How mature is Nonlinear Optimization?

Philippe Toint
(University of Namur, Belgium)
(with N. Gould)

Sydney, ICIAM, June 2003

Outline

- What is nonlinear optimization?
- A field no longer in infancy
- Some signs of maturity
  - an adequate theoretical view
  - improved software testing
  - a world of applications
- Is senility lurking?
  - impact of interior point methods
  - (revival of derivative free methods)
  - filter methods
- The future's challenge

The nonlinear optimization problem

The general nonlinear programming problem:

$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & c_{\mathcal{E}}(x) = 0 \\ & c_{\mathcal{I}}(x) \geq 0 \end{array}$$

for $x \in \mathrm{IR}^n$, f, $c_{\mathcal{E}}$ and $c_{\mathcal{I}}$ smooth.

$\Rightarrow$ iterative numerical algorithms

Simplified because

- smoothness
- convexity
- special cases (LP, QP, . . .)

Why are we interested?

- The world is nonlinear

- Nature and people like to optimize
  - energy or material ressources
  - costs
  - risks
  - time
  - ...

  descriptive modelling no longer sufficient!

- Intellectually challenging

A field no longer in infancy

- respectable history (over 200 years)

- emphasis more on "long term" issues (local vs. global convergence)

- internally structured (journals) . . .

- . . . but also conscious of the world
  - applications
  - software

- no longer interested in "toys"

Some signs of maturity

An adequate theoretical understanding:

- theory viewed as a necessary condition for good algorithms
- confluence of paradigms: trust-region and linesearch methods
- better understanding of the convex case: self-concordance and Newton's method

Some signs of maturity (cont'd)

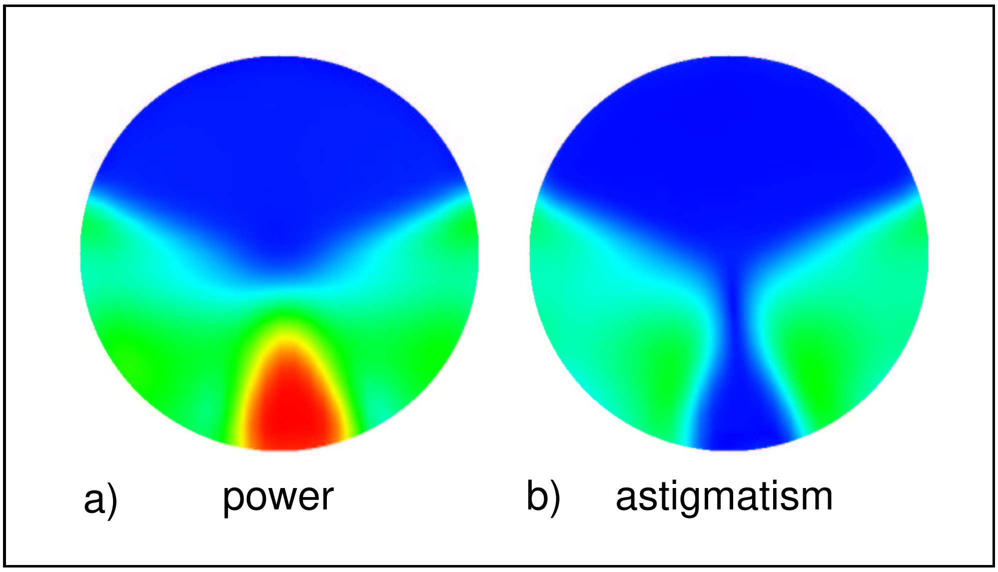
Improved software testing:

- increasing recognition
- better modelling languages
  (SIF(?), GAMS, AMPL)
- growing test cases collections
  (Argonne, HS, CUTEr, COPS, . . .)
  + independent testing
- improved reporting methodology
  (performance profiles)

Applications: PAL design (1)

Design of modern Progressive Adaptive Lenses:

