

6 Functions on several variables

To any ordered pair of real numbers (x, y) there is related one point in xy -plane. To any point in xy -plane there are related the coordinates of this point, that means the ordered pair of real numbers. It is said that between ordered pairs of real numbers and the points on xy -plane there is *one-to-one correspondence*.

The subset of the points of xy -plane is called the domain (region). We shall denote the domains by D . For example the domain

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

is the unit disk centered at the origin, which contains the circle surrounding this disk.

The curve bounding the domain is called the *boundary line* of this domain and the points on the boundary line are called *boundary points*. The points not laying on the boundary line are called *interior points*.

The domain containing all of its boundary points (that means the whole boundary line) is called *closed*.

The domain containing none of its boundary points is called *open* (if it contains some but not all of its boundary points, then it is neither open or closed).

If the domain contains its boundary line or a part of its boundary line, we sketch this line (part of the line) by the continuous line. If the domain does not contain its boundary line or a part of its boundary line, we sketch this line (part of the line) by the dashed line.



Figure 6.1. Open and closed disk

Any open disk centered at the given point is called the *neighborhood* of this point. If $\varepsilon > 0$ is whatever real number, then the ε -neighborhood of the point (x_0, y_0) is the open disk (without center)

$$U_\varepsilon(x_0, y_0) = \{(x, y) \mid 0 < (x - x_0)^2 + (y - y_0)^2 < \varepsilon^2\}$$

There exists the one-to-one correspondence between the triplets of real numbers (x, y, z) and the points in space. The subset of the (x, y, z) -space is called the *spatial region*.

The spatial region is separated from the rest of the space by a surface, which is called the *boundary surface*. The points on the boundary surface are called the boundary points and the points not laying on the boundary surface are called interior points.

The region is called closed, if it contains all of its boundary points and the region is called open, if it contains none of its boundary points.

Thus, the closed region is the region with the surface surrounding the region and the open region is the region without the surface surrounding the region.

The ε -neighborhood of the spatial point (x_0, y_0, z_0) is the open ball

$$U_\varepsilon(x_0, y_0, z_0) = \{(x, y, z) \mid 0 < (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 < \varepsilon^2\}$$

that means the open ball centered at (x_0, y_0, z_0) with radius ε . This open ball does not contain the sphere surrounding the ball and does not contain the center (x_0, y_0, z_0) .

6.1 Functions of two variables

Let D be some domain in the xy -plane (included the whole plane). A function of two variables is a function whose inputs are points $(x; y)$ in the xy -plane and whose outputs real numbers.

Definition 1. If to each point $(x; y) \in D$ there is related one certain value of the variable z , then z is called the function of two variables x and y and denoted

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

The function of two variables can be also denoted by $z = g(x, y)$, $z = F(x, y)$, $z = \varphi(x, y)$ or $z = z(x, y)$.

The variables x and y are the *independent variables* and z is the function or the *dependent variable*.

Whenever a quantity depend on two others we have a function of two variables. The area on the rectangle of length x and width y is $S = xy$. The number of items n which can be sold is the function of the price p and the advertising budget a that is $n = f(p, a)$. The force of the suns gravity F depends on an object mass m and the distance d : $F = F(m, d)$.

Further we shall consider the functions given implicitly. In those cases to each point $(x; y) \in D$ there can be related two or more values of the variable z . We talk about the two-valued functions, three-valued functions, etc.

In the graph of the function of two variables $z = f(x, y)$ is the spatial point with coordinates $(x, y, f(x, y))$. The set of all those point is the surface in space. Hence, the graph of the function of two variables $z = f(x, y)$ is the surface in x, y, z -coordinates.

Example 1. The graph of the function $z = 1 - x - y$ is the plane. The part of this plane, which is in the first octant of spatial coordinates is in Figure 6.2.

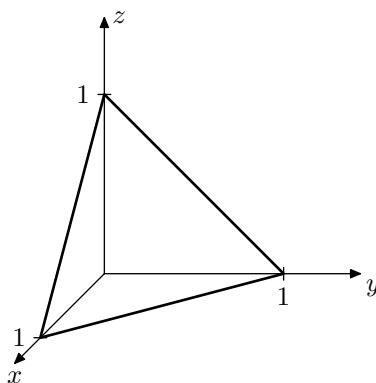


Figure 6.2. The graph of the function $z = 1 - x - y$ in I octant

Example 2. The graph of the function $z = x^2 + y^2$ is the paraboloid of revolution created by the rotation of the parabola $z = y^2$ around z -axis.

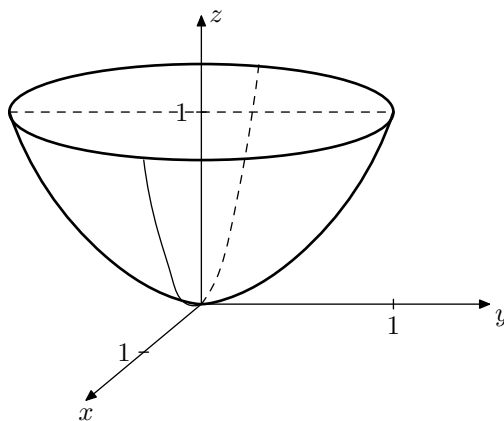


Figure 6.3. Paraboloid of revolution

The next surface is the graph of the function of two variables given im-

plicitly.

Example 3. The graph of the function $x^2 + y^2 + z^2 = r^2$ given implicitly is the sphere with radius r centered at the origin.

Solving this equation for the variable z , we obtain two one-valued functions of two variables $z = \sqrt{r^2 - x^2 - y^2}$ and $z = -\sqrt{r^2 - x^2 - y^2}$. The graph of the first function is the upper side and second function the lower side of the sphere.

Definition 2. The domain of the function of two variables $z = f(x, y)$ is the set of ordered pairs (x, y) (the points of the plane) for which by the given rule it is possible to compute the value of the function.

Example 4. Let us find the domain of the function $z = \ln(8 - x^2 - y^2) + \sqrt{2y - x^2}$ and sketch it in the coordinate plane.

The function is defined if there hold two conditions

$$\begin{cases} 8 - x^2 - y^2 > 0 \\ 2y - x^2 \geq 0. \end{cases}$$

The first condition yields $x^2 + y^2 < 8$ and the second $y \geq \frac{x^2}{2}$. The first condition holds for the points in xy -plane, which belong to the disk centered at the origin and with radius $2\sqrt{2}$. There is no equality to 8, therefore the circle surrounding the disk does not belong to the set and we sketch the circle with dashed line.

The second condition holds for the points in xy -plane, which are above the parabola $y = \frac{x^2}{2}$. This condition contains the equality, consequently the parabola belongs to the set and we sketch it with continuous line.

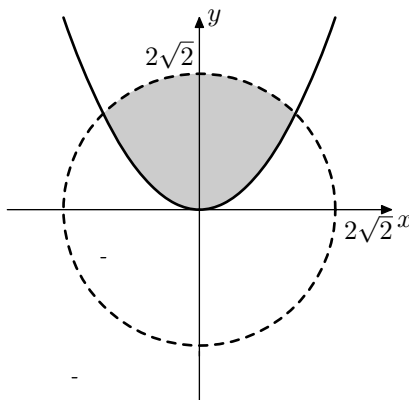


Figure 6.4. The domain of the function $z = \ln(8 - x^2 - y^2) + \sqrt{2y - x^2}$

6.2 Plane sections and level curves of graph of function of two variables

To get an idea how does the graph of the function of two variables looks like it is useful to sketch plane sections of this surface by the planes which are perpendicular to one of the coordinate axis (i.e. parallel to one of the coordinate planes). The equation of the yz -plane is $x = 0$, the equation of the xz -plane is $y = 0$ and the equation of the xy -plane is $z = 0$.

The plane $x = a$ is perpendicular to x -axes, i.e. parallel to yz -plane; the plane $y = b$ is perpendicular to y -axes i.e. parallel to xz -plane; the plane $z = c$ is perpendicular to z -axes i.e. parallel to xy -plane.

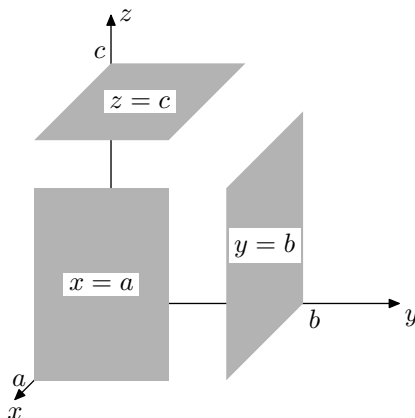


Figure 6.5. The planes $x = a$, $y = b$ and $z = c$

The intersections of the surface $z = f(x, y)$ with the planes $x = a$ are the curves

$$\begin{cases} z = f(x, y) \\ x = a, \end{cases}$$

The intersections of the surface $z = f(x, y)$ with the planes $y = b$ are the curves

$$\begin{cases} z = f(x, y) \\ y = b \end{cases}$$

The intersections of the surface $z = f(x, y)$ with the planes $z = c$ are the

$$\begin{cases} z = f(x, y) \\ z = c, \end{cases}$$

The projection of the resulting curve onto the xy -plane is called the *level curve*. A collection of level curves of a surface is called a *contour map*.

Example 1. Let us sketch the surface $x^2 + y^2 - z^2 = 0$, using the intersections with the planes $z = 0$, $z = \pm 1$, $z = \pm 2$ and $x = 0$. First five are the horizontal curves and the sixth is the intersection with the yz -plane.

The intersection of this surface with xy -plane $z = 0$ is actually one point determined by the equations $x^2 + y^2 = 0$, $z = 0$, which is the origin.

The intersection of this surface with the horizontal plane $z = 1$ is the circle $x^2 + y^2 = 1$, $z = 1$, the unit circle on the plane $z = 1$ centered at $(0; 0; 1)$.

The intersection of this surface with the horizontal plane $z = -1$ is the unit circle $x^2 + y^2 = 1$ again but centered at $(0; 0; -1)$.

The intersection of this surface with the horizontal plane $z = 2$ is the circle $x^2 + y^2 = 4$ centered at $(0; 0; 2)$ with radius 2.

The intersection of this surface with the horizontal plane $z = -2$ is the circle $x^2 + y^2 = 4$ centered at $(0; 0; -2)$ with radius 2. We sketch these circles in the spatial coordinates (Figure 6.6).

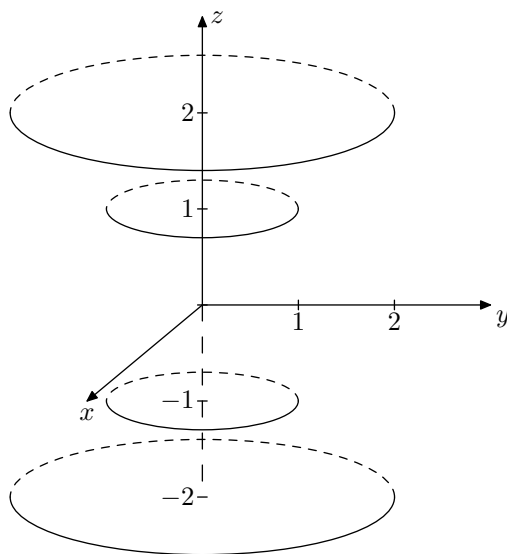


Figure 6.6. The level curves of surface $x^2 + y^2 - z^2 = 0$

The intersection of this surface with the vertical plane $x = 0$ is determined by $z^2 = y^2$, $x = 0$, that is two perpendicular lines $z = y$ and $z = -y$ on yz -plane. Adding these two lines to our sketch it turns obvious that the given surface is the cone, whose vertex is at the origin.

If we convert the function $x^2 + y^2 - z^2 = 0$ to the explicit form we obtain two one-valued functions $z = \sqrt{x^2 + y^2}$ and $z = -\sqrt{x^2 + y^2}$. The graph of

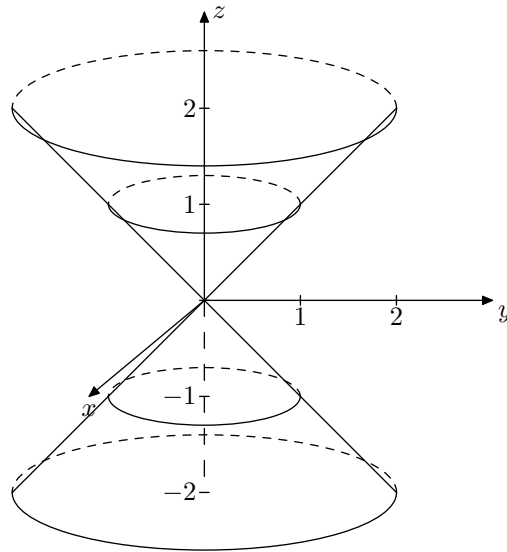


Figure 6.7. The cone

the first function is the upper part of the cone and the graph of the second function is the lower part of the cone.

Example 2. Let us sketch the surface $z = x^2 - y^2$, using the intersections with the planes $y = 0$, $x = \pm 1$, $x = \pm 0,5$, $x = 0$, $z = 0$ and $z = -0,44$.

In this example we draw the coordinate axes in an unusual way, taking the sheet of paper the xz -plane and directing y -axes backwards.

The intersection with the plane $y = 0$ is the parabola $z = x^2$, $y = 0$.

The intersections with the planes $x = \pm 1$ are the parabolas $z = 1 - y^2$, $x = 1$ and $z = 1 - y^2$, $x = -1$.

The intersections with the planes $x = \pm 0,5$ are the parabolas $z = 0,25 - y^2$, $x = 0,5$ and $z = 0,25 - y^2$, $x = -0,5$.

The intersections with the plane $z = 0$ are two perpendicular lines $y = x$ and $y = -x$ on the xy -plane.

The intersection with the plane $z = -0,44$ is the equilateral hyperbola $y^2 - x^2 = 0,44$, whose real axis is the y -axis.

The *level surfaces* of the graph of function of three variables $w = f(x, y, z)$ are the surfaces

$$\begin{cases} w = f(x, y, z) \\ w = c. \end{cases}$$

This system of equations yields the equation $f(x, y, z) = c$, the function of two variables given implicitly, whose graph is a surface in the space.

