MA 105 : Calculus
Division 3, Lecture 25

Prof. Sudhir R. Ghorpade
IIT Bombay

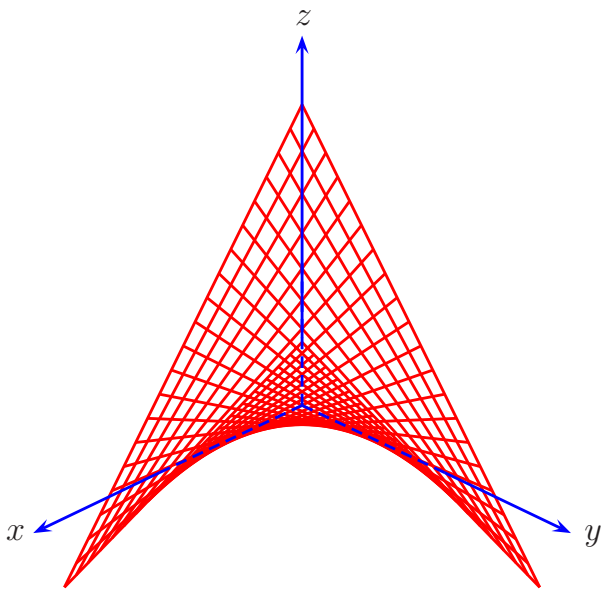
Recap of the previous lecture

- Local extremum
- Necessary condition for a local extremum
- Saddle point. Examples
- Bivariate Mean Value Theorem
- Extended Bivariate Mean Value Theorem
- Bivariate Taylor's theorem

