find an upper bound for the error in the approximation over the rectangle $R : |x - 3| \leq 0.1$, $|y - 2| \leq 0.1$.
Sol. We find

$$L(x, y) = f(x_0, y_0)+f_x(x_0, y_0)(x-x_0)+f_y(x_0, y_0)(y-y_0) = f(3, 2)+f_x(3, 2)(x-3)+f_y(3, 2)(y-2) = 4x-y-2.$$

We know that the error satisfies the inequality

$$|E(x, y)| \leq \frac{1}{2}M(|x - x_0| + |y - y_0|)^2.$$

To find a suitable value for M, we calculate $|f_{xx}| = 2$, $|f_{xy}| = 1$ and $|f_{yy}| = 1$. The largest of these is 2, so we may safely take M to be 2. With $(x_0, y_0) = (3, 2)$, over the given rectangle $R : |x - 3| \leq 0.1$, $|y - 2| \leq 0.1$, we find

$$|E(x, y)| \leq \frac{1}{2}(2)(|x - 3| + |y - 2|)^2 = 0.04.$$

Differentials

Suppose the first order partial derivatives of a function $f(x, y)$ exist at (x_0, y_0). If x and y change from x_0 and y_0 by small amounts dx and dy respectively, then the total differential

$$df = f_x(x_0, y_0)dx + f_y(x_0, y_0)dy$$

gives a good approximation of the resulting change in f.

Ex. Suppose that a cylindrical can is designed to have a redius of 1 in. and a height of 5 in., but that the radius and height are off by the amounts $dr = +0.03$ and $dh = -0.1$. Estimate the resulting absolute change in the volume of the can.

Sol. To estimate the absolute change in $V = \pi r^2 h$, we use

$$\Delta V \approx dV = V_r(r_0, h_0)dr + V_h(r_0, h_0)dh.$$

With $V_r = 2\pi rh$ and $V_h = \pi r^2$, we get

$$dV = 2\pi r_0 h_0 dr + \pi r_0^2(r_0, h_0)dh = 2\pi(1)(5)(0.03) + \pi(1)^2(-0.1) = 0.2\pi \approx 0.63 \text{ in}^3.$$

SECTION 14.7 (Extreme Values and Saddle Points)

Maxima and minima of a function of one variable

Let $f(x)$ be a function defined in some neighbourhood of a point $x = a$. Then $f(x)$ is said to have a local maximum at a if $f(x) < f(a)$ in some deleted neighbourhood of a. Equivalently, $f(x)$ has local maximum at a if there exists some $\epsilon > 0$ such that $f(a+h) - f(a) < 0$ for all $h \in (-\epsilon, \epsilon) - \{0\}$.

For example, $f(x) = \sqrt{1 - x^2}$ has local maximum at $x = 0$.

Likewise, $f(x)$ is said to have a local minimum at a if $f(x) > f(a)$ in some deleted neighbourhood of a. Equivalently, $f(x)$ has local minimum at a if there exists some $\epsilon > 0$ such that $f(a+h) - f(a) > 0$ for all $h \in (-\epsilon, \epsilon) - \{0\}$.

For example, $f(x) = x^2$ has local minimum at $x = 0$.

Further, if in every deleted neighbourhood of a, there are values of x for which $f(x) < f(a)$, and there are also values of x for which $f(x) > f(a)$, then $f(x)$ is said to have point of inflexion at a.

For example, $f(x) = x^3$ has point of inflexion at $x = 0$ because in every neighbourhood of 0, we have $f(x) < f(0)$ for $x < 0$, and $f(x) > f(0)$ for $x > 0$.

Critical point

A critical point of a function is a point where its derivative vanishes. So $x = a$ is a critical point of $f(x)$ if $f'(a) = 0$.

The points of maxima and minima of a function are always its critical points.

For, if $f(x)$ has local maximum at $x = a$, then $f(x) \leq f(a)$ throughout some nbd of a. So we have

$$f'(a) = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} \geq 0 \quad (\because \quad x - a < 0)$$

and

$$f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \leq 0. \quad (\because \quad x - a > 0)$$

It follows that $f'(a) = 0$. Likewise is the case of local minimum.

First derivative test for maxima and minima

Suppose $x = a$ is a critical point of $f(x)$. If there exists some neighbourhood of a in which $f'(x)$ changes its sign from positive to negative, then $x = a$ is a point of maxima. This happens because as x goes from left to right in the neighbourhood of a, the angle of the tangent to the curve changes from acute to obtuse, and consequently slope of the tangent, that is, $f'(x)$ changes from positive to negative. Likewise, if $f'(x)$ changes its sign from negative to positive in some neighbourhood of a, then $x = a$ is a point of minima.

