Optimization: an Overview

Moritz Diehl
University of Freiburg and University of Leuven

(some slide material was provided by W. Bangerth and K. Mombaur)

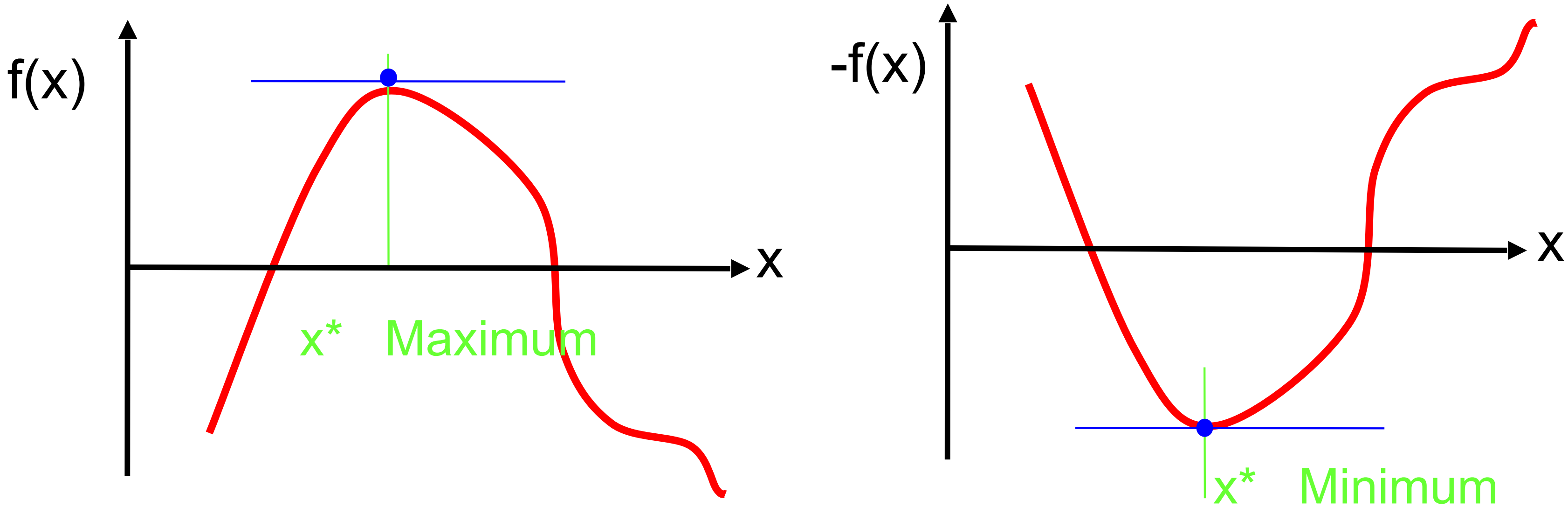
Overview of presentation

- **Optimization: basic definitions and concepts**
- Introduction to classes of optimization problems

What is optimization?

- Optimization = search for the best solution
- in mathematical terms:
  minimization or maximization of an objective function f (x)
  depending on variables x subject to constraints

Equivalence of maximization and minimization problems:
(*from now on only minimization)*



Constrained optimization

- Often variable x shall satisfy certain constraints, e.g.:
  - $x \geq 0$
  - $x_1^2 + x_2^2 = C$
- General formulation:

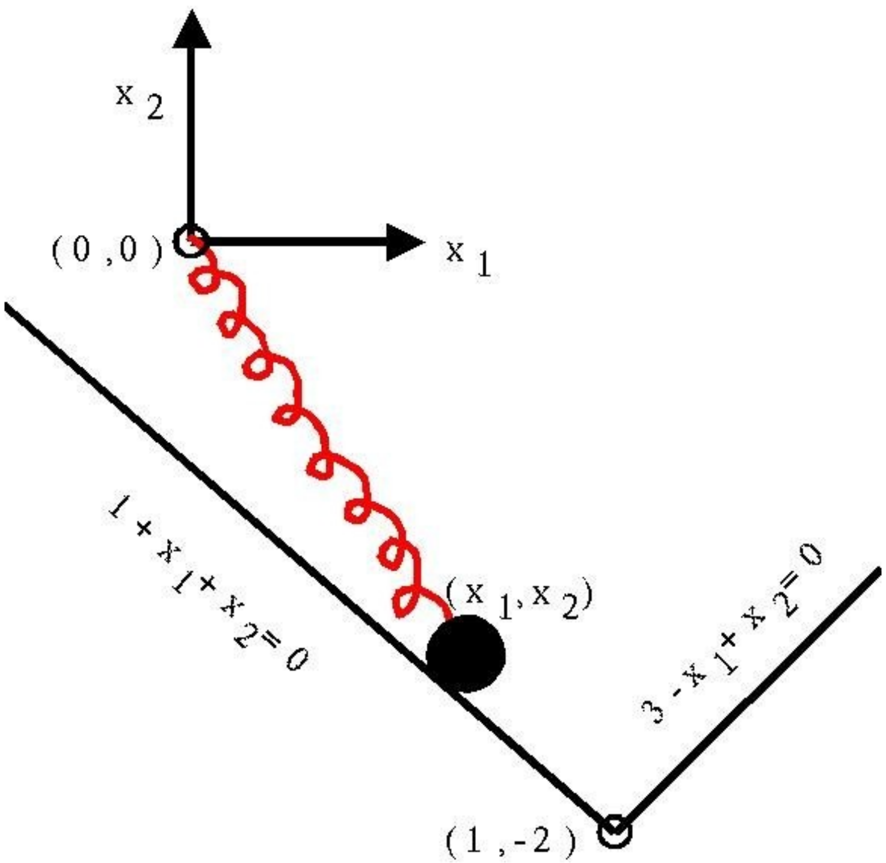
$$\min f(x)$$

$$\text{subject to (s.t.)}$$

$$\begin{aligned} g(x) &= 0 \\ h(x) &\geq 0 \end{aligned}$$

f objective function / cost function
g equality constraints
h inequality constraints

Simple example: Ball hanging on a spring



To find position at rest,
minimize potential energy!