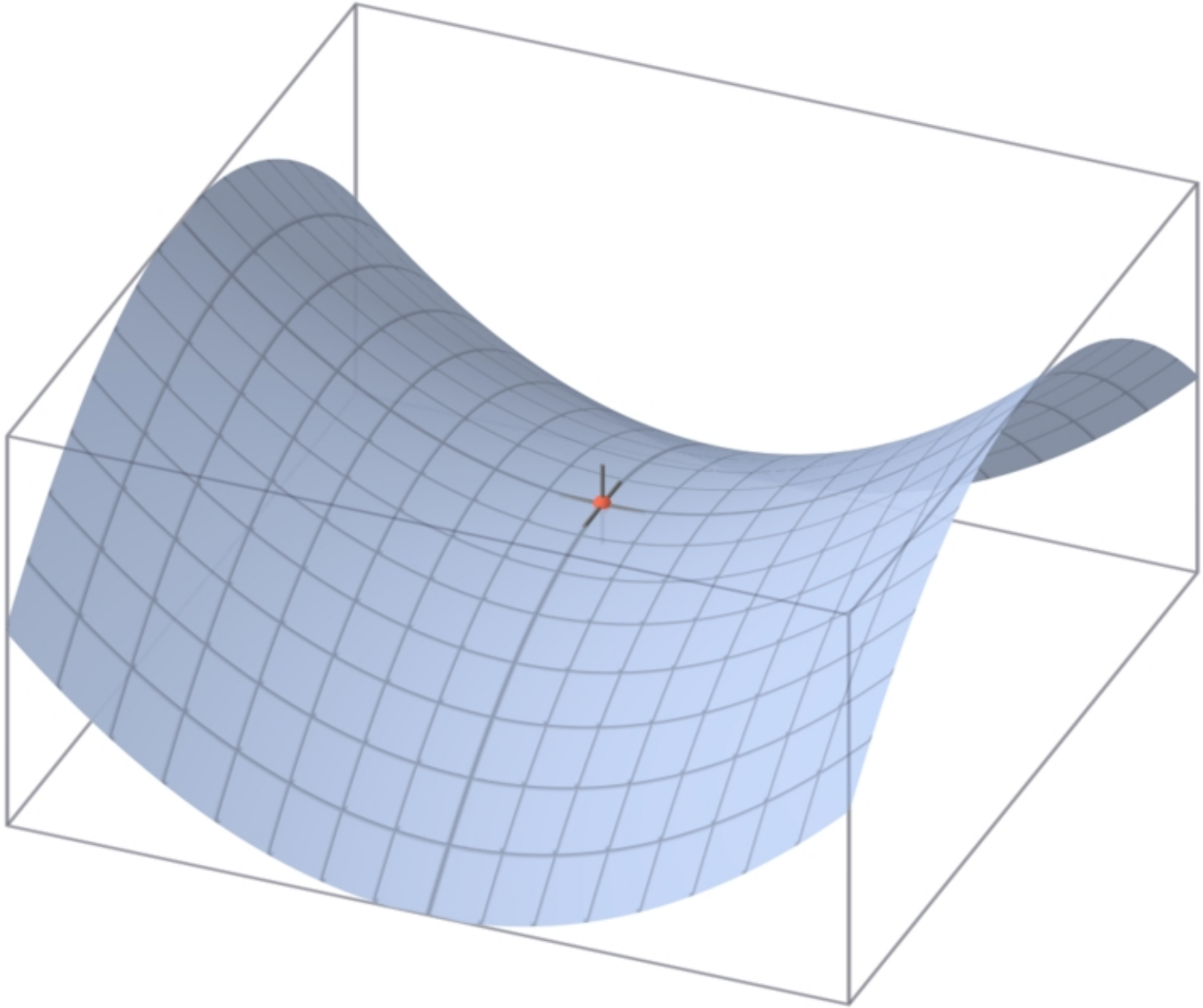
Saddle Point

Let (x_0, y_0) be an interior point of $D \subset \mathbb{R}^2$, and let $f : D \to \mathbb{R}$. We say that the function f has a **saddle point** at (x_0, y_0) if $(\nabla f)(x_0, y_0) = (0, 0)$, but f does not have a local extremum at (x_0, y_0).

Example: $(0, 0)$ is a saddle point of $f(x, y) := x\,y$, $(x, y) \in \mathbb{R}^2$. The surface given by $z = x\,y$ looks as follows.



Picture of a Saddle



Before considering sufficient conditions for a local extremum or for a saddle point, we consider the following useful results.

Let $D \subset \mathbb{R}^2$ be such that every point of D is an interior point, and the line segment joining any two points in D lies in D. Let $(x_0, y_0), (x_1, y_1) \in D$, and consider a function $f : D \to \mathbb{R}$.

(Bivariate Mean Value Theorem)

Suppose f has continuous partial derivatives on D. Then there is $(c, d) \in D$ lying on the open line segment joining (x_0, y_0) and (x_1, y_1) such that

$$f(x_1, y_1) = f(x_0, y_0) + (x_1 - x_0) f_x(c, d) + (y_1 - y_0) f_y(c, d).$$

Proof: Use the one variable MVT for $\phi : [0, 1] \to \mathbb{R}$ defined by $\phi(t) := f(x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0))$, noting that for $t \in (0, 1)$, $\phi'(t) = (x_1 - x_0) f_x((x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0)) + (y_1 - y_0) f_y((x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0))$ by the chain rule. $\square$

To simplify the statements of extensions of the above result, let us introduce a partial differential operator

$$\mathcal{D}_{h,k} := h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}, \quad \text{where } h, k \in \mathbb{R}.$$

It can be applied to a function f which has partial derivatives. Further, by the Mixed Partials Theorem, we obtain

$$\mathcal{D}^2_{h,k} f = \mathcal{D}_{h,k}(\mathcal{D}_{h,k} f) = h^2 f_{xx} + 2hk\ f_{xy} + k^2 f_{yy}$$

if a function f has continuous partial derivatives of order ≤ 2.

(Bivariate Extended Mean Value Theorem)

Suppose f has continuous partial derivatives of order ≤ 2 on D. Let $h := x_1 - x_0$ and $k := y_1 - y_0$. Then there is $(c, d) \in D$ lying on the open line segment joining (x_0, y_0) and (x_1, y_1) with

$$f(x_1, y_1) = f(x_0, y_0) + (\mathcal{D}_{h,k} f)(x_0, y_0) + \frac{1}{2}(\mathcal{D}^2_{h,k} f)(c, d).$$

Thus $f(x_0, y_0) + (\mathcal{D}_{h,k} f)(x_0, y_0)$, or in other words,

$$f(x_0, y_0) + (x_1 - x_0) f_x(x_0, y_0) + (y_1 - y_0) f_y(x_0, y_0)$$

can be considered as an approximation of $f(x_1, y_1)$ with $(1/2)(\mathcal{D}^2_{h,k} f)(c, d)$ as the error.