Second derivative test for maxima and minima

Suppose $x = a$ is a critical point of $f(x)$ so that $f'(a) = 0$. If $f''(a) < 0$, then by continuity $f''(x) < 0$ in some neighbourhood a, say $N(a)$. Let $a + h$ be some neighbouring point of a inside $N(a)$. Then by Taylor's formula[1], we have

$$f(a + h) - f(a) = \frac{1}{2!}h^2 f''(a + \theta h) < 0,$$

since $a + \theta h$ lies in $N(a)$, where $f''(x) < 0$. It follows that $x = a$ is a point of maxima when $f''(a) < 0$. Likewise, $x = a$ is a point of minima when $f''(a) > 0$. If $f''(a) = 0$ but $f'''(a) \neq 0$, then $x = a$ is a point of inflexion.

[1] **Taylor's formula for a function of one variable:** If a function $f(x)$ possesses continuous derivatives up to nth order in some neighbourhood of a point a, then for any point $a+h$ in the neighbourhood of a, there exists a real number $\theta \in (0, 1)$ such that

$$f(a + h) = f(a) + hf'(a) + \frac{1}{2!}h^2 f''(a) + ... + \frac{1}{(n-1)!}h^{n-1} f^{(n-1)}(a) + \frac{1}{n!}h^n f^{(n)}(a + \theta h).$$

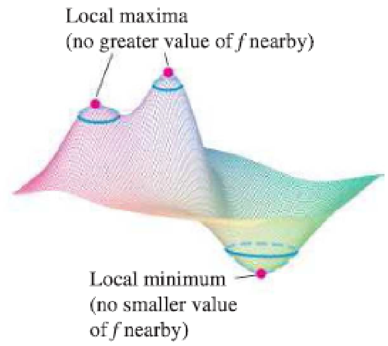
In particular, for $n = 2$, we have

$$f(a + h) = f(a) + hf'(a) + \frac{1}{2!}h^2 f''(a + \theta h)$$

Note: In case, the domain of the function is a closed interval, we can find its absolute maximum and minimum values by considering the values of the function at the end points of the interval and the critical points lying inside the interval.

Maxima and minima of a function of two variables

Let $f(x,y)$ be a function defined in some neighbourhood of a point (a,b). Then $f(x,y)$ is said to have a local maximum at (a,b) if $f(x,y) < f(a,b)$ in some deleted neighbourhood of (a,b). Equivalently, $f(x,y)$ has local maximum at (a,b) if there exists some $\epsilon > 0$ such that $f(a+h,b+k) - f(a,b) < 0$ for all $h \in (-\epsilon,\epsilon) - \{0\}$ and $k \in (-\epsilon,\epsilon) - \{0\}$.



For example, $f(x,y) = \sqrt{1-x^2-y^2}$ has local maximum at $(0,0)$.

Likewise, $f(x,y)$ is said to have a local minimum at (a,b) if $f(x,y) > f(a,b)$ in some deleted neighbourhood of (a,b). Equivalently, $f(x,y)$ has local minimum at (a,b) if there exists some $\epsilon > 0$ such that $f(a+h,b+k) - f(a,b) > 0$ for all $h \in (-\epsilon,\epsilon) - \{0\}$ and $k \in (-\epsilon,\epsilon) - \{0\}$.

For example, $f(x,y) = x^2 + y^2$ has local minimum at $(0,0)$.

Further, if in every deleted neighbourhood of (a,b), there are values of (x,y) for which $f(x,y) < f(a,b)$, and there are also values of (x,y) for which $f(x,y) > f(a,b)$, then $f(x,y)$ is said to have saddle point at (a,b).

Consider the point at the center of the surface of a saddle (seat used on a horse) placed on the back of a horse. It is easy to imagine that the surface points rise as we move on the surface from the center along the horse while the surface points fall as we move on the surface from the center perpendicular to the horse. Clearly, in every neighbourhood of the center of the this surface, there are points which are at more height than the center as well as there are points which are at less height than the center. It implies that center of the surface of saddle is a saddle point. (That is why the name saddle point!)

Critical point

A critical point of a function $f(x,y)$ is a point where both its partial derivatives vanish. So (a,b) is a critical point of $f(x,y)$ if $f_x(a,b) = 0$ and $f_y(a,b) = 0$.