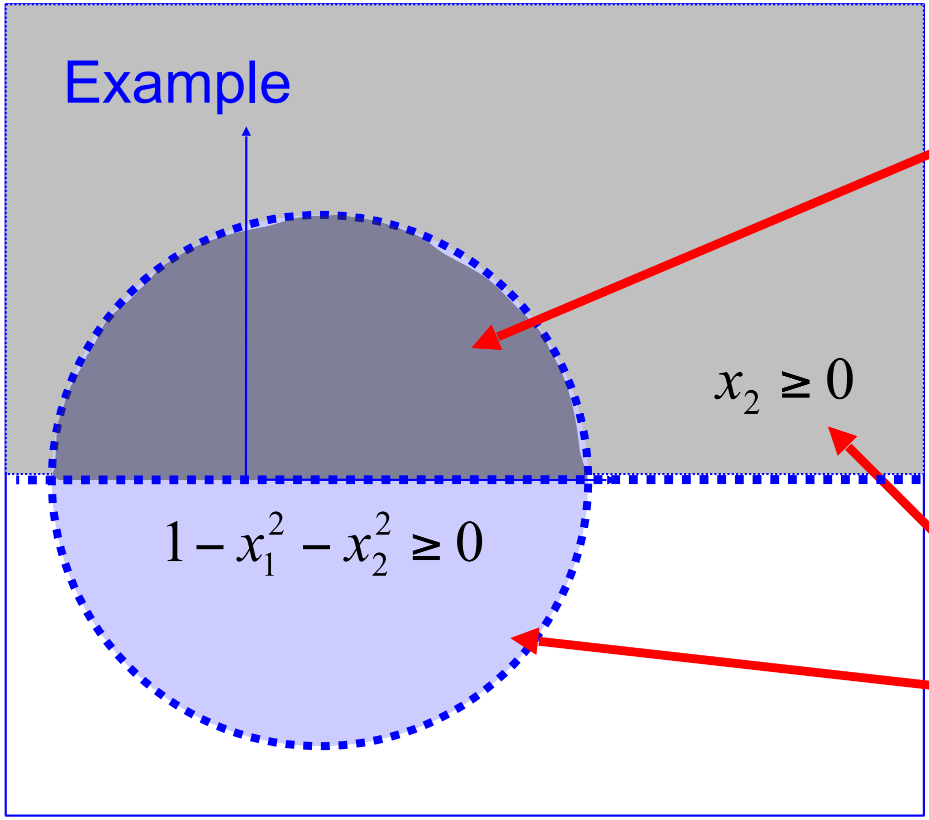
$$\min \underbrace{x_1^2 + x_2^2}_{\textit{spring}} + \underbrace{mx_2}_{\textit{gravity}}$$

$$1 + x_1 + x_2 \geq 0$$

$$3 - x_1 + x_2 \geq 0$$

Feasible set

Feasible set = collection of all points that satisfy all constraints:



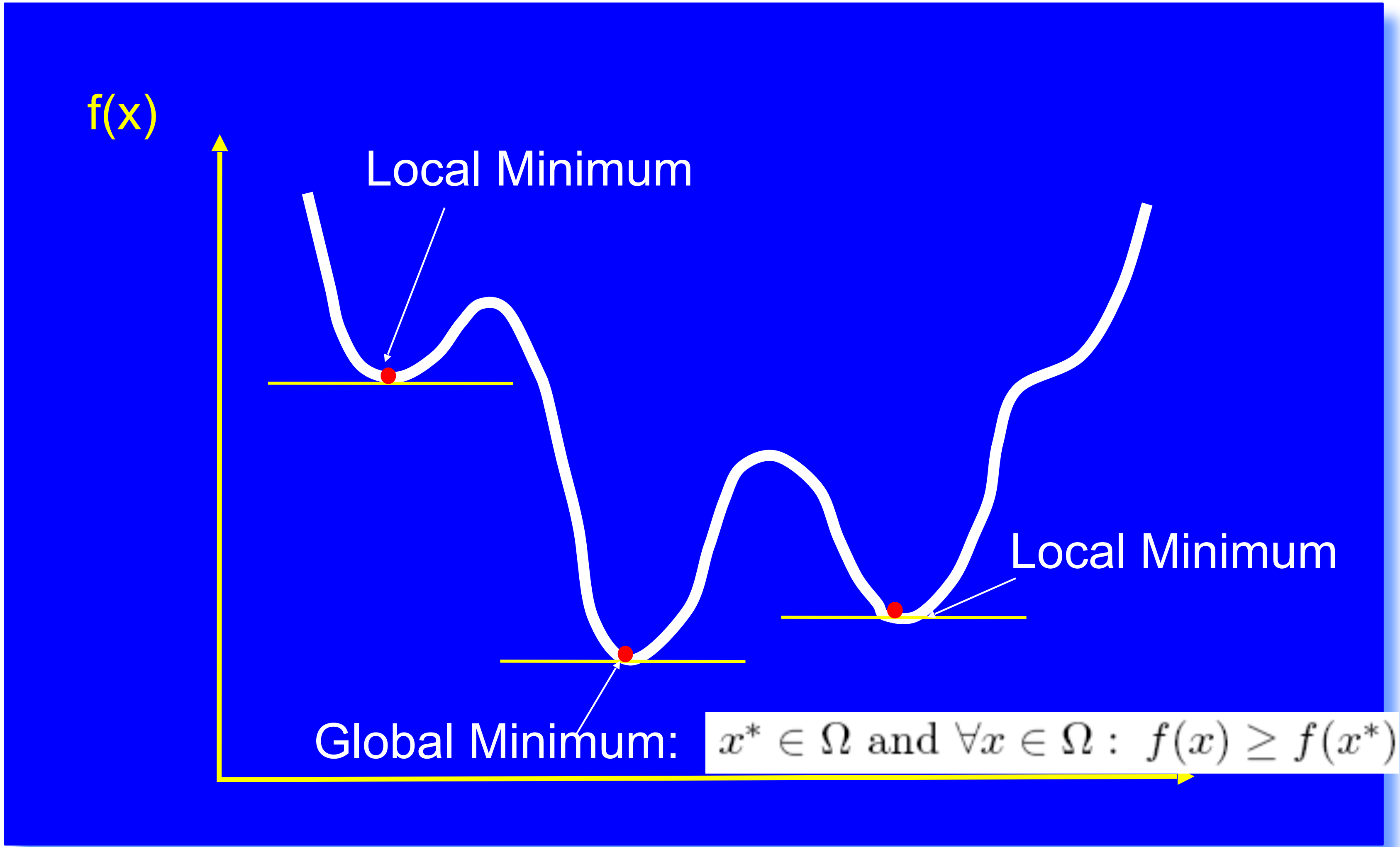
feasible set is intersection of grey and blue area

