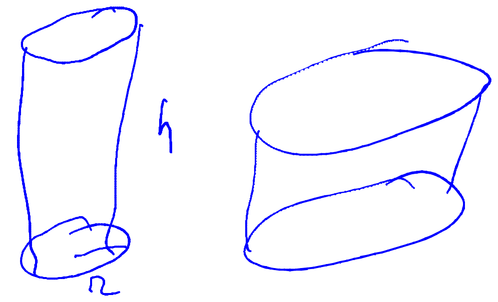
2.3 Extrem value problems under constraints

Question: What is the size of a metallic cylindrical can in order to minimize the material amount by given volume?

Ansatz for solution: Let $r > 0$ be the radius and $h > 0$ the height of the can. Then

$$\boxed{V = \pi r^2 h}$$

$$\boxed{O = 2\pi r^2 + 2\pi r h}$$



Top + bottom

Let $c \in \mathbb{R}_+$ be the given volume (with $x := r, y := h$),

$$\boxed{f(x,y) = 2\pi x^2 + 2\pi x y}$$

$$\boxed{g(x,y) = \pi x^2 y - c = 0}$$

given Volume c

Determine the minimum of the function $f(x,y)$ on the set

$$\boxed{G := \{(x,y) \in \mathbb{R}_+^2 \,|\, g(x,y) = 0\}}$$

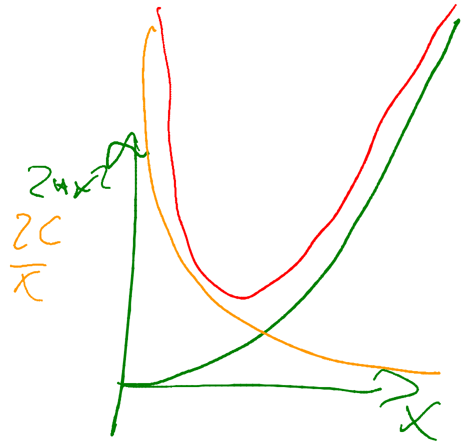
Solution of the constraint minimisation problem.

From $\boxed{g(x,y) = \pi x^2 y - c = 0}$ follows

$$y = \frac{c}{\pi x^2} = y(x)$$

We plug this into $f(x,y)$ and obtain $h(x) = f(x, y(x))$

$$h(x) := 2\pi x^2 + 2\pi x \frac{c}{\pi x^2} = 2\pi x^2 + \frac{2c}{x}$$



Determine the minimum of the function $h(x)$:

$$h'(x) = 4\pi x - \frac{2c}{x^2} = 0 \quad \Rightarrow \quad 4\pi x = \frac{2c}{x^2} \quad \Rightarrow \quad \boxed{x = \left(\frac{c}{2\pi}\right)^{1/3}}$$

Sufficient condition

$$h''(x) = 4\pi + \frac{4c}{x^3} \quad \Rightarrow \quad h''\left(\left(\frac{c}{2\pi}\right)^{1/3}\right) = 12\pi > 0$$

$$f(x_{min}, y(x_{min})) = 2\pi x_{min}^2 + 2\pi x_{min} \cdot y_{min} = 2\pi\left(\frac{c}{2\pi}\right)^{2/3} + 2\pi\left(\frac{c}{2\pi}\right)^{1/3} \cdot \frac{c}{\pi}\left(\frac{2\pi}{c}\right)^{2/3} = (2\pi)^{1/3}\left(c^{2/3} + 2c^{2/3}\right) = \boxed{3(2\pi)^{1/3} \cdot c^{2/3}}$$

General formulation of the problem.

Determine the extrem values of the function $f : \mathbb{R}^n \to \mathbb{R}$ under the constraint Skalar

$$\boxed{\mathbf{g}(\mathbf{x}) = 0}$$

m equations

where $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$.

The constraints are

$$\begin{aligned} g_1(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ g_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

Alternatively: Determine the extrem values of the function $f(\mathbf{x})$ on the set

$$G := \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$$

The Lagrange–function and the Lagrange–Lemma.

We define the Lagrange–function

$$F(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) \quad = G(x, \lambda)$$

and look for the extrem values of $F(\mathbf{x})$ for fixed $\lambda = (\lambda_1, \ldots, \lambda_m)^T$.

The numbers λ_i, $i = 1, \ldots, m$ are called Lagrange–multiplier.