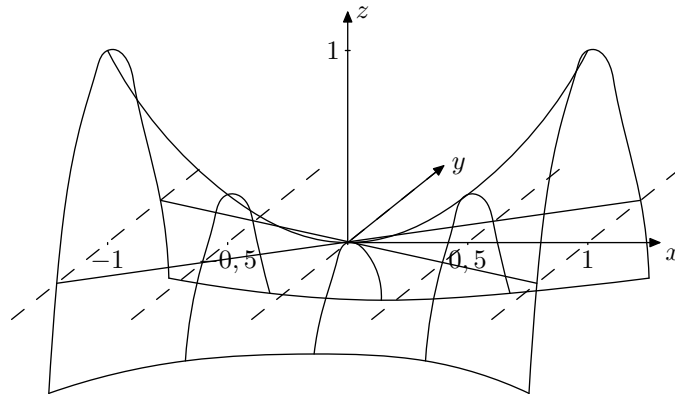


Figure 6.8. Hyperbolic paraboloid or saddle surface

Example 3. The level surfaces of the function $w = x^2 + y^2 + z^2$ are $x^2 + y^2 + z^2 = c$ provided $c > 0$. Those surfaces are the spheres centered at the origin with radius \sqrt{c} .

6.3 Increment of function of several variables

Let us fix one point $P(x, y)$ in the domain of the function $z = f(x, y)$. Changing the variable x by Δx and y by Δy , we obtain a point $Q(x + \Delta x, y + \Delta y)$. Assuming that the increments of the independent variables Δx and Δy are sufficiently small, that is Q is also in the domain of the function, we define the *total increment* of the function

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y) \quad (6.1)$$

Assuming that y is a constant or $\Delta y = 0$, we have the increment of the function

$$\Delta_x z = f(x + \Delta x, y) - f(x, y) \quad (6.2)$$

Assuming that x is a constant or $\Delta x = 0$, we have the increment of the function

$$\Delta_y z = f(x, y + \Delta y) - f(x, y) \quad (6.3)$$

One might guess that $\Delta z = \Delta_x z + \Delta_y z$ but as the following example proves, this is not true.

Example 1. For the function $z = xy$ let us find Δz and $\Delta_x z + \Delta_y z$ if $x = 2$, $y = 3$, $\Delta x = 0,2$ and $\Delta y = 0,1$.

First $\Delta_x z = (x + \Delta x)y - xy = y\Delta x = 3 \cdot 0, 2 = 0, 6$,
 second $\Delta_y z = x(y + \Delta y) - xy = x\Delta y = 2 \cdot 0, 1 = 0, 2$. Thus, $\Delta_x z + \Delta_y z = 0, 8$.

The total increment of the function $\Delta z = (x + \Delta x)(y + \Delta y) - xy = x\Delta y + y\Delta x + \Delta x\Delta y = 2 \cdot 0, 1 + 3 \cdot 0, 2 + 0, 2 \cdot 0, 1 = 0, 82$.

The total increment of the function of three variables $w = f(x, y, z)$ is defined as

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

If y and z are constants, we define

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

if x and z are constants, we define

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

and if x and y are constants, we define

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

6.4 Limit and continuity of functions of two variables

Suppose $P_0(x_0, y_0)$ is a fixed point in the domain of the function $z = f(x, y)$ and $P(x, y)$ is a moving point that approaches P_0 . We shall write $(x, y) \rightarrow (x_0, y_0)$ or $x \rightarrow x_0, y \rightarrow y_0$.

To find the limit of a function of one variable, we only needed to test the approach from the left and the approach from the right. If both approaches gave the same result, the function had a limit. To find the limit of a function of two variables however, we must show that the limit is the same no matter from which direction we approach (x_0, y_0)

The moving point P can approach the fixed point P_0 along whatever path: along the straight line, broken line, the arc of parabola etc. Independently of the path, the moving point P reaches to any neighborhood of $U_\delta(x_0, y_0)$ for arbitrary small $\delta > 0$.

Definition 1. The real number L is called the limit of the function $f(x, y)$ in limiting process $(x, y) \rightarrow (x_0, y_0)$, if $\forall \varepsilon > 0$ there exists the neighborhood $U_\delta(x_0, y_0)$ such that $|f(x, y) - L| < \varepsilon$ whenever $(x, y) \in U_\delta(x_0, y_0)$

In other words, L is the limit of the function $f(x, y)$ as $(x, y) \rightarrow (x_0, y_0)$, if the value of the function $f(x, y)$ can be made as close as desired to L by taking $P(x, y)$ in the neighborhood of $P_0(x_0, y_0)$ small enough.

This is denoted

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

Example 1. Find the limit $\lim_{(x,y) \rightarrow (0;0)} \frac{xy}{x^2 + y^2}$.

Let $(x, y) \rightarrow (0; 0)$ along the line $y = kx$. Then

$$\frac{xy}{x^2 + y^2} = \frac{x \cdot kx}{x^2 + k^2x^2} = \frac{k \cdot x^2}{x^2(1 + k^2)} = \frac{k}{1 + k^2}$$

This shows that the result depends on the choice of the slope of the line k . Therefore, the limit does not exist.

Often it is useful to convert the limit into polar coordinates, taking $x = \rho \cos \varphi$ and $y = \rho \sin \varphi$. Then $x^2 + y^2 = \rho^2$ and the limiting process $(x, y) \rightarrow (0; 0)$ is equivalent to $\rho \rightarrow 0$. In Example 1 we could write

$$\lim_{(x,y) \rightarrow (0;0)} \frac{xy}{x^2 + y^2} = \lim_{\rho \rightarrow 0} \frac{\rho \cos \varphi \rho \sin \varphi}{\rho^2 \cos^2 \varphi + \rho^2 \sin^2 \varphi} = \lim_{\rho \rightarrow 0} \frac{\rho^2 \cos \varphi \sin \varphi}{\rho^2} = \cos \varphi \sin \varphi$$

The result depends on the polar angle and this proves again that the limit does not exist.

Example 2. Find the limit $\lim_{(x,y) \rightarrow (0;0)} \frac{\sin(x^2 + y^2)}{x^2 + y^2}$.

Converting this limit into polar coordinates, we have

$$\lim_{(x,y) \rightarrow (0;0)} \frac{\sin(x^2 + y^2)}{x^2 + y^2} = \lim_{\rho \rightarrow 0} \frac{\sin(\rho^2 \cos^2 \varphi + \rho^2 \sin^2 \varphi)}{\rho^2 \cos^2 \varphi + \rho^2 \sin^2 \varphi} = \lim_{\rho \rightarrow 0} \frac{\sin \rho^2}{\rho^2} = 1$$

Definition 2. The function $f(x, y)$ is called continuous at the point $P_0(x_0, y_0)$, if

1. $\exists f(x_0, y_0)$
2. $\exists \lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$
3. $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = f(x_0, y_0)$

Definition 3. The function is called continuous in the domain D , if it is continuous at every point of this domain.

Let us denote the fixed point in Definition 2 by (x, y) and the moving point by $(x + \Delta x, y + \Delta y)$. Then $(x + \Delta x, y + \Delta y) \rightarrow (x, y)$ if and only if $(\Delta x, \Delta y) \rightarrow (0; 0)$. The third condition of continuity can be re-written

$$\lim_{(\Delta x, \Delta y) \rightarrow (0;0)} f(x + \Delta x, y + \Delta y) = f(x, y)$$

or

$$\lim_{(\Delta x, \Delta y) \rightarrow (0;0)} [f(x + \Delta x, y + \Delta y) - f(x, y)] = 0 \quad (6.4)$$

In square brackets of the last condition there is the total increment Δz of the function $z = f(x, y)$ and the condition of the continuity (6.4) at the point (x, y) is

$$\lim_{(\Delta x, \Delta y) \rightarrow (0, 0)} \Delta z = 0 \quad (6.5)$$

The equality (6.5) is called the **necessary and sufficient condition of continuity**.

6.5 Partial derivatives

Fix in the domain of the function of two variables $z = f(x, y)$ one point $P(x, y)$. Holding y constant and increasing the variable x by Δx we have the increment of the function $f(x, y)$

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

Definition 1. If there exists the limit

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \quad (6.6)$$

then this limit is called the *partial derivative* of the function $f(x, y)$ with respect to the variable x at the point (x, y) .

The partial derivative with respect to x is denoted also z'_x , $f'_x(x, y)$, $\frac{\partial f}{\partial x}$.

Holding x constant and increasing the variable y by Δy we have the increment of the function $f(x, y)$ as $\Delta_y z = f(x, y + \Delta y) - f(x, y)$.

Definition 2. If there exists the limit

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta_y z}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \quad (6.7)$$

then this limit is called $f'_y(x, y)$ the partial derivative of the function $f(x, y)$ with respect to the variable y at the point (x, y) .

The possible alternate notations for partial derivatives with respect to y are z'_y , $f'_y(x, y)$, $\frac{\partial f}{\partial y}$.

If we find the partial derivative with respect to the variable x the variable y is treated as constant. The only variable in Definition 1 is Δx . As well, finding the partial derivative with respect to the variable y the variable x is treated as constant. The only variable in Definition 2 is Δy . We need to pay very close attention to which variable we are differentiating with respect to. This is important because we are going to treat the other variable as

constant and then proceed with the derivative as if it was a function of a single variable. Consequently, all the rules of differentiation of functions of one variable hold if we find the partial derivatives.

Example 1. Find the partial derivatives with respect to both variables for the function $z = x^3y - x^2y^2$.

Finding the partial derivative with respect to x , y is treated as constant. Thus, by the difference rule and a constant rule we obtain

$$\frac{\partial z}{\partial x} = \frac{\partial}{\partial x}(x^3y) - \frac{\partial}{\partial x}(x^2y^2) = y \frac{\partial}{\partial x}(x^3) - y^2 \frac{\partial}{\partial x}(x^2) = y \cdot 3x^2 - y^2 \cdot 2x = 3x^2y - 2xy^2.$$

Finding the partial derivative with respect to y , x is treated as constant. By the rules of differentiation

$$\frac{\partial z}{\partial y} = \frac{\partial}{\partial y}(x^3y) - \frac{\partial}{\partial y}(x^2y^2) = x^3 \frac{\partial}{\partial y}(y) - x^2 \frac{\partial}{\partial y}(y^2) = x^3 - x^2 \cdot 2y = x^3 - 2x^2y$$

The chain rule is also still valid.

Example 2. Find the partial derivatives with respect to both variables for the function $z = \arctan \frac{x}{y}$.

The partial derivative with respect to x is

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

The partial derivative with respect to y is

$$\begin{aligned} \frac{\partial z}{\partial y} &= \frac{1}{1 + \left(\frac{x}{y}\right)^2} \cdot \frac{\partial z}{\partial y} \left(\frac{x}{y}\right) = \frac{y^2}{y^2 + x^2} \cdot x \frac{\partial}{\partial y} \left(\frac{1}{y}\right) \\ &= \frac{y^2}{x^2 + y^2} \cdot \left(-\frac{x}{y^2}\right) = -\frac{x}{x^2 + y^2} \end{aligned}$$

The partial derivatives of the function of three variables $w = f(x, y, z)$ with respect to variables x , y and z are defined as the limits

$$\begin{aligned} \frac{\partial w}{\partial x} &= \lim_{\Delta x \rightarrow 0} \frac{\Delta_x w}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x} \\ \frac{\partial w}{\partial y} &= \lim_{\Delta y \rightarrow 0} \frac{\Delta_y w}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y, z) - f(x, y, z)}{\Delta y} \end{aligned}$$

and

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

If we find the partial derivative with respect to one independent variable, the other independent variables are treated as constants.

Example 3. Find the partial derivatives with respect to all three independent variables for the function $w = x^{y^z}$.

Finding the partial derivative with respect to x , we have the power function with constant exponent y^z , therefore,

$$\frac{\partial w}{\partial x} = y^z x^{y^z-1}$$

To find the partial derivative with respect to y we use the chain rule. The outside function is the exponential function with constant base x and the variable exponent y^z , which is the power function with respect to y . By the chain rule

$$\frac{\partial w}{\partial y} = x^{y^z} \ln x \cdot z y^{z-1}$$

To find the partial derivative with respect to z we use the chain rule again. The outside function is the exponential function with constant base x . The inside function is another exponential function y^z with the constant base y . Thus

$$\frac{\partial w}{\partial z} = x^{y^z} \ln x \cdot y^z \ln y.$$

6.6 Total increment and total differential

Assume that the function $f(x, y)$ is continuous and has the continuous partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ at the point $P(x, y)$ and in some neighborhood of this point.

Adding and subtracting the same term $f(x, y + \Delta y)$ we represent the total increment of the function

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

In the first two terms y is constant and equal to $y + \Delta y$. In third and fourth term x is constant.

At the point P and in the neighborhood of this point there hold the assumptions of Lagrange's theorem. Consequently there exists $\bar{x} \in (x, x + \Delta x)$ such that

$$f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) = \frac{\partial f(\bar{x}, y + \Delta y)}{\partial x} \Delta x.$$

Next, by Lagrange's theorem there exists $\bar{y} \in (y, y + \Delta y)$ such that

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

Because of the continuity of partial derivatives

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\partial f(\bar{x}, y + \Delta y)}{\partial x} = \frac{\partial f(x, y)}{\partial x}$$

and

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\partial f(x, \bar{y})}{\partial y} = \frac{\partial f(x, y)}{\partial y}$$

The theorem about the infinitesimals yields

$$\frac{\partial f(\bar{x}, y + \Delta y)}{\partial x} = \frac{\partial f(x, y)}{\partial x} + \alpha$$

and

$$\frac{\partial f(x, \bar{y})}{\partial y} = \frac{\partial f(x, y)}{\partial y} + \beta$$

where α and β are the infinitesimals as $(\Delta x, \Delta y) \rightarrow (0; 0)$.

Now the total increment of the function converts to

$$\Delta z = \left(\frac{\partial f(x, y)}{\partial x} + \alpha \right) \Delta x + \left(\frac{\partial f(x, y)}{\partial y} + \beta \right) \Delta y$$

or

$$\Delta z = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \alpha \Delta x + \beta \Delta y. \quad (6.8)$$

In subsection 6.4 we have used the notation $\Delta \rho = \sqrt{\Delta x^2 + \Delta y^2}$. The conditions

$$\left| \frac{\Delta x}{\Delta \rho} \right| \leq 1$$

and

$$\left| \frac{\Delta y}{\Delta \rho} \right| \leq 1$$

mean that these are the bounded quantities. Thus, $\alpha \frac{\Delta x}{\Delta \rho}$ and $\beta \frac{\Delta y}{\Delta \rho}$ are infinitesimals as the products of the infinitesimals and a bounded quantities. The limit

$$\lim_{\Delta \rho \rightarrow 0} \frac{\alpha \Delta x + \beta \Delta y}{\Delta \rho} = \lim_{\Delta \rho \rightarrow 0} \alpha \frac{\Delta x}{\Delta \rho} + \lim_{\Delta \rho \rightarrow 0} \beta \frac{\Delta y}{\Delta \rho} = 0$$

means that $\alpha\Delta x + \beta\Delta y$ is an infinitesimal of the higher order with respect to $\Delta\rho$, i.e. Δx and Δy .

