Institute for Mathematics and Its Applications

Mathematics of Inverse Problems

IMA Summer School June 15-July 3
University of Delaware, Newark Delaware
Week 3: Inverse Problems as Optimization Problems

Jonathan Borwein, FRSC www.carma.newcastle.edu.au/~jb616

Canada Research Chair in Collaborative Technology

Laureate Professor University of Newcastle, NSW

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

Revised 04-07-09

Our Goals for the Week

A brief introduction to some key ideas from optimization that should be useful later in your careers

Convex Analysis
- Duality and Optimality Conditions
- Fixed Points and Monotone Mappings

Variational Principles
- Stability and Regularity

Models and Algorithm Design
- Some Concrete Examples
- Some Experimentation

Mathematics is not a spectator subject

It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment. When I have clarified and exhausted a subject, then I turn away from it, in order to go into darkness again; the never-satisfied man is so strange if he has completed a structure, then it is not in order to dwell in it peacefully, but in order to begin another. I imagine the world conqueror must feel thus, who, after one kingdom is scarcely conquered, stretches out his arms for others.

(Carl Friedrich Gauss 1777-1855)

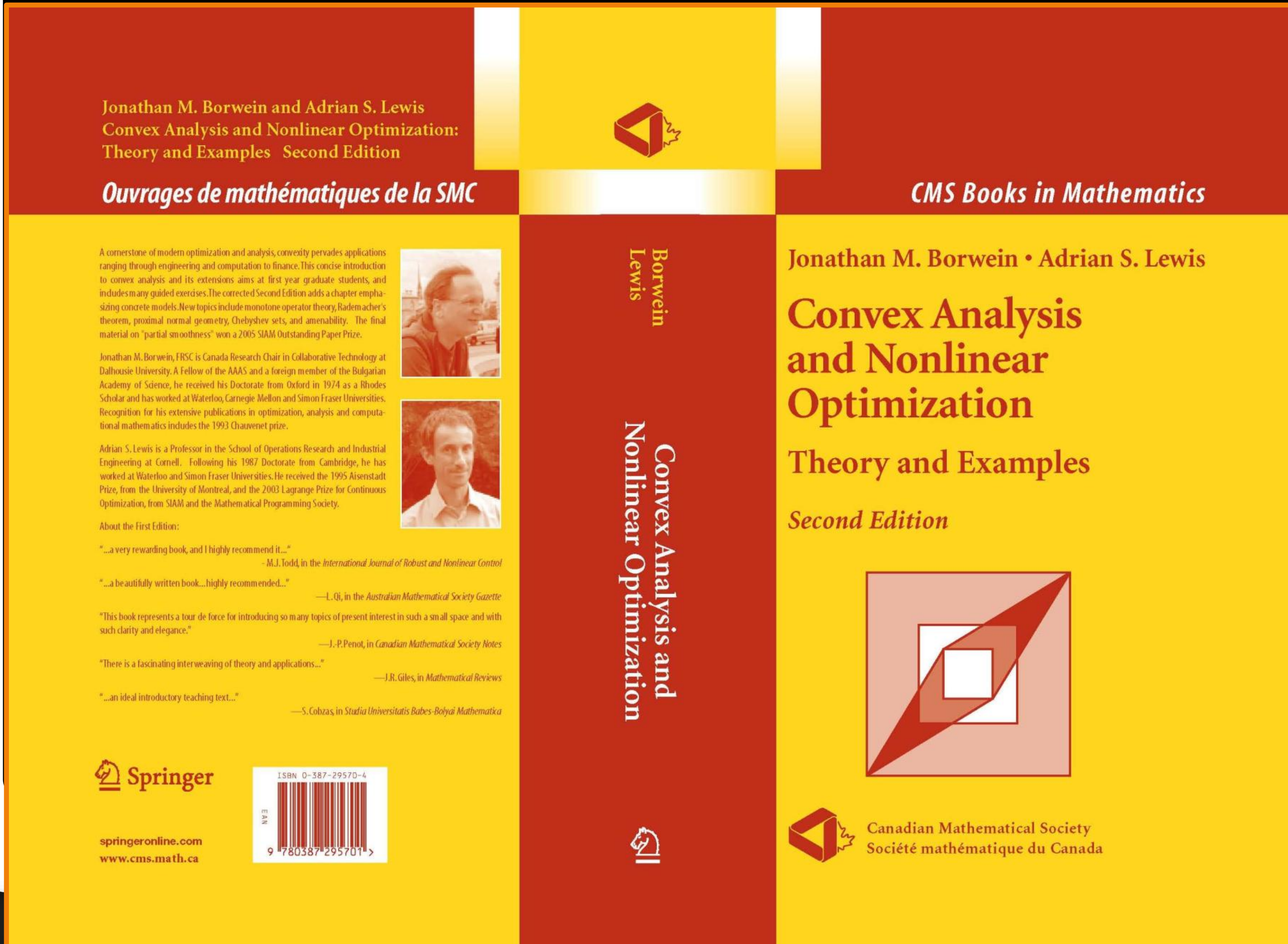
In an 1808 letter to his friend Farkas Bolyai
(the father of Janos Bolyai).

Outline of Week's Lectures

Day 1: A big-picture **Overview** of the Week
Day 2 Convex Duality and Applications
Day 3 Variational Principles & Applications
Day 4 Monotone & Non-expansive Maps
Day 5 Algebraic Reconstruction Methods and Interactive Geometry

Days will spill over...

The primary source is Chapters 3-5, 7-8 of



Contents

Assuring a minimum

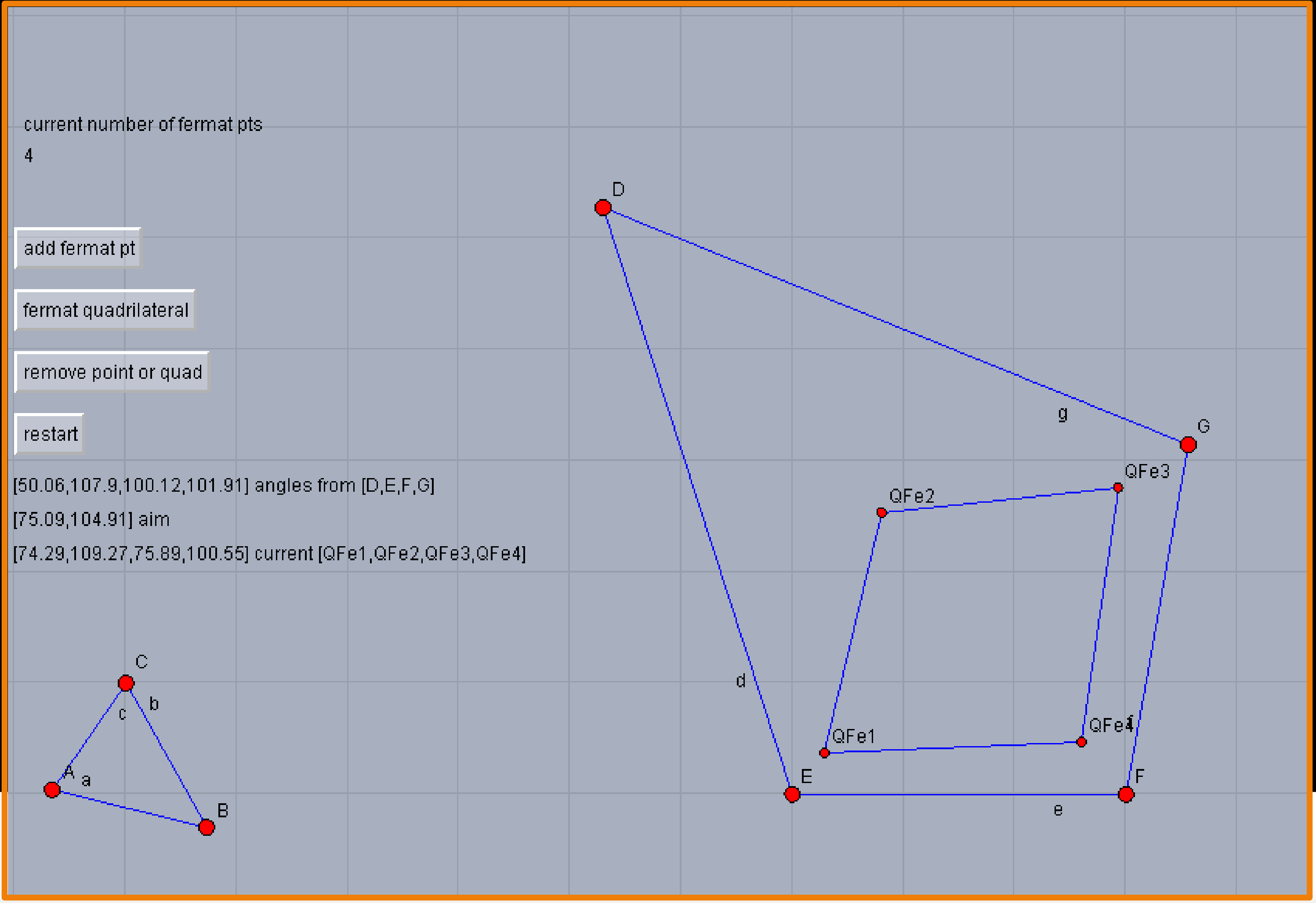
The basis of all optimization is:

Theorem (Weierstrass) A proper lower semicontinuous function on a compact set in a topological space achieves its infimum.

Hence a proper lower semicontinuous function on a weakly compact set in a Banach space achieves its infimum.

- When this holds we are in business.
- If not we have to work harder to establish the minimum exists
  - e.g., the isoperimetric problem (of Queen Dido).

The Fermat (location) problem with a twist

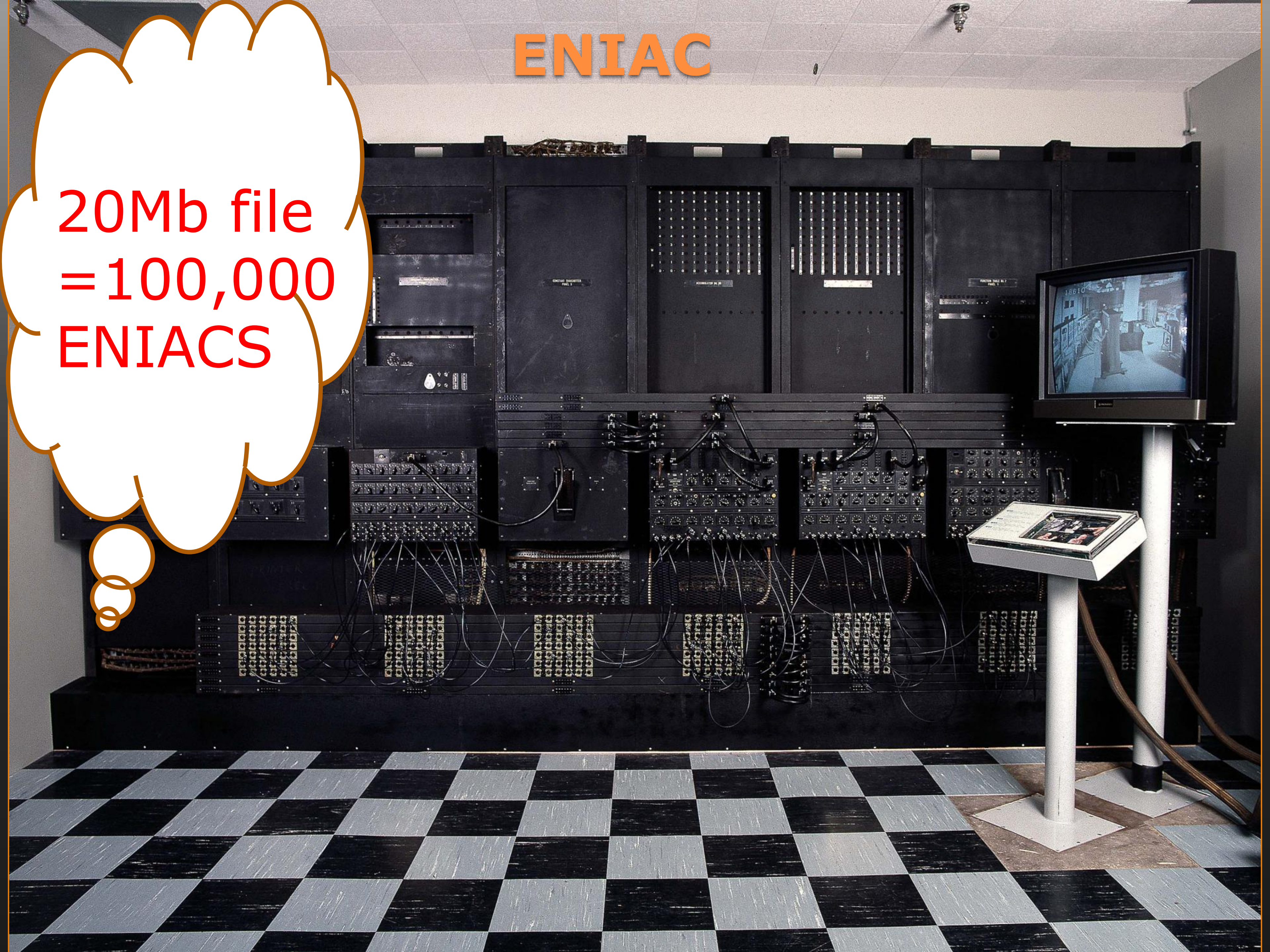


Day 1: An Overview of the Week and How to Maximize Surprise

"What it comes down to is our software is too hard and our hardware is too soft."