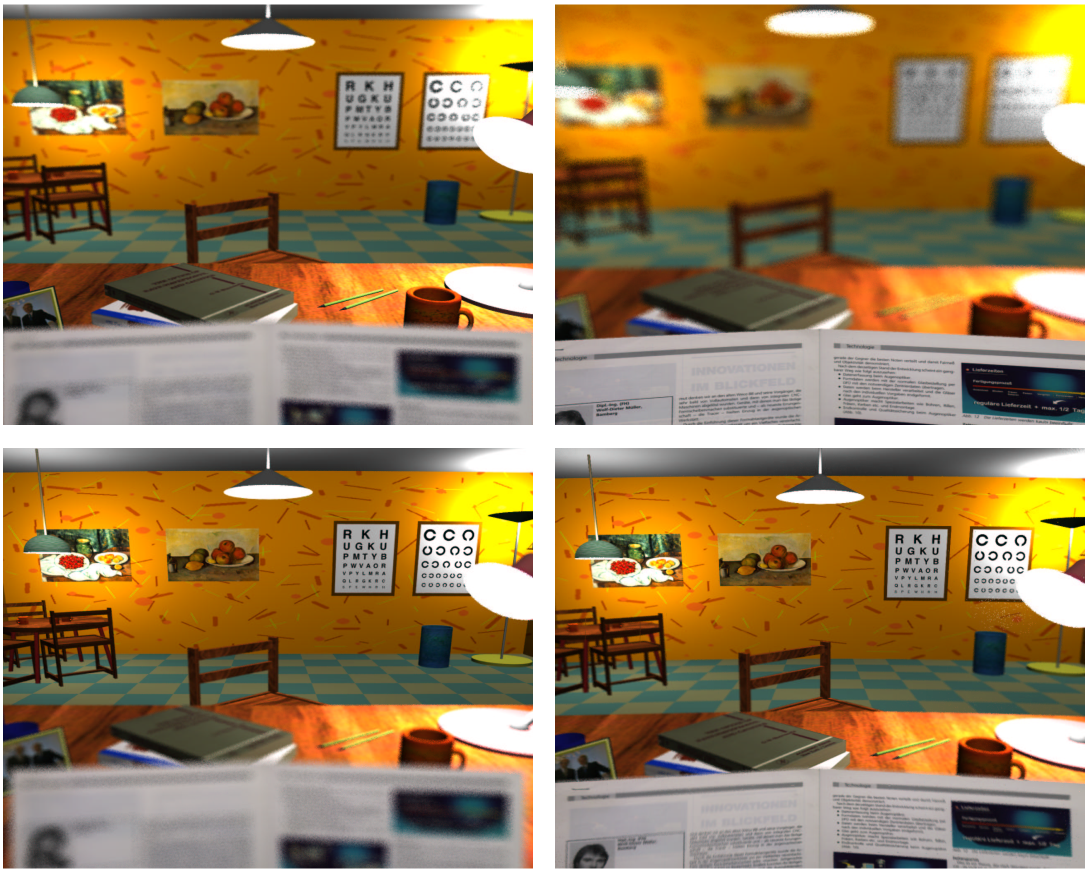
vary optical power of lenses
while minimizing astigmatism



(source: Loos, Greiner, Seidel 1997)

Applications: PAL design (2)

Achievements:



uncorrected long distance	short distance PAL

(source: Loos, Greiner, Seidel 1997)

Applications: PAL design (3)

Is this nonlinear?

Assume the lens surface is $z = z(x, y)$.
The optical power is

$$p(x,y) = \frac{N^3}{2}\left[\left(1+\left[\frac{\partial z}{\partial x}\right]^2\right)\frac{\partial^2 z}{\partial y^2}\right.$$

$$\left. + \left(1+\left[\frac{\partial z}{\partial y}\right]^2\right)\frac{\partial^2 z}{\partial x^2} - 2\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}\frac{\partial^2 z}{\partial x \partial y}\right],$$

where

$$N = N(x,y) = \frac{1}{\sqrt{1+\left[\frac{\partial z}{\partial x}\right]^2+\left[\frac{\partial z}{\partial y}\right]^2}}.$$

The surface astigmatism is then

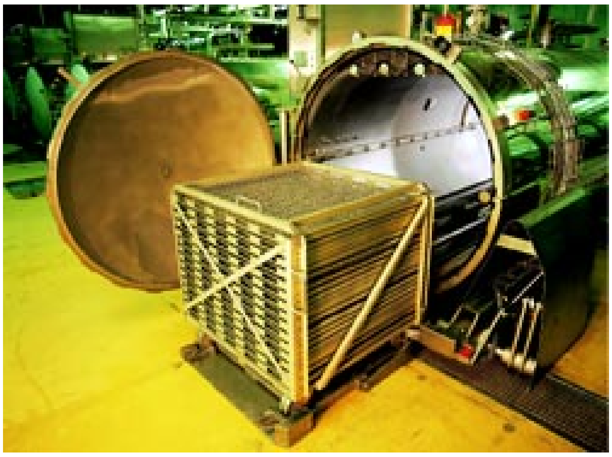
$$a(x,y) = -2\sqrt{p(x,y) - N^4\left(\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} - \left[\frac{\partial^2 z}{\partial x \partial y}\right]^2\right)}$$

Applications: Food sterilization (1)

A common problem in the food processing industry:

keep the vitamins while killing
a prescribed fraction of the bacteria

heating in steam/hot water autoclaves



(source: Sachs, 2003)

Applications: Food sterilization (2)

Model: coupled PDEs

Concentration of micro-organisms and other nutrients:

$$\frac{\partial C}{\partial t}(x,t) = -K[\theta(x,t)]C(x,t),$$

where $\theta(x,t)$ is the temperature, and where

$$K[\theta] = K_1 e^{-K_2\left(\frac{1}{\theta}-\frac{1}{\theta_r}\right)}$$

(Arrhenius equation).