Definition. If in (6.8) the first sum is linear with respect to Δx and Δy and the second sum is an infinitesimal with respect to the same variables, then the linear part is called the *total differential* of the function $z = f(x, y)$ and denoted by dz . According to the definition

$$dz = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y$$

For the function $z = x$ the partial derivatives $\frac{\partial z}{\partial x} = 1$, $\frac{\partial z}{\partial y} = 0$ and $dz = dx = \Delta x$.

For the function $z = y$ the partial derivatives $\frac{\partial z}{\partial x} = 0$, $\frac{\partial z}{\partial y} = 1$ and $dz = dy = \Delta y$.

Consequently for the independent variables x and y the notions of differential and increment coincide and the total differential can be re-written as

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy. \quad (6.9)$$

Example 1. Find the total differential for the function $z = \arctan \frac{x}{y}$.

Using the partial derivatives found in Example 6.5, we obtain

$$dz = \frac{y}{x^2 + y^2}dx - \frac{x}{x^2 + y^2}dy = \frac{ydx - xdy}{x^2 + y^2}.$$

Example 2. Evaluate the total increment and the total differential for the function $z = \sqrt{x^2 + y^2}$, if $x = 3$, $y = 4$, $\Delta x = 0,2$ and $\Delta y = 0,1$.

By the formula of the total increment of the function (6.1) we get

$$\Delta z = \sqrt{3,2^2 + 4,1^2} - \sqrt{3^2 + 4^2} = \sqrt{27,05} - \sqrt{25} = 0,20096$$

To evaluate the total differential we find

$$\frac{\partial z}{\partial x} = \frac{1}{2\sqrt{x^2 + y^2}} \cdot 2x = \frac{x}{\sqrt{x^2 + y^2}}$$

and

$$\frac{\partial z}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$$

Then

$$dz = \frac{3}{\sqrt{3^2 + 4^2}} \cdot 0,2 + \frac{4}{\sqrt{3^2 + 4^2}} \cdot 0,1 = \frac{0,6}{5} + \frac{0,4}{5} = 0,2$$

We see that the difference between the total increment and the total differential is less than 0,001, which is less by two orders of values with respect to Δx and Δy .

The last fact gives us the possibility to compute the approximate values of functions of two variables using the total differential. If Δx and Δy are sufficiently small, then Δz and dz differ by the quantity, which is the infinitesimal of a higher order with respect to Δx and Δy . We can write

$$\Delta z \approx dz$$

or

$$f(x + \Delta x, y + \Delta y) - f(x, y) \approx \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

This gives us the formula of approximate computation

$$f(x + \Delta x, y + \Delta y) \approx f(x, y) + \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y \quad (6.10)$$

Example 3. Using the total differential, compute $2,03^3 \cdot 0,96^2$.

Here we choose the function $f(x, y) = x^3 y^2$ and the values $x = 2, y = 1, \Delta x = 0,03$ and $\Delta y = -0,04$. The partial derivatives are

$$\frac{\partial f}{\partial x} = 3x^2 y^2$$

and

$$\frac{\partial f}{\partial y} = 2x^3 y$$

The value of the function at the point chosen $f(2, 1) = 8 \cdot 1 = 8$ and the values of partial derivatives are $\frac{\partial f}{\partial x} = 3 \cdot 4 \cdot 1 = 12$ and $\frac{\partial f}{\partial y} = 2 \cdot 8 \cdot 1 = 16$.

By the formula (6.13)

$$(2 + 0,03)^3 \cdot (1 - 0,04)^2 = 8 + 12 \cdot 0,03 - 16 \cdot 0,04 = 7,72$$

Suppose that the function of three variables $w = f(x, y, z)$ and the partial derivatives $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$ and $\frac{\partial w}{\partial z}$ are continuous at the point $P(x, y, z)$ and in some neighborhood of this point. Analogously to the formula (6.8) it is possible to prove that the total increment of the function can be expressed as

$$\Delta w = \frac{\partial w}{\partial x} \Delta x + \frac{\partial w}{\partial y} \Delta y + \frac{\partial w}{\partial z} \Delta z + \alpha \Delta x + \beta \Delta y + \gamma \Delta z, \quad (6.11)$$

where $\alpha\Delta x + \beta\Delta y + \gamma\Delta z$ is an infinitesimal of a higher order with respect to $\Delta\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$. The expression

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz \quad (6.12)$$

is called the total differential of the function $w = f(x, y, z)$. Again, for the independent variables x, y and z the notions of the increment and differential coincide, i.e. $dx = \Delta x$, $dy = \Delta y$ and $dz = \Delta z$.

Example 4. Find the total differential for the function $w = x^{y^z}$.

Using the partial derivatives found in Example 6.5, we obtain

$$\begin{aligned} dw &= y^z x^{y^z-1} dx + x^{y^z} \ln x \cdot zy^{z-1} dy + x^{y^z} \ln x \cdot y^z \ln y = \\ &= y^z x^{y^z} \left(\frac{dx}{x} + \frac{z \ln x dy}{y} + \ln x \ln y \right) \end{aligned}$$

As well as for the function of two variables there holds the formula of approximate computation

$$f(x + \Delta x, y + \Delta y, z + \Delta z) \approx f(x, y, z) + \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \frac{\partial f}{\partial z} \Delta z \quad (6.13)$$

6.7 Partial derivatives of implicit function

Consider the function

$$F(x, y) = 0 \quad (6.14)$$

given implicitly. This equation determines the variable y as the function of x (in general case not one-valued). The graph of this function is a curve in the plane (Figure 6.9). Suppose that the function $F(x, y)$ is continuous and it has the continuous partial derivatives at the point $P(x, y)$ and in some neighborhood of this point. In addition suppose that at $P(x, y)$ the partial derivative $F'_y(x, y) \neq 0$. Let us deduce the formula to find the derivative $\frac{dy}{dx}$, using the partial derivatives of the function $F(x, y)$.

Let us fix the point $P(x, y)$ on the graph of given function. The coordinates of this point satisfy the equation (6.14). Change x by Δx and find on the graph the value of $y + \Delta y$ related to $x + \Delta x$. As $Q(x + \Delta x, y + \Delta y)$ is a point on the graph again, the coordinates of this point also satisfy the equation

$$F(x + \Delta x, y + \Delta y) = 0. \quad (6.15)$$

Subtracting from the equation (6.15) the equation (6.14), we obtain

$$F(x + \Delta x, y + \Delta y) - F(x, y) = 0$$

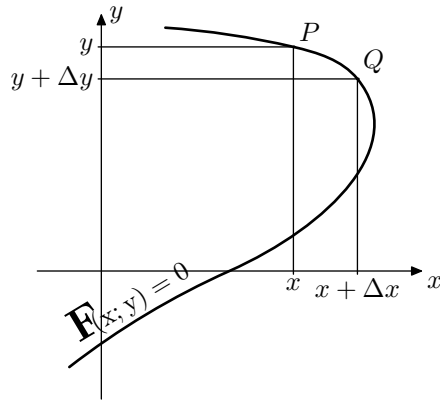


Figure 6.9. The graph of the function $F(x, y) = 0$

The left side of the last equality is the total increment of the function $F(x, y)$ and the equality can be re-written

$$\Delta F = 0$$

Because of the assumptions made in the beginning of this subsection this equality converts by (6.8) to

$$\frac{\partial F}{\partial x} \Delta x + \frac{\partial F}{\partial y} \Delta y + \alpha \Delta x + \beta \Delta y = 0$$

which yields

$$\left(\frac{\partial F}{\partial y} + \beta \right) \Delta y = - \left(\frac{\partial F}{\partial x} + \alpha \right) \Delta x$$

or

$$\frac{\Delta y}{\Delta x} = - \frac{\frac{\partial F}{\partial x} + \alpha}{\frac{\partial F}{\partial y} + \beta}$$

Find the limits of both sides of this equality as $\Delta x \rightarrow 0$. The limit of the left side is by the definition of the derivative $\frac{dy}{dx}$. The function is continuous, consequently if $\Delta x \rightarrow 0$ then $\Delta y \rightarrow 0$. Knowing that α and β are the infinitesimals as $(\Delta x, \Delta y) \rightarrow (0; 0)$, that is $\lim_{\Delta x \rightarrow 0} \alpha = 0$ and $\lim_{\Delta x \rightarrow 0} \beta = 0$, the limit of the right side of the equality is

$$- \frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}$$

Thus, to find the derivative of the function given implicitly we have the formula

$$\frac{dy}{dx} = -\frac{F'_x}{F'_y} \quad (6.16)$$

Example 1. Find $\frac{dy}{dx}$ for $x^4 + y^4 - a^2x^2y^2 = 0$.

Here $F(x, y) = x^4 + y^4 - a^2x^2y^2$, so $F'_x = 4x^3 - 2a^2xy^2$ and $F'_y = 4y^3 - 2a^2x^2y$. By the formula (6.16)

$$\frac{dy}{dx} = -\frac{4x^3 - 2a^2xy^2}{4y^3 - 2a^2x^2y} = -\frac{x(2x^2 - a^2y^2)}{y(2y^2 - a^2x^2)}.$$

The equation $F(x, y, z) = 0$ relates to pairs of (x, y) some value(s) of the variable z . In other words, this equation defines z as a function of x and y . Assume that the function $F(x, y, z)$ is continuous and has the continuous partial derivatives $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ and $\frac{\partial F}{\partial z}$ at the point $P(x, y, z)$ and in some neighborhood of this point. Moreover assume that $F'_z(x, y, z) \neq 0$ at $P(x, y, z)$.

If we find the partial derivative of the function z with respect to x the variable y is treated as constant. In this case in the equation $F(x, y, z) = 0$ there are only two variables x and z and by (6.16) we obtain

$$\frac{\partial z}{\partial x} = -\frac{F'_x}{F'_z} \quad (6.17)$$

If we repeat this reasoning for y we have

$$\frac{\partial z}{\partial y} = -\frac{F'_y}{F'_z} \quad (6.18)$$

Example 2. Find the partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for the function of two variables $x^2 + y^2 + z^2 = r^2$ given implicitly.

As $F'_x = 2x$, $F'_y = 2y$ and $F'_z = 2z$ we obtain by the formula (6.17) the partial derivative

$$\frac{\partial z}{\partial x} = -\frac{x}{z}$$

and by the formula (6.18) the partial derivative

$$\frac{\partial z}{\partial y} = -\frac{y}{z}$$

6.8 Partial derivatives of composite functions

Suppose that the variable z is a function of two variables u and v , denote $z = f(u, v)$, and u and v are the functions of two independent variables x and y , denote $u = \varphi(x, y)$ and $v = \psi(x, y)$. Then z is a composite function with respect to x and y , i.e.

$$z = f(\varphi(x, y), \psi(x, y)) = F(x, y)$$

Let us fix a point $P(x, y)$ in the common domain of the functions $u = \varphi(x, y)$ and $v = \psi(x, y)$. Then the related point (u, v) in the (u, v) -plane is also fixed. Suppose that the functions u and v are continuous and have the continuous partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$ and $\frac{\partial v}{\partial y}$ at the point $P(x, y)$ and in some neighborhood of this point. Also assume that the function z is continuous and has the continuous partial derivatives $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ at the related point (u, v) and in some neighborhood of this point.

Holding the argument y constant, change the argument x by Δx . The increments of the functions u and v with respect to x $\Delta_x u$ and $\Delta_x v$ are simultaneously the increments of the arguments the function $z = f(u, v)$. By the formula (6.8) the increment of the function z can be written as

$$\Delta z = \frac{\partial z}{\partial u} \Delta_x u + \frac{\partial z}{\partial v} \Delta_x v + \alpha \Delta_x u + \beta \Delta_x v$$

where α and β are the infinitesimals as $(\Delta_x u, \Delta_x v) \rightarrow (0; 0)$. Dividing both sides of the last equality by Δx , we obtain

$$\frac{\Delta z}{\Delta x} = \frac{\partial z}{\partial u} \frac{\Delta_x u}{\Delta x} + \frac{\partial z}{\partial v} \frac{\Delta_x v}{\Delta x} + \alpha \frac{\Delta_x u}{\Delta x} + \beta \frac{\Delta_x v}{\Delta x} \quad (6.19)$$

Find the limits of both sides as $\Delta x \rightarrow 0$. The limit of the left side of (6.19) is the partial derivative of z with respect to x because we have got the increment of z as the result of change of x while y is constant.

We have assumed the existence $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial x}$. The differentiability of functions u and v with respect to x yields the continuity of those functions by x . Therefore,

$$\lim_{\Delta x \rightarrow 0} \Delta_x u = 0$$

and

$$\lim_{\Delta x \rightarrow 0} \Delta_x v = 0$$

Consequently

$$\lim_{\Delta x \rightarrow 0} \alpha = \lim_{\Delta x \rightarrow 0} \beta = 0$$

Using the definitions of partial derivatives of the functions u and v with respect to x , the limit of the right side of (6.19) is

$$\frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x} + 0 \cdot \frac{\partial u}{\partial x} + 0 \cdot \frac{\partial v}{\partial x}$$

We conclude that the partial derivative of the composite function $z = F(x, y)$ with respect to x is

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x} \quad (6.20)$$

Holding the argument x constant and changing y by Δy , we obtain after the similar reasoning the partial derivative of the composite function z with respect to the variable y

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y} \quad (6.21)$$

Example 1. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for $z = \ln(u^2 + v)$, $u = e^{x+y^2}$ and $v = x^2 + y$.

According to the formulas (6.20) and (6.21) we have to find six partial derivatives

$$\begin{aligned} \frac{\partial z}{\partial u} &= \frac{2u}{u^2 + v}, & \frac{\partial z}{\partial v} &= \frac{1}{u^2 + v}; \\ \frac{\partial u}{\partial x} &= e^{x+y^2}, & \frac{\partial u}{\partial y} &= 2ye^{x+y^2}; \\ \frac{\partial v}{\partial x} &= 2x, & \frac{\partial v}{\partial y} &= 1; \end{aligned}$$

By (6.20) we have

$$\frac{\partial z}{\partial x} = \frac{2u}{u^2 + v} e^{x+y^2} + \frac{1}{u^2 + v} 2x = \frac{2}{u^2 + v} (ue^{x+y^2} + x)$$

and by (6.21)

$$\frac{\partial z}{\partial y} = \frac{2u}{u^2 + v} 2ye^{x+y^2} + \frac{1}{u^2 + v} = \frac{1}{u^2 + v} (4uye^{x+y^2} + 1)$$

Remark. If z is a function of three variables $z = f(u, v, w)$ and in addition to the u and v there is $w = \chi(x, y)$, then the partial derivatives of the composite function z with respect to the variables x and y can be found by the formulas

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x} + \frac{\partial z}{\partial w} \frac{\partial w}{\partial x} \quad (6.22)$$

and

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y} + \frac{\partial z}{\partial w} \frac{\partial w}{\partial y} \quad (6.23)$$

Next, let z be a function of three variables x , u and v $z = f(x, u, v)$, where $u = \varphi(x)$ and $v = \psi(x)$. In this case z is a composite function of one variable x

$$z = f(x, \varphi(x), \psi(x))$$

The derivative of that function $\frac{dz}{dx}$ we obtain using (6.22). As the derivative $\frac{dx}{dx} = 1$ and u and v are the functions of one variable, then

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial u} \frac{du}{dx} + \frac{\partial z}{\partial v} \frac{dv}{dx}. \quad (6.24)$$

The derivative in (6.24) is called the *total derivative*.

