Partial Derivatives

1 Functions of two or more variables

In many situations a quantity (variable) of interest depends on two or more other quantities (variables), e.g.

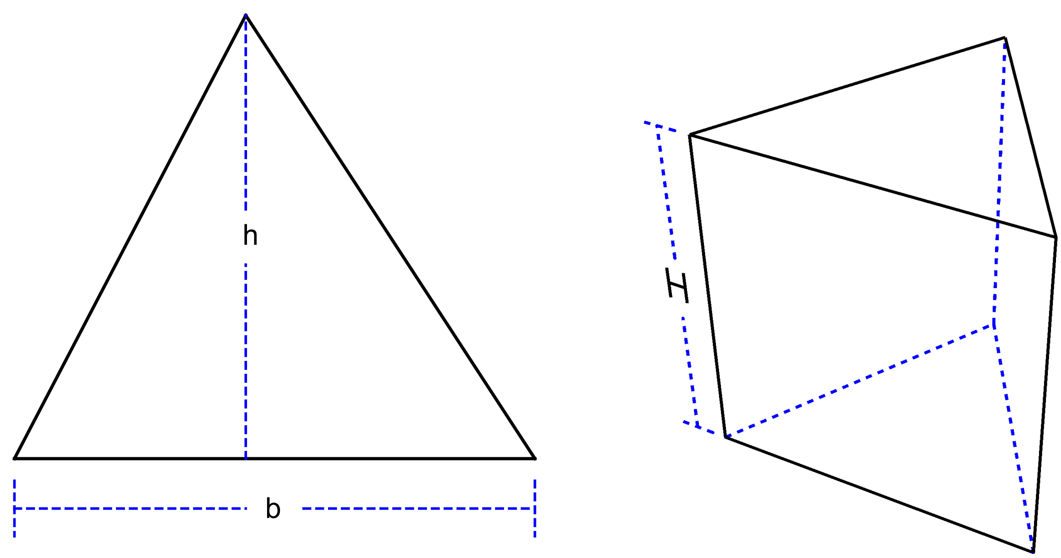


Figure 1: b is the base length of the triangle, h is the height of the triangle, H is the height of the cylinder.

The area of the triangle and the base of the cylinder: $A = \frac{1}{2}bh$

The volume of the cylinder: $V = AH = \frac{1}{2}bhH$

The arithmetic average $\bar{x}$ of n real numbers $x_1, \ldots, x_n$

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

We say

A is a function of the two variables b and h.
V is a function of the three variables b, h and H.
$\bar{x}$ is a function of the n variables $x_1, ..., x_n$.

The expression $z = f(x, y)$ means that z is a function of x and y;

$w = f(x, y, z)$; $u = f(x_1, x_2, \ldots, x_n)$.

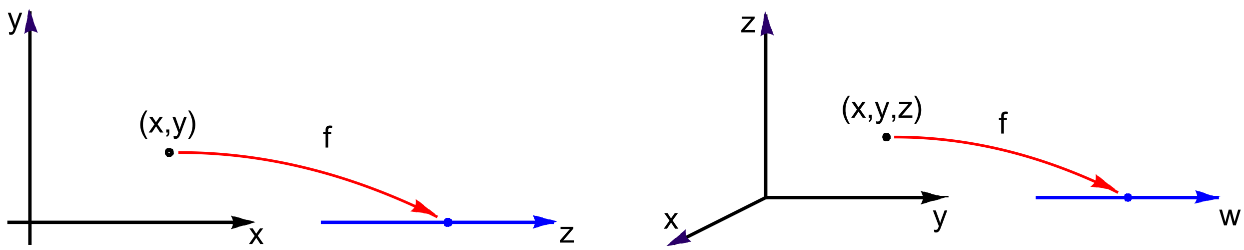


Figure 2: A function f assigns a unique number $z = f(x, y)$, or $w = f(x, y, z)$ to a point in (x, y)-plane or (x, y, z)-space.

The independent variables of a function may be restricted to lie in some set $\mathcal{D}$ which we call the **domain** of f, and denote $\mathcal{D}(f)$. The **natural domain** consists of all points for which a function defined by a formula gives a real number.

Definition. A function f of two variables, x and y, is a rule that assigns a unique real number $f(x, y)$ to each point (x, y) in some set $\mathcal{D}$ in the xy-plane.

A function f of n variables, $x_1, \ldots, x_n$, is a rule that assigns a unique real number $f(x_1, \ldots, x_n)$ to each point $(x_1, \ldots, x_n)$ in some set $\mathcal{D}$ in the n-dimensional $x_1 \ldots x_n$-space, denoted $\mathbb{R}^n$.

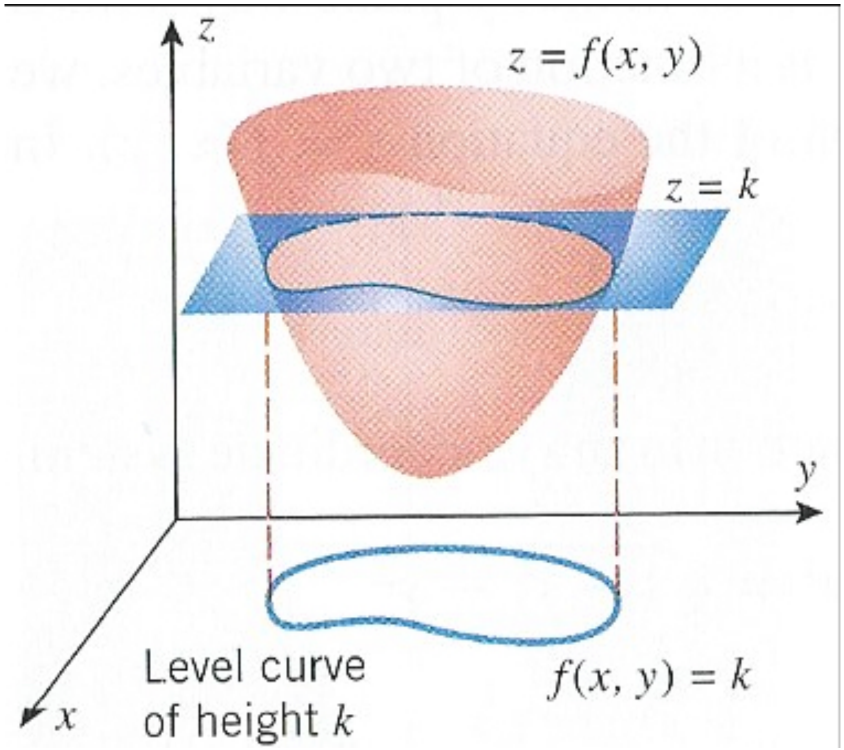
Definition. The **graph** of a function $z = f(x, y)$ in xyz-space is a set of points $P = \big(x, y, f(x, y)\big)$ where (x, y) belong to $\mathcal{D}(f)$.

In general such a graph is a surface in 3-space.

Examples. Find the natural domain of f, identify the graph of f as a surface in 3-space and sketch it.

1. $f(x, y) = 0$;
2. $f(x, y) = 1$;
3. $f(x, y) = x$;
4. $f(x, y) = ax + by + c$;
5. $f(x, y) = x^2 + y^2$;
6. $f(x, y) = \sqrt{1 - x^2 - y^2}$;
7. $f(x, y) = \sqrt{1 + x^2 + y^2}$;
8. $f(x, y) = \sqrt{x^2 + y^2 - 1}$;
9. $f(x, y) = -\sqrt{x^2 + y^2}$;

2 Level curves



If $z = f(x, y)$ is cut by $z = k$, then at all points on the intersection we have $f(x, y) = k$.

This defines a curve in the xy-plane which is the projection of the intersection onto the xy-plane, and is called the **level curve of height** k or the **level curve with constant** k.

A set of level curves for $z = f(x, y)$ is called a **contour plot** or **contour map** of f.

Examples.

1. $f(x, y) = ax + by + c$;
2. $f(x, y) = x^2 + y^2$;
3. $f(x, y) = \sqrt{1 - x^2 - y^2}$;
4. $f(x, y) = \sqrt{1 + x^2 + y^2}$;
5. $f(x, y) = \sqrt{x^2 + y^2 - 1}$;
6. $f(x, y) = -\sqrt{x^2 + y^2}$;

7. $f(x, y) = y^2 - x^2$. It is the hyperbolic paraboloid (saddle surface).

