MULTI VARIABLE CALCULUS

1. Introduction

It is now known to science that there are many more dimensions than the classical four. Scientists say that these don't normally impinge on the world because the extra dimensions are very small and curve in on themselves, and that since reality is fractal most of it is tucked inside itself. This means either that the universe is more full of wonders than we can hope to understand or, more probably, that scientists make things up as they go along. Terry Pratchett

Motivation :Multivariable calculus (also known as multivariate calculus) is the extension of calculus in one variable to calculus with functions of several variables: the differentiation and integration of functions involving several variables, rather than just one.
We have already studied function of one variable and their calculus. You developed knowledge of calculus of functions of type

$$y = f(x), \quad x \in I$$

where I is an interval in $\mathbb{R}$. In these lectures, we extend these ideas to functions of many variables. In particular, we will learn limits, continuity, derivatives and their properties.

2. Functions of Several Variable

A function of two variables (x, y) maps each ordered pair (x, y) in a subset D of the real plane $\mathbb{R}^2$ to a unique real number z. The set D is called the domain of the function. The function of the type

$$f(x, y) = z, \quad x \in D,$$

where $D(\subseteq \mathbb{R}^2)$ is the domain of f and f is real-valued. we write

$$f : D(\subseteq \mathbb{R}^2) \to \mathbb{R}$$

Let us illustrate with the following examples :

$$f(x, y) = x \cos y + 5xy \sin x, \qquad D = \mathbb{R}^2$$

$$f(x, y) = \sqrt{9 - x^2 - y^2}, \qquad D = \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 \leq 9\}$$

$$f(x, y) = \frac{1}{2x - y} \qquad D = \{(x, y) \in \mathbb{R}^2;\ y \neq 2x\}$$

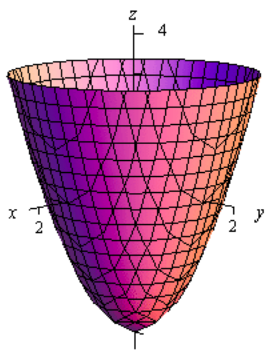
Graphing functions of two variables

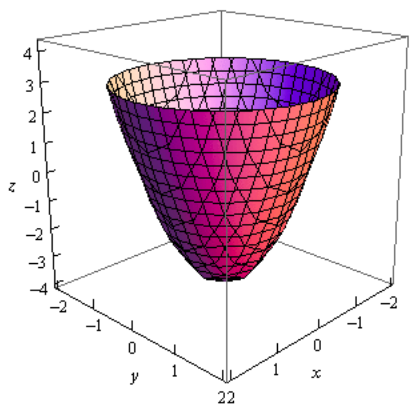
Suppose we wish to graph the function $z = f(x, y)$. This function has two independent variables (x and y) and one dependent variable (z). When graphing a function $y = f(x)$ of one variable, we use the Cartesian plane. We are able to graph any ordered pair (x, y) in the plane, and every point in the plane has an ordered pair (x, y) associated with it. With a function of two variables, each ordered pair (x, y) the domain of the function is mapped to a real number z. Therefore, the graph of the function f consists of ordered triples (x, y, z).

Definition 2.1. Surface The graph of a function $z = f(x, y)$ two variables is called a surface. In fact more than two variable is also called surface.

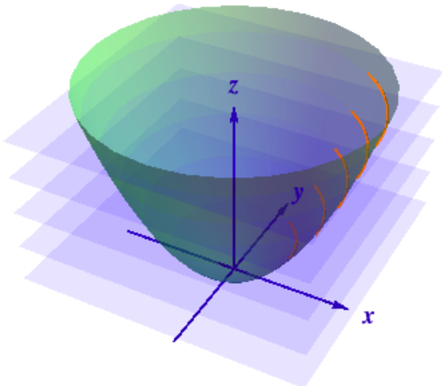
To understand more completely the concept of plotting a set of ordered triples to obtain a surface in three-dimensional space, imagine the (x, y) coordinate system laying flat. Then, every point in the domain of the function f has a unique z-value associated with it. If z is positive, then the graphed point is located above the xy-plane, if z is negative, then the graphed point is located below the xy-plane. The set of all the graphed points becomes the two-dimensional surface that is the graph of the function f.

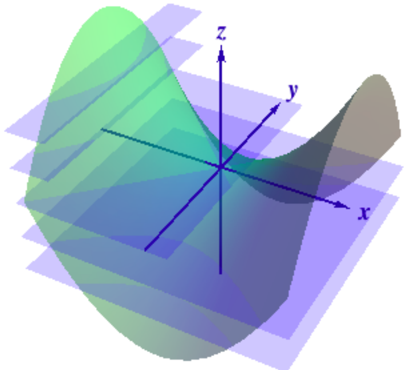
Example 2.1. *First, remember that graphs of functions of two variables, $z = f(x, y)$ are surfaces in three dimensional space. For example, here is the graph of $z = 2x^2 + 2y^2 - 4$*





Definition 2.2. Level Curves Given a function $z = f(x, y)$ and a number c in the range of f a level curve of a function of two variables for the value c is defined to be the set of points satisfying the equation $z = f(x, y) = c$.





We have now examined functions of more than one variable and seen how to graph them. In this section, we see how to take the limit of a function of more than one variable, and what it means for a function of more than one variable to be continuous at a point in its domain. It turns out these concepts have aspects that just don't occur with functions of one variable.

3. Limit of a Function of Two Variables

Recall the $(\varepsilon - \delta)$ definition of a limit of a function of one variable.
Before we can adapt this definition to define a limit of a function of two variables, we first need to see how to extend the idea of an open interval in one variable to an open interval in two variables.

Definition 3.1. Disk centered at (x_0, y_0) of radius $\delta > 0$ Consider a point (x_0, y_0) in $\mathbb{R}^2$. A disk centered at point (x_0, y_0) of radius $\delta > 0$, is defined to be

$$\{(x, y) \in \mathbb{R}^2 : (x - x_0)^2 + (y - y_0)^2 < \delta^2\} \tag{1}$$

Remark 3.1. The idea of a $\delta > 0$ disk appears in the definition of the limit of a function of two variables. If δ is small, then all the points (x_0, y_0) in the δ disk are close to (x_0, y_0). This is

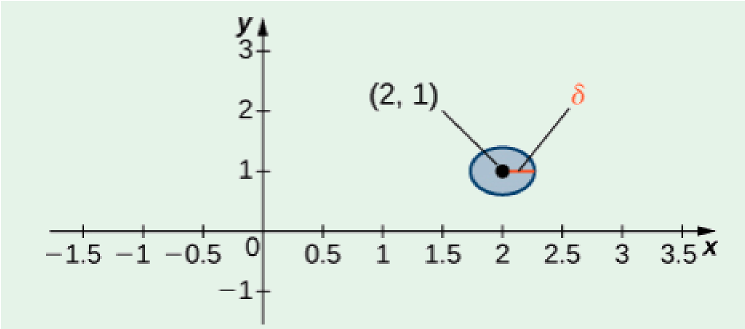


Figure 2.3.1: A δ disk centered around the point (2, 1).

completely analogous to x being close to a in the definition of a limit of a function of one variable. In one dimension, we express this restriction as

$$a - \delta < x < a + \delta$$

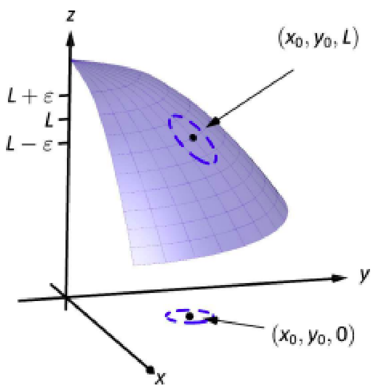
In more than one dimension, we use a $\delta > 0$ disk.

Definition 3.2. Limit of a function of two variables Let $f(x, y)$ be a function of two variables, x and y. The limit of $f(x, y)$ as (x, y) approaches (x_0, y_0) is L, written

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$$

if for each $\varepsilon > 0$ there exists a small enough $\delta > 0$ (depending on $\varepsilon > 0$) such that for all points (x, y) in a δ disk around (x_0, y_0) except possibly for (x_0, y_0) itself, the value of $f(x, y)$ no more than $\varepsilon > 0$ away from L.
In symbols, we write the following: for every $\varepsilon > 0$, there exists a $\delta > 0$ (depending on $\varepsilon > 0$) such that for all $(x, y) \in \mathbb{R}^2$ with

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta \quad \text{implies} \quad |f(x, y) - L| < \varepsilon \tag{2}$$



Remark 3.2. Let f be a real valued function defined on $D \subseteq \mathbb{R}^2$, except possibly at (x_0, y_0). We say that the limit of f is $L \in \mathbb{R}$ as (x, y) approaches to (x_0, y_0) (or at $(x, y) = (x_0, y_0)$), written $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$, if for every sequence $\{(x_n, y_n)\}$ in D with $(x_n, y_n) \neq (x_0, y_0)$ for all n and $(x_n, y_n) \to (x_0, y_0)$, we have $f(x_n, y_n) \to L$, as $n \to \infty$.

Remark 3.3. One can show as in the case of one variable calculus, the limit is unique. This means that the limit is independent of the path $(x_n, y_n) \to (x_0, y_0)$ or $(x, y) \to (x_0, y_0)$. This plays an important role in the existence of limits. We shall illustrate in the examples.

Example 3.1. *Let* $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ *defined by*

$$f(x,y) = \frac{4xy^2}{x^2+y^2}$$

Compute the limit at $(x_0, y_0) = (0,0)$ *using* $\varepsilon - \delta$ *definition. One can see immediately,*

$$|f(x,y) - 0| = \frac{4|x|y^2}{x^2+y^2} \leq \frac{2(x^2+y^2)\sqrt{(x^2+y^2)}}{x^2+y^2} = 2\sqrt{x^2+y^2}$$

Therefore when $\sqrt{x^2+y^2} < \varepsilon/2 = \delta_\varepsilon$*, then we will have*

$$|f(x,y) - 0| < \varepsilon,$$

Hence the limit $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Example 3.2. *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *defined by*

$$f(x,y) = \begin{cases} \frac{x^3-y^3}{x-y} \ x \neq y \\ 0 \ x = y \end{cases}$$

Compute the limit at $(x_0, y_0) = (0,0)$.
One can see immediately,

$$|f(x,y) - 0| = |\frac{x^3-y^3}{x-y}| \leq |(x^2+xy+y^2)| \leq |(x^2+2|x||y|+y^2)| = (|x|+|y|)^2 \leq 2(|x|^2+|y|^2)$$

Therefore when $\sqrt{x^2+y^2} < \sqrt{\varepsilon/2} = \delta_\varepsilon$ *with* $x \neq y$*, then we will have*

$$|f(x,y) - 0| < \varepsilon,$$

and when $x = y$*, then in that case also when* $\sqrt{x^2+y^2} = \sqrt{2x^2} < \sqrt{\varepsilon/2} = \delta_\varepsilon$*, then we will have*

$$|f(x,x) - 0| = |0 - 0| < \varepsilon,$$

in any case. Hence the limit $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Example 3.3. *Let* $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ *defined by*

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

Compute the limit at $(x,y) = (0,0)$*. We shall use polar coordinate :* $x = r\cos\theta, y = r\sin\theta$*, with* $r > 0$ *and* $0 \leq \theta < 2\pi$*. Then*

$$|f(x,y) - 0| := |f(r,\theta) - 0| = |r\cos^3\theta| \leq r \to 0, \text{as } r \to 0.$$

Hence the limit $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Or you can prove using $\varepsilon - \delta$ *definition.*

Remark 3.4. Algebra of Limits of functions We have

$$\lim_{(x,y)\to(x_0,y_0)} (f(x,y) \pm g(x,y)) = \lim_{(x,y)\to(x_0,y_0)} f(x,y) \pm \lim_{(x,y)\to(x_0,y_0)} g(x,y)$$
$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)g(x,y) = \lim_{(x,y)\to(x_0,y_0)} f(x,y) \lim_{(x,y)\to(x_0,y_0)} g(x,y)$$
$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)}{g(x,y)} = \frac{\lim_{(x,y)\to(x_0,y_0)} f(x,y)}{\lim_{(x,y)\to(x_0,y_0)} g(x,y)} \quad \text{if} \lim_{(x,y)\to(x_0,y_0)} g(x,y) \neq 0$$

Example 3.4. *Let $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ defined by*

$$f(x,y) = \frac{2x^2y}{x^4+y^2}$$

Compute the limit at $(x,y) = (1,1)$. Then

$$\lim_{(x,y)\to(1,1)} f(x,y) = \frac{\lim_{(x,y)\to(1,1)} 2x^2y}{\lim_{(x,y)\to(1,1)}(x^4+y^2)} = \frac{2}{2} = 1$$

Hence the limit $\lim_{(x,y)\to(0,0)} f(x,y) = 1$.

3.1. **Limits that fail to exist/ Dependent on Path.**

Example 3.5. *Let $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ defined by*

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

Required to prove $\lim_{(x,y)\to(0,0)} f(x,y)$ exists.
If we consider (x,y) is approaching to $(0,0)$ along the line $y = 0$ that is x-axis, Then

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,0)\to(0,0)} f(x,0) = 0.$$

If we consider (x,y) is approaching to $(0,0)$ along the line $y = mx$ then,

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

So it depends on m. So the limit does not exist.

Example 3.6. *Evaluate the limits $f : \mathbb{R}^2 \setminus \{(2,1)\} \to \mathbb{R}$ defined by*

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

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

If we consider the (x,y) approaching to $(2,1)$ along the line $y = k(x-2)+1$, then

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

Since the answer depends on k the limit fails to exist.

Example 3.7. *$f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ defined by*

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

Required to prove $\lim_{(x,y)\to(0,0)} f(x,y)$ exists.
If we consider (x,y) is approaching to $(0,0)$ along the line $y = 0$ that is x-axis, Then

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,0)\to(0,0)} f(x,0) = 0.$$

If we consider (x, y) is approaching to $(0, 0)$ along the line $y = mx^2$ then,

$$\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{(x,mx^2)\to(0,0)} f(x, mx^2) = \lim_{(x,mx^2)\to(0,0)} \frac{2mx^4}{(1 + m^2)x^4} = \frac{2m}{1 + m^2}.$$

So it depends on m. So the limit does not exist

4. Continuity of a function in two variables

In Continuity, we defined the continuity of a function of one variable and saw how it relied on the limit of a function of one variable. In particular, three conditions are necessary for $f(x)$ to be continuous at point $x = a$ if
1) $f(a)$ exists.
2) $\lim_{x\to a} f(x)$ exists and
3) $\lim_{x\to a} f(x) = f(a)$.

These three conditions are necessary for continuity of a function of two variables as well.

Definition 4.1. A function $f(x, y)$ is continuous at a point (x_0, y_0) in its domain if the following conditions are satisfied:
1) $f(x_0, y_0)$ exists.
2) $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ exists and
3) $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$.