$$h_1(x) := \qquad x_2 \geq 0$$

$$h_2(x) := \quad 1 - x_1^2 - x_2^2 \geq 0$$

The "feasible set" Ω is $\{x \in \mathbb{R}^n | g(x) = 0,\ h(x) \geq 0\}$.

Local and global optima



The point $x^* \in \mathbb{R}^n$ is a "local minimizer" iff $x^* \in \Omega$ and there exists a neighborhood $\mathcal{N}$ of x^* (e.g. an open ball around x^*) so that $\forall x \in \Omega \cap \mathcal{N}: \; f(x) \geq f(x^*)$.

Derivatives

- First and second derivatives of the objective function or the constraints play an important role in optimization

- The first order derivatives are called the gradient (of the resp. fct)

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)^T$$

- and the second order derivatives are called the Hessian matrix

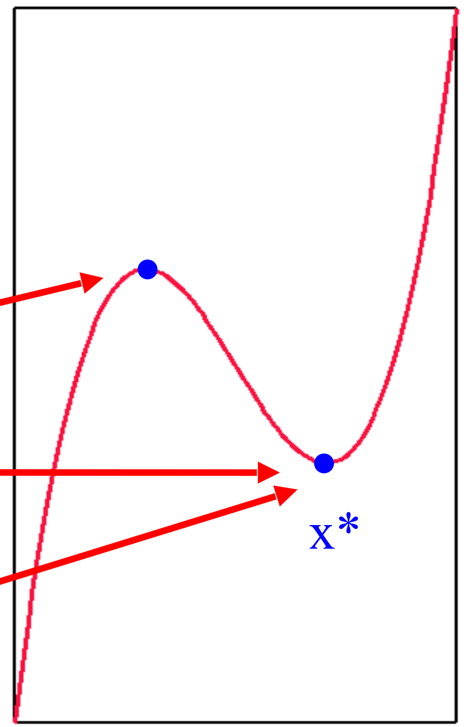
$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial {x_2}^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial {x_n}^2} \end{pmatrix}$$

Optimality conditions (unconstrained)

$$\min f(x) \quad x \in R^n$$

Assume that f is twice differentiable.
We want to test a point *x** for local optimality.

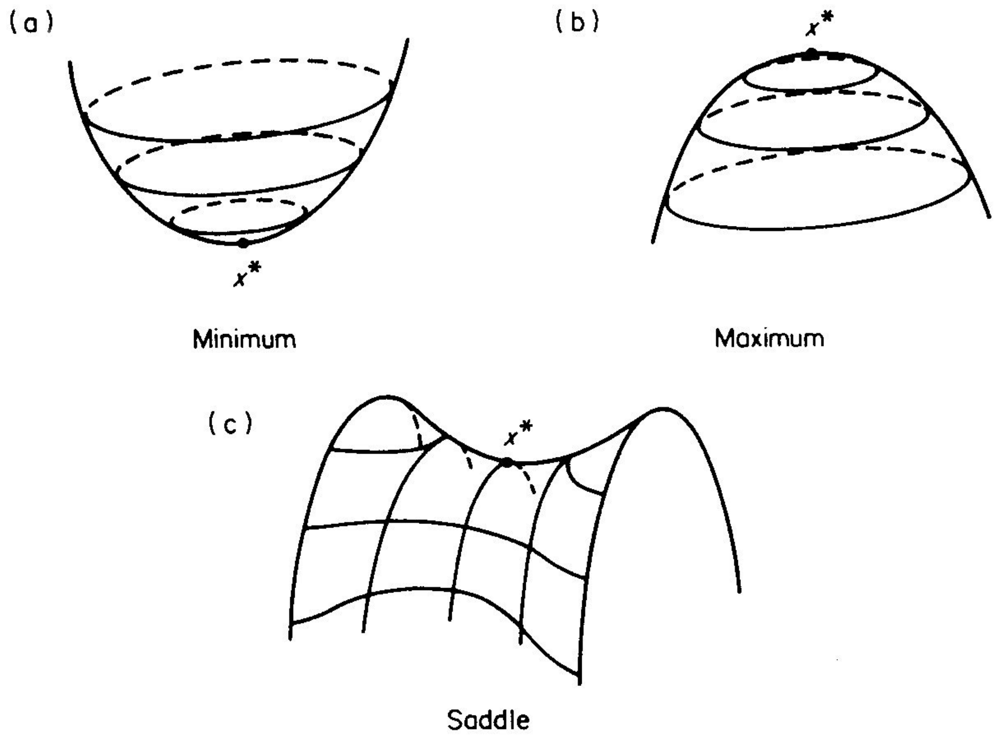
- *necessary condition:*
  $\nabla f(x^*)=0$ (*stationarity*)
- *sufficient condition:*
  *x** stationary and $\nabla^2 f(x^*)$ positive definite



