Unit 3: Partial Derivatives

Dr.P.M.Bajracharya

July 13, 2024

Contents

1 Functions of several variables

- Our first step is to explain what a function of more than one variable is, starting with functions of two independent variables. This step includes identifying
  - the domain and range of such functions and
  - learning how to graph them.
- We also examine ways to relate the graphs of functions in three dimensions to graphs of more familiar planar functions.

1.1 Functions of Two Variables

A function of two variables maps ordered pairs of real numbers to a real number.

A function of two variables

Let $D \subseteq \mathbb{R}^2$. If to each ordered pair (x, y) in D there corresponds a unique real number $f(x, y)$ then f is called a function of x and y defined on D. We write

$$f : D \to \mathbb{R}.$$

The set D is called the **domain** of the function.
The **range** of f is the set of all real numbers z that has at least one ordered pair $(x, y) \in D$ such that $f(x, y) = z$ as shown in the following figure.

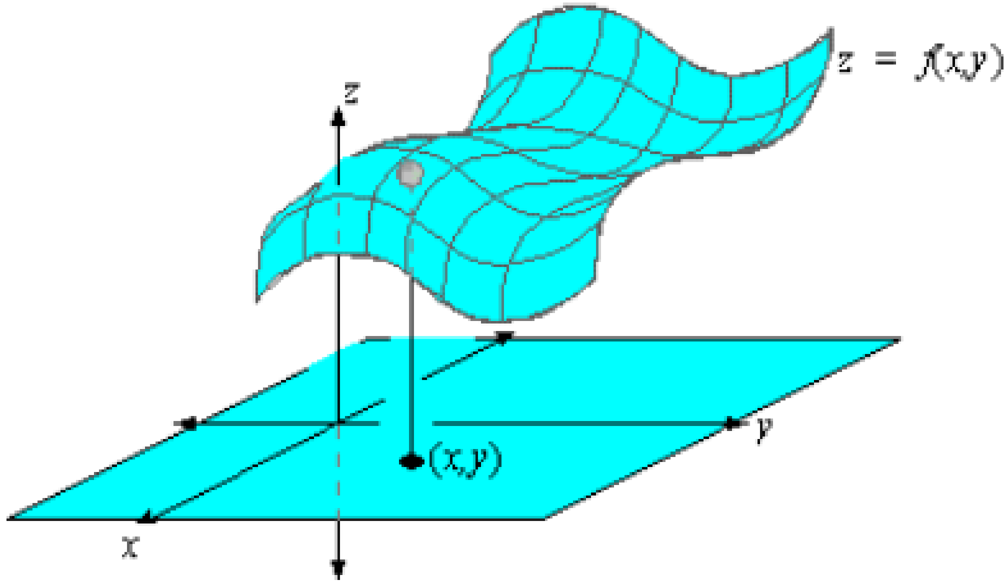


Figure 1: The domain of a function of two variables consists of ordered pairs (x, y).

1.1.1 Determining the domain and the range of a function of two variables

As with functions of one variable, the most common way to describe a function of two variables is with an equation, and unless it is otherwise restricted, you can assume that

Domain

The domain is the set of all points for which the equation is defined. In other words,

$$\text{Domain} = \{(x, y) \in \mathbb{R}^2 |\ f(x) \in \mathbb{R}\}.$$

For instance,

Example 1. 1. Consider a function:

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

What is the domain of the function f?

2. Consider a function:

$$g(x, y) = \ln xy.$$

What is the domain of the function g?

Ans.: 1. The entire -plane. 2. The set of all points in the first and third quadrants.

Example 2 (Domains)**.** Find the domain of each of the following functions

(a) $f(x, y) = 3x + 5y + 2$

(b) $g(x, y) = \sqrt{9 - x^2 - y^2}$

(c) $h(x, y) = \dfrac{\sqrt{x^2 + y^2 - 9}}{x}$

Solution.

(a) The function f is defined for all points $(x, y) \in \mathbb{R}^2$. Hence the domain of the function f is $\mathbb{R}^2$.

(b) The function g is defined for all points (x, y) and

$$x^2 + y^2 \leq 9.$$

So, the domain is the set of all points lying on or inside the circle $x^2 + y^2 = 9$. That is,

$$\text{Domain} = \{(x, y) \in \mathbb{R}^2 |\ x^2 + y^2 \leq 9\}.$$

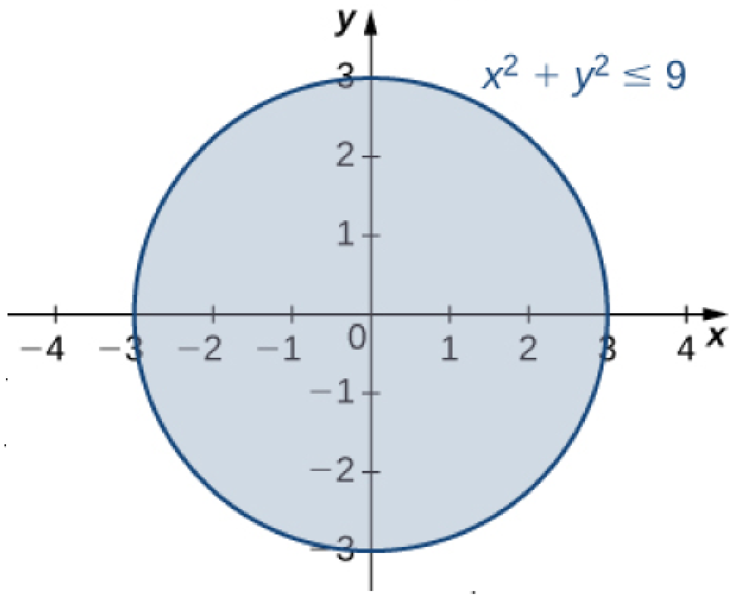


Figure 2: The domain of the function
$g(x, y) = \sqrt{9 - x^2 - y^2}$.

(c) The function h is defined for all points (x, y) such that $x \neq 0$ and

$$x^2 + y^2 \geq 9.$$

So, the domain is the set of all points lying on or outside the circle $x^2+y^2 = 9$ except those points on the y-axis. That is,

$$\text{Domain} = \{(x, y) \in \mathbb{R}^2 |\ x^2 + y^2 \geq 9, x \neq 0\}.$$

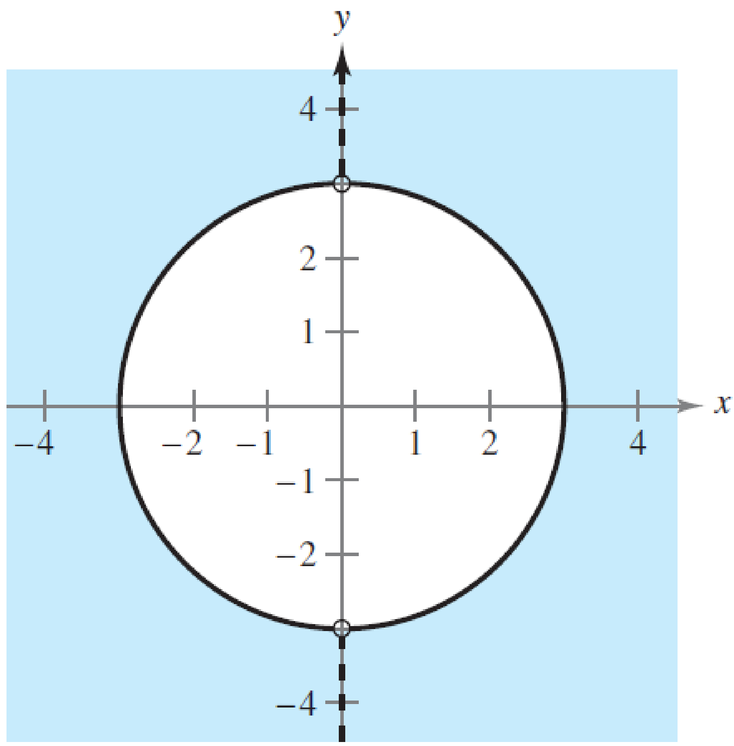


Figure 3: The domain of the function $h(x, y) = \frac{\sqrt{x^2+y^2-9}}{x}$.

1.1.2 Graph of a Function of Two Variables

As with functions of a single variable, you can learn a lot about the behavior of a function of two variables by sketching its graph.

The graph of a function of two variables is the set of all points (x, y, z) for which $z = f(x, y)$ and (x, y) is in the domain of f. This graph can be interpreted geometrically as a surface in space.

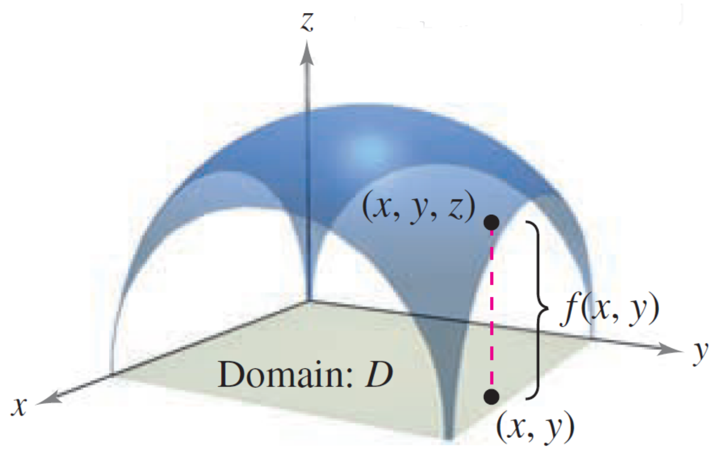


Figure 4: The surface $(x, y, f(x, y))$.

In the above Figure, note that the graph of is a surface whose projection onto the xy-plane is the domain of f. To each point (x, y) in D there corresponds a point (x, y, z) on the surface, and, conversely, to each point (x, y, z) on the surface there corresponds a point (x, y) in D.

Example 3 (Range)**.** Find the range of each of the following functions

(a) $f(x, y) = 3x + 5y + 2$

(b) $f(x, y) = \sqrt{9 - x^2 - y^2}$

(c) $f(x, y) = \sqrt{16 - 4x^2 - y^2}$.

Describe the graphs of f in each case.

Solution.

(a) Domain $= \mathbb{R}^2$.

Range $= \mathbb{R}$.

(b) The domain D of f is the set of all points (x, y) such that

$$9 - x^2 - y^2 \geq 0.$$

So, D is the set of all points lying on or inside the circle

$$x^2 + y^2 = 9.$$

<u>Range of f</u>

We have

$$z = \sqrt{9 - x^2 - y^2} = \sqrt{9 - (x^2 + y^2)}.$$

Since $x^2 + y^2 \geq 0$, we obtain

$$0 \leq z \leq \sqrt{9} = 3.$$

This is the range of f.

<u>Graph of f</u>

A point (x, y, z) is on the graph of f if and only if

$$\begin{aligned} z &= \sqrt{9 - x^2 - y^2} \\ i.e. \quad z^2 &= 9 - x^2 - y^2 \\ i.e. \quad x^2 + y^2 + z^2 &= 0, \end{aligned}$$

which is the equation of a sphere with centre at the origin and of radius 3. Also, we know that $0 \leq z \leq 4$. Hence the graph of f is the upper half of the sphere, as shown in the figure given below.

(c) The domain D implied by the equation of f is the set of all points (x, y) such that

$$16 - 4x^2 - y^2 \geq 0.$$

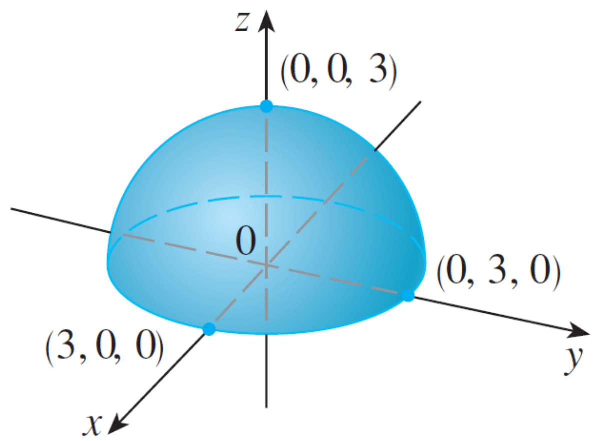


Figure 5: The graph of $f(x,y) = \sqrt{9 - x^2 - y^2}$ is the upper half of a sphere.

So, D is the set of all points lying on or inside the ellipse
$$\frac{x^2}{4} + \frac{y^2}{16} = 1.$$

Range of f

We have
$$z = \sqrt{16 - 4x^2 - y^2} = \sqrt{16 - (4x^2 + y^2)}.$$

Since $4x^2 + y^2 \geq 0$, we obtain
$$0 \leq z \leq \sqrt{16} = 4.$$

This is the range of f.

Graph of f

A point (x, y, z) is on the graph of f if and only if
$$\begin{aligned} z &= \sqrt{16 - 4x^2 - y^2} \\ i.e. \quad z^2 &= 16 - 4x^2 - y^2 \\ i.e. \quad \frac{x^2}{4} + \frac{y^2}{16} + \frac{z^2}{16} &= 1, \end{aligned}$$

where $0 \leq z \leq 4$. Hence the graph of f is the upper half of an ellipsoid, as shown in the figure given below.

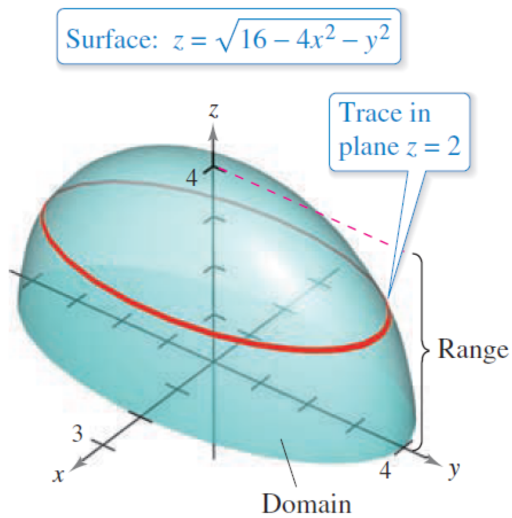


Figure 6: The graph of $f(x,y) = \sqrt{16 - 4x^2 - y^2}$ is the upper half of an ellipsoid.

Problem 1. Find the domain and range of the function $f(x,y) = \sqrt{36 - 9x^2 - 9y^2}$. Describe the graphs of f.

Problem 2. Sketch the graph of $f(x,y) = x^2 + y^2$.

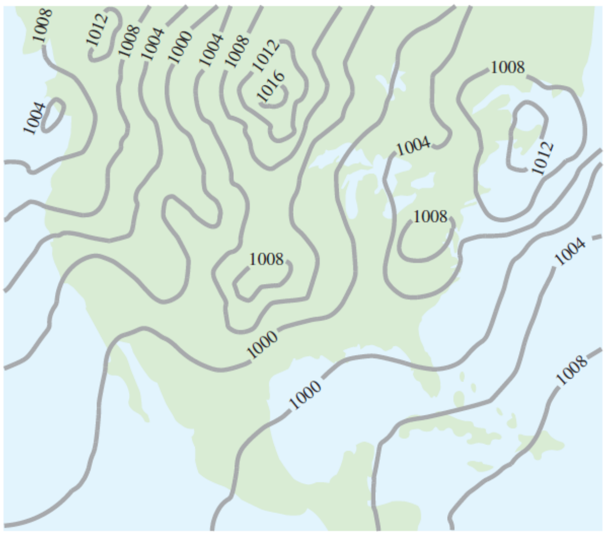
Problem 3. A profit function for a hardware manufacturer is given by $f(x,y) = 16 - (x-3)^2 - (y-2)^2$, where x is the number of nuts sold per month (measured in thousands) and y represents the number of bolts sold per month (measured in thousands). Profit is measured in thousands of dollars. Sketch a graph of this function.

1.1.3 Contour maps

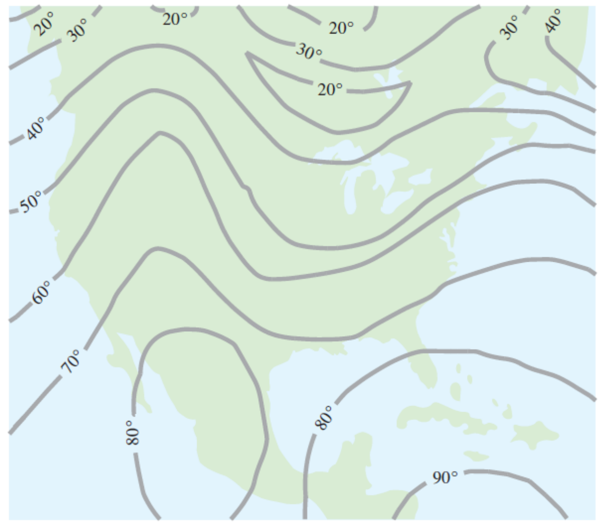
Another method for visualizing functions, borrowed from mapmakers, is a contour map on which points of constant elevation are joined to form *contour lines*, or *level curves*.

Contour map

The **level curves** of a function of two variables are the curves with equations $f(x,y) = c$, where c is a constant (in the range of f). A graph of the various level curves of a function is called a **contour map**.



Level curves show the lines of equal pressure (isobars), measured in millibars.



Level curves show the lines of equal temperature (isotherms), measured in degrees Fahrenheit.

Example 4. Create a contour map for the surface

$$f(x,y) = \sqrt{64 - x^2 - y^2}$$

corresponding to $c = 0, 1, 2, ..., 8$.

Solution.

For $c_1 = 0$, we have

$$\sqrt{64 - x^2 - y^2} = 0$$
$$\implies x^2 + y^2 = 64.$$

Hence the level curve for f using $c_1 = 0$ is given by

$$x^2 + y^2 = 8^2.$$

For $c_2 = 1$, we have

$$\begin{aligned} \sqrt{64 - x^2 - y^2} &= 1 \\ \Longrightarrow x^2 + y^2 &= 63. \end{aligned}$$

Hence the level curve for f using $c_2 = 1$ is given by

$$x^2 + y^2 = (\sqrt{63})^2.$$

For $c_3 = 2$, we have

$$\begin{aligned} \sqrt{64 - x^2 - y^2} &= 2 \\ \Longrightarrow x^2 + y^2 &= 62. \end{aligned}$$

Hence the level curve for f using $c_3 = 2$ is given by

$$x^2 + y^2 = (\sqrt{62})^2.$$

Similarly, we find level curves for f using $c_4 = 3, c_5 = 4, c_6 = 5, c_7 = 6, c_8 = 7$.

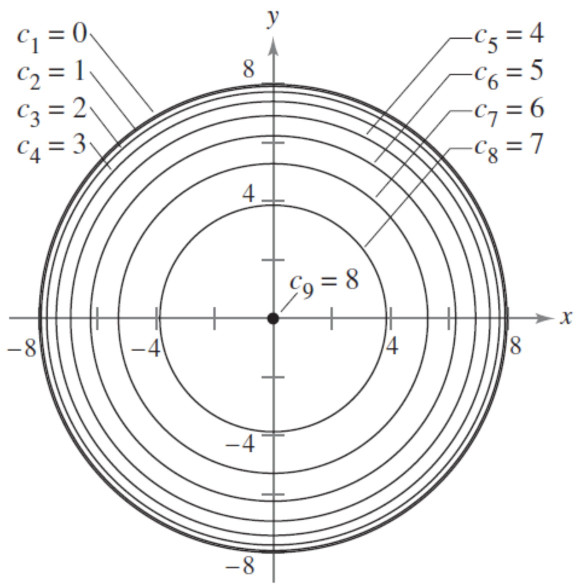
For $c_9 = 8$, we have

$$\begin{aligned} \sqrt{64 - x^2 - y^2} &= 8 \\ \Longrightarrow x^2 + y^2 &= 0 \\ \Longrightarrow x = 0,\ y &= 0. \end{aligned}$$

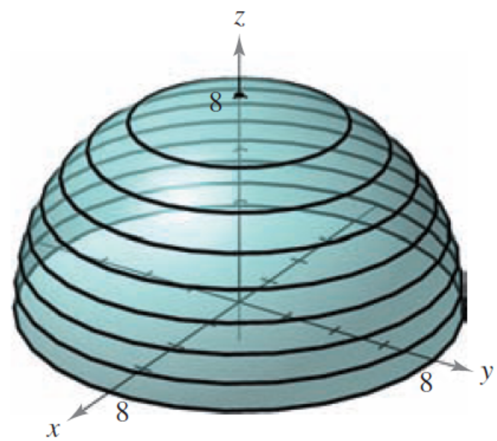
Hence the level curve for f using $c_9 = 8$ is given by

$$x = 0,\ y = 0.$$

It is a degenerate circle.



Contour map of the function $f(x, y) = \sqrt{64 - x^2 - y^2}$, using $c = 0, 1, 2, ..., 8$.



A Hemisphere with level curves.

Note that in the previous derivation it may be possible that we introduced extra solutions by squaring both sides. This is not the case here because the range of the square root function is nonnegative.

Problem 4. Create a contour map for the surface

$$f(x, y) = \sqrt{9 - x^2 - y^2}.$$

corresponding to $c = 0, 1, 2, 3$.

Problem 5. Given the function

$$f(x, y) = \sqrt{8 + 8x - 4y - 4x^2 - y^2},$$

find the level curve corresponding to $k = 0$. Then create a contour map for this function. What are the domain and range of f ?

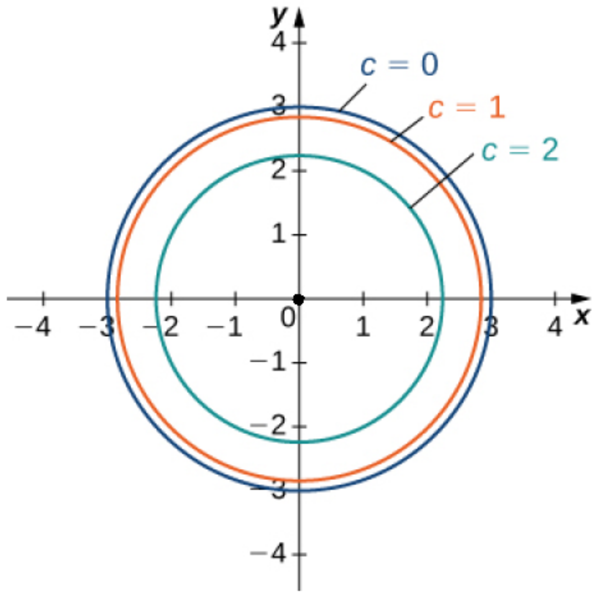


Figure 7: Level curves of the function $g(x,y) = \sqrt{9-x^2-y^2}$, using $c = 0, 1, 2, 3$.

1.2 Functions of Three Variables

Let us take a brief look at functions of three variables.

A function of three variables

A **function of three variables**, f, is a rule that assigns to each ordered triple (x, y, z) in a domain $D \subseteq \mathbb{R}^3$ a unique real number denoted by $f(x, y, z)$.

For instance, the temperature T at a point on the surface of the earth depends on the longitude x and latitude y of the point and on the time t, so we could write $T = f(x, y, t)$.

Example 5. Find the domain of a function f given by

$$f(x,y,z) = \frac{x}{\sqrt{9-x^2-y^2-z^2}}$$

Solution. The function f is defined for all points (x, y, z) such that

$$9 - x^2 - y^2 - z^2 > 0, i.e., x^2 + y^2 + z^2 < 9.$$

Consequently, the domain is the set of all points (x, y, z) lying inside a sphere of radius 3 with center at the origin.

Example 6. Find the domain of a function f given by

$$f(x, y, z) = \ln(z - y) + xy \sin z$$

Solution. $D = \{(x, y, z) \in \mathbb{R}^3 : \ z > y\}$.

This is a half-space consisting of all points that lie above the plane $z = y$.

Example 7. Find the domain of a function f given by

$$f(x, y, t) = \frac{\sqrt{2t - 4}}{x^2 - y^2}.$$

Solution. $D = \{(x, y, t) \in \mathbb{R}^3 : \ y \neq \pm x, t \geq 2\}$.

Problem 6. Find the domain of each of the following functions:

(a) $f(x, y, t) = (3t - 6)\sqrt{y - 4x^2 + 4}$

(b) $g(x, y, z) = \frac{3x - 4y + 2z}{\sqrt{9 - x^2 - y^2 - z^2}}$.