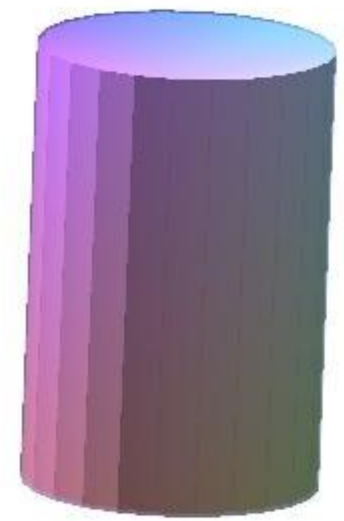
Day 2: Convex Duality and Applications

"It says it's sick of doing things like inventories and payrolls, and it wants to make some break-throughs in astrophysics."

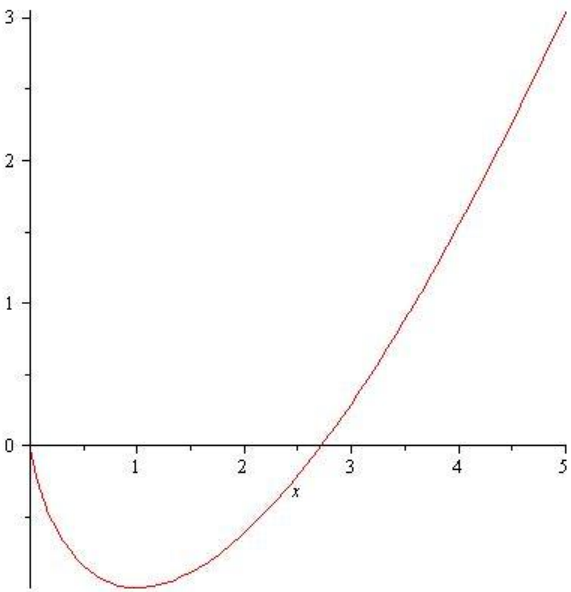

ENIAC
20Mb file
=100,000
ENIACS

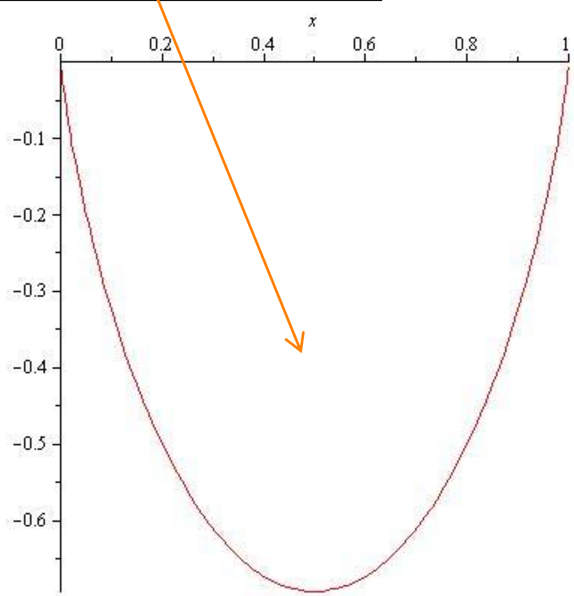
Convex functions

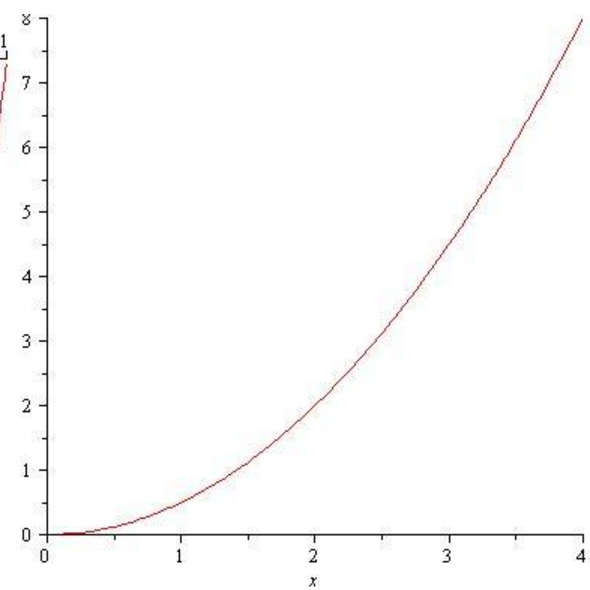
- A set is convex iff its **indicator function** is



- A function is convex iff the **epigraph** is

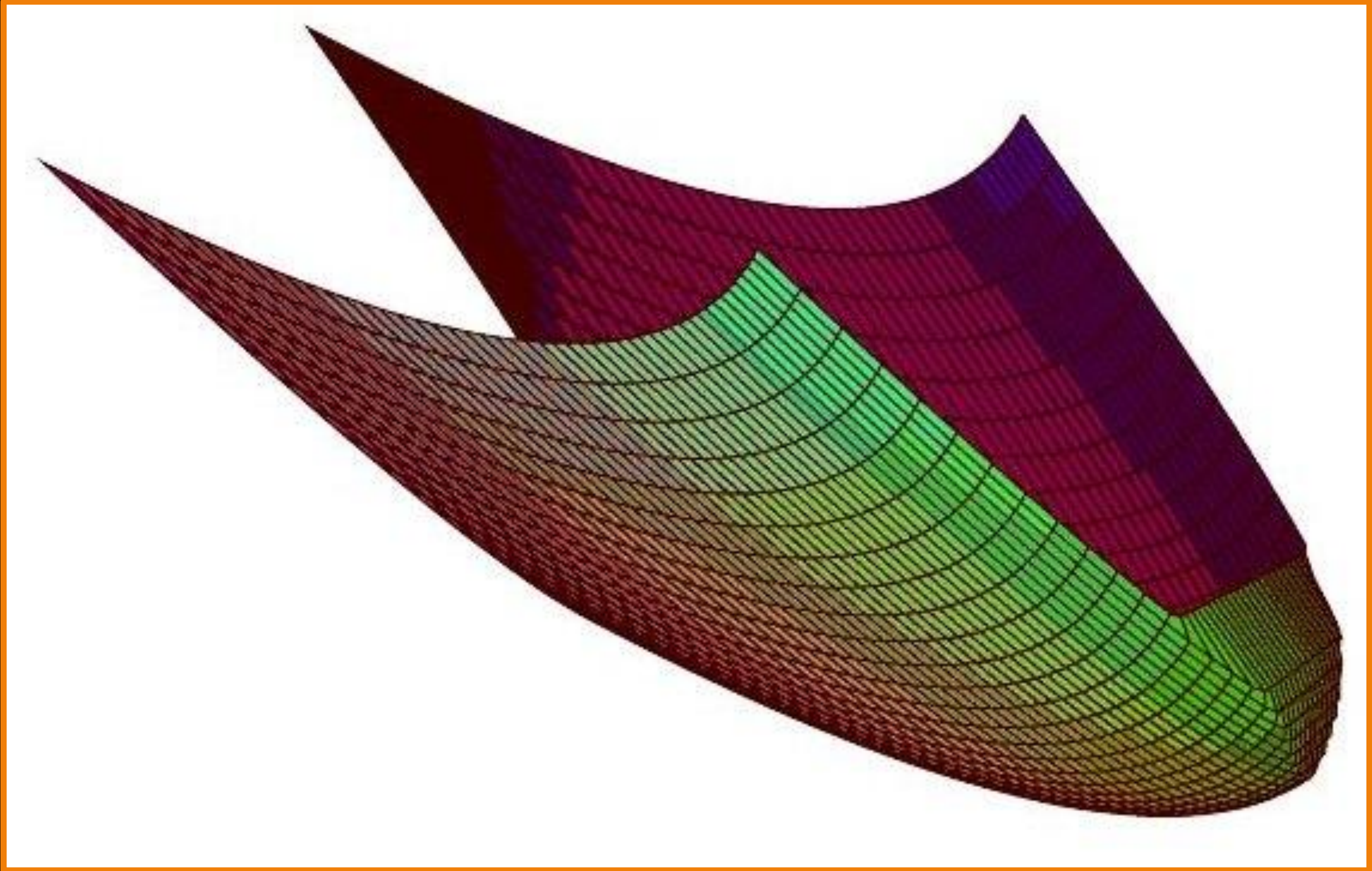






$x\log(x) - x$, $x\log(x) + (1-x)\log(1-x)$ and $x^2/2$ for $x > 0$

A subtle convex function



An essentially strictly convex function f which is not strictly convex and dom ∂f is non-convex:

$$(x, y) \mapsto \max\{(x-2)^2 + y^2 - 1, -(xy)^{1/4}\}$$

CANO2

CFC3

Topics for today: subgradients

- Symbolic convex analysis (in **Maple**)

- formalizing our model for the week (Potter and Arun)

Subgradients

CFC3

Max formula = subgradient and Fenchel duality

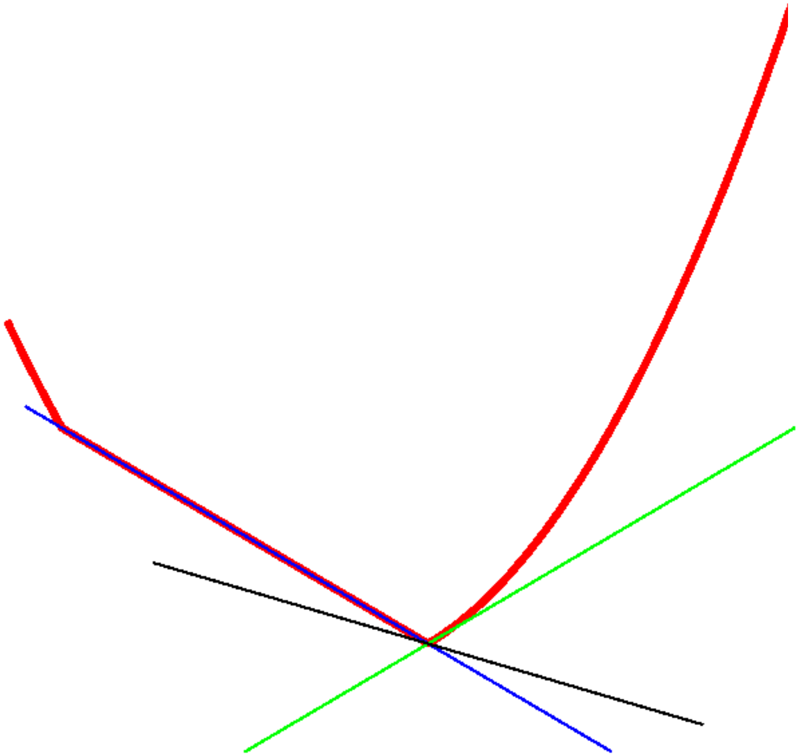


Figure 2.2: The Max formula (2.1.3) for $\max\{x^2, -x\}$.

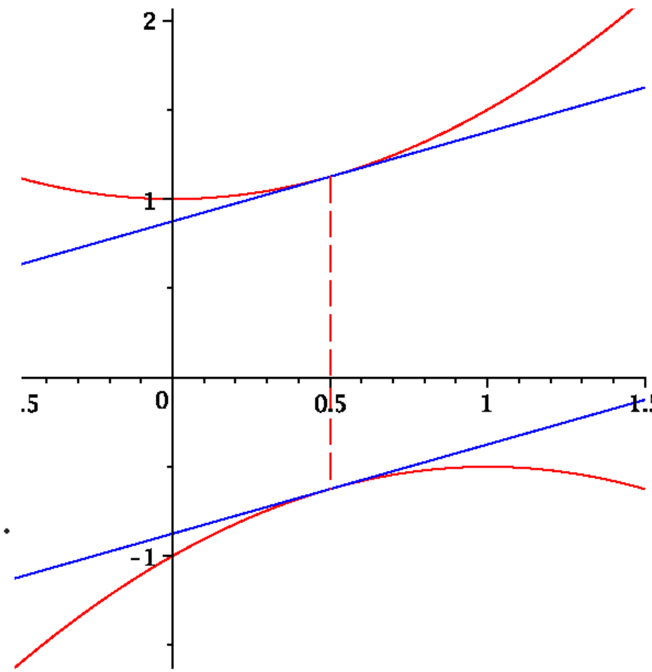
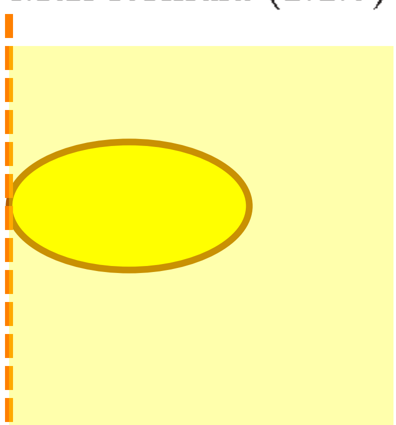


Figure 2.6: Fenchel duality (Theorem 2.3.4) illustrated for $x^2/2+1$ and $-(x$ $1)^2/2 - 1/2$. The minimum gap occurs at 1/2 with value 7/4.