Remark 4.1. Or you can check whether $\lim_{(h,k)\to(0,0)} f(x_0 + h, y_0 + k) = f(x_0, y_0)$ to see whether a function $f(x, y)$ is continuous at a point (x_0, y_0).

Example 4.1. *Let D be any subset of $\mathbb{R}^2$. So the map $f : D \to \mathbb{R}$ such that $f(x, y) = c\ \forall(x, y) \in D$ is continuous.*

Example 4.2. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

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

So for all (x, y) with $\sqrt{x^2 + y^2} = \delta = 2\varepsilon$, then

$$|f(x, y) - f(0, 0)| = |f(x, y) - 0| = |f(x, y)| = \frac{|xy|}{\sqrt{x^2 + y^2}} \leq \frac{1}{2}\sqrt{x^2 + y^2} < \varepsilon$$

Example 4.3. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

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

Required to prove $\lim_{(x,y)\to(0,0)} f(x, y) = f(0, 0) = 0$.
If we consider (x, y) is approaching to $(0, 0)$ along the line $y = x$, Then

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

So f is not continuous at $(0, 0)$.

Example 4.4. *Consider the function* $f : \mathbb{R}^2 \to \mathbb{R}$ *defined by*

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Required to prove $\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0) = 0.$
If we consider (x,y) *is approaching to* $(0,0)$ *along the line* $y-$*axis, Then*

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,0)\to(0,0)} \frac{x^2.0}{x^4+0} = 0.$$

If we consider (x,y) *is approaching to* $(0,0)$ *along the line* $x-$*axis, Then*

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

If we consider (x,y) *is approaching to* $(0,0)$ *along the line* $y = mx$*, Then*

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,mx)\to(0,0)} \frac{mx^3}{x^4+(mx)^2} = \lim_{x\to 0} \frac{mx}{x^2+m^2} = 0.$$

But if we consider (x,y) *is approaching to* $(0,0)$ *along the parabola* $y = x^2$*, Then*

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,x^2)\to(0,0)} \frac{x^4}{x^4+x^4} = \frac{1}{2} \neq 0 = f(0,0).$$

f *is not continuous at* $(x,y) = (0,0)$.

Remark 4.2. Algebra of continuous functions If f and g are two continuous functions at (x_0, y_0). We have

$$\lim_{(x,y)\to(x_0,y_0)} (f(x,y) \pm g(x,y)) = f(x_0,y_0) \pm g(x_0,y_0)$$
$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)g(x,y) = f(x_0,y_0)g(x_0,y_0)$$
$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)}{g(x,y)} = \frac{f(x_0,y_0)}{g(x_0,y_0)} \quad \text{if } g(x_0,y_0) \neq 0$$

5. Partial Derivatives

Now that we have examined limits and continuity of functions of two variables, we can proceed to study derivatives. Finding derivatives of functions of two variables is the key concept in this chapter, with as many applications in mathematics, science, and engineering as differentiation of single-variable functions. However, we have already seen that limits and continuity of multi variable functions have new issues and require new terminology and ideas to deal with them. This carries over into differentiation as well.

Derivatives of a Function of Two Variables When studying derivatives of functions of one variable, we found that one interpretation of the derivative is an instantaneous rate of change of y as a function of x. Leibnitz notation for the derivative is $\frac{df}{dx}$ which implies that y is the dependent variable and x is the independent variable. For a function $z = f(x, y)$ two variables, x and y are the independent variables and z is the dependent variable. This raises two questions right away: How do we adapt Leibniz notation for functions of two variables? Also, what is an interpretation of the derivative? The answer lies in partial derivatives.

Definition 5.1. Let $z = f(x, y)$ a function of two variables.
Then the partial derivative of f with respect to x, written as $\frac{\partial f}{\partial x}$ or f_x is defined as

$$\frac{\partial f}{\partial x}(x,y) = f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h} \tag{3}$$

Or the partial derivative of f at (x_0, y_0) with respect to x or the first variable

$$\frac{\partial f}{\partial x}(x_0,y_0) = f_x(x_0,y_0) = \lim_{h\to 0} \frac{f(x_0+h,y_0) - f(x_0,y_0)}{h} \tag{4}$$

The partial derivative of f with respect to y, written as $\frac{\partial f}{\partial y}$ or f_y is defined as

$$\frac{\partial f}{\partial y}(x,y) = f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h} \tag{5}$$

Or the partial derivative of f at (x_0, y_0) with respect to y or the first variable

$$\frac{\partial f}{\partial y}(x_0,y_0) = f_y(x_0,y_0) = \lim_{k\to 0} \frac{f(x_0,y_0+k) - f(x_0,y_0)}{k} \tag{6}$$

Remark 5.1. This definition shows two differences already. First, the notation changes, in the sense that we still use a version of Leibniz notation, but the d in the original notation is replaced with the symbol ∂. (This rounded "d" is usually called "partial," so $\frac{\partial f}{\partial x}$ spoken as the "partial of f with respect to x") This is the first hint that we are dealing with partial derivatives. Second, we now have two different derivatives we can take, since there are two different independent variables. Depending on which variable we choose, we can come up with different partial derivatives altogether, and often do.

Example 5.1. *Use the definition of the partial derivative as a limit to calculate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for the function*

$$f(x, y) = x^2 - 3xy + 2y^2 - 4x + 5y - 12$$

Now $f(x + h, y) = x^2 - 3xy + 2y^2 - 4x + 5y - 12 + 2xh + h^2 - 3yh - 4h$. Hence

$$\begin{aligned} f(x+h,y) - f(x,y) &= h(2x + h - 3y - 4) \\ \frac{f(x+h,y) - f(x,y)}{h} &= 2x + h - 3y - 4 \\ \frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h} &= 2x - 3y - 4 \end{aligned}$$

Geometrical interpretation Partial Derivatives can also be interpreted as rates of change. If $z = f(x, y)$ then $\frac{\partial f}{\partial x}$ is the rate of change of z with respect to x when y is fixed. Similarly, $\frac{\partial f}{\partial y}$ is the rate of change of z with respect to y when x is fixed.

or else

Example 5.2. *Consider the function $z = f(x, y) = x^2 + 2y^2$, as graphed in Figure 12.11(a). By fixing $y = 2$, we focus our attention to all points on the surface where the y-value is 2, shown in both parts (a) and (b) of the figure. These points form a curve in space: $z = f(x, 2) = x^2 + 8$ which is a function of just one variable. We can take the derivative of z with respect to x along this curve and find equations of tangent lines, etc.*

The key notion to extract from this example is: by treating y as constant (it does not vary) we can consider how z changes with respect to x. In a similar fashion, we can hold x constant and consider how z changes with respect to y. This is the underlying principle of partial derivatives. We state the formal, limit–based definition first, then show how to compute these partial derivatives without directly taking limits.

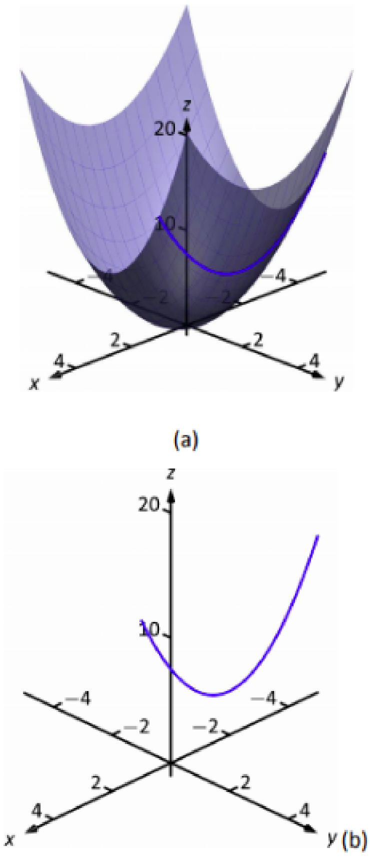


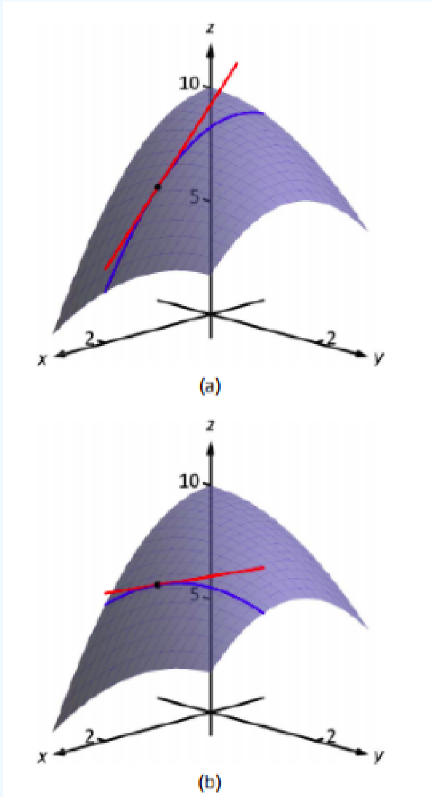
Figure 12.11: By fixing y=2, the surface $f(x, y) = x^2 + 2y^2$ is a curve in space.

Example 5.3. *Let $z = f(x, y) = -x^2 - \frac{1}{2}y^2 + xy + 10$. Find $f_x(2, 1)$ and $f_y(2, 1)$, and interpret their meaning.*
We begin by computing $f_x(x, y) = -2x + y$ and $f_y(x, y) = -y + x$.

$$f_x(2, 1) = -3 \quad \textit{and} \quad f_y(2, 1) = 1.$$

It is also useful to note that $f(2, 1) = 7.5$. Consider $f_x(2, 1) = -3$ along with Figure 12.12(a). If one "stands" on the surface at the point $(2, 1, 7.5)$ and moves parallel to the x-axis (i.e, only the x -value changes, not the y-value), then the instantaneous rate of change is -3 . Increasing the

x-value will decrease the z-value; decreasing the x-value will increase the z-value.
Now consider $f_y(2,1) = 1$, illustrated in Figure 12.12(b). Moving along the curve drawn on the surface, i.e., parallel to the y -axis and not changing the x-values, increases the z-value instantaneously at a rate of 1.



re 12.12: Illustrating the meaning of partial derivatives.

Example 5.4. Partial Derivatives may exist without being Continuous *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$f(x,y) = \{ \begin{array}{cc} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{array}$$

Now $\frac{f(h,0)-f(0,0)}{h} = 0$

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

Similarly $\frac{\partial f}{\partial x}(0,0) = 0$. *But f is not continuous at $(0,0)$. (Example 4.3)*

Example 5.5. Do partial derivatives always exist? *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = |x| + |y|$. The function is continuous at $(0,0)$:*

$$|f(x,y) - f(0,0)| = ||x| + |y| - 0| = |\sqrt{(|x| + |y|)^2}| \leq \sqrt{2}\sqrt{(x-0)^2 + (y-0)^2}$$

Hence

$$|f(x,y) - f(0,0)| < \varepsilon \quad whenever \quad \sqrt{(x-0)^2 + (y-0)^2} < \frac{\varepsilon}{2} = \delta_\varepsilon$$

As $\varepsilon > 0$ is arbitary, so The function is continuous at $(0,0)$. Now we have

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

does not exist. Similarly $\frac{\partial f}{\partial y}(0,0)$ does not exist.
So the function is continuous at $(0,0)$ although the partial derivatives does not exits.

Theorem 5.1. Sufficient condition for continuity: *Suppose one of the partial derivatives exist at (x_0, y_0) and the other partial derivative is bounded in a neighborhood of (x_0, y_0). Then $f(x,y)$ is continuous at (x_0, y_0).*

Proof. Let f_y exists at (x_0, y_0). Consider

$$\frac{f(x_0, y_0+k) - f(x_0, y_0)}{k} - f_y(x_0, y_0) := \varepsilon_1$$

then $|\varepsilon_1| \to 0$ as $k \to 0$. Since f_x exists and bounded in a neighborhood of at (x_0, y_0), So there exists a real number M such that $|f_x(u,v)| \le M$ for all $(u,v) \in N_\delta$. Hence for all $(x_0+h, y_0+k) \in N_\delta$, we consider $g(x) = f(x, y_0+k)$ in N_δ, and g is differentiable on (x_0+h, x_0) then applying Mean Value theorem in (x_0+h, x_0)

$$\begin{aligned}
|f(x_0+h, y_0+k) - f(x_0, y_0)| &= |f(x_0+h, y_0+k) - f(x_0, y_0+k) + f(x_0, y_0+k) - f(x_0, y_0)| \\
&= |hf_x(x_0+h\theta, y_0+k) + kf_y(x_0, y_0) + k\varepsilon_1| \\
&, \le |h|M + |k||f_y(x_0, y_0)| + |k||\varepsilon_1| \ \to 0 \text{ as } (h,k) \to (0,0).
\end{aligned}$$

So $\lim_{(h,k)\to(0,0)} f(x_0+h, y_0+k) = f(x_0, y_0)$. Hence $f(x,y)$ is continuous at (x_0, y_0). □

6. Differentiability

Let D be an open subset of $\mathbb{R}^2$. Then

Definition 6.1. A function $f(x,y): D \to \mathbb{R}$ is differentiable at a point (x_0, y_0) of D if there exist $\alpha = (\alpha_1, \alpha_2)$ and $\varepsilon_1 = \varepsilon_1(h,k)$, $\varepsilon_2 = \varepsilon_2(h,k)$ such that

$$f(x_0+h, y_0+k) - f(x_0,y_0) = h\alpha_1 + k\alpha_2 + h\varepsilon_1(h,k) + k\varepsilon_2(h,k) \tag{7}$$

where $\varepsilon_1(h,k), \varepsilon_2(h,k) \to 0$ as $(h,k) \to (0,0)$.

Theorem 6.1. *Suppose f is differentiable at a point (x_0,y_0). Then the partial derivatives $\frac{\partial f}{\partial x}$ or f_x and $\frac{\partial f}{\partial y}$ or f_y exist at (x_0,y_0). Then $\alpha = (\alpha_1,\alpha_2) = (f_x(x_0,y_0), f_y(x_0,y_0))$ in the above definition (7).*

Proof. Since f is differentiable at a point (x_0,y_0) of D then there exist $\alpha = (\alpha_1,\alpha_2)$ and $\varepsilon_1 = \varepsilon_1(h,k)$, $\varepsilon_2 = \varepsilon_2(h,k)$ such that

$$f(x_0+h, y_0+k) - f(x_0,y_0) = h\alpha_1 + k\alpha_2 + h\varepsilon_1(h,k) + k\varepsilon_2(h,k)$$

Substituting $k = 0$ in the above

$$f(x_0+h, y_0) - f(x_0,y_0) = h\alpha_1 + h\varepsilon_1(h,k)$$

which in turn implies

$$\lim_{h\to 0} \frac{f(x_0+h,y_0) - f(x_0,y_0)}{h} = \alpha_1$$

And by definition $f_x(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0) = \alpha_1$.