Evolution of temperature:

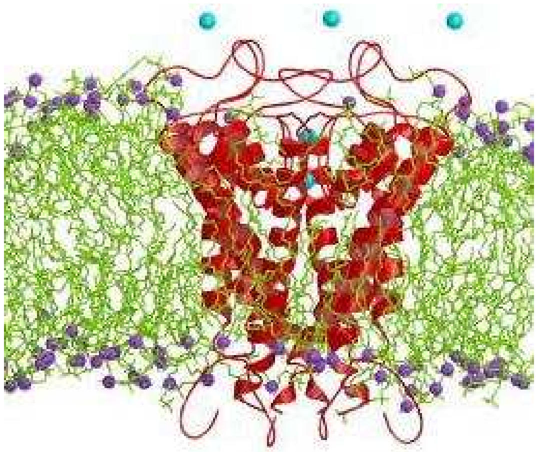
$$\rho c(\theta)\frac{\partial \theta}{\partial t} = \nabla \cdot [k(\theta)\nabla\theta],$$

(with suitable boundary conditions)

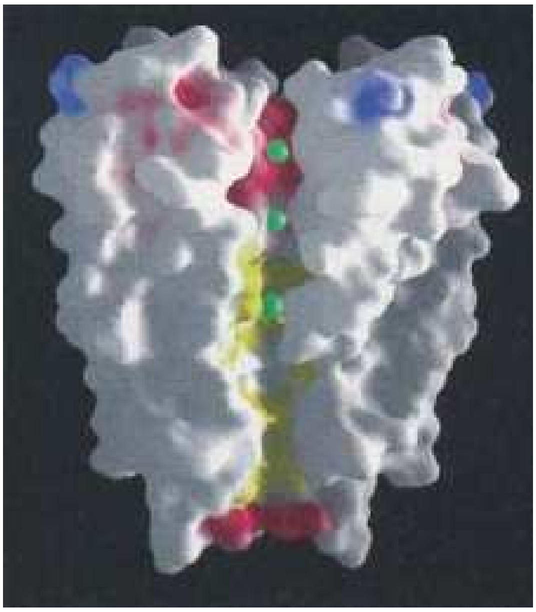
Symmetry $\Rightarrow$ 2D problem.

Applications: biological parameter estimation

K-channel in a the model of a neuron membrane:



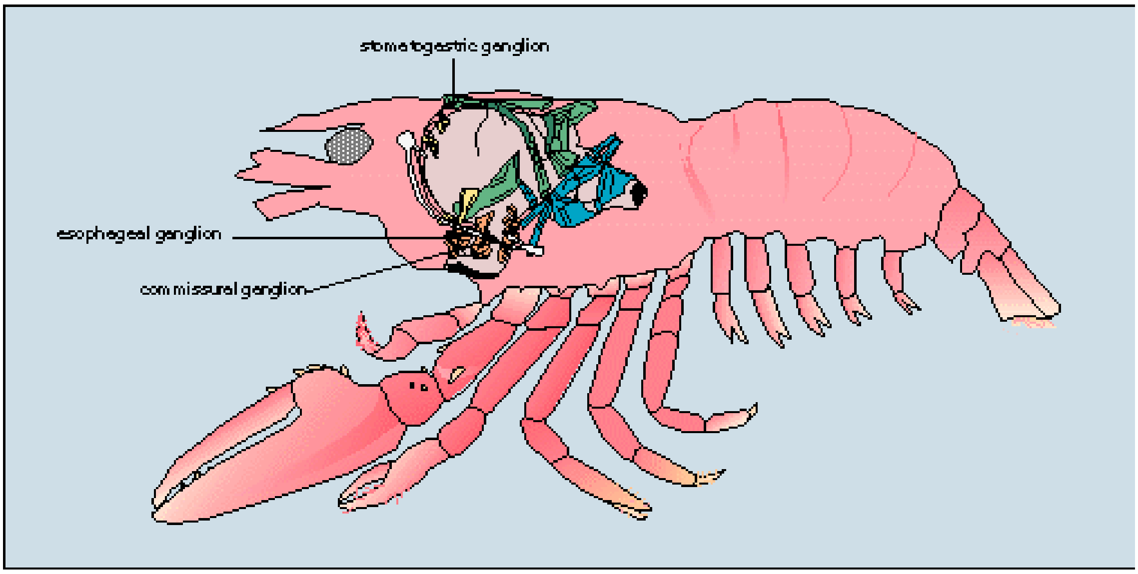
(source: Sansom, 2001)



(source: Doyle *et al.*, 1998)

Applications: biological parameter estimation

Where are these neurons?



(source: Simmers, Meyrand and Moulins, 1995)

in a Pacific spiny lobster!

Applications: biological parameter estimation

After gathering experimental data
(applying a current to the cell):

estimate the biological model
parameters that best fit experiments

Model:

- Activation: p independent gates
- Deactivation: n_h gates with different dynamics
- n_h+2 coupled ODEs for the voltage, the activation level, the partial inactivations levels
- 5-points BDF for $\approx$ 50000 time steps

Applications: shock simulation in video games

Formulate the motion of multiple rigid bodies in space as a "complementarity problem" (Anitescu and Potra, 1996):

$$\nabla_q \Phi[q(t)]v(t) \geq 0$$

$$\Phi(q(t)) \geq 0$$

($q(t)$ = positions, $v(t) = \frac{dq}{dt}(t)$ = velocities)

$\Rightarrow$ system of inequalities and equalities

used in realtime for video animation

(demo thanks to Vivid Image and MathEngine)

Is senility lurking?