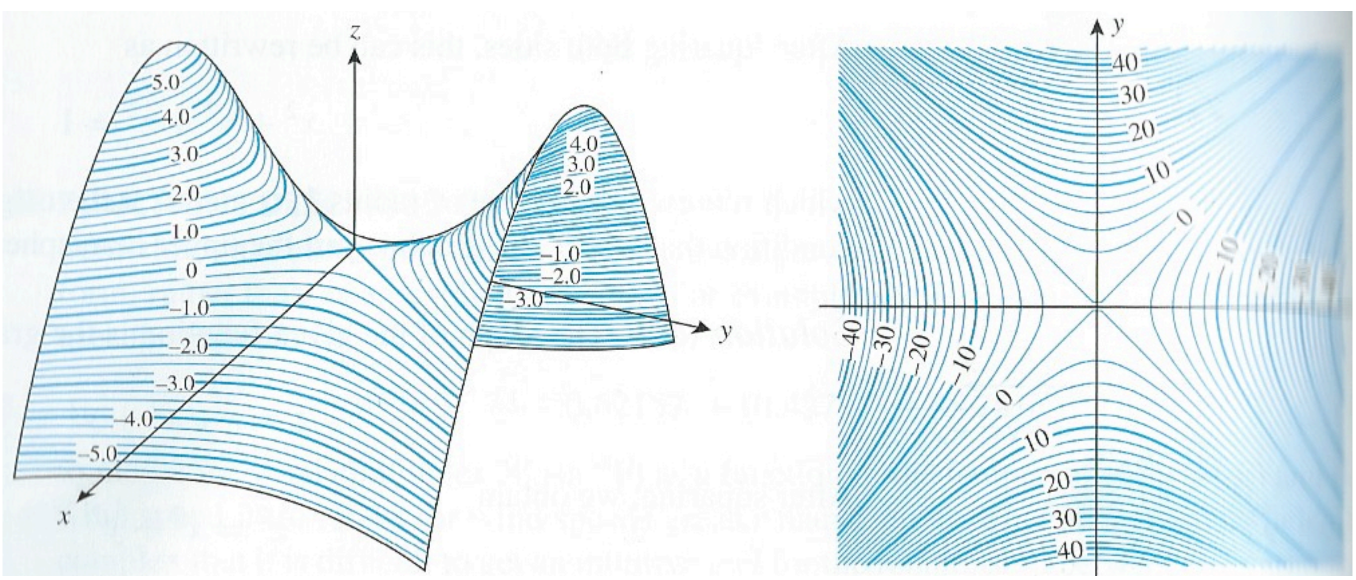


Figure 3: The hyperbolic paraboloid and its contour map.

There is no "direct" way to graph a function of three variables. The graph would be a curved 3-dimensional space (a 3-dim *manifold* if it is smooth), in 4-space. But $f(x, y, z) = k$ defines a surface in 3-space which we call the **level surface with constant** k.

Examples.

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

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

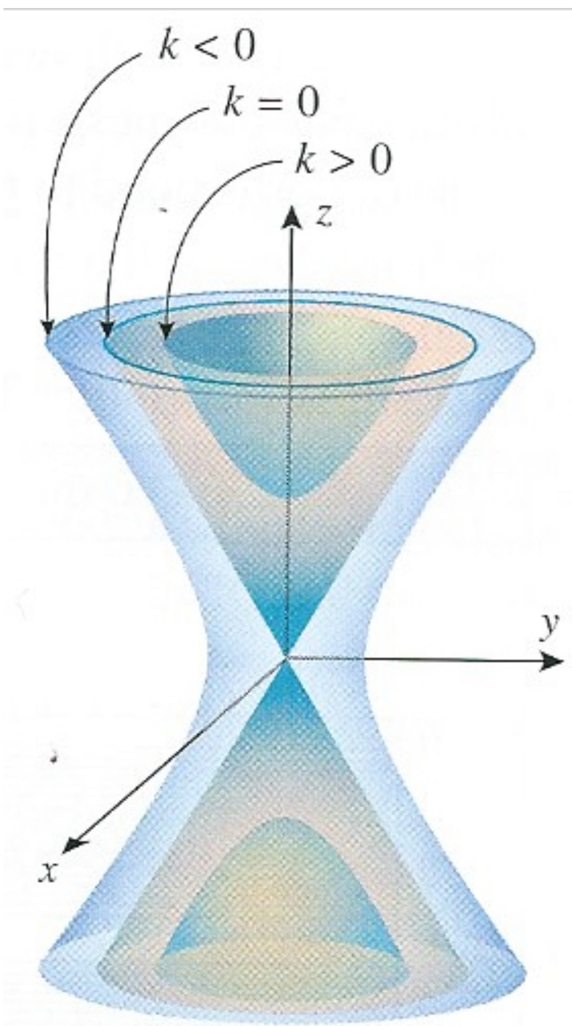
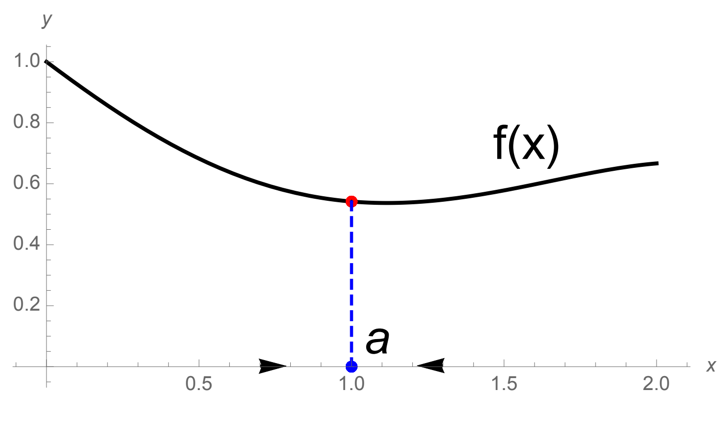
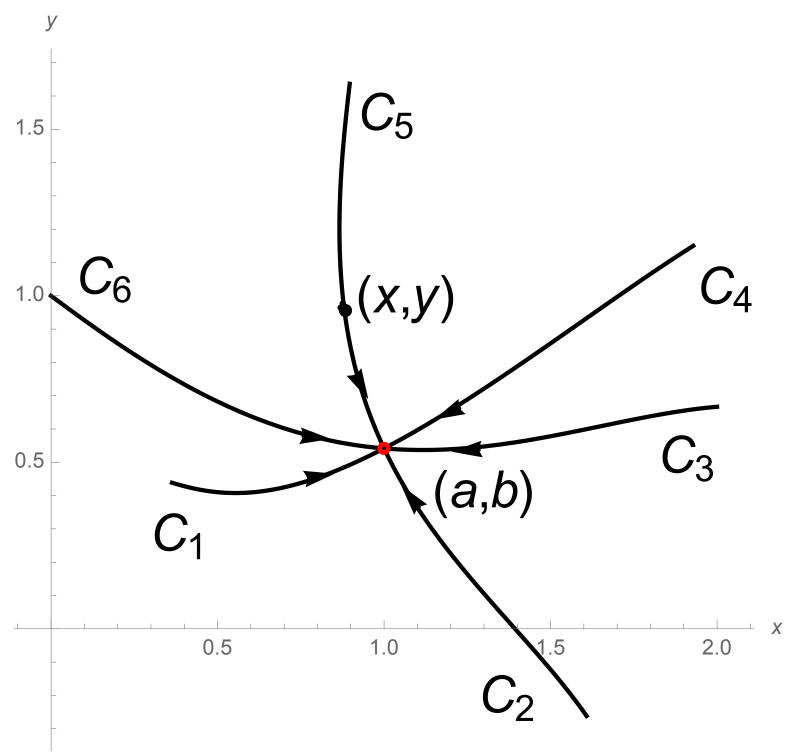


Figure 4: Level surfaces of $f(x, y, z) = z^2 - x^2 - y^2$

3 Limits and Continuity



There are two one-sided limits for $y = f(x)$.



For $z = f(x, y)$ there are infinitely many curves along which one can approach (a, b).
This leads to the notion of the limit of $f(x, y)$ along a curve C.
If all these limits coincide then $f(x, y)$ has a limit at (a, b), and the limit is equal to $f(a, b)$ then f is continuous at (a, b).

4 Partial Derivatives

Recall that for a function $f(x)$ of a single variable the derivative of f at $x = a$

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

is the instantaneous rate of change of f at a, and is equal to the slope of the tangent line to the graph of $f(x)$ at $(a, f(a))$.

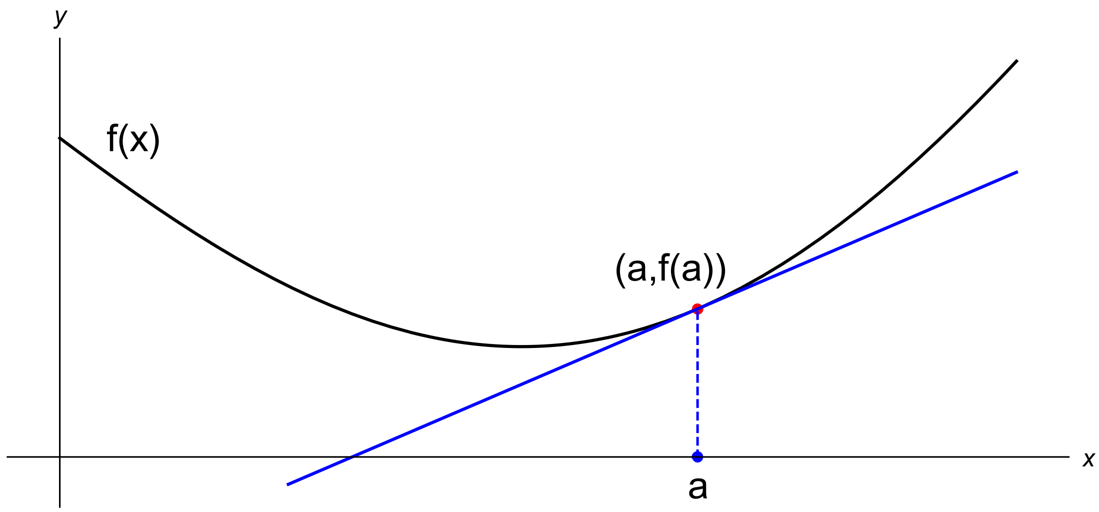


Figure 5: Equation of the tangent line: $y = f(a) + f'(a)(x - a)$.