Similarly, we can prove $f_y(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = \alpha_2$ □

Remark 6.1. Suppose f is differentiable at a point (x_0,y_0). Then there exist $\varepsilon_1 = \varepsilon_1(h,k)$, $\varepsilon_2 = \varepsilon_2(h,k)$

$$f(x_0+h, y_0+k) - f(x_0,y_0) = hf_x(x_0,y_0) + kf_y(x_0,y_0) + h\varepsilon_1(h,k) + k\varepsilon_2(h,k). \tag{8}$$

where $\varepsilon_1, \varepsilon_2 \to 0$ as $(h,k) \to (0,0)$.

Theorem 6.2. *Suppose f is differentiable at a point (x_0,y_0), then f is continuous at (x_0,y_0).*

Proof. We have

$$\begin{aligned} |f(x_0+h,y_0+k) - f(x_0,y_0)| &= |hf_x(x_0,y_0) + kf_y(x_0,y_0) + h\varepsilon_1(h,k) + k\varepsilon_2(h,k)| \\ &\le |h||f_x(x_0,y_0)| + |k||f_y(x_0,y_0)| + |h||\varepsilon_1(h,k)| + |k||\varepsilon_2(h,k)| \to 0 \end{aligned}$$

as $(h,k) \to (0,0)$. Hence f is continuous at (x_0,y_0). □

Example 6.1. *Consider the function $f(x,y) = x^2 + y^2 + xy$. Then $f_x(0,0) = f_y(0,0) = 0$. Also*

$$f(h,k) - f(0,0) = h^2 + k^2 + hk = 0.h + 0.k + h\varepsilon_1(h,k) + k\varepsilon_2(h,k)$$

where $\varepsilon_1 = h+k$, and $\varepsilon_2 = k$. So $\varepsilon_1, \varepsilon_2 \to 0$ as $(h,k) \to (0,0)$. Therefore f is differentiable at $(0,0)$.

Example 6.2. *Show that the following function $f(x,y)$ is is not differentiable at $(0,0)$. Consider the function $f:\mathbb{R}^2\to\mathbb{R}$ defined by*

$$f(x,y)=\ \{\quad \begin{matrix} x\sin\frac{1}{x}+y\sin\frac{1}{y} & (x,y)\neq(0,0) \\ 0 & xy=0 \end{matrix}$$

Using the boundedness of $\sin$ *and* $\cos$*, we get $|f(x,y)|\leq|x|+|y|\leq 2\sqrt{x^2+y^2}$ implies that f is continuous at $(0,0)$.*
Also it is easy to prove $f_x(0,0)=f_y(0,0)=0$.

If $f(x,y)$ is differentiable at $(0,0)$ then there exist $\varepsilon_1=\varepsilon_1(h,k)$, $\varepsilon_2=\varepsilon_2(h,k)$ such that

$$f(h,k)-f(0,0)=h\varepsilon_1(h,k)+k\varepsilon_2(h,k)$$

and $\varepsilon_1,\varepsilon_2\to 0$ as $(h,k)\to(0,0)$. Taking $(h,k)\to(0,0)$ along the line $h=k$, we have

$$\begin{aligned} f(h,h)-f(0,0)&=h\varepsilon_1(h,h)+h\varepsilon_2(h,h)\\ 2h\sin\frac{1}{h}&=h(\varepsilon_1(h,h)+\varepsilon_2(h,h))\\ 2\sin\frac{1}{h}&=(\varepsilon_1(h,h)+\varepsilon_2(h,h))\to 0 \quad \textit{as } h\to 0 \end{aligned}$$

But $\sin\frac{1}{h}$ does not goes to 0 as $h\to 0$. In fact the limit does not exist. So our assumption is wrong. Hence $f(x,y)$ is not differentiable at $(0,0)$.

Example 6.3. *Show that the function $f(x,y)=\sqrt{|xy|}$ is not differentiable at the origin.*
Easy to check the continuity (take $\delta=\varepsilon$).
$f_x(0,0)=\lim_{h\to 0}\frac{0-0}{h}=0$, and similar calculation shows $f_y(0,0)=0$. So if f is differentiable at $(0,0)$, then there exist,ε_1 and ε_2 such that

$$f(h,k)-f(0,0)=h\varepsilon_1(h,k)+k\varepsilon_2(h,k)$$

and $\varepsilon_1,\varepsilon_2\to 0$ as $(h,k)\to(0,0)$. Taking $(h,k)\to(0,0)$ along the line $h=k$, we have

$$\begin{aligned} &f(h,h)-f(0,0)=h\varepsilon_1(h,h)+h\varepsilon_2(h,h)\\ &\Rightarrow\ |h|=h(\varepsilon_1(h,h)+\varepsilon_2(h,h))\\ &\Rightarrow\ \frac{|h|}{h}=\varepsilon_1(h,h)+\varepsilon_2(h,h)\to 0 \quad \textit{as } h\to 0 \end{aligned}$$

But $\frac{|h|}{h}$ does not goes to 0 as $h\to 0$. In fact the limit does not exist. So our assumption is wrong. Hence $f(x,y)$ is not differentiable at $(0,0)$.

6.1. **Equivalent condition for Differentiability. Notations:** 1. $\Delta f(x_0,y_0)=f(x_0+h,y_0+k)-f(x_0,y_0)$,the total variation of f.

2. $df(x_0,y_0)=hf_x(x_0,y_0)+kf_y(x_0,y_0)$, the total differential of f.

3. $\rho=\sqrt{h^2+k^2}$.

Theorem 6.3. Equivalent condition for differentiability: *f is differentiable at $(x_0,y_0)\iff \lim_{\rho\to 0}\frac{\Delta f(x_0,y_0)-df(x_0,y_0)}{\rho}=0$.*

Proof. Suppose f is differentiable at a point (x_0,y_0). Then then there exist $\varepsilon_1=\varepsilon_1(h,k)$, $\varepsilon_2=\varepsilon_2(h,k)$

$$f(x_0+h,y_0+k)-f(x_0,y_0)=hf_x(x_0,y_0)+kf_y(x_0,y_0)+h\varepsilon_1(h,k)+k\varepsilon_2(h,k).$$

and $\varepsilon_1, \varepsilon_2 \to 0$ as $(h,k) \to (0,0)$.

$$\Delta f(x_0,y_0) - df(x_0,y_0) = h\varepsilon_1(h,k) + k\varepsilon_2(h,k).$$

Now $\frac{|h|}{\sqrt{h^2+k^2}} \leq 1$ and $\frac{|k|}{\sqrt{h^2+k^2}} \leq 1$. So

$$\begin{aligned}
|\frac{\Delta f(x_0,y_0) - df(x_0,y_0)}{\rho}| &= |\frac{h}{\rho}\varepsilon_1(h,k) + \frac{k}{\rho}\varepsilon_2(h,k)| \\
&\leq \frac{|h|}{\rho}|\varepsilon_1(h,k)| + \frac{|k|}{\rho}|\varepsilon_2(h,k)| \\
. &\leq |\varepsilon_1(h,k)| + |\varepsilon_2(h,k)| \quad \to 0 \ \text{ as } \ (h,k) \to (0,0)
\end{aligned}$$

Now $\rho \to 0 \Leftrightarrow (h,k) \to (0,0)$, hence

$$\lim_{\rho \to 0} \frac{\Delta f(x_0,y_0) - df(x_0,y_0)}{\rho} = 0.$$

Conversely, let $\lim_{\rho \to 0} \frac{\Delta f(x_0,y_0) - df(x_0,y_0)}{\rho} = 0$. Required to prove f is differentiable at (x_0,y_0). Then

$$\begin{aligned}
\Delta f(x_0,y_0) &= df(x_0,y_0) + \rho\varepsilon(\rho) \qquad \text{where } \varepsilon(\rho) \to 0 \text{ as } \rho \to 0 \\
&= df(x_0,y_0) + \rho\varepsilon(h,k) \to 0 \text{ as } (h,k) \to (0,0).
\end{aligned}$$

since $\rho \to 0 \Leftrightarrow (h,k) \to (0,0)$. Now for $(h,k) \neq (0,0)$, we have

$$\begin{aligned}
\rho\varepsilon(h,k) &= \rho\varepsilon(h,k) \\
&= \sqrt{h^2+k^2}\varepsilon(h,k) \\
&\leq \sqrt{h^2+k^2+2|h||k|}\varepsilon(h,k) \\
&= (|h|+|k|)\varepsilon(h,k) \\
&= h(sgn(h)\varepsilon(h,k)) + k(sgn(k)\varepsilon(h,k))
\end{aligned}$$

So we take $\varepsilon_1(h,k) = \varepsilon(h,k)sgn(h)$ and $\varepsilon_2(h,k) = \varepsilon(h,k)sgn(k)$. And as $\varepsilon(\rho) = \varepsilon(h,k) \to 0$ as $\rho \to 0$, so are $\varepsilon_1(h,k)$ and $\varepsilon_2(h,k)$. □

Example 6.4. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

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

Partial derivatives exist at $(0,0)$ and $f_x(0,0) = 0$, $f_y(0,0) = 0$, so $df(0,0) = 0$. Hence

$$\Delta f(0,0) = f(h,k) - f(0,0) = \frac{h^2k^2}{h^2+k^2} = \frac{h^2k^2}{\rho^2}$$

By taking $h = \rho\cos\theta$, $k = \rho\sin\theta$, we get

$$\frac{\Delta f(0,0) - df(0,0)}{\rho} = \frac{\rho^4\sin^2\theta\cos^2\theta}{\rho^3} = \rho\sin^2\theta\cos^2\theta \ \to 0 \ \textit{as } \rho \to 0$$

Hence $\lim_{\rho \to 0} \frac{\Delta f(0,0) - df(0,0)}{\rho} = 0$. Therefore by previous theorem f is differentiable at $(0,0)$.

Example 6.5. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

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

Partial derivatives exist at $(0,0)$ and $f_x(0,0) = 0$, $f_y(0,0) = 0$, so $df(0,0) = 0$. Hence

$$\Delta f(0,0) = f(h,k) - f(0,0) = \frac{h^2k}{h^2+k^2} = \frac{h^2k}{\rho^2}$$

By taking $h = \rho\cos\theta$, $k = \rho\sin\theta$, we get

$$\frac{\Delta f(0,0) - df(0,0)}{\rho} = \frac{\rho^3 \sin\theta\cos^2\theta}{\rho^3} = \sin\theta\cos^2\theta$$

The limit does not exist. Therefore by previous theorem f is not differentiable at $(0,0)$.

6.2. **The Sufficient condition for differentiability:** Following theorem is on Sufficient condition for differentiability:

Theorem 6.4. *Suppose $f_x(x,y)$ and $f_y(x,y)$ exist in an open neighborhood containing (x_0,y_0) and both functions are continuous at (x_0,y_0). Then f is differentiable at (x_0,y_0).*

Proof. Since $\frac{\partial f}{\partial y}$ is continuous at (x_0,y_0), there exists a neighborhood N_δ(say) of (x_0,y_0) at every point of which f_y exists. We take (x_0+h, y_0+k), a point of this neighborhood so that $(x_0+h, y_0), (x_0, y_0+k)$ also belongs to N_δ. We write

$$f(x_0+h,y_0+k) - f(x_0,y_0) = f(x_0+h,y_0+k) - f(x_0+h,y_0) + f(x_0+h,y_0) - f(x_0,y_0).$$

Consider a function of one variable $\phi(y) = f(x_0+h, y_0)$. Since f_y exists in N_δ, $\phi(y)$ is differentiable with respect to y in the closed interval $[y_0, y_0+k]$ and as such we can apply Lagrange's Mean Value Theorem, for function of one variable y in this interval and thus obtain

$$\phi(y_0+k) - \phi(y_0) = k\phi'(y_0+k\theta) = kf_y(x_0+h, y_0+k\theta)$$

where $0 < \theta < 1$. Hence

$$f(x_0+h,y_0+k) - f(x_0+h,y_0) = kf_y(x_0+h,y_0+k\theta), \quad 0 < \theta < 1.$$

Now, if we write

$$f_y(x_0+h,y_0+k\theta) - f_y(x_0,y_0) = \varepsilon_2(h,k)$$

then from the fact that f_y is continuous at (x_0,y_0). We obtain $\varepsilon_2(h,k) \to 0$ as $(h,k) \to (0,0)$.

Again because f_x exists at (x_0,y_0) implies

$$f(x_0+h,y_0) - f(x_0,y_0) = hf_x(x_0,y_0) + h\varepsilon_1(h,k),$$

where $\varepsilon_1(h,k) \to 0$ as $(h,k) \to (0,0)$. Combining all these we get

$$\begin{aligned} f(x_0+h,y_0+k) - f(x_0,y_0) &= k[f_y(x_0,y_0) + \varepsilon_2(h,k)] + h[f_x(x_0,y_0) + \varepsilon_2(h,k)] \\ &= hf_x(x_0,y_0) + kf_y(x_0,y_0) + h\varepsilon_1(h,k) + k\varepsilon_2(h,k) \end{aligned}$$

where $\varepsilon_1, \varepsilon_2(h,k) \to 0$ as $(h,k) \to (0,0)$. This proves that $f(x,y)$ is differentiable at (x_0,y_0). $\square$

Remark 6.2. The above proof still holds if f_y is continuous and f_x exists at (x_0,y_0). Converse is not true: There are functions which are Differentiable but the partial derivatives need not be continuous. For example,

Example 6.6. *$f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$f(x,y) = \{ \begin{array}{cc} x^3 \sin \frac{1}{x^2} + y^3 \sin \frac{1}{y^2} & (x,y) \neq (0,0) \\ 0 & xy = 0 \end{array}$$

Here

$$f_x(x,y) = \{ \begin{array}{cc} 3x^2 \sin \frac{1}{x^2} - 2 \cos \frac{1}{x^2} & (x,y) \neq (0,0) \\ 0 & xy = 0 \end{array}$$

Although $f_x(0,0) = 0$, but $\lim_{(x,y)\to(0,0)} f_x(x,y)$ does not exist, so the partial derivative is not continuous. Moreover

$$f(h,k) - f(0,0) = f(h,k) = h.0 + k.0 + h^3 \sin \frac{1}{h^2} + k^3 \sin \frac{1}{k^2} = h\varepsilon_1 + k\varepsilon_2$$

where $\varepsilon_1 = h^3 \sin \frac{1}{h^2}$ and $\varepsilon_2 = k^3 \sin \frac{1}{k^2}$ and both goes to 0 as $(h,k) \to (0,0)$.

7. Chain Rule

The general Chain Rule with two variables We the following general Chain Rule is needed to find derivatives of composite functions in the form $z = f(x(t), y(t))$ or $z = f(x(s,t), y(s,t))$ in cases where the outer function f has only a letter name. We begin with functions of the first type.

Theorem 7.1. (The Chain Rule) *Suppose that $x = x(t)$ and $y = y(t)$ are differentiable functions of t and $z = f(x, y)$ has partial derivatives with respect to x and y. Then $z = f(x(t), y(t))$ is a differentiable function of t and*

$$\frac{d}{dt}[f(x(t), y(t))] = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \tag{9}$$

assuming f_x and f_y are continuous.