The points of maxima and minima of a function $f(x,y)$ are always its critical points.

For, if f has a local extremum at (a,b), then the function $g(x) = f(x,b)$ has a local extremum at

$x = a$ (Figure 11). Therefore, $g'(a) = 0$. Now $g'(a) = f_x(a, b)$, so $f_x(a, b) = 0$. A similar argument with the function $h(y) = f(a, y)$ shows that $f_y(a, b) = 0$.

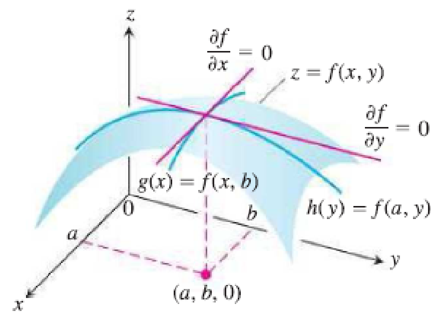


Figure 12:

Test for maxima and minima of $f(x, y)$

Suppose (a, b) is a critical point of $f(x, y)$ so that $f_x(a, b) = 0$ and $f_y(a, b) = 0$. If $h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b) \neq 0$, then by continuity $h^2 f_{xx}(x, y) + 2hk f_{xy}(x, y) + k^2 f_{yy}(x, y) \neq 0$ in some neighbourhood (a, b), say $N(a, b)$. Let $(a + h, b + k)$ be some neighbouring point of (a, b) inside $N(a, b)$. Then by Taylor's formula[2], we have

$$f(a+h, b+k) - f(a, b) = \frac{1}{2!}[h^2 f_{xx}(a+\theta h, b+\theta k) + 2hk f_{xy}(a+\theta h, b+\theta k) + k^2 f_{yy}(a+\theta h, b+\theta k)]$$

Since $(a + \theta h, b + \theta k)$ lies inside $N(a, b)$, sign of the right hand side expression would be same as of $h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b)$, which can be rewritten as

$$F = h^2 f_{xx}(a, b) + 2hk f_{xy}(a, b) + k^2 f_{yy}(a, b) = \frac{1}{f_{xx}}[(h f_{xx} + k f_{xy})^2 + k^2(f_{xx} f_{yy} - f_{xy}^2)]_{(a,b)}.$$

Now, we have the following cases:

(i) If $f_{xx} f_{yy} - f_{xy}^2 > 0$ and $f_{xx} > 0$, then $F > 0$ and consequently $f(a + h, b + k) - f(a, b) > 0$. Therefore, $f(x, y)$ has local minimum at (a, b).

[2]**Taylor's formula for a function of two variables:** If a function $f(x, y)$ possesses continuous partial derivatives up to nth order in some neighbourhood of a point (a, b), then for any point $(a + h, b + k)$ in the neighbourhood of (a, b), there exists a real number $\theta \in (0, 1)$ such that

$$f(a+h, b+k) = f(a, b) + Lf(a, b) + \frac{1}{2!}L^2 f(a, b) + \ldots\ldots\ldots\ldots + \frac{1}{(n-1)!}L^{n-1}f(a, b) + \frac{1}{n!}L^n f(a+\theta h, b+\theta h),$$

where $L = h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}$. In particular, for $n = 2$, we have

$$f(a+h, b+k) = f(a, b) + hf_x(a, b) + kf_y(a, b) + \frac{1}{2!}[h^2 f_{xx}(a+\theta h, b+\theta k) + 2hk f_{xy}(a+\theta h, b+\theta k) + k^2 f_{yy}(a+\theta h, b+\theta k)]$$

(ii) If $f_{xx}f_{yy} - f_{xy}^2 > 0$ and $f_{xx} < 0$, then $F < 0$ and consequently $f(a+h, b+k) - f(a,b) < 0$. Therefore, $f(x,y)$ has local maximum at (a,b).

(iii) If $f_{xx}f_{yy} - f_{xy}^2 < 0$, then F will attain positive as well as negative values in every neighbourhood of (a,b) (why?). So $f(x,y)$ has saddle point at (a,b).

(iv) If $f_{xx}f_{yy} - f_{xy}^2 = 0$, then we can not decide in general and further investigation is required. In this case, we directly check the sign of $f(a+h, b+k) - f(a,b)$ in the neighbourhood of (a,b) and take the decision accordingly.

Ex. Find the local maximum and minimum values of the function $f(x,y) = \sqrt{1 - x^2 - y^2}$.