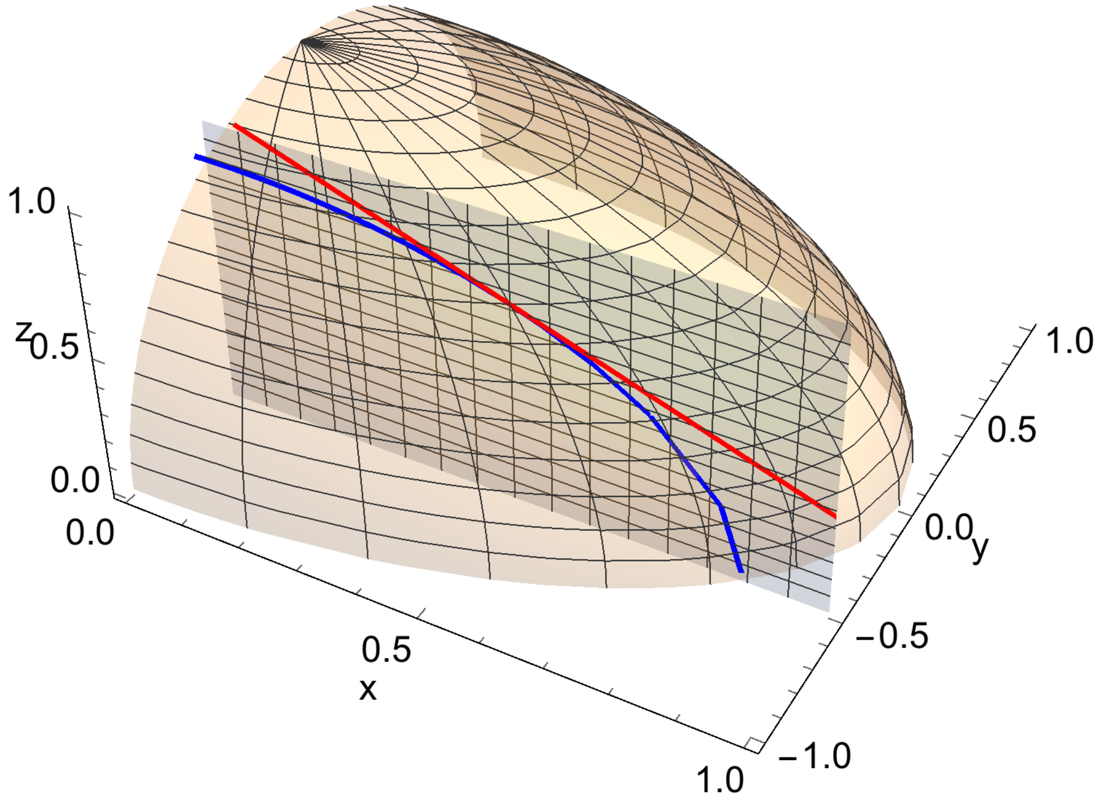
Consider $f(x, y)$. If we fix $y = b$ where b is a number from the domain of f then $f(x, b)$ is a function of a single variable x and we can calculate its derivative at some $x = a$. This derivative is called the **partial derivative** of $f(x, y)$ with respect to x at (a, b) and is denoted by

$$f_x(a, b) \quad \text{or by} \quad \frac{\partial f(a, b)}{\partial x}$$

$$f_x(a, b) = \frac{\partial f(a, b)}{\partial x} = \frac{d}{dx}\big[f(x, b)\big]\Big|_{x=a} = \lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h}$$

$$\text{If} \quad f(x, y) = x \quad \text{then} \quad \frac{\partial x}{\partial x} = 1\,, \quad \text{and if} \quad f(x, y) = y \quad \text{then} \quad \frac{\partial y}{\partial x} = 0$$

Geometrically, given the surface $z = f(x, y)$, we consider its intersection with the plane $y = b$ which is a curve. This curve is the graph of the function $f(x, b)$, and therefore the partial derivative $f_x(a, b)$ is the slope of the tangent line to the curve at $(a, b, f(a, b))$



Equation of the tangent line: $x = t$, $y = b$, $z = f(a, b) + f_x(a, b)(t - a)$

We call $f_x(a, b)$ the **slope of the surface in the x-direction** at (a, b)

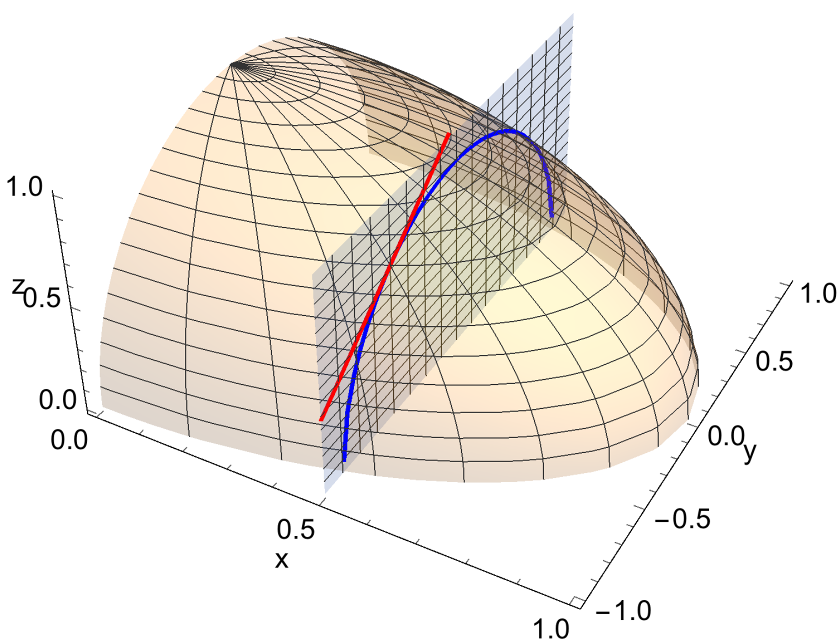
Similarly, if we fix $x = a$ where a is a number from the domain of f then $f(a, y)$ is a function of a single variable y and we can calculate its derivative at some $y = b$. This derivative is called the **partial derivative** of $f(x, y)$ with respect to y at (a, b) and is denoted by

$$f_y(a,b) \text{ or by } \frac{\partial f(a,b)}{\partial y}$$

$$f_y(a,b) = \frac{\partial f(a,b)}{\partial y} = \frac{d}{dy}\big[f(a,y)\big]\Big|_{y=b} = \lim_{h\to 0}\frac{f(a,b+h)-f(a,b)}{h}$$

$$\text{If } f(x,y) = x \text{ then } \frac{\partial x}{\partial y} = 0, \text{ and if } f(x,y) = y \text{ then } \frac{\partial y}{\partial y} = 1$$

The intersection of the surface $z = f(x, y)$ with the plane $x = a$ is a curve which is the graph of the function $f(a, y)$, and therefore the partial derivative $f_y(a, b)$ is the slope of the tangent line to the curve at $(a, b, f(a, b))$



Equation of the tangent line: $x = a$, $y = t$, $z = f(a, b) + f_y(a, b)(t - a)$

We call $f_y(a, b)$ the **slope of the surface in the y-direction** at (a, b)

If we allow (a,b) to vary, the partial derivatives become functions of two variables:

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

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

Partial derivative notation: if $z = f(x,y)$ then

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial z}{\partial x} = \partial_x f = \partial_x z\,, \quad f_y = \frac{\partial f}{\partial y} = \frac{\partial z}{\partial y} = \partial_y f = \partial_y z$$

Example.

$$z = f(x,y) = \ln \frac{\sqrt[3]{2x^2 - 3xy^2 + 3\cos(2x+3y) - 3y^3 + 18}}{2}$$

Find $f_x(x,y)$, $f_y(x,y)$, $f(3,-2)$, $f_x(3,-2)$, $f_y(3,-2)$

For $w = f(x,y,z)$ there are three partial derivatives $f_x(x,y,z)$, $f_y(x,y,z)$, $f_z(x,y,z)$

Example.

$$f(x,y,z) = \sqrt{z^2 + y - x + 2\cos(3x - 2y)}$$

Find

$f_x(x,y,z)$, $f_y(x,y,z)$, $f_z(x,y,z)$,
$f(2,3,-1)$, $f_x(2,3,-1)$, $f_y(2,3,-1)$, $f_z(2,3,-1)$

In general, for $w = f(x_1, x_2, \ldots, x_n)$ there are n partial derivatives:

$$\frac{\partial w}{\partial x_1}, \quad \frac{\partial w}{\partial x_2}, \quad \cdots \quad , \quad \frac{\partial w}{\partial x_n}$$

Example.

$$r = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Find

$$\frac{\partial r}{\partial x_1}, \quad \frac{\partial r}{\partial x_2}, \quad \frac{\partial r}{\partial x_9}, \quad \frac{\partial r}{\partial x_i}, \quad \frac{\partial r}{\partial x_{n-1}}, \quad n \geq 9\,, \; i \leq n$$

Second-order derivatives: f_{xx}, f_{xy}, f_{yx}, f_{yy}

$$\begin{tikzcd}
 & & f_{xx} \\
 & f_x \arrow[ur] \arrow[r] & f_{xy} \\
f \arrow[ur] \arrow[dr] & & \\
 & f_y \arrow[r] \arrow[dr] & f_{yx} \\
 & & f_{yy}
\end{tikzcd}$$

Notation

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

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

f_{xy} and f_{yx} are called the **mixed second-order partial derivatives**. f_x and f_y can be called first-order partial derivative.