Theorem: (Lagrange–Lemma) If $\mathbf{x}^0$ minimizes (or maximizes) the Lagrange–function $F(\mathbf{x})$ (for a fixed λ) on D and if $\mathbf{g}(\mathbf{x}^0) = 0$ holds, then $\mathbf{x}^0$ is the minimum (or maximum) of $f(\mathbf{x})$ on $G := \{\mathbf{x} \in D \,|\, \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$.

Proof: For an arbitrary $\mathbf{x} \in D$ we have minimum $F(x^0) \le F(x)$

$$F(x^0) = \boxed{f(\mathbf{x}^0)} + \lambda^T \underbrace{\mathbf{g}(\mathbf{x}^0)}_{=0} \le f(\mathbf{x}) + \lambda^T \underbrace{\mathbf{g}(\mathbf{x})}_{=0} = F(x)$$

If we choose $\mathbf{x} \in G$, then $\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}^0) = \mathbf{0}$, thus $f(\mathbf{x}^0) \le f(\mathbf{x})$.

If F max/min and $g(x)=0$ $\Longrightarrow$ f is max/min under $g(x)=0$

If f extreme value + $g(x)=0$ + rank is ... $\Longrightarrow$ grad $F = 0$

A necessary condition for local extrema.

Let f and g_i, $i = 1, \ldots, m$, $\mathcal{C}^1$–functions, then a necessary condition for an extrem value $\mathbf{x}^0$ of $F(\mathbf{x})$ is given by

x n, λ_i m } $n+m$ unknowns

$$\boxed{\text{grad}\, F(\mathbf{x})} = \text{grad}\, f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \text{grad}\, g_i(\mathbf{x}) = \mathbf{0}$$

n, m } equations $n+m$

Together with the constraints $\boxed{\mathbf{g}(\mathbf{x}) = 0}$ we obtain a set of (non-linear) equations with $(n+m)$ equations and $(n+m)$ unknowns $\mathbf{x}$ and λ.

The solutions $(\mathbf{x}^0, \lambda^0)$ are the candidates for the extrem values, since these solutions satisfy the above necessary condition.

Alternatively: Define a Langrange–function

$$0 = \text{grad}_{(x,\lambda)}\, G = \begin{pmatrix} \text{grad}_x F \\ g_1 \\ \vdots \\ g_m \end{pmatrix} = 0$$

$$\boxed{G(\mathbf{x}, \lambda)} := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

and look for the extrem values of $G(\mathbf{x}, \lambda)$ with respect to $\mathbf{x}$ **and** λ.

Some remarks on sufficient conditions.

1. We can formulate a **sufficient** condition:
   If the functions f and $\mathbf{g}$ are $\mathcal{C}^2$–functions and if the Hesse–matrix $\mathbf{H}F(\mathbf{x}^0)$ of the Lagrange–function is positiv (negativ) definit, then $\mathbf{x}^0$ is a strict local minimum (maximum) of $f(\mathbf{x})$ on G.
2. In most of the applications the necessary condition are **not** satisfied, allthough $\mathbf{x}^0$ is a strict local extremum.
3. And from the indefinitness of the Hesse–matrix $\mathbf{H}F(\mathbf{x}^0)$ we **cannot** conclude, that $\mathbf{x}^0$ is not an extremum.
4. We have a similar problem with the necessary condition which is obtained from the Hesse–matrix of the Lagrange–function $G(\mathbf{x}, \lambda)$ with respect to **x and** λ.

$$f(x,y) = 2\pi x^2 + 2\pi xy$$
$$g(x,y) = \pi x^2 y - c = 0$$

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

$$\operatorname{grad} F(x,y) = \begin{pmatrix} 4\pi x + 2\pi y + 2\lambda\pi xy \\ 2\pi x + \lambda \pi x^2 \\ \pi x^2 y - c \end{pmatrix} = 0 \implies (2+\lambda x)x = 0$$