It's very difficult to visualize a function f of three variables by its graph, since that would lie in a four-dimensional space. However, we do gain some insight into f by examining its **level**

surfaces, which are the surfaces with equations $f(x, y, z) = k$, where k is a constant. If the point (x, y, z) moves along a level surface, the value of $f(x, y, z)$ remains fixed.

Example 8. Find the level surfaces of the function

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

Solution. The level surfaces form a family of concentric spheres with radius $\sqrt{k}$.

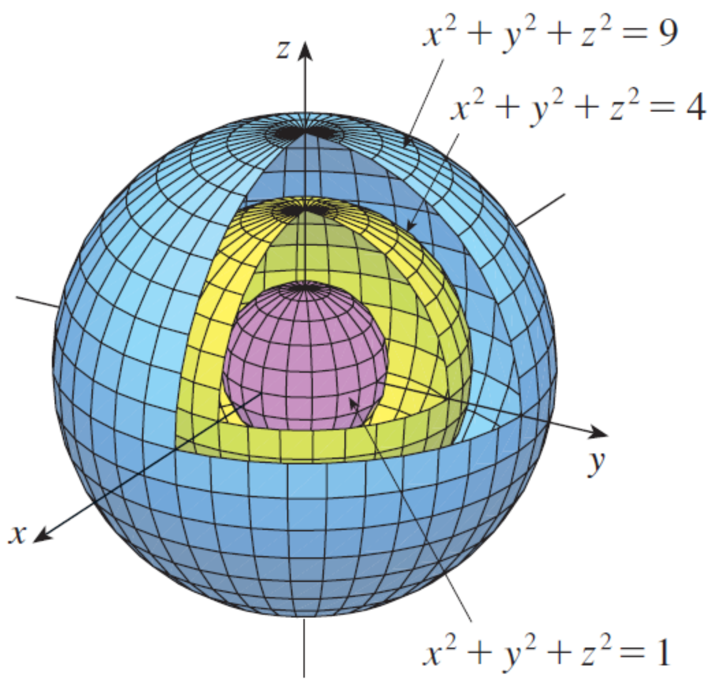


Figure 8: Level surfaces of a sphere.

Problem 1. Find the level surface for the function and describe the surface, if possible.

(a) $f(x,y,z) = 4x^2 + 9y^2 - z^2$ corresponding to $k = 1$.

(b) $f(x,y,z) = x^2 + y^2 + z^2 - 2x + 4y - 6z$ corresponding to $k = 2$.

1.3 Functions of n Variables

Functions of any number of variables can be considered. A function of n variables is a rule that assigns a number $z = f(x_1, x_2, ..., x_n)$ to a n-tuple $(x_1, x_2, ..., x_n)$ of real numbers. In other words,

$$f : \mathbb{R}^n \to \mathbb{R}.$$

For example, if a company uses n different ingredients in making a food product, c_i is the cost per unit of the ith ingredient, and x_i units of the ingredient are used, then the total cost C of the ingredients is a function of the n variables $x_1, x_2, ..., x_n$:

$$C = f(x_1, x_2, ..., x_n) = c_1x_1 + c_2x_2 + \cdot + c_nx_n. \tag{1}$$

Sometimes we will use vector notation to write such functions more compactly: If $x = (x_1, x_2, ..., x_n)$, we often write $f(x)$ in place of $f(x_1, x_2, ..., x_n)$. With this notation we can rewrite the function defined in Equation (1) as

$$f(x) = c \cdot x,$$

where $c = (c_1, c_2, ..., c_n)$ and $c \cdot x$ denotes the dot product of the vectors c and x in V_n.

In view of the one-to-one correspondence between points $(x_1, x_2, ..., x_n)$ in $\mathbb{R}^n$ and their position vectors $x = (x_1, x_2, ..., x_n)$ in $\mathbb{R}^n$, we have three ways of looking at a function f defined on a subset of $\mathbb{R}^n$:

1. As a function of n real variables $x_1, x_2, ..., x_n$.
2. As a function of a single point variable $(x_1, x_2, ..., x_n)$
3. As a function of a single vector variable $x = (x_1, x_2, ..., x_n)$

We will see that all three points of view are useful.

Functions of several variables can be combined in the same ways as functions of single variables. For instance, you can form the sum, difference, product, and quotient of two functions of two variables as follows.

Algebraic operations

$$\begin{aligned}(f \pm g)(x, y) &= f(x, y) \pm g(x, y)\\ (fg)(x, y) &= f(x, y)g(x, y)\\ \frac{f}{g}(x, y) &= \frac{f(x, y)}{g(x, y)}\end{aligned}$$

We cannot form the composite of two functions of several variables. We can, however, form the composite function $(g \circ f)(x, y)$ where g is a function of a single variable and f is a function of two variables.

The composite function $g \circ f$

$$\boxed{(g \circ f)(x, y) := g(f(x, y)).}$$

The domain of this composite function consists of all (x, y) in the domain of f such that $f(x, y)$ is in the domain of g. That is,

$$\boxed{\text{Domain}_{g\circ f} = \{(x, y) \in \text{Domain}_f : \; f(x, y) \in \text{Domain}_g\}}$$

Example 9. A function $h(x, y) = \sqrt{16 - x^2 - 4y^2}$. It can be viewed as the composite of the function f of two variables given by

$$f(x, y) = 16 - x^2 - 4y^2$$

and the function g of a single variable given by

$$g(u) = \sqrt{u}.$$

Thus,

$$h(x, y) = g(f(x, y)) = (g \circ f)(x, y).$$

That is,

$$h = g \circ f.$$

The domain of this function is the set of all points lying on or inside the ellipse $x^2 + 4y^2 = 16$.

2 Limits and continuity

Limits

Let $u = (x, y) \in \mathbb{R}^2$. Then we write

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

As you know, this is the Euclidean norm of u.

Let $D \subseteq \mathbb{R}^2$ and $p = (a, b) \in \mathbb{R}^2$. The point p is called a limitpoint or accumulation point of D if D includes points $u = (x, y)$ arbitrarily close to p, i.e.,

$$\forall\, r > 0\ \exists\, u \in D:\ 0 < \|p - u\| < r.$$

Limit at a point

Let f be a real valued function defined on $D \subseteq \mathbb{R}^2$ and $p = (a, b) \in \mathbb{R}^2$ be a limitpoint of D. Then we say that L is the limit of $f(u)$ as $u = (x, y) \in D$ tends to p if for every number $\varepsilon > 0$ there is a number $\delta > 0$ such that

$$\forall\, u = (x, y) \in D,\ 0 < \|u - p\| < \delta \Rightarrow |f(x, y) - L| < \varepsilon.$$

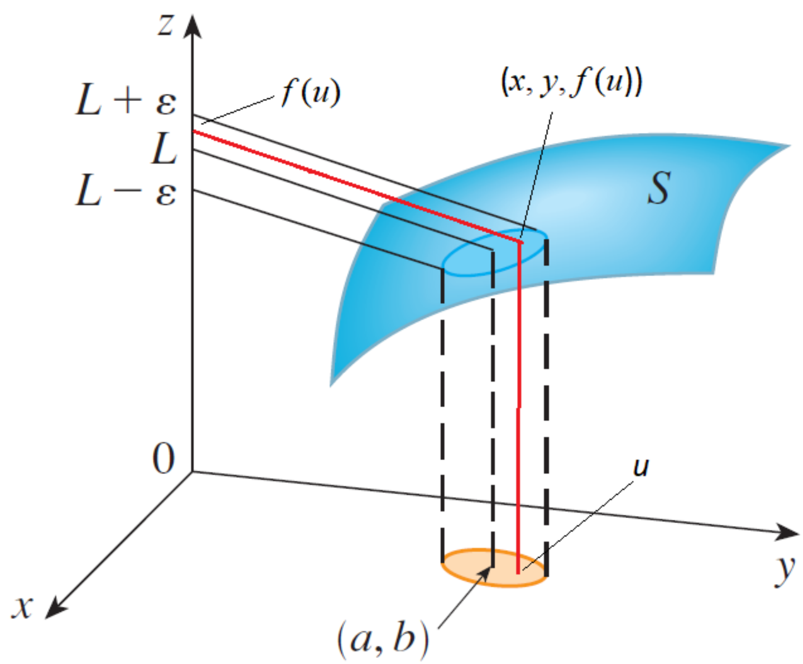


Figure 9: Limit at a point u.

In this case, we write

$$\lim_{u \to p} f(u) = L.$$

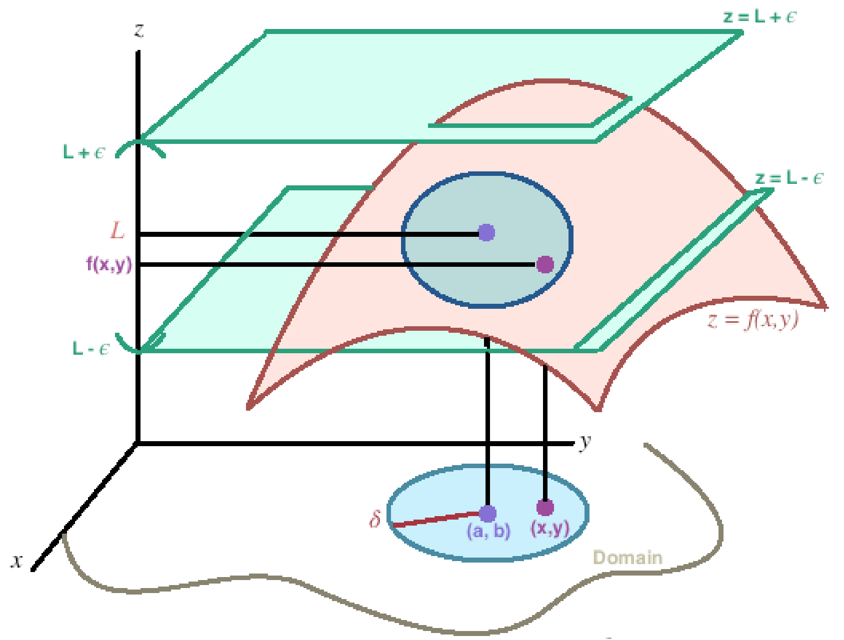


Figure 10: Limit at a point u.

Other notations for the limit in the above definition are

$$\lim_{(x,y)\to(a,b)} f(x,y) = L \quad \text{and} \quad f(x,y) \to L \text{ as } (x,y) \to (a,b).$$

The above definition says that

Intuitively

The distance between $f(u)$ and L can be made arbitrarily small by making the distance from u to p sufficiently small (but not 0).

Verifying a Limit by the Definition

Example 10. Show that

$$\lim_{(x,y)\to(a,b)} f(x,y) = a.$$

Solution. Let $f(x,y) = x$ and $L = a$. We need to show that for each $\varepsilon > 0$ if $(x,y) \neq (a,b)$ and

$$\|(x,y) - (a,b)\| < \delta, i.e., \sqrt{(x-a)^2 + (y-b)^2} < \delta,$$

then

$$|f(x,y) - L| < \varepsilon.$$

Now, if $(x,y) \neq (a,b)$ and

$$\|(x,y) - (a,b)\| < \delta,$$

then we have

$$\begin{aligned} |f(x,y) - L| &= |x-a| = \sqrt{(x-a)^2} \\ &\leq \sqrt{(x-a)^2 + (y-b)^2} \\ &= \|(x,y) - (a,b)\| \\ &< \delta. \end{aligned}$$

So, choosing $\delta = \varepsilon$, we obtain

$$|f(x,y) - L| < \varepsilon. \qquad \blacktriangleleft$$

Limit laws:

Limits of functions of several variables have the same properties regarding sums, differences, products, and quotients as do limits of functions of single variables.

Let $u = (x,y)$ and $p = (a,b)$. If $\lim_{u \to p} f(u) = L$, $\lim_{u \to p} g(u) = M$, then the following sum, product, and quotient rules, and squeeze theorem hold.

Limit laws

(a) $\lim_{u\to p}(f(u)+g(u)) = L+M$

(b) $\lim_{u\to p}(f(u)g(u)) = LM$

(c) $\lim_{u\to p}\dfrac{f(u)}{g(u)} = \dfrac{L}{M}$ $(M\neq 0)$.

Some of these properties are used in the next example.

Example 11. Evaluate

$$\lim_{(x,y)\to(1,2)} \frac{3x^2y}{x^2+y^2}.$$

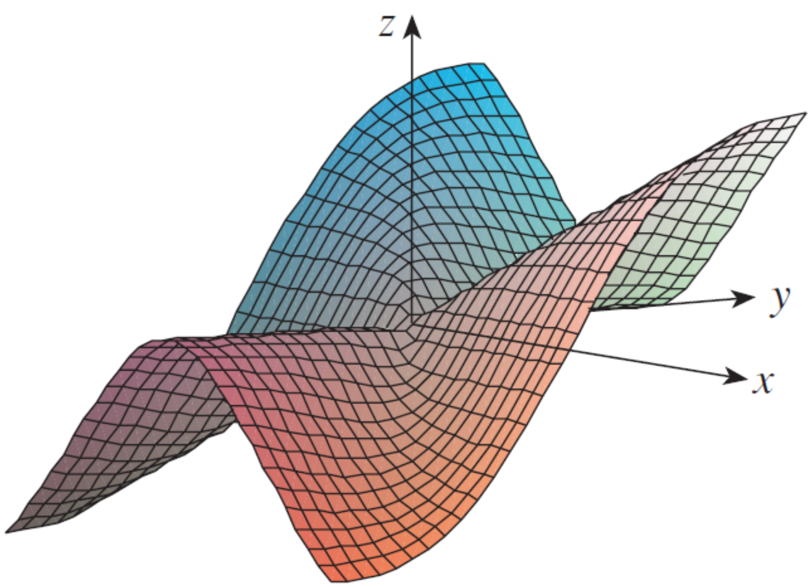


Figure 11: $\dfrac{3x^2y}{x^2+y^2}$

.

Solution.

We observe that

$$\lim_{(x,y)\to(1,2)} 3x^2y = 5(1)^2(2) = 6,$$
$$\lim_{(x,y)\to(1,2)} (x^2+y^2) = 1^2+2^2 = 5.$$

Because the limit of a quotient is equal to the quotient of the limits (and the denominator is not 0), we have

$$\lim_{(x,y)\to(1,2)} \frac{3x^2y}{x^2+y^2} = \frac{6}{5}. \qquad \blacktriangleleft$$

Example 12. Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2}$, if exists.

Solution. In this case, the limits of the numerator and of the denominator are both 0, and so we cannot evaluate the limit by using operations on limits as in the previous example. However, it seems reasonable that the limit might be 0, because if $y = ax$, then

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = \lim_{(x,y)\to(0,0)} \frac{3ax^3}{x^2(1+a^2)} = 0.$$

So, we try applying the definition to $L = 0$.

For it we observe that

$$\frac{x^2}{x^2+y^2} \leq 1, \qquad y \leq |y| \leq \sqrt{x^2+y^2}.$$

Now, suppose that $\|(x,y)-(0,0)\| < \delta$. Then

$$\sqrt{x^2+y^2} = \|(x,y)\| < \delta.$$

Now, put $f(x,y) = \frac{3x^2y}{x^2+y^2}$. For $(x,y) \neq (0,0)$ we have

$$
\begin{aligned}
|f(x,y) - 0| &= \left|\frac{3x^2y}{x^2+y^2}\right| = 3|y|\frac{x^2}{x^2+y^2} \\
&\leq 3|y| = 3\sqrt{y^2} \\
&\leq 3\sqrt{x^2+y^2} \\
&< 3\delta \\
&= \varepsilon. \qquad (\text{choosing } \delta = \varepsilon/3)
\end{aligned}
$$

Therefore,

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = 0. \qquad \blacktriangleleft$$

Squeeze theorem

It makes easier sometimes to find the limit applying the following theorem, in the case when it is not possible to find the limit by using the above operations.

Squeeze theorem

Let

$$\lim_{u\to p} f(u) = L, \quad \lim_{u\to p} g(u) = M.$$

If

$$\lim_{u\to p} f(u) = \lim_{u\to p} g(u) \text{ and } f(u) \leq h(u) \leq g(u),$$

then $\lim_{u\to p} h(u)$ exists and equals L which equals M.

Example 13. Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2}$, if exists.

Solution. Let $y = ax$. Then

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = \lim_{(x,y)\to(0,0)} \frac{3ax^3}{x^2(1+a^2)} = 0.$$

This shows that the limit along any line through the origin is 0. Although this doesn't prove that the given limit is 0, we begin to suspect that the limit does exist and is equal to 0. we prove it.

Since $y^2 \geq 0$, we have

$$x^2 \leq x^2 + y^2 \Rightarrow \frac{x^2}{x^2+y^2} \leq 1$$

$$\Rightarrow \quad 0 \leq \frac{3x^2|y|}{x^2+y^2} \leq 3\sqrt{y^2}$$

$$\Rightarrow \quad \lim_{(x,y)\to(0,0)} 0 \leq \lim_{(x,y)\to(0,0)} \frac{3x^2|y|}{x^2+y^2} \leq \lim_{(x,y)\to(0,0)} 3\sqrt{y^2}$$

However, we have

$$\lim_{(x,y)\to(0,0)} 0 = 0, \qquad \lim_{(x,y)\to(0,0)} 3\sqrt{y^2} = 0.$$

Therefore, by the Squeeze Theorem,

$$\lim_{(x,y)\to(0,0)} \frac{3x^2|y|}{x^2+y^2} = 0.$$

Using the property that $-|c| \leq c \leq |c|$ for any real number c, we obtain

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = 0. \qquad \blacktriangleleft$$

The definition refers only to the distance between u and p. It does not refer to the direction of approach. That means, if the limit exists, then $f(u)$ must approach the same limit no matter how u approaches p.

For some functions, it is easy to recognize that a limit does not exist. For instance, it is clear that the limit

$$\lim_{(x,y)\to(0,0)} \frac{1}{x^2+y^2}$$

does not exist because the values of $f(x, y)$ increase without bound as (x, y) approaches along any path (see the figure given below.

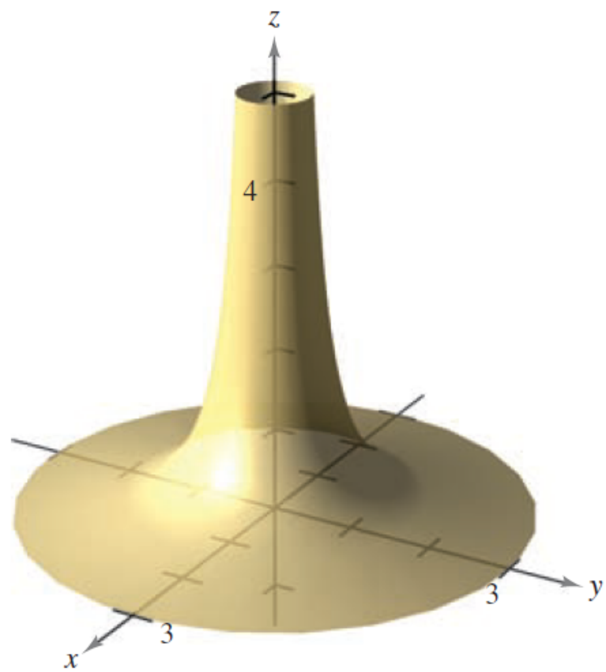


Figure 12: $\lim_{(x,y)\to(0,0)} \frac{1}{x^2+y^2}$ does not exist.

For other functions, it is not so easy to recognize that a limit does not exist. However, in many cases, the following criterion is very helpful.

Nonexistence criterion for limits

Let $f(u) \to L_1$ as $u \to p$ along a path C_1 and $f(u) \to L_2$ as $u \to p$ along a path C_2. If $L_1 \neq L_2$, then the limit $\lim_{u \to p} f(u)$ does not exist.

Thus, by this criterion, if we can find two different paths of approach along which the function $f(u)$ has different limits, then it follows that $\lim_{u \to p} f(u)$ does not exist.

Example 14. Show that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2}$ does not exist.

Solution. Let's approach $(0, 0)$ first along the x-axis and then along the y-axis.

Example 15. If $f(x, y) = \dfrac{xy}{x^2 + y^2}$, does the limit $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y)$ exist?

Solution. Let's approach $(0, 0)$ first along the x-axis and along the y-axis. Then approach $(0, 0)$ along another line, say, $x = y$ for all $x \neq 0$. No!

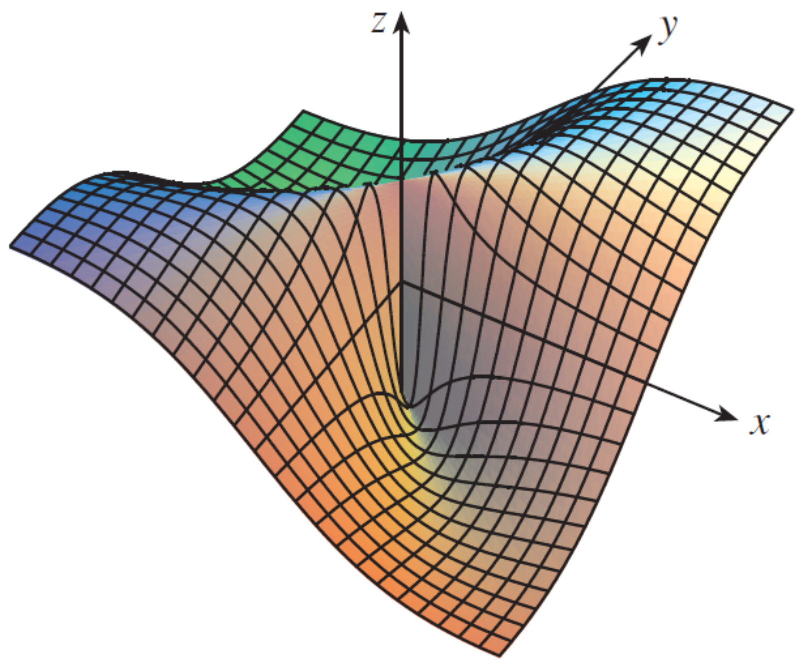


Figure 13: $f(x,y) = \dfrac{xy}{x^2+y^2}$

Example 16. If $f(x,y) = \dfrac{xy^2}{x^2+y^4}$, does the limit $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ exist?

Solution. Let's approach $(0,0)$ along any nonvertical line $y = mx$ through the origin. But if we approach $(0,0)$ Then approach $(0,0)$ along along the parabola $x = y^2$, then $f(x,y) = 1/2$. No!

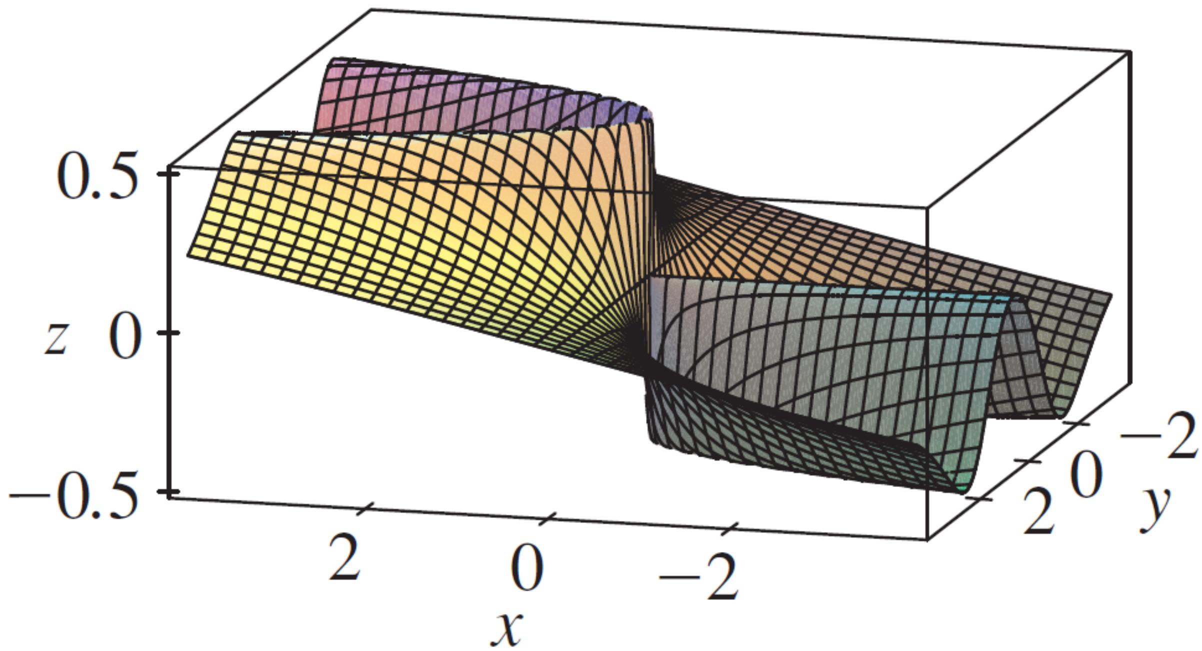


Figure 14: $f(x,y) = \dfrac{xy^2}{x^2+y^4}$

Continuity

Continuity at a point

A function f of two variables is called **continuous at a point** (a,b) in a set $D \subseteq \mathbb{R}^2$ if the following conditions are satisfied:

1. $f(a,b)$ exists.
2. $\displaystyle\lim_{(x,y)\to(a,b)} f(x,y)$ exists.
3. $\displaystyle\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$.