Tangent cone to ellipse includes vertical line



CFC3

SCA

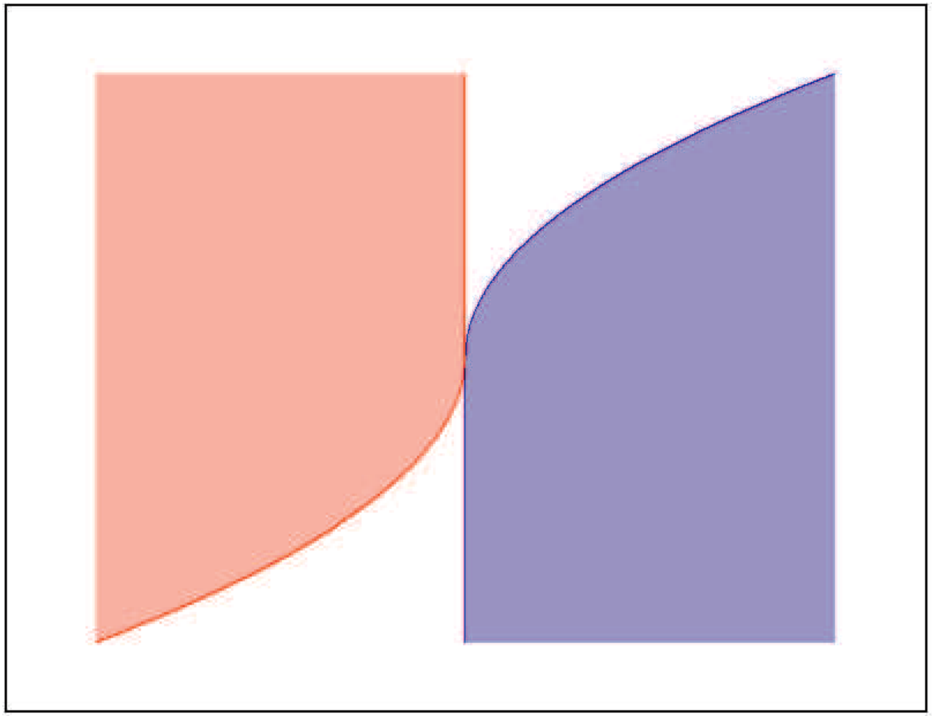
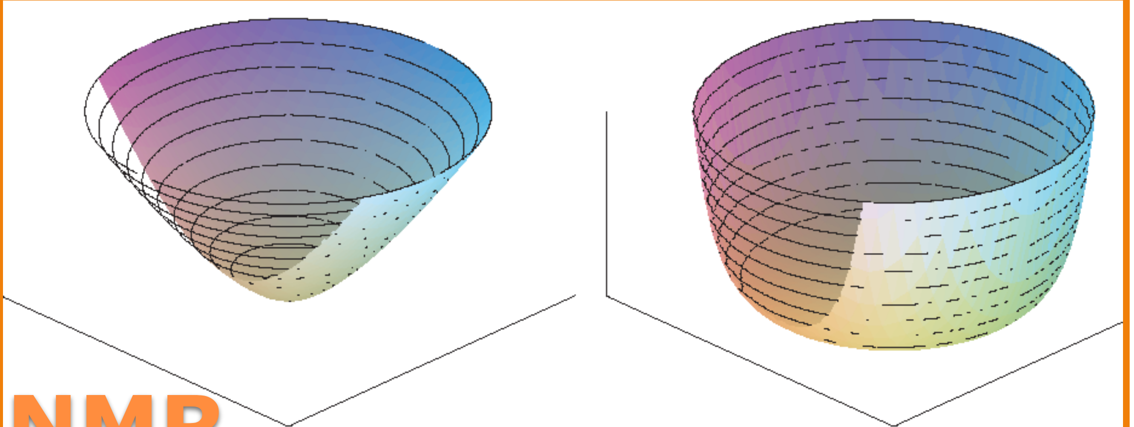


Figure 2.15: **Failure of Fenchel duality for $\sqrt{x}$ and $-\sqrt{-x}$.**

2.3.32 (Symbolic convex analysis).[†] It is possible to perform a significant amount of convex analysis in a computer algebra system—both in one and several dimensions. Some of the underlying ideas are discussed in [87]. The computation of $f_5 := x \mapsto |x| - 2\sqrt{1-x}$ in Chris Hamilton's *Maple* package SCAT (`http://flame.cs.dal.ca/~chamilto/files/scat.mpl`) is shown in Figure 2.11. Figure 2.12 and Figure 2.13 illustrate computing $sdf5 = \partial f_5$ and $g_5 = f_5^*$ respectively. Note the need to deal carefully with piecewise smooth functions. Figure 2.13 also confirms the convexity of f_5. The plots of f_5 and ∂f_5 are shown in Figure 2.14.

2.3.31 (NMR entropy).[†] Let $z = (x, y)$ and let $|z|$ denote the Euclidean norm. Show that $z \mapsto \cosh(|z|)$ and $z \mapsto |z| \log\left(|z| + \sqrt{1+|z|^2}\right) - \sqrt{1+|z|^2}$ are mutually conjugate convex function on $\mathbb{R}^2$ (or $\mathbb{C}$). They are shown in Figure 2.10. The latter is the building block of the *Hoch-Stern information function* introduced for entropy-based magnetic resonancing [102].



NMR

Figure 2.10: **The NMR entropy and its conjugate.**

```
piecewise(-3<=x and x<=1,abs(x)-2*sqrt(1-x),infinity)
f5 := convert(%,PWF);
```

$$f5 := \begin{cases} \infty, & x < -3 \\ 1, & x = -3 \\ -2\sqrt{1-x} - x, & (-3 < x) \textbf{ and } (x < 0) \\ -2, & x = 0 \\ -2\sqrt{1-x} + x, & (0 < x) \textbf{ and } (x < 1) \\ 1, & x = 1 \\ \infty, & 1 < x \end{cases}$$

Figure 2.11: **Symbolic convex analysis of f_5.**

```
Plot(sdf5,-3..1,view=[-3..1,-3..5],axes=none),
```

$$sdf5 := \begin{cases} \{\}, & x < -3 \\ [-\infty, -\frac{1}{2}], & x = -3 \\ \{\frac{(-1+\sqrt{1-x})\sqrt{1-x}}{x-1}\}, & (-3 < x) \textbf{ and } (x < 0) \\ [0, 2], & x = 0 \\ \{-\frac{(1+\sqrt{1-x})\sqrt{1-x}}{x-1}\}, & (0 < x) \textbf{ and } (x < 1) \\ \{\}, & x = 1 \\ \{\}, & 1 < x \end{cases}$$

Figure 2.12: **Symbolic convex analysis of ∂f_5.**

$$g5 := \begin{cases} -3y + 1, & y < -\frac{1}{2} \\ \frac{5}{2}, & y = \frac{-1}{2} \\ \frac{y^2+2y+2}{1+y}, & (\frac{-1}{2} < y) \text{ and } (y < 0) \\ 2, & y = 0 \\ 2, & (0 < y) \text{ and } (y < 2) \\ 2, & y = 2 \\ \frac{y^2-2y+2}{-1+y}, & 2 < y \end{cases}$$

```
Conj(g5,x):
(f5,F5);
```

true

Figure 2.13: **Symbolic convex analysis of $g_5 = f_5^*$.**

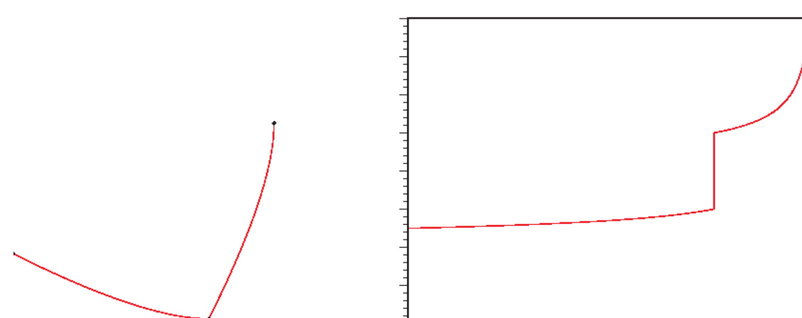


Figure 2.14: **Plots f_5 and ∂f_5.**

(back)

Application to Reflections

R_C **is nonexpansive.** First $P_C(x)$ exists and is unique. Then $\min_{c\in C} \frac{1}{2}\|x-c\|^2 = \min_{y\in H} \frac{1}{2}\|y-x\|^2 - \iota_C(y)$, and so for $p := P_C(x)$

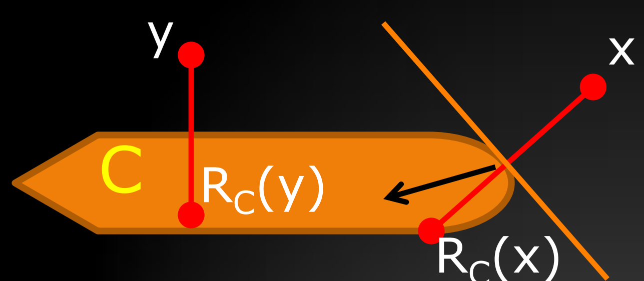
$$0 \in \partial\left(\frac{1}{2}\|\cdot - x\|^2 - \iota_C\right)(p) \Leftrightarrow \boxed{\langle x-p, p-c\rangle \le 0, \forall c \in C.}$$

characterizes p= $P_C(x)$

Hence with $q := P_C(y)$ we have

$$\langle x-p, p-q\rangle \le 0, \langle y-q, q-p\rangle \le 0 \Rightarrow \langle x-y, p-q\rangle \ge \langle p-q, p-q\rangle$$

$$\Leftrightarrow \|R_C(x) - R_C(y)\| \le \|x-y\|, \forall x, y \in X.$$



P_C is "firmly nonexpansive": for some nonexpansive T

$$P = \frac{T+I}{2}$$

Assuming $\|A\| \cdot\ 1$ we check $M_A := AP_CA^*$ is also firmly nonexpansive

Dualizing Potter and Arun

Potter and Arun show

$$v(b) := \min_{c\in C}\{\|z\|: \quad Az = b\}$$

is solved by

$$M_A(x) = A\left(P_C(A^*x)\right) = b$$

assuming $b \in \operatorname{int} A(C)$, the (CQ).

—as we saw.

- **Check** this follows from $-\lambda \in \partial v(b)$ and *projection characterization*.
- Thursday we will look at solving

$$\boxed{M_A x = b}$$

with M_A firmly nonexpansive.

(back)

Day 3: Variational Principles and Applications

"What I appreciate even more than its remarkable speed and accuracy are the words of understanding and compassion I get from it."

2009 Record Pi Computation

Parallel Implementation of Multiple-Precision Arithmetic and $1,649,267,440,000$ Decimal Digits of π Calculation

Daisuke Takahashi

Graduate School of Systems and Information Engineering, University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan
daisuke@cs.tsukuba.ac.jp

Abstract. We present efficient parallel algorithms for multiple-precision arithmetic operations of more than several million decimal digits on distributed-memory parallel computers. A parallel implementation of floating-point real FFT-based multiplication is used because a key operation in fast multiple-precision arithmetic is multiplication. We also parallelized an operation of releasing propagated carries and borrows in multiple-precision addition, subtraction and multiplication. More than 1.6 trillion decimal digits of π were computed on 256 nodes of Appro Xtreme-X3 (648 nodes, 147.2 GFlops/node, 95.4 TFlops peak performance) with a computing elapsed time of 137 hours 42 minutes which includes the time for verification.

Previous Record: December 2002
1.4 trillion decimals
(1 trillion hex digits)
by Kanada and his team.
They used Machin formulae
confirmed by BBP hex formula.

5.3 Results of π $1,649,267,440,000$ Decimal Digit Calculation

The calculations of π by Gauss-Legendre algorithm and Borweins' quartically convergent algorithm were performed on 256 nodes of Appro Xtreme-X3 supercomputer.

All routines were written in FORTRAN 77 with MPI and OpenMP. Main program and verification program were run on 1,024 MPI processes, i.e. each node has 4 MPI processes. Due to the time limit of a job, two programs were performed in 10 steps, respectively.