More generally, we can prove the following result.

(Bivariate Taylor Theorem)

Let $n \in \mathbb{N}$. Suppose f has continuous partial derivatives of order $\leq (n+1)$ on D. Then there is $(c, d) \in D$ lying on the open line segment joining (x_0, y_0) and (x_1, y_1) such that

$$f(x_1, y_1) = \sum_{i=0}^{n} \frac{1}{i!} (\mathcal{D}^i_{h,k} f)(x_0, y_0) + R_n,$$

where $R_n := \dfrac{1}{(n+1)!} (\mathcal{D}^{n+1}_{h,k} f)(c, d)$, $h := x_1 - x_0$, $k := y_1 - y_0$.

The bivariate Taylor theorem can be proved by considering the function $\phi : [0,1] \to \mathbb{R}$ defined by

$$\phi(t) := f(x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0)),$$

and by using the one variable Taylor theorem for ϕ, as in the case of the bivariate mean value theorem. (Note: The proof involves repeated use of the Mixed Partials Theorem.)

The polynomial

$$P_n(x,y) \;:=\; \sum_{i=0}^{n} \frac{1}{i!} (\mathcal{D}^i_{x-x_0,y-y_0} f)(x_0, y_0)$$

of degree $\leq n$ is an 'approximation' of $f(x,y)$ whenever $(x,y) \in D$. It is called the nth **bivariate Taylor polynomial** of f around (x_0, y_0), and

$$R_n(x,y) := \frac{1}{(n+1)!} (\mathcal{D}^{n+1}_{x-x_0,y-y_0} f)(c,d)$$

is called the nth **Lagrange remainder** of f around (x_0, y_0).

Using the Extended Bivariate Mean Value Theorem, we obtain

Proposition (Discriminant Test)

Let $D \subset \mathbb{R}^2$ and let (x_0, y_0) be an interior point of D. Suppose $f : D \to \mathbb{R}$ is such that the first-order and second-order partial derivatives of f exist and are continuous in a neighbourhood of (x_0, y_0), and $(\nabla f)(x_0, y_0) = (0, 0)$. Consider the **discriminant**

$$(\Delta f)(x_0, y_0) := f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2$$

of f at (x_0, y_0).

(i) If $(\Delta f)(x_0, y_0) > 0$ and $f_{xx}(x_0, y_0) < 0$, then f has a local maximum at (x_0, y_0).

(ii) If $(\Delta f)(x_0, y_0) > 0$ and $f_{xx}(x_0, y_0) > 0$, then f has a local minimum at (x_0, y_0).

(iii) If $(\Delta f)(x_0, y_0) < 0$, then f has a saddle point at (x_0, y_0).

The discrimimant test is inconclusive if $(\Delta f)(x_0, y_0) = 0$.

Examples:

(i) Let $f(x,y) := x^2 + y^2$ for $(x,y) \in \mathbb{R}^2$. Then $(\nabla f)(x,y) = (2x, 2y) = (0,0) \iff (x,y) = (0,0)$. Also, $(\Delta f)(0,0) = 4$ and $f_{xx}(0,0) = 2 > 0$. Hence f has a local minimum at $(0,0)$.