We say f is **continuous on a set** $D \subseteq \mathbb{R}^2$ if f is continuous at every point (a,b) in D.

The intuitive meaning of continuity is that if the point (x,y) changes by a small amount, then the value of $f(x,y)$ changes by a small amount. This means that a surface that is the graph of a continuous function has no hole or break.

Using the properties of limits, you can see that sums, differences, products, and quotients of continuous functions are continuous on their domains. Moreover, the following property related to a composition of two functions also holds:

Interchange of the limit and function

Let $\lim_{(x,y)\to(a,b)} f(x,y) = L$. If $F(t)$ is a continuous function at $t = L$, then

$$\lim_{(x,y)\to(a,b)} F(f(x,y)) = F(L) = F(\lim_{(x,y)\to(a,b)} f(x,y)).$$

That is, for continuous functions, we may interchange the limit and function composition operations.

Let's use these facts to give examples of continuous functions.

It is easy to show that

$$\lim_{(x,y)\to(a,b)} x = a, \qquad \lim_{(x,y)\to(a,b)} y = b, \qquad \lim_{(x,y)\to(a,b)} c = c.$$

These limits show that the functions $f(x,y) = x$, $g(x,y) = y$, and $h(x,y) = c$ are continuous everywhere on $\mathbb{R}^2$.

A **polynomial function of two variables** is a sum of terms of the form $cx^m y^n$, where c is a constant and m and n are nonnegative integers. It follows that all polynomials are continuous on $\mathbb{R}^2$.

Likewise, any rational function $f(x,y) = \dfrac{P(x,y)}{Q(x,y)}$, where $P(x,y), Q(x,y)$ are polynomials and $Q(x,y) \neq 0$ is continuous on its domain because it is a quotient of continuous functions

$P(x, y)$ and $Q(x, y)$.

Example 17. Evaluate $\lim_{(x,y)\to(a,b)} (x^2y^3 - x^3y^2 + 3x + 2y)$.

Solution. ...

Example 18. Discuss the continuity of the function

$$f(x, y) = \frac{3x^2y}{x^2 + y^2}.$$

Solution. ...

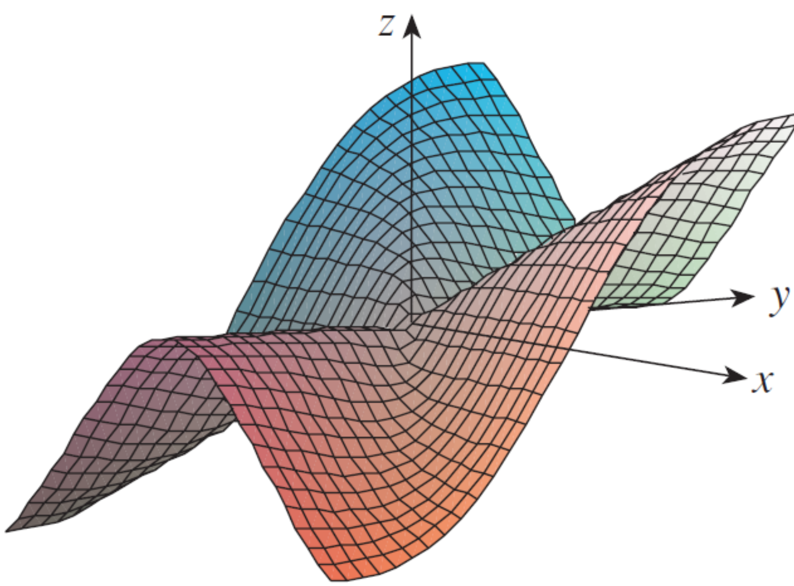


Figure 15: $f(x, y) = \dfrac{3x^2y}{x^2 + y^2}$.

Example 19. Discuss the continuity of the function

$$f(x,y) = \begin{cases} \dfrac{3x^2y}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Solution. ...

Example 20. If $f(x,y) = \dfrac{xy}{x^2+y^2}$, does the limit $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ exist?

Solution. ...

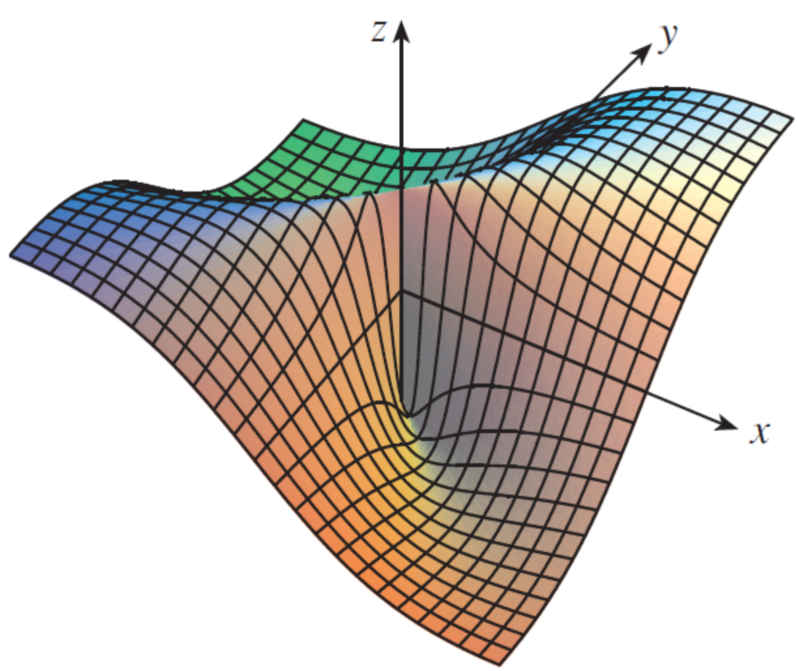


Figure 16: $f(x,y) = \dfrac{xy}{x^2+y^2}$

Example 21. Where is the function $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$ continuous?

Solution. ...

Example 22. Show that the function

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

is discontinuous at the origin.

Solution.

Example 23. Where is the function $h(x, y) = \arctan(y/x)$ continuous?

Solution. The function $f(x, y) = y/x$ is a rational function and therefore continuous except on the line $x = 0$. The function $g(t) = \arctan t$ is continuous everywhere. So the composite function

$$g(f(x, y)) = \arctan(y/x) = h(x, y)$$

is continuous except where $x = 0$. The graph in the following figure shows the break in the graph of h above the x-axis.

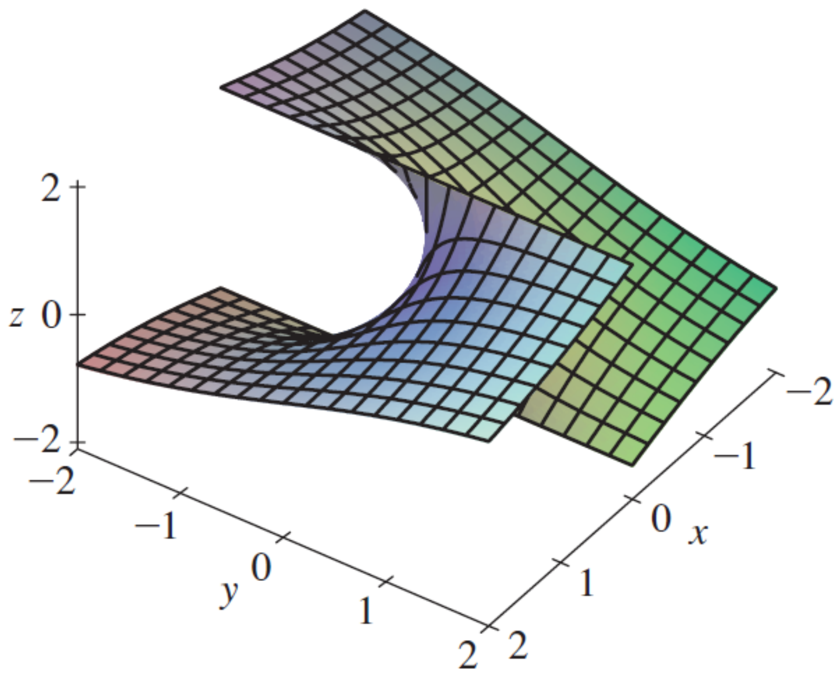


Figure 17: The function $h(x, y) = arctan(y/x)$
is discontinuous where $x = 0$.

Everything that we have done in this section can be extended to functions of three or more variables. The notation

$$\lim_{(x,y,z)\to(a,b,c)} f(x, y, z) = L$$

means that the values of $f(x, y, z)$ approach the number as the point (x, y, z) approaches the point (a, b, c) along any path in the domain of f. The function f is continuous at (a, b, c) if

$$\lim_{(x,y,z)\to(a,b,c)} f(x, y, z) = f(a, b, c)$$

For instance, the function

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

is a rational function of three variables and so is continuous at every point in $\mathbb{R}^2$ except where $x^2 + y^2 + z^2 = 1$. In other words, it is discontinuous on the sphere with center at the origin and of radius 1.

3 Partial derivatives

It is worthwhile to note that a function f of two or more variables does not have a unique rate of change because each variable may affect f in different ways.

For example, the current I in a circuit is a function of both voltage V and resistance R given by Ohm's Law:

$$I(V, R) = \frac{V}{R}.$$

The current I is increasing as a function of V but decreasing as a function of R.

The partial derivatives are the rates of change with respect to each variable separately. A function $f(x, y)$ of two variables has two partial derivatives, denoted f_x and f_y , defined by the following limits (if they exist):

Partial derivatives

$$f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}$$
$$f_y(a, b) = \lim_{k \to 0} \frac{f(a, b + k) - f(a, b)}{k}$$

Thus, f_x is the derivative of $f(x, b)$ as a function of x alone, and f_y is the derivative of $f(a, y)$ as a function of y alone. The Leibniz notation for partial derivatives is

$$\frac{\partial f}{\partial x} = D_x f = f_x, \qquad \frac{\partial f}{\partial y} = D_y f = f_y,$$

$$\left.\frac{\partial f}{\partial x}\right|_{(a,b)} = f_x(a,b), \qquad \left.\frac{\partial f}{\partial y}\right|_{(a,b)} = f_y(a,b).$$

To compute partial derivatives, all we have to do is remember that the partial derivative with respect to x is just the ordinary derivative of the function f of a single variable that we get by keeping y fixed. Thus we have the following rule.

Rules for Finding Partial Derivatives

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

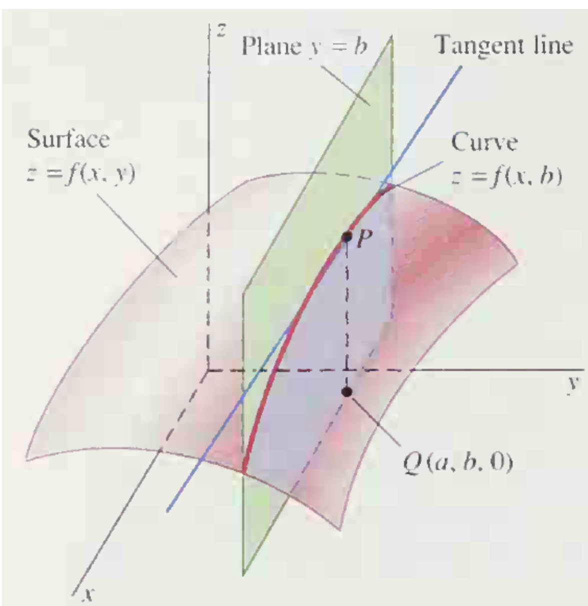
1. To find f_x, regard y as a constant and differentiate $f(x, y)$ with respect to x.
2. To find f_y, regard x as a constant and differentiate $f(x, y)$ with respect to y.

3.1 Interpretations of Partial Derivatives

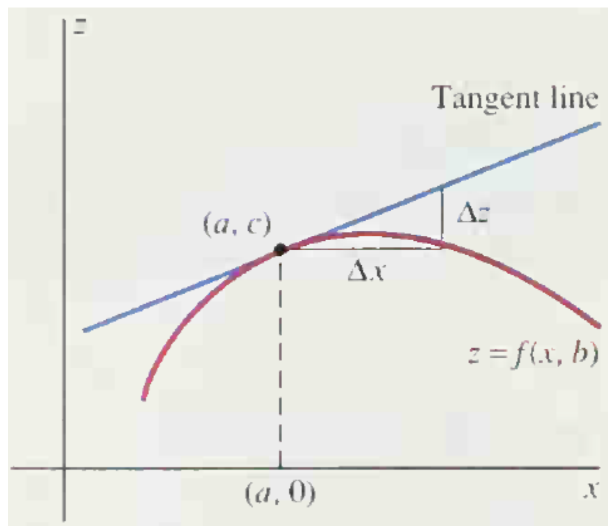
The intersection of a surface $z = f(x, y)$ with a vertical plane $y = b$ that is parallel to the xz-coordinate plane. Along the intersection curve, the x-coordinate varies but the y-coordinate is constant: $y = b$ at each point, because the curve lies in the vertical plane $y = b$.

x-curve on a surface

A curve of intersection of $z = f(x, y)$ with a vertical plane parallel to the xz-plane is called an x-curve on the surface.



(a) An x-curve and its tangent line.



(b) Projection into the xz-plane of the x-curve through $P(a, b, c)$ and its tangent line.

Figure (a) shows a point $P(a.b, c)$ in the surface $z = f(x, y)$, the x-curve through P and the line tangent to this x-curve at P. Figure (b) shows the parallel projection of the vertical plane $y = b$ onto the xz-plane itself. We can now "ignore" the presence of $y = b$ and regard $z = f(x, b)$ as a function of the single variable x. The slope of the line tangent to the original x-curve through P (see Fig. (a)) is equal to the slope of the tangent line in Fig. (b). But by familiar single-variable calculus, this latter slope is given by

$$\frac{\partial z}{\partial x} = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h} = f_x(a,b).$$

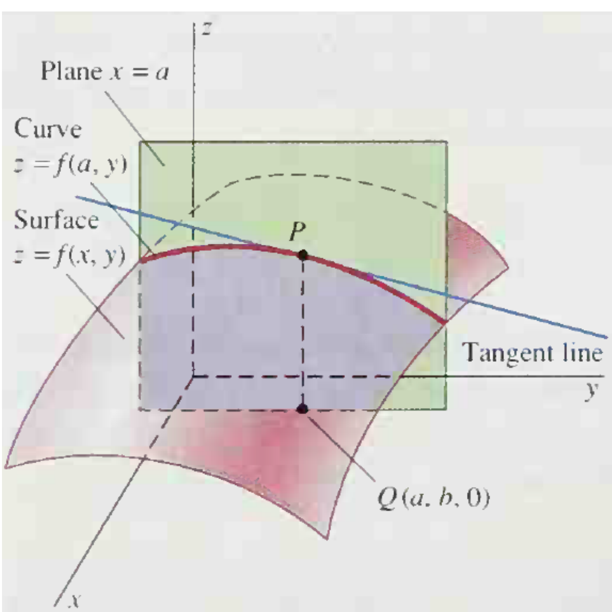
Thus, we see that the geometric meaning of f_x is this:

The value $\partial z/\partial y = f_x(a,b)$ is the slope of the line tangent at $P(a,b,c)$ to the x-curve through P on the surface $z = f(x,y)$ as shown in Figure (a).

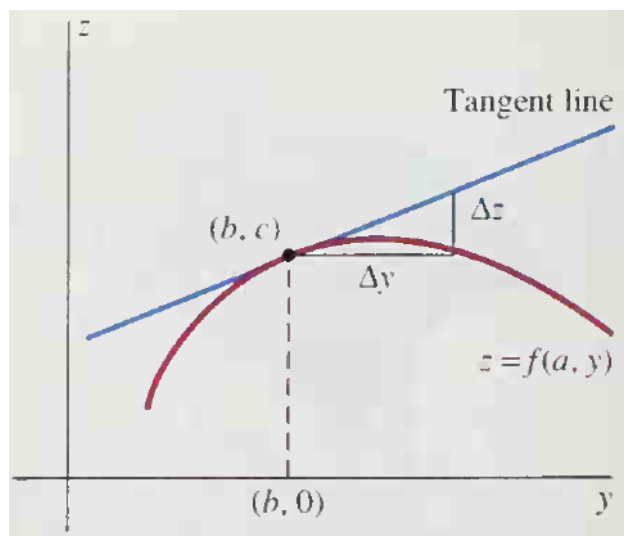
We proceed in much the same way to investigate the geometric meaning of partial derivative f_x.

y-curve on a surface

A curve of intersection of a surface $z = f(x,y)$ with a vertical plane parallel to the yz-plane is called a y-curve on the surface.



(c) A y-curve and its tangent line.



(d) Projection into the yz-plane of the y-curve through $P(a,b,c)$ and its tangent line.

Figure (c) shows a point $P(a.b,c)$ in the surface $z = f(x,y)$, the

y-curve through P and the line tangent to this y-curve at P. Figure (d) shows the parallel projection of the vertical plane $x = a$ onto the yz-plane itself. We can now "ignore" the presence of $x = a$ and regard $z = f(a, y)$ as a function of the single variable y. The slope of the line tangent to the original y-curve through P (see Fig. (c)) is equal to the slope of the tangent line in Fig. (d). But by familiar single-variable calculus, this latter slope is given by

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

Thus, we see that the geometric meaning of f_x is this:

If $x = x_0$ then $z = f(x_0, y)$ represents the curve formed by intersecting the surface $z = f(x, y)$ with the plane $x = x_0$ as shown in Figure (c).

We put the two geometric interpretations together for the comparison purpose.

Geometric interpretation

- If $y = y_0$ then $z = f(x, y_0)$ represents the curve formed by intersecting the surface $z = f(x, y)$ with the plane $y = y_0$.
- If $x = x_0$ then $z = f(x_0, y)$ represents the curve formed by intersecting the surface $z = f(x, y)$ with the plane $x = x_0$.

Partials evaluated at a point

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} \quad - \text{ the slope of the tangent line } L_1 \text{ to the curve } f(x, y_0) \text{ at } (x_0, y_0).$$

$$\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} \quad - \text{ the slope of the tangent line } L_2 \text{ to the curve } f(x_0, y) \text{ at } (x_0, y_0).$$

Informally, the values of f_x and f_y at a point denote the slopes of the surface in the x- and y-directions at the point, respectively.

Problem 7.

(a) For $f(x, y) = 9x^2y - 3x^5y$, find f_x and f_y.

(b) For $f(x, y) = xe^{x^2y}$, find f_x and f_y, and evaluate each at the point $(1, \ln 2)$.

Problem 8. Let $f(x, y) = \sqrt{3x + 2y}$.

(a) Find the slope of the surface $z = f(x, y)$ in the x-direction at the point $(4, 2)$.

(b) Find the slope of the surface $z = f(x, y)$ in the y-direction at the point $(4, 2)$.

Problem 9. Let $z = \sin(y^2 - 4x)$.

(a) Find the rate of change of z with respect to x at the point $(2, 1)$ with y held fixed.

(b) Find the rate of change of z with respect to y at the point $(2, 1)$ with x held fixed.

Example 24 (Implicit partial differentiation:)**.** Find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ if z is defined implicitly as a function of x and y by the equation

$$x^3 + y^3 + z^3 + 6xyz = 1.$$

Solution. ...

3.2 Functions of More Than Two Variables

Problem 10. Find f_x, f_y, and f_z, if $f(x, y, z) = z\ln(x^2 y\cos z)$.

3.3 Higher Derivatives

Suppose that f is a function of two variables x and y. Since the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are also functions of x and y, these functions may themselves have partial derivatives. This gives rise to four possible second-order partial derivatives of f, which are defined by

Differentiate twice with respect to x.

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = f_{xx}$$

Differentiate twice with respect to y.

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = f_{yy}$$

Differentiate first with respect to y and then with respect to x.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = f_{xy}$$

Differentiate first with respect to x and then with respect to y.

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = f_{yx}$$

The last two cases are called the mixed second-order partial derivatives or the mixed second partials. Also, the derivatives $\partial f/\partial x$ and $\partial f/\partial y$ are often called the first-order partial derivatives when it is necessary to distinguish them from higher-order partial derivatives.

Similar conventions apply to the second-order partial derivatives of a function of three variables.

Warning

Observe that the two notations for the mixed second partials have opposite conventions for the order of differentiation. In the "∂" notation the derivatives are taken right to left, and in the "subscript" notation they are taken left to right.

Example 25. Find the second partial derivatives of

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

Solution. ...

The following theorem, which was discovered by the French mathematician Alexis Clairaut (1713 - 1765), gives conditions under which we can assert that $f_{xy} = f_{yx}$.

Theorem 3.1. Suppose f is defined on a disk D that contains the point (a, b). If the functions f_{xy} and f_{yx} are both continuous on D, then
$$f_{xy}(a, b) = f_{yx}(a, b).$$

Proof.

For small values of $h \neq 0$, consider the difference

$$\Delta(h) = [f(a+h, b+h) - f(a+h, b)] - [f(a, b+h) - f(a, b)].$$

Notice that if we let $g(x) = f(x, b+h) - f(x, b)$, then
$$\Delta(h) = g(a+h) - g(a).$$
By the Mean Value Theorem, there is a number c between a and $a+h$ such that

$$g(a+h) - g(a) = g'(c)h = h[f_x(c, b+h) - f_x(c, b)].$$

Applying the Mean Value Theorem again, this time to f_x, we get a number d between b and $b+h$ such that

$$f_x(c, b+h) - f_x(c, b) = f_{xy}(c, d)h$$

Combining these equations, we obtain

$$\Delta(h) = h^2 f_{xy}(c, d)$$

If $h \to 0$, then $(c, d) \to (a, b)$, so the continuity of f_{xy} at (a, b) gives

$$\lim_{h \to 0} \frac{\Delta(h)}{h^2} = \lim_{(c,d) \to (a,b)} f_{xy}(c, d) = f_{xy}(a, b)$$

Similarly, by writing

$$\Delta(h) = [f(a+h, b+h) - f(a, b+h)] - [f(a+h, b) - f(a, b)].$$

and using the Mean Value Theorem twice and the continuity of f_{yx} at (a, b) , we obtain

$$\lim_{h \to 0} \frac{\Delta(h)}{h^2} = f_{yx}(a, b)$$

It follows that

$$f_{xy}(a, b) = f_{yx}(a, b). \qquad \blacktriangleleft$$

Problem 11. Let $f(x, y) = e^x \cos y$. Confirm that the mixed second-order partial derivatives of f are the same

Partial derivatives of order 3 or higher

Example 26. Calculate f_{xxyz} if $f(x, y, z) = \sin(3x + yz)$.

4 Partial Differential Equations

A **partial differential equation** (PDE) is a differential equation involving functions of several variables and their partial derivatives. Many important and interesting phenomena are modelled by functions of several variables that satisfy certain partial differential equations. We mention three particular partial differential equations that arise frequently in mathematics and the physical sciences.

- Laplace's equation

- Wave equation
- Heat equation.

We will also introduce the Cobb-Douglas Production model and the associated partial differential equation.

4.1 Laplace's equation

Laplace's equation

The partial differential equation of the form:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is called **Laplace's equation**. Solutions of this equation are called **harmonic functions**.