Example.
$$z = 2e^{y-\frac{\pi}{2}}\sin x - 3e^{x-\frac{\pi}{4}}\cos y$$
Find
$$\frac{\partial z}{\partial x}, \quad \frac{\partial z}{\partial y}, \quad \frac{\partial^2 z}{\partial x^2}, \quad \frac{\partial^2 z}{\partial x \partial y}, \quad \frac{\partial^2 z}{\partial y^2}, \quad \frac{\partial^2 z}{\partial y \partial x},$$
$$\frac{\partial z}{\partial x}(\frac{\pi}{4}, \frac{\pi}{2}), \quad \frac{\partial z}{\partial y}(\frac{\pi}{4}, \frac{\pi}{2}), \quad \frac{\partial^2 z}{\partial x \partial y}(\frac{\pi}{4}, \frac{\pi}{2}), \quad \frac{\partial^2 z}{\partial y \partial x}(\frac{\pi}{4}, \frac{\pi}{2})$$

Equality of mixed partial derivatives

Theorem. Let f be a function of two variables. If f_{xy} and f_{yx} are continuous on some open disc, then $f_{xy} = f_{yx}$ on that disc.

Higher-order derivatives

Third-order, fourth-order, and higher-order derivatives are obtained by successive differentiation.
$$f_{xxx} = \frac{\partial^3 f}{\partial x^3} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right), \quad f_{xyy} = \frac{\partial^3 f}{\partial y^2 \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y \partial x}\right)$$
$$f_{xyxz} = \frac{\partial^4 f}{\partial z \partial x \partial y \partial x} = \frac{\partial}{\partial z}\left(\frac{\partial^3 f}{\partial x \partial y \partial x}\right)$$

For higher-order derivatives the equality of mixed partial derivatives also holds if the derivatives are continuous.

In what follows we always assume that the order of partial derivatives is irrelevant for functions of any number of independent variables.

5 Differentiability, differentials and local linearity

For $f(x, y)$, the symbol Δf, called the **increment** of f, denotes the change

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

For small Δx, Δy

$$\Delta f \approx f_x(a, b)\Delta x + f_y(a, b)\Delta y$$

Definition. A function $f(x, y)$ is said to be **differentiable** at (a, b) provided $f_x(a, b)$ and $f_y(a, b)$ both exist and

$$\lim_{(\Delta x, \Delta y) \to (0,0)} \frac{\Delta f - f_x(a, b)\Delta x - f_y(a, b)\Delta y}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} = 0$$

For $f(x, y, z)$

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

For small Δx, Δy, Δz

$$\Delta f \approx f_x(a, b, c)\Delta x + f_y(a, b, c)\Delta y + f_z(a, b, c)\Delta z$$

and $f(x, y, z)$ is **differentiable** at (a, b, c) if

$$\lim_{(\Delta x, \Delta y, \Delta z) \to (0,0,0)} \frac{\Delta f - f_x(a, b, c)\Delta x - f_y(a, b, c)\Delta y - f_z(a, b, c)\Delta z}{\sqrt{(\Delta x)^2 + (\Delta y)^2 + +(\Delta z)^2}} = 0$$

Theorem. If a function is differentiable at a point, then it is continuous at that point.

Theorem. If all first-order derivatives of f exist and are continuous at a point, then f is differentiable at a point.

Differentials

If $z = f(x, y)$ is differentiable at (a, b) we let

$$dz = f_x(a, b)dx + f_y(a, b)dy$$

denote a new function with dependent variable dz and independent variables dx, dy. It is called the **total differential of z (or f) at** (a, b). It is a linear function of dx and dy.

Note that $\Delta z \approx dz$ if $\Delta x = dx$ and $\Delta y = dy$

If we allow (a, b) to vary, the differential becomes a function of four variables, dx, dy, x, y:

$$a \to x\ ,\ \ b \to y\ \ \Rightarrow\ \ dz = f_x(x, y)dx + f_y(x, y)dy$$

Definition. If $f(x, y)$ is differentiable at (a, b) then

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

is called the **local linear approximation of f at** (a, b).
Its graph is the tangent plane to the surface $z = f(x, y)$ at $(a, b, f(a, b))$

Example. $f(x, y) = \sqrt{x^2 + y^2}$. Compute $f(3.04, 3.98)$, and estimate the error if a calculator gives $f(3.04, 3.98) \approx 5.00819$

If $w = f(x, y, z)$, the total differential of w (or f) at (a, b, c) is

$$dw = f_x(a, b, c)dx + f_y(a, b, c)dy + f_z(a, b, c)dz$$

or if $a \to x\ ,\ \ b \to y\ ,\ \ c \to z$

$$dw = f_x(x, y, z)dx + f_y(x, y, z)dy + f_z(x, y, z)dz$$

The local linear approximation of f at (a, b, c) is

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

6 The Chain Rule

Recall

$$y = f(x(t)) \;\Rightarrow\; \frac{dy}{dt} = \frac{dy}{dx}\,\frac{dx}{dt}$$

because

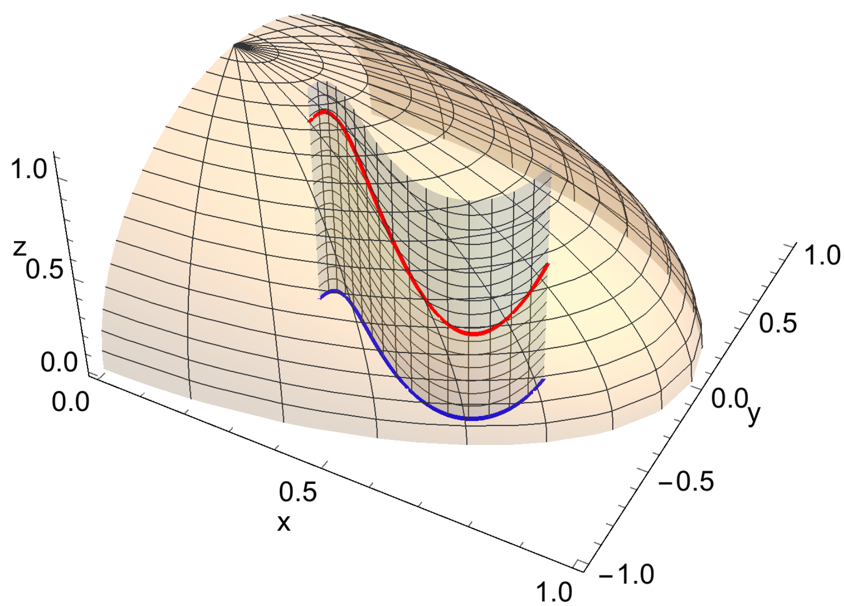
$$\Delta y \approx \frac{dy}{dx}\Delta x\,, \quad \Delta x \approx \frac{dx}{dt}\Delta t$$

Let $z = f(x, y)$ and $x = x(t)$, $y = y(t)$. Then $z = f(x(t), y(t))$ is a function of the single variable t.

$$\Delta z \approx \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y\,, \quad \Delta x \approx \frac{dx}{dt}\Delta t\,,\; \Delta y \approx \frac{dy}{dt}\Delta t$$

and therefore

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



Example. $z = \sqrt{4 - x^2 - y^2}$, $x = 1 + \cos t$, $y = \sin t$

Similarly, if $w = f(x, y, z)$ and $x = x(t)$, $y = y(t)$, $z = z(t)$. Then $w = f(x(t), y(t), z(t))$ is a function of the single variable t, and

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

In general, if $w = f(x_1, x_2, \ldots, x_n)$ and $x_1 = x_1(t)$, $x_2 = x_2(t)$, ... , $x_n = x_n(t)$, then

$$\frac{dw}{dt} = \frac{\partial w}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial w}{\partial x_2}\frac{dx_2}{dt} + \cdots \frac{\partial w}{\partial x_n}\frac{dx_n}{dt} = \sum_{i=1}^{n}\frac{\partial w}{\partial x_i}\frac{dx_i}{dt}$$

Implicit differentiation

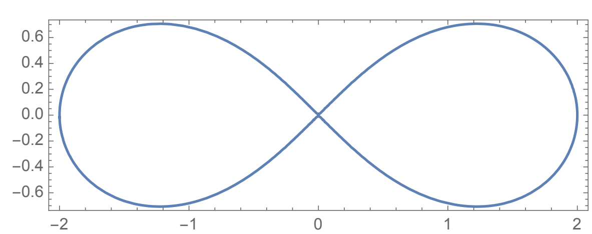
Let $z = f(x, y)$ and $y = y(x)$. Then

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

Suppose $y(x)$ is such that $f(x, y(x)) = \text{const}$. Then, $\frac{dz}{dx} = 0$ and

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0 \;\Rightarrow\; \frac{dy}{dx} = -\frac{f_x}{f_y} \quad \text{if} \quad f_y \neq 0$$

Example. The lemniscate is defined by the equation



$$(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$$

Find dy/dx.