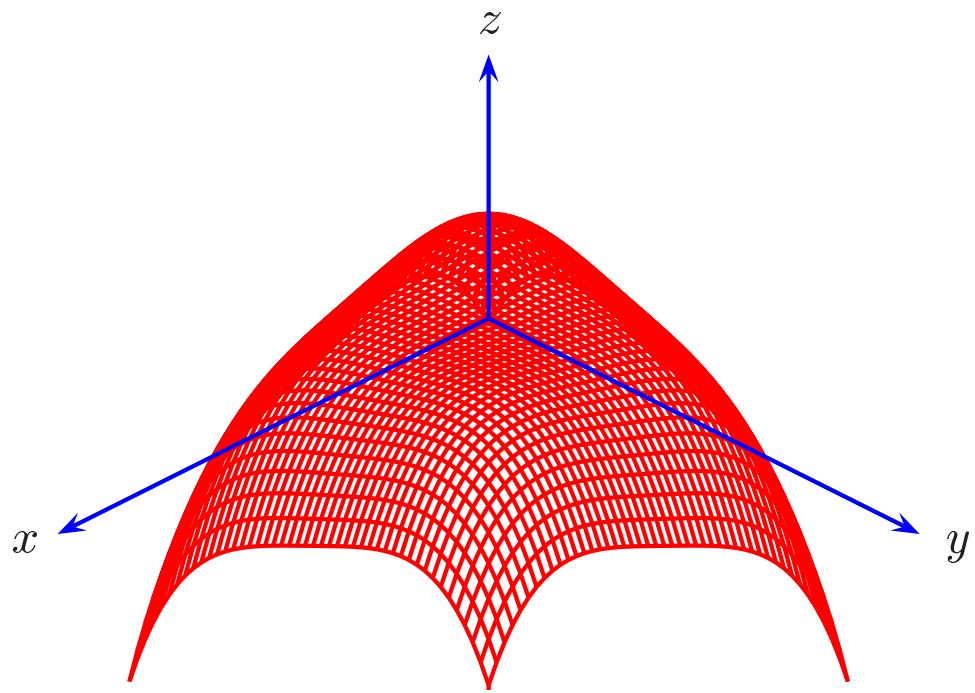
(ii) Let $f(x,y) := -x^2 - y^2$ for $(x,y) \in \mathbb{R}^2$. Then $(\nabla f)(x,y) = (-2x, -2y) = (0,0) \iff (x,y) = (0,0)$. Also, $(\Delta f)(0,0) = 4$ and $f_{xx}(0,0) = -2 < 0$. Hence f has a local maximum at $(0,0)$.

(iii) For $(x,y) \in \mathbb{R}^2$, let $f(x,y) := xy$. Then $(\nabla f)(x,y) = (y,x) = (0,0) \iff (x,y) = (0,0)$. Also, $(\Delta f)(0,0) = -1$. Hence f has a saddle point at $(0,0)$.

(iv) For $(x,y) \in \mathbb{R}^2$, let $f_1(x,y) := x^4 + y^4$, $f_2(x,y) := -x^4 - y^4$, and $f_3(x,y) := x^3 y^3$. Then $(\nabla f_i)(0,0) = (0,0)$ and $(\Delta f_i)(0,0) = 0$ for $i = 1,2,3$. We observe that f_1 has a local minimum, f_2 has a local maximum and f_3 has a saddle point at $(0,0)$.

(v) Let $f(x, y) := 4xy - x^4 - y^4$ for $(x, y) \in \mathbb{R}^2$. Then
$(\nabla f)(x, y) = (4(y - x^3), 4(x - y^3)) = (0, 0)$
$\iff (x, y) = (0, 0),\ (1, 1),$ or $(-1, -1)$.

Further, $(\Delta f)(x, y) = (-12x^2)(-12y^2) - 16 = 16(9x^2y^2 - 1)$ for $(x, y) \in \mathbb{R}^2$. In particular, $(\Delta f)(0, 0) = -16 < 0$ and $(\Delta f)(1, 1) = (\Delta f)(-1, -1) = 128 > 0$. Also, $f_{xx}(1, 1) = f_{xx}(-1, -1) = -12 < 0$. Hence f has a saddle point at $(0, 0)$ and local maxima at $(1, 1)$ and $(-1, -1)$.



(vi) Let $f(x, y) := x \sin y$ for $(x, y) \in \mathbb{R}^2$. Then $(\nabla f)(x, y) = (\sin y,\ x \cos y) = (0, 0) \iff (x, y) = (0, k\pi)$ for $k \in \mathbb{Z}$.

Further, $(\Delta f)(x, y) = 0(-x \sin y) - \cos^2 y = -\cos^2 y$ for $(x, y) \in \mathbb{R}^2$.

In particular, $(\Delta f)(0, k\pi) = -1 < 0$ for each $k \in \mathbb{Z}$.

Hence f has a saddle point at $(0, k\pi)$ for each $k \in \mathbb{Z}$.

Closed Sets and Boundary Points

Let $D \subset \mathbb{R}^2$. We say that D is **closed** if every sequence in D that converges in $\mathbb{R}^2$ converges to a point of D itself.

A point $(x_0, y_0) \in \mathbb{R}^2$ is called a **boundary point** of D if there is a sequence in D that converges to (x_0, y_0) and also a sequence in the complement of D that converges to (x_0, y_0).