Proof. We assume in this theorem and its applications that $x = x(t)$ and $y = y(t)$ have first derivatives at t and that $z = f(x, y)$ has continuous first-order derivatives in an open circle centered at $(x(t), y(t))$. Equation (9) can be read as the following statement: the t-derivative of the composite function equals the x-derivative of the outer function $z = f(x, y)$ at the point $(x(t), y(t))$ multiplied by the t-derivative of the inner function $x = x(t)$, plus the y-derivative of the outer function at $(x(t), y(t))$ multiplied by the t-derivative of the inner function $y = y(t)$.

We fix t and set $(x, y) = (x(t), y(t))$. We consider nonzero Δt so small that $(x(t+\Delta t), y(t+\Delta t))$ is in the circle where f has continuous first derivatives and set $\Delta x = x(t + \Delta t) - x(t)$ and $\Delta y = y(t + \Delta t) - y(t)$. Then, by the definition of the derivative,

$$\begin{aligned}
\frac{d}{dt}[f(x(t), y(t))] &= \lim_{\Delta t \to 0} \frac{f(x(t + \Delta t), y(t + \Delta t)) - f(x(t), y(t))}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{f(x(t) + \Delta x, y(t) + \Delta y) - f(x(t), y(t))}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{f(x(t) + \Delta x, y(t) + \Delta y) - f(x(t), y(t))}{\Delta t} \qquad (10)
\end{aligned}$$

We express the change $f(x(t)+\Delta x, y(t)+\Delta y) - f(x(t), y(t))$ in the value of $z = f(x, y)$ from (x, y) to $x+\Delta x, y+\Delta y$ as the change in the x-direction from (x, y) to $(x+\Delta x, y)$ plus the change in the y-direction from $(x + \Delta x, y)$ to $(x + \Delta x, y + \Delta y)$.

(11)
$$f(x + \Delta x, y + \Delta y) - f(x, y) = [f(x + \Delta x, y) - f(x, y)] + [f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)].$$

(Notice that the terms $f(x + \Delta x, y)$ and $-f(x + \Delta x, y)$ on the right side of (11) cancel to give the left side.

We can apply the Mean Value Theorem to the expression in the first set of square brackets on the right of (11) where y is constant and to the expression in the second set of square brackets where x is constant. We conclude that there is a number $c_1 \in (x, x + \Delta x)$ and a number $c_2 \in (y, y + \Delta y)$ such that

$$\begin{aligned}
&f(x + \Delta x, y) - f(x, y) = f_x(c_1, y)\Delta x \\
&f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y) = f_y(x + \Delta x, c_2)\Delta y. \qquad (12)
\end{aligned}$$

We combine equations (11) and (12) and divide by Δt to obtain

$$\frac{f(x + \Delta x, y + \Delta y) - f(x(t), y(t))}{\Delta t} = f_x(c_1, y)\frac{\Delta x}{\Delta t} + f_y(x + \Delta x, c_2)\frac{\Delta y}{\Delta t}. \tag{13}$$

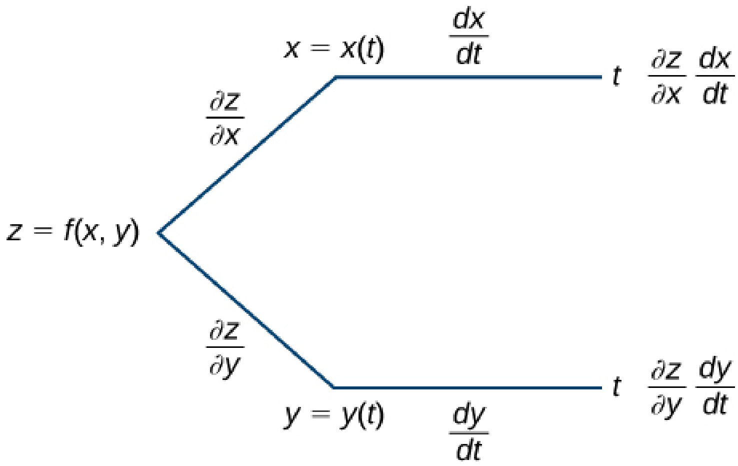
The functions $x = x(t)$ and $y = y(t)$ are continuous at t because they have derivatives at that point. Consequently, as $\Delta t \to 0$, the numbers $\Delta x = x(t + \Delta t) - x(t)$ and $\Delta y = y(t + \Delta t) - y(t)$ both tend to zero. Because the partial derivatives of f are continuous, so $f_x(c_1, y) \to f_x(x, y)$ and $f_y(x + \Delta x, c_2) \to f_y(x, y)$ as $\Delta t \to 0$.

Moreover $\frac{\Delta x}{\Delta t} \to\to x'(t)$ as $\Delta t \to 0$. And $\frac{\Delta y}{\Delta t} \to\to y'(t)$ as $\Delta t \to 0$.
So equation (10) and (13) gives

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

□

It is often useful to create a visual representation of Equation (14) for the chain rule. This is called a tree diagram for the chain rule for functions of one variable and it provides a way to remember the formula (Figure below). This diagram can be expanded for functions of more than one variable, as we shall see very shortly.



Example 7.1. *Use the Chain Rule to find the derivative of*

$$w = xy$$

with respect to t along the path $x = \cos t$, $y = \sin t$. What is the derivative's value $t = \pi/2$.
We apply the Chain Rule to find $\frac{dw}{dt}$ as follows:

$$\begin{aligned}
\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \\
&= \frac{\partial w}{\partial x}\frac{d\cos t}{dt} + \frac{\partial w}{\partial y}\frac{d\sin t}{dt} \\
&= y(-\sin t) + x(\cos t) \\
&= -\sin^2 t + \cos^2 t = \cos 2t
\end{aligned}$$

In this example, we can check the result with a more direct calculation. As a function of t,

$$w = \cos t \sin t = \frac{1}{2}\sin 2t$$

$$\frac{dw}{dt} = \cos 2t$$

In either case, at the given value of t,

$$\frac{dw}{dt}|_{\pi/2} = -1$$

Theorem 6.1 can be applied to find the s- and t-derivatives of a function of the form $z = f(x(s,t), y(s,t))$ because in taking the derivative with respect to s or t, the other variable is constant. We obtain the following.

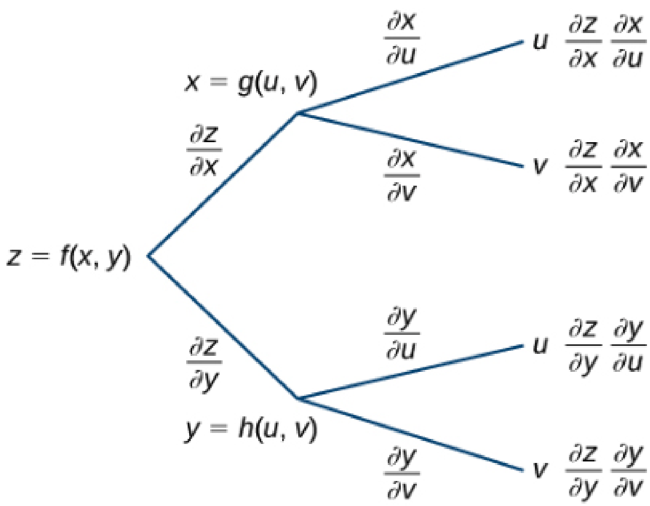
Theorem 7.2. (The Chain Rule) *The s- and t-derivatives of the composite function $z = f(x(s,t), y(s,t))$ are*

$$\frac{\partial}{\partial s}[f(x(s,t),y(s,t))] = f_x(x(s,t),y(s,t))x_s(s,t) + f_y(x(s,t),y(s,t))y_s(s,t),$$

$$\frac{\partial}{\partial t}[f(x(s,t),y(s,t))] = f_x(x(s,t),y(s,t))x_t(s,t) + f_y(x(s,t),y(s,t))y_t(s,t), \tag{14}$$

We assume in this theorem and its applications that the functions involved have continuous first derivatives in the open sets where they are considered. Formulas (14) are easier to remember without the values of the variables in the form,

$$f_s = f_x x_s + f_y y_s$$
$$f_t = f_x x_t + f_y y_t$$



Example 7.2. *Let $z = \log(x^2+y^2)$ and $x(s,t) = e^{s+t^2}$ and $y(s,t) = s^2 + t$. Then*

$$\frac{\partial z}{\partial x} = \frac{2x}{x^2+y^2} \quad \textit{and} \quad \frac{\partial z}{\partial y} = \frac{2y}{x^2+y^2}$$
$$\frac{\partial x}{\partial s} = e^{s+t^2} \quad \textit{and} \quad \frac{\partial y}{\partial s} = 2s$$

Hence

$$\frac{\partial z}{\partial s} = \frac{2x}{x^2+y^2}e^{s+t^2} + \frac{4sy}{x^2+y^2}$$

8. Derivatives of Implicitly Defined function

The two-variable Chain Rule in leads to a formula that takes some of the algebra out of implicit differentiation. Suppose that
1. The function $F(x, y)$ is differentiable and
2. The equation $F(x, y) = 0$ defines y implicitly as a differentiable function of x, say $y = h(x)$.

Theorem 8.1. A Formula for Implicit Differentiation *Suppose that $F(x, y)$ is differentiable and that the equation $F(x, y) = 0$ defines y as a differentiable function of x. Then at any point where $F_y \neq 0$, then*

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

Proof. Since $w = F(x, y) = 0$, the derivative $\frac{dw}{dx}$ must be zero. Computing the derivative from the Chain Rule 9, with ($t = x$ and $f = F$)

$$0 = \frac{dw}{dx} = F_x \frac{dx}{dx} + F_y \frac{dy}{dx} = F_x.1 + F_y \frac{dy}{dx}$$

This implies

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

□

Example 8.1. *The function $y(x)$ defined implicitly as $e^y - e^x + xy = 0$. Let $F(x, y) = e^y - e^x + xy$. Then*

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

Example 8.2. *Consider $2x^2 - \sin y = y^2$. Prove that $\frac{dy}{dx} = \frac{4x}{2y+\cos y}$.*

Example 8.3. *Find the slope of the tangent line to the curve $x^2 + y^2 = 25$ at the point $(3, -4)$. Let $F(x, y) = x^2 + y^2 - 25$. And $F(x, y) = 0$ with $F_x(x, y) = 2x$ and $F_y(x, y) = 2y$. By the formula for implicit differentiation,*

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{2x}{2y}. \tag{15}$$

Because the slope of the tangent line to a curve is the derivative with respect to x, (15) yields $\frac{dy}{dx}$ at $(3, -4)$ is $\frac{3}{4}$.

Example 8.4. *A 28-foot ladder is leaning against a wall of a building. It starts to slide down the building at a rate of 4ft/s. How fast is the base of the ladder moving away from the wall when the top of the ladder has slid down 8 ft from its initial position?*

We know that the length of the ladder is 28 ft. We have a right triangle setup with the given situation. Thus, we know that $x^2 + y^2 = 28$. We are also looking for the moment when $y = 20$ and $\frac{dy}{dt} = -4$ ft/s. So $F(x, y) = x^2 + y^2 - 28^2$.
Compute the derivative using implicit differentiation. Using the chain rule, we have

$$\frac{d}{dt}[F(x, y) = 0]$$
$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0$$

When $y = 20$, we have

$$x^2 + 400 = 28^2$$
$$x = 8\sqrt{6}$$

Using all of the information gathered so far, we can substitute into our implicit derivative and solve for $\frac{dx}{dt}$. Hence

$$28\sqrt{6}\frac{dx}{dt} + 2.20(-4) = 0$$
$$16\sqrt{6}\frac{dx}{dt} = 160$$
$$\frac{dx}{dt} = 4.082$$

Example 8.5. *The pressure P (in Kilopascals), volume V (in Litres) and temperature T (in Kelvins) of a mole of an ideal gas are related by $PV = 8.31T$. Find the rate at which the pressure is changing when the temperature is $300K$ and increasing at a rate of 0.1 K/s and the volume is $100L$ and increasing at a rate of $0.2L/s$.*

9. Directional derivatives

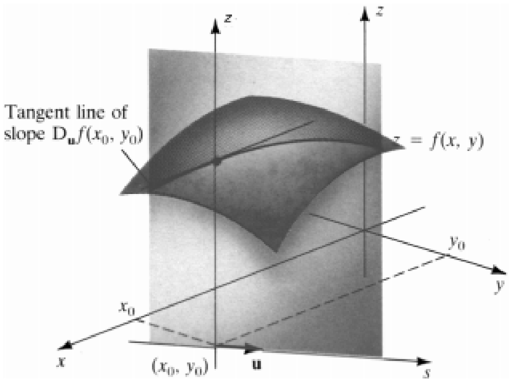
Definition 9.1. Let $\tilde{p} = p_1 i + p_2 j$ be any unit vector. Then the directional derivative of $f(x, y)$ at (x_0, y_0) in the direction of $\tilde{p}$ is

$$D_{\tilde{p}}(x_0, y_0) = \lim_{s \to 0} \frac{f(x_0 + sp_1, y_0 + sp_2) - f(x_0, y_0)}{s}$$

We can replace $\tilde{p} = p_1 i + p_2 j$ by $\tilde{p} = \cos\theta i + \sin\theta j$, and obtain

$$D_{\tilde{p}}(x_0, y_0) = \lim_{s \to 0} \frac{f(x_0 + s\cos\theta, y_0 + s\sin\theta) - f(x_0, y_0)}{s}$$

Geometrical Interpretation Consider $F(s) := f(x_0 + s\cos\theta, y_0 + s\sin\theta)$, so $F(0) = f(x_0, y_0)$ Then the graph of $z = F(s)$ the intersection of the surface $z = f(x, y)$ with the sz-plane. The



Tangent line of slope

$F'(0) = D_{\mathbf{u}} f(x_0, y_0)$

directional derivative ($D_{\tilde{p}}(x_0, y_0) = F'(0)$) of $z = f(x, y)$ is the slope of the tangent line to this curve in the positive s-direction at $s = 0$, which is at the point $(x_0, y_0, f(x_0, y_0))$.