NO because. . .

- important challenges persist
  - problem size
  - problem nonlinearity
  - global optimization
  - nonlinear problems in discrete variables

- very active subfields
  (ex: interior point methods,
  derivative free optimization)

- new ideas (ex: filter)

(But . . . some "navel contemplation")

Interior point methods: a simple case

Consider

$$\min f(x) \text{ under the constraint that } x \geq 0$$

and build

$$\phi^{\log}(x,\mu) \stackrel{\text{def}}{=} f(x) - \mu\langle e, \log(x)\rangle = f(x) - \mu \sum_{i=1}^{n} \log(x_i)$$

using a barrier function

Under acceptable assumptions,

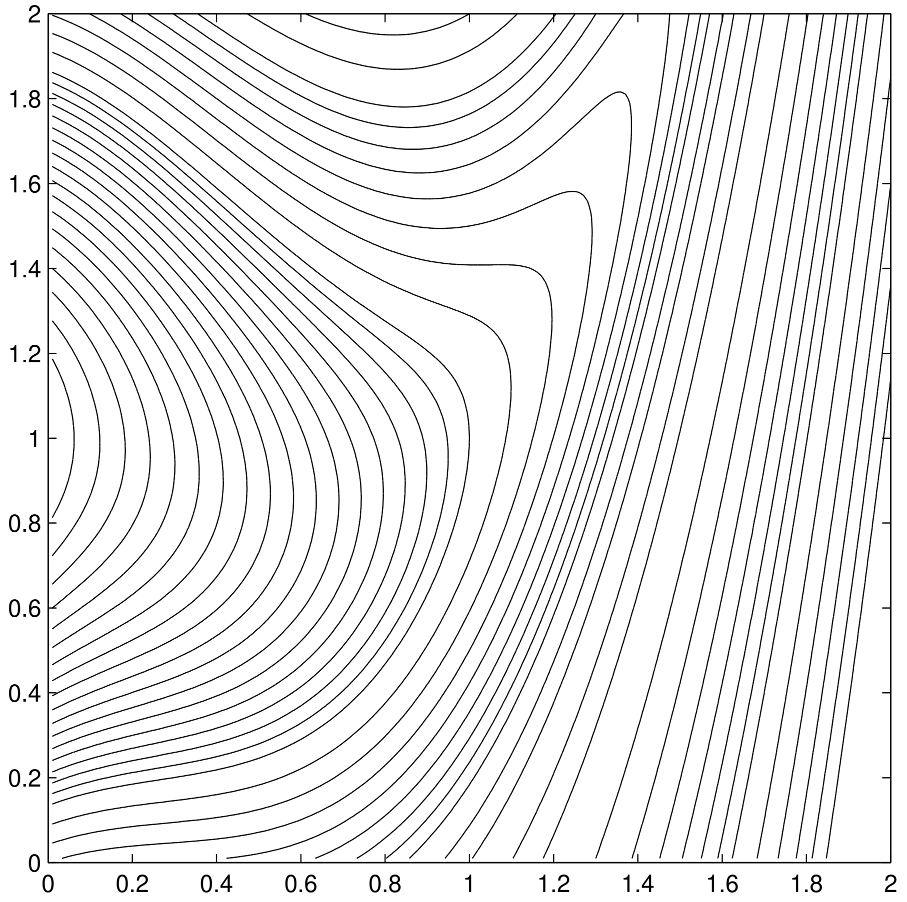
$$x_*(\mu) = \arg\min_x \phi^{\log}(x,\mu)$$

converge to the solution of the problem

$$\min_{x\in\mathcal{C}} f(x)$$

when $\mu \searrow 0$.