Laplace's equations play a role in problems of heat conduction, fluid flow, and electric potential.

Example 27. Show that the function $u(x, y) = e^x \sin y$ is a solution of Laplace's equation.

Solution. ...

4.2 Wave equation

Wave equation

The partial differential equation of the form:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

is called a **wave equation**.

This equation describes the motion of a waveform, which could be an ocean wave, a sound wave, a light wave, or a wave traveling along a vibrating string.

If t measures time, then $f(x - ct)$ represents a waveform travelling to the right along the x-axis with speed c depending on the density of the string and on the tension in the string.. (See the figure given below.)

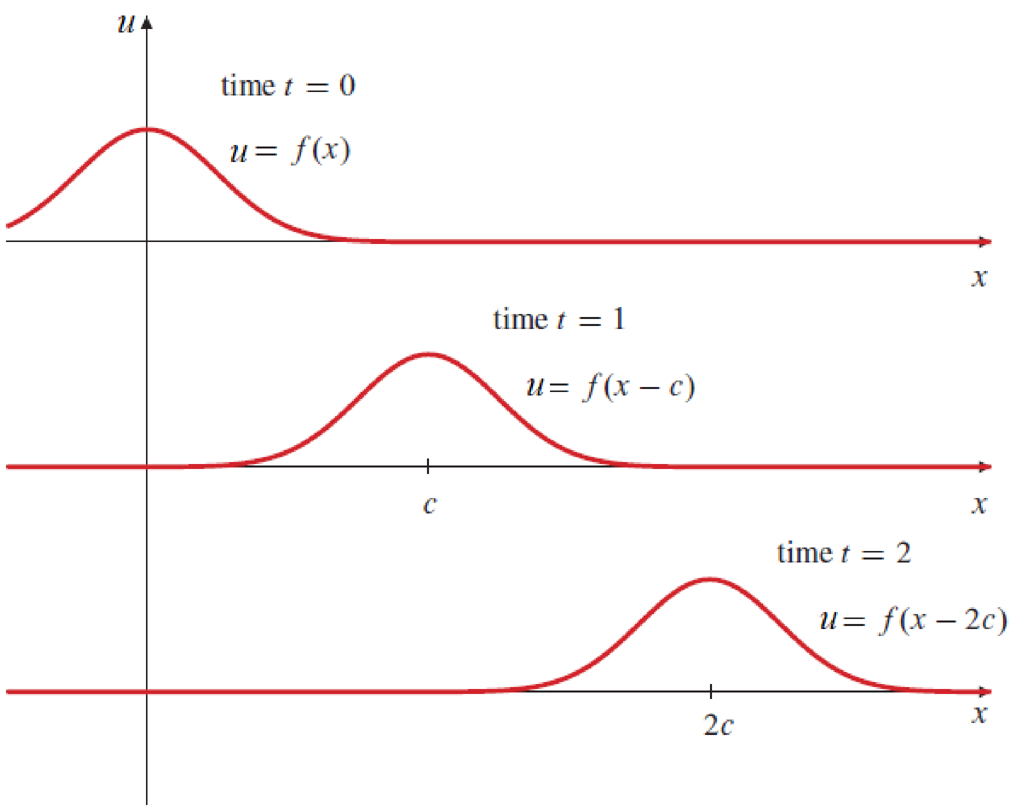


Figure 18: $u = f(x - ct)$ represents a waveform moving to the right with speed c.

Similarly, $g(x + ct)$ represents a waveform travelling to the left with speed c. Unlike the solutions of Laplace's equation that must be infinitely differentiable, solutions of the wave equation need only have enough derivatives to satisfy the differential equation. The functions f and g are otherwise arbitrary.

Example 28. Verify that the function $u(x, y) = \sin(x - ct)$ satisfies the wave equation.

Solution. ...

4.2.1 Heat (diffusion) equation

Wave equation

The partial differential equation of the form:

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$$

is called a **heat equation** or **diffusion equation**.

Example 29. Verify that the function $u(x, y) = e^{-a^2 t} \sin(ax)$ satisfies the heat equation with $c = 1$.

Solution. ...

4.3 Cobb-Douglas Production Function

Let

P : total production of an economic system

L : the amount of labor required to produce P

K : the capital investment required to produce P.

Then the total production P can be described as a function of L and K.

Let the production function be denoted by $P = P(L, K)$. Then

- The partial derivative $\partial P/\partial L$ is the rate at which production changes with respect to the amount of labor. Economists call it the **marginal production with respect to labor** or the **marginal productivity of labor**.
- The partial derivative $\partial P/\partial K$ is the rate of change of production with respect to capital and is called the **marginal productivity of capital**.

Problem 12. The assumptions made by Cobb and Douglas can be stated as follows.

(i) If either labor or capital vanishes, then so will production.

(ii) The marginal productivity of labor is proportional to the amount of production per unit of labor.

(iii) The marginal productivity of capital is proportional to the amount of production per unit of capital.

Construct the Cobb–Douglas production model.

Solution. Because the production per unit of labor is P/L, assumption (ii) says that

$$\frac{\partial P}{\partial L} = \alpha \frac{P}{L}$$

for some constant α. Assumption (i) shows that $\alpha > 0$. If we keep $K = K_0$ constant , then this partial differential equation becomes an ordinary differential equation:

$$\frac{dP}{dL} = \alpha \frac{P}{L}.$$

If we solve this separable differential equation, we get

$$P(L, K_0) = C_1(K_0)L^{\alpha}. \tag{2}$$

Notice that we have written the constant C_1 as a function of K_0 because it could depend on the value of K_0.

Similarly, assumption (iii) says that

$$\frac{\partial P}{\partial K} = \beta \frac{P}{K},$$

for some constant β. Assumption (i) shows that $\beta > 0$. Keeping L constant($L = L_0$), and we can solve this differential equation to get

$$P(L_0, K) = C_2(L_0)K^{\beta}. \tag{3}$$

Comparing Equations (2) and (3), we have

$$P(L, K) = cL^{\alpha}K^{\beta}, \tag{4}$$

where c is a constant that is independent of both L and K.

Thus,

Cobb-Douglas production model

Let

P: total production of an economic system

L: amount of labor required to produce P

K: capital investment required to produce P.

Also, let c be a constant independent of both L and K, and let $\alpha, \beta > 0$. The function P given by the equation

$$P(L, K) = cL^{\alpha}K^{\beta},$$

is called the **Cobb-Douglas production model**.

Example 30. Consider the Cobb-Douglas production model given by the formula $P = 1.01L^{0.75}K^{0.25}$. Its level curves are shown below.
In the figure level curves are labeled with the value of the production P. For instance, the level curve labeled 140 shows all values of the labor L and capital investment K that result in a production of $P = 140$. We see that, for a fixed value of P, as L increases K decreases, and vice versa.

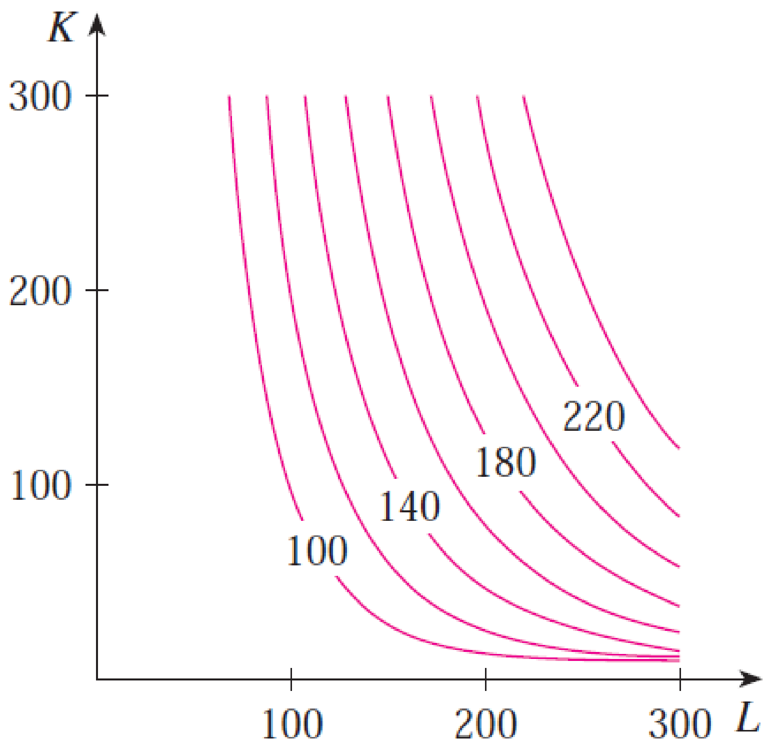


Figure 19: Level curves for $P = 1.01L^{0.75}K^{0.25}$.

Example 31. Show that the Cobb-Douglas production function $P = cL^{\alpha}K^{\beta}$ satisfies the equation

$$L\frac{\partial P}{\partial L} + K\frac{\partial P}{\partial K} = (\alpha + \beta)P.$$

Notice from Equation (4) that if labor and capital are both increased by a factor m, then

$$P(mL, mK) = b(mL)^{\alpha}(mK)^{\beta} = m^{\alpha+\beta}P(L, K)$$

If $\alpha + \beta = 1$, then $P(mL, mK) = mP(L, K)$, which means that production is also increased by a factor of m. Hence Cobb and Douglas assumed that $\alpha + \beta = 1$ and therefore

$$P(L, K) = cL^{\alpha}K^{1-\alpha}.$$

Note that if $\alpha+\beta = 1$, then the production function P satisfies the differential equation:

$$L\frac{\partial P}{\partial L} + K\frac{\partial P}{\partial K} = P.$$

Cobb-Douglas production function

Let

P: total production of an economic system

L: amount of labor required to produce P

K: capital investment required to produce P.

Also, let c be a constant independent of both L and K, and let $0 < \alpha < 1$. The **Cobb-Douglas production function** P is given by the equation

$$P(L, K) = cL^{\alpha}K^{1-\alpha}.$$

5 Tangent planes and linear approximation

In this section we will extend the notion of differentiability to functions of two variables. Our definition of differentiability will be based on the idea that a function is differentiable at a point provided it can be very closely approximated by a linear function near that point. In the process, we will expand the concept of a "differential" to functions of more than one variable and define the "local linear approximation" of a function. Thus, we

- Determine the equation of a plane tangent to a given surface at a point.
- Use the tangent plane to approximate a function of two variables at a point.
- Explain when a function of two variables is differentiable.
- Use the total differential to approximate the change in a function of two variables.

Tangent Planes

A function of one variable: $y = f(x)$.

The slope of the tangent line at the point $x = a$: $m = f'(a)$.

The equation of the tangent line at the point $x = a$:

$$y = f(a) + f'(a)(x - a).$$

What is the slope of a tangent plane?

Tangent plane

Let $P_0 = (x_0, y_0, z_0)$ be a point on a surface S, and let C be any curve passing through P_0 and lying entirely in S. If the tangent lines to all such curves C at P_0 lie in the same plane, then this plane is called the **tangent plane** to S at P_0.

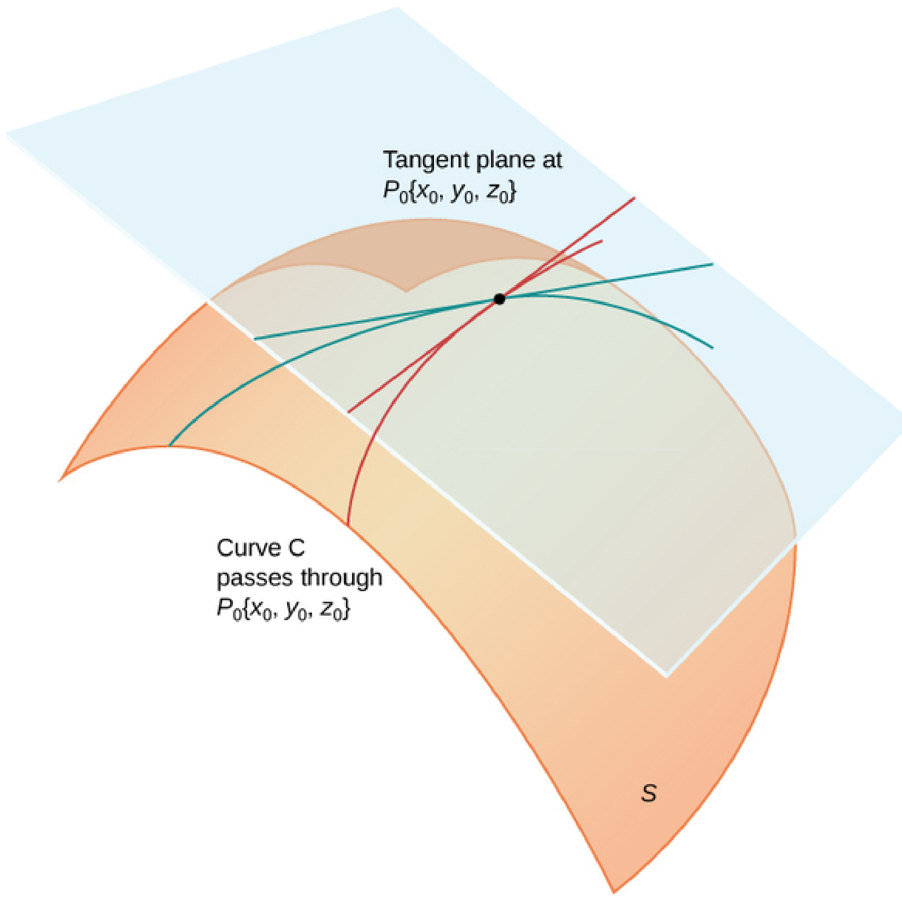


Figure 20: The tangent plane to a surface at a point contains all the tangent lines to curves in that pass through.

Equation of a tangent plane

We know that any plane passing through the point (x_0, y_0, z_0) has an equation of the form

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \tag{5}$$

Dividing this equation by C, we can write it in the form

$$z - z_0 = a(x - x_0) + b(y - y_0). \tag{6}$$

Equation of a tangent plane

Let S be a surface defined by a function $z = f(x, y)$ and $P_0 = (x_0, y_0)$ a point in the domain of f. Suppose that f has continuous partial derivatives. Then the **tangent plane** to S at P_0 is given by the equation

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \tag{7}$$

Example 32. Find the tangent plane to the elliptic paraboloid $z = 2x^2 + y^2$ at the point $(1, 1, 3)$.

Solution. ...

Linear Approximations

One nice use of tangent planes is they give us a way to approximate a surface near a point. As long as we are near to the point (x_0, y_0) then the tangent plane should nearly approximate the function at that point.

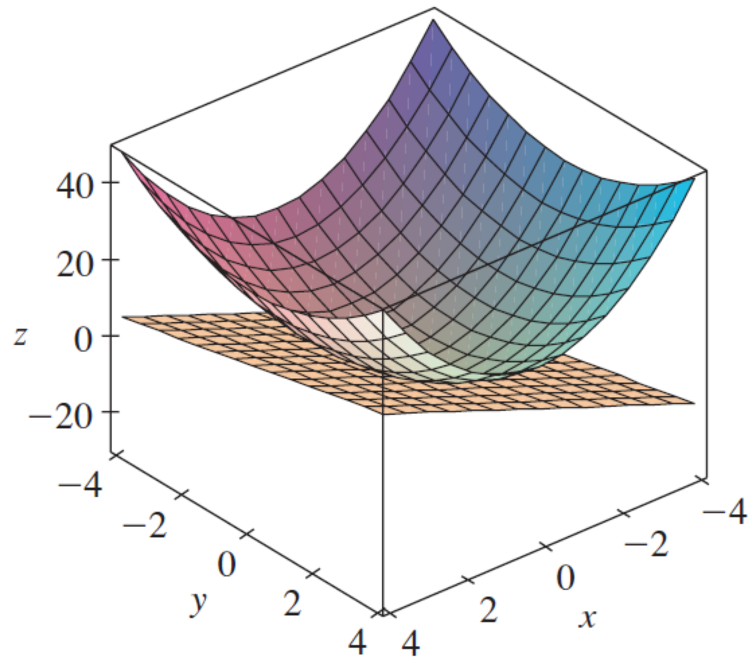


Figure 21: The tangent plane to $z = 2x^2 + y^2$ at the point $(1, 1, 3)$.

In Example 32, we found that an equation of the tangent plane to the graph of the function $f(x, y) = 2x^2 + y^2$ at the point $(1, 1, 3)$ is

$$z = 4x + 2y - 3.$$

Set

$$L(x, y) = 4x + 2y - 3.$$

This is a linear function in two variables

For instance, at the point $(1.1, 0.95)$ we have

$$f(1.1, 0.95) = 2(1.1)^2 + (0.95)^2 = 3.3225,$$
$$L(1.1, 0.95) = 4(1.1) + 2(0.95) - 3 = 3.3$$

Clearly, $f(x, y) \approx L(x, y)$. But if we take a point farther away from $(1, 1)$, such as $(2, 3)$, we no longer get a good approximation.

In fact,

$$f(2,3) = 2(2)^2 + (3)^2 = 17,$$
$$L(2,3) = 4(2) + 2(3) - 3 = 11.$$

Thus, $L(x, y)$ is a good approximation to $f(x, y)$ when (x, y) is near (1, 1). The function L is called the **linearization** of f at $(1, 1)$ and the approximation

$$f(x, y) \approx 4x + 2y - 3$$

is called the **linear approximation** or **tangent plane approximation** of f at $(1, 1)$.

Because of this we define the linear approximation to be,

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

Linear approximation

In general, if f has continuous partial derivatives, then the equation of a tangent plane to the graph of a function $f(x, y)$ at the point $(a, b, f(a, b))$ is

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Put

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

The expression $L(x, y)$ is called the **linear approximation** or the **tangent plane approximation** of f at (a, b). Indeed,

$$f(x, y) \approx L(x, y) \text{ near } (a, b).$$

We have defined tangent planes for surfaces $z = f(x, y)$, where f has continuous first partial derivatives. What happens if f_x and f_y are not continuous?

Example 33. Consider the following function:

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

We can verify that its partial derivatives exist at the origin and, in fact, $f_x(0,0) = f_y(0,0) = 0$, but f_x and f_y are not continuous. The linear approximation would be $f(x,y) \approx 0$, but $f(x,y) = \frac{1}{2}$ at all points on the line $y = x$. So a function of two variables can behave badly even though both of its partial derivatives exist.

To rule out such behavior as in the above example, we formulate the idea of a differentiable function of two variables.

Differentiable functions of two variables

Let us return to the one-dimensional case. If Δx is an increment in x, then the increment in y, Δy, is defined as

$$\Delta y = f(x + \Delta x) - f(x).$$

By definition, the derivative of $f : \mathbb{R} \to \mathbb{R}$ is given by

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

Put

$$\frac{\Delta y}{\Delta x} - f'(x) = \varepsilon. \tag{8}$$

Clearly,

$$\varepsilon \to 0 \text{ as } \Delta x \to 0.$$

From Equation (8), we get

$$\Delta y = f'(x)\Delta x + \varepsilon \Delta x, \quad \text{where } \varepsilon \to 0 \text{ as } \Delta x \to 0. \tag{9}$$

Thus, a function $y = f(x)$ has a derivative (or is differentiable) if (9) holds true.

Now consider a function of two variables, $z = f(x, y)$, and suppose that Δx is an increment in x and Δy an increment in y. Then the corresponding increment of z is given by

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y)$$

By analogy with (9) we define the differentiability of a function of two variables as follows.

Differentiable functions

Let $z = f(x, y)$. We say that f is differentiable at (a, b) if Δz can be expressed in the form

$$\Delta z = f_x(a, b)\Delta x + f_y(a, b)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y,$$

where $\varepsilon_1, \varepsilon_2 \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$.

Example 34. Show that the function $f(x, y) = x^2 + 3y$ is differentiable at every point in the plane.

Solution. ...

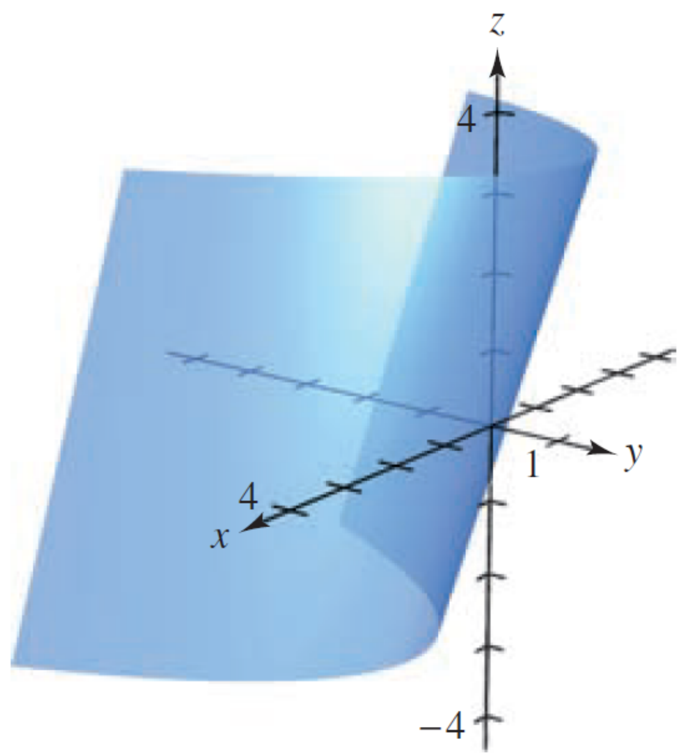


Figure 22: The surface $f(x, y) = x^2 + 3y$.

The above definition says that a differentiable function f is one for which the linear approximation

$$\begin{aligned} f(x, y) &\approx L(x, y) \\ &= f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \end{aligned}$$

is a good approximation when (x, y) is near (a, b). In other words, the tangent plane approximates the graph of f well near the point of tangency.

It's sometimes hard to use above definition directly to check the differentiability of a function, but the next theorem provides a convenient sufficient condition for differentiability.

Theorem 5.1. If the partial derivatives f_x and f_y exist near (a, b) and are continuous at (a, b), then f is differentiable at (a, b).

Example 35 (Using a linearization to estimate a function value)**.** Show that $f(x,y) = xe^{xy}$ is differentiable at $(1,0)$ and find its linearization there. Then use it to approximate $f(1.1,-0.1)$.

Solution. The partial derivatives are

$$f_x(x,y) = e^{xy} + xye^{xy}, \qquad f_y(x,y) = x^2e^{xy}.$$

This implies that

$$f_x(1,0) = 1, \qquad f_y(1,0) = 1.$$

Both f_x and f_y are continuous functions, so f is differentiable. The linearization is

$$\begin{aligned} L(x,y) &= f(1,0) + f_x(1,0)(x-1) + f_y(1,0)(y-0) \\ &= 1 + 1(1-0) + 1 \cdot y = x + y. \end{aligned}$$

The corresponding linear approximation is

$$xe^{xy} \approx x + y$$

and so

$$f(1.1,-0.1) \approx 1.1 - 0.1 = 1.$$

Compare this with the actual value of

$$f(1.1,-0.1) = 1.1e^{-0.11} \approx 0.98542.$$

Theorem 5.2 (Differentiability Implies Continuity)**.** If a function $z = f(x,y)$ is differentiable at a point, then it is continuous at the point.

Proof.

Assume that f is differentiable at a point (a, b). To prove that f is continuous at (a, b) we must show that

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b).$$

Let

$$\Delta x = x - a, \quad \Delta y = y - b,$$

Then

$$\Delta z = f_x(a, b)\Delta x + f_y(a, b)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y,$$

where $\varepsilon_1, \varepsilon_2 \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. However, by definition, we know that

$$\Delta z = f(a + \Delta x, b + \Delta y) - f(a, b)$$

Setting $x = a + \Delta x$ and $y = b + \Delta y$ produces

$$\begin{aligned} & f(x, y) - f(a, b) \\ =& f_x(a, b)\Delta x + f_y(a, b)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y \\ =& f_x(a, b)(x - a) + f_y(a, b)(y - b) + \varepsilon_1 (x - a) + \varepsilon_2 (y - b) \end{aligned}$$

Taking the limit as $(x, y) \to (a, b)$, we have

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b).$$

which means that f is continuous at (a, b). ◀

Differentials

Recall: one dimensional case.