$$x \neq 0 \quad \lightning \qquad x = -\frac{2}{\lambda}$$

$$y = \frac{c}{\pi x^2} = \frac{c\lambda^2}{\pi\, 4}$$

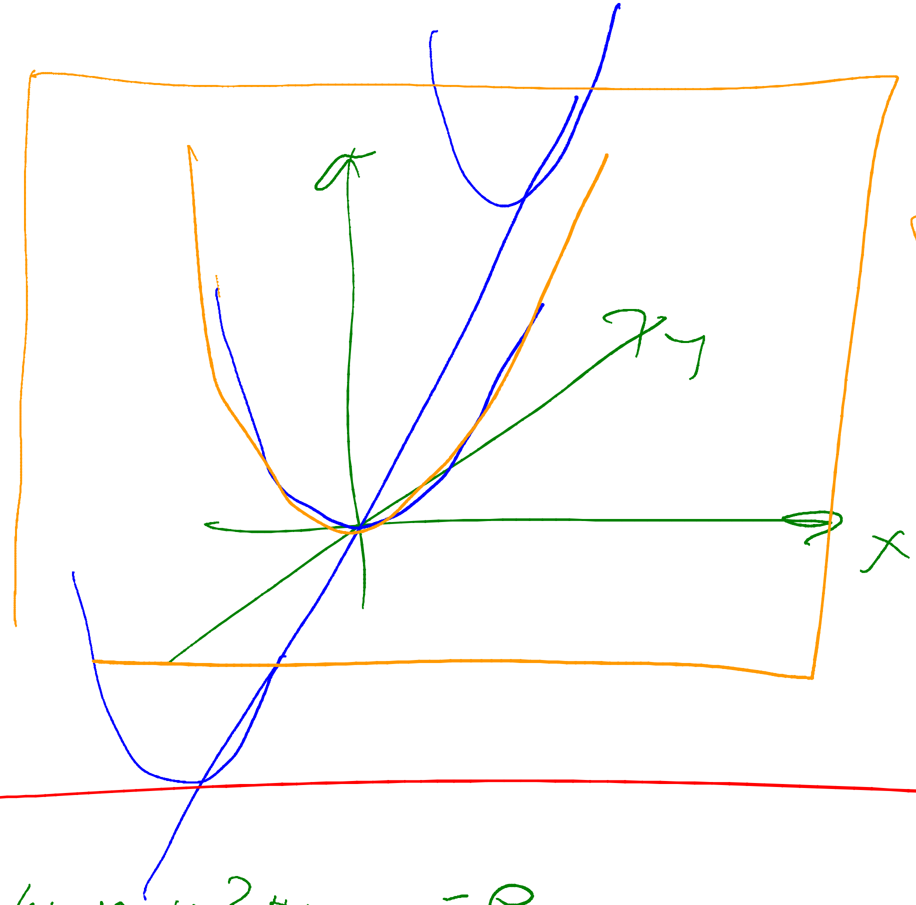
$$0 = 2x + y + \lambda xy = 2x + y \cancel{- \frac{2}{\lambda}\lambda xy} = 2x - y = 0$$

$$y = 2x$$

$$\pi x^2 2x = c \qquad x = \left(\frac{c}{2\pi}\right)^{1/3}$$

$$HF = \begin{pmatrix} 4\pi + 2\lambda\pi y & 2\pi + 2\pi\lambda x \\ 2\pi + 2\pi\lambda x & 0 \end{pmatrix} = 2\pi \begin{pmatrix} 2 + \lambda y & 1 + \lambda x \\ 1 + \lambda x & 0 \end{pmatrix}$$

$$\det HF = -(1+\lambda x)^2 \leq 0 \implies \text{no conclusioni}$$



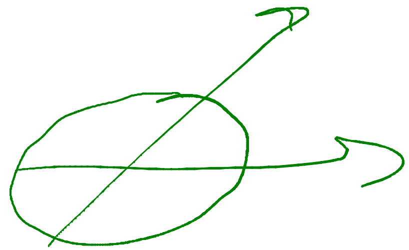
$f(x,y) = \cancel{y} = 0$

$$4a\,x + 2b\,y = 0$$
$$2a\,x = 0 \Rightarrow x = 0 \Rightarrow y = 0$$
$$a\,x^2 y - c = 0$$

↯

An example of a minimisation problem with constraints.

We look for extrem values of $\boxed{f(x, y) := xy}$ on the disc



$$K := \{(x, y)^T \,|\, \underline{x^2 + y^2 \leq 1}\}$$

Since the function f is continuous and $K \subset \mathbb{R}^2$ compact we conclude from the min–max–property the existence of global maxima and minima on K.

We consider first the interior K^0 of K, i.e. the open set

1. step

$$K^0 := \{(x, y)^T \,|\, x^2 + y^2 < 1\}$$

interior

The necessary condition for an extrem value is given by

$$\operatorname{grad} f = (y, x) = \mathbf{0}$$

Thus the origin $\mathbf{x}^0 = \mathbf{0}$ is a candidate for a (local) extrem value.

continuation of the example.