Sol. Let $g(x,y) = f^2(x,y) = 1 - x^2 - y^2$. Then maxima and minima of $g(x,y)$ are same as of $f(x,y)$. The critical points of $g(x,y)$ are given by $g_x = -2x = 0$ and $g_y = -2y = 0$. So we get $(0,0)$ as the critical point. Next at $(0,0)$, we have $g_{xx} = -2$, $g_{xy} = 0$ and $g_{yy} = -2$. Therefore, $g_{xx}g_{yy} - g_{xy}^2 = 4 > 0$ and $g_{xx} = -2 < 0$. It implies that $g(x,y)$, and hence $f(x,y)$ has maximum value at $(0,0)$, and the maximum value reads as $f(0,0) = 1$.

Ex. Show that the hyperbolic paraboloid $z = x^2 - y^2$ possesses a saddle point at $(0,0)$.

Sol. Please try yourself.

Note: In case, the domain of the function is a closed and bounded region R, its absolute maximum/minimum values are the maximum/minimum of the values of the function at the critical points within R and at the boundary points of R as illustrated in the following example.

Ex. Find the absolute maximum and minimum values of

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

on the triangular region in the first quadrant bounded by the lines $x = 0$, $y = 0$, $y = 9 - x$.

Sol. The given triangular region is shown in Figure 12.

We find $f_x = 2 - 2x$ and $f_y = 2 - 2y$. So $f_x = 0$ and $f_y = 0$ yield the only critical point $(1,1)$ in the interior of R, and $f(1,1) = 4$.

Now we consider the three sides of the triangular region OAB one by one.

Along OA, $y = 0$. So we have $f(x,0) = 2 + 2x - x^2$, which may be treated a function of single variable x on the interval $[0,9]$. Its extreme values may occur at the end points $x = 0$, $x = 9$ and at the interior

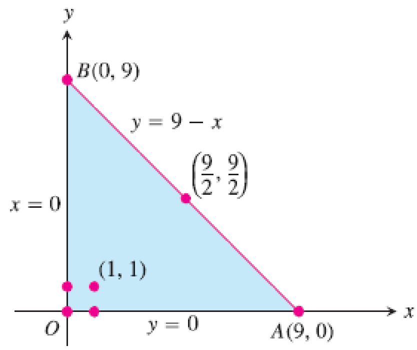


Figure 13:

points, where $f'(x,0) = 2 - 2x = 0$ or $x = 1$. The values are $f(0,0) = 2$, $f(9,0) = -61$ and $f(1,0) = 3$.

Along OB, $x = 0$. So we have $f(0,y) = 2 + 2y - y^2$, which may be treated a function of single variable y on the interval $[0,9]$. Its extreme values may occur at the end points $y = 0$, $y = 9$ and at the interior points, where $f'(0,y) = 2 - 2y = 0$ or $y = 1$. The values are $f(0,0) = 2$, $f(0,9) = -61$ and $f(0,1) = 3$.

Along AB, $y = 9 - x$. So we have $f(x,y) = -61 + 18x - 2x^2$, which may be treated a function of single variable x on the interval $[0,9]$. Its extreme values may occur at the end points $x = 0$, $x = 9$ and at the interior points, where $f'(x, 9-x) = 18 - 4x = 0$ or $x = 9/2$, and therefore $y = 9 - 9/2 = 9/2$. The values are $f(0,0) = 2$, $f(0,9) = -61$ and $f(9/2, 9/2) = -41/2$.

Finally we list all the values: 4, 2, −61, 3, −(41/2). So the absolute maximum value is 4, which f assumes at (1, 1). The absolute minimum value is −61, which f assumes at (0, 9) and (9, 0).

SECTION 14.8 (Lagrange Multipliers)

Constrained Maxima and Minima

We first consider a problem where a constrained maximum and minimum can be found by eliminating a variable.

Ex. Find the dimensions of a box open at the top with volume 32 cubic units requiring the least material for its construction.

Sol. Let x, y, z be length, breadth and height of the box so that its volume is $xyz = 32$ and surface area is $xy + 2yz + 2zx$. So we need to minimize $xy + 2yz + 2zx$ subject to the constraint $xyz = 32$. Using $z = \frac{32}{xy}$ into the surface area function, we get the two variable function

$$f(x,y) = xy + \frac{64}{x} + \frac{64}{y}.$$

Applying the maxima/minima test (please try yourself), we find minimum value of $f(x,y)$ for $x = 4$,

$y = 4$. Thus, the required dimensions of the box are 4, 4, 2.