Types of stationary points

(a)-(c) x^* is stationary: $\nabla f(x^*)=0$

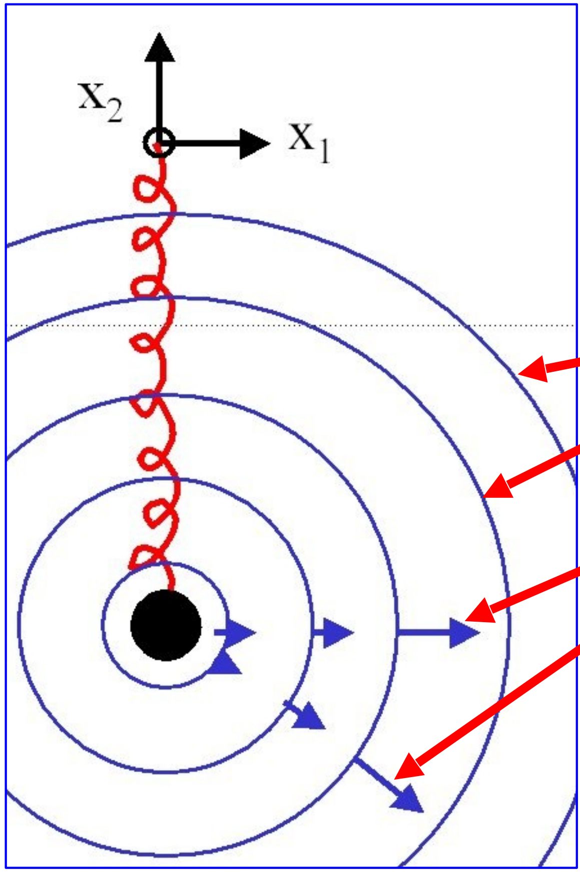
$\nabla^2 f(x^*)$ positive definite: local minimum