The chain rule for partial derivatives

1. Let $y = f(x)$ and $x = x(u, v)$
Then $y = f(x(u, v))$ is a function of u and v, and

$$\Delta y \approx \frac{dy}{dx}\Delta x\,, \quad \Delta x \approx \frac{\partial x}{\partial u}\Delta u + \frac{\partial x}{\partial v}\Delta v$$

Thus,

$$\frac{\partial y}{\partial u} = \frac{dy}{dx}\frac{\partial x}{\partial u}\,, \qquad \frac{\partial y}{\partial v} = \frac{dy}{dx}\frac{\partial x}{\partial v}$$

2. Let $z = f(x, y)$ and $x = x(u, v)$, $y = y(u, v)$
Then $x = f(x(u, v), y(u, v)$ is a function of u and v, and

$$\Delta z \approx \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y\,, \quad \Delta x \approx \frac{\partial x}{\partial u}\Delta u + \frac{\partial x}{\partial v}\Delta v\,, \quad \Delta y \approx \frac{\partial y}{\partial u}\Delta u + \frac{\partial y}{\partial v}\Delta v$$

Thus,

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}\,, \qquad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

3. Let $w = f(x, y, z)$ and $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$

$$\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u}\,, \qquad \frac{\partial w}{\partial v} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial v}$$

4. Let $w = f(x_1, ..., x_n)$ and $x_1 = x_1(u_1, ..., u_m)$, ... , $x_n = x_n(u_1, ..., u_m)$

$$\frac{\partial w}{\partial u_\alpha} = \sum_{i=1}^{n} \frac{\partial w}{\partial x_i}\frac{\partial x_i}{\partial u_\alpha}\,, \quad \alpha = 1, \ldots, m$$

Example. Find $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ where

$$z = \cos\frac{x}{2}\sin 2y\,; \quad x = 3u - 2v\,, \;\; y = u^2 - 2v^3$$

Example. The wave equation: Consider a string of length L that is stretched taut between $x = 0$ and $x = L$ on an x-axis, and suppose that the string is set into vibratory motion by "plucking" it at time $t = 0$. The displacement of a point on the string depends both on x and t: $u(x,t)$. One-dimensional wave equation for small displacements

$$\frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} = 0$$

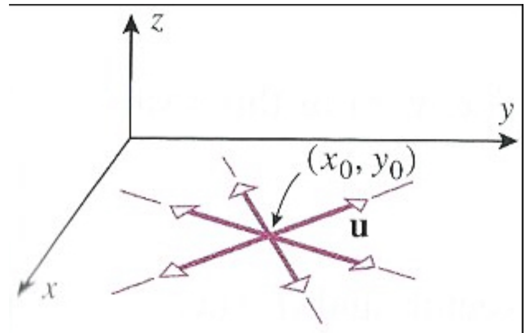
Show that

$$u(x,t) = f(x - ct) + g(x + ct)$$

is a solution to the equation. In fact it is the general solution.

7 Directional Derivatives and Gradients

Suppose we need to compute the rate of change of $f(x,y)$ with respect to the distance from a point (a,b) in some direction. Let $\vec{u} = u_1\vec{i} + u_2\vec{j}$ be the unit vector that has its initial point at (a,b) and points in the desired direction. It determines a line in the xy-plane:



$$x = a + s\,u_1\,, \quad y = b + s\,u_2$$

where s is the arc length parameter that has its reference point at (a,b) and has positive values in the direction of $\vec{u}$.

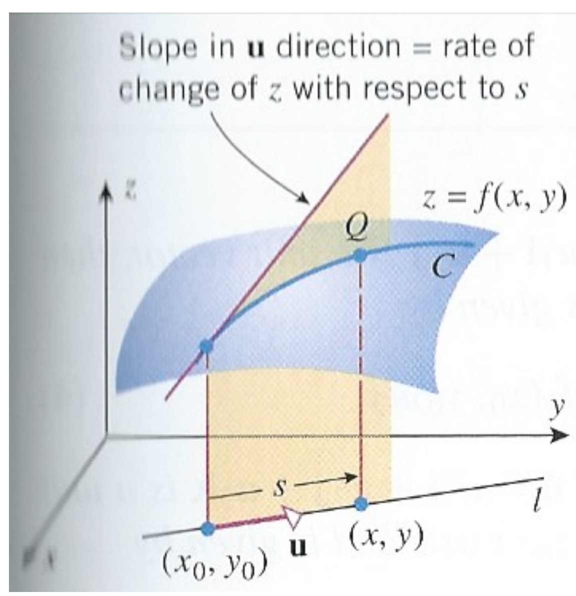
Definition. The **directional derivative** of $f(x,y)$ in the direction of $\vec{u}$ at (a,b) is denoted by $D_{\vec{u}}f(a,b)$ and is defined by

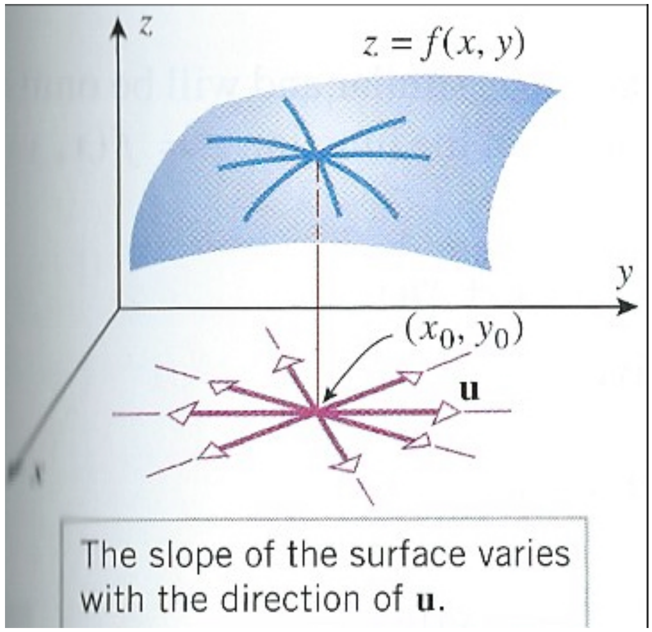
$$D_{\vec{u}}f(a,b) = \frac{d}{ds}\left[f(a + s\,u_1\,,\, b + s\,u_2)\right]\Big|_{s=0} = f_x(a,b)\,u_1 + f_y(a,b)\,u_2$$

provided this derivative exists.

Analytically, $D_{\vec{u}}f(a,b)$ is the instantaneous rate of change of $f(x,y)$ with respect to the distance in the direction of $\vec{u}$ at the point (a,b).

Geometrically, $D_{\vec{u}}f(a,b)$ is the slope of the surface $z = f(x,y)$ in the direction of $\vec{u}$ at the point $(a,b,f(a,b))$.





Generalisation to $f(x, y, z)$ (and $f(x_1, \dots, x_n)$) is straightforward.

Definition. Let $\vec{u} = u_1 \vec{i} + u_2 \vec{j} + u_3 \vec{k}$ be a unit vector.
The **directional derivative** of $f(x, y, z)$ in the direction of $\vec{u}$ at (a, b, c) is denoted by $D_{\vec{u}} f(a, b, c)$ and is defined by

$$\begin{aligned} D_{\vec{u}} f(a, b, c) &= \frac{d}{ds} \left[f(a + s\, u_1 \,,\, b + s\, u_2 \,,\, c + s\, u_3) \right] \Big|_{s=0} \\ &= f_x(a, b, c)\, u_1 + f_y(a, b, c)\, u_2 + f_z(a, b, c)\, u_3 \end{aligned}$$

Example. Find $D_{\vec{u}} f(2, 1)$ in the direction of $\vec{a} = 3\,\vec{i} + 4\,\vec{j}$

$$f(x, y) = \ln \left(\frac{1}{2} e^{2/3} \sqrt[3]{12 \sin(x - 2y) + 8y^2 - x^3 - 6x^2 y + 32} \right)$$

Answer: $D_{\vec{u}} f(2, 1) = -5/3$

The gradient

Note that

$$D_{\vec{u}}f = f_x\,u_1 + f_y\,u_2 + f_z\,u_3 = (f_x\,\vec{i} + f_y\,\vec{j} + f_z\,\vec{k}) \cdot (u_1\,\vec{i} + u_2\,\vec{j} + u_3\,\vec{k})$$