Example 2. Find $\frac{dz}{dx}$ for $z = x^2 + \sqrt{y}$ and $y = x^2 + 1$.

Here z is the function of two variables x and y , where y is the function of the variable x . In this case the formula (6.24) gives

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \frac{dy}{dx} = 2x + \frac{1}{2\sqrt{y}} \cdot 2x = x \left(2 + \frac{1}{\sqrt{y}} \right) = x \left(2 + \frac{1}{\sqrt{x^2 + 1}} \right).$$

Let us find the total differential of the composite function $z = f(u, v)$, $u = \varphi(x, y)$ and $v = \psi(x, y)$ considered in the beginning of this subsection

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \quad (6.25)$$

Substituting by (6.20) and (6.21) the partial derivatives, we obtain

$$dz = \left(\frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x} \right) dx + \left(\frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y} \right) dy$$

or after conversion

$$dz = \frac{\partial z}{\partial u} \left(\frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \right) + \frac{\partial z}{\partial v} \left(\frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy \right)$$

In the parenthesis there are the total differentials

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

and

$$dv = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy$$

Hence,

$$dz = \frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv \tag{6.26}$$

Comparing the expressions of total differentials (6.25) and (6.26), we notice that the total differential has the same form. It is not important whether u and v are independent variables or functions of other variables x and y . This is called the *invariance property of a total differential*.

6.9 Higher order partial derivatives

As we have seen in many examples, the partial derivatives of the function $z = f(x, y)$ $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are in general functions of two variables again. Thus, it is possible to differentiate both of them with respect to x and y .

Definition 1. The partial derivative with respect to x of the partial derivative $\frac{\partial z}{\partial x}$ is called the *second order partial derivative with respect to x* and denoted $\frac{\partial^2 z}{\partial x^2}$ (to be read *de-squared-zed de-ex-squared*), that means

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right)$$

Definition 2. The partial derivative with respect to y of the partial derivative $\frac{\partial z}{\partial x}$ is called the *second order partial derivative with respect to x and y* and denoted $\frac{\partial^2 z}{\partial x \partial y}$ (to be read *de-squared-zed de-ex-de-y*). By this definition

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right)$$

Definition 3. The partial derivative with respect to x of the partial derivative $\frac{\partial z}{\partial y}$ is called the *second order partial derivative with respect to y and x* and denoted $\frac{\partial^2 z}{\partial y \partial x}$, that is

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right)$$

Definition 4. The partial derivative with respect to y of the partial derivative $\frac{\partial z}{\partial y}$ is called the *second order partial derivative with respect to y* and denoted $\frac{\partial^2 z}{\partial y^2}$, i.e

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial y} \right)$$

The second and third second order partial derivatives are often called *mixed partial derivatives* since we are taking derivatives with respect to more than one variable.

The second order partial derivatives are denoted also z''_{xx} , z''_{xy} , z''_{yx} and z''_{yy} or $f''_{xx}(x, y)$, $f''_{xy}(x, y)$, $f''_{yx}(x, y)$ and $f''_{yy}(x, y)$.

The second order partial derivatives are the functions of two variables x and y again. Hence, all four second order partial derivatives can be differentiated with respect to x and y . So we define eight third order partial derivatives

$$\begin{aligned} \frac{\partial^3 z}{\partial x^3} &= \frac{\partial}{\partial x} \left(\frac{\partial^2 z}{\partial x^2} \right), & \frac{\partial^3 z}{\partial x^2 \partial y} &= \frac{\partial}{\partial y} \left(\frac{\partial^2 z}{\partial x^2} \right) \\ \frac{\partial^3 z}{\partial x \partial y \partial x} &= \frac{\partial}{\partial x} \left(\frac{\partial^2 z}{\partial x \partial y} \right), & \frac{\partial^3 z}{\partial x \partial y^2} &= \frac{\partial}{\partial y} \left(\frac{\partial^2 z}{\partial x \partial y} \right) \\ \frac{\partial^3 z}{\partial y \partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial^2 z}{\partial y \partial x} \right), & \frac{\partial^3 z}{\partial y \partial x \partial y} &= \frac{\partial}{\partial y} \left(\frac{\partial^2 z}{\partial y \partial x} \right) \\ \frac{\partial^3 z}{\partial y^2 \partial x} &= \frac{\partial}{\partial x} \left(\frac{\partial^2 z}{\partial y^2} \right), & \frac{\partial^3 z}{\partial y^3} &= \frac{\partial}{\partial y} \left(\frac{\partial^2 z}{\partial y^2} \right) \end{aligned}$$

Example 1. Find all second order partial derivatives for $z = \arctan \frac{x}{y}$.

In Example 2 of subsection 6.5 we have found

$$\frac{\partial z}{\partial x} = \frac{y}{x^2 + y^2} \quad \text{and} \quad \frac{\partial z}{\partial y} = -\frac{x}{x^2 + y^2}$$

We find

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

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{y}{x^2 + y^2} \right) = \frac{x^2 + y^2 - y \cdot 2y}{(x^2 + y^2)^2} = \frac{x^2 - y^2}{(x^2 + y^2)^2}$$

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial x} \left(-\frac{x}{x^2 + y^2} \right) = -\frac{x^2 + y^2 - x \cdot 2x}{(x^2 + y^2)^2} = \frac{x^2 - y^2}{(x^2 + y^2)^2}$$

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

These results suggest a question, are the mixed second order partial derivatives

$$\frac{\partial^2 z}{\partial x \partial y} \quad \text{and} \quad \frac{\partial^2 z}{\partial y \partial x}$$

equal. The next theorem says that if the function is smooth enough this will always be the case.

Theorem. If the function $z = f(x, y)$ and its partial derivatives $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, $\frac{\partial^2 z}{\partial x \partial y}$ and $\frac{\partial^2 z}{\partial y \partial x}$ are continuous at the point P and on some neighborhood of this point, then at the point P

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$$

This theorem says that if the partial derivatives to be evaluated are continuous, then the result of repeated differentiation is independent of the order in which it is performed.

Therefore, if the partial derivatives involved are continuous, the also holds

$$\frac{\partial^4 z}{\partial x \partial y \partial x \partial y} = \frac{\partial^4 z}{\partial x^2 \partial y^2} = \frac{\partial^4 z}{\partial y^2 \partial x^2}$$

Analogous theorem is valid also for the functions of three etc. variables.

Example 2. Find the third order partial derivatives $\frac{\partial^3 w}{\partial x \partial y \partial z}$ and $\frac{\partial^3 w}{\partial z \partial x \partial y}$ for the function of three variables $w = e^x \sin(yz)$.

First we find

$$\frac{\partial w}{\partial x} = e^x \sin(yz)$$

second

$$\frac{\partial^2 w}{\partial x \partial y} = e^x \cos(yz) \cdot z = z e^x \cos(yz)$$

and third

$$\frac{\partial^3 w}{\partial x \partial y \partial z} = e^x \cos(yz) + z(-e^x \sin(yz)) \cdot y = e^x [\cos(yz) - yz \sin(yz)]$$

To find the second third order partial derivative, we find

$$\frac{\partial w}{\partial z} = y e^x \cos(yz)$$

next

$$\frac{\partial^2 w}{\partial z \partial x} = ye^x \cos(yz)$$

and finally

$$\frac{\partial^3 w}{\partial z \partial x \partial y} = e^x \cos(yz) - ye^x \sin(yz) \cdot z = e^x [\cos(yz) - yz \sin(yz)]$$

6.10 Tangent line of space curve

In this subsection we want to look at an application of derivatives for vector functions. In Mathematical analysis I we've used the fact that the derivative of a function is the slope of the tangent line. For vector functions we get the similar result.

Consider the space curve given by parametric equations

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases} \quad (6.27)$$

or using the vector notation

$$\mathbf{r}(t) = (x(t), y(t), z(t))$$

If we fix a value of the parameter t we get the fixed point on the space curve $P(x(t); y(t); z(t))$. Assume that the functions of the parameter t (6.27) are differentiable at P .

Changing the fixed value t by Δt , we get the value $t + \Delta t$ and to this value there corresponds another point $Q(x(t + \Delta t), y(t + \Delta t), z(t + \Delta t))$ on space curve

The vector $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ is the direction vector of the secant line PQ and the coordinates of this vector are $\overrightarrow{PQ} = (\Delta x, \Delta y, \Delta z)$, where

$$\begin{aligned} \Delta x &= x(t + \Delta t) - x(t) \\ \Delta y &= y(t + \Delta t) - y(t) \\ \Delta z &= z(t + \Delta t) - z(t) \end{aligned}$$

Multiplying this vector by scalar $\frac{1}{\Delta t}$, we obtain the vector

$$\frac{1}{\Delta t} \overrightarrow{PQ} = \left(\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t} \right)$$

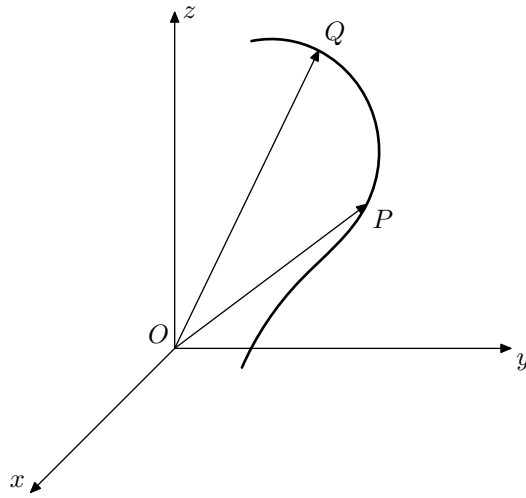


Figure 6.10.

which has the same direction as \overrightarrow{PQ} .

Since the functions (6.27) are differentiable at t then the limits of the coordinates of vector $\frac{1}{\Delta t}\overrightarrow{PQ}$ as $\Delta t \rightarrow 0$ are the derivatives of coordinate functions by parameter at the point P

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} = \dot{x}$$

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta y}{\Delta t} = \dot{y}$$

and

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta z}{\Delta t} = \dot{z}$$

Denoting by $\dot{\mathbf{r}} = (\dot{x}, \dot{y}, \dot{z})$, we obtain

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t}\overrightarrow{PQ} = \dot{\mathbf{r}}$$

The differentiability at P yields the continuity at this point. Thus,

$$\lim_{\Delta t \rightarrow 0} x(t + \Delta t) = x(t)$$

$$\lim_{\Delta t \rightarrow 0} y(t + \Delta t) = y(t)$$

and

$$\lim_{\Delta t \rightarrow 0} z(t + \Delta t) = z(t)$$

i.e. the point Q approaches along the curve to the point P and the direction vector \overrightarrow{PQ} approaches to the direction vector of the tangent line drawn to the curve at P . Since $\frac{1}{\Delta t}\overrightarrow{PQ}$ has always the same direction as \overrightarrow{PQ} despite of how close the point Q is to the point P then

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t}\overrightarrow{PQ}$$

is the direction vector of the tangent line at P , that means the direction vector of the tangent line at P is

$$\dot{\mathbf{r}} = (\dot{x}, \dot{y}, \dot{z})$$

Since at P the value of the parameter t has been fixed, then $x(t)$, $y(t)$ and $z(t)$ are the coordinates of the fixed point and the canonical equations of the tangent line at P are

$$\frac{x - x(t)}{\dot{x}} = \frac{y - y(t)}{\dot{y}} = \frac{z - z(t)}{\dot{z}} \quad (6.28)$$

Example. Find the canonical equations of the screw line $x = a \cos t$, $y = a \sin t$, $z = bt$ at the point where $t = \frac{\pi}{3}$.

First we find the coordinates of the corresponding point on the screw line $x\left(\frac{\pi}{3}\right) = \frac{a}{2}$, $y\left(\frac{\pi}{3}\right) = \frac{a\sqrt{3}}{2}$ and $z\left(\frac{\pi}{3}\right) = \frac{b\pi}{3}$.

Next we find the derivatives by parameter $\dot{x} = -a \sin t$, $\dot{y} = a \cos t$ and $\dot{z} = b$. Hence, at the given point the direction vector of the tangent line is

$$\dot{\mathbf{r}} = \left(-a \sin \frac{\pi}{3}, a \cos \frac{\pi}{3}, b\right) = \left(-\frac{a\sqrt{3}}{2}, \frac{a}{2}, b\right)$$

and by (6.28) the canonical equations of the tangent line are

$$\frac{x - \frac{a}{2}}{-\frac{a\sqrt{3}}{2}} = \frac{y - \frac{a\sqrt{3}}{2}}{\frac{a}{2}} = \frac{z - \frac{b\pi}{3}}{b}$$

or

$$\frac{2x - a}{-a\sqrt{3}} = \frac{2y - a\sqrt{3}}{a} = \frac{3z - b\pi}{3b}$$

The parametrical equations of this tangent line are

$$\begin{aligned}x &= \frac{a}{2} - \frac{a\sqrt{3}}{2}t \\y &= \frac{a\sqrt{3}}{2} + \frac{a}{2}t \\z &= \frac{b\pi}{3} + bt\end{aligned}$$

6.11 Tangent plane and normal line of surface

The graph of the function of two variables is a surface in space. Consider the function of two variable given implicitly

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

This is an equation of some surface. Let us fix one point $P(x, y, z)$ on this surface.

Definition 1. The point of the surface P is called the *regular point* if there exist all three partial derivatives $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ and $\frac{\partial F}{\partial z}$ and at least one of them does not equal to zero. The point of the surface is called *singular* if all three partial derivatives equal to zero or at least one of them does not exist.

We can draw infinitely many curves on the surface that pass the point P .

Definition 2. The tangent line of whatever curve on the surface passing the point P is called the tangent line of the surface at the point P .

Theorem. If P is a regular point of the surface, then all tangent lines of the surface at P lie on the same plane.