The Hesse–matrix at the origin is given by

$$\mathbf{H}f(\mathbf{0}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and is indefinit. Thus $\mathbf{x}^0$ is a saddel point.

Therefore the extrem values have to be on the boundary which is represented by a constraint equation:

$$\boxed{g(x,y) = x^2 + y^2 - 1 = 0}$$

Therefore we look for the extrem values of $f(x,y) = xy$ under the constraint $g(x,y) = 0$.

The Lagrange–function is given by

$$\rightarrow \qquad F(x,y) = xy + \lambda(x^2 + y^2 - 1)$$

Completion of the example.

We obtain the non-linear system of equations

$2\lambda = -\frac{y}{x}$

$$\operatorname{grad} F = \left\{ \begin{array}{rcl} y + 2\lambda x & = & 0 \\ x + 2\lambda y & = & 0 \end{array} \right.$$

$$g(x,y) = \quad x^2 + y^2 = 1$$

$x \neq \frac{y}{x} y - \frac{1}{x}(x^2 - y^2) = 0$

$2x^2 = 1 \qquad x = \pm\sqrt{\frac{1}{2}} = y$

$\lambda = \pm\frac{1}{2}$

with the four solution

$$\lambda = \frac{1}{2} \quad : \quad \mathbf{x}^{(1)} = (\sqrt{1/2}, -\sqrt{1/2})^T \quad \mathbf{x}^{(2)} = (-\sqrt{1/2}, \sqrt{1/2})^T$$

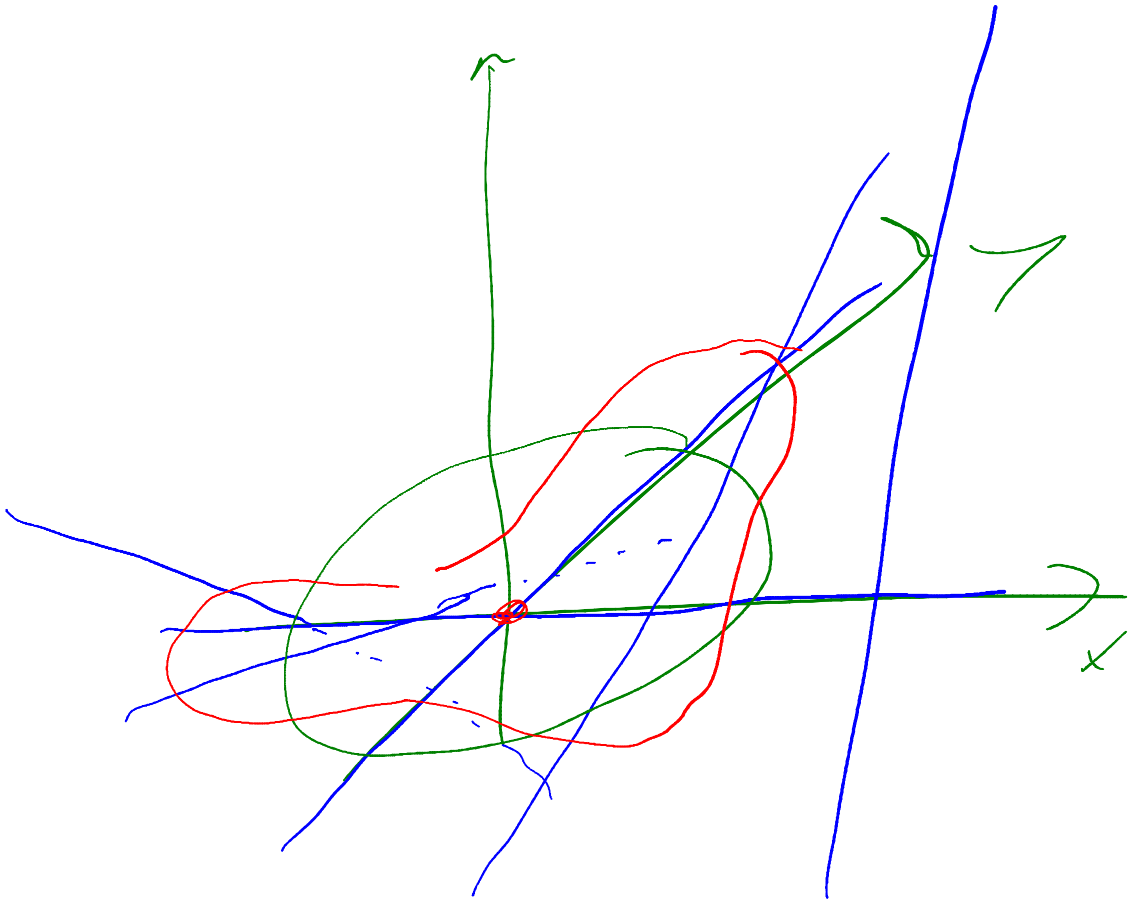
$$\lambda = -\frac{1}{2} \quad : \quad \mathbf{x}^{(3)} = (\sqrt{1/2}, \sqrt{1/2})^T \quad \mathbf{x}^{(4)} = (-\sqrt{1/2}, -\sqrt{1/2})^T$$