Example 9.1. *Let $\theta = arc\cos(3/5)$. Find the directional derivative of $f(x, y) = x^2 - xy + 3y^2$ in the direction of $u = (\cos\theta)\,\hat{\mathbf{i}} + (\sin\theta)\,\hat{\mathbf{j}}$.*
First of all, since $\cos\theta = 3/5$. So

$$\sin\theta = \sqrt{1 - \left(\frac{3}{5}\right)^2} = \sqrt{\frac{16}{25}} = \frac{4}{5}.$$

Using $f(x, y) = x^2 - xy + 3y^2$, we will find $f(x + s\cos\theta, y + s\sin\theta)$:

$$\begin{aligned}
f(x + s\cos\theta, y + s\sin\theta) &= (x + s\cos\theta)^2 - (x + s\cos\theta)(y + s\sin\theta) + 3(y + s\sin\theta)^2 \\
&= x^2 + 2xs\cos\theta + s^2\cos^2\theta - xy - xs\sin\theta - ys\cos\theta \\
&\quad - s^2\sin\theta\cos\theta + 3y^2 + 6ys\sin\theta + 3s^2\sin^2\theta \\
&= x^2 + 2xh(\frac{3}{5}) + \frac{9s^2}{25} - xy - \frac{4xs}{5} - \frac{3ys}{5} - \frac{12s^2}{25} + 3y^2 + 6ys(\frac{4}{5}) + 3s^2(\frac{16}{25}) \\
&= x^2 - xy + 3y^2 + \frac{2xs}{5} + \frac{9s^2}{5} + \frac{21ys}{5}.
\end{aligned}$$

Hence we have

$$\begin{aligned}
D_u f(x,y) &= \lim_{s\to 0} \frac{f(x+s\cos\theta, y+s\sin\theta)-f(x,y)}{s} \\
&= \lim_{s\to 0} \frac{(x^2-xy+3y^2+\frac{2xs}{5}+\frac{9s^2}{5}+\frac{21yh}{5})-(x^2-xy+3y^2)}{s} \\
&= \lim_{s\to 0} \frac{\frac{2xs}{5}+\frac{9s^2}{5}+\frac{21ys}{5}}{h} \\
&= \lim_{s\to 0} \frac{2x}{5}+\frac{9s}{5}+\frac{21y}{5} \\
&= \frac{2x+21y}{5}.
\end{aligned}$$

We have

$$D_u f(-1,2) = \frac{2(-1)+21(2)}{5} = \frac{-2+42}{5} = 8.$$

Example 9.2. *$f(x,y)=x^2+xy$ at $P(1,2)$ in the direction of unit vector $\tilde{p}=\frac{i}{\sqrt{2}}+\frac{i}{\sqrt{2}}$.*

$$\begin{aligned}
D_{\tilde{p}}(1,2) &= \lim_{s\to 0} \frac{f(1+\frac{s}{\sqrt{2}}, b+\frac{s}{\sqrt{2}})-f(1,2))}{s} \\
&= \lim_{s\to 0} \frac{\left(s^2+s(2\sqrt{2}+\frac{1}{\sqrt{2}})\right)}{s} = 2\sqrt{2}+\frac{1}{\sqrt{2}}
\end{aligned}$$

The existence of partial derivatives does not guarantee the existence of directional derivatives in all directions. For example take

Example 9.3. *Consider the function $f:\mathbb{R}^2\to\mathbb{R}$ defined by*

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

For example take any $\tilde{p}=p_1 i+p_2 j$ any unit vector. The directional derivative of f alon $\tilde{p}$ at $(0,0)$ is

$$D_{\tilde{p}}(0,0) = \lim_{s\to 0} \frac{f(sp_1,sp_2)-f(0,0)}{s} = \lim_{s\to 0} \frac{p_1p_2}{s(p_1^2+p_1^2)}$$

exists iff either $p_1=0$ or $p_2=0$.

Example 9.4. *Consider the function $f:\mathbb{R}^2\to\mathbb{R}$ defined by*

$$f(x,y)=\sqrt{|xy|}$$

For example take any $\tilde{p}=p_1 i+p_2 j$ any unit vector. The directional derivative of f alon $\tilde{p}$ at $(0,0)$ is

$$D_{\tilde{p}}(0,0) = \lim_{s\to 0} \frac{f(sp_1,sp_2)-f(0,0)}{s} = \lim_{s\to 0} \frac{|h|}{h}\sqrt{p_1p_2}$$

exists iff either $p_1=0$ or $p_2=0$. Moreover, f is continuous at $(0,0)$. Indeed,

$$|f(x,y)-f(0,0)| = |f(x,y)-0| = \sqrt{|xy|} \leq \frac{1}{\sqrt{2}}\sqrt{(x-0)^2+(y-0)^2} < \varepsilon$$

whenever $\sqrt{(x-0)^2+(y-0)^2} < \sqrt{2}\varepsilon = \delta_\varepsilon$

10. Directional Derivatives and Gradient Vectors

Let us recall the definition of differentiable function. A function f is said to be differentiable at the point (x_0, y_0) if there exist and ε_1 and ε_2 such that

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0)).(h, k) + h\varepsilon_1(h, k) + k\varepsilon_2(h, k).$$

and $\varepsilon_1, \varepsilon_2 \to 0$ as $(h, k) \to (0, 0)$.

Definition 10.1. The **gradient vector (or gradient)** of $f(x, y)$ is the vector

$$\nabla f := (f_x, f_y) = f_x\hat{i} + f_y\tilde{j} \in \mathbb{R}^2$$

Gradient of f at a point (x_0, y_0) is defined to be a vector in $\mathbb{R}^2$

$$\nabla f|_{(x_0,y_0)} = (f_x(x_0, y_0), f_y(x_0, y_0)) = f_x(x_0, y_0)\hat{i} + f_y(x_0, y_0)\hat{j}$$

The gradient vector is drawn as an arrow with its base at (x_0, y_0).

Example 10.1. *Find the gradient of each of the following functions:*

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

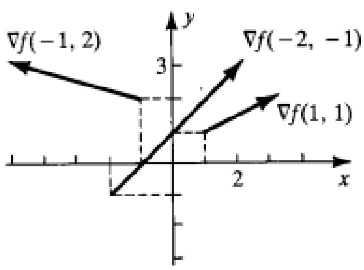
at (x, y) *Here* $f_x(x, y) = 2x - y$ *and* $f_y(x, y) = -x + 6y$ *so*

$$\begin{aligned}\vec{\nabla} f(x, y) &= f_x(x, y)\,\hat{\mathbf{i}} + f_y(x, y)\,\hat{\mathbf{j}} \\ &= (2x - y)\,\hat{\mathbf{i}} + (-x + 6y)\,\hat{\mathbf{j}}.\end{aligned}$$

Example 10.2. *Draw* $\vec{\nabla} f(1, 1)$ *and* $\vec{\nabla} f(-1, 2)$ *and* $\vec{\nabla} f(-2, -1)$ *for* $f(x, y) = x^2y$.

Here $f_x(x, y) = 2xy$ *and* $f_y(x, y) = x^2$ *so*

$$\begin{aligned}\vec{\nabla} f(x, y) &= f_x(x, y)\,\hat{\mathbf{i}} + f_y(x, y)\,\hat{\mathbf{j}} \\ &= 2xy\,\hat{\mathbf{i}} + x^2\,\hat{\mathbf{j}}.\end{aligned}$$



So

$$\vec{\nabla} f(1, 1) = 2\,\hat{\mathbf{i}} + 1\,\hat{\mathbf{j}}.$$

and

$$\vec{\nabla} f(-1, 2) = -4\,\hat{\mathbf{i}} + 1\,\hat{\mathbf{j}}.$$

And

$$\vec{\nabla} f(-2, -1) = 4\,\hat{\mathbf{i}} + 4\,\hat{\mathbf{j}}.$$

Remark 10.1. The gradient vector has lot of geometric significance. Moreover it is evident from the definition that the gradient vector may exists even even when the function is not differentiable at some point. If the function is differentiable at some point then we have the following proposition.

Theorem 10.1. The Directional Derivative Is a Dot Product *If $f(x, y)$ is a differentiable function in an open region containing $P(x_0, y_0)$ and f_x and f_y exist at (x_0, y_0), then the directional derivative in the direction $\overrightarrow{p} = p_1 \overrightarrow{i} + p_2 \overrightarrow{j}$ at (x_0, y_0) is*

$$D_{\overrightarrow{p}} f(x_0, y_0) = \nabla f(x_0, y_0).\overrightarrow{p}$$

Proof. Applying the definition of a directional derivative in the direction of $\overrightarrow{p} = p_1\hat{i} + p_2\hat{j}$ at a point (x_0, y_0) in the domain of f can be written

$$D_{\overrightarrow{p}} f(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + tp_1, y_0 + tp_2) - f(x_0, y_0)}{t}$$

We take the line $x(s) = x_0 + tp_1$ and $y(s) = y_0 + tp_2$ and $x(0) = x_0$ and $y(0) = y_0$, then define $g(t) = f(x(t), y(t))$. Since f_x and f_y both exist, we can use the chain rule for functions of two variables to calculate $g'(t)$

$$g'(t) = f_x(x, y)\frac{dx}{dt} + f_y(x, y)\frac{dy}{dt} = f_x(x, y)p_1 + f_y(x, y)p_2.$$

So we have

$$g'(0) = f_x(x_0, y_0)p_1 + f_y(x_0, y_0)p_2 = (f_x(x_0, y_0), f_y(x_0, y_0))(p_1, p_2) = \nabla f(x_0, y_0).\overrightarrow{p}$$

By the definition of $g'(t)$ it is also true that

$$g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \lim_{t\to 0} \frac{f(x_0 + tp_1, y_0 + tp_2) - f(x_0, y_0)}{t} = D_{\overrightarrow{p}} f(x_0, y_0).$$

Hence

$$D_{\overrightarrow{p}} f(x_0, y_0) = \nabla f(x_0, y_0).\overrightarrow{p}$$

□

Example 10.3. *Let $\theta = \arccos(3/5)$. Find the directional derivative of $f(x, y) = x^2 - xy + 3y^2$ is the direction of $\overrightarrow{u} = (\cos\theta)\,\hat{\mathbf{i}} + (\sin\theta)\,\hat{\mathbf{j}}$. Find $D_{\overrightarrow{u}} f(-1, 2)$.*
First, we must calculate the partial derivatives of f:

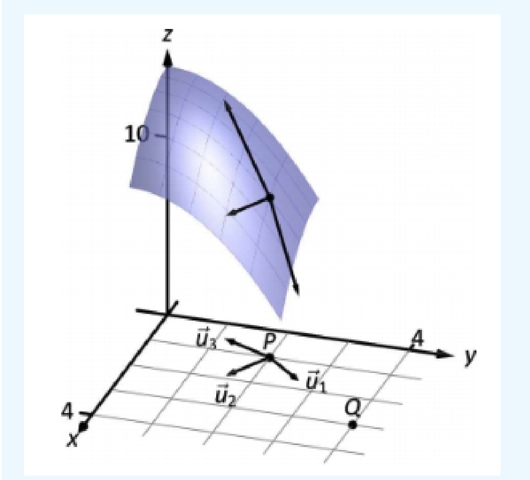
$$f_x(x, y) = 2x - y$$
$$f_y(x, y) = -x + 6y,$$

Here we use previous result

$$\begin{aligned} D_{\overrightarrow{u}} f(x, y) &= f_x(x, y)\cos\theta + f_y(x, y)\sin\theta \\ &= (2x - y)\frac{3}{5} + (-x + 6y)\frac{4}{5} \\ &= \frac{6x}{5} - \frac{3y}{5} - \frac{4x}{5} + \frac{24y}{5} \\ &= \frac{2x + 21y}{5}. \end{aligned}$$

$$D_{\overrightarrow{u}} f(-1, 2) = \frac{2(-1) + 21(2)}{5} = \frac{-2 + 42}{5} = 8$$

Example 10.4. *Let $z = 14 - x^2 - y^2$ and $P = (1, 2)$. Then Find the directional derivative of f at P, in the following directions:*
1) toward the point $Q = (3, 4)$.
2) in the direction of $\langle 2, -1\rangle$.
3) toward the origin.
The surface is plotted in the figure, where the point $P = (1, 2)$ is indicated in the $x - y$ plane as well as the point $(1, 2, 9)$ which lies on the surface of f. We find that $f_x(x, y) = -2x$ and $f_y(x, y) = -2y$. So $f_x(1, 2) = -2$ and $f_x(1, 2) = -4$.



1) Let u_1 be the unit vector that points from the point $P = (1, 2)$ to the point $Q = (3, 4)$. as shown in the figure. The vector $\bar{PQ} = \langle < 2, 2\rangle$; the unit vector in this direction is $u_1 = \langle < \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\rangle$. Thus the directional derivative of f at P in the direction of u_1 is

$$D_{u_1} f(1, 2) = -2(1/\sqrt{2}) + (-4)(1/\sqrt{2}) = -6/\sqrt{2} \approx -4.24.$$

2) We seek the directional derivative in the direction of $\langle 2, -1\rangle$. The unit vector in this direction is $u_2 = \langle 2/\sqrt{5}, -1/\sqrt{5}\rangle$, then the directional derivative of f at P in the direction of u_1 is

$$D_{u_2} f(1, 2) = -2(2/\sqrt{5}) + (-4)(-1/\sqrt{5}) = 0.$$

Starting on the surface of $z = f(x, y)$ at P and moving in the direction of $\langle 2, -1\rangle$ or u_2 results in no instantaneous change in z -value. This is analogous to standing on the side of a hill and choosing a direction to walk that does not change the elevation. One neither walks up nor down, rather just "along the side" of the hill.

3) At P the direction towards the origin is given by the vector $\langle -1, -2\rangle$, the unit vector in this direction is $u_3 = \langle -1/\sqrt{5}, -2/\sqrt{5}\rangle$. The directional derivative of f at P in the direction of the origin is

$$D_{u_3} f(1, 2) = -2(-1/\sqrt{5}) + (-4)(-2/\sqrt{5}) = 10/\sqrt{5} \approx 4.47.$$

Moving towards the origin means "walking uphill" quite steeply, with an initial slope of about 4.47.

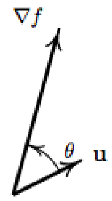
10.1. **Properties of Directional Derivatives and Gradient.** The gradient has some important properties. We have already seen one formula that uses the gradient: the formula for the directional derivative. Recall from The Dot Product that if the angle between two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ is ϕ, then $\overrightarrow{a}.\overrightarrow{b} = \|a\|\|b\| \cos\phi$.