Method of Lagrange Multipliers

Let $f(x, y, z)$ be differentiable in an open region containing a smooth curve C given by $\overrightarrow{r}(t) = x(t)\hat{i} + y(t)\hat{j} + z(t)\hat{k}$. Then on the curve C, we have

$$\frac{df}{dt} = \frac{df}{dx}\frac{dx}{dt} + \frac{df}{dy}\frac{dy}{dt} + \frac{df}{dz}\frac{dz}{dt} = \nabla f.\frac{d\overrightarrow{r}}{dt}.$$

At a point $P(x(t), y(t), z(t))$ of local maxima or minima of f on the curve C, $\frac{df}{dt} = 0$. So $\nabla f.\frac{d\overrightarrow{r}}{dt} = 0$, which in turn implies that ∇f is orthogonal to C at the point of maxima or minima.

Now suppose a differentiable function $f(x, y, z)$ takes a local maximum or minimum value at some point P_0 of a surface $g(x, y, z) = 0$, where $g(x, y, z)$ is differentiable function and $\nabla g \neq 0$. Then f takes a local maximum or minimum value at point P_0 relative to its values on every differentiable curve passing through P_0 and lying on the surface $g(x, y, z) = 0$. So ∇f is orthogonal to every differentiable curve C at the point P_0. It implies that ∇f is normal to the surface $g(x, y, z) = 0$ at the point P_0. Also, ∇g is normal to the surface $g(x, y, z) = 0$ at every point. It implies that the normal vectors ∇f and ∇g at P_0 are parallel vectors, and therefore $\nabla f = \lambda \nabla g$ for some non-zero constant λ, known as Lagrange multiplier. Thus, to find the local maximum and minimum values of f subject to the constraint $g(x, y, z) = 0$, we find the values of x, y, z and λ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g, \qquad g(x, y, z) = 0.$$

Note: It must be noted that Lagrange multiplier method gives the points related to the extremum (maximum or minimum) values of the function. It does not make distinction between points of maxima and minima. In other words, it does not tells us the exact nature of the point. We have to judge the nature of the point from the nature of the problem under consideration. It may be treated as a drawback of the Lagrange multiplier method.

Ex. Find the dimensions of a box open at the top with volume 32 cubic units requiring the least material for its construction.

Sol. Let x, y, z be length, breadth and height of the box so that its volume is $xyz = 32$ and surface area is $xy + 2yz + 2zx$. So we need to minimize $xy + 2yz + 2zx$ subject to the constraint $xyz = 32$. We model the given problem as a Lagrange multiplier problem with

$$f(x, y, z) = xy + 2yz + 2zx, \quad g(x, y, z) = xyz - 32$$

and look for the values of x, y, z and λ that satisfy the equations

$$\nabla f = \lambda \nabla g, \qquad g(x, y, z) = 0.$$

These are, in fact, four equations given by

$y + 2z - \lambda yz = 0$

$x + 2z - \lambda xz = 0$

$2x + 2y - \lambda xy = 0$

$xyz - 32 = 0$

Multiplying the first three equations by x, y and z, respectively, we get

$xy + 2xz - \lambda xyz = 0$

$xy + 2yz - \lambda xyz = 0$

$2xz + 2yz - \lambda xyz = 0$

Now subtracting the first two and last two of the above equations, we get $x = y$ and $y = 2z$, respectively. Using these in $xyz - 32 = 0$, we get $x = 4$, $y = 4$ and $z = 2$.

Ex. Find the maximum volume of the cuboid that can be inscribed inside the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Sol. Let $2x$, $2y$ and $2z$ be dimensions of the cuboid that can be inscribed inside the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. So volume of the cuboid is $8xyz$. Thus, we need to maximize

$$f(x, y, z) = 8xyz \tag{5}$$

subject to the constraint

$$g(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0. \tag{6}$$

So as per the method of Lagrange multipliers, in order to find the local maximum and minimum values of f subject to the constraint $g(x, y, z) = 0$, we find the values of x, y, z and λ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g, \qquad g(x, y, z) = 0.$$

Now, $\nabla f = \lambda \nabla g$ leads to the following equations:

$$4yz + \lambda \frac{x}{a^2} = 0, \;\; 4zx + \lambda \frac{y}{b^2} = 0, \;\; 4xy + \lambda \frac{z}{c^2} = 0.$$

Multiplying these equations respectively by x, y, z, and adding the resulting equations in view of (6), we get

$$12xyz + \lambda = 0.$$

Solving this equation with $4yz + \lambda \frac{x}{a^2} = 0$, we get $x = \frac{a}{\sqrt{3}}$. Likewise, we get $y = \frac{b}{\sqrt{3}}$ and $z = \frac{c}{\sqrt{3}}$. So in view of (5), maximum volume of the cuboid is $\frac{8abc}{3\sqrt{3}}$.