Main program run:

Job start	: 2nd January 2009 21:17:26 (JST)
Job end	: 5th January 2009 23:56:23 (JST)
Total elapsed time	: 64 hours 14 minutes
Main memory	: 6732 GB
Algorithm	: Gauss-Legendre algorithm

Verification program run:

Job start	: 6th January 2009 01:36:45 (JST)
Job end	: 26th January 2009 08:45:58 (JST)
Total elapsed time	: 73 hours 28 minutes
Main memory	: 6348 GB
Algorithm	: Borweins' quartically convergent algorithm

The decimal numbers of π and $1/\pi$ from $1,649,267,439,951$-st to $1,649,267,440,000$-th digits are:

π : 7712856414 0105560548 9805732574 3212539317 0912654849
$1/\pi$: 7726694296 8436590719 4549360485 5555663940 4302590248.

The main computation took 40 iterations of the improved Gauss-Legendre algorithm for π, to yield $3 \times 2^{39} = 1,649,267,441,664$ digits of π. This computation was checked with 20 iterations of improved Borweins' quartically convergent algorithm for $1/\pi$, followed by a reciprocal operation.

A comparison of these output results gave no discrepancies except for the last 139 digits due to the normal truncation errors.

Topics for today

CANO2 +...

- tangency of convex sets and

- an application of metric regularity
- two smooth variational principles

- computing projections with KKT multipliers

Tangency

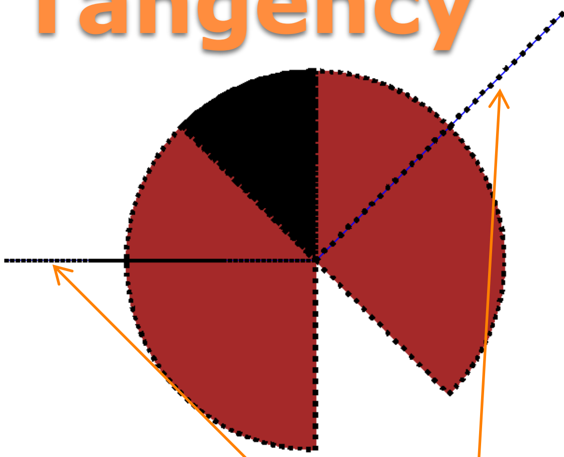


Fig. 5.2. A Pacman set, star and normal cone.

Regular limiting normals

$T_C(x) = N_C(x)^o$
shrinks as $K_C(x)$ grows
- a nonconvex pathology

(back)

Theorem 6.3.6 (Tangent cones) *Suppose the point x lies in a set S in* $\mathbf{E}$.

(a) *The contingent cone $K_S(x)$ consists of those vectors h in $\mathbf{E}$ for which there are sequences $t_r \downarrow 0$ in $\mathbf{R}$ and $h^r \to h$ in $\mathbf{E}$ such that $x + t_r h^r$ lies in S for all r.*

(b) *The Clarke tangent cone $T_S(x)$ consists of those vectors h in $\mathbf{E}$ such that for any sequences $t_r \downarrow 0$ in $\mathbf{R}$ and $x^r \to x$ in S, there is a sequence $h^r \to h$ in $\mathbf{E}$ such that $x^r + t_r h^r$ lies in S for all r.*

Intuitively, the contingent cone $K_S(x)$ consists of limits of directions to points near x in S, while the Clarke tangent cone $T_S(x)$ "stabilizes" this tangency idea by allowing perturbations of the base point x.

We call the set S *tangentially regular* at the point $x \in S$ if the contingent and Clarke tangent cones coincide (which clearly holds if the distance function d_S is regular at x). The convex case is an example.

Corollary 6.3.7 (Convex tangent cone) *If the point x lies in the convex set $C \subset \mathbf{E}$, then C is tangentially regular at x with*

$$T_C(x) = K_C(x) = \mathrm{cl}\,\mathbf{R}_+(C - x).$$

Proof. The regularity follows from Theorem 6.2.2 (Regularity of convex functions). The identity $K_C(x) = \mathrm{cl}\,\mathbf{R}_+(C - x)$ follows easily from the contingent cone characterization in Theorem 6.3.6. □

Our very first optimality result (Proposition 2.1.1) required the condition $-\nabla f(x) \in N_C(x)$ if the point x is a local minimizer of a differentiable function f on a convex set $C \subset \mathbf{E}$. If the function $f : \mathbf{E} \to (\infty, +\infty]$ is convex and continuous at $x \in C$, then in fact a necessary and sufficient condition for global minimality is

$$0 \in \partial(f + \delta_C)(x) = \partial f(x) + N_C(x),$$

Ekeland's principle in Euclidean Space

Theorem 3.5.1 (Ekeland's variational principle). *Let E be Euclidean and let $f\colon E \to (-\infty, +\infty]$ be a lsc function bounded from below. Suppose that $\varepsilon > 0$ and $z \in E$ satisfy*

$$f(z) < \inf_E f + \varepsilon.$$

Suppose $\lambda > 0$ is given, then there exists $y \in E$ such that

(a) $\|z - y\| \leq \lambda$, *(b)* $f(y) + (\varepsilon/\lambda)\|z - y\| \leq f(z)$, *and*

(c) $f(x) + (\varepsilon/\lambda)\|x - y\| > f(y)$, *for all* $x \in E \setminus \{y\}$.

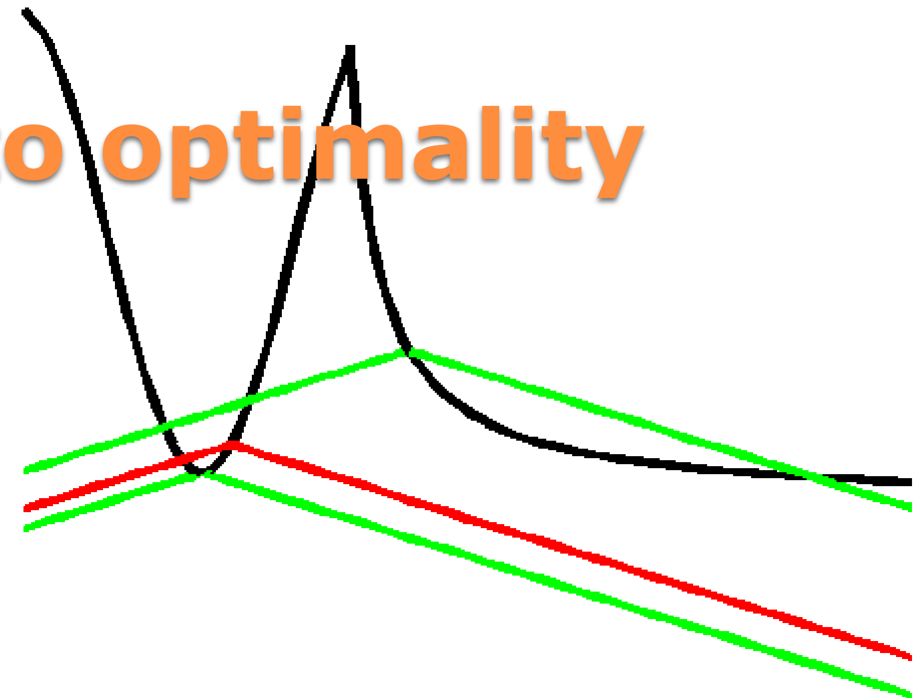
Proof. Let g be defined by $g(x) := f(x) + (\varepsilon/\lambda)\|x - z\|$. Then g is lsc and coercive and so achieves its minimum at a point y. Hence

$$f(x) + (\varepsilon/\lambda)\|x - z\| \geq f(y) + (\varepsilon/\lambda)\|z - y\| \tag{3.5.1}$$

for all $x \in E$. In particular $\inf_E f + \varepsilon > f(z) \geq f(y) + (\varepsilon/\lambda)\|z - y\|$, whence (a) and (b) follow. The triangle inequality applied to (3.5.1) gives (c). $\square$

Ekeland's principle is Pareto optimality for an ice-cream (second-order) cone

Pareto optimality



x_0 is ε-optimal, x_1 is an improvement, and x is a point guaranteed by EVP

Fig. 2.1. Ekeland variational principle. Top cone: $f(x_0) - \varepsilon|x - x_0|$; Middle cone: $f(x_1) - \varepsilon|x - x_1|$; Lower cone: $f(y) - \varepsilon|x - y|$.

the osculating function is nonsmooth at the important point:

This can be fixed as I discovered in 1986

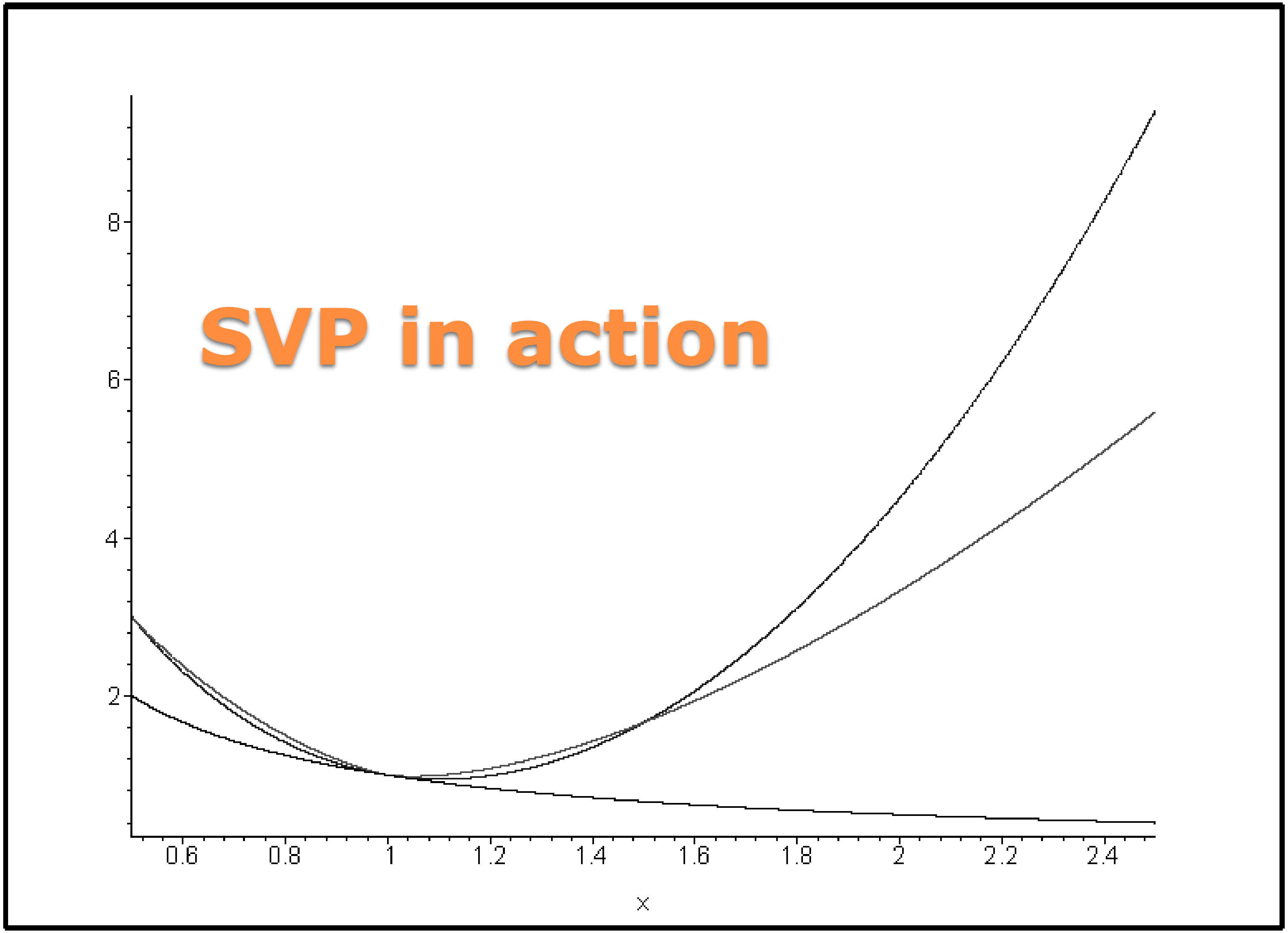


Fig. 2.5. Smooth attained perturbations of $1/x$

Definition 2.5.1 *Let (X,d) be a metric space. We say that a continuous function $\rho\colon X\times X\to[0,\infty]$ is a* gauge-type function *on a complete metric space (X,d) provided that* (usually a norm)

(i) $\rho(x,x)=0$, *for all* $x\in X$,
(ii) *for any* $\varepsilon>0$ *there exists* $\delta>0$ *such that for all* $y,z\in X$ *we have* $\rho(y,z)\le\delta$ *implies that* $d(y,z)<\varepsilon$.

Theorem 2.5.2 (Borwein–Preiss Variational Principle) *Let (X,d) be a complete metric space and let $f\colon X\to\mathbb{R}\cup\{+\infty\}$ be a lsc function bounded from below. Suppose that ρ is a gauge-type function and $(\delta_i)_{i=0}^{\infty}$ is a sequence of positive numbers, and suppose that $\varepsilon>0$ and $z\in X$ satisfy*

(back)

$$f(z)\le\inf_X f+\varepsilon.$$

2 SVPs

Then there exist y and a sequence $\{x_i\}\subset X$ such that

(i) $\rho(z,y)\le\varepsilon/\delta_0$, $\rho(x_i,y)\le\varepsilon/(2^i\delta_0)$,
(ii) $f(y)+\sum_{i=0}^{\infty}\delta_i\rho(y,x_i)\le f(z)$, *and*
(iii) $f(x)+\sum_{i=0}^{\infty}\delta_i\rho(x,x_i)>f(y)+\sum_{i=0}^{\infty}\delta_i\rho(y,x_i)$, *for all* $x\in X\backslash\{y\}$.