$\nabla^2 f(x^*)$ negative definite: local maximum

$\nabla^2 f(x^*)$ indefinite: saddle point

Ball on a spring without constraints



$$\min_{x \in R^2} x_1^2 + x_2^2 + m x_2$$

contour lines of *f(x)*

gradient vector

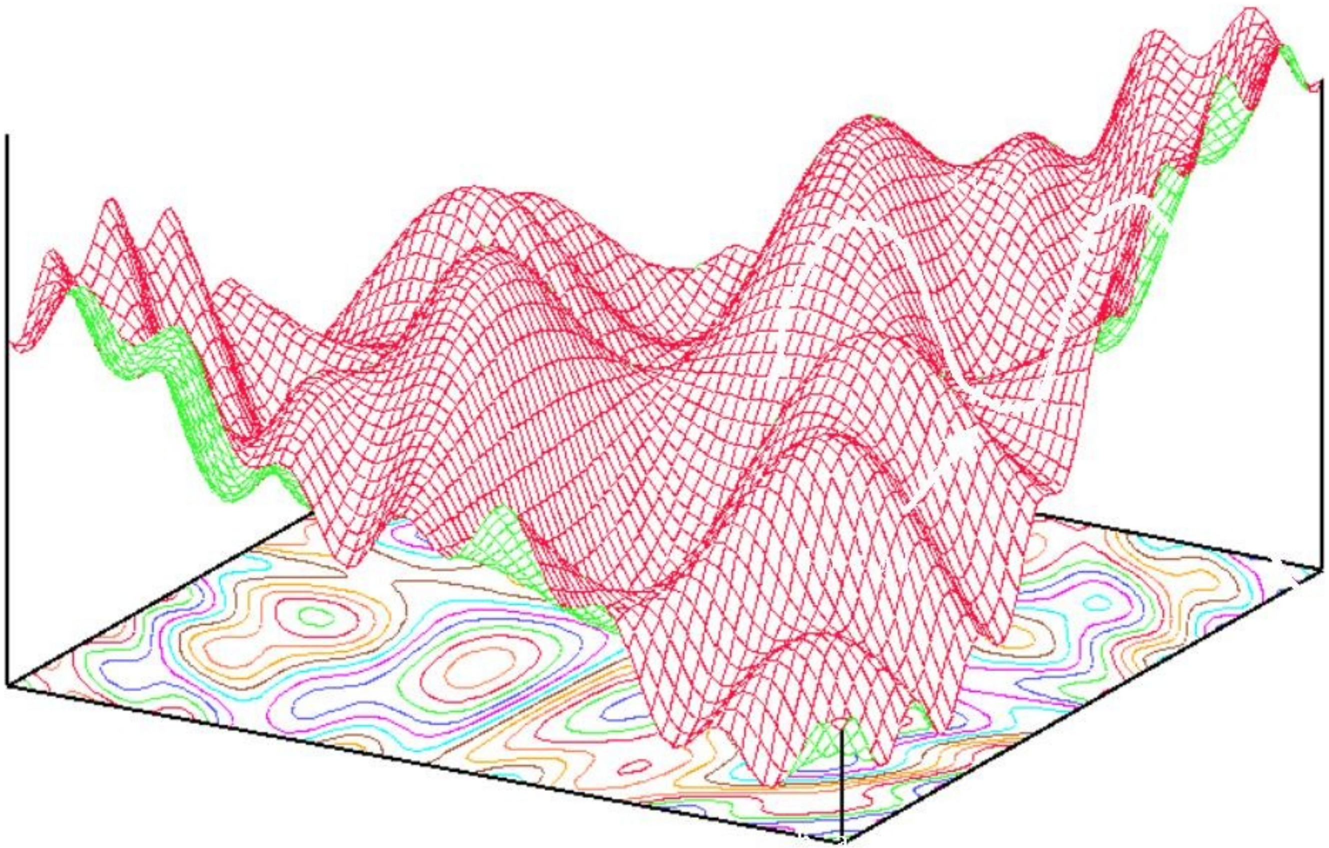
$$\nabla f(x) = (2x_1, 2x_2 + m)$$

unconstrained minimum:

$$0 = \nabla f(x^*) \Leftrightarrow (x_1^*, x_2^*) = (0, -\frac{m}{2})$$

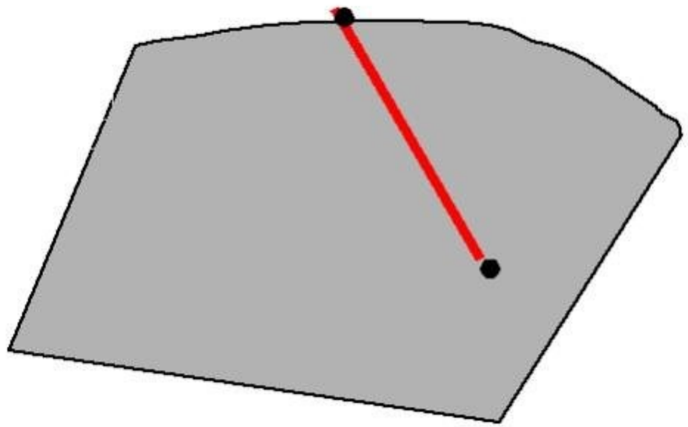
Sometimes there are many local minima

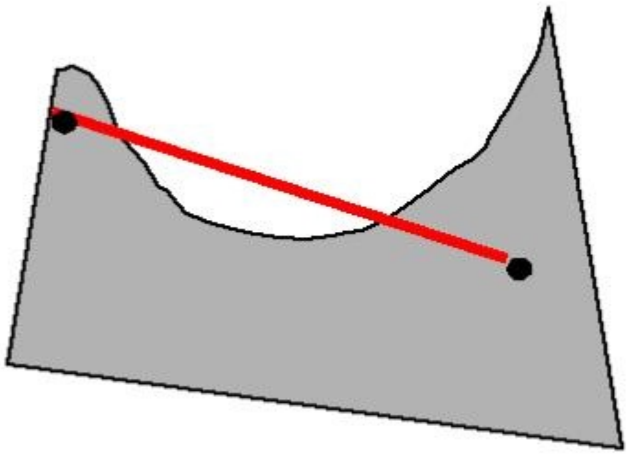
e.g. potential energy
of macromolecule



Global optimization is a very hard issue - most algorithms find only the next local minimum. But there is a favourable special case...

Convex feasible sets





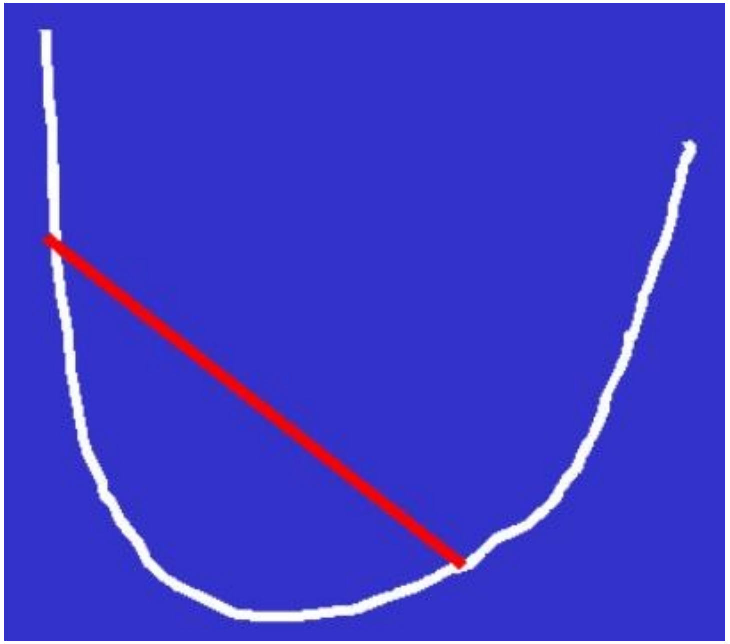
Convex: all connecting lines between feasible points are in the feasible set