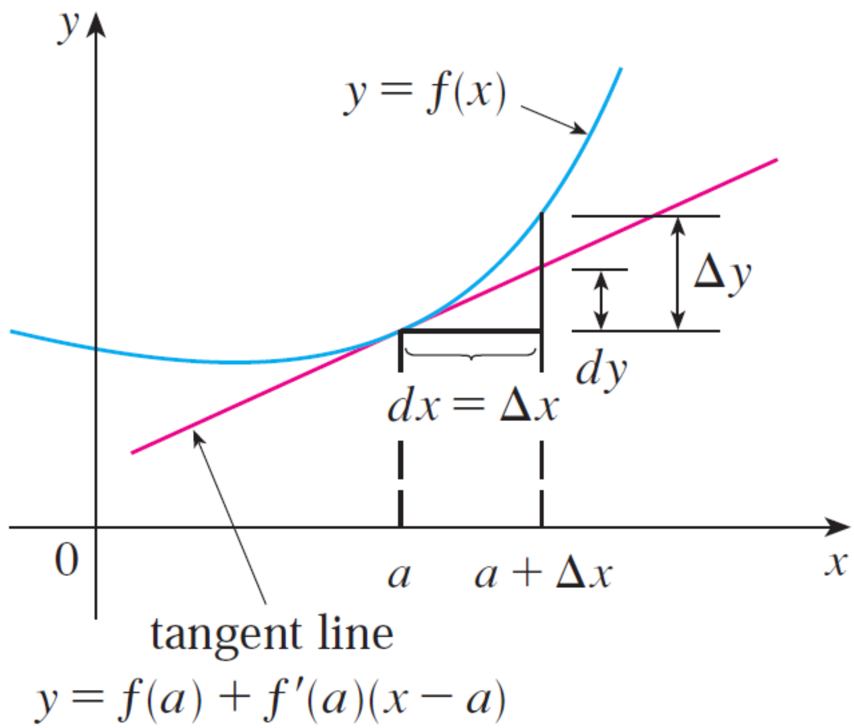


Figure 23: The tangent line

For a differentiable function of two variables, $z = f(x, y)$, we define the differentials dx and dy to be independent variables; that is, they can be given any values.

Total differential

Let $z = f(x, y)$. If Δx and Δy are increments in x and y, the differentials of the independent variables and are

$$dx = \Delta x, \ dy = \Delta y.$$

The **total differential**, dz, of the dependent variable z is defined by

$$dz = f_x(x, y)dx + f_y(x, y)dy = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy.$$

We know that

$$\begin{aligned} f(x,y) &\approx L(x,y) \\ &= f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b). \end{aligned}$$

If we take

$$dx = \Delta x = x - a, \quad dy = \Delta y = y - b,$$

then the differential dz is

$$dz = f_x(x,y)(x-a) + f_y(x,y)(y-b).$$

So, the linear approximation can be written as

$$\boxed{f(x,y) \approx f(a,b) + dz = L(x,y).}$$

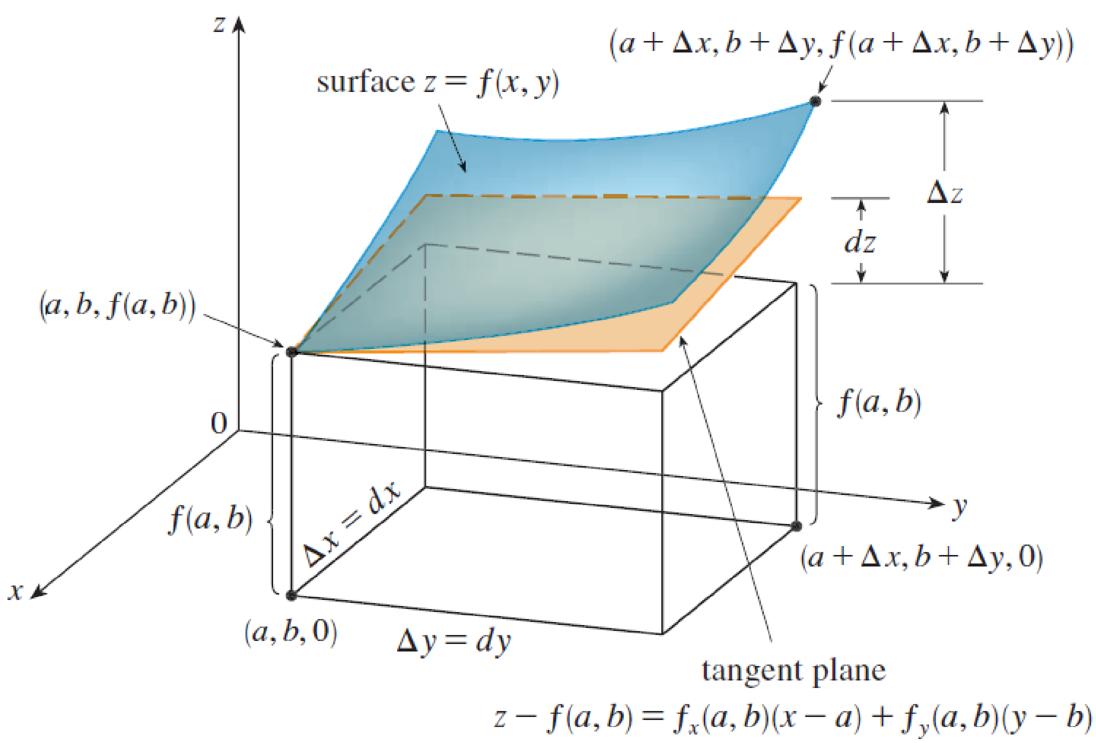


Figure 24: The tangent plane

Figure 24 is the three-dimensional counterpart of Figure 23 and shows the geometric interpretation of the differential and the increment : represents the change in height of the tangent plane, whereas represents the change in height of the surface when changes from (a, b) to $(a + \Delta x, b + \Delta y)$.

Example 36 (Differentials versus increments)**.**

(a) If $z = f(x, y) = x^2 + 3xy - y^2$, find the differential dz.

(b) If x changes from 2 to 2.05 and y changes from 3 to 2.96, compare the values of Δz and dz.

Solution. (a)

(b) Putting

$$x = 2, dx = \Delta x = 0.05, \quad y = 3, \quad dy = \Delta y = -0.04,$$

we get

$$dz = 2(2) + 3(3)0.05[3(2) - 2(3)](-0.04) = 0.65$$

The increment of z is

$$\begin{aligned}\Delta z &= f(2.05, 2.96) - f(2, 3) \\ &= [2.05^2 + 3(2.05)(2.96) - 2.96^2] - [2^2 + 3(2)(3) - 3^2] \\ &= 0.6449.\end{aligned}$$

Therefore, $\Delta z \approx dz$.

Example 37 (Using differentials to estimate an error:)**.** The base radius and height of a right circular cone are measured as 10 cm and 25 cm, respectively, with a possible error in measurement of as much as 0.1 cm in each. Use differentials to estimate the maximum error in the calculated volume of the cone.

Solution. We know that the volume of a cone with base radius r and height h is

$$V = \frac{1}{3}\pi r^2 h.$$

So

$$dV = \frac{\partial V}{\partial r}dr + \frac{\partial V}{\partial h}dh = \frac{2\pi r h}{3}dr + \frac{\pi r^2}{3}dh.$$

Since each error is at most 0.1 cm, we have

$$|\Delta r| \leq 0.1, \quad |\Delta h| \leq 0.1$$

To find the largest error in the volume we take the largest error in the measurement of r and of h. Therefore we take $dr = 0.1$ and $dh = 0.1$ along with $r = 10$, $h = 25$. This gives

$$dV = \frac{500\pi}{3}(0.1) + \frac{100\pi}{3}(0.1) = 20\pi.$$

Thus the maximum error in the calculated volume is about

$$20\pi \text{ cm}^3 = 63 \text{ cm}^3.$$

Functions of Three or More Variables

Linear approximation:

$$\begin{aligned} f(x,y,z) &\approx L(x,y,z) \\ &= f(a,b,c) + f_x(a,b,c)(x-a) + f_y(a,b,c)(y-b) \\ &\qquad + f_x(a,b,c)(z-c). \end{aligned}$$

Differentiability

Let $u = f(x,y,z)$. We say that f is differentiable at (a,b,c) if Δu can be expressed in the form

$$\begin{aligned} \Delta u = f_x(a,b,c)\Delta x + f_y(a,b,c)\Delta y + f_z(a,b,c)\Delta z + \varepsilon_1 \Delta x \\ + \varepsilon_2 \Delta y + \varepsilon_3 \Delta z, \end{aligned}$$

where $\varepsilon_1, \varepsilon_2, \varepsilon_3 \to 0$ as $(\Delta x, \Delta y, \Delta z) \to (0,0,0)$.

Total differentials

The differential du, also called the **total differential**, is defined by

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial yz}dz.$$

Example 38. The dimensions of a rectangular box are measured to be 75 cm, 60 cm, and 40 cm, and each measurement is correct to within 0.2 cm. Use differentials to estimate the largest possible error when the volume of the box is calculated from these measurements.

Solution. If the dimensions of the box are x, y, and z, its volume is

$$v = xyz$$

and so

$$\begin{aligned} du &= \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial yz}dz \\ &= yzdx + xzdy + xydz. \end{aligned}$$

We are given that

$$\Delta x, \Delta y, \Delta z \leq 0.2.$$

To find the largest error in the volume, we therefore use

$$dx = dy = dz = 0.2$$

together with

$$x = 75, y = 60, z = 40.$$

We have

$$\Delta V = dV = (60)(40)(0.2)+(75)(40)(0.2)+(75)(60)(0.2) = 1980$$

Thus an error of only 0.2 cm in measuring each dimension could lead to an error of as much as 1980 cm in the calculated volume! This may seem like a large error, but it's only about 1% of the volume of the box.

6 Chain rules

Chain rule I

Let $x = x(t)$, $y = y(t)$ and $z = f(x, y)$ be differentiable functions. Then $z = f(x(t), y(t))$ is a differentiable function of t and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt},$$

where the ordinary derivatives are evaluated at t and the partial derivatives are evaluated at (x, y).

Proof.

A change of Δt in t produces changes of Δx in x and Δy in y . These, in turn, produce a change of Δz in z. Since $z = f(x, y)$ is differentiable, we have

$$\Delta z = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y,$$

where $\varepsilon_1, \varepsilon_2 \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. Dividing both sides of this equation by Δt, we have

$$\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t} + \varepsilon_1 \frac{\Delta x}{\Delta t} + \varepsilon_2 \frac{\Delta y}{\Delta t},$$

If we now let $\Delta t \to 0$, then

$$\Delta x = x(t + \Delta t) - x(t) \to 0,$$

because x is differentiable, therefore, continuous. Similarly,

$$\Delta y \to 0.$$

This, in turn, means that $\varepsilon_1, \varepsilon_2 \to 0$, so

$$\begin{aligned}
\frac{dz}{dt} &= \lim_{\Delta t \to 0} \frac{\Delta z}{\Delta t} \\
&= \frac{\partial f}{\partial x} \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y} \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} + \lim_{\Delta t \to 0} \varepsilon_1 \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} \\
&\quad + \lim_{\Delta t \to 0} \varepsilon_2 \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} \\
&= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} + 0 \cdot \frac{dx}{dt} + 0 \cdot \frac{dy}{dt} \\
&= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.
\end{aligned}$$

Hence $z = f(x(t), y(t))$ is a differentiable function of t and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

◀

Example 39. If $z = x^2y + 3xy^4$, where $x = \sin 2t$ and $y = \cos t$, find dz/dt when $t = 0$.

The derivative in Example(39) can be interpreted as the rate of change of z with respect to t as the point (x, y) moves along the curve C with parametric equations $x = \sin 2t$, $y = \cos t$, (See the figure given below.)

In particular, when $t = 0$, the point (x, y) is (0,1) and $dz/dt = 6$ is the rate of change as we move along the curve C through (0,1). If, for instance,

$$z = T(x, y) = x^2 + 3xy^4$$

represents the temperature at the point (x, y), then the composite function $z = T(\sin 2t, \cos t)$ represents the temperature

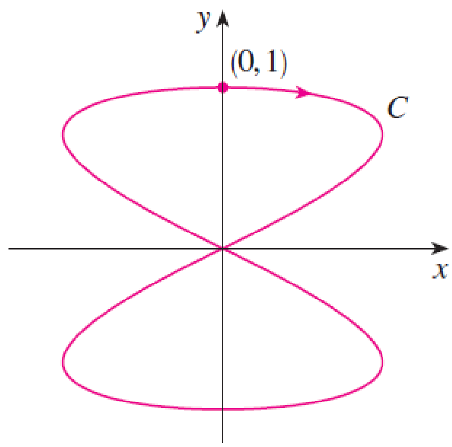


Figure 25: The curve $x = \sin 2t$, $y = \cos t$

at points on C and the derivative dz/dt represents the rate at which the temperature changes along C.

Chain rule II

Let $x = x(s,t)$, $y = y(s,t)$ and $z = f(x,y)$ be differentiable functions. Then $z = f(x(s,t), y(s,t))$ be a differentiable function of s, t. And

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$$
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

Example 40. If $z = e^x \sin y$, where $x = st^2, y = s^2t$, find $\partial z/\partial s, \partial z/\partial t$.

It is easy to extend the chain rule to the general situation in which a dependent variable z is a function of n intermediate variables $x_1, ..., x_n$ each of which is, in turn, a function of m independent variables $t_1, ..., t_m$.

Example 41. Write out the Chain Rule for the case, where $w = f(x, y, z, t)$ and $x = x(u, v), y = y(u, v), z = z(u, v), t = t(u, v)$.

Example 42. If $g(s, t) = f(s^2 - t^2, t^2 - s^2)$ and f is differentiable, show that g satisfies the equation:

$$t\frac{\partial g}{\partial s} + s\frac{\partial g}{\partial t} = 0.$$

Example 43. If $z = f(x, y)$ and f has continuous second-order partial derivatives and $x = r^2 + s^2, y = 2rs$, find $\partial z/\partial r, \partial^2 z/\partial r^2$.

Implicit Differentiation

(as an application of the chain rule.)

Case I: $F(x, y) = 0$, where $y = f(x)$.

Implicit Function Theorem I

If F is defined on a disk containing (a, b), where $F(a, b) = 0, F_y(a, b) \neq 0$, and F_x, F_y are continuous on the disk, then the equation $F(x, y) = 0$ defines y as a function of x near the point (a, b) and the derivative of this function is given by

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{F_x}{F_y}.$$

Proof.

Consider the function

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

By the chain rule, we have

$$\frac{dz}{dx} = \frac{\partial z}{\partial x}\frac{dx}{dx} + \frac{\partial z}{\partial y}\frac{dy}{dx} = F_x(x, y)\frac{dx}{dx} + F_y(x, y)\frac{dy}{dx}.$$

Since $z = F(x, y) = 0$ for all x in the domain of f, we obtain

$$\frac{dz}{dx} = 0$$

and we have

$$F_x(x, y)\frac{dx}{dx} + F_y(x, y)\frac{dy}{dx} = 0.$$

Therefore, if $F_y(x, y) \neq 0$, then

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)}.$$

◀

Example 44. Find y' if $x^3 + y^3 = 6xy$.

Case II: $F(x, y, z) = 0$, where $z = f(x, y)$.

Implicit Function Theorem II

If F is defined within a sphere containing (a, b, c), where $F(a, b, c) = 0, F_z(a, b, c) \neq 0$, and F_x, F_y, F_z are continuous inside the sphere, then the equation $F(x, y, z) = 0$ defines z as a function of x, y near the point (a, b, c) and the derivative of this function is differentiable, with partial derivatives given by

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \quad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}.$$

Example 45. Find $\partial z/\partial x$ and $\partial z/\partial y$ if $x^3 + y^3 + z^3 + 6xyz = 1$.

7 Directional derivatives and gradient vector

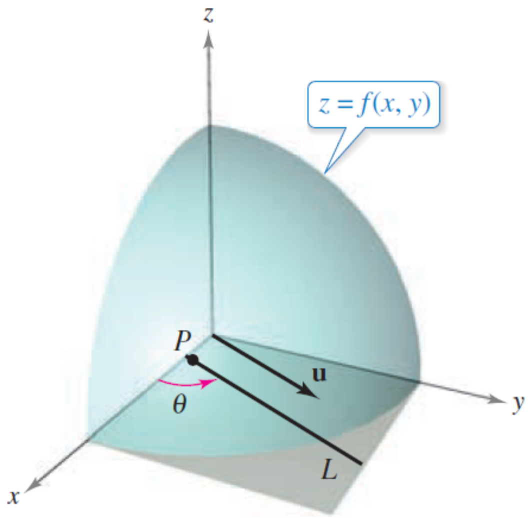
The partial derivatives of a function f tell us the rate of change of f in the direction of the coordinate axes.

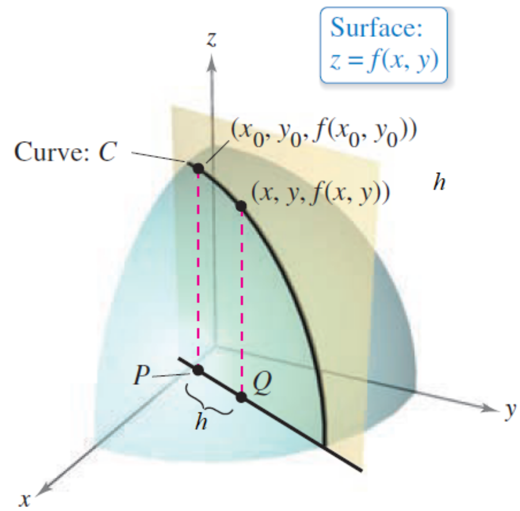
How can we measure the rate of change of f
in other directions?

In order to formally define the derivative in a particular direction of motion, we want to represent the change in f for a given unit change in the direction of motion.

We can represent this unit change in direction with a unit vector, say $u = (a, b)$. This unit vector helps us to "mark off" units on the line. A vector equation for the line through (x_0, y_0) in this direction is

$$v(h) = (x_0 + ha, y_0 + hb).$$





Because u is a unit vector, the value of h is precisely the distance along the line from (x_0, y_0) to (x_0+ha, y_0+hb). Indeed,

$$\begin{aligned}
&\|(x_0 + ha, y_0 + hb) - (x_0, y_0)\| \\
=&\|(ha, hb)\| \\
=&|h|\|(a, b)\| \\
=&|h|. \qquad (\because (a, b) \text{ is a unit vector})
\end{aligned}$$

If we move a distance in the direction of u from a fixed point (x_0, y_0), we then arrive at the new point $(x_0 + ha, y_0 + hb)$. It now follows that the slope of the secant line to the curve on the surface through (x_0, y_0) in the direction of u through the points (x_0, y_0) and $(x_0 + ha, y_0 + hb)$ is

$$m_{\text{sec}} = \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}. \tag{1}$$

To get the instantaneous rate of change of f in the direction $u = (a, b)$, we must take the limit of the quantity in Equation (1) as $h \to 0$. Doing so results in the formal definition of the directional derivative.

Directional derivatives

The directional derivative of f at (x_0, y_0) in the direction of a unit vector $u = (a, b)$ is

$$D_u f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

Partial derivatives and directional derivatives

If $u = i = (1, 0)$, then

$$D_i f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = f_x(x_0, y_0).$$

If $u = j = (0, 1)$, then

$$D_j f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} = f_y(x_0, y_0).$$

Thus,

$$D_i f = f_x, \qquad D_j f = f_y.$$

In other words, the partial derivatives of f with respect to x and y are just special cases of the directional derivative.

Method of finding directional derivatives

It is time consuming to find the directional derivative using the above definition. However, we can find a way to evaluate directional derivatives without resorting to the limit definition.

Theorem 7.1. If f is a differentiable function of x and y, then f has a directional derivative at (x_0, y_0) in the direction of a **unit vector** $u = (a, b)$ and

$$D_u f(x_0, y_0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b. \tag{10}$$

Proof.

For a fixed point (x_0, y_0) let

$$x = x_0 + ha, \qquad y = y_0 + hb.$$

Then differentiating both x and y with respect to h, we obtain

$$x' = a, \qquad y' = b.$$

Let

$$g(h) = f(x_0 + ha, y_0 + hb) = f(x, y).$$

Because f is differentiable, we can apply the Chain Rule to obtain

$$g'(h) = f_x(x, y)x'(h) + f_y(x, y)y'(h) = f_x(x, y)a + f_y(x, y)b.$$

If $h = 0$, then $x = x_0, y = y_0$, and so,

$$g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b.$$

By the definition of $g'(h)$, it is also true that

$$\begin{aligned} g'(0) &= \lim_{h \to 0} \frac{g(h) - g(0)}{h} \\ &= \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h} \\ &= D_u f(x_0, y_0). \end{aligned}$$

Therefore, we see that

$$D_u f(x_0, y_0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b. \qquad \blacktriangleleft$$

Remark 1.

To use the theroem, we must have a unit vector in the direction of motion. In the event that we have a direction prescribed by a non-unit vector, we must first scale the vector to have length 1.

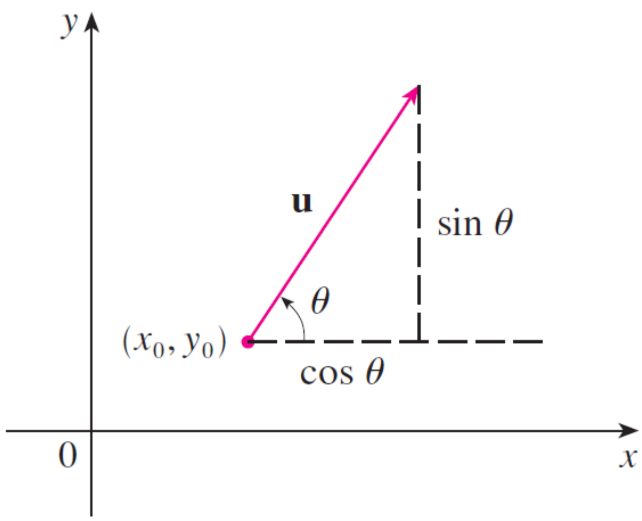


Figure 26: A unit vector $\vec{u} = (a, b) = (\cos\theta, \sin\theta)$

If the unit vector u makes an angle θ with the positive x-axis (as in Figure 26), then we can write $u = (\cos\theta, \sin\theta)$ and Formula (10) becomes

$$D_u f(x_0, y_0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta.$$

Example 46. Find the directional derivative $D_u f(x, y)$ if

$$f(x, y) = x^3 - 3xy + 4y^2$$

and u is the unit vector given by angle $\theta = \pi/6$. What is $D_u f(1, 2)$?

Example 47. Find the directional derivative of

$$f(x, y) = x^2 \sin 2y$$

at $(1, \pi/2)$ in the direction of

$$v = 3i - 4j.$$

The Gradient Vector

Notice that the directional derivative of a differentiable function can be written as the dot product of two vectors:

$$\begin{aligned} D_u f(x, y) &= f_x(x, y)a + f_y(x, y)b \\ &= (f_x(x, y), f_y(x, y)) \cdot (a, b) \\ &= (f_x(x, y), f_y(x, y)) \cdot u. \end{aligned}$$

The first vector in this dot product is called the **gradient** of f and is denoted by

$$\operatorname{grad} f \quad \text{or} \quad \nabla f.$$

Gradient

If f is a function of two variables x and y, then the gradient of f is the vector function ∇f defined by

$$\nabla f(x, y) = (f_x(x, y), f_y(x, y)) = f_x i + f_y j.$$

Example 48. Find the gradient f if

$$f(x, y) = \sin x + e^{xy}.$$

What is $\nabla f(0, 1)$?