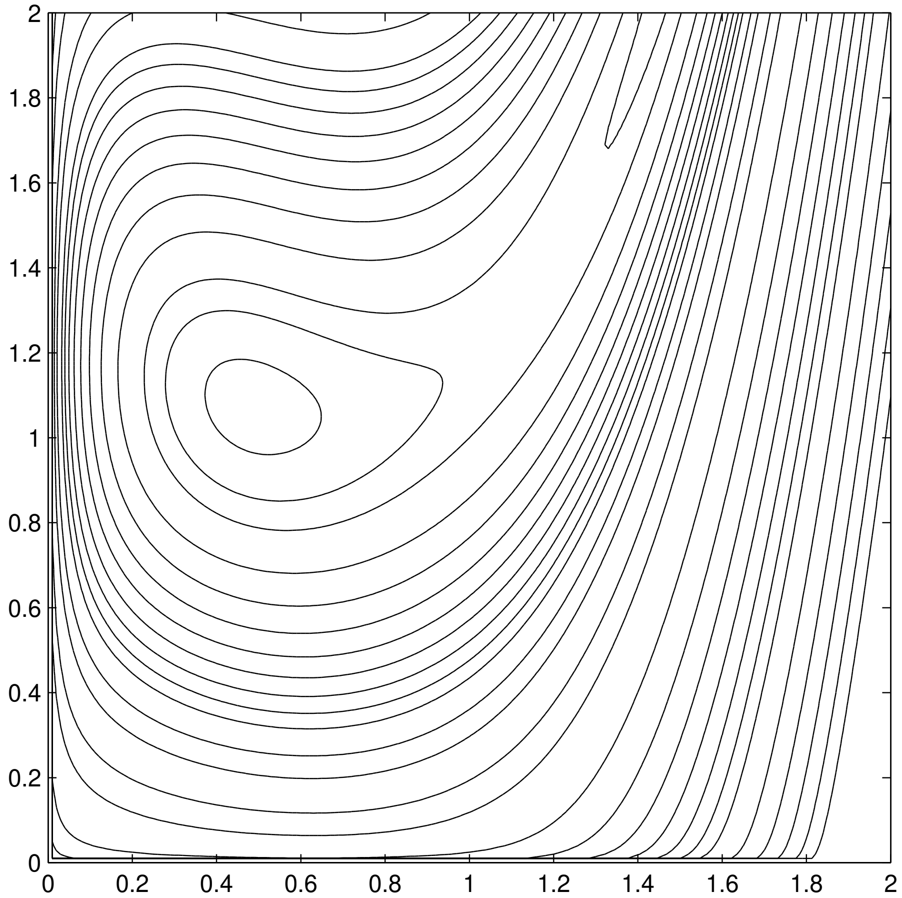
Interior point methods: how it works. . . (1)



Example:

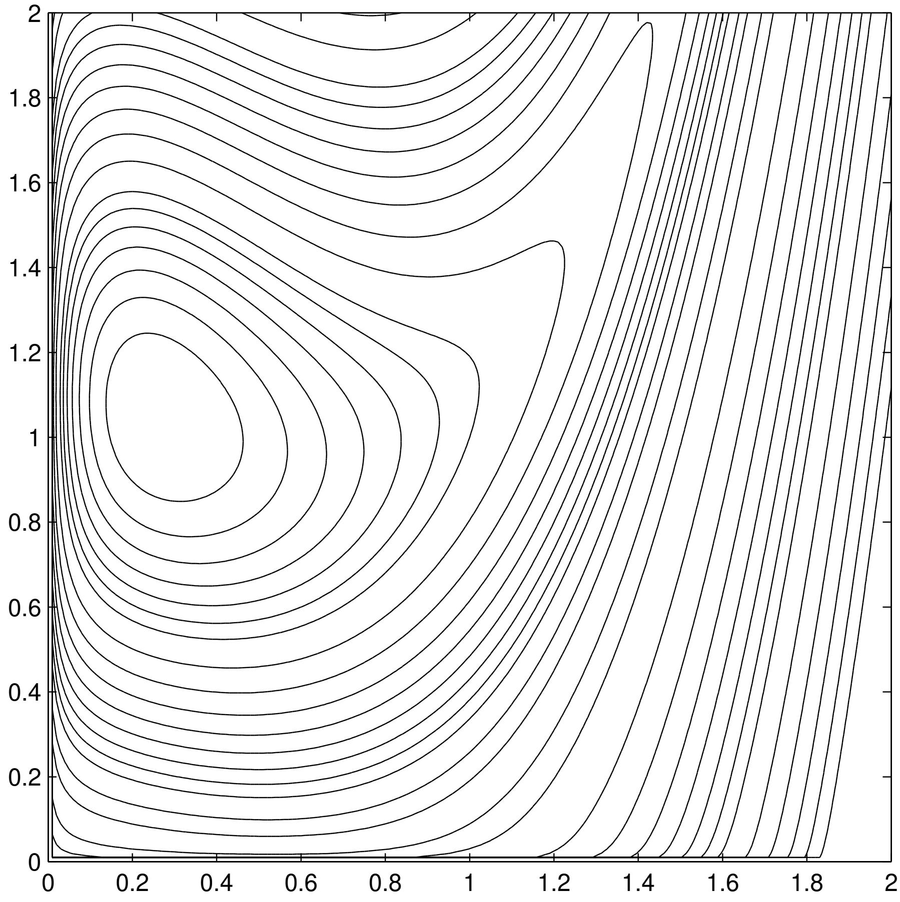
$$\min_{x_1, x_2 \geq 0} 120[x_1^2(x_1-1)-x_2+1]^2+10(4+x_1)^2-150.$$

Interior point methods: how it works. . . (2)