Proof. Choose whatever curve on the surface passing P . Suppose the parametrical equations of this curve are $x = x(t)$, $y = y(t)$ and $z = z(t)$. Since the curve is on the surface, then any point of the curve is on the surface and substituting the equations of the curve into the equation of the surface gives the identity

$$F(x(t), y(t), z(t)) \equiv 0$$

We differentiate both sides of this identity by the parameter using the chain rule

$$\frac{\partial F}{\partial x} \frac{dx}{dt} + \frac{\partial F}{\partial y} \frac{dy}{dt} + \frac{\partial F}{\partial z} \frac{dz}{dt} = 0$$

or

$$\frac{\partial F}{\partial x} \dot{x} + \frac{\partial F}{\partial y} \dot{y} + \frac{\partial F}{\partial z} \dot{z} = 0$$

The left side of this equation is the scalar product of two vectors

$$\vec{n} = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right)$$

and

$$\dot{\mathbf{r}} = (\dot{x}, \dot{y}, \dot{z})$$

As we know from the previous subsection, the vector $\dot{\mathbf{r}}$ is the direction vector of the curve at P . Since

$$\vec{n} \cdot \dot{\mathbf{r}} = 0$$

and we chose the random curve on the surface going through P then the tangent line of the whatever curve on the surface at P is perpendicular to the same vector \vec{n} . Hence all the tangent lines of the surface at P are perpendicular to the same vector, that means they lie on the same plane.

Definition 3. The plain containing all tangent lines of the surface passing P is called the *tangent plane of the surface at P* . The vector

$$\vec{n} = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right)$$

is called the *normal vector of the surface*.

Denote the fixed point on the surface by $P_0(x_0, y_0, z_0)$. Assuming that P_0 is a regular point, we can compute the normal vector \vec{n} at this point. If $P(x, y, z)$ is a random point on the tangent plane then $\vec{n} \cdot \overrightarrow{P_0P} = 0$. The coordinates of the vector $\overrightarrow{P_0P}$ are $\overrightarrow{P_0P} = (x - x_0, y - y_0, z - z_0)$ and the equation of the tangent plane

$$\frac{\partial F}{\partial x}(x - x_0) + \frac{\partial F}{\partial y}(y - y_0) + \frac{\partial F}{\partial z}(z - z_0) = 0 \quad (6.29)$$

Example 1. Find the equation of the tangent plane of the sphere $x^2 + y^2 + z^2 = 3$ at the point $P_0(1; 1; 1)$.

It is easy to check that the given point is on the sphere. The equation of the sphere $x^2 + y^2 + z^2 - 3 = 0$ determines the function $F(x, y, z) = x^2 + y^2 + z^2 - 3$. Find the partial derivatives

$$\frac{\partial F}{\partial x} = 2x; \quad \frac{\partial F}{\partial y} = 2y; \quad \frac{\partial F}{\partial z} = 2z$$

At the point $P_0(1; 1; 1)$

$$\frac{\partial F}{\partial x} = 2; \quad \frac{\partial F}{\partial y} = 2; \quad \frac{\partial F}{\partial z} = 2$$

Thus, the normal vector is $\vec{n} = (2; 2; 2)$ and the equation of the tangent plane (6.29)

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

After dividing by 2 on removing the parenthesis we get

$$x + y + z = 3$$

At the singular point the tangent plane may not exist. One example of the singular point is the vertex $O(0; 0; 0)$ of the cone $z^2 = x^2 + y^2$ because in this case

$$\begin{aligned} x^2 + y^2 - z^2 &= 0 \\ \frac{\partial F}{\partial x} &= 2x; \quad \frac{\partial F}{\partial y} = 2y; \quad \frac{\partial F}{\partial z} = -2z \end{aligned}$$

At the point $O \vec{n} = (0; 0; 0)$ and there does not exist uniquely determined tangent plane.

If the surface is a graph of the explicit function $z = f(x, y)$ then we can write this function implicitly $f(x, y) - z = 0$. Denoting $F(x, y, z) = f(x, y) - z$, we get

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x}; \quad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y}; \quad \frac{\partial F}{\partial z} = -1$$

and the equation of the tangent plane (6.29)

$$\frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) - (z - z_0) = 0$$

or

$$z - z_0 = \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) \quad (6.30)$$

Example 2. Find the tangent plane of the surface $z = 3x^2 - xy$ at the point $(1; 2; 1)$.

In this example $f(x, y) = 3x^2 - xy$, hence $f'_x = 6x - y$, $f'_y = -x$ and the values of these partial derivatives at the point $(1; 2; 1)$ are $f'_x(1; 2) = 4$ and $f'_y(1; 2) = -1$. The equation of the tangent plane (6.30) is

$$z - 1 = 4(x - 1) - 1(y - 2)$$

or

$$4x - y - z = 1$$

Denoting in (6.30) $x - x_0 = \Delta x$ and $y - y_0 = \Delta y$, we get the equation

$$z - z_0 = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y$$

whose right side is the total differential of the function $z = f(x, y)$ at P_0 .

This yields the geometric meaning of the total differential of the function of two variables. The total differential of the function of two variables equals to the increment of the third coordinate z on the tangent plane if x changes by Δx and y changes by Δy .

Definition 4. The line perpendicular to the tangent plane of the surface at the point P_0 is called the *normal line*.

The direction vector of the normal line of the surface $F(x, y, z) = 0$ is the normal vector of the surface (i.e. the normal vector of the tangent plane) $\vec{n} = (F'_x, F'_y, F'_z)$. Denoting the coordinates of the point P_0 on the surface by $P_0(x_0, y_0, z_0)$, we obtain the canonical equations of the normal line

$$\frac{x - x_0}{\frac{\partial F}{\partial x}} = \frac{y - y_0}{\frac{\partial F}{\partial y}} = \frac{z - z_0}{\frac{\partial F}{\partial z}} \quad (6.31)$$

Example 3. Find the canonical equations of the normal line of sphere $x^2 + y^2 + z^2 = 3$ at the point $P_0(1; 1; 1)$.

According to the data obtained in Example 1 we can write

$$\frac{x - 1}{2} = \frac{y - 1}{2} = \frac{z - 1}{2}$$

or

$$x = y = z$$

6.12 Directional derivative

Up to now for the function of two variables $z = f(x, y)$ we've only looked at the two partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$. Recall that these derivatives represent the rate of change of f as we vary x (holding y fixed) and as we vary y (holding x fixed) respectively. We now need to discuss how to find the rate of change of $f(x, y)$ if we allow both x and y to change simultaneously. In other words how to find the rate of change of $f(x, y)$ in the direction of vector $\vec{s} = (\Delta x, \Delta y)$.

The goal is to obtain the formula to compute the derivative of the function $z = f(x, y)$ at the point $P(x, y)$ in the direction of the vector $\vec{s} = (\Delta x, \Delta y)$ (Figure 6.10).

Assume that the function $z = f(x, y)$ and its partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are continuous at P and in some neighborhood of this point.

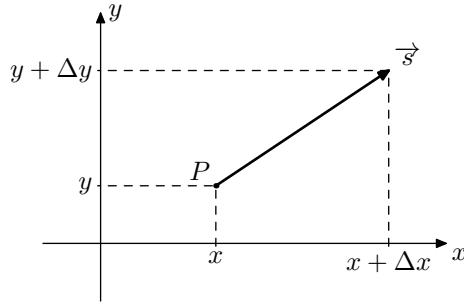


Figure 6.11.

Denote the length of the vector \vec{s} by $\Delta s = \sqrt{\Delta x^2 + \Delta y^2}$. By the (6.8) the total increment of the function has the form

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

where ε_1 and ε_2 are infinitesimals as $\Delta s \rightarrow 0$. Dividing the last equality by the length of the vector \vec{s} gives

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

The ratios $\frac{\Delta x}{\Delta s}$ and $\frac{\Delta y}{\Delta s}$ are the coordinates of the unit vector \vec{s}° in direction of the vector \vec{s} . Denoting by α and β the angles that \vec{s} forms with the coordinate axes, it's obvious that

$$\frac{\Delta x}{\Delta s} = \cos \alpha \quad \text{and} \quad \frac{\Delta y}{\Delta s} = \cos \beta$$

Therefore, these ratios, i.e. the coordinates of the unit vector in direction of the vector \vec{s} are called the *directional cosines* of that vector.

Definition. The limit

$$\lim_{\Delta s \rightarrow 0} \frac{\Delta z}{\Delta s}$$

is called the derivative of z at the point P in the direction of the vector \vec{s} and denoted $\frac{\partial z}{\partial \vec{s}}$. Since

$$\lim_{\Delta s \rightarrow 0} \left(\varepsilon_1 \frac{\Delta x}{\Delta s} + \varepsilon_2 \frac{\Delta y}{\Delta s} \right) = 0$$

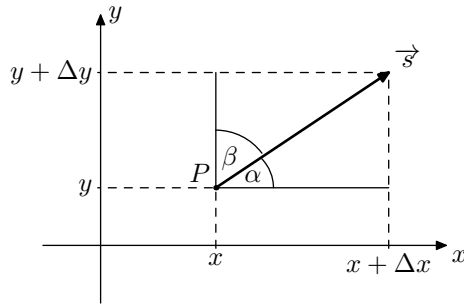


Figure 6.12.

we have the formula to compute the directional derivative

$$\frac{\partial z}{\partial \vec{s}} = \frac{\partial z}{\partial x} \cos \alpha + \frac{\partial z}{\partial y} \cos \beta \quad (6.32)$$

Example 1. Find the derivatives of the function $z = x^2 + y^2$ at the point $P(1; 1)$ in directions of vectors $\vec{s}_1 = (1; 1)$ and $\vec{s}_2 = (1; -1)$.

First we evaluate the partial derivatives of z at P

$$\frac{\partial z}{\partial x} = 2x \Big|_P = 2$$

and

$$\frac{\partial z}{\partial y} = 2y \Big|_P = 2$$

The length of the vector \vec{s}_1 is $\Delta s_1 = \sqrt{2}$, the directional cosines are $\cos \alpha = \frac{1}{\sqrt{2}}$ and $\cos \beta = \frac{1}{\sqrt{2}}$. Hence,

$$\frac{\partial z}{\partial \vec{s}_1} = 2 \cdot \frac{1}{\sqrt{2}} + 2 \cdot \frac{1}{\sqrt{2}} = 2\sqrt{2}$$

The length of the vector \vec{s}_2 is $\Delta s_2 = \sqrt{2}$, the directional cosines are $\cos \alpha = \frac{1}{\sqrt{2}}$ and $\cos \beta = -\frac{1}{\sqrt{2}}$. Thus,

$$\frac{\partial z}{\partial \vec{s}_2} = 2 \cdot \frac{1}{\sqrt{2}} - 2 \cdot \frac{1}{\sqrt{2}} = 0$$

Thus, starting from the same point in the xy -plane and moving in different directions, we get the different results. The directional derivative gives us

the instantaneous rate of change of the given function of two variables at a certain point in the pre-scribed direction.

Partial derivatives with respect to x and y are special cases of the directional derivative. If the given vector \vec{s} points in direction of x -axis then $\alpha = 0$, $\beta = \frac{\pi}{2}$, $\cos \alpha = 1$ and $\cos \beta = 0$. Hence,

$$\frac{\partial z}{\partial \vec{s}} = \frac{\partial z}{\partial x}$$

If the given vector \vec{s} points in direction of y -axis then $\alpha = \frac{\pi}{2}$, $\beta = 0$, $\cos \alpha = 0$ and $\cos \beta = 1$. It follows

$$\frac{\partial z}{\partial \vec{s}} = \frac{\partial z}{\partial y}$$

Thus, the directional derivative in the direction of x -axis is the partial derivative with respect to x and the directional derivative in the direction of y -axis is the partial derivative with respect to y .

The directional derivative of the function of three variables $w = f(x, y, z)$ at the point $P(x, y, z)$ in the direction of the vector $\vec{s} = (\Delta x, \Delta y, \Delta z)$ can be found by the similar formula. Let α , β and γ denote the angles between the vector \vec{s} and x -axis, y -axis and z -axis respectively. Then the directional cosines of the vector \vec{s} are $\cos \alpha$, $\cos \beta$ and $\cos \gamma$. The directional derivative is computed by the formula

$$\frac{\partial w}{\partial \vec{s}} = \frac{\partial w}{\partial x} \cos \alpha + \frac{\partial w}{\partial y} \cos \beta + \frac{\partial w}{\partial z} \cos \gamma \quad (6.33)$$

Example 2. Find the directional derivative of the function $w = xy + xz + yz$ at the point $P(1; 1; 2)$ in the direction of the vector that makes with the coordinate axes the angles 60° , 60° and 45° respectively.

Find the partial derivatives at the point P

$$\frac{\partial w}{\partial x} = y + z \Big|_P = 3, \quad \frac{\partial w}{\partial y} = x + z \Big|_P = 3$$

and

$$\frac{\partial w}{\partial z} = x + y \Big|_P = 2$$

and the directional cosines

$$\vec{s}^\delta = (\cos 60^\circ; \cos 60^\circ; \cos 45^\circ) = \left(\frac{1}{2}; \frac{1}{2}; \frac{\sqrt{2}}{2} \right).$$

By the formula (6.33) we obtain

$$\frac{\partial w}{\partial \vec{s}} = 3 \cdot \frac{1}{2} + 3 \cdot \frac{1}{2} + 2 \cdot \frac{\sqrt{2}}{2} = 3 + \sqrt{2}$$

6.13 Gradient

The function of two variables $z = f(x, y)$ associates to any point $P(x, y)$ in the domain of that function D one value of the dependent variable z or a scalar. To any point in the domain of the function there is related a scalar. Hence, the function of two variables creates a scalar field in the plane.

The function of two variables $w = f(x, y, z)$ associates to any point $P(x, y, z)$ in its domain V a scalar, i.e. creates a scalar field in the domain V . Examples used in physics include the temperature distribution throughout space, the pressure distribution in a fluid or in a gas. Scalar fields are contrasted with other physical quantities such as vector fields, which associate a vector to every point of a region.

Definition 1.

$$\text{grad } z = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right) \quad (6.34)$$

is called the *gradient* of the scalar field $z = f(x, y)$.

Definition 2. The vector

$$\text{grad } w = \left(\frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z} \right) \quad (6.35)$$

is called the gradient of the scalar field $w = f(x, y, z)$.

In the first case there is defined a vector field in the plane and in the second case a vector field in the space. These are called the *gradient field*.

If $\vec{s}^\delta = (\cos \alpha, \cos \beta)$ denotes the unit vector in the direction of the vector \vec{s} , the formula (6.32) can be written as the scalar product of the gradient and the unit vector \vec{s}^δ

$$\frac{\partial z}{\partial \vec{s}} = \text{grad } z \cdot \vec{s}^\delta$$

Since $\vec{s}^\delta = \frac{\vec{s}}{\Delta s}$, then

$$\frac{\partial z}{\partial \vec{s}} = \text{grad } z \cdot \frac{\vec{s}}{\Delta s} = |\text{grad } z| \frac{\text{grad } z \cdot \vec{s}}{|\text{grad } z| \Delta s}$$

where $|\text{grad } z|$ is the length of the gradient vector. Denoting by φ the angle between the gradient and the vector \vec{s} we obtain

$$\cos \varphi = \frac{\text{grad } z \cdot \vec{s}}{|\text{grad } z| \Delta s}$$

and

$$\frac{\partial z}{\partial \vec{s}} = |\text{grad } z| \cos \varphi. \quad (6.36)$$

Now we formulate this result as a theorem.