Definition. Let $\vec{e}_i$ be the standard orthonormal coordinate basis of $\mathbb{R}^n$, so that $\vec{r} = \sum_{i=1}^{n} x_i \vec{e}_i$.
The **gradient** of $f(x_1, \cdots, x_n)$ is defined by

$$\vec{\nabla} f(x_1, \cdots, x_n) = \sum_{i=1}^{n} \frac{\partial f(x_1, \cdots, x_n)}{\partial x_i} \vec{e}_i$$

In particular

$$\vec{\nabla} f(x, y) = f_x(x, y)\,\vec{i} + f_y(x, y)\,\vec{j}$$

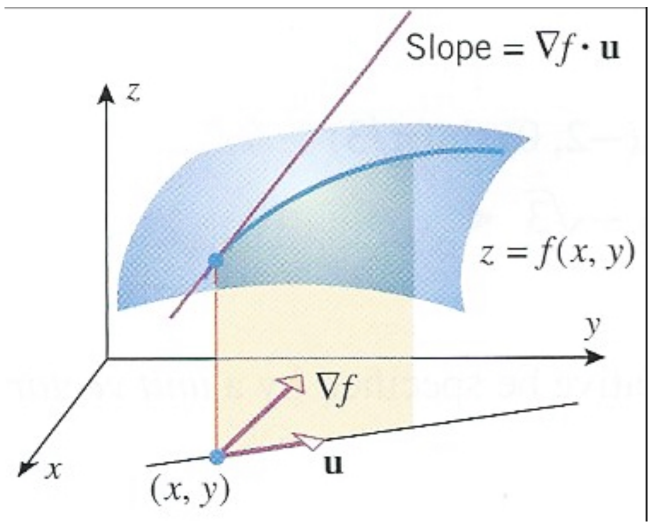
$$\vec{\nabla} f(x, y, z) = f_x(x, y, z)\,\vec{i} + f_y(x, y, z)\,\vec{j} + f_z(x, y, z)\,\vec{k}$$

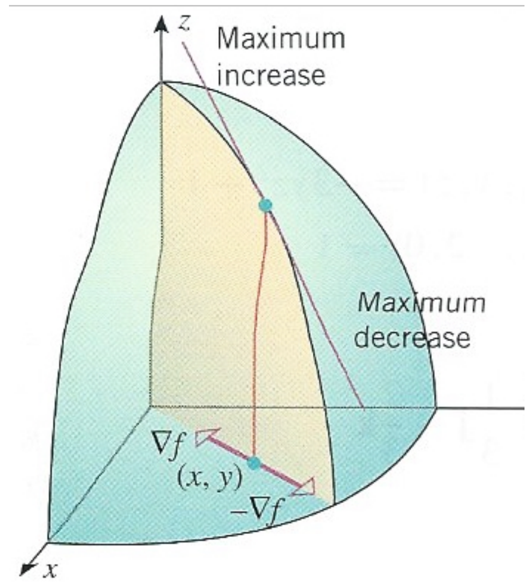
The symbol $\vec{\nabla}$ is read as either “nabla” (from ancient Hebrew) or “del” (it is inverted Δ).

$$D_{\vec{u}}f(a, b) = \vec{\nabla} f(a, b) \cdot \vec{u}\,, \quad D_{\vec{u}}f(a, b, c) = \vec{\nabla} f(a, b, c) \cdot \vec{u}\,, \quad D_{\vec{u}}f = \vec{\nabla} f \cdot \vec{u}$$

Example. Find $\vec{\nabla} r$; $r = \sqrt{x^2 + y^2 + z^2}$ and $D_{\vec{u}}r(1, 1, 1)$ in the direction of $\vec{a} = \vec{i} + 2\,\vec{j} + 2\,\vec{k}$.

Properties of the gradient





$$D_{\vec{u}}f(a,b) = \vec{\nabla}f(a,b) \cdot \vec{u} = |\vec{\nabla}f(a,b)|\,|\vec{u}|\,\cos\theta = |\vec{\nabla}f(a,b)|\,\cos\theta$$

Since $-1 \le \cos\theta \le 1$, if $|\vec{\nabla}f(a,b)| \neq 0$ then the maximum value of $D_{\vec{u}}f(a,b)$ is $|\vec{\nabla}f(a,b)|$ and it occurs when $\theta = 0$, that is, when $\vec{u}$ is in the direction of $\vec{\nabla}f(a,b)$.

Geometrically, the maximum slope of the surface $z = f(x,y)$ at (a,b) is in the direction of the gradient and is equal to $|\vec{\nabla}f(a,b)|$.

If $|\vec{\nabla}f(a,b)| = 0$ then $D_{\vec{u}}f(a,b) = 0$ in all directions at (a,b).
It occurs where the surface $z = f(x,y)$ has a relative maximum or minimum or a saddle point.

Since $D_{\vec{u}} f(x_1, \ldots, x_n) = |\vec{\nabla} f(x_1, \ldots, x_n)| \cos\theta$, these properties hold for functions of any number of variables.

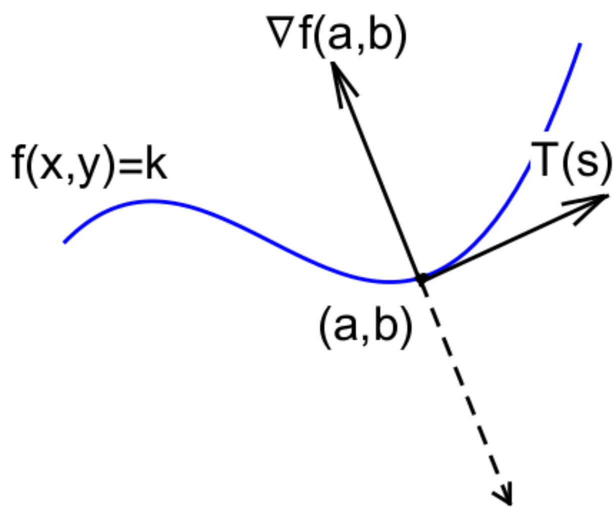
Theorem. Let f be a function differentiable at a point P.

1. If $\vec{\nabla} f = \vec{0}$ at P then all directional derivatives of f at P are 0.
2. If $\vec{\nabla} f \neq \vec{0}$ at P then the derivative in the direction of $\vec{\nabla} f$ at P has the largest value equal to $|\vec{\nabla} f|$ at P.
3. If $\vec{\nabla} f \neq \vec{0}$ at P then the derivative in the direction opposite to that of $\vec{\nabla} f$ at P has the smallest value equal to $-|\vec{\nabla} f|$ at P.

Example. The point $P = (2, 3, -1)$

$$f(x, y, z) = \sqrt{2xy + 3z^4 - 6\cos(3x - 2y)}$$

Gradients are normal to level curves and level surfaces



Level curve C: $f(x,y)=k$.

Let C be smoothly parametrised as $x=x(s)$, $y=y(s)$ where s is an arc length parameter. The unit tangent vector to C is

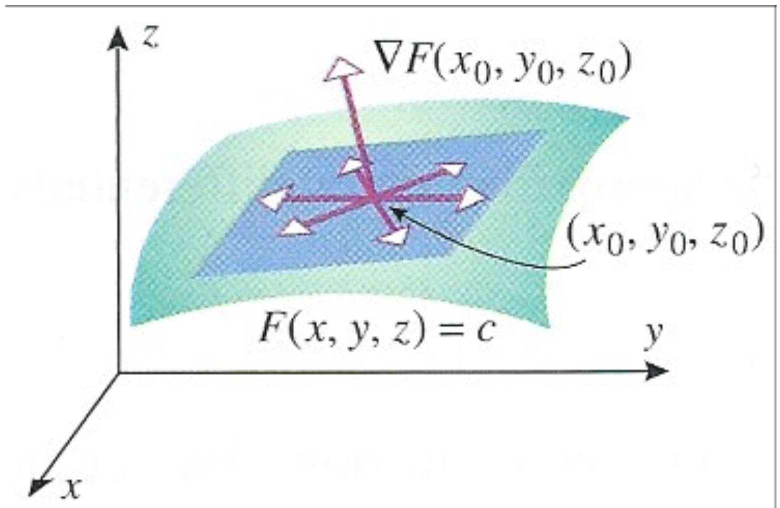
$$\vec{T}(s)=\frac{dx}{ds}\,\vec{i}+\frac{dy}{ds}\,\vec{j}$$

Since $f(x,y)$ is constant on C we expect $D_{\vec{T}}f(x,y)=0$. Indeed

$$\begin{aligned} D_{\vec{T}}f(x,y) &= \vec{\nabla}f\cdot\vec{T}=(f_x\,\vec{i}+f_y\,\vec{j})\cdot(\frac{dx}{ds}\,\vec{i}+\frac{dy}{ds}\,\vec{j}) \\ &= f_x\frac{dx}{ds}+f_y\frac{dy}{ds}=\frac{d}{ds}f(x(s),y(s))=0\ \Rightarrow\ \vec{\nabla}f\perp\vec{T} \end{aligned}$$

Thus if (a,b) belongs to the level curve, and $\vec{\nabla}f(a,b)\neq\vec{0}$ then $\vec{\nabla}f(a,b)$ is normal to $\vec{T}$ at (a,b) and therefore to the level curve.