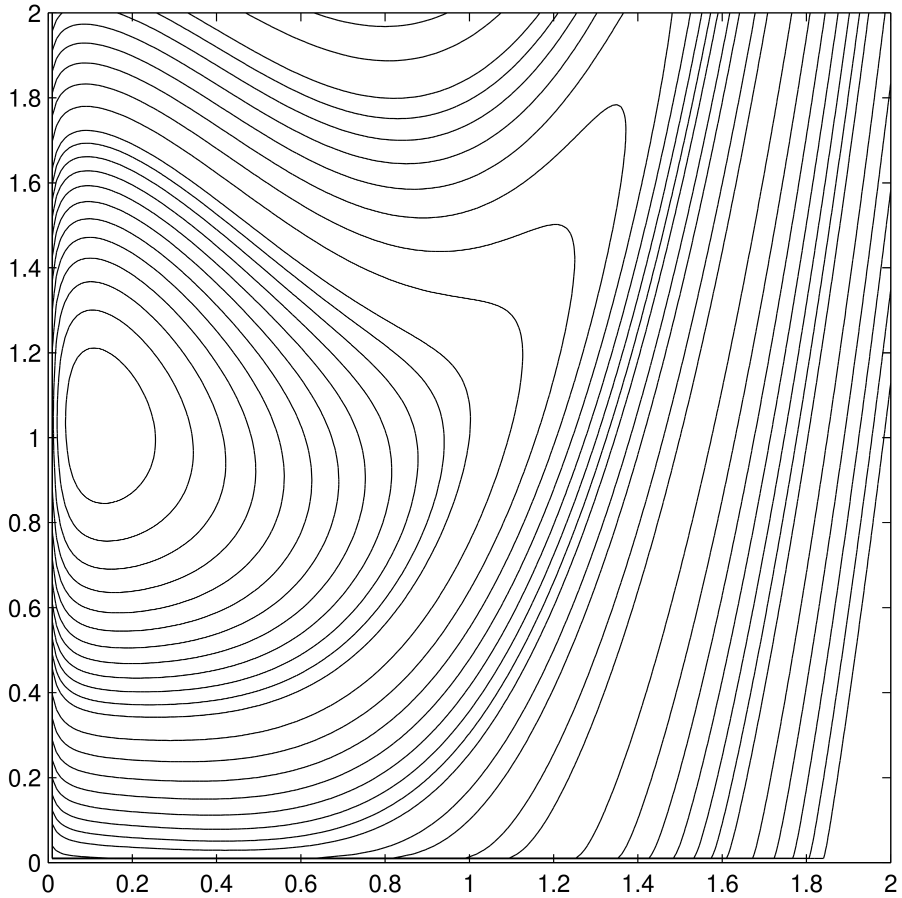
$\mu = 50$

Interior point methods: how it works. . . (3)



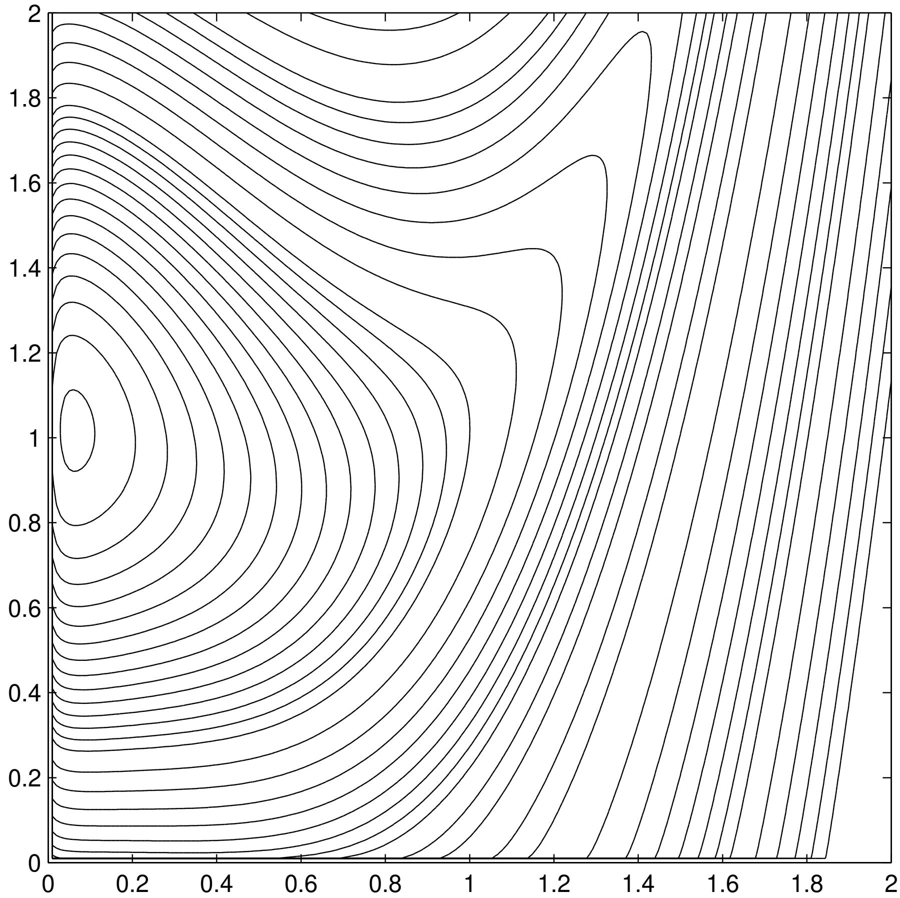
$$\mu = 25$$

Interior point methods: how it works. . . (4)