Theorem 1. The directional derivative of the function $z = f(x, y)$ equals to the projection of the gradient vector onto vector \vec{s} .

Two important conclusions of this theorem.

Conclusion 1. The directional derivative in direction perpendicular to the gradient equals to zero.

This conclusion is obvious because in our case $\varphi = \frac{\pi}{2}$ and $\frac{\partial z}{\partial \vec{s}} = 0$.

Conclusion 2. The directional derivative has the greatest value in the direction of the gradient and equals to the length of the gradient.

It's enough to recall that the cosine function obtains its greatest value 1 if $\varphi = 0$. Thus, the direction of fastest change for a function is given by the gradient vector at that point.

Example 1. Find the greatest rate of growth of the function $z = x^2 + y^2$ at the point $P(1; 1)$.

The directional derivative gives the instantaneous rate of change at the given point. The greatest instantaneous rate of change equals to the length of the gradient. We find the gradient vector at the point P

$$\text{grad } z = (2x, 2y) \Big|_P = (2; 2)$$

and its length $|\text{grad } z| = 2\sqrt{2}$.

This result is the same as the result in Example 1 of the previous subsection, where we have found the directional derivative in direction of the vector \vec{s}_1 . This is natural because the vector $\vec{s}_1 = (1; 1)$ and the gradient have the same directions.

Theorem 2. The gradient is perpendicular to the tangent of level curve.

Proof. The projection of the level curve of the surface $z = f(x, y)$ onto xy -plane is $f(x, y) = c$. This is a function given implicitly and the slope of the tangent line is $\frac{dy}{dx} = -\frac{f'_x}{f'_y}$. Hence, the direction vector of the tangent line is

$$\vec{s} = \left(1; -\frac{f'_x}{f'_y} \right) = \frac{1}{f'_y} (f'_y, -f'_x)$$

The scalar product of the gradient vector and the direction vector of the tangent line

$$\text{grad } z \cdot \vec{s} = f'_x f'_y - f'_y f'_x = 0$$

which means that these two vectors are perpendicular.

Now the Conclusion 1 gives us.

Conclusion 3. The derivative in the direction of the tangent line of the level curve equals to zero.

In Example 1 of the previous subsection the vector \vec{s}_2 has the same direction as the tangent line of the level curve. Thus, by Conclusion 3 it is natural that the derivative in the direction of this vector equals to zero.

Definition 3. A vector field $\vec{F} = (X(x, y), Y(x, y))$ is called a *conservative vector field* if there exists a scalar field $z = f(x, y)$ such that $\vec{F} = \text{grad } z$. If \vec{F} is a conservative vector field then the function $f(x, y)$ is called a *potential function* for \vec{F} .

All this definition is saying is that a vector field is conservative if it is also a gradient vector field for some scalar field.

Example 2. The vector field $\vec{F} = (2xy, x^2)$ is conservative because there exists the scalar field $z = x^2y$ such that $\text{grad } z = \vec{F}$ and x^2y is the potential function for \vec{F} .

6.14 Divergence and curl

The gradient vector field is just one example of vector fields. More generally, a vector field $\vec{F} = (X(x, y, z); Y(x, y, z); Z(x, y, z))$ is an assignment of a vector to each point (x, y, z) in a subset of space. Vector fields are often used to model, for example, the strength and direction of some force, such as the magnetic or gravitational force, as it changes from point to point or the speed and direction of a moving fluid throughout space.

Definition 1. The scalar

$$\text{div } \vec{F} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \quad (6.37)$$

is called the *divergence* of the vector field \vec{F} at the point $P(x, y, z)$.

Definition 2. The vector

$$\text{rot } \vec{F} = \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}; \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}; \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right) \quad (6.38)$$

is called the *curl* (or rotor) of the vector field \vec{F} at the point $P(x, y, z)$.

Example 1. Find the divergence and curl of the vector field $\vec{F} = \left(xyz; x^2 + z^2; \frac{xy}{z} \right)$.

In this example $X = xyz$, $Y = x^2 + z^2$, $Z = \frac{xy}{z}$, thus, $\frac{\partial X}{\partial x} = yz$, $\frac{\partial Y}{\partial y} = 0$ and $\frac{\partial Z}{\partial z} = -\frac{xy}{z^2}$. Hence, the divergence

$$\operatorname{div} \vec{F} = yz - \frac{xy}{z^2}$$

The components of the curl vector

$$\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z} = \frac{x}{z} - 2z$$

$$\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x} = xy - \frac{y}{z}$$

and

$$\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} = 2x - xz$$

Consequently,

$$\operatorname{rot} \vec{F} = \left(\frac{x}{z} - 2z; xy - \frac{y}{z}; 2x - xz \right)$$

If the vector field represents the velocity of a moving flow in space, then the divergence of a vector field \vec{F} at point $P(x, y, z)$ represents a measure of the rate at which the flow diverges (spreads away) from P . That is, $\operatorname{div} \vec{F}|_P$ is the limit of the flow per unit volume out of the infinitesimal sphere centered at P . The curl represents the rotation of a flow, i.e. $\operatorname{rot} \vec{F}|_P$ measures the extent to which the vector field \vec{F} rotates around P .

In field theory there is used a formal vector.

Definition 3. The vector

$$\nabla = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right)$$

is called *Hamilton nabla vector* or *Hamilton nabla operator*.

The coordinates of this vector are not numbers but some operators. The first coordinate means that we find the partial derivative with respect to x for some function etc.

If we treat this vector as an usual vector, we can write for the scalar field $w = f(x, y, z)$

$$\nabla w = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right) w = \left(\frac{\partial w}{\partial x}; \frac{\partial w}{\partial y}; \frac{\partial w}{\partial z} \right) = \operatorname{grad} w$$

Here we have the formal scalar multiplication of ∇ and w . The order of factors is important. The quantities on which ∇ acts must appear to the right of ∇ .

The scalar product of ∇ and the vector field $\vec{F} = (X(x, y, z); Y(x, y, z); Z(x, y, z))$ is

$$\nabla \cdot \vec{F} = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right) \cdot (X; Y; Z) = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = \text{div } \vec{F}$$

The vector product of ∇ and the vector field $\vec{F} = (X(x, y, z); Y(x, y, z); Z(x, y, z))$ is

$$\nabla \times \vec{F} = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right) \times (X; Y; Z) = \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}; \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}; \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right) = \text{rot } \vec{F}$$

Hence, using the nabla operator, we can write

$$\text{grad } w = \nabla w$$

$$\text{div } \vec{F} = \nabla \cdot \vec{F}$$

$$\text{rot } \vec{F} = \nabla \times \vec{F}$$

Definition 4. The scalar product of nabla vector by itself $\nabla^2 = \nabla \cdot \nabla$ is called Laplacian operator and denoted

$$\Delta = \nabla^2$$

The scalar product of nabla vector by itself is not a real quantity

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

but applying this operator to some function, we obtain at every point of the space a scalar.

Example 2. Find the Laplacian operator for the function $w = e^x \sin(yz)$. First we find the first-order partial derivatives

$$\frac{\partial w}{\partial x} = e^x \sin(yz)$$

$$\frac{\partial w}{\partial y} = z e^x \cos(yz)$$

$$\frac{\partial w}{\partial z} = y e^x \cos(yz)$$

and next

$$\begin{aligned}
\Delta w &= \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \\
&= e^x \sin(yz) - z^2 e^x \sin(yz) - y^2 e^x \sin(yz) \\
&= e^x \sin(yz)(1 - z^2 - y^2) = w(1 - z^2 - y^2)
\end{aligned}$$

Finally we prove some equalities that hold for the scalar field $w = f(x, y, z)$ and vector field $\vec{F} = (X; Y; Z)$.

Corollary 1. $\operatorname{div} \operatorname{grad} w = \Delta w$

Proof. We write

$$\operatorname{div} \operatorname{grad} w = \nabla \cdot \operatorname{grad} w = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right) \cdot \left(\frac{\partial w}{\partial x}; \frac{\partial w}{\partial y}; \frac{\partial w}{\partial z} \right) = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}$$

Corollary 2. $\operatorname{div} w \vec{F} = \operatorname{grad} w \cdot \vec{F} + w \operatorname{div} \vec{F}$

Proof. Using the expression of the divergence by nabla vector, we have

$$\begin{aligned}
\operatorname{div} w \vec{F} &= \nabla \cdot w \vec{F} = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right) \cdot (wX; wY; wZ) \\
&= \frac{\partial}{\partial x}(wX) + \frac{\partial}{\partial y}(wY) + \frac{\partial}{\partial z}(wZ) \\
&= \frac{\partial w}{\partial x}X + w \frac{\partial X}{\partial x} + \frac{\partial w}{\partial y}Y + w \frac{\partial Y}{\partial y} + \frac{\partial w}{\partial z}Z + w \frac{\partial Z}{\partial z} \\
&= \left(\frac{\partial w}{\partial x}; \frac{\partial w}{\partial y}; \frac{\partial w}{\partial z} \right) \cdot (X; Y; Z) + w \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \right) \\
&= \operatorname{grad} w \cdot \vec{F} + w \operatorname{div} \vec{F}
\end{aligned}$$

Corollary 3. $\operatorname{rot} w \vec{F} = \operatorname{grad} w \times \vec{F} + w \operatorname{rot} \vec{F}$

Proof. Writing the curl vector via nabla, we obtain

$$\begin{aligned}
\operatorname{rot} w \vec{F} &= \nabla \times w \vec{F} = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z} \right) \times (wX; wY; wZ) \\
&= \left(\frac{\partial(wZ)}{\partial y} - \frac{\partial(wY)}{\partial z}; \frac{\partial(wX)}{\partial z} - \frac{\partial(wZ)}{\partial x}; \frac{\partial(wY)}{\partial x} - \frac{\partial(wX)}{\partial y} \right) \\
&= \left(\frac{\partial w}{\partial y}Z + w \frac{\partial Z}{\partial y} - \frac{\partial w}{\partial z}Y - w \frac{\partial Y}{\partial z}; \frac{\partial w}{\partial z}X + w \frac{\partial X}{\partial z} - \frac{\partial w}{\partial x}Z - w \frac{\partial Z}{\partial x}; \frac{\partial w}{\partial x}Y + w \frac{\partial Y}{\partial x} - \frac{\partial w}{\partial y}X - w \frac{\partial X}{\partial y} \right) \\
&= \left(\frac{\partial w}{\partial y}Z - \frac{\partial w}{\partial z}Y; \frac{\partial w}{\partial z}X - \frac{\partial w}{\partial x}Z; \frac{\partial w}{\partial x}Y - \frac{\partial w}{\partial y}X \right) \\
&+ w \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}; \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}; \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right) \\
&= \operatorname{grad} w \times \vec{F} + w \operatorname{rot} \vec{F}
\end{aligned}$$

6.15 Taylor's formula of function of two variables

When we dealt with the function of one variable, the goal was to represent the function by a polynomial in powers $x - a$ with sufficient accuracy. If the function $f(x)$ and its derivatives up to $n + 1$ degree are continuous at a and in some neighborhood of this point, then

$$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n(x) \quad (6.39)$$

The Lagrange form of the remainder of Taylor's formula is

$$R_n(x) = \frac{(x - a)^{n+1}}{(n + 1)!} f^{(n+1)}[a + \Theta(x - a)] \quad (6.40)$$

where $0 < \Theta < 1$, i.e. $a + \Theta(x - a)$ is a point between a and x . If we approximate the value of the function by Taylor's polynomial at some point x , the remainder shows an approximation error.

For a function of two variables we restrict ourselves with Taylor's polynomial of the second degree. Suppose the function $f(x, y)$ and its partial derivatives up to third order are continuous at the point $A(a, b)$ and in some neighborhood of this point. Let us fix the value of y and write by (6.39) the Taylor's formula of the second degree for the function of one variable x in powers $x - a$

$$f(x, y) = f(a, y) + \frac{f'_x(a, y)}{1!}(x - a) + \frac{f''_{xx}(a, y)}{2!}(x - a)^2 + R_{2,1}(x, y) \quad (6.41)$$

where the remainder

$$R_{2,1}(x, y) = \frac{(x - a)^3}{3!} f'''_{xxx}(\xi, y)$$

and ξ is some point between a and x .

Next we expand $f(a, y)$, $f'_x(a, y)$ and $f''_{xx}(a, y)$, using the Taylor's formula (6.39) in powers $y - b$. First

$$f(a, y) = f(a, b) + \frac{f'_y(a, b)}{1!}(y - b) + \frac{f''_{yy}(a, b)}{2!}(y - b)^2 + R_{2,2}(a, y)$$

where

$$R_{2,2}(a, y) = \frac{(y - b)^3}{3!} f'''_{yyy}(a, \eta_1)$$

and η_1 is some point between b and y . Second

$$f'_x(a, y) = f'_x(a, b) + \frac{f''_{xy}(a, b)}{1!}(y - b) + R_{2,3}(a, y)$$

where $R_{2,3}(a, y) = \frac{(y-b)^2}{2!} f'''_{xyy}(a, \eta_2)$ and η_2 is some point between b and y .

Third

$$f''_{xx}(a, y) = f''_{xx}(a, b) + R_{2,4}(a, y)$$

where $R_{2,4}(a, y) = \frac{y-b}{1!} f'''_{xxy}(a, \eta_3)$ and η_3 is some point between b and y .