Theorem 4.3.6 (Smooth variational principle). *Let X be a Banach space that admits a Lipschitz function with bounded nonempty support that is Fréchet differentiable (resp. Gâteaux differentiable). Then for every proper lsc bounded below function f defined on X and every $\varepsilon>0$, there exists a function g which is Lipschitz and Fréchet differentiable (resp. Gâteaux differentiable) on X such that $\|g\|_\infty\le\varepsilon$, $\|g'\|_\infty\le\varepsilon$ and $f+g$ attains its strong minimum on X.*

(back)

Error bounds and the distance to the intersection of two convex sets

Theorem [Distance to the intersection] Suppose that C and D are closed convex sets in a Banach space and that

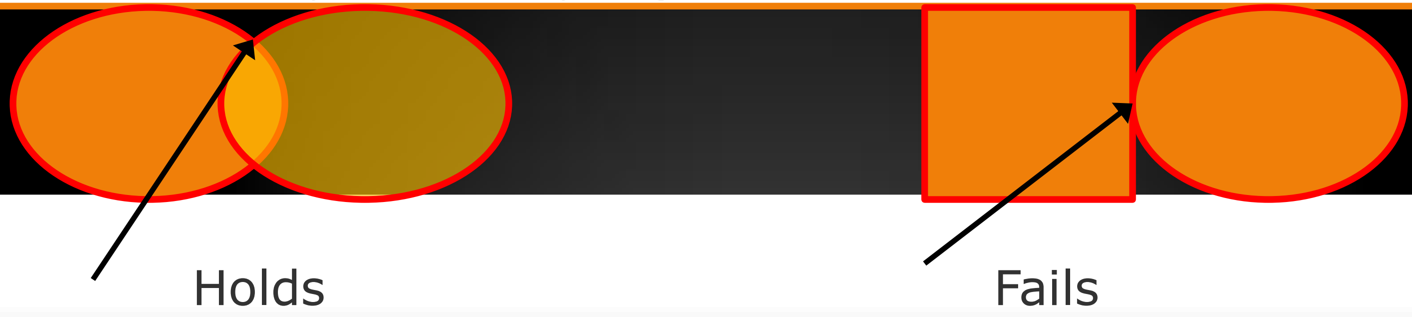
$$0 \in \operatorname{int}\{C - D\}.$$

Then for any $M > 0$ there is $k > 0$ such that

$$d_{C\cap D}(x) \leq k\{d_C(x) + d_D(x)\}$$

for all x with $\|x\| \leq M$.

Proof sketch The multifunction $\Omega(x) := x - D$ for $x \in C$, and empty elsewhere, is weakly metrically regular.



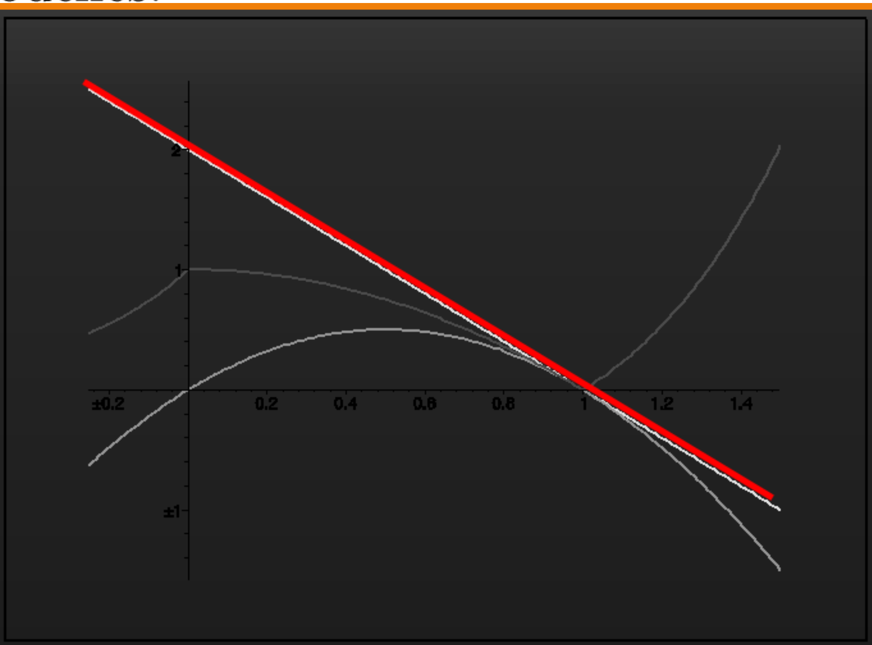
(back)

Asplund spaces

An important corollary is

Theorem If a Banach space X has a Fréchet (respectively Gateaux) differentiable norm then every continuous convex function on X is densely Fréchet (respectively Gateaux) differentiable.

Proof. A concave function must be differentiable at any point at which a smooth minorant touches.



(back)

Computing Projections using Multipliers

Remark 2 (Nearest point to an ellipse) Consider the ellipse

$$E := \left\{(x,y) \colon \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\right\}$$

in standard form. The best approximation $P_E(u,v) = \left(\frac{a^2 u}{a^2 - t}, \frac{b^2 v}{b^2 - t}\right)$ where t solves $\frac{a^2 u^2}{(a^2-t)^2} + \frac{b^2 v^2}{(b^2-t)^2} = 1$. This generalizes neatly to a hyperbola (one solves the general quartic $x^4 - ux^3 + vx - 1 = 0$ and $[x, 1/x]$ is the nearest point.)

Remark 3 (Nearest point to the p-sphere) For $0 < p < \infty$, consider the p-sphere in two dimensions

$$S_p := \{(x,y) \colon |x|^p + |y|^p = 1\}.$$

Let $z^* := (1 - z^p)^{1/p}$. For $uv \neq 0$, the best approximation $P_{S_p}(u,v) = (\operatorname{sign}(u) z, \operatorname{sign}(v) z^*)$ where either $z = 0, 1$ or $0 < z < 1$ solves

$${z^*}^{p-1}(z - |u|) - z^{p-1}(z^* - |v|) = 0.$$

[Then one computes the two or three distances and select the point yielding the least value. It is instructive to make a plot, say for $p = 1/2$.] This extends to the case where $uv = 0$. Note that this also yields the nearest point formula for the p-ball.

The Raleigh Quotient

Let $A : H \to H$ be symmetric, compact and linear. Consider the convex *maximization*

$$\mu := \max\{\langle Ax, x\rangle \colon \|x\|^2 \leq 1\}$$

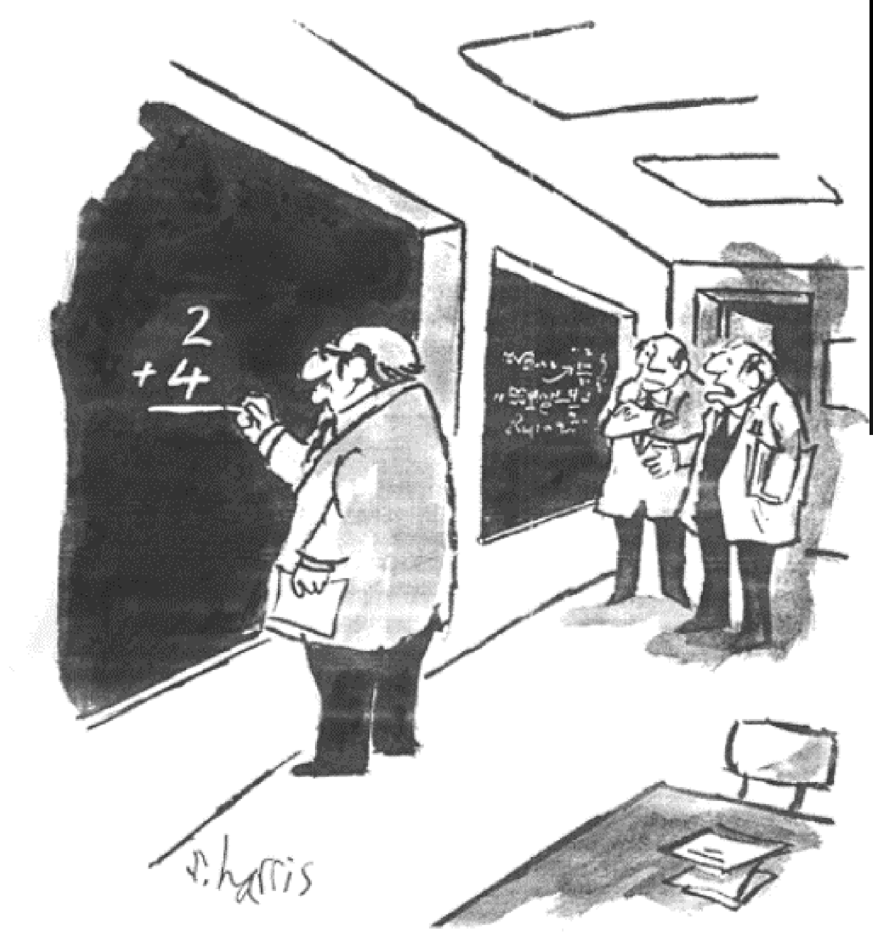
- The max exists since the A is compact.
- The KKT theorem applies, yielding $\lambda > 0$ and $\|x^*\| = 1$:

$$Ax^* = \lambda x^*$$

and provides a maximal eigenvector.

(back)

Day 4: Monotonicity & Applications



"He was very big in Vienna."

Goethe about Us

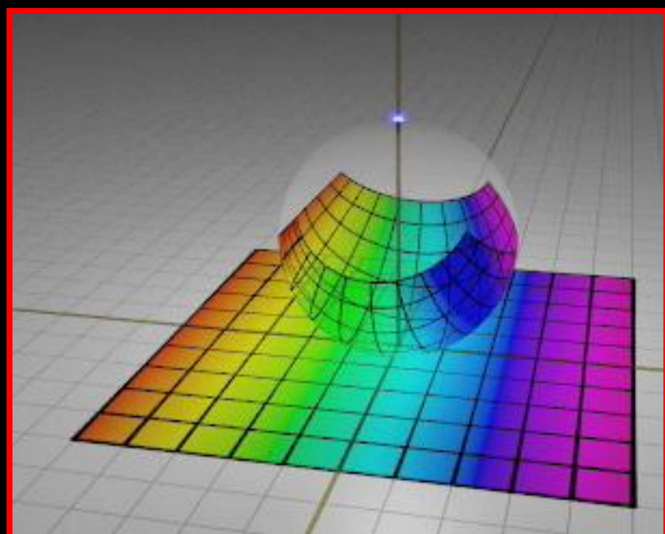
“Die Mathematiker sind eine Art Franzosen; redet man mit ihnen, so übersetzen sie es in ihre Sprache, und dann ist es alsobald ganz etwas anderes. [Mathematicians are a kind of Frenchman: whatever you say to them they translate into their own language, and right away it is something entirely different.]”

(Johann Wolfgang von Goethe, 1748-1932)

Maximen und Reflexionen, no. **1279**, p.160 Penguin Classic ed.

Visualization of Moebius transformations and Chemical potentials

Rogness and Arnold
at **IMA**



Helaman Ferguson
Sculpture

Topics for today

- Cuscos and Fenchel Duality as decoupling
- Multifunction Section 5.1.4 from TOVA
- Sum theorem for maximal monotones
- Monotonicity of the Laplacian
- Potter and Arun revisited
  - iterates of firmly non-expansive mappings
  - implementing our model for the week (Potter and Arun)

Minimal and non-minimal Cuscos

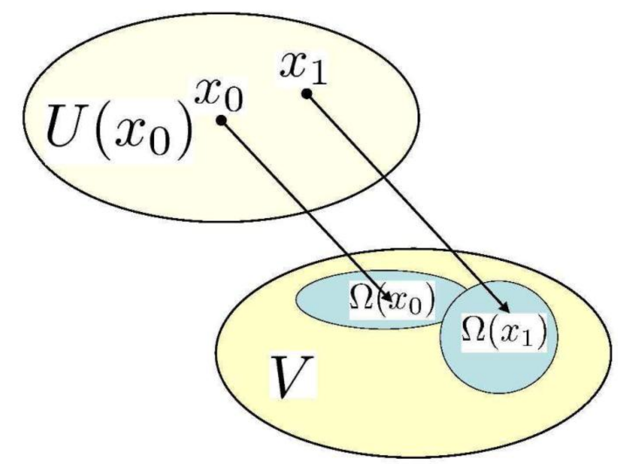


Figure 6.1: A cusco with $\Omega(x_0) \subset V$.