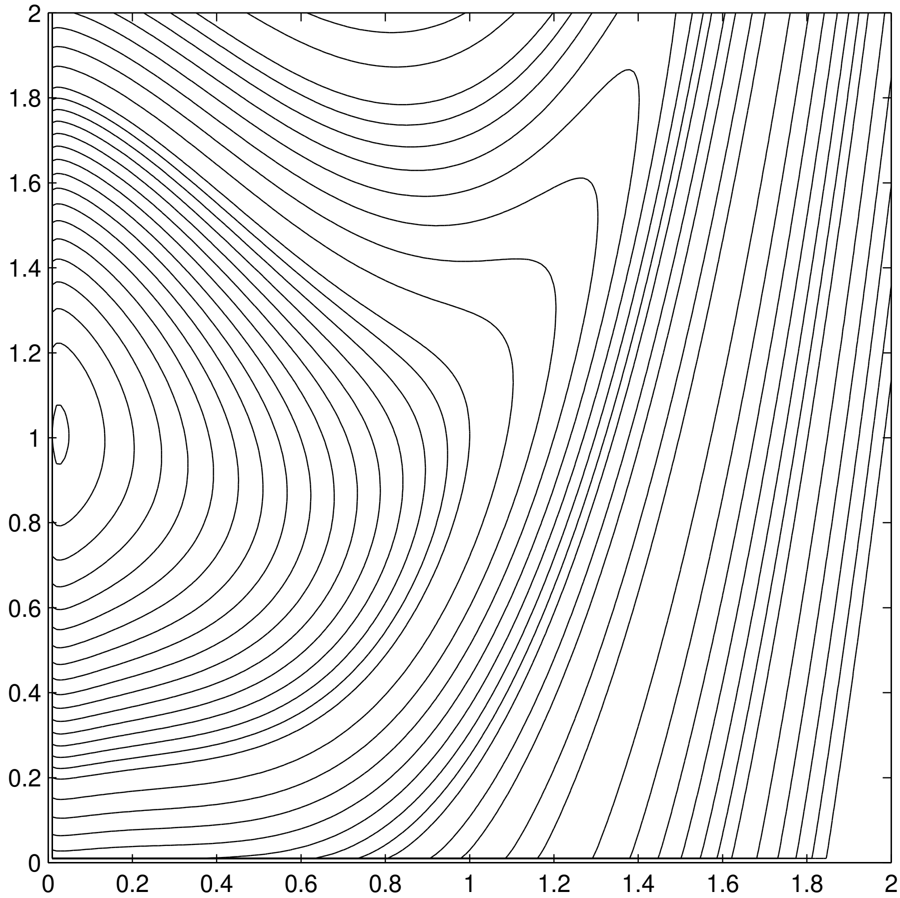
$\mu = 10$

Interior point methods: how it works. . . (5)



$$\mu = 5$$

Interior point methods: how it works. . . (6)



$$\mu = 2$$

The impact of interior point methods. . .

- massive impact in linear programming
- birth of semi-definite programming (matrices constrained to be positive semi-definite)
- well adapted and studied for convex problems
- under active development for nonconvex problems (KNITRO, LOQO, IPOPT, GALAHAD)

A new approach: filter methods

Conflicting objectives and dominance:
at the same time:

- reduce the objective function $f(x)$

- reduce constraint violation $\theta(x)$

$\Rightarrow$ conflict

Classical question: what is a better point?

Rephrase (Fletcher, Leyffer, 1997) as:

What is a worse point?