Substituting those three to (6.41) gives us

$$\begin{aligned} f(x, y) &= f(a, b) + \frac{f'_y(a, b)}{1!}(y-b) + \frac{f''_{yy}(a, b)}{2!}(y-b)^2 + R_{2,2}(a, y) \\ &+ \frac{x-a}{1!} \left[f'_x(a, b) + \frac{f''_{xy}(a, b)}{1!}(y-b) + R_{2,3}(a, y) \right] \\ &+ \frac{(x-a)^2}{2!} [f''_{xx}(a, b) + R_{2,4}(a, y)] + R_{2,1}(x, y) \end{aligned}$$

or after removing the square brackets and re-arranging the terms

$$\begin{aligned} f(x, y) &= f(a, b) + \frac{f'_x(a, b)}{1!}(x-a) + \frac{f'_y(a, b)}{1!}(y-b) \\ &+ \frac{f''_{xx}(a, b)}{2!}(x-a)^2 + \frac{f''_{xy}(a, b)}{1!}(x-a)(y-b) + \frac{f''_{yy}(a, b)}{2!}(y-b)^2 \\ &+ R_{2,1}(x, y) + R_{2,2}(a, y) + (x-a)R_{2,3}(a, y) + \frac{(x-a)^2}{2!}R_{2,4}(a, y) \end{aligned}$$

The second order Taylor polynomial for the function $f(x, y)$ in powers $x-a$ and $y-b$ is

$$\begin{aligned} P_2(x, y) &= f(a, b) + \frac{f'_x(a, b)}{1!}(x-a) + \frac{f'_y(a, b)}{1!}(y-b) + \\ &+ \frac{1}{2!} [f''_{xx}(a, b)(x-a)^2 + 2f''_{xy}(a, b)(x-a)(y-b) + f''_{yy}(a, b)(y-b)^2] \end{aligned} \quad (6.42)$$

Denoting the remainder of the Taylor's formula of second degree by

$$R_2 = R_{2,1}(x, y) + R_{2,2}(a, y) + (x-a)R_{2,3}(a, y) + \frac{(x-a)^2}{2!}R_{2,4}(a, y)$$

we obtain the Taylor's formula of the second degree for the function $f(x, y)$ in the neighborhood of the point $A(a, b)$ (in powers $x-a$ and $y-b$).

$$\begin{aligned} f(x, y) &= f(a, b) + \frac{f'_x(a, b)}{1!}(x-a) + \frac{f'_y(a, b)}{1!}(y-b) \\ &+ \frac{1}{2!} [f''_{xx}(a, b)(x-a)^2 + 2f''_{xy}(a, b)(x-a)(y-b) + f''_{yy}(a, b)(y-b)^2] + R_2 \end{aligned}$$

If we denote in the expression of the remainder

$$R_2 = \frac{(x-a)^3}{3!} f'''_{xxx}(\xi, y) + \frac{(y-b)^3}{3!} f'''_{yyy}(a, \eta_1) \\ + (x-a) \frac{(y-b)^2}{2!} f'''_{xyy}(a, \eta_2) + \frac{(x-a)^2}{2!} (y-b) f'''_{xxy}(a, \eta_3)$$

$x-a = \Delta x$ and $y-b = \Delta y$, we obtain

$$R_2 = \frac{1}{3!} [(\Delta x)^3 f'''_{xxx}(\xi, y) + 3(\Delta x)^2 \Delta y f'''_{xxy}(a, \eta_3) + 3\Delta x (\Delta y)^2 f'''_{xyy}(a, \eta_2) + (\Delta y)^3 f'''_{yyy}(a, \eta_1)]$$

or using the traditional notation $\Delta \rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$

$$R_2 = \frac{(\Delta \rho)^3}{3!} \left[\frac{(\Delta x)^3}{(\Delta \rho)^3} f'''_{xxx}(\xi, y) + \frac{3(\Delta x)^2 \Delta y}{(\Delta \rho)^3} f'''_{xxy}(a, \eta_3) + \frac{3\Delta x (\Delta y)^2}{(\Delta \rho)^3} f'''_{xyy}(a, \eta_2) + \frac{(\Delta y)^3}{(\Delta \rho)^3} f'''_{yyy}(a, \eta_1) \right]$$

According to our assumptions the third order partial derivatives are continuous in the neighborhood of the point $A(a, b)$. Therefore, they are bounded in this neighborhood. In addition $\left| \frac{\Delta x}{\Delta \rho} \right| \leq 1$ and $\left| \frac{\Delta y}{\Delta \rho} \right| \leq 1$. Hence, in the remainder R_2 the multiplier of $(\Delta \rho)^3$ is a bounded quantity. Denoting this bounded quantity by α_0 , we have $R_2 = \alpha_0 (\Delta \rho)^3$ and the Taylor's formula of function of two variables converts to

$$f(x, y) = f(a, b) + \frac{f'_x(a, b)}{1!} (x-a) + \frac{f'_y(a, b)}{1!} (y-b) \\ + \frac{1}{2!} [f''_{xx}(a, b)(x-a)^2 + 2f''_{xy}(a, b)(x-a)(y-b) + f''_{yy}(a, b)(y-b)^2] + \alpha_0 (\Delta \rho)^3 \quad (6.43)$$

Example. Find the second degree Taylor polynomial for the function $f(x, y) = xe^y$ in the neighborhood of the point $(1; 0)$.

To apply the formula (6.42) we have to evaluate the function and all the partial derivatives up to second order at the point $(1; 0)$. The value of the function

$$f(1; 0) = 1 \cdot e^0 = 1 \cdot 1 = 1$$

The values of the first order partial derivatives $f'_x = e^y$ and $f'_y = xe^y$ are

$$f'_x(1; 0) = 1 \quad \text{and} \quad f'_y(1; 0) = 1 \cdot 1 = 1$$

The values of the second order partial derivatives $f''_{xx} = 0$, $f''_{xy} = e^y$ and $f''_{yy} = xe^y$ are

$$f''_{xy}(1; 0) = 1 \quad \text{and} \quad f''_{yy}(1; 0) = 1 \cdot 1 = 1$$

Now, by (6.42) we get

$$\begin{aligned} P_2(x, y) &= 1 + 1 \cdot (x - 1) + 1 \cdot y + \frac{1}{2} [0 \cdot (x - 1)^2 + 2 \cdot 1 \cdot (x - 1)y + 1 \cdot y^2] = \\ &= x + y + xy - y + \frac{1}{2}y^2 = x + xy + \frac{1}{2}y^2 \end{aligned}$$

6.16 Local extrema of function of two variables

The theory of maxima and minima for the functions of two variables is similar to the theory for one variable.

Definition 1. It is said that the function of two variables $f(x, y)$ has a local maximum at the point $P_1(x_1, y_1)$, if there exists a neighborhood of this point $U_\varepsilon(x_1, y_1)$ such that for any $P(x, y) \in U_\varepsilon(x_1, y_1)$

$$f(x, y) < f(x_1, y_1)$$

Definition 2. It is said that the function of two variables $f(x, y)$ has a local minimum at the point $P_2(x_2, y_2)$, if there exists a neighborhood of this point $U_\varepsilon(x_2, y_2)$ such that for any $P(x, y) \in U_\varepsilon(x_2, y_2)$

$$f(x, y) > f(x_2, y_2)$$

Local extremum is either a local maximum or a local minimum.

Example 1. By Definition 2 the function $z = x^2 + y^2$ has the local minimum at the point $P_0(0; 0)$ because $f(0; 0) = 0$ and for any point $P(x, y)$ different of P_0 there holds $f(x, y) = x^2 + y^2 > 0$.

Example 2. The function $z = x^2 - y^2$ has no local extremum at the point $P_0(0; 0)$. We have $f(0; 0) = 0$ and any neighborhood $U_\varepsilon(0; 0)$ contains the points of x -axis and y -axis. At the points on x -axis $y = 0$ and $z = x^2 > 0$, at the points of y -axis $x = 0$ and $z = -y^2 < 0$.

If the function of two variables has local extremum at the point $P_0(x_0, y_0)$ then the intersection curve of surface (the graph of the function of two variables) and the plain $y = y_0$ has local extremum at x_0 . Hence, the function of one variable $z = f(x, y_0)$ has local extremum at x_0 . It follows that at the point P_0 either $\frac{\partial z}{\partial x} = 0$ or does not exist.

As well, the intersection curve of surface and the plain $x = x_0$ has local extremum at y_0 . The function of one variable $z = f(x_0, y)$ has local extremum at y_0 . Then at the point P_0 either $\frac{\partial z}{\partial y} = 0$ or does not exist.

Definition 3. The points, where $\frac{\partial z}{\partial x} = 0$ or does not exist and $\frac{\partial z}{\partial y} = 0$ or does not exist, are called the *critical points* of the function of two variables.

Now we can formulate the theorem.

Theorem 1. (Necessary condition for existence of local extremum). If the function $z = f(x, y)$ has local extremum at the point P_0 , then P_0 is the critical point of this function.

This theorem says that the function of two variables has a local extremum only at the critical point of this function. But the condition given in this theorem is not sufficient for the function to have a local extremum. For instance the point $O(0; 0)$ is the critical point of the function $z = x^2 - y^2$ because the partial derivatives $\frac{\partial z}{\partial x} = 2x$ and $\frac{\partial z}{\partial y} = 2y$ both equal to zero at this point, but as we know by Example 2, this function has no local maximum and local minimum at $O(0; 0)$.

Because of this theorem we know that if we have all the critical points of a function then we also have every possible local extremum for the function. The fact tells us that all local extrema must be at the critical points so we know that if the function does have local extrema then they must be in the set of all the critical points. However, it will be completely possible that at least at one of the critical points the function hasn't a local extremum.

So the question is how to determine whether the function of two variables has a local extremum at the critical point or not and if it has, is at that point a local maximum or a local minimum.

In the following we consider only the critical points where both partial derivatives equal to zero, i.e. the system of equations

$$\begin{cases} \frac{\partial z}{\partial x} = 0 \\ \frac{\partial z}{\partial y} = 0 \end{cases} \quad (6.44)$$

The solutions of this system of equations are called the *stationary points* of the function $z = f(x, y)$. Every stationary point is also a critical point of the function of two variables but not vice versa. There exist the critical points that are not the stationary points. For instance, for the function $z = \sqrt{x^2 + y^2}$ the partial derivatives

$$\frac{\partial z}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}$$

and

$$\frac{\partial z}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$$

are never simultaneously zero, however they both don't exist at $O(0; 0)$. Therefore, $O(0; 0)$ is a critical point and a possible extremum. The graph of

$z = \sqrt{x^2 + y^2}$ is a cone opening upwards with vertex at the origin. Therefore, at $O(0; 0)$ this function has a local minimum at $O(0; 0)$.

We find the sufficient conditions for existence of the local extremum at the stationary points. Let P_0 be a stationary point of the function $z = f(x, y)$. Evaluate the second order partial derivatives at P_0 and denote

$$A = \frac{\partial^2 z}{\partial x^2} \Big|_{P_0} \quad B = \frac{\partial^2 z}{\partial x \partial y} \Big|_{P_0} \quad \text{and} \quad C = \frac{\partial^2 z}{\partial y^2} \Big|_{P_0}$$

Theorem 2 (sufficient conditions for existence of a local extremum). Let P_0 be a stationary point of the function $z = f(x, y)$.

1. If $AC - B^2 > 0$ and $A < 0$ then the function $z = f(x, y)$ has a local maximum at P_0 .
2. If $AC - B^2 > 0$ and $A > 0$ then the function $z = f(x, y)$ has a local minimum at P_0 .
3. If $AC - B^2 < 0$ then the function $z = f(x, y)$ has no local extremum at P_0 .

Definition 4. If $AC - B^2 < 0$ then the stationary point P_0 is called the *saddle point* of the function $z = f(x, y)$.

Proof. We prove only the first assertion of this theorem. We use the second degree Taylor formula (6.43) taking $a = x_0$, $b = y_0$, $x - x_0 = \Delta x$, $y - y_0 = \Delta y$, i.e. $x = x_0 + \Delta x$ and $y = y_0 + \Delta y$. By our assumption P_0 is a stationary point of the function $z = f(x, y)$ thus,

$$f'_x(x_0, y_0) = 0$$

and

$$f'_y(x_0, y_0) = 0$$

Now (6.43) converts to

$$f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) + \frac{1}{2!} [A(\Delta x)^2 + 2B\Delta x\Delta y + C(\Delta y)^2] + \alpha_0(\Delta \rho)^3$$

where $\Delta \rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$.

Converting the last equality gives

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{(\Delta \rho)^2}{2!} \left[A \frac{(\Delta x)^2}{(\Delta \rho)^2} + 2B \frac{\Delta x \Delta y}{(\Delta \rho)^2} + C \frac{(\Delta y)^2}{(\Delta \rho)^2} + \alpha_0 \Delta \rho \right] \quad (6.45)$$

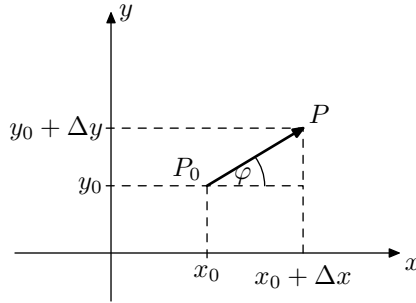


Figure 6.13.

Here $P(x_0 + \Delta x, y_0 + \Delta y)$ is an arbitrary point in the neighborhood of $P_0(x_0, y_0)$. Denote the angle between the vector $\overrightarrow{P_0P}$ and x -axis by φ -ga (Figure 6.12). Then $\cos \varphi = \frac{\Delta x}{\Delta \rho}$ and $\sin \varphi = \frac{\Delta y}{\Delta \rho}$ or $\Delta x = \Delta \rho \cos \varphi$ and $\Delta y = \Delta \rho \sin \varphi$.

Substituting these into (6.45), we obtain

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{(\Delta \rho)^2}{2!} [A \cos^2 \varphi + 2B \cos \varphi \sin \varphi + C \sin^2 \varphi + \alpha_0 \Delta \rho]$$

Multiplying and dividing three first terms in square brackets by A gives

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{(\Delta \rho)^2}{2!} \left[\frac{A^2 \cos^2 \varphi + 2AB \cos \varphi \sin \varphi + AC \sin^2 \varphi}{A} + \alpha_0 \Delta \rho \right]$$

If we add to the numerator of the fraction obtained $0 = B^2 \sin^2 \varphi - B^2 \sin^2 \varphi$ then

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{(\Delta \rho)^2}{2!} \left[\frac{(A \cos \varphi + B \sin \varphi)^2 + (AC - B^2) \sin^2 \varphi}{A} + \alpha_0 \Delta \rho \right]$$

Since $A \cos \varphi + B \sin \varphi$ and $\sin \varphi$ cannot simultaneously equal to zero, by assumption $AC - B^2 > 0$ the numerator of this fraction is positive and due to the assumption $A < 0$ the fraction itself is negative. For sufficiently small Δx and Δy , i.e. for sufficiently small $\Delta \rho$ the expression in the square brackets is negative. Hence,

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) < 0$$

or

$$f(x_0 + \Delta x, y_0 + \Delta y) < f(x_0, y_0)$$

which proves that the function $f(x, y)$ has at $P_0 = (x_0; y_0)$ a local maximum.

We obtain the stationary point $P_0(0; 0)$ of the function $z = x^2 + y^2$ as the solution of the system of equations (6.44)

$$\begin{cases} 2x = 0 \\ 2y = 0 \end{cases}$$

We find

$$A = \frac{\partial^2 z}{\partial x^2} = 2$$
$$B = \frac{\partial^2 z}{\partial x \partial y} = 0$$

and

$$C = \frac{\partial^2 z}{\partial y^2} = 2$$

Hence, $AC - B^2 = 4 > 0$ and $A > 0$. Consequently, by Theorem 2 the function $z = x^2 + y^2$ has at stationary point $P_0(0; 0)$ a local minimum.