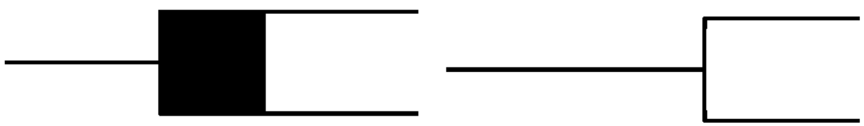


Figure 6.2: Two nonminimal cuscos on $[-2, 2]$.

Figure 6.3: A minimal cusco and smaller minimal usco on $[-2, 2]$.

Fenchel Duality as Decoupling

Lemma 4.3.1 (Decoupling Lemma) *Let X and Y be Banach spaces, let the functions $f\colon X \to \mathbb{R}$ and $g\colon Y \to \mathbb{R}$ be convex and the map $A\colon X \to Y$ be linear and bounded. Suppose that f, g and A satisfy either the condition*

$$0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f) \tag{4.3.1}$$

and both f and g are lsc, or the condition

$$A \ \operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset. \tag{4.3.2}$$

Then there is a $y^* \in Y^*$ such that for any $x \in X$ and $y \in Y$,

$$p \leq [f(x) - \langle y^*, Ax\rangle] + [g(y) + \langle y^*, y\rangle], \tag{4.3.3}$$

where $p = \inf_X \{f(x) + g(Ax)\}$.

(back)

The Sum Theorem

(back)

$*$**Exercise 5.1.44** Let X be a reflexive Banach space. Prove that a monotone mapping $T : X \to 2^X$ is maximal if and only if the mapping $T(\cdot + x) + J$ is surjective for all x in X. References: [33, 240].

$*$**Exercise 5.1.45** Prove the following theorem.

Theorem 5.1.35 *Let X be a reflexive space, let T be maximal and let f be closed and convex. Suppose that*

$$0 \in \operatorname{core}\{\operatorname{conv}\operatorname{dom}(T) - \operatorname{conv}\operatorname{dom}\partial(f)\}.$$

Then

(a) $\partial f + T + J$ is surjective.
(b) $\partial f + T$ is maximal monotone.
(c) ∂f is maximal monotone.

Hint: Consider the Fitzpatrick function $F_T(x, x^*)$ and further introduce $f_J(x) := f(x) + 1/2\|x\|^2$. Let $G(x, x^*) := -f_J(x) - f_J^*(-x^*)$. Observe that

$$F_T(x, x^*) \geq \langle x, x^* \rangle \geq G(x, x^*)$$

pointwise thanks to the *Fenchel-Young inequality*. Now apply the decoupling result in Lemma 4.3.1 and Exercise 5.1.44.

$*$**Exercise 5.1.46** Deduce the following result in [240] as a corollary of Theorem 5.1.35.

Theorem 5.1.36 *Let X be a reflexive Banach space, let $T, S : X \to 2^X$ be maximal monotone operators. Suppose that*

$$0 \in \operatorname{core}[\operatorname{conv}\operatorname{dom}(T) - \operatorname{conv}\operatorname{dom}(S)].$$

Then $T + S$ is maximal monotone.

Hint: Apply Theorem 5.1.35 to $T(x, y) := (T_1(x), T_2(y))$ and the indicator function $f(x, y) = \iota_{\{x'=y'\}}(x, y)$.

Laplacians as Maximal Monotone Operators

9.3.2 (Elliptic partial differential equations [195, 131, 273]).† Much early impetus for the study of maximal monotone operators came out of partial differential equations and takes place within the confines of Sobolev space —and so we content ourselves with an example of what is possible.

As an application of their study of existence of eigenvectors of second order nonlinear elliptic equations in $L_2(\Omega)$, the authors of [273] assume that $\Omega \subset \mathbb{R}^n$, $(n > 1)$ is a bounded open set with boundary belonging to $C^{2,\alpha}$ for some $\alpha > 0$. They assume that one has functions $|a_i(x,u)| \le \nu$ $(1 \le i \le n)$ and $|a_0(x,u)| \le \nu|u| + a(x)$ for some $a \in L_2(\Omega)$ and $\nu > 0$; where all a_i are measurable in x and continuous in u (a.e. x). They then consider the normalized eigenvalue problem

$$\boxed{\triangle u + \lambda \left\{ \sum_{i=1}^{n} a_i(x,u) \frac{\partial u}{\partial x_i} + a_0(x,u) \right\} = 0, \quad x \in \Omega,} \tag{9.3.8}$$

where $\triangle u = -\nabla^2 u = -\sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}$ is the classical *Laplacian* . To make this accessible to Sobolev theory, a weak solution is requested to (9.3.8) for $0 < \lambda \le 1$ when $u \in W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega)$. In this setting, a solution of

$$\triangle u + \tau u = f(x)$$

for all $\tau > 0$ and all $f \in L_2$ (and with $\|u\|_2 = 1$) is assured. Minty's surjectivity condition (Proposition 9.3.1) implies $T := \triangle$ is linear and maximal monotone on $L_2(\Omega)$ with domain $W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega)$. Of course, one must first check monotonicity of $\triangle$ using integration by parts in the form

$$\int_\Omega \langle v, \triangle u \rangle = \int_\Omega \langle \nabla v, \nabla u \rangle,$$

for all $v \in W^{-1,2}(\Omega), u \in C_0^\infty(\Omega) \subset W_0^{1,2}(\Omega)$. One is now able to provide a Fredholm alternative type result for (9.3.8) [273, Theorem 10]. In like-fashion one can make sense of the assertion that for $2 \le p < \infty$ the *p-Laplacian* $\triangle_p$ is maximal monotone: $\triangle_p u$ is given by

$$\boxed{\triangle_p u := -\mathrm{div}(|\nabla u|^{p-2} \nabla u) \in W^{-1,q}(\Omega)}$$

for $u \in W^{1,p}(\Omega)$ with $1/p + 1/q = 1$.

(back)

Nonexpansive Maps

In Hilbert space: P nonexpansive implies $I - P$ is monotone.
$\|Px - Py\| \leq \|x - y\| \Rightarrow \langle (I-P)(x) - (I-P)(y), x - y \rangle \geq 0.$

A monotone, continuous M is maximal: for $t > 0$

$$t^{-1}\langle M(x+th) - y, x + th - x\rangle = \langle M(x+th) - y, h\rangle \geq 0, \forall h.$$

Thus, $M(x_n) \to_{w^*} y, x_n \to x \Rightarrow y = M(x)$.

Suppose P nonexpansive and $F :=\mathrm{Fix}(T) \neq \emptyset$. For $p \in F, x \in H$

$$d_F(P(x)) \leq \|p - P(x)\| = \|P(p) - P(x)\| \leq \|x - p\| \qquad (*)$$

and so

$$d_F(P(x)) \leq d_F(x).$$

- (*) is a very strong property called *Féjer monotonicity*.

(back)

(Firmly) Nonexpansive Maps

demiclosure. $(I-P)(x_n) = y_n$
$y_n \to 0, x_n \to_w x \Rightarrow x = P(x)$
because $I - P$ is *maximal* monotone.

Facts. (Krasnoselski) Let P be non-expansive and set $x_{n+1} := P(x_n)$. Then $\lim_n \|x_{n+1} - x_n\| =: \sigma \geq 0$.
$\sigma = 0 \Leftrightarrow \|x_n - x\| \to 0$ for some $x \in F$.
(Takahashi) If P is firmly nonexpansive then $\sigma = 0$.

and so iteration provide solutions to Potter and Arun's formulation

$$T_\tau := I + \tau(b - AP_C A^*)$$
is firmly nonexpansive for $0 < \tau \leq 2$

(back)

Day 5: Closing the Circle: Interactive Algorithmic Analysis

A really new work: based on Potter and Arun's model

arXiv.org > math > arXiv:0907.0436v2

Search or Article-id (Help | Advanced search)

All papers Go!

Mathematics > Optimization and Control

Dualization of Signal Recovery Problems

Patrick L. Combettes, Dinh Dung, Bang Cong Vu

(Submitted on 2 Jul 2009 (v1), last revised 3 Jul 2009 (this version, v2))

In convex optimization, duality theory can sometimes lead to simpler solution methods than those resulting from direct primal analysis. In this paper, this principle is applied to a class of composite variational problems arising in signal recovery. These problems are not easily amenable to solution by current methods but they feature Fenchel-Moreau-Rockafellar dual problems that can be solved reliably by forward-backward splitting and allow for a simple construction of primal solutions from dual solutions. The proposed framework is shown to capture and extend several existing duality-based signal recovery methods and to be applicable to a variety of new problems beyond their scope.

Subjects: **Optimization and Control (math.OC)**; Functional Analysis (math.FA)

MSC classes: 90C25,

Cite as: **arXiv:0907.0436v2 [math.OC]**

Download:

- PostScript
- PDF
- Other formats

Current browse context:

math.OC

< prev | next >

new | recent | 0907

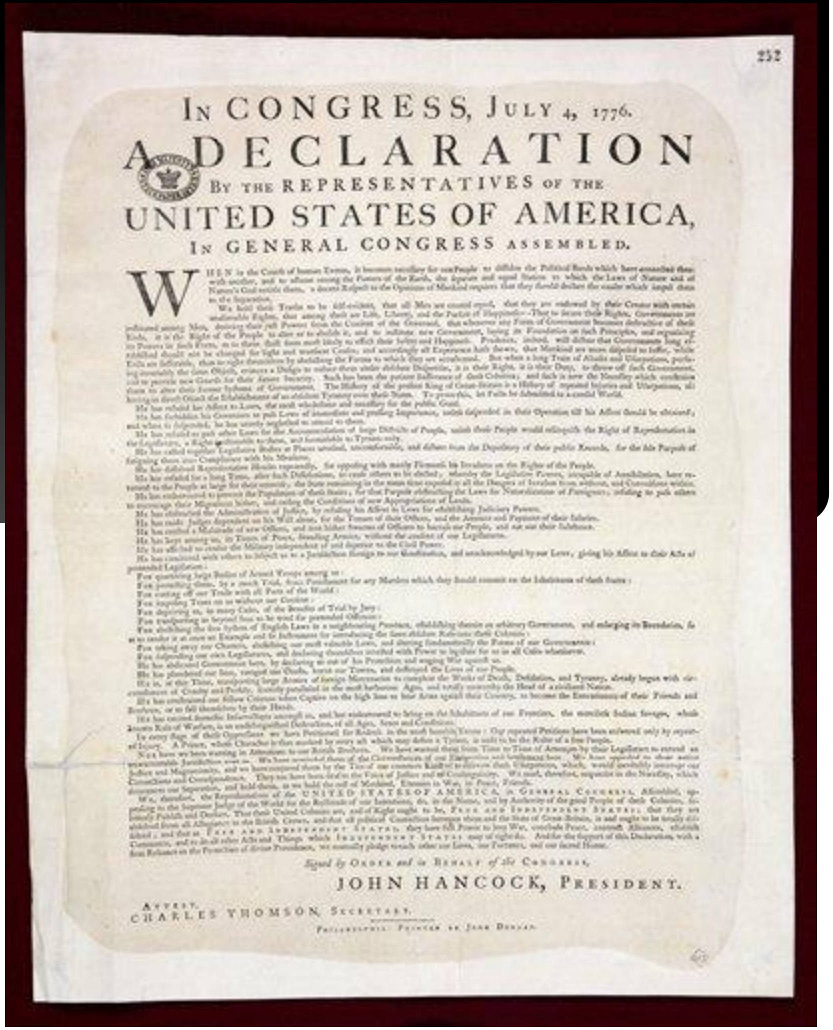
Change to browse by:

math

math.FA

Bookmark (what is this?)

The Old Math


IN CONGRESS, JULY 4, 1776.

A DECLARATION

BY THE REPRESENTATIVES OF THE

UNITED STATES OF AMERICA,

IN GENERAL CONGRESS ASSEMBLED.

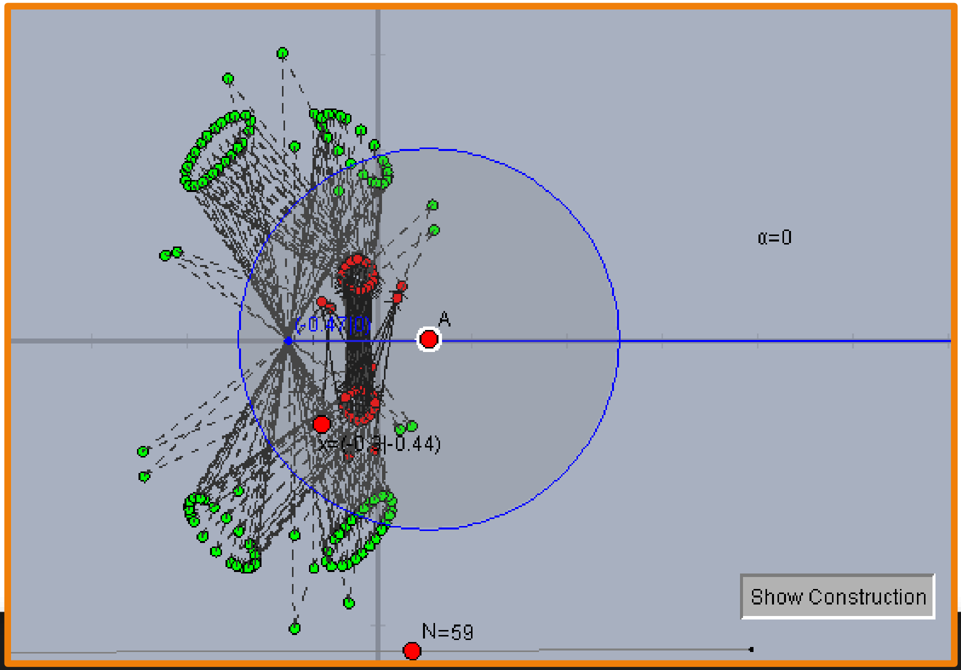
JOHN HANCOCK, PRESIDENT.