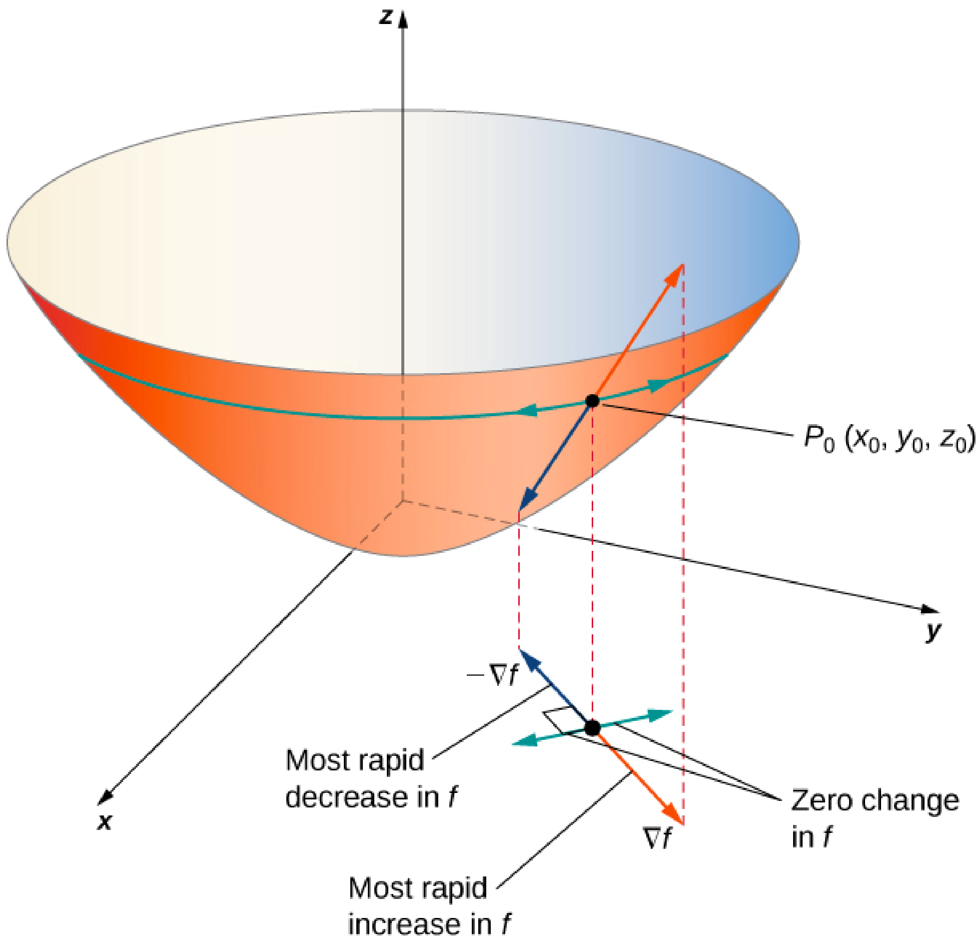
Therefore, if the angle between $\nabla f(x_0, y_0)$ and $\overrightarrow{u} = u_1 \overrightarrow{i} + u_2 \overrightarrow{j}$ (unit vector) is ϕ, we have

$$D_{\overrightarrow{u}} f(x_0, y_0) = \|\overrightarrow{\nabla} f(x_0, y_0)\|\|u\| \cos\phi = \|\overrightarrow{\nabla} f(x_0, y_0)\| \cos\phi$$

Therefore, the directional derivative is equal to the magnitude of the gradient evaluated at (x_0, y_0) multiplied by $\cos\phi$. Recall that $\cos\phi$ ranges from -1 to 1.

(1) If $\phi = 0$, then $\cos\phi = 1$ and $\nabla f(x_0, y_0)$ and $\overrightarrow{u}$ both point in the same direction, then we say $D_{\overrightarrow{u}} f(x_0, y_0)$ is **maximised**.
(2) If $\phi = \pi$, then $\cos\phi = -1$ and $\nabla f(x_0, y_0)$ and $\overrightarrow{u}$ are in opposite directions, then we say $D_{\overrightarrow{u}} f(x_0, y_0)$ is **minimised**.
(3) We can also see that if $\overrightarrow{\nabla} f(x_0, y_0) = 0$ then $D_{\overrightarrow{p}} f(x_0, y_0) = 0$.





Example 10.5. *Find the direction for which the directional derivative of $f(x,y) = 3x^2 - 4xy + 2y^2$ at $(-2,3)$ is a maximum. What is the maximum value?*
The maximum value of the directional derivative occurs when $\vec{\nabla} f$ and the unit vector point in the same direction. Therefore, we start by calculating $\vec{\nabla} f$:

$$f_x(x,y) = 6x - 4y \text{ and } f_y(x,y) = -4x + 4y$$

and

$$\vec{\nabla} f(x,y) = f_x(x,y)\,\hat{\mathbf{i}} + f_y(x,y)\,\hat{\mathbf{j}} = (6x-4y)\,\hat{\mathbf{i}} + (-4x+4y)\,\hat{\mathbf{j}}.$$

Next, we evaluate the gradient at $(-2,3)$:

$$\vec{\nabla} f(-2,3) = (6(-2)-4(3))\,\hat{\mathbf{i}} + (-4(-2)+4(3))\,\hat{\mathbf{j}} = -24\,\hat{\mathbf{i}} + 20\,\hat{\mathbf{j}}.$$

We need to find a unit vector that points in the same direction as $\vec{\nabla} f(-2,3)$*, so the next step is to divide* $\vec{\nabla} f(-2,3)$ *by its magnitude, which is* $\sqrt{-24)^2+(20)^2} = \sqrt{976} = 4\sqrt{61}$*. Therefore,*

$$\begin{aligned}\frac{f(-2,3)}{\|\vec{\nabla} f(-2,3)\|} &= \frac{-24}{4\sqrt{61}}\hat{\mathbf{i}} + \frac{20}{4\sqrt{61}}\hat{\mathbf{j}} \\ &= -\frac{6\sqrt{61}}{61}\,\hat{\mathbf{i}} + \frac{5\sqrt{61}}{61}\,\hat{\mathbf{j}}.\end{aligned}$$

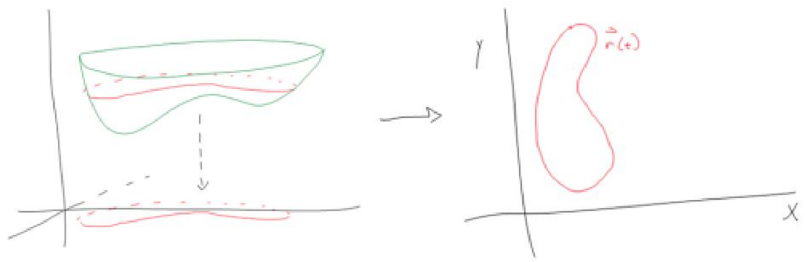
This is the unit vector that points in the same direction as $\vec{\nabla} f(-2,3)$*. To find the angle corresponding to this unit vector, we solve the equations*

$$\cos\theta = \frac{-6\sqrt{61}}{61} \text{ and } \sin\theta = \frac{5\sqrt{61}}{61}$$

for θ*. Since cosine is negative and sine is positive, the angle must be in the second quadrant. Therefore* $\theta = \pi - \arcsin((5\sqrt{61})/61) = 2.45$
The maximum value of the directional derivative at $(-2,3)$ *is* $\|\vec{\nabla} f(-2,3)\| = 4\sqrt{61}$

Theorem 10.2. *At* $P = (x_0, y_0)$*,* $\vec{\nabla} f(x_0,y_0)$ *is orthogonal to the level curve passing through* $(x_0, y_0, f(x_0,y_0))$*.*
Or Suppose the function $z = f(x,y)$ *has continuous first-order partial derivatives in an open disk centered at a point* $P = (x_0, y_0)$*. If* $\vec{\nabla} f(x_0,y_0) \neq 0$*, then* $\vec{\nabla} f(x_0,y_0)$ *is normal to the level curve of* f *at* $P = (x_0, y_0)$*.*

Proof. If the graph of $z = f(x,y)$ is a sufficiently nice surface, we can think about its level curves at height $z = c$, and $f(x,y) = c$. A slightly different way to think about a chosen level curve $f(x,y) = c$ is to think about a vector function $r(t) = x(t)\bar{i} + y(t)\bar{j}$ that traces out the same curve in the xy plane. (Recall that a curve is defined parametrically by the function pair $(x(t), y(t))$ or a vector function $\bar{r}(t) = x(t)\bar{i} + y(t)\bar{j}$, then the vector $\bar{r'}(t) = x'(t)\bar{i} + y'(t)\bar{j}$ is tangent to the curve for every value of t in the domain.)

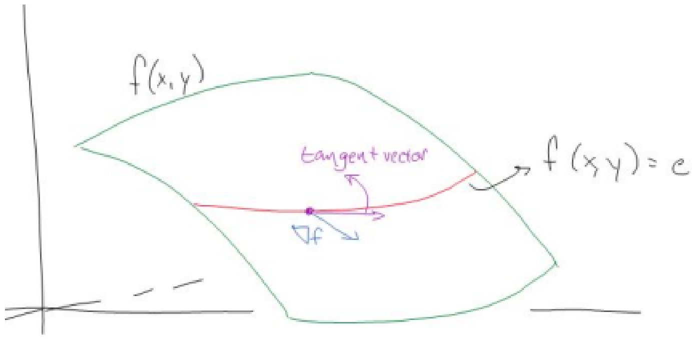


In particular, this allows us to think of $f(x,y) = c$ as $f(x(t), y(t)) = c$; in other words, we turn f into a function of one variable, t. Let's differentiate the relationship with respect to t (apply chain rule):

$$\frac{d}{dt}[f(x(t),y(t))] = f_x(x(t),y(t))x'(t) + f_y(x(t),y(t))y'(t) \tag{16}$$
$$= (f_x(x(t),y(t))\bar{i} + f_y(x(t),y(t))\bar{j})(x'(t)\bar{i} + y'(t)\bar{j}) \tag{17}$$

Since Since $f(x(t), y(t)) = c$ and $\frac{d\ c}{dt} = 0$ (c is a constant), we have

$$\begin{aligned} 0 = \frac{d}{dt}[f(x(t),y(t))] &= (f_x(x(t),y(t))\bar{i} + f_y(x(t),y(t))\bar{j}) \cdot (x'(t)\bar{i} + y'(t)\bar{j}) \\ &= \vec{\nabla} f(x(t),y(t)) \cdot (x'(t)\bar{i} + y'(t)\bar{j})\end{aligned}$$



Now $\frac{d\bar{r}}{dt}$, which is a vector that is tangent to the level curve and pointing in the direction of increasing t. Now $\overrightarrow{\nabla} f$ and $\frac{d\bar{r}}{dt}$ are always orthogonal.

□

Example 10.6. *For the function $f(x,y) = 2x^2 - 3xy + 8y^2 + 2x - 4y + 4$, find a tangent vector to the level curve at point $(-2,1)$. Graph the level curve corresponding to $f(x,y) = 18$ and draw in $\overrightarrow{\nabla} f(-2,1)$ and a tangent vector.*

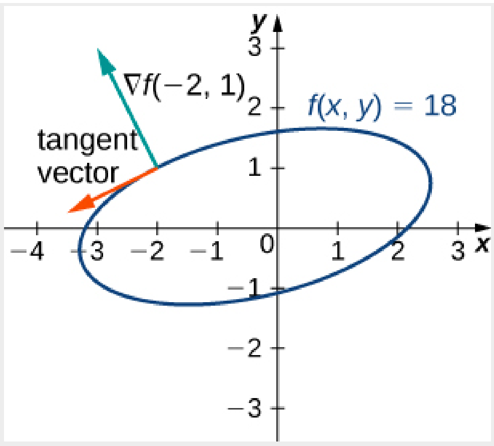


Figure 4.45 A rotated ellipse with equation f(x, y) = 18. At the point (–2, 1) on the ellipse, there are drawn two arrows, one tangent vector and one normal vector. The normal vector is marked ∇f(–2, 1) and is perpendicular to the tangent vector.

First we calculate $\overrightarrow{\nabla} f(x,y)$ Now

$$f_x(x,y) = 4x - 3y + 2 \text{ and } f_y(x,y) = -3x + 16y - 4$$

Hence $\overrightarrow{\nabla} f(-2,1) = -9\bar{i} + 18\bar{j}$

This vector is orthogonal to the curve at point $(-2,1)$. We can obtain a tangent vector by reversing the components and multiplying either one by -1. Thus, for example, $18\bar{i} - 9\bar{j}$ is a tangent vector (see the following graph).

11. MVT for multivariable

Mean Value Theorem : We will present the MVT for functions of several variables which is a consequence of MVT for functions of one variable.

Theorem 11.1. *Suppose $f : D \to \mathbb{R}$ be a function which is differentiable. Suppose $(x_0, y_0) \in D$ and $(x_0 + h, y_0 + k) \in D$, then there exists a $c \in (0, 1)$ such that*

$$f((x_0 + h, y_0 + k) - f(x_0, y_0) = hf_x(x_0 + ch, y_0 + ck) + kf_y(x_0 + ch, y_0 + ck)$$

Proof. Consider the function $\phi : [0, 1] \to \mathbb{R}$ by

$$\phi(t) = f((x_0 + th, y_0 + tk)$$

By Chain Rule

$$\frac{d\phi}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} = hf_x + kf_y$$

So by MVT, there exists $c \in (0, 1)$ such that

$$\phi(1) - \phi(0) = \phi'(c).$$

This proves the result. □

Remark 11.1. If $f(x, y)$ is constant if and only if $f_x = 0$ and $f_y = 0$.

12. Mixed Derivatives

The first-order partial derivatives $f_x = \frac{\partial f}{\partial x}$ and $f_y = \frac{\partial f}{\partial y}$ of $z = f(x, y)$ can be differentiated with respect to x and y to obtain the second x-derivative and the second y-derivative

$$f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \tag{18}$$

$$f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{19}$$

and the mixed second derivatives

$$f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \tag{20}$$

$$f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \tag{21}$$

But $f_{xy} \neq f_{yx}$ is not always true. Consider the example

Example 12.1. $f : \mathbb{R}^2 \to \mathbb{R}$ *defined by*

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

This function is continuous everywhere. We now compute the first order partial derivatives. Now $f_y(0, 0) = 0$

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

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

Similarly, $f_{xy}(0, 0) = -1$*. So* $f_{yx}(0, 0) \neq f_{xy}(0, 0)$.