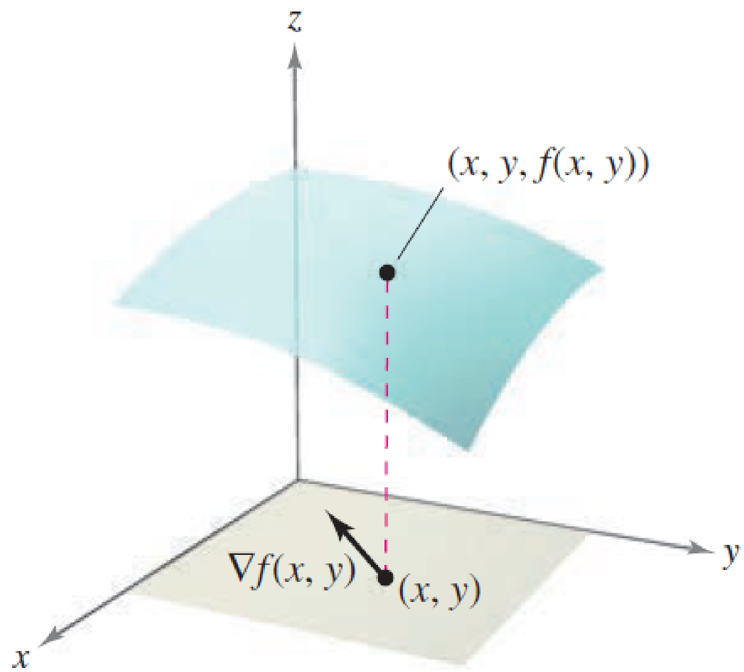


Figure 27: The gradient of f: a vector in the xy- plane.

With this notation for the gradient vector, we can rewrite the expression (10) for the directional derivative of a differentiable function as

$$\boxed{D_u f(x, y) = \nabla f(x, y) \cdot u.}$$

This expresses the directional derivative in the direction of u as the scalar projection of the gradient vector onto u.

Example 49 (Using a gradient vector to find a directional derivative). Find the directional derivative of the function $f(x, y) = x^2y^3 - 4y$ at the point $(2, -1)$ in the direction of the vector $v = 2i + 5j$.

Functions of Three Variables

For functions of three variables we can define directional derivatives in a similar manner. Again $D_u f(x, y, z)$ can be interpreted as the rate of change of the function in the direction of a unit vector $\vec{u}$.

Directional derivative

The directional derivative of f at $\vec{x}_0 = (x_0, y_0, z_0)$ in the direction of a unit vector $\vec{u} = (a, b, c)$ is

$$D_u f(\vec{x}_0) = \lim_{h \to 0} \frac{f(\vec{x}_0 + h\vec{u}) - f(\vec{x}_0)}{h}$$

if this limit exists.

This is reasonable because the vector equation of the line through $\vec{x}_0$ in the direction of the vector $\vec{u}$ is given by

$$\vec{x}_0 + t\vec{u}$$

and so $f(\vec{x}_0 + h\vec{u})$ represents the value of f at a point $\vec{x}_0$ on this line.

If $f(x, y, z)$ is differentiable and $\vec{u} = (a, b, c)$, then we can prove by the same method as in the case of a function of two variables that

$$D_{\vec{u}} f(x, y, z) = f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c. \tag{11}$$

For a function f of three variables, the gradient vector is

$$\nabla f(x, y, z) = (f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)).$$

or

$$\nabla f = (f_x, f_y, f_z) = \frac{\partial f}{\partial x} i + \frac{\partial f}{\partial y} j + \frac{\partial f}{\partial z} k$$

Then, just as with functions of two variables, Formula (11) for the directional derivative can be rewritten as

$$D_{\vec{u}} f(x, y, z) = \nabla f_x(x, y, z) \cdot \vec{u}.$$

Example 50. If $f(x, y, z) = x \sin yz$,

(a) find the gradient of f

(b) find the directional derivative of f at (1, 3, 0) in the direction of $v = i + 2j - k$.

Maximizing the Directional Derivative

Suppose we have a function of two or three variables and we consider all possible directional derivatives of f at a given point. These give the rates of change of f in all possible directions. We can then ask the questions:

- In which of these directions f does change fastest or which one is the direction of maximum increase of f and
- What is the maximum rate of change?

The answers are provided by the following theorem.

Theorem 7.2. Suppose f is a differentiable function of two or three variables. Then the maximum value of the directional derivative $D_{\vec{u}f(x)}$ is $\|\nabla f(x)\|$ and it occurs when $\vec{u}$ has the same direction as the gradient vector $\nabla f(x)$.

Proof.

We have

$$\begin{aligned} D_{\vec{u}} f &= \nabla f \cdot \vec{u} \\ &= \|\nabla f\| \|\vec{u}\| \cos\theta \\ &= \|\nabla f\| \cos\theta, \end{aligned}$$

where θ is the angle between ∇f and $\vec{u}$. The maximum value of $\cos\theta$ is 1 and this occurs when $\theta = 0$. Therefore the maximum value of $D_{\vec{u}} f$ is $\|\nabla f\|$ and it occurs when $\theta = 0$, that is, when $\vec{u}$ has the same direction as ∇f. ◀

Example 51. Let $f(x, y) = xe^y$.

(a) find the directional derivative of f at $P(2, 0)$ in the direction from P to $Q(1/2, 2)$.

(b) In what direction does f have the maximum rate of change? What is this maximum rate of change?

Example 52. Suppose that the temperature in degrees Celsius on the surface of a metal plate is given by

$$T(x, y) = 20 - 4x^2 - y^2,$$

where x and y are measured in centimeters. In what direction from $(2, -3)$ does the temperature increase most rapidly? What is this rate of increase?

Properties of the Gradient

We are now in a position to draw some interesting and important conclusions about the gradient.

First, suppose that $\nabla f_P \neq 0$ and let u be a unit vector (see the figure given below).

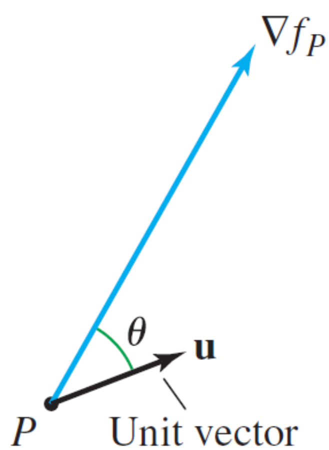


Figure 28: $D_u f(P) = \|\nabla f_P\| \cos\theta$.

We know that

$$D_u f(P) = \nabla f_P \cdot u = \|\nabla f_P\| \cos\theta,$$

where θ is the angle between ∇f_P and u. In other words, The rate of change in a given direction varies with the cosine of the angle θ between the gradient and the direction.

Because the cosine takes values between -1 and 1, we have

$$-\|\nabla f_P\| \leq D_u f(P) \leq \|\nabla f_P\|.$$

Since $\cos\theta = 1$, the maximum value of $D_u f(P)$ occurs for $\theta = 0$ – that is, when u points in the direction of ∇f_P. In other words,

The gradient vector ∇f_P points in the direction of the maximum rate of increase, and this maximum rate is $\|\nabla f_P\|$.

Similarly, f decreases most rapidly in the opposite direction, $-\nabla f_P$, because $\cos\theta = -1$ for $\theta = \pi$. The rate of maximum decrease is $-\|\nabla f_P\|$. The directional derivative is zero in directions orthogonal to the gradient because $\cos(\pi/2) = \theta$.

Another key property is that

Gradient vectors are normal to level curves.

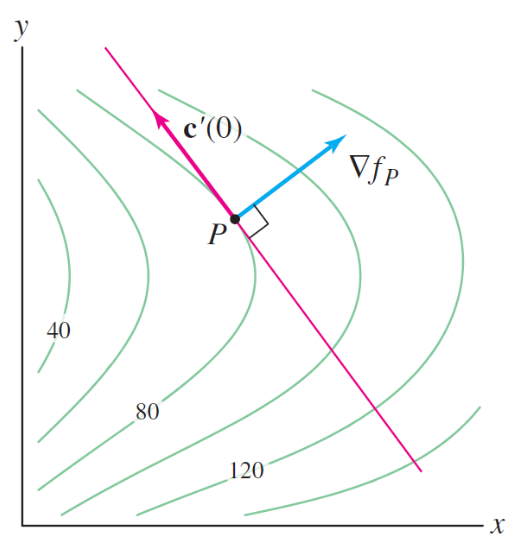


Figure 29: Contour map of $f(x, y)$. The gradient at P is orthogonal to the level curve through P.

Proof.

To prove this, suppose that P lies on the level curve $f(x, y) = k$. We parametrize this level curve by a path $c(\mathrm{t})$ such that $c(0) = P$ and $c'(0) \neq 0$ (this is possible whenever $\nabla f_P \neq 0$). Then $f(c(t)) = k$ for all t , so by the Chain Rule,

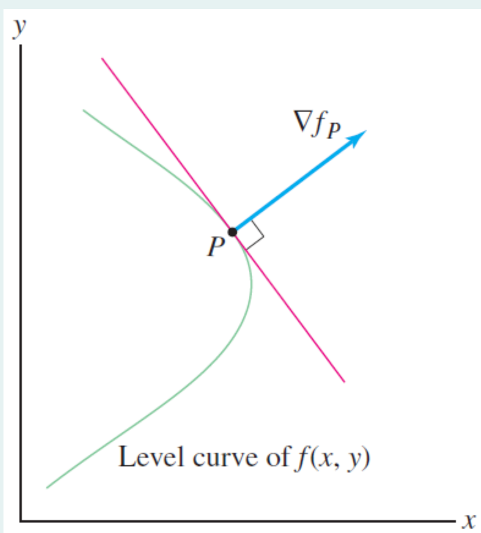
$$\nabla f_P \cdot c'(0) = \frac{d}{dt} f(c(0)) = \frac{d}{dt} k = 0.$$

This proves that ∇f_P is orthogonal to $c'(0)$, and since $c'(0)$ is tangent to the level curve, we conclude that ∇f_P is normal to the level curve. ◀

For functions of three variables, a similar argument shows that ∇f_P is normal to the level surface $f(x, y, z) = k$ through P.

Graphical insight

At each point P, there is a unique direction in which $f(x, y)$ increases most rapidly (per unit distance). This chosen direction is perpendicular to the level curves and that it is specified by the gradient vector.



For most functions, however, the direction of maximum rate of increase varies from point to point.

In summary,

- $D_u f(P) = \nabla f_P \cdot u = \|\nabla f_P\| \cos\theta$.
  That is, the rate of change in a given direction varies with the cosine of the angle θ between the gradient and the direction.

- The gradient vector ∇f_P points in the direction of the maximum rate of increase, and this maximum rate is $\|\nabla f_P\|$.

- The gradient vector $-\nabla f_P$ points in the direction of the maximum rate of decrease, and this maximum rate of decrease is $-\|\nabla f_P\|$.

- Gradient vector ∇f_P is normal to level curve (or surface) of f at P.

Tangent Planes to Level Surfaces

Suppose S is a surface with equation $F(x, y, z) = k$, that is, it is a level surface of a function F of three variables, and let $P(x_0, y_0, z_0)$ be a point on S. Let C be any curve that lies on the surface S and passes through the point P. The curve C is described by a continuous vector function $\vec{r}(t) = (x(t), y(t), z(t))$. Let t_0 be the parameter value corresponding to P; that is, $\vec{r}(t_0) = (x_0, y_0, z_0)$. Since C lies on s, any point $(x(t), y(t), z(t))$ must satisfy the equation of s, that is,

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

If x, y, z are differentiable functions of t and F is also differentiable, then we can use the Chain Rule to differentiate its both sides to obtain

$$\nabla F \cdot \vec{r}\,'(t) = 0.$$

In particular, when $t = t_0$ we have $\vec{r}(t_0) = (x_0, y_0, z_0)$, so

$$\nabla F(x_0, y_0, z_0) \cdot \vec{r}\,'(t_0) = 0.$$

This equation says that the gradient vector $\nabla F(x_0, y_0, z_0)$ at P is perpendicular to the tangent vector $\vec{r}\,'(t_0)$ to any curve C on S that passes through P. (See the figure given below.)

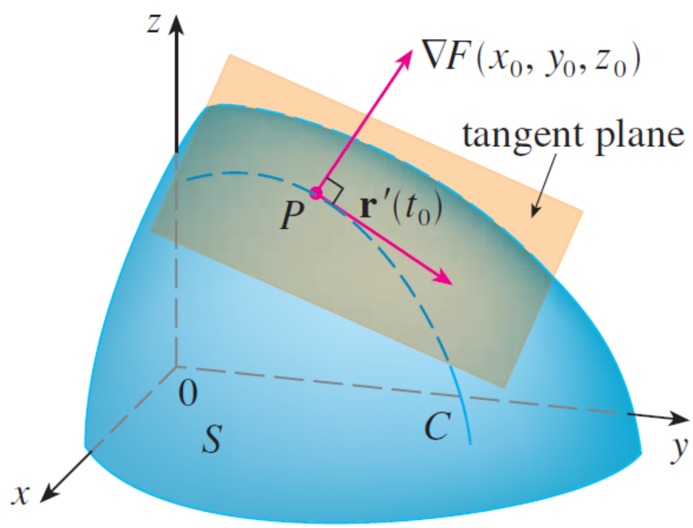


Figure 30

If $\nabla F(x_0, y_0, z_0) \neq 0$, it is therefore natural to define the **tangent plane to the level surface** $F(x_0, y_0, z_0) = k$ at $P(x_0, y_0, z_0)$ as the plane that passes through P and has normal vector $\nabla F(x_0, y_0, z_0)$. Using the standard equation of a plane, we can write the equation of this tangent plane as

$$\begin{aligned} F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) \\ + F_z(x_0, y_0, z_0)(z - z_0) = 0. \end{aligned} \tag{12}$$

The normal line to S at P is the line passing through P and perpendicular to the tangent plane. The direction of the normal line is therefore given by the gradient vector $\nabla F(x_0, y_0, z_0)$ and so, its symmetric equations are

$$\boxed{\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)}}$$

In the special case in which the equation of a surface S is of the form $z = f(x, y, z)$ (that is, is the graph of a function f of two variables), we can rewrite the equation as

$$F(x, y, z) = f(x, y, z) - z = 0$$

and regard S as a level surface (with $k = 0$) of F. Then

$$\begin{aligned} F_x(x_0, y_0, z_0) &= f_x(x_0, y_0) \\ F_y(x_0, y_0, z_0) &= f_y(x_0, y_0) \\ F_z(x_0, y_0, z_0) &= -1 \end{aligned}$$

so Equation (12) becomes

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

Thus our new, more general, definition of a tangent plane is consistent with the definition that was given earlier.

Example 53.

Find the equations of the tangent plane and normal line at the point $(-2, 1, -3)$ to the ellipsoid

$$\frac{x^2}{4} + y^2 + \frac{z^2}{9} = 3.$$

8 Maximum and minimum values

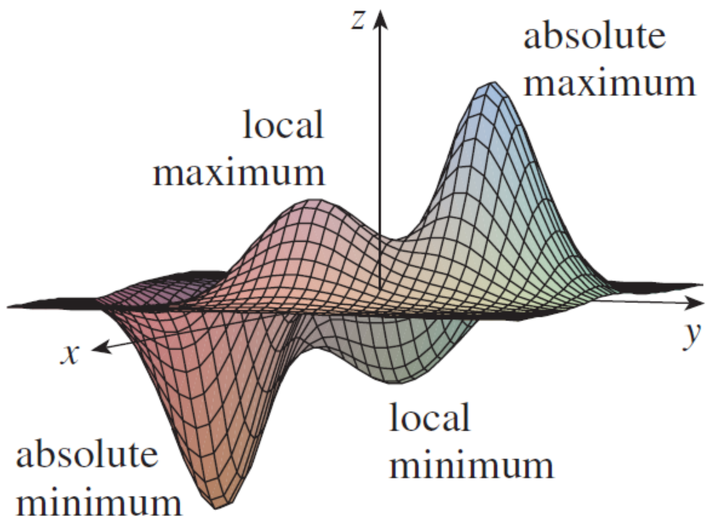


Figure 31

Local extrema

Let f be a function of two variables x and y.

- The function f has a **local maximum** at a point (a, b) provided that $f(x, y) \leq f(a, b)$ for all points (x, y) in some disk with center at (a, b). In this situation we say that $f(a, b)$ is a **local maximum value**.

- The function f has a **local minimum** at a point (a, b) provided that $f(x, y) \geq f(a, b)$ for all points (x, y) in some disk with center at (a, b). In this situation we say that $f(a, b)$ is a **local minimum value**.

We use the term **extremum point** to refer to any point (a, b) at which f has a maximum or minimum. In addition, the function value $f(a, b)$ at an extremum is called an **extremal value**.

Absolute extrema

Let f be a function of two variables x and y.

- An **absolute maximum point** is a point (a, b) for which $f(x, y) \leq f(a, b)$ for all points (x, y) in the domain of f. The value of f at an absolute maximum point is the **maximum value** of f.
- An **absolute minimum point** is a point such that $f(x, y) \geq f(a, b)$ for all points (x, y) in the domain of f. The value of f at an absolute minimum point is the **maximum value** of f.

Critical points

A **critical point** (a, b) of a function $f = f(x, y)$ is a point in the domain of f at which one of the following is true:

1. $f_x(a, b) = 0$ and $f_y(a, b) = 0$,
2. $f_x(a, b)$ or $f_y(a, b)$ fails to exist.

Stationary points

A **stationary point** (a, b) of a function $f = f(x, y)$ is a point in the domain of f at which $f_x(a, b) = f_y(a, b) = 0$.

Theorem 8.1 (Fermat)**.** If f has a local maximum or minimum at (a, b) and the first-order partial derivatives of f exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Proof.

Let $g(x) = f(x, b)$. If f has a local maximum (or minimum) at (a, b), then g has a local maximum (or minimum) at a, so by Fermat's Theorem for functions of one variable,

$$g'(a) = 0.$$

But

$$g'(a) = f_x(a, b)$$

and so

$$f_x(a, b) = 0.$$

Similarly, by applying Fermat's Theorem to the function $G(y) = f(a, y)$, we obtain

$$f_y(a, b) = 0.$$

◀

This theorem says that if f has a local maximum or minimum at (a, b), then (a, b) is a critical point of f. At a critical point, a function could have a local maximum or a local minimum or neither.

Geometric interpretation of Fermat's theorem

If we put $f_x(a, b) = 0$ and $f_y(a, b) = 0$ in the equation of a tangent plane:

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

we get

$$z = z_0.$$

Thus the geometric interpretation of Fermat's theorem is as follows:

Geometric interpretation of Fermat's theorem

If the graph of f has a tangent plane at a local maximum or minimum, then the tangent plane must be horizontal.

Example 54. Investigate the critical points of the function:

$$f(x, y) = x^2 + y^2 - 2x - 6y + 14.$$

Solution. We have

$$f_y(x, y) = 2x - 2 \qquad f_y(x, y) = 2y - 6.$$

These partial derivatives are equal to 0 when $x = 1$ and y=3, so the only critical point is (1,3). By completing the square, we find that

$$f(x, y) = 4 + (x - 1)^2 + (y - 3)^2.$$

Since $(x - 1)^2, (y - 3)^2 \geq 0$, we have $f(x, y) \geq 4$ for all values of x, y. Therefore $f(1, 3) = 4$ is a local minimum, and in fact it is the absolute minimum of f. This can be confirmed geometrically from the graph of f which is the elliptic paraboloid with vertex shown in the figure geven below.

Example 55.

A function with no extreme values: Investigate the extreme values of $f(x, y) = xy$.

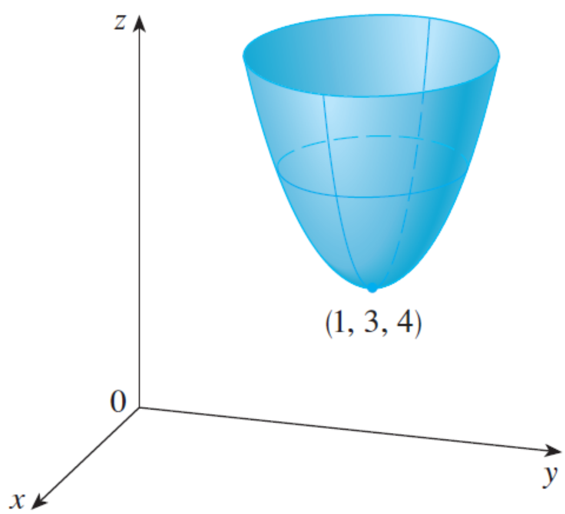


Figure 32: The paraboloid $f(x, y) = x^2 + y^2 - 2x - 6y + 14$

This example illustrates the fact that a function need not have a maximum or minimum value at a critical point. Figure 33 shows how this is possible. The graph of f is the hyperbolic paraboloid $z = xy$, which has a horizontal tangent plane $(z = 0)$ at the origin. You can see that $f(0, 0) = 0$ is a maximum in the direction of the line $y = x$ but a minimum in the direction of the line $y = -x$. Near the origin the graph has the shape of a saddle and so the point $(0, 0)$is called a *saddle point* of f.

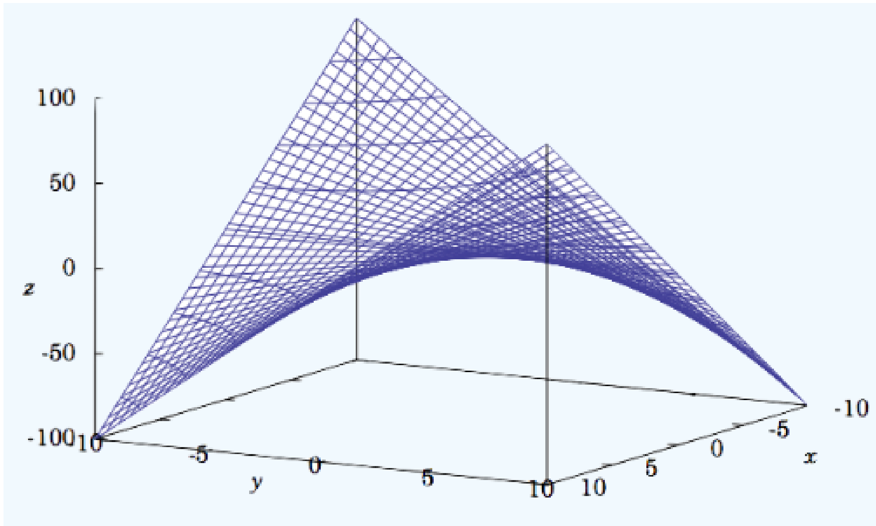


Figure 33: The hyperbolic paraboloid $f(x, y) = xy$

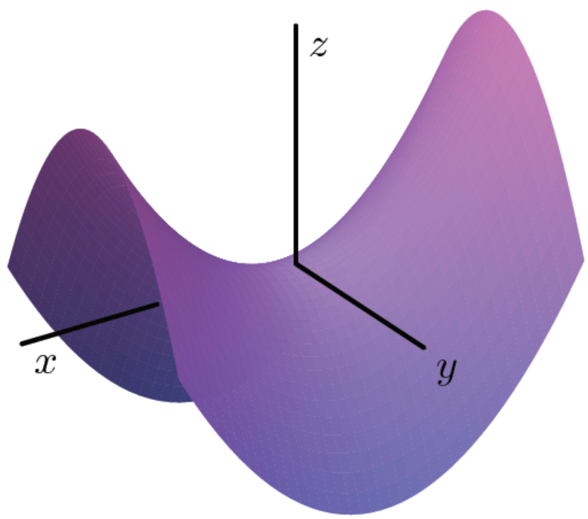


Figure 34: (0,0): saddle point

Saddle points

Given the function $z = f(x, y)$, the point $(a, b, f(a, b))$ is a **saddle point** if there are two distinct vertical planes through this point such that the intersection of the surface with one of the planes has a relative maximum at (a, b) and the intersection wtih the other has a relative minimum at (a, b).

In other words, the point $(a, b, f(a, b))$ is a saddle point if both

$f_x(a,b) = 0$ and $f_y(a,b) = 0$, but f does not have a local extremum at (a,b).

The Second Derivative Test.
Suppose the second partial derivatives of f are continuous on a disk with center (a,b) and suppose that $f_x(a,b) = 0$ and $f_y(a,b) = 0$. Let D be the quantity defined by

$$D = f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2.$$

1. If $D > 0$ and $f_{xx}(a,b) < 0$, then f has a local maximum at (a,b).
2. If $D > 0$ and $f_{xx}(a,b) > 0$, then f has a local minimum at (a,b).
3. If $D < 0$, then f has a saddle point at (a,b).
4. If $D = 0$, then this test yields no information about what happens at (a,b).

The quantity D is called the **discriminant** of the function f at (a,b).

Remark

It's helpful to write D as a determinant:

$$D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - (f_{xy})^2.$$

The matrix

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

is called the **Hessian matrix** of f.