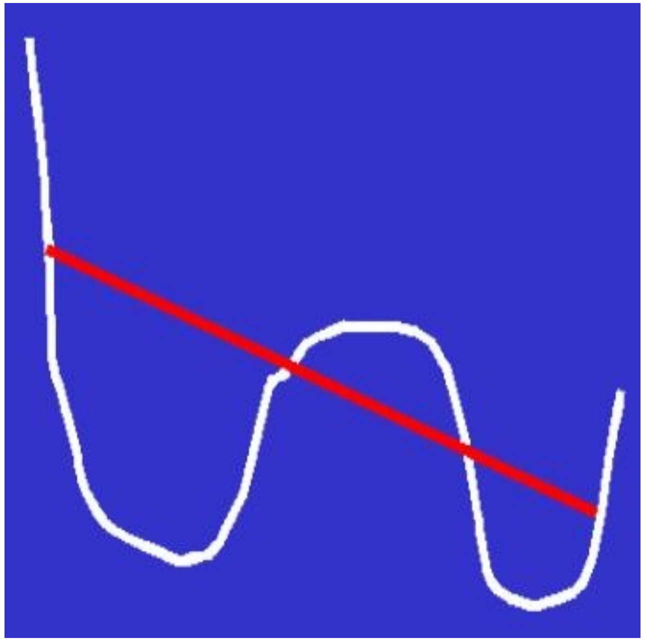
Non-convex: some connecting line between two feasible points is not in the feasible set

A set $\Omega \subset \mathbb{R}^n$ *is convex if*

$$\forall x, y \in \Omega, t \in [0,1] : \; x + t(y - x) \in \Omega.$$

Convex functions





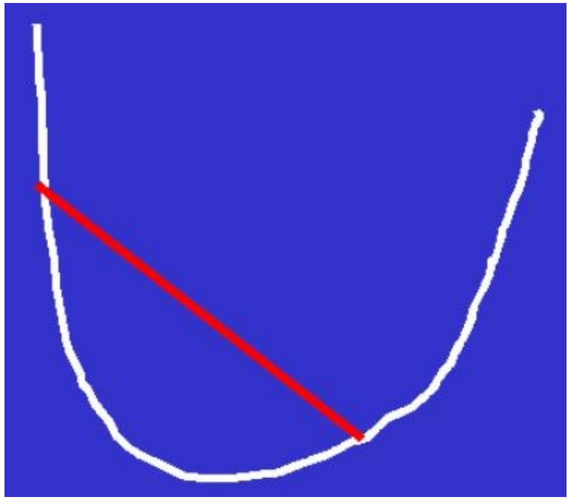
Convex: all connecting lines are above graph

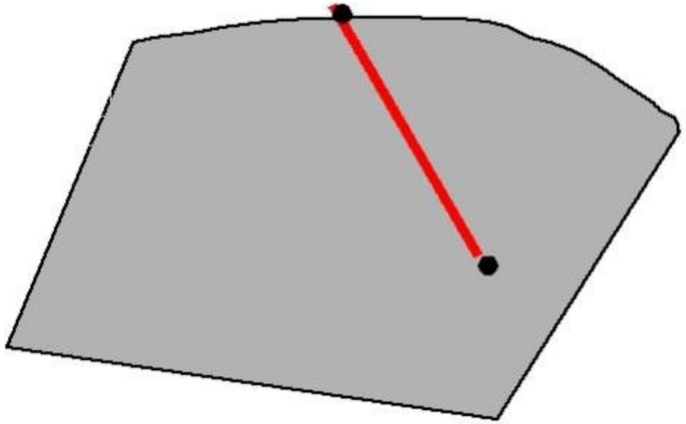
Non-convex: some connecting lines are not above graph

(Convex Function) *A function* $f : \Omega \to \mathbb{R}$ *is convex, if* Ω *is convex and if*

$$\forall x, y \in \Omega, t \in [0,1] : \; f(x + t(y - x)) \leq f(x) + t(f(y) - f(x)).$$

Convex problems





Convex problem if

f(x) is convex **<u>and</u>** the feasible set is convex

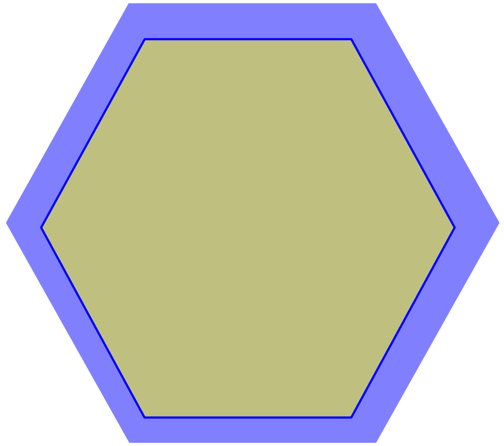
One can show:

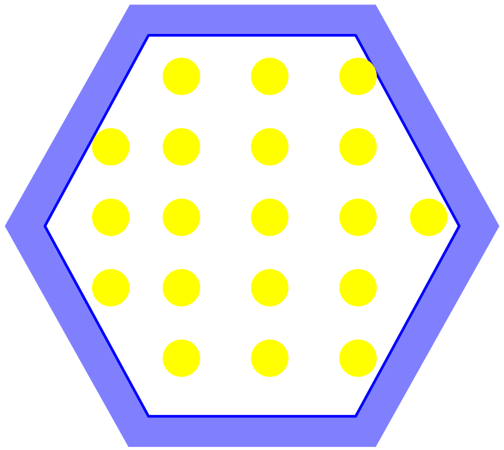
For convex problems, every local minimum is also a global minimum.