Definition. A vector is called normal to a surface at (a, b, c) if it is normal to a tangent vector to any curve on the surface through (a, b, c).



Level surface σ: $F(x, y, z) = k$

Let C, smoothly parametrised as $x = x(s)$, $y = y(s)$, $z = z(s)$

be any curve on σ through (a, b, c). The unit tangent vector to C is

$$\vec{T}(s) = \frac{dx}{ds}\vec{i} + \frac{dy}{ds}\vec{j} + \frac{dz}{ds}\vec{k}$$

and $D_{\vec{T}}F(x, y, z)$ is

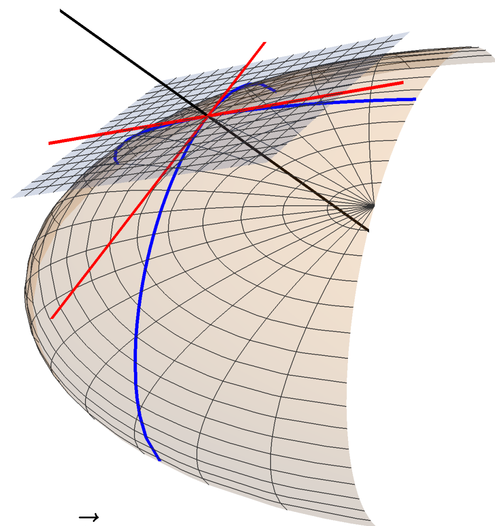
$$D_{\vec{T}}F(x, y, z) = \vec{\nabla}F \cdot \vec{T} = (F_x\vec{i} + F_y\vec{j} + F_z\vec{k}) \cdot (\frac{dx}{ds}\vec{i} + \frac{dy}{ds}\vec{j} + \frac{dz}{ds}\vec{k})$$

$$= F_x\frac{dx}{ds} + F_y\frac{dy}{ds} + F_z\frac{dz}{ds} = \frac{d}{ds}F(x(s), y(s), z(s)) = 0 \Rightarrow \vec{\nabla}F \perp \vec{T}$$

Thus, $\vec{\nabla}F(a, b, c)$ is normal to $\vec{T}$ at (a, b, c) and therefore to σ.

Tangent planes

Consider a level surface σ: $F(x, y, z) = k$,
and let $P = (a, b, c)$ belong to σ.
Since $\vec{\nabla} F(a, b, c)$ is normal to tangent vectors to curves on σ through P,
all these tangent vectors belong to one and the same plane.
This plane is called the **tangent plane** to the surface σ at P.



To find an equation of the tangent plane we use that if we know a vector $\vec{n}$ normal to a plane through a point $\vec{r}_0 = a\,\vec{i} + b\,\vec{j} + c\,\vec{k}$ then an equation of the plane is

$$\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0 \quad \Leftrightarrow \quad n_1(x - a) + n_2(y - b) + n_3(z - c) = 0$$

because $\vec{r} - \vec{r}_0$ is parallel to the plane and therefore normal to $\vec{n}$.

Choosing $\vec{n} = \vec{\nabla} F(a, b, c)$, we get the equation of the tangent plane to the level surface σ at $P = (a, b, c)$

$$F_x(a, b, c)(x - a) + F_y(a, b, c)(y - b) + F_z(a, b, c)(z - c) = 0$$

The line through P parallel to $\vec{\nabla} F(a, b, c)$ is perpendicular to the tangent plane, and is called the **normal line** to the surface σ at P. Its parametric equations are

$$x = a + F_x(a, b, c)t\,, \quad y = b + F_y(a, b, c)t\,, \quad z = c + F_z(a, b, c)t$$

Example. $4x^2 + y^2 + z^2 = 18$ at $(2, 1, 1)$.
Tangent plane, normal line, the angle the tangent plane makes with the xy-plane?

Tangent planes to $z = f(x, y)$

The graph of a function $z = f(x, y)$ can be thought of as the level surface of the function $F(x, y, z) = f(x, y) - z$ with constant 0.

We find

1. the gradient

$$\vec{\nabla} F(a, b, c) = f_x(a, b)\,\vec{i} + f_y(a, b)\,\vec{j} - \vec{k}\,, \quad c = f(a, b)$$

2. the equation of the tangent plane to the surface $z = f(x, y)$ at $(a, b, f(a, b))$

$$f_x(a, b)(x - a) + f_y(a, b)(y - b) - (z - c) = 0 \;\Rightarrow$$
$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

that is the local linear approximation of f at (a, b),

3. the parametric equations of the normal line to the surface $z = f(x, y)$ at $(a, b, f(a, b))$

$$x = a + f_x(a, b)\,t\,, \quad y = b + f_y(a, b)\,t\,, \quad z = f(a, b) - t$$

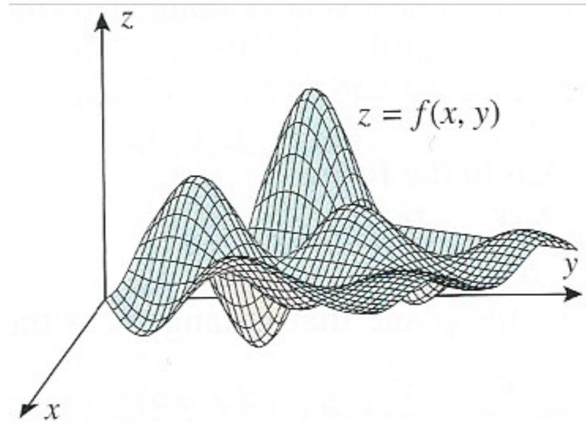
Example. Consider the surface

$$z = f(x, y) = \ln\left(\frac{1}{2} e^{2/3} \sqrt[3]{12\sin(x - 2y) + 8y^2 - x^3 - 6x^2 y + 32}\right)$$

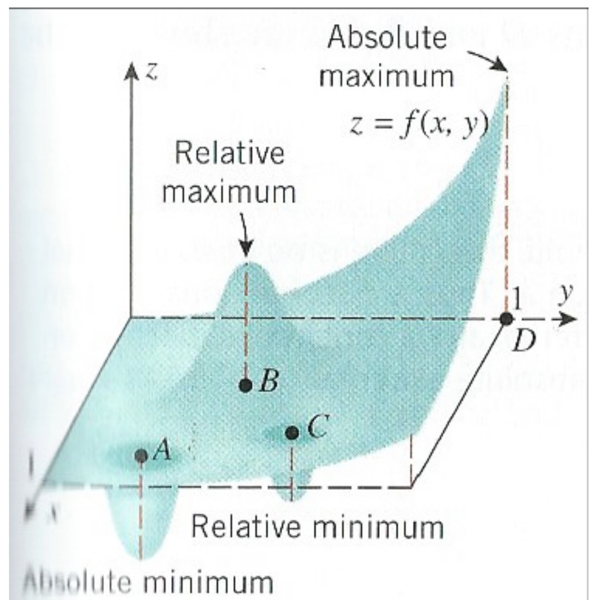
1. Find an equation for the tangent plane and parametric equations for the normal line to the surface at the point $P = (2, 1, z_0)$ where $z_0 = f(2, 1)$.

2. Find points of intersection of the tangent plane with the x-, y- and z-axes. Sketch the tangent plane, and show the point P on it. Sketch the normal line to the surface at P.

8 Maxima and minima of functions of two variables



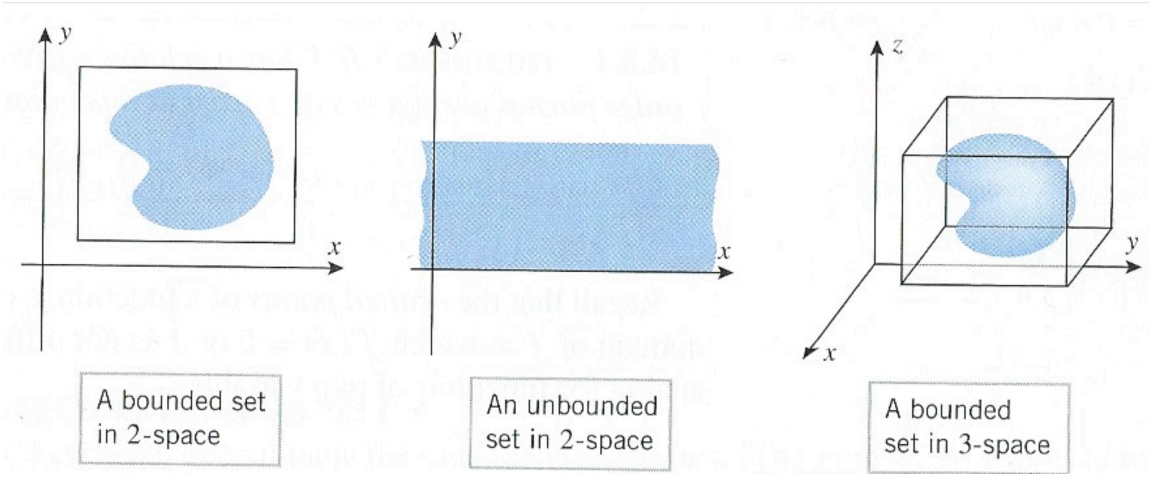
Definition. A function f of two variables is said to have a **relative maximum (minimum)** at a point (a, b) if there is a disc centred at (a, b) such that $f(a, b) \geq f(x, y)$ $(f(a, b) \leq f(x, y))$ for all points (x, y) that lie inside the disc.