Theorem 12.1. Clairaut's theorem *If $f(x,y)$ and its partial derivatives f_x, f_y, f_{xy}, f_{yx} are defined in a neighborhood of (x_0,y_0) and all are continuous at (x_0,y_0) then $f_{xy}(x_0,y_0)=f_{yx}(x_0,y_0)$.*

Proof. We have

$$\begin{aligned}
f_{xy}(x_0,y_0)=(f_x)_y(x_0,y_0) &= \lim_{k\to 0}\frac{1}{k}\left(f_x(x_0,y_0+k)-f_x(x_0,y_0)\right)\\
&= \lim_{k\to 0}\frac{1}{k}\lim_{h\to 0}\frac{1}{h}\left([f(x_0+h,y_0+k)-f(x_0,y_0+k)]-[f(x_0+h,y_0)-f(x_0,y_0)]\right)\\
&= \lim_{(h,k)\to 0}\frac{1}{hk}\left([f(x_0+h,y_0+k)-f(x_0,y_0+k)]-[f(x_0+h,y_0)-f(x_0,y_0)]\right)\\
&= \lim_{(h,k)\to 0}\frac{1}{hk}\left([f(x_0+h,y_0+k)-f(x_0+h,y_0))]-[(f(x_0,y_0+k)-f(x_0,y_0))]\right)
\end{aligned}$$

Now, using that f_y is defined in the neighborhood, applying the mean value theorem to $g(y)=f(x_0+h,y)$, we have $f(x_0+h,y_0+k)-f(x_0+h,y_0))=kf_y(x_0+h,y_0+\zeta(k))$ where $y_0+\zeta(k)\in(y_0,y_0+k)$ or in (y_0,y_0+k).
Again using that f_y is defined in the neighborhood and applying the mean value theorem to $h(y)=f(x_0,y)$, we have $f(x_0,y_0+k)-f(x_0,y_0))=kf_y(x_0,y_0+\psi(k))$ where $y_0+\psi(k)\in(y_0,y_0+k)$ or in (y_0+k,y_0).
Hence

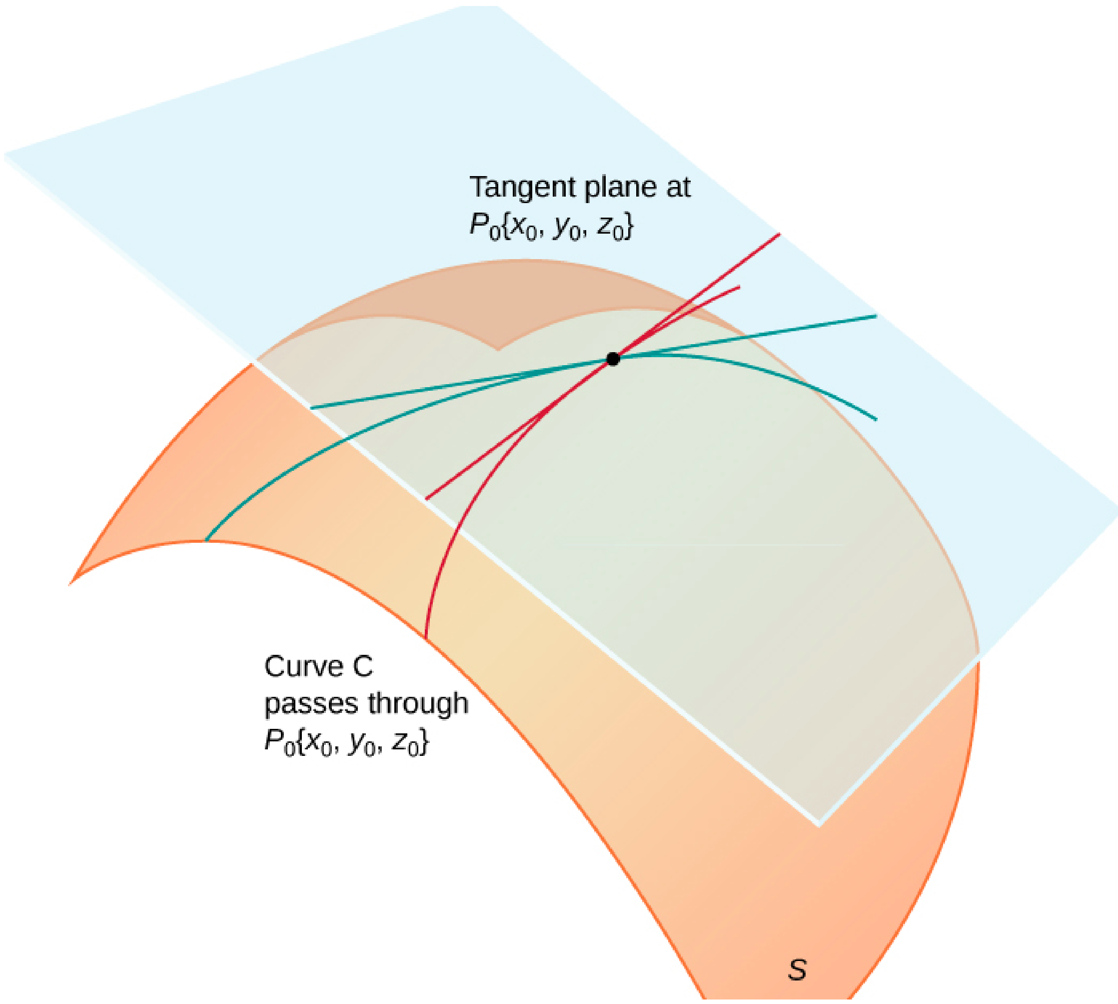
$$\begin{aligned}
f_{xy}(x_0,y_0) &= \lim_{(h,k)\to 0}\frac{1}{h}[f_y(x_0+h,y_0+\zeta(k))-f_y(x_0,y_0+\psi(k))]\\
&= \lim_{h\to 0}\frac{1}{h}\lim_{k\to 0}[f_y(x_0+h,y_0+\zeta(k))-f_y(x_0,y_0+\psi(k))]\\
&= \lim_{h\to 0}\frac{1}{h}[f_y(x_0+h,y_0)-f_y(x_0,y_0)] \text{ as } \ f_y \text{ is a continuous function}\\
&= (f_y)_x(x_0,y_0)=f_{yx}(x_0,y_0)
\end{aligned}$$

Hence the result. $\square$

13. Tangent Plane

Tangent Planes Intuitively, it seems clear that, in a plane, only one line can be tangent to a curve at a point. However, in three-dimensional space, many lines can be tangent to a given point. If these lines lie in the same plane, they determine the tangent plane at that point. A tangent plane at a regular point contains all of the lines tangent to that point. A more intuitive way to think of a tangent plane is to assume the surface is smooth at that point (no corners). Then, a tangent line to the surface at that point in any direction does not have any abrupt changes in slope because the direction changes smoothly.

Definition 13.1. Tangent Plane Let $P = (x_0, y_0, z_0 = f(x_0, y0))$ be a point on a surface $z = f(x, y)$, and let C be any curve passing through P 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.



Equation of Tangent Plane Let S be a surface defined by a differentiable function $z = f(x, y)$, and let $P_0 = (x_0, y_0, z_0)$ be a point on S. Then, the equation of the tangent plane to S at P_0 is given by

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

Example 13.1. *Find the equation of the tangent plane to the surface* $z = x^2 + y^2$ *at* $(1, 2, 5)$*.*

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

$f_x(x, y) = 2x$ *and* $f_y(x, y) = 2y$ *so the equation of the tangent plane at the point* $(1, 2, 5)$ *is*

$$2(1)(x - 1) + 2(2)(y - 2) - z + 5 = 0$$

14. Taylor's theorem

Remark 14.1. The linearization of $f(x,y)$ at a point $P_0(x_0,y_0)$ is

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

. We can approximate $f(x,y)$ with this tangent plane in a small neighbourhood of (x_0,y_0). Similarly, we have quadratic approximation, and so on. The quadratic approximation,

$$Q(x,y) = f(x_0,y_0)+f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-y_0)+\frac{1}{2!}(x-x_0)^2 f_{xx}(x_0,y_0)+(x-x_0)(y-y_0)f_{xy}(x_0,y_0)+\frac{1}{2!}(y-$$

We have studied extensively Taylor's theorem for one variable calculus. Now we shall discuss the Taylor's theorem for more than one variable functions. Essentially following the same idea we can state and proof the theorem. We state the following theorem :

Theorem 14.1. (Taylor's Theorem) *Suppose $f(x,y)$ and its partial derivatives up to $(n+1)$-th are continuous throughout an open rectangular region R centred at a point (x_0,y_0). Then for all points in R we have*

$$f(x_0+h,y_0+k) = f(x_0,y_0) + (hf_x(x_0,y_0) + kf_y(x_0,y_0) + \frac{1}{2!}(h^2 f_{xx}(x_0,y_0) + 2hkf_{xy}(x_0,y_0) + k^2 f_{yy}(x_0,y_0))$$
$$+ \cdots + \frac{1}{n!}(hf_x(x,y) + kf_y(x,y))^{(n)}|_{(x_0,y_0)} + \frac{1}{(n+1)!}(hf_x(x,y) + kf_y(x,y))^{(n+1)}|_{(a+ch,b+ck)}$$

where (x_0+ch, y_0+ck) is a point on the line segment joining (x_0,y_0) and (x_0+h,y_0+k).

Proof. The proof follows immediately from one variable calculus. Apply Taylor's theorem (one variable) and Chain rule to the function,

$$\phi(t) = f(x_0+ht, y_0+kt) \quad \text{with} \quad t = 0$$

That is first apply Taylor's theorem to $\phi(t)$ around $t=0$. Then apply chain rule to calculate $\phi'(t), \phi^{(2)}(t), \cdots, \phi^{(n+1)}(t)$. □

Example 14.1. *Expand $f(x,y) = e^{x+2y}$ by Taylor's Theorem around $(0,0)$. Bound the error of this approximation if $|x| \le 0.1$ and $|y| \le 0.1$*
As $f(0,0) = 1$, $f(x,y) = 1 + x + 2y + \frac{1}{2}x^2 + 2xy + 2y^2 + E_2(x,y)$ with $|E_2(x,y)| \le \frac{(0.3)^3}{3!}e^{0.3}$ when $|x| \le 0.1$ and $|y| \le 0.1$. (Of course there are many other correct bounds on $|E_2(x,y)|$

15. Maxima and Minima

To find the local extreme values of a function of a single variable, we look for points where the graph has a horizontal tangent line. At such points, we then look for local maxima, local minima, and points of inflection. For a function $f(x,y)$ of two variables, we look for points where the surface $z = f(x,y)$ has a horizontal tangent plane. At such points, we then look for local maxima, local minima, and saddle points. We begin by defining maxima and minima.

Definition 15.1. Let $f(x,y)$ be defined on a region R containing the point (x_0,y_0). Then
1. $f(x_0,y_0)$ is a local maximum value of f if $f(x_0,y_0) \ge f(x,y)$ for all domain points (x,y) in an open disk centered at (x_0,y_0).
2. $f(x_0,y_0)$ is a local minimum value of f if $f(x_0,y_0) \le f(x,y)$ for all domain points (x,y) in an open disk centered at (x_0,y_0). Local extrema are also called **relative extrema**.

Theorem 15.1. First Derivative Test for Local Extreme Values *If $f(x,y)$ has a local maximum or minimum value at an interior point (x_0,y_0) of its domain and if the first partial derivatives exist there, then $f_x(x_0,y_0) = 0$ and $f_y(x_0,y_0) = 0$.*

Proof. If f has a local extremum at (x_0, y_0), then the function $g(x) = f(x, y_0)$ has a local extremum at $x = x_0$ (now it is a one variable function). Therefore, $g'(x_0) = 0$. Now $g'(x_0) = f_x(x_0, y_0)$, so $f_x(x_0, y_0) = 0$. A similar argument with the function $h(y) = f(x_0, y)$ shows $f_y(x_0, y_0) = 0$. □

Example 15.1. *Find the local extreme values of $f(x, y) = x^2 + y^2 - 4y + 9$.*
The domain of f is the entire plane (so there are no boundary points) and the partial derivatives $f_x = 2x$ and $f_y = 2y - 4$ exist everywhere. Therefore, local extreme values **can (not necessarily)** *occur only where $f_x = 2x = 0$ and $f_y = 2y - 4 = 0$. The only possibility is the point $(0, 2)$, where the value of f is 5. Since $f(x, y) = x^2 + (y - 2)^2 + 5 \geq 5 = f(0, 2)$ for all $(x, y) \in \mathbb{R}^2$, we see that the critical point $(0, 2)$ gives a local minimum.*

Definition 15.2. Critical Point An interior point of the domain of a function $f(x, y)$ where both f_x and f_y are zero or where one or both of f_x and f_y do not exist is a critical point of f.

Definition 15.3. Saddle Point A differentiable function $f(x, y)$ has a saddle point at a critical point (x_0, y_0) if in every open disk centered at (x_0, y_0) there are domain points (x, y) where $f(x, y) \geq f(x_0, y_0)$ and domain points (x, y) where $f(x, y) \leq f(x_0, y_0)$. The corresponding point $(a, b, (x_0, y_0))$ on the surface $z = (x, y)$ is called a saddle point of the surface.

Example 15.2. *Find the local extreme values (if any) of $f(x, y) = y^2 - x^2$.*
Solution The domain of f is the entire plane (so there are no boundary points) and the partial derivatives $f_x(x, y) = -2x$ and $f_y(x, y) = 2y$ exist everywhere. Therefore, local extrema **can (it may not occur at that point also)** *occur only at the origin $(0, 0)$ where $f_x = 0$ and $f_y = 0$. Along the positive (negative also) x-axis, however, f has the value $f(x, 0) = -x^2 \neq 0$; along the positive (negative also) y-axis, f has the value $f(0, y) = y^2 \neq 0$. Therefore, every open disk in the xy-plane centered at $(0, 0)$ contains points where the function is positive (that is $f(0, y) \geq 0 = f(0, 0)$) and points where it is negative (that is $f(x, 0) \leq 0 = f(0, 0)$) . The function has a saddle point at the origin and no local extreme values.*

Theorem 15.2. Second Derivative Test for Local Extreme Values *Suppose that $f(x, y)$ and its first and second partial derivatives are continuous throughout a disk centered at (x_0, y_0) and that $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$.*
i) f has a **local maximum** *at (x_0, y_0) if $f_{xx} < 0$ and $D = f_{xx}f_{yy} - f_{xy}^2 > 0$ at (x_0, y_0).*
ii) f has a **local minimum** *at (x_0, y_0) if $f_{xx} > 0$ and $D = f_{xx}f_{yy} - f_{xy}^2 > 0$ at (x_0, y_0).*
iii) f has a **saddle point** *at (x_0, y_0) if $D = f_{xx}f_{yy} - f_{xy}^2 < 0$ at (x_0, y_0).*
iv) the **test is inconclusive** *at (x_0, y_0) if $D = f_{xx}f_{yy} - f_{xy}^2 = 0$ at (x_0, y_0). In this case, we must find some other way to determine the behavior of f at (x_0, y_0).*

Proof. R □

Example 15.3. *1. The functions $f_1(x, y) = -(x^4 + y^4)$ and $f_2(x, y) = x^4 + y^4$ satisfy the above equation $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ for $(x_0, y_0) = (0, 0)$ but f_1 has a local maximum at $(0, 0)$ and f_2 has a local minimum at $(0, 0)$.*
2. Consider the function $f(x, y) = (x + y)^2 - x^4$. This function satisfies the above equation $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ for $(x_0, y_0) = (0, 0)$ but it has neither a local maximum nor a local minimum at $(0, 0)$. Moreover $f_{xx}f_{yy} - f_{xy}^2 = 0$ at $(0, 0)$ In fact, $(0, 0)$ is a saddle point. This can be verified as follows. Note that for $0 < x < 1, f(x, x) > 0 = f(0, 0)$ and $f(x, -x) < 0 = f(0, 0)$.

3. Let $f(x, y) = x \sin y$. Here $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ for $(x_0, y_0) = (0, n\pi)$ where $n \in \mathbb{N}$. Note that $f_{xx}f_{yy} - f_{xy}^2 < 0$ at $(x_0, y_0) = (0, n\pi)$. Therefore, the points $(x_0, y_0) = (0, n\pi)$ where $n \in \mathbb{N}$ are saddle points.