The set of all boundary points of D in $\mathbb{R}^2$ is called the **boundary** of D (in $\mathbb{R}^2$). We shall denote it by ∂D.

One can show that D is closed $\iff$ $\partial D \subset D$ (Tutorial 9).

Examples: Let $D_1 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ and $D_2 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$. Then D_1 is closed, but D_2 is not. In fact, $\partial D_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} = \partial D_2$.

Let $D := [a, b] \times [c, d]$. Then D is a closed subset of $\mathbb{R}^2$ and $\partial D = \big\{(x, y) \in D : x \in \{a, b\}$ or $y \in \{c, d\}\big\}$, which is the union of the 4 sides of D.

Finding Global Extrema

Let $D \subset \mathbb{R}^2$. We say that D is **bounded** if there is $\alpha > 0$ such that $\|(x, y)\| \leq \alpha$ for all $(x, y) \in D$.

Theorem (Extreme Value Theorem)

Let D be a nonempty closed and bounded subset of $\mathbb{R}^2$, and let $f : D \to \mathbb{R}$ be continuous. Then f is bounded and attains its bounds on D, that is, there are $(x_1, y_1), (x_2, y_2) \in D$ such that

$$\begin{aligned} f(x_1, y_1) &= \min\{f(x, y) : (x, y) \in D\}, \\ f(x_2, y_2) &= \max\{f(x, y) : (x, y) \in D\}. \end{aligned}$$

As in the case of the extreme value theorem for functions of one variable, the proof is quite involved, and we shall skip it.

Critical Points and Global Exrema

Let $D \subset \mathbb{R}^2$, and let $f : D \to \mathbb{R}$. An interior point (x_0, y_0) of D is called a **critical point** of f if either $(\nabla f)(x_0, y_0)$ does not exist, or $(\nabla f)(x_0, y_0)$ exists and is equal to $(0, 0)$.

Proposition

Let D be a nonempty, closed and bounded subset of $\mathbb{R}^2$, and $f : D \to \mathbb{R}$ be continuous. Then the global minimum m as well as the global maximum M of f is attained either at a critical point of f, or at a boundary point of D.

Proof. Let $(x_1, y_1) \in D$ be such that $f(x_1, y_1) = m$. If $(x_1, y_1) \in \partial D$, then we are done. Otherwise (x_1, y_1) is an interior point of D, and f has a local minimum at (x_1, y_1). If $(\nabla f)(x_1, y_1)$ does not exist, then (x_1, y_1) is a critical point of f. If $(\nabla f)(x_1, y_1)$ exists, then necessarily $(\nabla f)(x_1, y_1) = (0, 0)$, and so (x_1, y_1) is a critical point of f. Similarly, for M. □

Finding Global Extrema

Let D be a nonempty closed and bounded subset of $\mathbb{R}^2$, and let $f : D \to \mathbb{R}$ be a continuous function. Here is a procedure for finding the global extrema of f on D.

(i) Determine the critical points of f in D. (Do not wonder whether f has a local extremum or a saddle point.)

(ii) Determine the boundary ∂D of D. Often ∂D consists of closed and bounded 'one dimensional pieces'.

(iii) Restrict the function f to each 'one dimensional piece' of the boundary of D. Use methods of one variable calculus to find the global extrema of f on that piece.

(iv) Compare the values of f at the critical points of f and at the extreme values of f on the one dimensional piece(s) of ∂D. The largest (resp., smallest) among these values is the global maximum (resp., minimum) of f on D.

We shall later give an alternative to (iii) above, called the Orthogonal Gradient Theorem.

Finding Global Extrema: Example

Example:
Let $D := \{(x, y) \in \mathbb{R}^2 : x,\ y \geq 0 \text{ and } x + y \leq 9\}$, and let $f(x, y) := 2 + 2x + 2y - x^2 - y^2$ for $(x, y) \in D$.

(i) (1,1) is the only critical point of f, and $f(1, 1) = 4$.

(ii) The boundary of D has three closed one dimensional pieces, one each on the lines $x = 0$, $y = 0$ and $y = 9 - x$.

(iii) $f(0, y) = 2 + 2y - y^2$ for $y \in [0, 9]$. At the critical point $y = 1$, $f(0, 1) = 3$. Also, $f(0, 0) = 2$ and $f(0, 9) = -61$.