Lagrange Multipliers with Two Constraints

Suppose f assumes a local maximum or minimum value at a point P_0 lying on two surfaces $g_1(x, y, z) = 0$ and $g_2(x, y, z) = 0$. Then P_0 lies on the curve, say C of the intersection of the two surfaces. It follows that ∇f, ∇g_1 and ∇g_2 all are normal to C at P_0, and hence lie in the same plane so that $\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$ for some non-zero constants λ and μ, known as Lagrange multipliers. Thus, to find the local maximum and minimum values of f subject to two constraints $g_1(x, y, z) = 0$ and $g_2(x, y, z) = 0$, we find the values of x, y, z, λ and μ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2, \qquad g_1(x, y, z) = 0, \qquad g_2(x, y, z) = 0.$$

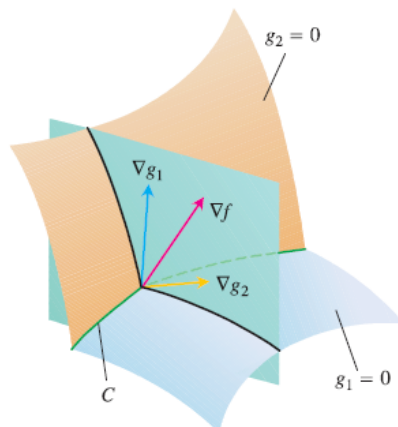


Figure 14:

Ex. The plane $x + y + z = 1$ cuts the cylinder $x^2 + y^2 = 1$ in an ellipse (Figure 14). Find the points on the ellipse that lie closest to and farthest from the origin.

Sol. We find the extreme values of

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

(the square of the distance from (x, y, z) to the origin) subject to the constraints

$$g_1(x, y, z) = x^2 + y^2 - 1 = 0,$$

$$g_2(x, y, z) = x + y + z - 1 = 0.$$

The gradient equation $\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$ leads to the following equations:

$2x = 2\lambda x + \mu,$

$2y = 2\lambda y + \mu,$

$2z = \mu.$

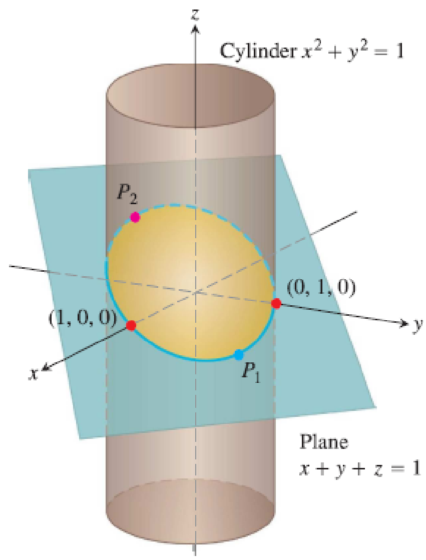


Figure 15:

Eliminating μ, we get the following equations:

$(1-\lambda)x = z$,

$(1-\lambda)y = z$.

Obviously, these two equations are satisfied simultaneously provided either $\lambda = 1$ and $z = 0$ or $\lambda \neq 1$ and $x = y = z/(1-\lambda)$.

In case $z = 0$, the solution of the equations $x^2 + y^2 - 1 = 0$ and $x + y + z - 1 = 0$ yields the two points $(1, 0, 0)$ and $(0, 1, 0)$.

In case $x = y$, the solution of the equations $x^2 + y^2 - 1 = 0$ and $x + y + z - 1 = 0$ gives the two points

$$P_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1 - \sqrt{2}\right), \quad P_2 = \left(\frac{-\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 1 + \sqrt{2}\right).$$

The distance of the points $(1, 0, 0)$ and $(0, 1, 0)$ from the origin is 1 unit while the distances of P_1 and P_2 from the origin are $4 - 2\sqrt{2}$ and $4 + 2\sqrt{2}$ units respectively. So the points on the ellipse closest to the origin are $(1, 0, 0)$ and $(0, 1, 0)$, and the point farthest from the origin is P_2.