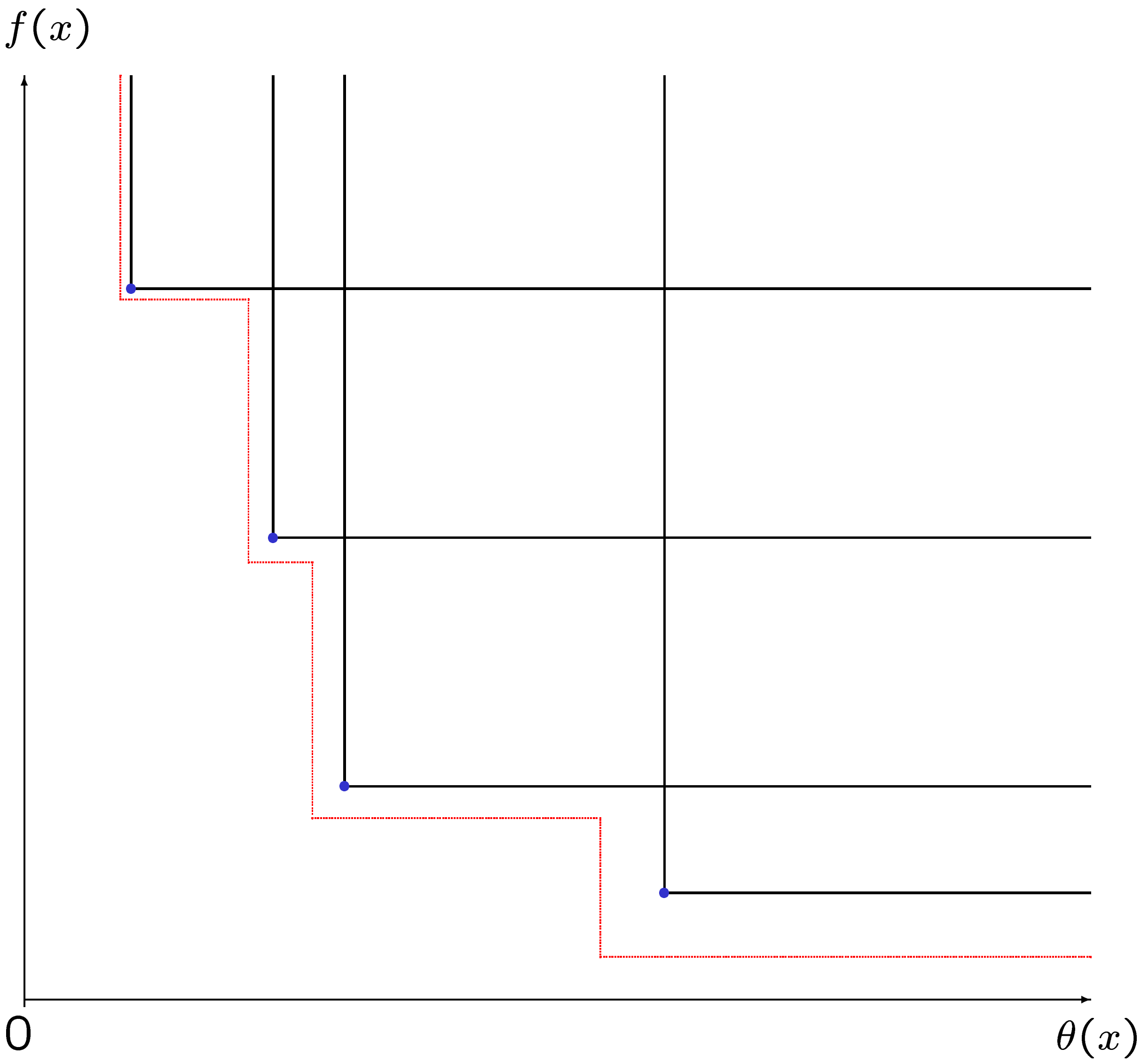
$$f(x) \leq f(y) \text{ and } \theta(x) \leq \theta(y)$$

(y dominated by x)

Filter algorithms

Idea: Build a numerical method where non-dominated points are accepted.

remember accepted points $\Rightarrow$ filter ($\mathcal{F}$)

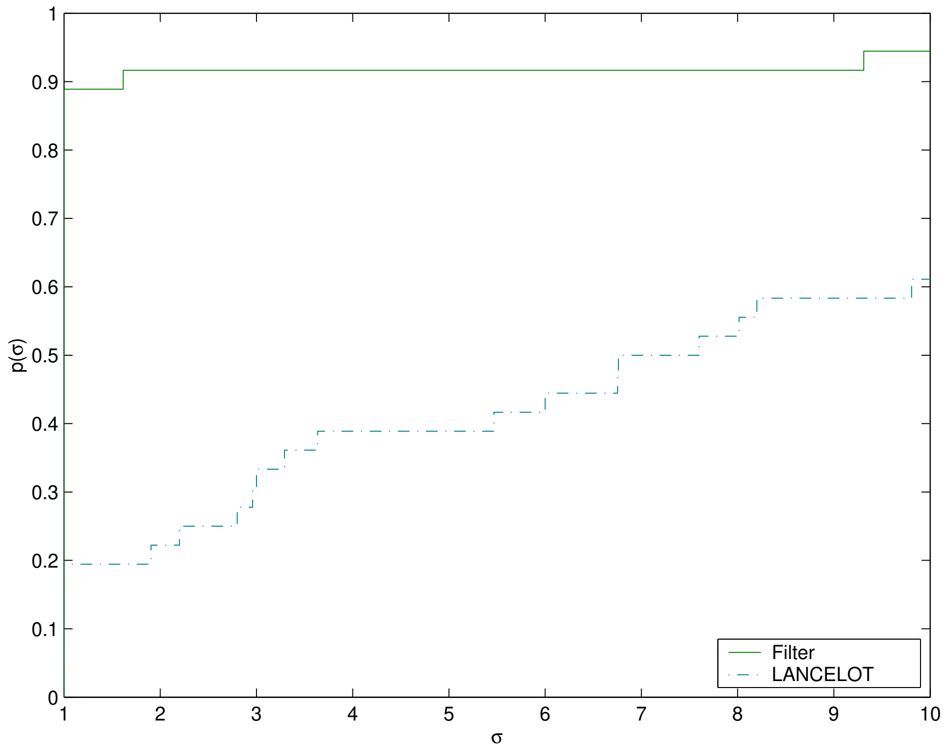


The efficiency of a filter method

A performance profile for comparing

$\Rightarrow$ FILTRANE vs. LANCELOT B $\Leftarrow$

for the nonlinear feasibility problem



time

(a knock-out!)

The future's challenges

- continuing improvement
  - algorithms
  - software

- further developments in specific areas:
  - MPECs
  - DAE constrained problems
  - mixed integer nonlinear programming
  - ...

***MUCH* remains to be done...**

Thank you for your attention