CHARLES THOMSON, SECRETARY.

And the New?

Topics: Algebraic Phase Reconstruction and Discovery

- Alternating Projections and Reflections
- Parallelization
- Related ODES and Linearizations
- Proofs and a final Variation on a Theme



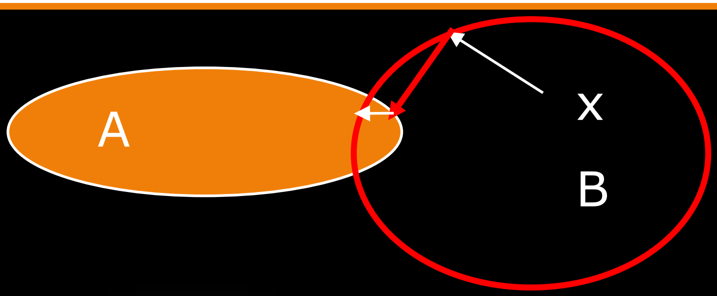
Periodicity with **reflections** on half line and circle

'2=N': Inverse Problems as Feasibility Problems

To find $x \in A \cap B$ one may use the *method of alternating projections*

$$y_n := P_B(x_n), \quad x_{n+1} := P_A(y_n)$$

(back)



This parallelizes efficiently to products of sets (project and average):

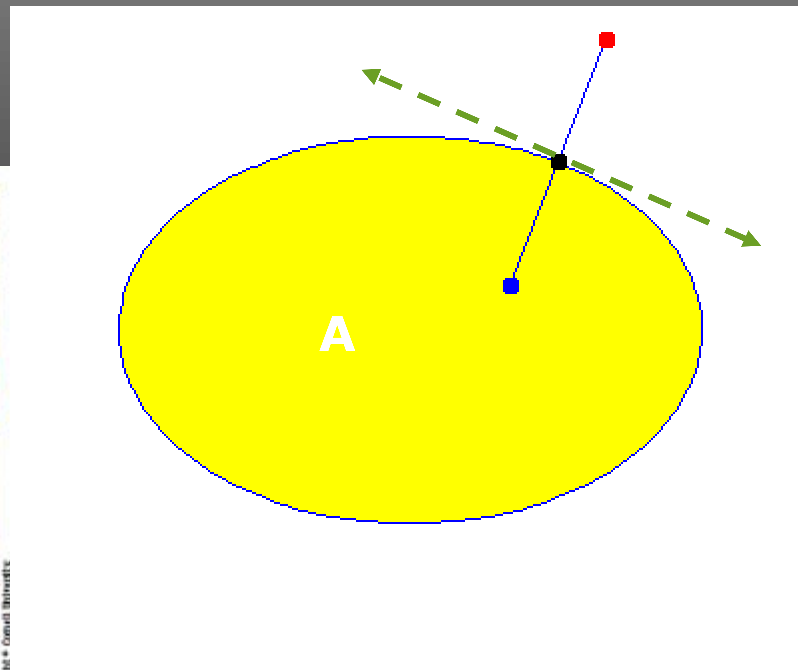
$$A := \prod_{k=1}^{M} A_k, \qquad B := \{(x_1, x_2, \cdots, x_N) | x_1 = x_2 = \cdots = x_N\}$$

$$P_A = (P_{A_1}, P_{A_2}, \cdots, P_{A_N}), \qquad P_B\,(x_1, x_2, \cdots, x_N)_i = \frac{\sum_{k=1}^{N} x_k}{N}.$$

- The theory is rich for convex sets.

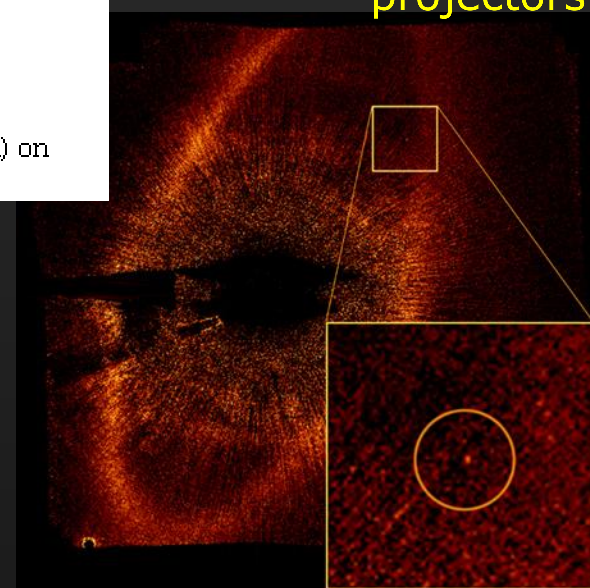
Algebraic Phase Reconstruction

Projectors and Reflectors: $P_A(x)$ is the metric projection or nearest point and $R_A(x)$ reflects in the tangent: x is red



projection (black) and reflection (blue) of point (red) on boundary (blue) of ellipse (yellow)

2008 Finding exoplanet Fomalhaut in Piscis with projectors



Veit Elser, Ph.D.

2007 Solving Sudoku with reflectors

"All physicists and a good many quite respectable mathematicians are contemptuous about proof."
- G. H. Hardy (1877-1947)

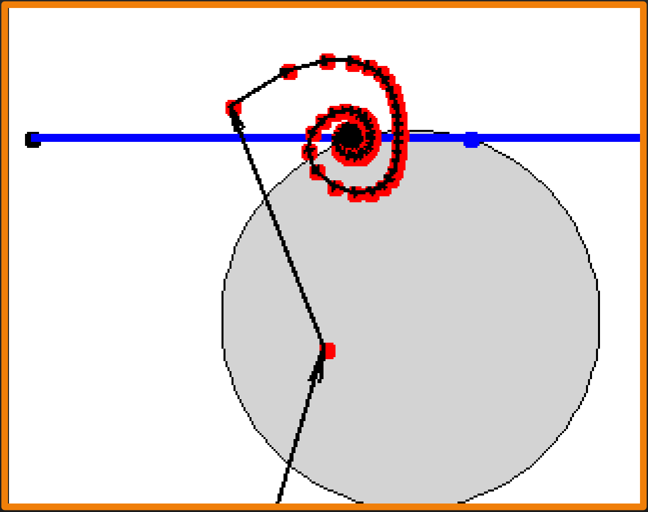
APR: Why does it work?

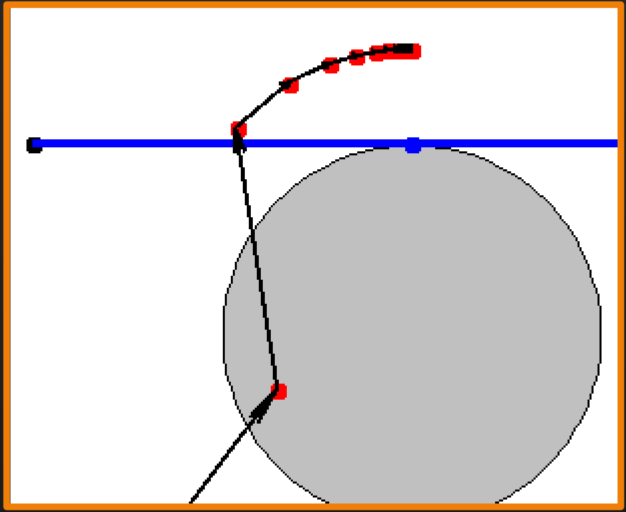
In a wide variety of problems (protein folding, 3SAT, Sudoku) the set B is non-convex but "divide and concur" works better than theory can explain. It is: $\boxed{R_A(x) := 2\,P_A(x) - x \text{ and } x \to \frac{x+R_A(R_B(x))}{2}}$

Consider the simplest case of a line **A** of height **h** and the unit circle **B**. With $\boxed{z_n := (x_n, y_n)}$ the iteration becomes

$$\boxed{x_{n+1} := \cos\theta_n, y_{n+1} := y_n + \alpha - \sin\theta_n, \qquad (\theta_n := \arg z_n)}$$

For h=0 we will prove convergence to one of the two points in A ∩ B iff we do not start on the vertical axis (where we have **chaos**). For h>1 (infeasible) it is easy to see the iterates go to infinity (vertically). For h=1 we converge to an infeasible point. For h in (0,1) the pictures are lovely but full proofs escape us. Two representative pictures follow:





An ideal problem to introduce early graduates to research, with many open accessible extensions in 2 or 3 dimensions

Interactive APR in Cinderella

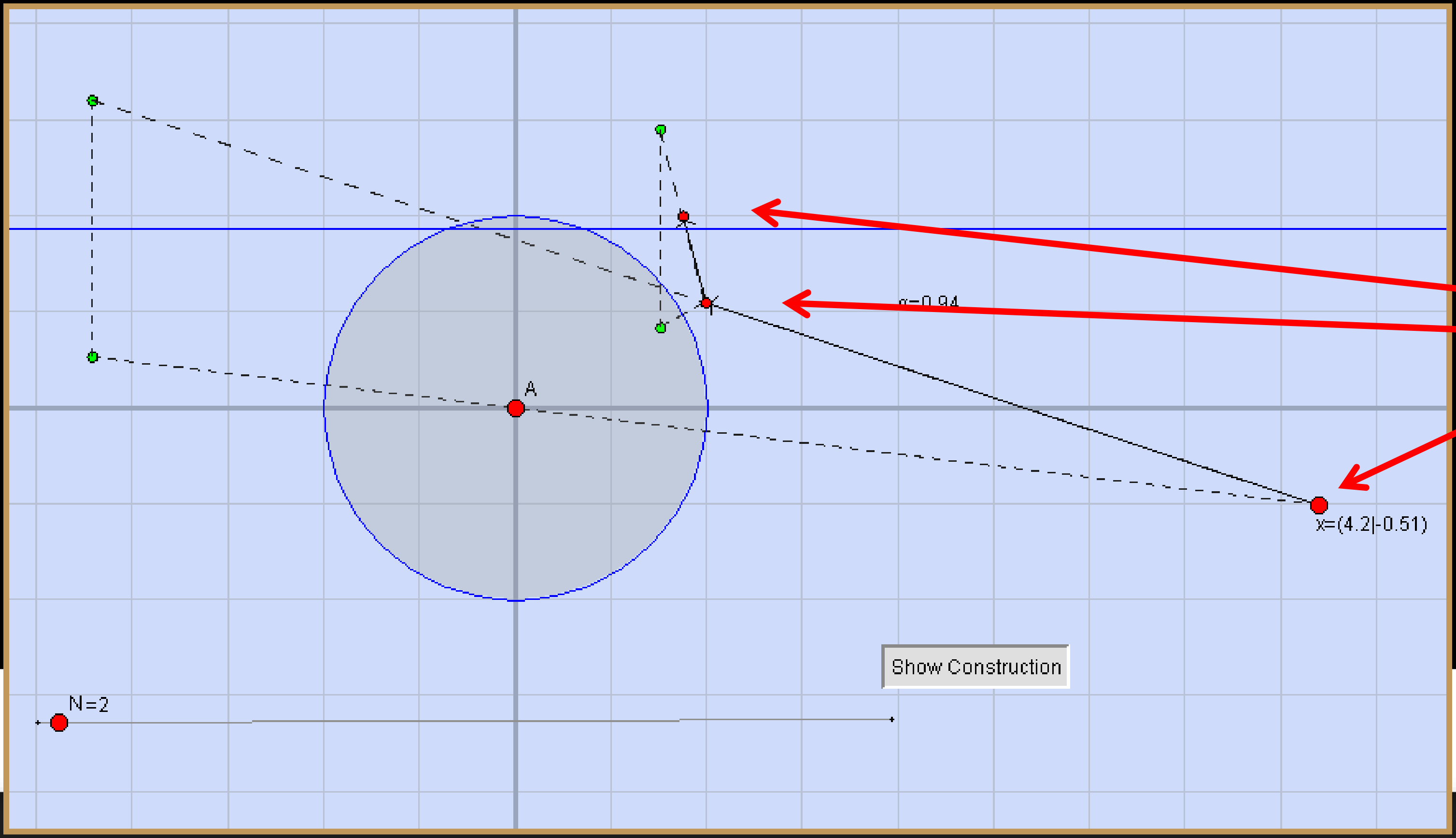
Recall the simplest case of a line A of height **h** and unit circle B.