Example 15.4. *Find the local Maximum and Minimum and saddle points of the function* $f(x,y) = x^4 + y^4 - 4xy + 1$.
First find the critical values.

$$f_x(x,y) = 4x^3 - 4y \text{ and } f_y(x,y) = 4y^3 - 4x$$
$$f_x(x,y) = 4x^3 - 4y = 0 \text{ and } f_y(x,y) = 4y^3 - 4x = 0$$

which implies $x^3 = y$ *and* $y^3 = x$. *Combining this two equations, we have* $x^9 - x = 0$. *Solutions are* $x = 0$, $x = 1$ *and* $x = -1$. *So the critical points are* $(0,0)$, $(1,1)$ *and* $(-1,-1)$. *Let us calculate the second derivatives,*

$$f_{xx}(x,y) = 12x^2 \text{ and } f_{yy}(x,y) = 12y^2 \text{ and } f_{xy}(x,y) = f_{yx}(x,y) = -4.$$

So $D = f_{xx}(x,y)f_{yy}(x,y) - f_{xy}^2(x,y) = 144x^2y^2 - 16$.
At $(x,y) = (0,0)$, $D|_{(0,0)} = -16 < 0$ *so* $(0,0)$ *is a saddle point.*
At $(x,y) = (1,1)$, $D|_{(1,1)} = -144 - 16 > 0$ *and* $f_{xx}(1,1) = 12 > 0$, *so* $(1,1)$ *is a point of local minimum.*
At $(x,y) = (-1,-1)$, $D|_{(-1,-1)} = -144 - 16 > 0$ *and* $f_{xx}(1,1) = 12 > 0$, *so* $(-1,-1)$ *is a point of local minimum.*

Example 15.5. *Find the shortest distance from the point* $(1,0,-2)$ *to the plane* $x + 2y + z = 4$.
The distance between $(1,0,-2)$ *to any arbitary point* (x,y,z) *of the plane is*

$$d(x,y) = \sqrt{(x-1)^2 + y^2 + (z+2)^2} = \sqrt{(x-1)^2 + y^2 + (4-x-2y+2)^2}$$
$$d^2(x,y) = (x-1)^2 + y^2 + (6-x-2y)^2$$

Let us calculate f_x *and* f_y

$$f_x(x,y) = 4x + 4y - 14 \text{ and } f_y(x,y) = 10y + 4x - 24$$
$$f_{xx}(x,y) = 4 \text{ and } f_{yy}(x,y) = 10 \text{ and } f_{xy}(x,y) = f_{yx}(x,y) = 4$$

Now $f_x(x,y) = 0 = f_y(x,y)$ *will give rise to* $x = 11/6$ *and* $y = 5/3$. *S0* $(11/6, 5/3)$ *is the critical point of the function.*
Now $f_{xx}(11/6, 5/3)) = 4 > 0$ *and* $f_{yy}((11/6, 5/3)) = 10$ *and* $f_{xy}((11/6, 5/3)) = f_{yx}((11/6, 5/3)) = 4$. *Hence* $D_{(11/6,5/3)} = 24 > 0$. *So by second derivative test,* $(11/6, 5/3)$ *is a point of local minimum for* $d^2(x,y)$ *and the minimum value is ....*

Example 15.6. *A rectangular box without lid is to be made from* $12m^2$ *of cardboard. Find the maximum volume of such box.*
Suppose x, y *and* z *are the length, breadth and height of the box. As it is without lid, so the total area is* $2xz + 2yz + xy$

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

Volume is $V(x,y) = xyz = xy\frac{12-xy}{2(x+y)}$.

$$V_x(x,y) = y^2\frac{(12 - 2xy - x^2)}{2(x+y)^2}$$
$$V_y(x,y) = x^2\frac{(12 - 2xy - y^2)}{2(x+y)^2}$$

So $V_x(x,y) = 0 = V_y(x,y)$ gives two solutions for critical points:
1) $x = y = 0$ but length can not be 0, so we discard it.
2) $x = y$, gives $x = \pm 2$, as length can not be negative, so we discard $x = -2$ but take $x = 2$. So $x = 2 = y$.
By second derivative test, prove $(2,2)$ gives the maximum volume. $z = 1$ is the height and 8 is the maximum volume.

Example 15.7. *Find all local maxima and minima of $f(x,y) = (x^2+y^2)e^{-(x^2+y^2)}$*

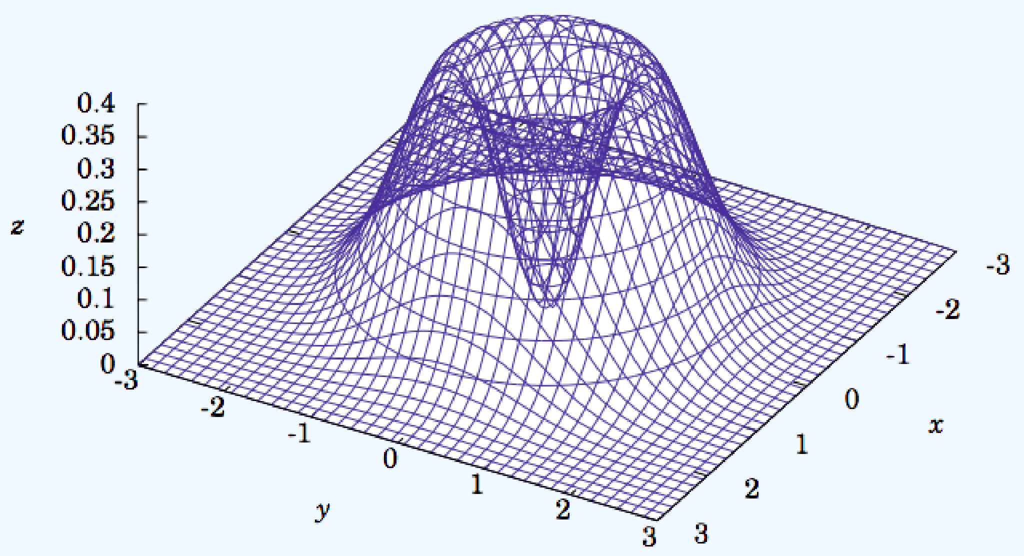


Figure 2.5.2 $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$

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

$$f_x(x,y) = 2x(1-(x^2+y^2))e^{-(x^2+y^2)} \tag{22}$$

$$f_y(x,y) = 2y(1-(x^2+y^2))e^{-(x^2+y^2)}$$

then the critical points are $(0,0)$ and all points (x,y) on the unit circle $x^2+y^2=1$.
we need the second-order partial derivatives:

$$f_{xx}(x,y) = 2[1-(x^2+y^2)-2x^2-2x^2(1-(x^2+y^2))]e^{-(x^2+y^2)}$$
$$f_{yy}(x,y) = 2[1-(x^2+y^2)-2y^2-2y^2(1-(x^2+y^2))]e^{-(x^2+y^2)}$$
$$f_{xy}(x,y) = -4xy[2-(x^2+y^2)]e^{-(x^2+y^2)} \tag{23}$$

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

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

and so the test fails. If we look at the graph of $f(x,y)$, as shown in 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: $g(r) = r^2e^{-r^2}$. Then $g'(r) = 2r(1-r^2)e^{-r^2}$, 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.

16. Lagrange Multipliers

For a rectangle whose perimeter is 20 m, find the dimensions that will maximize the area. The area A of a rectangle with width x and height y is $A = xy$. The perimeter P of the rectangle is then given by the formula $2x + 2y = P$. Since we are given that the perimeter $P = 20$, this problem can be stated as:

$$\begin{aligned} \text{Maximize} &: f(x, y) = xy \\ \text{given} &: 2x + 2y = 20 \end{aligned} \tag{24}$$

The reader is probably familiar with a simple method, using single-variable calculus, for solving this problem. Since we must have $2x + 2y = 20$, then we can solve for, say, y in terms of x using that equation. This gives $y = 10 - x$ which we then substitute into f to get $f(x, y) = xy = x(10 - x) = 10x - x^2$ now a function of x alone, so we now just have to maximize the function $f(x) = 10x - x^2$ on the interval $[0, 10]$. Since $f'(x) = 10 - 2x = 0 \Rightarrow x = 5$ and $f''(5) = -2 < 0$, then the Second Derivative Test tells us that $x = 5$ is a local maximum for f, and hence $x = 5$ must be the global maximum on the interval $[0, 10]$ (since $f = 0$ at the endpoints of the interval). So since $y = 10-x = 5$, then the maximum area occurs for a rectangle whose width and height both are 5 m.

Notice in the above example that the ease of the solution depended on being able to solve for one variable in terms of the other in the equation $2x + 2y = 20$. But what if that were not possible (which is often the case)? In this section we will use a general method, called the Lagrange multiplier method, for solving constrained optimization problems:

$$\begin{aligned} \text{Maximize (or minimize)} &: f(x, y) \quad (\text{or } f(x, y, z)) \\ \text{given} &: h(x, y) = c \quad (\text{or } h(x, y, z) = c) \text{ for some constant } c \end{aligned} \tag{25}$$

The equation $h(x, y) = c$ is called the constraint equation, and we say that x and y are constrained by $h(x, y) = c$. Points (x, y) which are maxima or minima of $f(x, y)$ with the condition that they satisfy the constraint equation $h(x, y) = c$ are called constrained maximum or constrained minimum points, respectively. Similar definitions hold for functions of three variables.

The Lagrange multiplier method for solving such problems can now be stated:

Theorem 16.1. *Let f and g be functions of two variables with continuous partial derivatives at every point of some open set containing the smooth curve $g(x, y) = 0$. Suppose that f, when restricted to points on the curve $g(x, y) = 0$, has a local extremum at the point (x_0, y_0) and that $\overrightarrow{\nabla} g(x_0, y_0) \neq 0$. Then there is a number λ called a Lagrange multiplier, for which*

$$\overrightarrow{\nabla} f(x_0, y_0) = \lambda \overrightarrow{\nabla} g(x_0, y_0).$$

Example 16.1.

$$\begin{aligned} \textit{Maximize (and minimize)} &: f(x, y) = xy \\ \textit{given} &: h(x, y) = 2x + 2y = 20 \end{aligned}$$

Now $g(x, y) = 2x + 2y - 20$, Solve

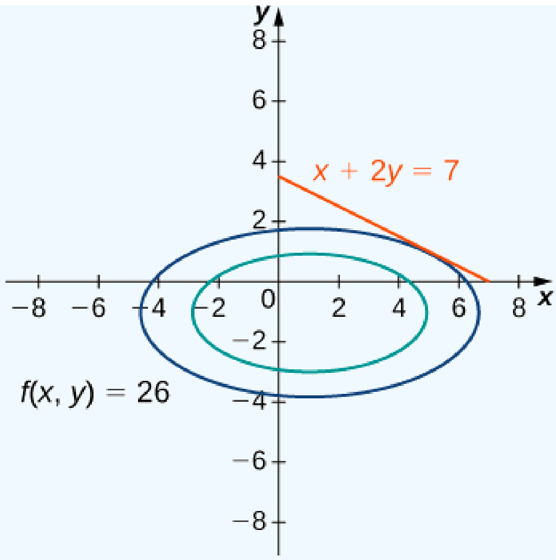
$$\overrightarrow{\nabla} f(x_0, y_0) = \lambda \overrightarrow{\nabla} g(x_0, y_0).$$

$$y_0 = 2\lambda \tag{26}$$

$$x_0 = 2\lambda \tag{27}$$

So $x = y$. Putiing this in $g(x, y) = 0$ gives $4x = 20$ so $x = y = 5$. and $f(5, 5) = 25$ TTo ensure this corresponds to a maximum value on the constraint function, let's try some other points on the constraint such as the $(1, 9)$ and $(9, 1)$. In both cases $f(9, 1) = f(1, 9) = 9 < 25 = f(5, 5)$. The maximum area occurs for a rectangle whose width and height both are 5 m.

Example 16.2. *Use the method of Lagrange multipliers to find the minimum value of $f(x, y) = x^2 + 4y^2 - 2x + 8y$ subject to the constraint $x + 2y = 7$.*
The objective function is $f(x, y) = x^2 + 4y^2 - 2x + 8y$. To determine the constraint function, we must first subtract 7 from both sides of the constraint. This gives $x + 2y - 7 = 0$. The constraint function is equal to the left-hand side, so $g(x, y) = x + 2y - 7$. The problem asks us to solve for the minimum value of f, subject to the constraint



3: Graph of level curves of the function $f(x, y) = x^2 + 4y^2 - 2x + 8y$ corresponding to $c = 10$ and 26. The red graph is the constraint functio

We then must calculate the gradients of both f and g:

$$\bar{\nabla} f(x, y) = (2x - 2)\,\hat{\mathbf{i}} + (8y + 8)\,\hat{\mathbf{j}}$$
$$\bar{\nabla} g(x, y) = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}.$$

$$\bar{\nabla} f(x, y) = (2x - 2)\,\hat{\mathbf{i}} + (8y + 8)\,\hat{\mathbf{j}}$$
$$\bar{\nabla} g(x, y) = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}.$$

The equation $\vec{\nabla} f(x_0, y_0) = \lambda \vec{\nabla} g(x_0, y_0)$ becomes

$$(2x_0 - 2)\,\hat{\mathbf{i}} + (8y_0 + 8)\,\hat{\mathbf{j}} = \lambda\left(\hat{\mathbf{i}} + 2\hat{\mathbf{j}}\right),$$

which can be rewritten as

$$(2x_0 - 2)\,\hat{\mathbf{i}} + (8y_0 + 8)\,\hat{\mathbf{j}} = \lambda\hat{\mathbf{i}} + 2\lambda\hat{\mathbf{j}}.$$

Next, we set the coefficients of $\bar{i}$ and $\bar{j}$ equal to each other:

$$2x_0 - 2 = \lambda$$
$$8y_0 + 8 = 2\lambda.$$

The equation $g(x,y)=0$ *becomes* $x+2y-7=0.$. *Therefore, the system of equations that needs to be solved is*

$$\begin{aligned} 2x_0-2&=\lambda\\ 8y_0+8&=2\lambda\\ x_0+2y_0-7&=0. \end{aligned}$$

This is a linear system of three equations in three variables. We start by solving the second equation for λ *and substituting it into the first equation. This gives* $\lambda=4y_0+4$ *, so substituting this into the first equation gives*

$$2x_0-2=4y_0+4.$$

Solving this equation for x_0 *gives* $x_0=2y_0+3$. *We then substitute this into the third equation:*

$$\begin{aligned} (2y_0+3)+2y_0-7=0&\\ &4y_0-4=0\\ &y_0=1. \end{aligned}$$

Since $x_0=2y_0+3$, $x_0=5$. *Next, we evaluate* $f(x,y)$ *at the point* $(5,1)$,

$$f(5,1)=5^2+4(1)^2-2(5)+8(1)=27.$$

To ensure this corresponds to a minimum value on the constraint function, let's try some other points on the constraint from either side of the point $(5,1)$, *such as the intercepts of* $g(x,y)=0$ *which are* $(7,0)$ *and* $(0,3.5)$. *Now* $f(7,0)=35>27=f(5,1)$ *and* $f(0,3.5)=77>27=f(5,1)$.