Minima and Maxima can be concluded from the values of the function

$$f(\mathbf{x}^{(1)}) = f(\mathbf{x}^{(2)}) = -1/2 \qquad f(\mathbf{x}^{(3)}) = f(\mathbf{x}^{(4)}) = 1/2$$

i.e. minima are $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, maxima are $\mathbf{x}^{(3)}$ and $\mathbf{x}^{(4)}$.

$f = xy$



Lagrange–multiplier–rule.

Satz: Let $f, g_1, \ldots, g_m : D \to \mathbb{R}$ be $\mathcal{C}^1$–functions, und let $\mathbf{x}^0 \in D$ a local extrem value of $f(\mathbf{x})$ under the constraint $\mathbf{g}(\mathbf{x}) = \mathbf{0}$. In addition let the regularity condition

$$\operatorname{rang}\left(\underbrace{\mathbf{J}\,\mathbf{g}(\mathbf{x}^0)}_{m \times n}\right) = m$$

Matrix $\quad m < n$

hold true. Then there exist Lagrange–multiplier $\lambda_1, \ldots, \lambda_m$, such that for the Lagrange function

$$F(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

the following **first order necessary condition** holds true:

$$\operatorname{grad} F(\mathbf{x}^0) = \mathbf{0}$$

Necessary condition of second order and sufficient condition.

Theorem: 1) Let $\mathbf{x}^0 \in D$ a local minimum of $f(\mathbf{x})$ under the constraint $\mathbf{g}(\mathbf{x}) = 0$, let the regularity condition be satisfied and let $\lambda_1, \ldots, \lambda_m$ be the related Lagrange–multiplier. Then the Hesse–matrix $\mathbf{H}F(\mathbf{x}^0)$ of the Lagrange–function is positiv semi-definit on the tangential space

$$TG(\mathbf{x}^0) := \{\mathbf{y} \in \mathbb{R}^n \,|\, \operatorname{grad} g_i(\mathbf{x}^0) \cdot \mathbf{y} = 0 \text{ for } i = 1, \ldots, m\}$$

i.e. it is $\mathbf{y}^T\ \mathbf{H}F(\mathbf{x}^0)\ \mathbf{y} \geq 0$ for all $\mathbf{y} \in TG(\mathbf{x}^0)$.

2) Let the regularity condition for a point $\mathbf{x}^0 \in G$ be staisfied. If there exist Lagrange–multiplier $\lambda_1, \ldots, \lambda_m$, such that $\mathbf{x}^0$ is a stationary point of the related Lagrange–function. Let the Hesse–matrix $\mathbf{H}F(\mathbf{x}^0)$ be positiv definit on the tangential space $TG(\mathbf{x}^0)$, i.e. it holds

$$\mathbf{y}^T\ \mathbf{H}F(\mathbf{x}^0)\ \mathbf{y} > 0 \quad \forall\, \mathbf{y} \in TG(\mathbf{x}^0) \setminus \{0\},$$