We obtain the stationary point $P_0(0; 0)$ of the function $z = x^2 - y^2$ as the solution of the system of equations (6.44)

$$\begin{cases} 2x = 0 \\ -2y = 0 \end{cases}$$

We find

$$A = \frac{\partial^2 z}{\partial x^2} = 2$$
$$B = \frac{\partial^2 z}{\partial x \partial y} = 0$$

and

$$C = \frac{\partial^2 z}{\partial y^2} = -2$$

Thus, $AC - B^2 = -4 < 0$. Consequently, by Theorem 2 the function $z = x^2 - y^2$ has't a local extremum at the stationary point $P_0(0; 0)$. In other words: the point $P_0(0; 0)$ is the saddle point of the function $z = x^2 - y^2$.

Example 3. Find the local extrema of the function $f(x, y) = 4 + x^3 + y^3 - 3xy$.

The first order partial derivatives are

$$\frac{\partial f}{\partial x} = 3x^2 - 3y \quad \text{and} \quad \frac{\partial f}{\partial y} = 3y^2 - 3x$$

To find the stationary points we solve the system of equations (6.44)

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}$$

or

$$\begin{cases} x^2 - y = 0 \\ y^2 - x = 0 \end{cases}$$

The first equation gives $y = x^2$. Substituting this into second equation gives $x^4 - x = 0$ or $x(x^3 - 1) = 0$, whose solutions are $x_1 = 0$ and $x_2 = 1$. Since $y = x^2$, we have two stationary points $P_1(0; 0)$ and $P_2(1; 1)$. Next we find the second order partial derivatives

$$\frac{\partial^2 f}{\partial x^2} = 6x \quad \frac{\partial^2 f}{\partial x \partial y} = -3 \quad \text{and} \quad \frac{\partial^2 f}{\partial y^2} = 6y$$

Since at the first stationary point $P_1(0; 0)$ the values $A = 0$, $B = -3$ and $C = 0$ and

$$AC - B^2 = 0 \cdot 0 - (-3)^2 = -9$$

the point $P_1(0; 0)$ is the saddle point of the given function.

At the second stationary point $P_2(1; 1)$ the values $A = 6$, $B = -3$ and $C = 6$ and

$$AC - B^2 = 6 \cdot 6 - (-3)^2 = 27 > 0$$

As well $A = 6 > 0$ and by Theorem 2 the given function has a local minimum at the point $P_2(1; 1)$ and this local minimum equals to

$$z_{min} = 4 + 1^3 + 1^3 - 3 \cdot 1 \cdot 1 = 3$$

Remark If in Theorem 2 $AC - B^2 = 0$ then anything is possible. More advanced methods are required to classify the stationary point properly.

6.17 Global extrema of function of two variables

The greatest and the least value of the function of two variables are called the *global extrema*.

Suppose that the function $z = f(x, y)$ is continuous in the closed region D . There hold two properties similar to the function of one variable.

Property 1. The continuous at every point of a closed region D function $f(x, y)$ has a maximum and a minimum in D .

Property 2. The continuous at every point of a closed region D function $f(x, y)$ acquires the greatest and the least value at a critical point in D or at a boundary point of D .

These two properties provide us the instructions how to find the greatest and the least value of a continuous function in a closed region D .

1. Find the critical points of the function $f(x, y)$ P_1, P_2, \dots in the region D and find the values of the function at these points $f(P_1), f(P_2), \dots$
2. Find the greatest and the least values of the function on the boundary of D or on the separate parts of the boundary. This can be done by solving for z as a function of x or y alone and using the method for one variable.
3. The largest of the values of steps 1 and 2 is the maximum value and the least is the minimum value.

Example. Find the maximum and the minimum value of the function $z = x^2 + 2xy - 4x - 2y$ in the triangle bounded by the lines $x = 0$, $y = 0$ and $x + y = 4$.

First we sketch the triangle (Figure 6.13).

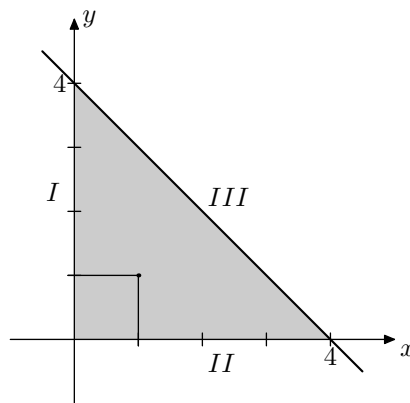


Figure 6.14.

The partial derivatives are $\frac{\partial z}{\partial x} = 2x + 2y - 4$ and $\frac{\partial z}{\partial y} = 2x - 2$.

The system of equations

$$\begin{cases} 2x + 2y - 4 = 0 \\ 2x - 2 = 0 \end{cases}$$

gives one critical point $P_0(1; 1)$ and this point belongs to the triangle given. The value of the function at this point is $f(1; 1) = -3$.

The boundary consists of three line segments. The first line segment is determined by $x = 0$ and $0 \leq y \leq 4$. On this part of boundary we have to find the greatest and the least value of the function $z = -2y$ on the closed

interval $[0; 4]$. Since $z' = -2$ there is no critical points. The values of the function at the endpoints are $z(0) = 0$ and $z(4) = -8$.

The second part of the boundary is determined by $y = 0$ and $0 \leq x \leq 4$. We have to find the greatest and the least value of the function $z = x^2 - 4x$ on the closed interval $[0; 4]$. The equation $z' = 0$ gives $2x - 4 = 0$, i.e. the critical point is $x = 2$. This belongs to the interval $[0; 4]$ and the value of the function at this point is $z(2) = -4$. The values of the function at the endpoints are $z(0) = 0$ and $z(4) = 0$.

The third part of the boundary is determined by $y = 4 - x$ and $0 \leq x \leq 4$. On this part of the boundary we have to find the greatest and the least value of the function $z = -x^2 + 6x - 8$ on the closed interval $[0; 4]$. The equation $z' = 0$, i.e. $-2x + 6 = 0$ gives us the critical point $x = 3$. The value of the function at this point is $z(3) = 1$. The endpoints of the interval are two vertexes of the triangle, where the values of the function have been found already ($z(0) = -8$ and $z(4) = 0$).

Thus, the least value of the function is -8 and the function acquires this value at the point $(0; 4)$. The greatest value is 1 and the function acquires this value at the point $(3; 1)$. The conclusion is

$$\begin{aligned} z_{min} &= z(0; 4) = -8 \\ z_{max} &= z(3; 1) = 1 \end{aligned}$$

6.18 Conditional extrema of function of several variables

One of the common problems in mathematical analysis is that of finding maxima or minima of a function subject to fixed outside conditions or constraints. The classical problem on conditional extrema is the problem of determining the maximum volume.

Example 1. From the sheet-metal with area $2a$ has to be made a rectangular closed box. We have to find the dimensions of maximum volume.

Denote the dimension of the rectangular box x , y and z . We have to find the dimensions of maximum volume $V = xyz$ subject to the constraint $2xy + 2xz + 2yz = 2a$

We return to this example after setting up the theory required.

Let us start with the simpler problem. Suppose we want to maximize or minimize the function $z = f(x, y)$ subject to equation $\varphi(x, y) = 0$. The equation $\varphi(x, y) = 0$ is called the *constraint* and this is a function of one variable given implicitly. The derivative of this function is $\frac{dy}{dx} = -\frac{\varphi'_x}{\varphi'_y}$.

Since the constraint defines y as a function of x then z is the composite function with respect to x . By the formula of the total derivative (6.24)

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

Substituting $\frac{dy}{dx}$ gives

$$\frac{dz}{dx} = f'_x + f'_y \left(-\frac{\varphi'_x}{\varphi'_y} \right)$$

Since at the extremum point $\frac{dz}{dx} = 0$ then

$$f'_x = f'_y \frac{\varphi'_x}{\varphi'_y}$$

Assuming that $f'_y \neq 0$, we obtain the system of equations to determine the point of the conditional extremum

$$\begin{cases} \frac{f'_x}{f'_y} = \frac{\varphi'_x}{\varphi'_y} \\ \varphi(x, y) = 0 \end{cases} \quad (6.46)$$

Example 2. Find the points of conditional extrema for the function $z = x^2 + y^2$ subject to the constraint $x + y = 1$.

In this example $f(x, y) = x^2 + y^2$ and $\varphi(x, y) = x + y - 1$. The partial derivatives $f'_x = 2x$, $f'_y = 2y$, $\varphi'_x = 1$ and $\varphi'_y = 1$. To determine the points of conditional extremum we compose the system of equations (6.46)

$$\begin{cases} \frac{2x}{2y} = \frac{1}{1} \\ x + y - 1 = 0 \end{cases} \quad (6.47)$$

The first equation gives $y = x$ and substituting y into the second equation, we get $2x = 1$, i.e. $x = \frac{1}{2}$. Thus, the conditional point of extremum is

$\left(\frac{1}{2}; \frac{1}{2}\right)$ and the conditional extremum $z = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$. To make sure, is this point a point of conditional maximum or minimum, we choose a nearby point $(0, 6; 0, 4)$ satisfying the constraint and compute the value of the function at this point $z = 0, 6^2 + 0, 4^2 = 0, 52$. Since this value is greater than the value of function at the conditional point of extremum we conclude

that the the function $z = x^2 + y^2$ has at the point $\left(\frac{1}{2}; \frac{1}{2}\right)$ the conditional minimum subject to constraint $x + y = 1$.

This problem has quite simple geometrical meaning. The graph of the function $z = x^2 + y^2$ is the paraboloid of revolution. The constraint is the plain parallel to z -axis and and we have to find the minimum and maximum points on the intersection curve.

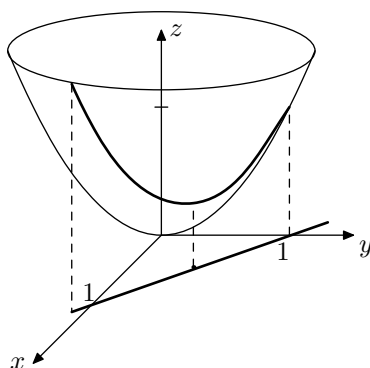


Figure 6.15.

Using the system of equations (6.46), we can find the points of extremum for the function of two variables subject to one constraint. If we have to find conditional extrema for the function of three or more variables, then more general method can be used, the Method of *Lagrange multiplier(s)*. To determine the points at which the function $f(x, y)$ attains constrained local extrema we form an auxiliary function, Lagrange function,

Suppose again we want to find the points of extremum for the function $z = f(x, y)$ subject to constraint $\varphi(x, y) = 0$. We introduce a new variable λ called a *Lagrange multiplier* and study the *Lagrange function* defined by

$$F(x, y, \lambda) = f(x, y) + \lambda\varphi(x, y)$$

The conditional extremum points are the stationary points of this function of three variables, i.e. the solutions of the system of equations

$$\begin{cases} F'_x = 0 \\ F'_y = 0 \\ F'_\lambda = 0 \end{cases} \quad (6.48)$$

Show that this system of equations is equivalent to the system of equations (6.46). Since $F'_\lambda = \varphi(x, y)$, the third equation of the system is the constraint

$\varphi(x, y) = 0$. The first equation is $f'_x + \lambda\varphi'_x = 0$ and the second $f'_y + \lambda\varphi'_y = 0$. To eliminate the variable λ we solve both equations for λ $\lambda = -\frac{f'_x}{\varphi'_x}$ and $\lambda = -\frac{f'_y}{\varphi'_y}$. Those two equations yield

$$\frac{f'_x}{\varphi'_x} = \frac{f'_y}{\varphi'_y}$$

which is equivalent to the first equation of (6.46).

Now it is easy to solve the more general conditional extremum problems. To find the conditional extrema of the function $w = f(x, y, z)$ subject to constraint $\varphi(x, y, z) = 0$, we introduce the Lagrange multiplier λ and define the Lagrange function

$$F(x, y, z, \lambda) = f(x, y, z) + \lambda\varphi(x, y, z)$$

The conditional points of extremum are the solutions of the system of equations

$$\begin{cases} F'_x = 0 \\ F'_y = 0 \\ F'_z = 0 \\ F'_\lambda = 0 \end{cases} \quad (6.49)$$

Now we are ready to solve the problem of maximal volume set up in Example 1. Find the maximum of the volume $V = xyz$ subject to constraint $xy + xz + yz = a$. Define the Lagrange function

$$F(x, y, z, \lambda) = xyz + \lambda(xy + xz + yz - a)$$

and compose the system of equations (6.49)

$$\begin{cases} yz + \lambda(y + z) = 0 \\ xz + \lambda(x + z) = 0 \\ xy + \lambda(x + y) = 0 \\ xy + xz + yz - a = 0 \end{cases} \quad (6.50)$$

The first equation gives $\lambda = -\frac{yz}{y+z}$, the second $\lambda = -\frac{xz}{x+z}$ and the third $\lambda = -\frac{xy}{x+y}$. The first and the second equations obtained give

$$\frac{yz}{y+z} = \frac{xz}{x+z}$$

and the first and the third give

$$\frac{yz}{y+z} = \frac{xy}{x+y}$$

We divide the first equation by z -ga and the second by y -ga. None of these variables cannot equal to zero because otherwise the volume should be zero.

The first equation gives $y(x+z) = x(y+z)$, that is $xy + yz = xy + xz$ or $yz = xz$, i.e. $y = x$. The second equation gives $z(x+y) = x(y+z)$ or $xz + yz = xy + xz$ or $yz = xy$, i.e. $z = x$. Hence, $z = y = x$, that means the dimensions of this box have to be equal and the rectangular box is actually the cube. Substituting z and y into the last equation of (6.50), we obtain

$$x^2 + x^2 + x^2 = a, \text{ which gives } x = \sqrt{\frac{a}{3}}.$$

Consequently, if we have a sheet metal with area $2a$ the dimensions of the rectangular box of maximal volume have to be $x = y = z = \sqrt{\frac{a}{3}}$ and the

$$\text{maximal volume is } V_{max} = \frac{a}{3} \sqrt{\frac{a}{3}}.$$

To find the conditional extrema for the function $w = f(x, y, z)$ subject to constraints $\varphi(x, y, z) = 0$ and $\psi(x, y, z) = 0$ we introduce two Lagrange multipliers λ and μ and define the Lagrange function

$$F(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda\varphi(x, y, z) + \mu\psi(x, y, z)$$

The conditional extremum points are the solutions of the system of equations

$$\begin{cases} F'_x = 0 \\ F'_y = 0 \\ F'_z = 0 \\ F'_\lambda = 0 \\ F'_\mu = 0 \end{cases} \quad (6.51)$$