A function f is said to have an **absolute maximum (minimum)** at (a, b) if $f(a, b) \geq f(x, y)$ $(f(a, b) \leq f(x, y))$ for all points (x, y) that lie inside in the domain of f.

If f has a relative (absolute) maximum or minimum at (a, b) then we say that f has a **relative (absolute) extremum** at (a, b).

relative $\leftrightarrow$ local



The extreme-value theorem. If $f(x, y)$ is continuous on a closed and bounded set R, then f has both absolute maximum and an absolute minimum on R.

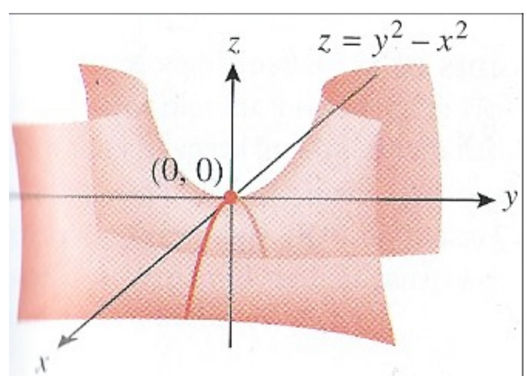
Finding relative extrema

Theorem. If f has a relative extremum at (a, b), and if the first-order derivatives of f exist at this point, then

$$f_x(a, b) = 0 \quad \text{and} \quad f_y(a, b) = 0$$

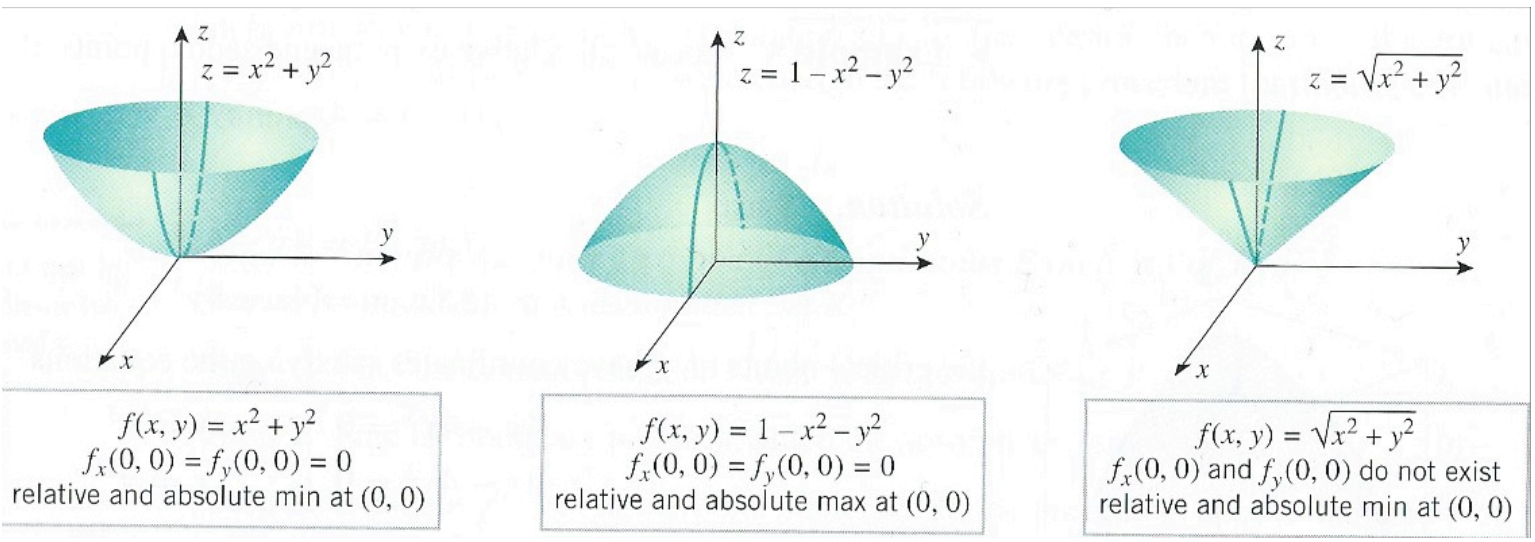
Definition. A point (a, b) in the domain of $f(x, y)$ is called a **critical point** of f if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, or if one or both partial derivatives do not exist at (a, b).

Example. $f(x, y) = y^2 - x^2$ is a hyperbolic paraboloid.
$f_x = -2x$, $f_y = 2y \Rightarrow (0, 0)$ is critical but it is not a relative extremum.
It is a **saddle point**.



We say that a surface $z = f(x, y)$ has a **saddle point** at (a, b) if there are two distinct vertical planes through this point such that the trace of the surface in one of the planes has a relative maximum at (a, b), and the trace in the other has a relative minimum at (a, b).

Example.



How to determine whether a critical point is a max or min?

The second partials test

Theorem. Let $f(x,y)$ have continuous second-order partial derivatives in some disc centred at a critical point (a,b), and let

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

1. If $D > 0$ and $f_{xx}(a,b) > 0$, then f has a relative minimum at (a,b).
2. If $D > 0$ and $f_{xx}(a,b) < 0$, then f has a relative maximum at (a,b).
3. If $D < 0$, then f has a saddle point at(a,b).
4. If $D = 0$, then no conclusion can be drawn.

Example.

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

How to find the absolute extrema of a continuous function of two variables on a closed and bounded set R?

1. Find the critical points of f that lie in the interior of R.
2. Find all the boundary points at which the absolute extrema can occur.
3. Evaluate $f(x,y)$ at the found points. The largest of these values is the absolute maximum, and the smallest the absolute minimum.

Example.

$$f(x,y) = 3x + 6y - 3xy - 7\,, \quad R \text{ is the triangle } (0,0), (0,3), (5,0)$$

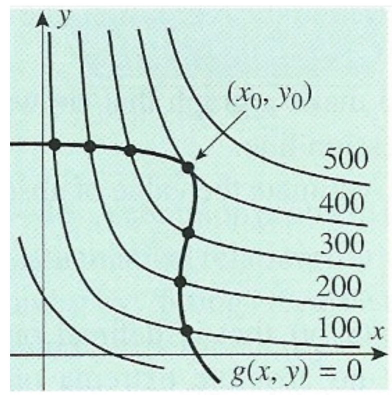
Lagrange multipliers

Extremum problems with constraints:

Find max or min of the function $f(x_1, \ldots, x_n)$ subject to constraints

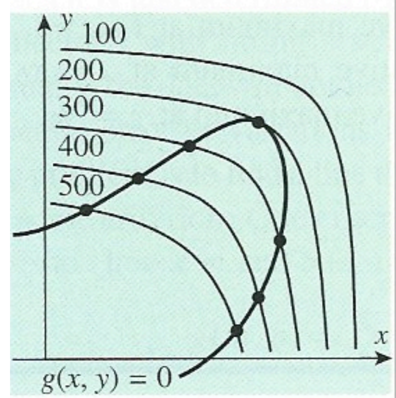
$g_\alpha(x_1, \ldots, x_n)$, $\alpha = 1, \ldots, m$

Consider $f(x, y)$ and $g(x, y) = 0$.
The graph of $g(x, y) = 0$ is a curve.
Consider level curves of f: $f(x, y) = k$.
At (a, b) the curves just touch, and thus have a common tangent line at (a, b). Since $\vec{\nabla} f(a, b)$ is normal to the level curve at (a, b), and $\vec{\nabla} g(a, b)$ is normal to the constraint curve at (a, b), we get $\vec{\nabla} f(a, b) || \vec{\nabla} g(a, b)$



Maximum of $f(x, y)$ is 400

(*a*)



Minimum of $f(x, y)$ is 200

$$\vec{\nabla} f(a, b) = \lambda \vec{\nabla} g(a, b)$$

for some scalar λ called the Lagrange multiplier.

Proof. Parametrise $g(x, y) = 0$.
Then, $f(x, y) = f(x(t), y(t))$ is a function of t and its local extrema are at

$$\frac{d}{dt} f(x(t), y(t)) = \frac{\partial f}{\partial x} x' + \frac{\partial f}{\partial y} y'$$
$$= \vec{\nabla} f \cdot (x' \vec{i} + y' \vec{j}) = \vec{\nabla} f \cdot \vec{T}$$

Thus, both $\vec{\nabla} f$ and $\vec{\nabla} g$ are $\perp$ to $\vec{T}$.

In general, we introduce a Lagrange multiplier λ_α for each of the constraint g_α, and the equations are

$$\vec{\nabla} f = \sum_{\alpha=1}^{m} \lambda_\alpha \, \vec{\nabla} g_\alpha \, .$$

Example. Find the points on the sphere $x^2 + y^2 + z^2 = 36$ that are closest to and farthest from the point $(1, 2, 2)$.