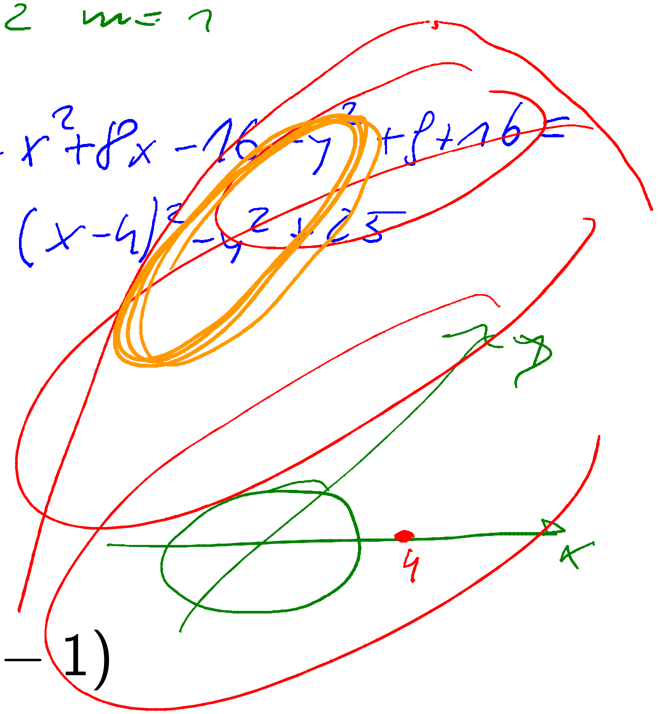
then $\mathbf{x}^0$ is a strict local minimum of $f(\mathbf{x})$ under the constraint $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Example.

Determine the global maximum of the function $n=2 \quad m=1$

$$f(x,y) = -x^2 + 8x - y^2 + 9 = -x^2+8x-16-y^2+9+16 =$$

$$= -(x-4)^2 - y^2 + 25$$



under the constraint

$$g(x,y) = x^2 + y^2 - 1 = 0$$

The Lagrange–function is given by

$$F(x) = -x^2 + 8x - y^2 + 9 + \lambda(x^2 + y^2 - 1)$$

From the necessary condition we obtain the non-linear system

$$0 = \operatorname{grad} F \Rightarrow \left\{ \begin{array}{rcl} -2x + 8 & = & -2\lambda x \\ -2y & = & -2\lambda y \end{array} \right.$$

$$g(x,y) = 0 \Rightarrow \quad x^2 + y^2 = 1$$

Continuation of the example.

From the necessary condition we obtain the non-linear system

$$\begin{aligned} -2x+8 &= -2\lambda x && \Rightarrow \lambda \neq 1 \\ -2y &= -2\lambda y && \Rightarrow y = 0 \\ x^2+y^2 &= 1 && x = \pm 1 \end{aligned}$$

The first equation gives $\lambda \neq 1$. Using this in the second equation we get $y = 0$. From the third equation we obtain $x = \pm 1$.

Therefore the two points $\boxed{(x,y) = (1,0)}$ and $\boxed{(x,y) = (-1,0)}$ are candidates for a global maximum. Since