Example 56.

Find the local maximum and minimum values and saddle points of

$$f(x, y) = x^4 + y^4 - 4xy + 1.$$

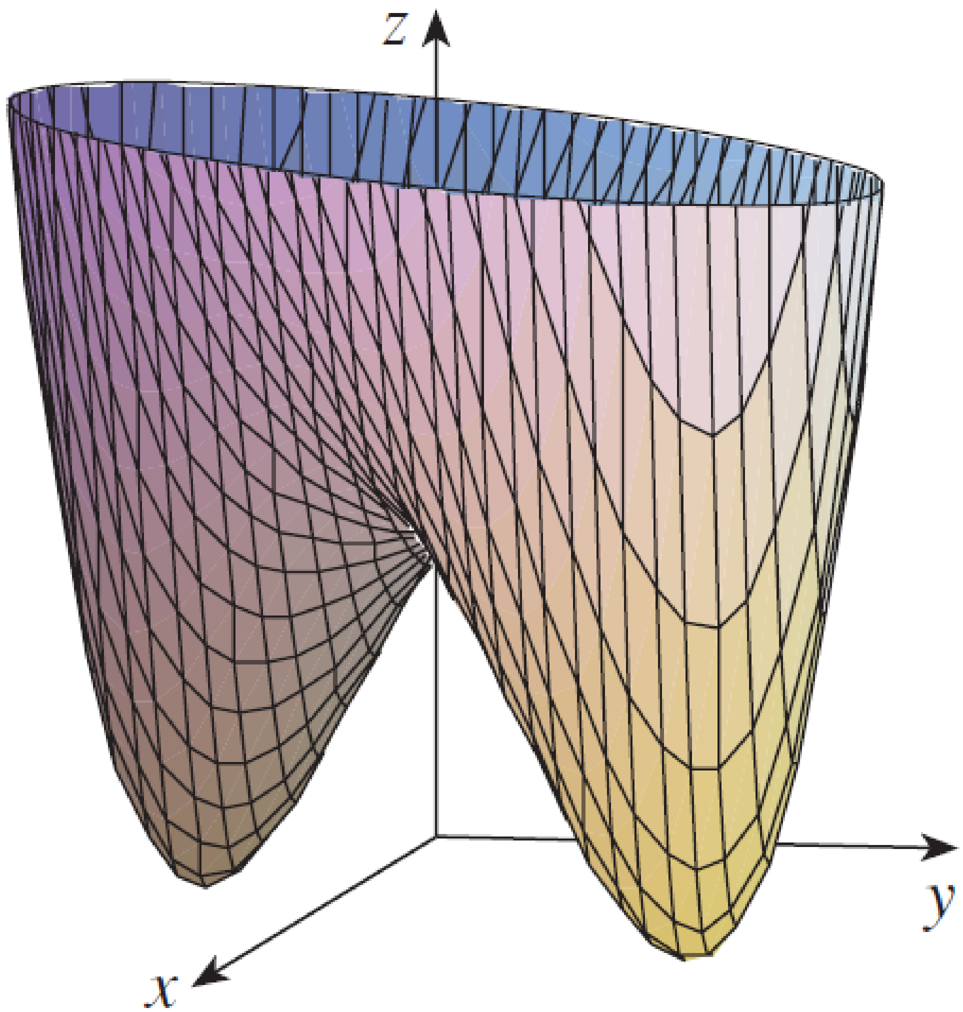


Figure 35: $f(x, y) = x^4 + y^4 - 4xy + 1$

Example 57.

Find all local maxima and minima of

$$f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}.$$

Solution. First find the critical points, i.e. where $\nabla f = 0$. Since

$$\frac{\partial f}{\partial x} = 2x(1 - (x^2 + y^2))e^{-(x^2+y^2)}$$
$$\frac{\partial f}{\partial y} = 2y(1 - (x^2 + y^2))e^{-(x^2+y^2)},$$

then the critical points are $(0, 0)$ and all points (x, y) on the unit circle $x^2 + y^2 = 1$.

Now, the second-order partial derivatives are:

$$\frac{\partial^2 f}{\partial x^2} = 2[1 - (x^2 + y^2) - 2x^2 - 2x^2(1 - (x^2 + y^2))]e^{-(x^2+y^2)}$$
$$\frac{\partial^2 f}{\partial y^2} = 2[1 - (x^2 + y^2) - 2y^2 - 2y^2(1 - (x^2 + y^2))]e^{-(x^2+y^2)}$$
$$\frac{\partial^2 f}{\partial y \partial x} = -4xy[2 - (x^2 + y^2)]e^{-(x^2+y^2)}$$

At $(0, 0)$, we have $D = 4 > 0$ and $\frac{\partial^2 f}{\partial x^2}(0, 0) = 2 > 0$, so $(0, 0)$ is a local minimum. However, for points (x, y) on the unit circle $x^2 + y^2 = 1$, we have

$$D = (-4x^2 e^{-1})(-4y^2 e^{-1}) - (-4xye^{-1})^2 = 0$$

and so the test fails.

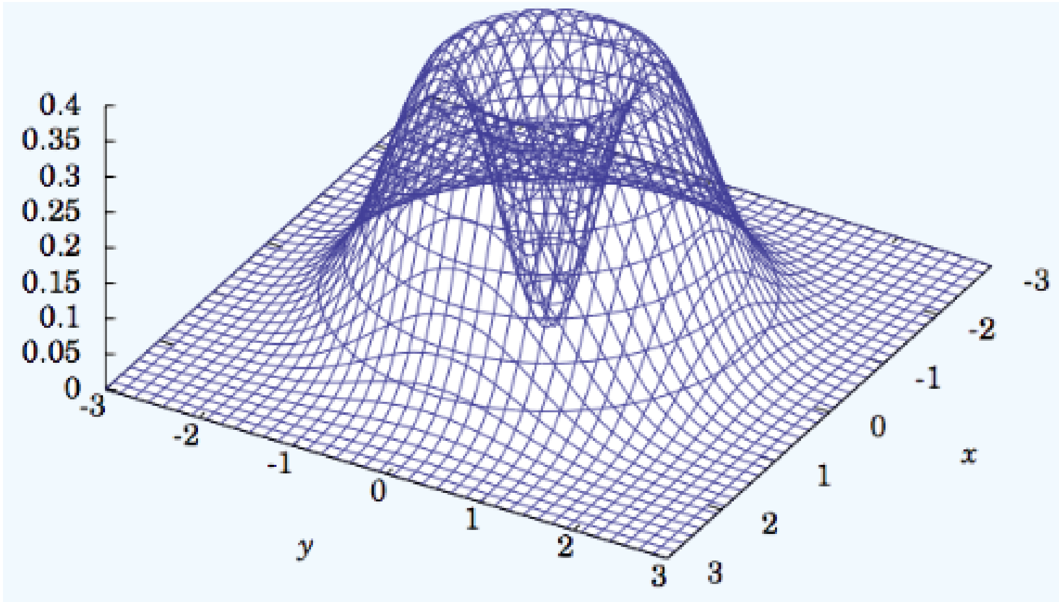


Figure 36: $f(x,y) = (x^2 + y^2)e^{-(x^2+y^2)}$

If we look at the graph of $f(x,y)$, as shown in the above figure, it looks like we might have a local maximum for (x,y) on the unit circle $x^2 + y^2 = 1$. If we switch to using polar coordinates (r,θ) instead of (x,y) in $\mathbb{R}^2$, where $r^2 = x^2 + y^2$, then we see that we can write $f(x,y)$ as a function $g(r)$ of the variable r alone:

$$g(r) = r^2 e^{-r^2}.$$

Then

$$g'(r) = r^2(1 - r^2)e^{-r^2},$$

so it has a critical point at $r = 1$, and we can check that

$$g''(1) = -4e^{-1} < 0,$$

so the Second Derivative Test from single-variable calculus says that $r = 1$ is a local maximum. But $r = 1$ corresponds to the unit circle $x^2 + y^2 = 1$. Thus, the points (x,y) on the unit circle $x^2 + y^2 = 1$ are local maximum points for f .

Example 58.

Find and classify all the critical points of $f(x, y) = 4 + x^3 + y^3 - 3xy$.

Solution. We first need all the first order (to find the critical points) and second order (to classify the critical points) partial derivatives so let's get those.

$f_x = 3x^2 - 3y$ $\qquad$ $f_y = 3y^2 - 3x$

$f_{xx} = 6x$ $\qquad$ $f_{yy} = 6y$ $\qquad$ $f_{xy} = -3$

$$\begin{aligned} D(x, y) &= f_{xx}(x, y)f_{yy}(x, y) - f_{xy}(x, y)^2 \\ &= 6x \times 6y - (-3)^2 \\ &= 36xy - 9. \end{aligned}$$

Let's first find the critical points. Critical points will be solutions to the system of equations:

$$\begin{aligned} f_x &= 3x^2 - 3y = 0 \\ f_y &= 3y^2 - 3x = 0. \end{aligned}$$

This is a non-linear system of equations and these can, on occasion, be difficult to solve. However, in this case it's not too bad. We can solve the first equation for y as follows:

$$3x^2 - 3y = 0 \Rightarrow y = x^2.$$

Plugging this into the second equation gives,

$$3x^4 - 3x = 3x(x^3 - 1) = 0$$

From this we can see that we must have $x = 0$ or $x = 1$. Now use the fact that $y = x^2$ to get the critical points.

$$x = 0 \Rightarrow y = 0$$
$$x = 1 \Rightarrow y = 1.$$

So, we get two critical points:

$$(0,0), (1,1).$$

All we need to do now is classify them. To do this we will need the sign of D at critical points. We have

$$D(0,0) = -9 < 0.$$

So, the critical point (0,0) must be a saddle point.

We also have

$$D(1,1) = 36(1)(1) - 9 = 27 > 0$$

and

$$f_{xx}(1,1) = 6(1) = 6 > 0.$$

Therefore, f has a local minimum at (1,1).

Thus,

Saddle point at $(0,0)$,

Relative minimum at $(1,1)$.

Example 59. Find and classify all the critical points for

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

Example 60. Find the shortest distance from the point $(1,0,-2)$ to the plane $x + 2y + z = 4$.

Solution. Here, we need to first come up with the equation that we are going to have to work with.

First, let's suppose that (x, y, z) is any point on the plane. The distance between this point and the point in question, $(1, 0, -2)$, is given by the formula,

$$d = \sqrt{(x-1)^2 + y^2 + (z+2)^2}.$$

What we are then asked to find is the minimum value of d. The point (x, y, z) that gives the minimum value of d will be the point on the plane that is closest to $(1, 0, -2)$.

There are a couple of issues with this equation. <u>First</u>, it is a function of x, y and z and we can only deal with functions of x and y at this point. However, this is easy to fix. We can solve the equation of the plane to see that,

$$z = 4 - x - 2y.$$

Plugging this into the distance formula gives

$$\begin{aligned} d &= \sqrt{(x-1)^2 + y^2 + (4 - x - 2y + 2)^2} \\ &= \sqrt{(x-1)^2 + y^2 + (6 - x - 2y)^2} \end{aligned}$$

Now, the next issue is that there is a square root in this formula and we know that we're going to be differentiating this eventually. So, in order to make our life a little easier let's notice that finding the minimum value of d will be equivalent to finding the minimum value of d^2.

So, let's instead find the minimum value of

$$d^2 = (x-1)^2 + y^2 + (6 - x - 2y)^2.$$

We can minimize d by minimizing the simpler expression

$$f(x, y) = (x-1)^2 + y^2 + (6 - x - 2y)^2.$$

So, let's go through the process. We have

$$\begin{aligned}
f_x &= 2(x-1) - 2(6-x-2y) = 4x + 4y - 14 = 0 \\
f_y &= 2y - 4(6-x-2y) = 4x + 10y - 24 = 0.
\end{aligned}$$

Solving these equations, we find that the only critical point is $(11/6, 5/13)$.

Since

$$f_{xx} = 4, f_{xy} = 4, f_{yy} = 10,$$

we have

$$\begin{aligned}
&D(x,y) = f_{xx}f_{yy} - (f_{xy})^2 = 24 > 0, \\
&f_{xx} > 0,
\end{aligned}$$

so by the Second Derivatives Test f has a local minimum at $(11/6, 5/13)$. Intuitively, we can see that this local minimum is actually an absolute minimum because there must be a point on the given plane that is closest to (1,0,-2).

Therefore,

$$\begin{aligned}
d &= \sqrt{(x-1)^2 + y^2 + (6-x-2y)^2} \\
&= \sqrt{(5/6)^2 + (5/3)^2 + (5/6)^2} = \frac{5}{6}\sqrt{6}.
\end{aligned}$$

Example 61.

A rectangular box without a lid is to be made from 12 m of cardboard. Find the maximum volume of such a box.

Solution. Let the length, width, and height of the box (in meters) be x, y, and z. Then the volume of the box is

$$V = xyz.$$

We can express V as a function of just two variables and by using the fact that the area of the four sides and the bottom of the box is

$$2xz + 2yz + xy = 12 \Rightarrow z = \frac{12 - xy}{2(x+y)}.$$

Thus,

$$V = xy\frac{12 - xy}{2(x+y)} = \frac{12xy - x^2y^2}{2(x+y)}.$$

Now, let's go through the process. We have ...

Example 62.

Find the local extrema of

$$f(x, y) = x^3 + x^2y - y^2 - 4y.$$

Solution. **Step 1:** Find the critical points. The derivative of f is

$$Df(x, y) = \begin{bmatrix} 3x^2 + 2xy & x^2 - 2y - 4 \end{bmatrix}.$$

$Df(x, y) = [0 \quad 0]$ means both components must be zero simultaneously. We need

$$x(3x + 2y) = 0 \qquad (1)$$

and

$$x^2 - 2y - 4 = 0. \qquad (2)$$

We need to solve two equations for the two unknowns x and y.

Equation (1) is satisfied if either $x = 0$ or if $3x + 2y = 0$, i.e., if $x = 0$ or if $y = -3x/2$. We consider these two solutions as two separate cases. For each case, we will find soultions for equation (2).

Case 1: Let $x = 0$. Then we know equation (1) is satisfied. We plug $x = 0$ into equation (2), which becomes $0 - 2y - 4 = 0$, i.e., $y = -2$. If $x = 0$ and $y = -2$, then both equation (1) and equation (2) are satisfied. Therefore the point $(0, -2)$ is a critical point.

Case 2: Let $y = -3x/2$. Then we know that Equation (1) is satisfied. We plug $y = -3x/2$ into Equation (2) and simplify:

$$\begin{aligned} & x^2 - 2(-3x/2) - 4 = 0 \\ \Rightarrow\ & x^2 + 3x - 4 = 0 \\ \Rightarrow\ & (x-1)(x+4) = 0 \\ \Rightarrow\ & x = 1 \text{ or } x = -4. \end{aligned}$$

So, we have two solutions of equation (2) for case 2. The first solution is when $x = 1$, which means $y = -3x/2 = -3/2$. If $x = 1$ and $y = -3/2$, then both equation (1) and equation (2) are satisfied. Therefore the point $(1, -3/2)$ is a critical point.

The second solution for case 2 is when $x = -4$, which means $y = -3x/2 = 6$. Therefore, the point $(-4, 6)$ is a critical point.

To summarize the results from both case 1 and case 2, we conclude that $f(x, y)$ has three critical points: $(0, -2)$, $(1, -3/2)$, and $(-4, 6)$.

Step 2: Classify the critical points.

The Hessian matrix is

$$Hf(x,y) = \begin{bmatrix} 6x+2y & 2x \\ 2x & -2 \end{bmatrix}$$

We need to check the definiteness of the $Hf(x,y)$ at the critical points $(0,-2)$, $(1,-3/2)$, and $(-4,6)$.

<u>For the critical point $(0,-2)$,</u>

$$Hf(0,-2) = \begin{bmatrix} -4 & 0 \\ 0 & -2 \end{bmatrix}$$

$h_{11} = -4 < 0$ and $\det(Hf) = 8 > 0$. This means $Hf(0,-2)$ is negative definite and f has a local maximum at $(0,-2)$.

<u>For the critical point $(1,-3/2)$,</u>

$$Hf(-4,6) = \begin{bmatrix} -12 & -8 \\ -8 & -2 \end{bmatrix}.$$

$h_{11} = 3 > 0$ and $\det(Hf) = -6-4 = -10 < 0$. This means $Hf(1,-3/2)$ is indefinite and f has a saddle at $(1,-3/2)$.

<u>For the critical point $(-4,6)$,</u>

$$Hf(-4,6) = \begin{bmatrix} -12 & -8 \\ -8 & -2 \end{bmatrix}.$$

$h_{11} = -12 < 0$ and $\det(Hf) = 24-64 = -40 < 0$. This means $Hf(-4,6)$ is indefinite and f has a saddle at $(-4,6)$.

Example 63. Identify the local extrama of

$$f(x,y) = (x^2+y^2)e^{-y}.$$

Solution. **Step 1:** *Find the critical points.*

The derivative of f is

$$Df(x,y) = \begin{bmatrix} 2xe^{-y} & (2y - x^2 - y^2)e^{-y} \end{bmatrix}$$

$Df(x,y) = [0 \quad 0]$ means that $2x = 0$ and $2y - x^2 - y^2 = 0$, i.e., $x = 0$ and $y(2-y) = 0$.

The critical points are therefore $(0,0)$ and $(0,2)$.

Step 2: *Classify the critical points.*

The Hessian matrix is

$$Hf(x,y) = \begin{bmatrix} 2e^{-y} & -2xe^{-y} \\ -2xe^{-y} & (2 - 4y + y^2 + x^2)e^{-y} \end{bmatrix}$$

At the critical point $(0,0)$

$$Hf(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$h_{11} = 2 > 0$ and $\det(Hf) = 4 > 0$, so $(0,0)$ is a local mininum.

At the critical point $(0,2)$

$$Hf(0,2) = \begin{bmatrix} e^{-2} & 0 \\ 0 & -2e^{-2} \end{bmatrix}$$

$h_{11} = e^{-2} > 0$ and $\det(Hf) = -2e^{-4} < 0$ so $(0,2)$ is a saddle point.

9 Lagrange multipliers

Consider a function $f(x,y) = \sqrt{x^2 + y^2}$. Clearly, it attains its minimum at $(x,y) = (0,0)$. But what if we are allowed to take points (x, y) satisfying the following equation:

$$g(x, y) = 2x + 3y - 6 = 0?$$

Thus, our problem is

$$\begin{array}{l} \text{Minimize } f(x, y) = \sqrt{x^2 + y^2} \\ \text{subject to } g(x, y) = 2x + 3y - 6 = 0. \end{array}$$

Here, $f(x, y)$ is called an **objective function** and $g(x, y)$ is called a **constraint function** (or simply, **constraint**).

Recall that

- Each element in the gradient ∇f is one of the function's first order partial derivatives.
- The gradient ∇f always points in the direction of the function's steepest slope at a given point.

Take a look at the drawing below. It illustrates how gradients work for a two-variable function of x_1 and x_2.

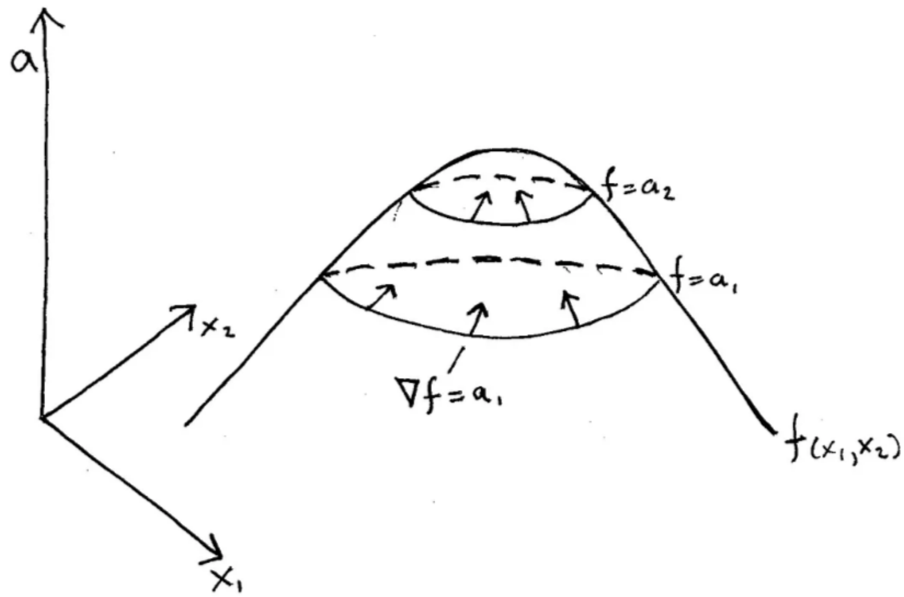


Figure 37

The function f in the figure forms a hill. Toward the peak there are two regions where we hold the height of f constant at some level a. These are level curves of f, and they're marked $f = a_1$, and $f = a_2$.

Imagine yourself standing on one of those level curves. Think of a hiking trail on a mountainside. Standing on the trail,

in what direction is the mountain steepest?

Clearly

the steepest direction is straight up the hill,
perpendicular to the trail.

In the drawing, these paths of steepest ascent are marked with arrows. These are the gradients ∇f at various points along the level curves. Just as the steepest hike is always perpendicular to our trail, the gradients of f are always perpendicular to its level curves.

That's the key idea here:

level curves are where $f = a$ and $\nabla f \perp f = a$.

How the Method Works

To see how Lagrange multipliers work, take a look at the drawing below. The function f is redrawn from above, along with a constraint $g = c$. In the drawing, the constraint is a plane that cuts through our hillside. A couple level curves of f are drawn as well. Our goal here is to climb as high on the hill as we can,

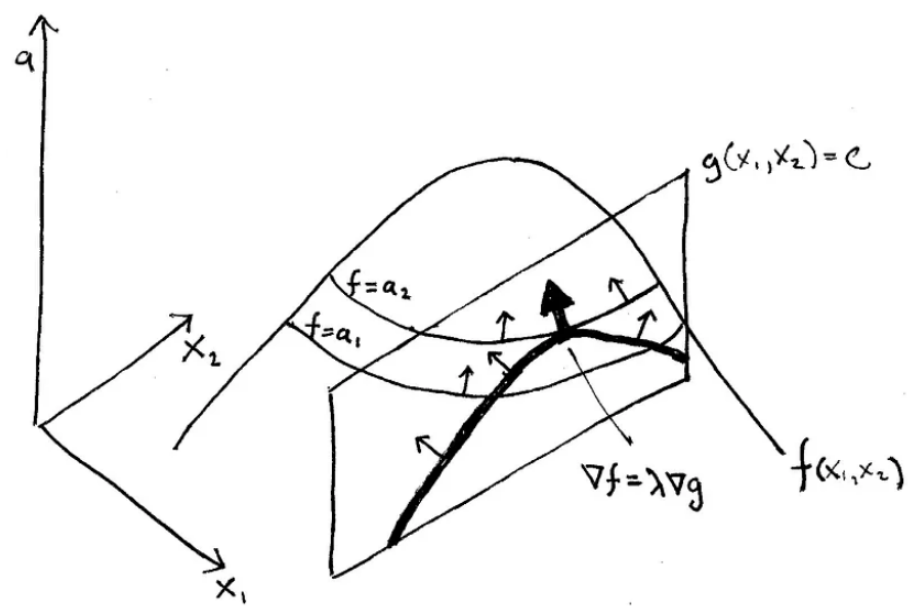


Figure 38

given that we can't move any higher than where the constraint $g = c$ cuts the hill.

In the drawing, the boundary where the constraint cuts the function is marked with a heavy line. Along that line are the highest points we can reach without stepping over our constraint.

That's an obvious place to start looking for a constrained maximum.

Imagine hiking from left to right on the constraint line. As we gain elevation, we walk through various level curves of f. I've marked two in the picture. At each level curve, imagine checking its slope – that is, the slope of a tangent line to it – and comparing that to the slope on the constraint where we're standing.

If our slope is greater than the level curve, we can reach a higher point on the hill if we keep moving right. If our slope is less than the level curve – say, toward the right where our constraint line is declining – we need to move backward to the left to reach a higher point.

When we reach a point where the slope of the constraint line just equals the slope of the level curve, we've moved as high as we can. That is, we've reached our constrained maximum. Any movement from that point will take us downhill. In the figure, this point is marked with a large arrow pointing toward the peak.