$f(x, 0) = 2 + 2x - x^2$ for $x \in [0, 9]$. At the critical point $x = 1$, $f(1, 0) = 3$. Also, $f(0, 0) = 2$ and $f(9, 0) = -61$.

$f(x, 9 - x) = -2x^2 + 18x - 61$ for $x \in [0, 9]$. At the critical point $x = 9/2$, $f(9/2, 9/2) = -41/2$. Also, $f(0, 9) = -61 = f(9, 0)$.

Finding Global Extrema

(iv) The above values of f are tabulated as follows:

(x, y)	(1,1)	(0,1)	(0,0)	(0,9)	(1,0)	(9,0)	(9/2,9/2)
$f(x, y)$	4	3	2	-61	3	-61	-41/2

Thus, the global maximum of f on D is 4, and it is attained at $(1, 1)$, while the global minimum of f on D is -61, and it is attained at $(0, 9)$ as well as at $(9, 0)$.

Aliter: Clearly, $f(x, y) = 4 - (x - 1)^2 - (y - 1)^2 \leq 4$ for all $(x, y) \in \mathbb{R}^2$. Also, since $(9, 0)$ and $(0, 9)$ are the farthest points of D from the point $(1, 1)$, we see that $f(x, y) \geq 4 - (9 - 1)^2 - (0 - 1)^2 = -61$ for all $(x, y) \in D$.

In fact, $f(1, 1) = 4$ and $f(9, 0) = f(0, 9) = -61$.

If D is a closed and bounded subset of $\mathbb{R}^2$, and $f : D \to \mathbb{R}$ is continuous, then the values of f on the boundary ∂D of D play a special role when we find the global extrema of f on D.

Since the boundary of D often consists of one or more curves, we need to find the extreme values of f, when it is restricted to a curve in D. In this connection the following result is useful.

Proposition (Orthogonal Gradient Theorem)

Let $E \subset \mathbb{R}^2$ and (x_0, y_0) be an interior point of E. Let C be a parametrized curve in E passing through (x_0, y_0) given by $(x(t), y(t))$, where $x, y : [\alpha, \beta] \to E$, and let $t_0 \in (\alpha, \beta)$ with $x(t_0) = x_0$, $y(t_0) = y_0$. Suppose a function $f : E \to \mathbb{R}$, when restricted to C, has a local extremum at (x_0, y_0). If x and y are differentiable at t_0, and f is differentiable at (x_0, y_0), then $(\nabla f)(x_0, y_0) \cdot (x'(t_0), y'(t_0)) = 0$, that is, the gradient of f at (x_0, y_0) is perpendicular to the tangent vector of C at (x_0, y_0).

Proof. Define $\phi : [\alpha, \beta] \to \mathbb{R}$ by $\phi(t) := f(x(t), y(t))$. Then ϕ has a local extremum at t_0, and so $\phi'(t_0) = 0$. But by the chain rule (ii), $\phi'(t_0) = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0)$. □

Example: Let $E := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 2\}$, and $f(x, y) := x\,y$ for $(x, y) \in E$. Let us find the global extrema of f on $D := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

- $(\nabla f)(x, y) = (y, x)$ if $x^2 + y^2 < 1$. Hence $(0, 0)$ is the only critical point of f, and $f(0, 0) = 0$.
- The boundary of D is given by $(x(t), y(t))$, where $x(t) := \cos t$, $y(t) := \sin t$ for $t \in [-\pi, \pi]$. Each point of D is an interior point of E.
- $(\nabla f)(x(t), y(t)) \cdot (x'(t), y'(t)) = -\sin^2 t + \cos^2 t = 0$ $\iff$ $t = \pm\pi/4, \pm 3\pi/4$, that is, $(x(t), y(t))$ $= \big(\pm 1/\sqrt{2},\ \pm 1/\sqrt{2}\big)$.
- $f(1/\sqrt{2}, 1/\sqrt{2}) = f(-1/\sqrt{2}, -1/\sqrt{2}) = 1/2$ and $f(-1/\sqrt{2}, 1/\sqrt{2}) = f(1/\sqrt{2}, -1/\sqrt{2}) = -1/2$.

So the maximum of f is $1/2$ and the minimum of f is $-1/2$.