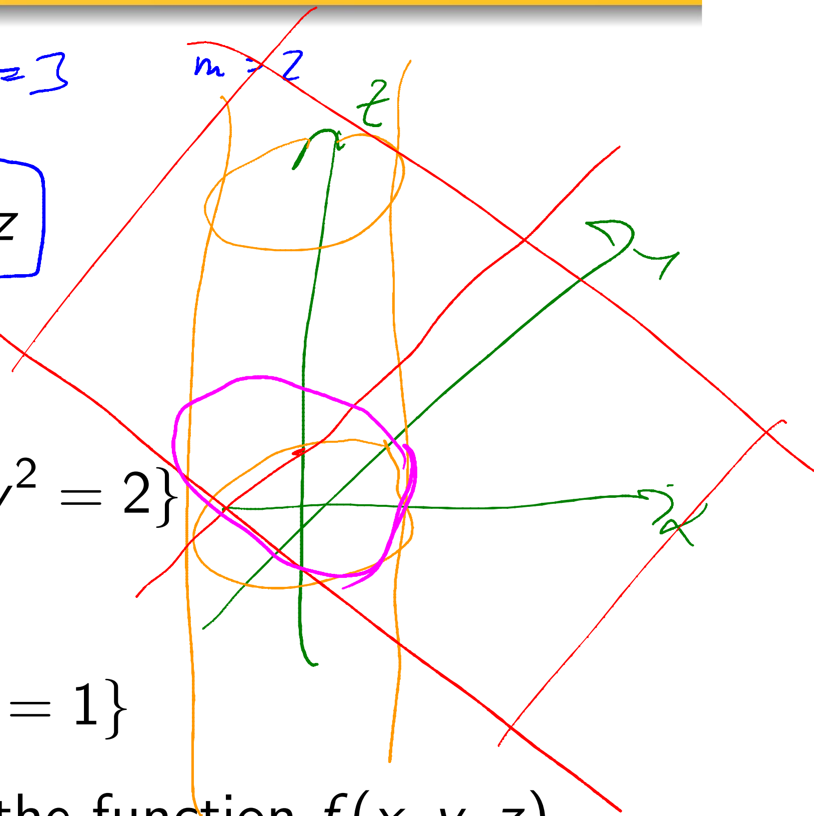
$$f(1,0) = 16 \qquad f(-1,0) = 0$$

the global maximum of $f(x,y)$ under the constraint $g(x,y) = 0$ is given at the point $(x,y) = (1,0)$.

Another example.

Determine the local extrem values of

n=3 m=2



$$\boxed{f(x,y,z) = 2x + 3y + 2z}$$

on the intersection of the cylinder surface

$$M_Z := \{(x,y,z)^T \in \mathbb{R}^3 \,|\, x^2 + y^2 = 2\}$$

with the plane

$$E := \{(x,y,z)^T \in \mathbb{R}^3 \,|\, x + z = 1\}$$

Reformulation: Determine the extrem values of the function $f(x,y,z)$ under the constraint

$$\begin{cases} g_1(x,y,z) := x^2 + y^2 - 2 = 0 \\ g_2(x,y,z) := x + z - 1 = 0 \end{cases}$$

Continuation of the example.

The Jacobi–matrix

$$\mathbf{Jg}(\mathbf{x}) = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$\operatorname{rang} Jg(x) \overset{?}{=} 2$

has rank 2, i.e. we can determine extrem values using the Lagrange–function:

$$F(x, y, z) = 2x + 3y + 2z + \underline{\underline{\lambda_1}}(x^2 + y^2 - 2) + \underline{\underline{\lambda_2}}(x + z - 1)$$

The necessary condition gives the non-linear system

$x, y, z, \lambda_1, \lambda_2$

$$\operatorname{grad} F = 0 \Rightarrow \left\{ \begin{array}{rcl} 2 + 2\lambda_1 x + \lambda_2 & = & 0 \\ 3 + 2\lambda_1 y & = & 0 \\ 2 + \lambda_2 & = & 0 \end{array} \right.$$

$$\begin{array}{l} f_1 = 0 \\ f_2 = 0 \end{array} \quad \left\{ \begin{array}{rcl} x^2 + y^2 & = & 2 \\ x + z & = & 1 \end{array} \right.$$

Continuation of the example.

The necessary condition gives the non-linear system

$$\begin{aligned} 2 + 2\lambda_1 x + \lambda_2 &= 0 \\ 3 + 2\lambda_1 y &= 0 \\ 2 + \lambda_2 &= 0 \\ x^2 + y^2 &= 2 \\ x + z &= 1 \end{aligned}$$

$2\lambda_1 x = 0$ $\quad x = 0$

$\Rightarrow \lambda_1 \neq 0$ $\Uparrow$

$y = \pm\sqrt{2}$

$z = 1$

From the first and the third equation it follows

$$2\lambda_1 x = 0$$

From the second equation it follows $\lambda_1 \neq 0$, i.e. $x = 0$.
Thus we have possible extrem values

$$\boxed{(x, y, z) = (0, \sqrt{2}, 1) \quad (x, y, z) = (0, -\sqrt{2}, 1)}$$

Completion if the example.

The possible extrem values are

$$(x, y, z) = (0, \sqrt{2}, 1) \quad (x, y, z) = (0, -\sqrt{2}, 1)$$

and lie on the cylinder surface M_Z of the cylinder Z with

$$\begin{aligned} Z &= \{(x, y, z)^T \in \mathbb{R}^3 \,|\, x^2 + y^2 \leq 2\} \\ M_Z &= \{(x, y, z)^T \in \mathbb{R}^3 \,|\, x^2 + y^2 = 2\} \end{aligned}$$

We calculate the related functiuon values

$$\begin{aligned} f(0, \sqrt{2}, 1) &= 3\sqrt{2} + 2 \\ f(0, -\sqrt{2}, 1) &= -3\sqrt{2} + 2 \end{aligned}$$

Thus the point $(x, y, z) = (0, \sqrt{2}, 1)$ is a maximum an the point $(x, y, z) = (0, -\sqrt{2}, 1)$ a minimum.