At that point, the level curve $f = a_2$ and the constraint have the same slope. That means they're parallel and point in the same direction. That is, the gradients of f and g both point in the same direction, and differ at most by a scalar. Let's call that scalar "lambda." Then we have

$$\nabla f = \lambda \nabla(g).$$

We have the following theorem.

Theorem 9.1. [**Lagrange Multipliers**] Assume that $f(x, y)$ and $g(x, y)$ are differentiable functions. If $f(x, y)$ has a local minimum or a local maximum on the constraint curve $g(x, y) = 0$ at $P = (a, b)$, and if $\nabla g_P \neq 0$, then there is a scalar λ such that

$$\nabla f_P(x, y) = \lambda \nabla g_P(x, y). \tag{13}$$

Proof.

Let $c(t)$ be a parametrization of the constraint curve $g(x, y) = 0$ near P, chosen so that

$$c(0) = P \text{ and } c'(0) \neq 0.$$

Then

$$f(c(0)) = f(P),$$

and by assumption, $f(c(t))$ has a local min or max at $t = 0$. Thus, $t = 0$ is a critical point of $f(c(t))$ and by the Chain Rule

$$\frac{d}{dt} f(c(0)) = \nabla f_P \cdot c'(0) = 0$$

This shows that ∇f_P is orthogonal to the tangent vector $c'(0)$ to the curve $g(x, y) = 0$.

The gradient $\nabla g_P$is also orthogonal to $c'(0)$ (because ∇g_P is orthogonal to the level curve $g(x, y) = 0$ at P). We conclude that ∇f_P and ∇g_P are parallel, and hence ∇f_P is a multiple of ∇g_P as claimed. ◀

We refer to Equation (13) as the **Lagrange condition**. When we write this condition in terms of components, we obtain the **Lagrange equations**:

$$f_x(a,b) = \lambda g_x(a,b)$$
$$f_y(a,b) = \lambda g_y(a,b)$$

A point $P = (a,b)$ satisfying these equations is called a **critical point** for the optimization problem with constraint and $f(a,b)$ is called a **critical value**.

Remark: The system of equations from the method actually has three equations. To see this let's take the first equation and put in the definition of the gradient vector to see what we get.

$$(f_x, f_y) = \lambda(g_x, g_y) = (\lambda g_x, \lambda g_y)$$

In order for these two vectors to be equal the individual components must also be equal. So, we actually have two equations here.

$$f_x = \lambda g_x, \quad f_y = \lambda g_y.$$

These two equations along with the constraint

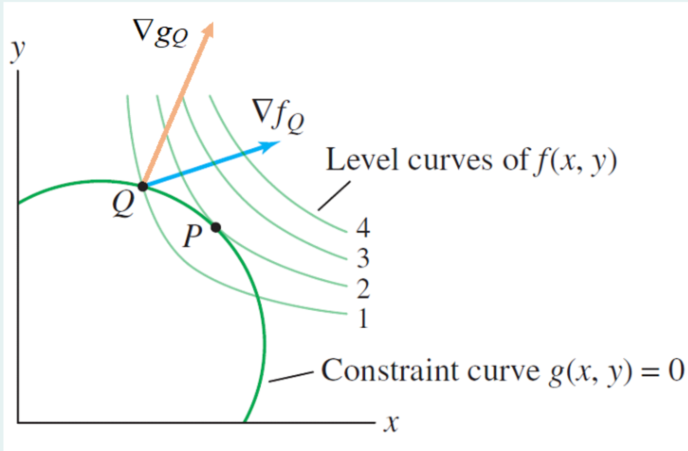
$$g(x,y) = c,$$

give three equations with three unknowns x, y, and λ.

How does the method of Lagrange multipliers work?

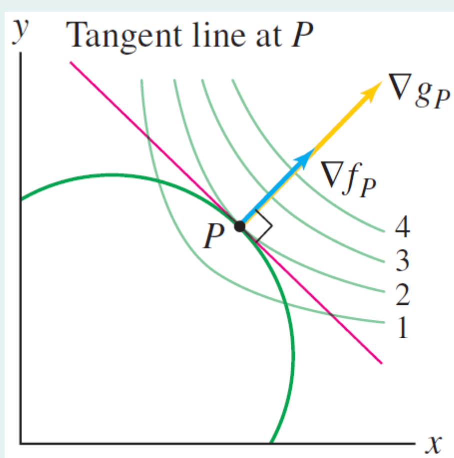
Graphical insight

Imagine standing at point Q in the figure given below.



We want to increase the value of f while remaining on the constraint curve. The gradient vector ∇f_Q points in the direction of maximum increase, but we cannot move in the gradient direction because that would take us off the constraint curve. However, the gradient points to the right, and so we can still increase f somewhat by moving to the right along the constraint curve.

We keep moving to the right until we arrive at the point P, where ∇f_P is orthogonal to the constraint curve as in the figure given below.



Once at P, we cannot increase f further by moving either to the right or to the left along the constraint curve. Thus $f(P)$

is a local maximum subject to the constraint.
Now, the vector ∇g_P is also orthogonal to the constraint curve, so ∇f_P and ∇g_P must point in the same or opposite directions. In other words,

$$\nabla g_P = \lambda \nabla g_P$$

for some scalar λ (called a Lagrange multiplier). Graphically, this means that a local max subject to the constraint occurs at points P where the level curves of f and g are tangent.

So the bottom line is that

Conclusion

The method of Lagrange multipliers is really just an algorithm that finds where the gradient of a function points in the same direction as the gradients of its constraints, while also satisfying those constraints.

Problem-Solving Strategy

1. Determine the objective function $f(x, y)$ and the constraint function $g(x, y)$. Does the optimization problem involve maximizing or minimizing the objective function?

2. Set up the following system of equations:

$$\begin{aligned} \nabla f(x, y) &= \lambda \nabla g(x, y) \\ g(x, y) &= 0, \end{aligned}$$

that is,

$$\begin{aligned} f_x(x, y) &= \lambda g_x(x, y) \\ f_x(x, y) &= \lambda g_x(x, y) \\ g(x, y) &= 0. \end{aligned}$$

3. Solve for x and y.

4. The largest of the values of f at the solutions found in step 3 maximizes f; the smallest of those values minimizes f.

Example 64. Find the maximum and minimum of the function $x^2 - 10x - y^2$ on the ellipse whose equation is $x^2 + 4y^2 = 16$.

Solution. For this problem the objective function is

$$f(x, y) = x^2 - 10x - y^2$$

and the constraint function is

$$g(x, y) = x^2 + 4y^2 - 16.$$

To apply the method of Lagrange multipliers we need ∇f and ∇g. So we start by computing the first order partial derivatives of these functions.

$$f_x = 2x - 10 \qquad f_y = -2y \qquad g_x = 2x \qquad g_y = 8y$$

So, according to the method of Lagrange multipliers, we need to find all solutions to

$$\begin{aligned} 2x - 10 &= \lambda(2x) \\ -2y &= \lambda(8y) \\ x^2 + 4y^2 - 16 &= 0 \end{aligned}$$

Rearranging these equations gives

$$(\lambda - 1)x = -5 \tag{E1}$$

$$(4\lambda + 1)y = 0 \tag{E2}$$

$$x^2 + 4y^2 - 16 = 0 \tag{E3}$$

From (E2), we see that $\lambda = -\frac{1}{4}$ or $y = 0$.

- If $\lambda = -\frac{1}{4}$, (E1) gives $-\frac{5}{4}x = -5$, i.e. $x = 4$.
- If $y = 0$, then $(E3)$ gives $x = \pm 4$.

So the method of Lagrange multipliers gives that the only possible locations of the maximum and minimum of the function f are $(4, 0)$ and $(-4, 0)$. To complete the problem, we only have to compute f at those points.

point	$(4, 0)$	$(-4, 0)$
value of f	-24	56
	min	max

Hence the maximum value of $x^2 - 10x - y^2$ on the ellipse is 56 and the minimum value is -24.

In the previous example, the objective function and the constraint were specified explicitly. That will not always be the case. In the next example, we have to do a little geometry to extract them.

Example 65. Find the rectangle of largest area (with sides parallel to the coordinates axes) that can be inscribed in the ellipse $x^2 + 2y^2 = 1$.

Solution. Since this question is so geometric, it is best to start by drawing a picture.

Call the coordinates of the upper right corner of the rectangle (x, y), as in the figure above. The four corners of the rectangle are $(\pm x, \pm y)$ so the rectangle has width $2x$ and height $2y$ and we have the problem:

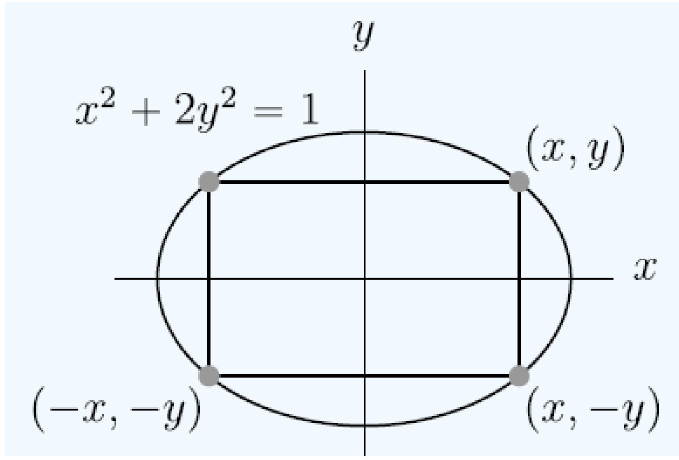


Figure 39

Maximize $f(x, y) = 4xy$
subject to $g(x, y) = x^2 + 2y^2 - 1$.

Again, to use Lagrange multipliers we need the first order partial derivatives.

$$f_x = 4y \qquad f_y = 4x \qquad g_x = 2x \qquad g_y = 4y$$

So, according to the method of Lagrange multipliers, we need to find all solutions to

$$4y = \lambda(2x) \tag{E1}$$
$$4x = \lambda(4y) \tag{E2}$$
$$x^2 + 2y^2 - 1 = 0 \tag{E3}$$

Equation (E1) gives $y = \frac{1}{2}\lambda x$. Substituting this into equation (E2) gives

$$4x = 2\lambda^2 x \qquad \text{or} \qquad 2x\big(2 - \lambda^2\big) = 0.$$

So (E2) is satisfied if either $x = 0$ or $\lambda = \pm\sqrt{2}$.

- If $x = 0$, then (E1) gives $y = 0$ too. But $(0, 0)$ violates the constraint equation (E3). Note that, to have a solution, *all* of the equations (E1), (E2) and (E3) must be satisfied.

- If $\lambda = \sqrt{2}$, then
  - (E2) gives $x = \sqrt{2}y$ and then
  - (E3) gives $2y^2 + 2y^2 = 1$ or $y^2 = \frac{1}{4}$ so that
  - $y = \pm\frac{1}{2}$ and $x = \sqrt{2}y = \pm\frac{1}{\sqrt{2}}$.
- If $\lambda = -\sqrt{2}$, then
  - (E2) gives $x = -\sqrt{2}y$ and then
  - (E3) gives $2y^2 + 2y^2 = 1$ or $y^2 = \frac{1}{4}$ so that
  - $y = \pm\frac{1}{2}$ and $x = -\sqrt{2}y = \mp\frac{1}{\sqrt{2}}$.

We now have four possible values of (x, y), namely $\big(\frac{1}{\sqrt{2}}, \frac{1}{2}\big)$, $\big(-\frac{1}{\sqrt{2}}, -\frac{1}{2}\big)$, $\big(\frac{1}{\sqrt{2}}, -\frac{1}{2}\big)$ and $\big(-\frac{1}{\sqrt{2}}, \frac{1}{2}\big)$. They are the four corners of a single rectangle. We said that we wanted (x, y) to be the upper right corner, i.e. the corner in the first quadrant. It is $\big(\frac{1}{\sqrt{2}}, \frac{1}{2}\big)$.

Example 66. Find the ends of the major and minor axes of the ellipse $3x^2 - 2xy + 3y^2 = 4$. They are the points on the ellipse that are farthest from and nearest to the origin.

Solution. The equation $3x^2 - 2xy + 3y^2 = 4$ represents an ellipse with center at the origin. Let (x, y) be a point on it. This point is at the end of a major axis when it maximizes its distance from the centre, $(0, 0)$ of the ellipse. It is at the end of a minor axis when it minimizes its distance from $(0, 0)$. So we wish to Maximize and minimize the distance

$$\sqrt{x^2 + y^2}.$$

It is equivalent to maximizing/minimizing its square

$$\left(\sqrt{x^2+y^2}\right)^2 = x^2 + y^2$$

and we are free to choose the objective function. So, our problem is

Maximize and minimize the distance

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

subject to the constraint

$$g(x,y) = 3x^2 - 2xy + 3y^2 - 4 = 0.$$

Now maximizing/minimizing $\sqrt{x^2+y^2}$ is equivalent to maximizing/minimizing its square $\left(\sqrt{x^2+y^2}\right)^2 = x^2+y^2$. So we are free to choose the objective function

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

which we will do, because it makes the derivatives cleaner. Again, we use Lagrange multipliers to solve this problem, so we start by finding the partial derivatives.

$$\begin{aligned} f_x(x,y) &= 2x \quad f_y(x,y) = 2y \\ g_x(x,y) &= 6x - 2y \quad g_y(x,y) = -2x + 6y \end{aligned}$$

We need to find all solutions to

$$\begin{aligned} 2x &= \lambda(6x - 2y) \\ 2y &= \lambda(-2x + 6y) \\ 3x^2 - 2xy &+ 3y^2 - 4 = 0 \end{aligned}$$

Dividing the first two equations by 2, and then collecting together the x's and the y's gives

$$
\begin{aligned}
(1-3\lambda)x+\lambda y&=0 && \text{(E1)}\\
\lambda x+(1-3\lambda)y&=0 && \text{(E2)}\\
3x^2-2xy+3y^2-4&=0 && \text{(E3)}
\end{aligned}
$$

To start, let's concentrate on the first two equations. Pretend, for a couple of minutes, that we already know the value of λ and are trying to find x and y.

Note that λ cannot be zero because if it is, (E1) forces $x=0$ and (E2) forces $y=0$ and $(0,0)$ is not on the ellipse, i.e. violates (E3). So we may divide by λ and (E1) gives

$$y=-\frac{1-3\lambda}{\lambda}x$$

Subbing this into (E2) gives

$$\lambda x-\frac{(1-3\lambda)^2}{\lambda}x=0$$

Again, x cannot be zero, since then $y=-\frac{1-3\lambda}{\lambda}x$ would give $y=0$ and $(0,0)$ is still not on the ellipse.

So we may divide $\lambda x-\frac{(1-3\lambda)^2}{\lambda}x=0$ by x, giving

$$
\begin{aligned}
\lambda-\frac{(1-3\lambda)^2}{\lambda}=0 &\Leftrightarrow (1-3\lambda)^2-\lambda^2=0\\
&\Leftrightarrow 8\lambda^2-6\lambda+1=(2\lambda-1)(4\lambda-1)=0
\end{aligned}
$$

We now know that λ must be either $\frac{1}{2}$ or $\frac{1}{4}$. Subbing these into either (E1) or (E2) gives

$$\begin{aligned}
\lambda = \frac{1}{2} \quad &\Longrightarrow \quad -\frac{1}{2}x + \frac{1}{2}y = 0 \quad \Longrightarrow \quad x = y \\
&\Longrightarrow \quad 3x^2 - 2x^2 + 3x^2 = 4 \quad \Longrightarrow \quad x = \pm 1 \\
\lambda = \frac{1}{4} \quad &\Longrightarrow \quad \frac{1}{4}x + \frac{1}{4}y = 0 \quad \Longrightarrow \quad x = -y \\
&\Longrightarrow \quad 3x^2 + 2x^2 + 3x^2 = 4 \quad \Longrightarrow \quad x = \pm\frac{1}{\sqrt{2}}
\end{aligned}$$

We now have $(x, y) = \pm(1, 1)$, from $\lambda = \frac{1}{2}$, and $(x, y) = \pm\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ from $\lambda = \frac{1}{4}$. The distance from $(0, 0)$ to $\pm(1, 1)$, namely $\sqrt{2}$, is larger than the distance from $(0, 0)$ to $\pm\big(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\big)$, namely 1. So the ends of the minor axes are $\pm\big(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\big)$ and the ends of the major axes are $\pm(1, 1)$. Those ends are sketched in the figure on the left below. Once we have the ends, it is an easy matter to sketch the ellipse as in the figure on the right below.

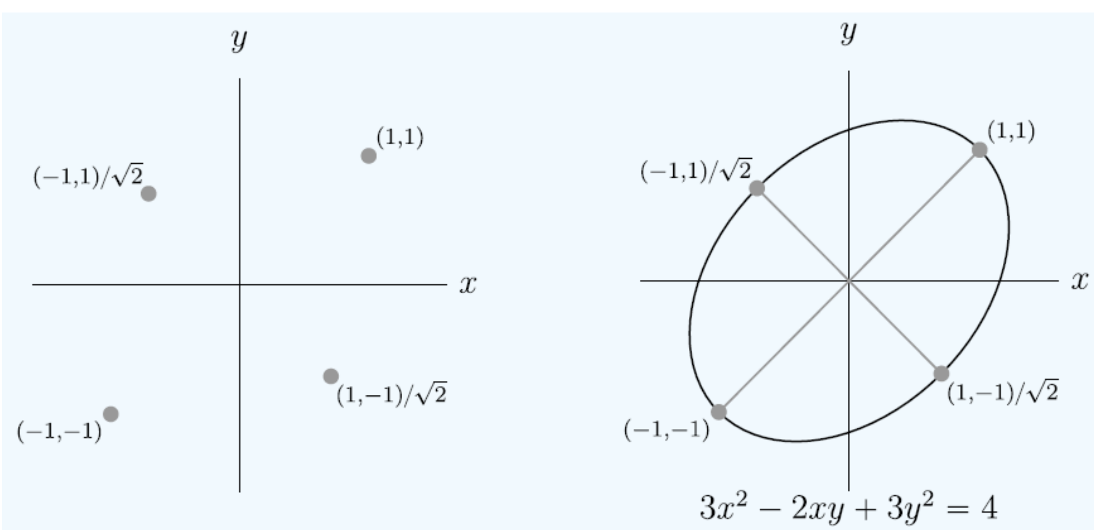


Figure 40

Method of Lagrange Multipliers for functions of three variables.

Method of Lagrange Multipliers To find the maximum and minimum values of $f(x, y, z)$ subject to the constraint $g(x, y, z) = k$ [assuming that these extreme values exist and $\nabla g \neq 0$ on the surface $g(x, y, z) = k$]:

(a) Find all values of x, y, z and λ solving the system of equations:

$$\begin{aligned} &\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \\ &g(x, y, z) = k. \end{aligned}$$

(b) Evaluate f at all the points (x, y, z) that result from step (a). The largest of these values is the maximum value of f, the smallest is the minimum value of f.

Notice that

- The system of equations from the method actually has four equations. To see this let's take the first equation and put in the definition of the gradient vector to see what we get.

  $$(f_x, f_y, f_z) = \lambda(g_x, g_y, g_z) = (\lambda g_x, \lambda g_y, \lambda g_z)$$

  In order for these two vectors to be equal the individual components must also be equal. So, we actually have three equations here.

  $$f_x = \lambda g_x \quad f_y = \lambda g_y \quad f_z = \lambda g_z$$

  These three equations along with the constraint

  $$g(x, y, z) = c,$$

give four equations with four unknowns x, y, z, and λ.

- If we only have functions of two variables then we won't have the third component of the gradient and so will only have three equations in three unknowns x, y, and λ.

Proof. Suppose that (a, b, c) is a point of S and that $f(x, y, z) \geq f(a, b, c)$ for all points (x, y, z) on S that are close to (a, b, c). That is (a, b, c) is a local minimum for f on S. Of course, the argument for a local maximum is virtually identical.
Imagine that we go for a walk on S, with the time t running, say, from $t = -1$ to $t = +1$ and that at time $t = 0$ we happen to be exactly at (a, b, c). Let's say that our position is $\big(x(t), y(t), z(t)\big)$ at time t. (We are always on S, so $g\big(x(t), y(t), z(t)\big) = 0$ for all t.)
Write

$$F(t) = f\big(x(t), y(t), z(t)\big)$$

So $F(t)$ is the value of f that we see on our walk at time t. Then for all t close to 0, the point $\big(x(t), y(t), z(t)\big)$ is close to

$$\big(x(0), y(0), z(0)\big) = (a, b, c)$$

so that

$$\begin{aligned} F(0) = f\big(x(0), y(0), z(0)\big) &= f(a, b, c) \\ \leq f\big(x(t), y(t), z(t)\big) &= F(t) \end{aligned}$$

for all t close to zero. So $F(t)$ has a local minimum at $t = 0$ and consequently $F'(0) = 0$.

By the chain rule,

$$\begin{aligned}F'(0) &= \frac{d}{dt}f\big(x(t),y(t),z(t)\big)\Big|_{t=0}\\ &= f_x\big(a,b,c\big)x'(0)+f_y\big(a,b,c\big)y'(0)+f_z\big(a,b,c\big)z'(0)\\ &= 0\end{aligned}$$

We may rewrite this as a dot product:

$$\begin{aligned}0 = F'(0) &= \nabla f(a,b,c)\cdot (x'(0)\,,\ y'(0)\,,\ z'(0))\\ \Longrightarrow \nabla f(a,b,c) &\perp (x'(0)\,,\ y'(0)\,,\ z'(0))\end{aligned}$$

This is true for all paths on S that pass through (a,b,c) at time 0. In particular it is true for all vectors $(x'(0)\,,\ y'(0)\,,\ z'(0))$ that are tangent to S at (a,b,c). So $\nabla f(a,b,c)$ is perpendicular to S at (a,b,c).

But we already know that $\nabla g(a,b,c)$ is also perpendicular to S at (a,b,c). So $\nabla f(a,b,c)$ and $\nabla g(a,b,c)$ have to be parallel vectors. That is,

$$\nabla f(a,b,c) = \lambda \nabla g(a,b,c)$$

for some number λ. That's the Lagrange multiplier rule of our theorem. ◀

Example 67. Find the point on the sphere $x^2+y^2+z^2=1$ farthest from $(1,2,3)$.

Solution. As before, we simplify the algebra by maximizing the square of the distance rather than the distance itself. So we are to maximize

$$f(x,y,z) = (x-1)^2+(y-2)^2+(z-3)^2$$

subject to the constraint

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

We have

$$\begin{aligned} &f_x = 2(x-1), \quad f_y = 2(y-2), \quad f_z = 2(z-3) \\ &g_x = 2x, \qquad g_y = 2y, \qquad g_z = 2z. \end{aligned}$$

We need to find all solutions to the system:

$$\begin{aligned} &2(x-1) = \lambda(2x) \Leftrightarrow x = \frac{1}{1-\lambda} \\ &2(y-2) = \lambda 2y \Leftrightarrow y = \frac{2}{1-\lambda} \\ &2(z-3) = \lambda 2z \Leftrightarrow z = \frac{3}{1-\lambda} \\ &x^2 + y^2 + z^2 - 1 = 0. \end{aligned}$$

Solving x, y, z from the first three equations, we get

$$\frac{1+4+9}{(1-\lambda)^2} - 1 = 0 \Rightarrow 1 - \lambda = \pm\sqrt{14}.$$

We can then substitute these two values of λ back into the expressions for in terms of x, y, z to get the two points

$$\frac{1}{\sqrt{14}}(1,2,3), \qquad -\frac{1}{\sqrt{14}}(1,2,3).$$

We thus obtain two vectors:

one from $\frac{1}{\sqrt{14}}(1,2,3)$ to $(1,2,3)$ –

$$\left[1 - \frac{1}{\sqrt{14}}\right](1,2,3)$$

and the other one from $-\frac{1}{\sqrt{14}}(1,2,3)$ to $(1,2,3)$ –

$$\left[1+\frac{1}{\sqrt{14}}\right](1,2,3).$$

Clearly, the first vector is shorter than the second one. Hence the nearest point is $\frac{1}{\sqrt{14}}(1,2,3)$ and the farthest point is $-\frac{1}{\sqrt{14}}(1,2,3)$.