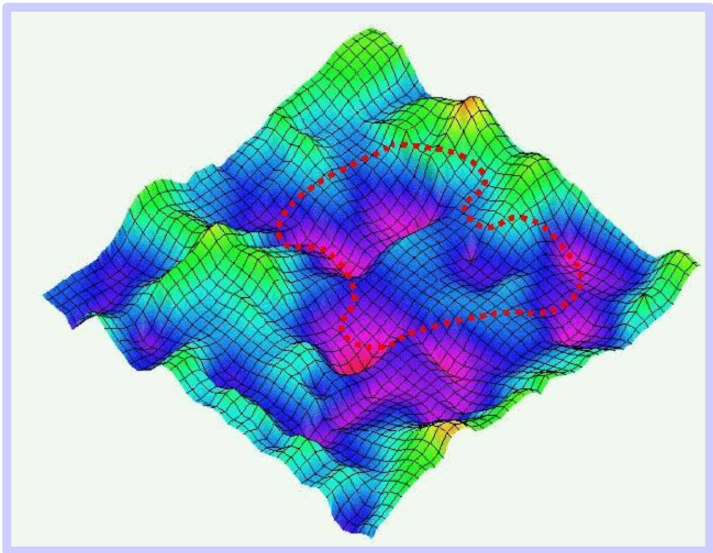
It is sufficient to find local minima!

Characteristics of optimization problems 1

- size / dimension of problem n ,
  i.e. number of free variables
- continuous or discrete search space
- number of minima





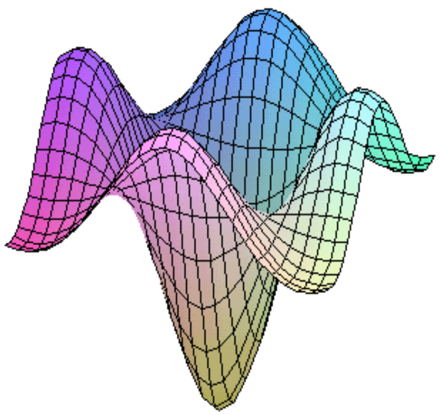


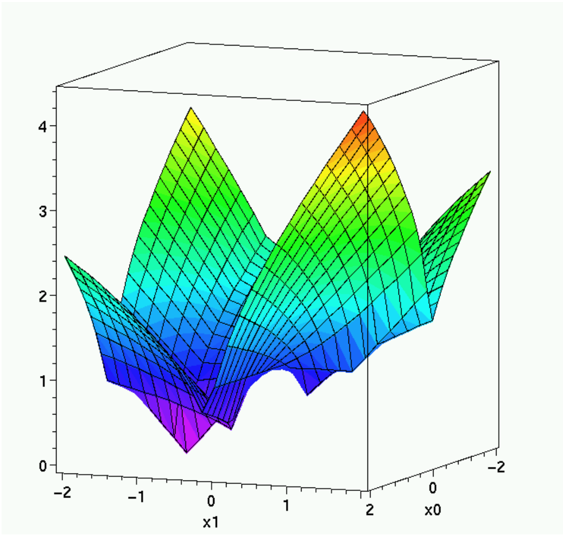
Characteristics of optimization problems 2

- Properties of the objective function:
  - type: linear, nonlinear, quadratic ...
  - smoothness: continuity, differentiability

- Existence of constraints

- Properties of constraints:
  - equalities / inequalities
  - type: „simple bounds“, linear, nonlinear, dynamic equations (optimal control)
  - smoothness





Overview of presentation

- Optimization: basic definitions and concepts
- **Introduction to classes of optimization problems**

Problem Class 1: Linear Programming (LP)

- Linear objective,
  linear constraints:
  Linear Optimization Problem
  (convex)

$$
\begin{aligned}
\min_{x} \quad & c^T x \\
\text{s. t.} \quad & Ax = b \\
& x \geq 0
\end{aligned}
$$

- Example: **Logistics Problem**
  - shipment of quantities $a_1, a_2, \dots a_m$ of a product from m locations
  - to be received at n destinations in quantities $b_1, b_2, \dots b_n$
  - shipping costs c_{ij}
  - determine amounts x_{ij}

⇨ Origin of linear programming in 2nd world war

Problem Class 2: Quadratic Programming (QP)

- Quadratic objective and linear constraints:
  Quadratic Optimization Problem (convex, if Q pos. def.)

$$
\begin{aligned}
\min_{x} \quad & c^T x + \frac{1}{2} x^T Q x \\
\text{s. t.} \quad & Ax = b \\
& Cx \geq d
\end{aligned}
$$

- Example: Markovitz mean variance portfolio optimization
  - quadratic objective: portfolio variance (sum of the variances and covariances of individual securities)
  - linear constraints specify a lower bound for portfolio return
- QPs are at the core of **Linear Model Predictive Control (MPC)**
- QPs play an important role as **subproblems in nonlinear optimization (and Nonlinear MPC)**

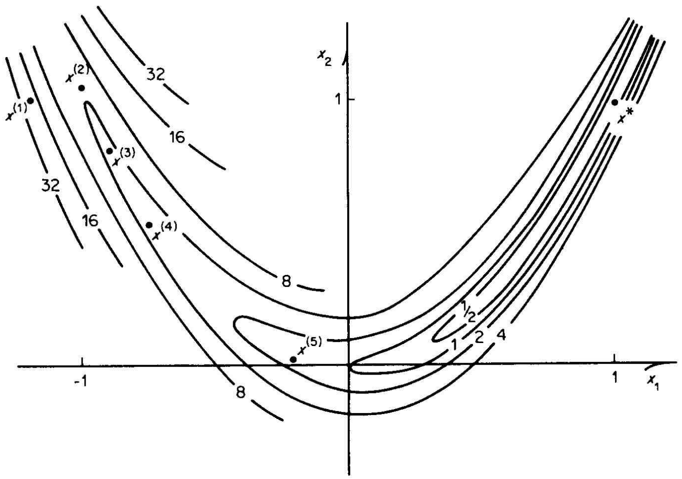
Problem Class 3: Nonlinear Programming (NLP)

- Nonlinear Optimization Problem (smooth, but in general nonconvex)

$$
\begin{array}{ll}
\min\limits_{x} & f(x) \\
\text{s. t.} & h(x) = 0 \\
 & g(x) \geq 0
\end{array}
$$

- E.g. the famous nonlinear Rosenbrock function

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$



Problem Class 4: Non-smooth optimization

- objective function or constraints are non-differentiable or not continuous e.g.

$$f(x) = |x|$$

$$f(x) = \max_i f_i(x), \quad i = 1, ..n$$

$$f(x) = \begin{cases} \cos x & \text{für } x \leq \frac{\pi}{2} \\ 0 & \text{für } x > \frac{\pi}{2} \end{cases}$$

$$f(x) = i \quad \text{for} \quad i \leq x < i+1, \; i = 0, 1, 2, \ldots$$

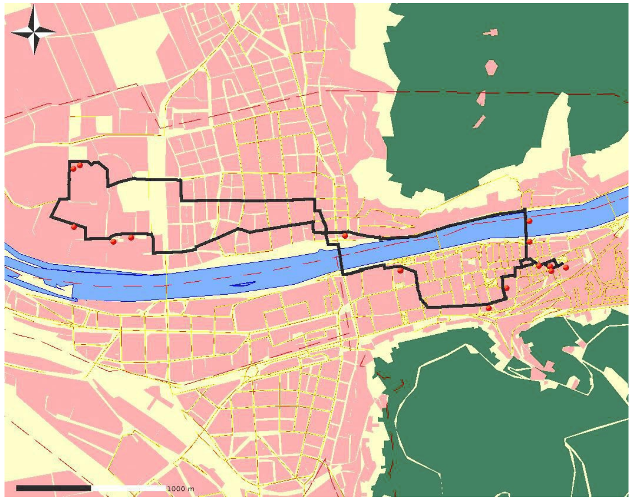
Problem Class 5: Integer Programming (IP)

- Some or all variables are integer (e.g. linear integer problems)

$$
\begin{array}{ll}
\min\limits_{x} & c^T x \\
\text{s. t.} & Ax = b \\
 & x \in Z_+^n
\end{array}
$$

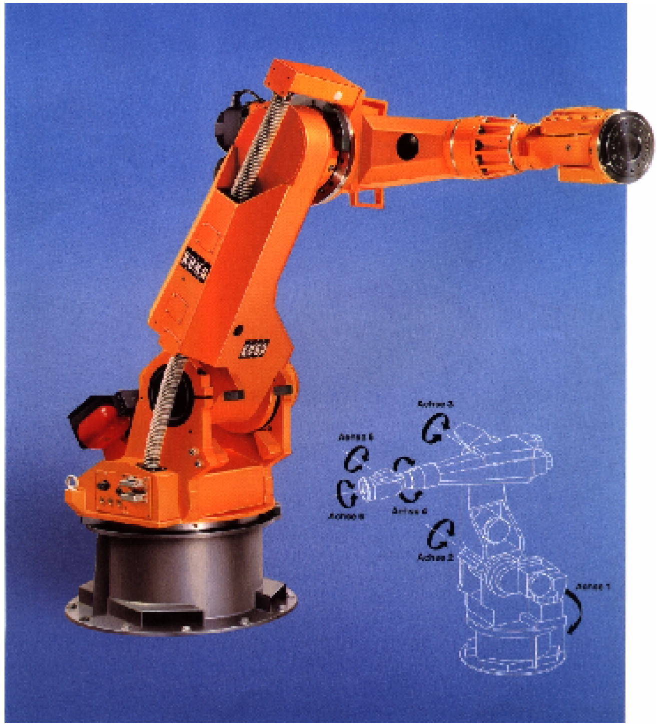
- Special case: combinatorial optimization problems -- feasible set is finite

- Example: traveling salesman problem
  - determine fastest/shortest round trip through n locations



Problem Class 6: Optimal Control

- Optimization problems including dynamics in form of **differential equations** (infinite dimensional)



Variables $x(t), u(t), p$ (partly ∞-dim.)

$$\min_{,x,u,p} \int_0^T \phi(t, x(t), u(t), p)dt$$

$$\text{s. t.} \quad \dot{x} = f(t, x(t), u(t), p)$$

$$\ldots.$$

THIS COURSE'S MAIN TOPIC!

Summary: Optimization Overview

Optimization problems can be:

- unconstrained or constrained
- convex or non-convex
- linear or non-linear
- differentiable or non-smooth
- continuous or integer or mixed-integer
- finite or infinite dimensional
- …

The great watershed

"The great watershed in optimization isn't
between linearity and nonlinearity,
but convexity and nonconvexity”

R. Tyrrell Rockafellar

- For convex optimization problems we can efficiently find global minima.
- For non-convex, but smooth problems we can efficiently find local minima.

Literature

- J. Nocedal, S. Wright: Numerical Optimization, Springer, 1999/2006
- P. E. Gill, W. Murray, M. H. Wright: Practical Optimization, Academic Press, 1981
- R. Fletcher, Practical Methods of Optimization, Wiley, 1987
- D. E. Luenberger: Linear and Nonlinear Programming, Addison Wesley, 1984
- S. Boyd, L. Vandenberghe: Convex Optimization, Cambridge University Press, 2004 (PDF freely available at:

  http://web.stanford.edu/~boyd/cvxbook/
- M. Diehl: Script on Optimal Control and Estimation:

  http://syscop.de/teaching/numerical-optimal-control/