With $z_n := (x_n, y_n)$

$$\boxed{x_{n+1} := \cos\theta_n, y_{n+1} := y_n + \alpha - \sin\theta_n, \quad (\theta_n := \arg z_n)}$$

The pictures are lovely but full proofs escape us.

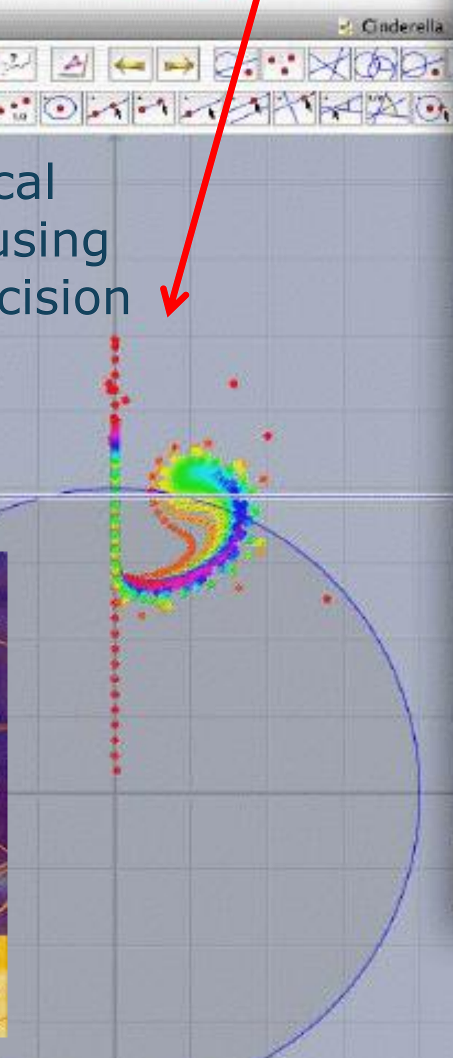
A *Cinderella* picture of two steps from (4.2,-0.51)



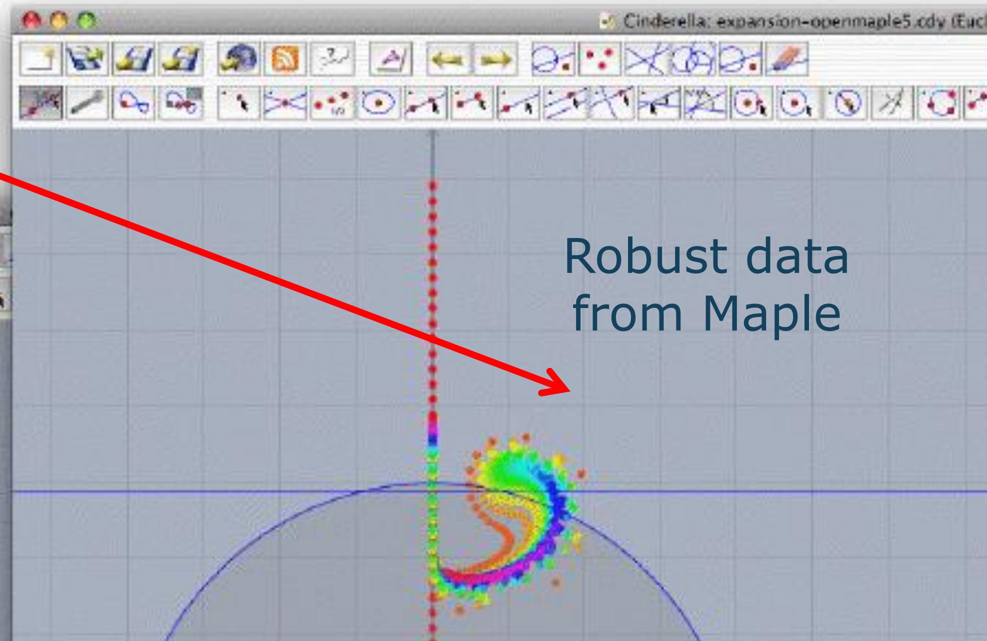
CAS+IGP: the Grief is in the GUI

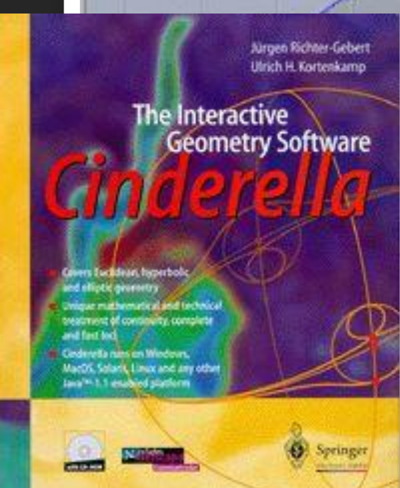
Divide –and-Concur
before and after accessing numerical output from Maple

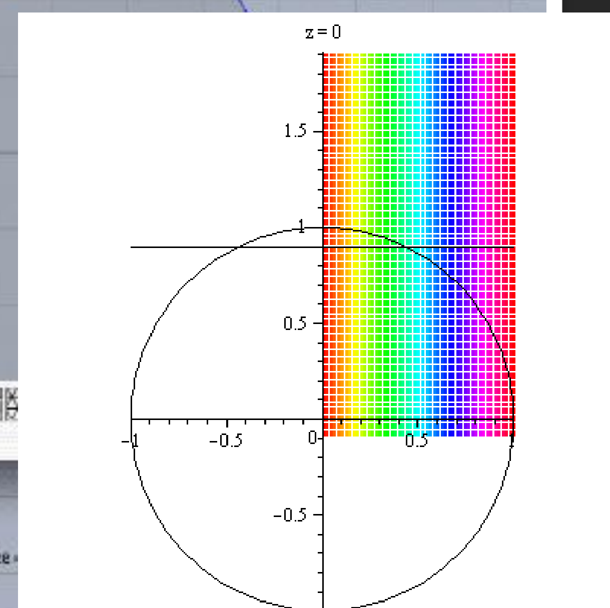
Numerical errors in using double precision



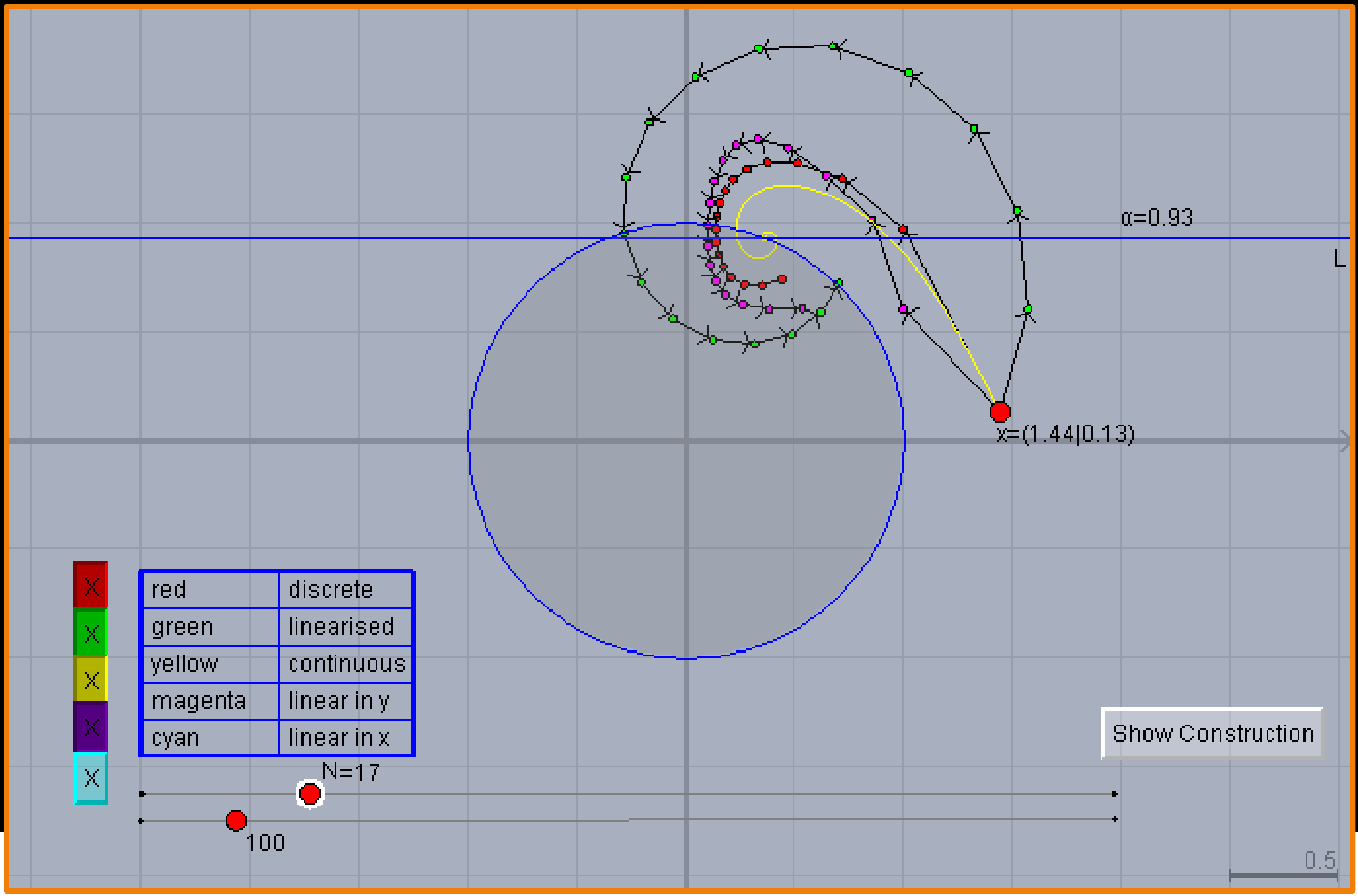
Robust data from Maple







Divide-and-concur and variations



(back)

We considered the analogous differential equation since asymptotic techniques for such differential equations are better developed.

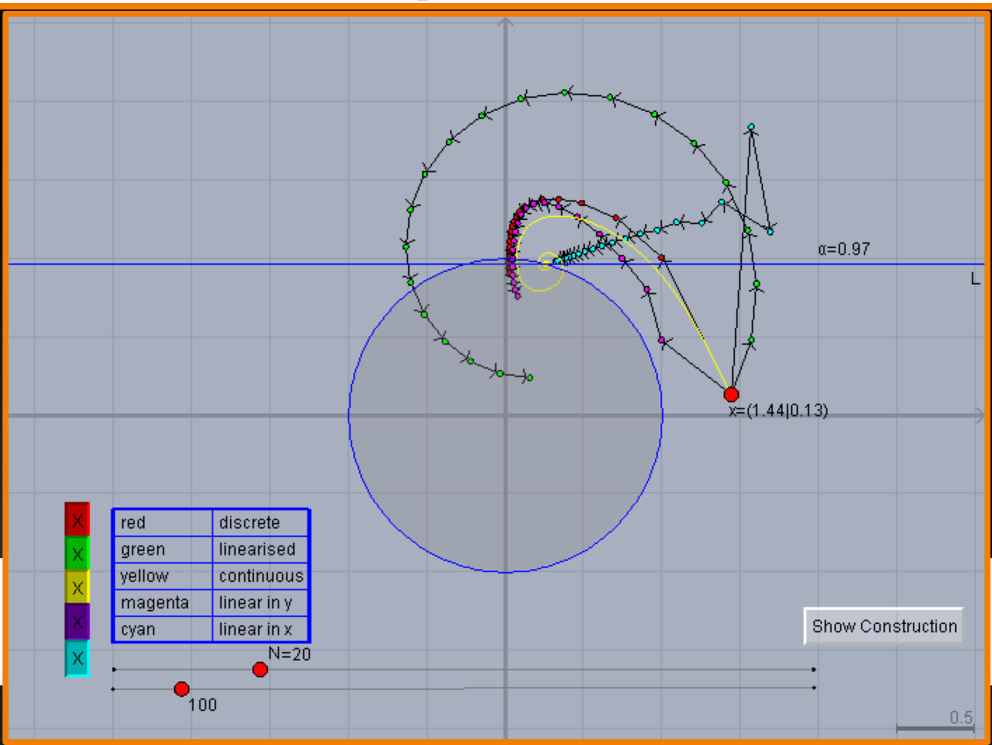
We decided

$$x'(t) = \frac{x(t)}{r(t)} - x(t) \quad [r(t) := \sqrt{x(t)^2 + y(t)^2}]$$

$$y'(t) = \alpha - \frac{y(t)}{r(t)}$$

is a reasonable counterpart to the Cartesian formulation–we replaced the difference $x_{n+1} - x_n$ by $x'(t)$, etc.— as shown in the picture below:

This convinced us that a local convergence result was possible



Perron's Theorem

Theorem 0.1 (Perron, see for example [2]). *If* $f : \boldsymbol{N} \times \boldsymbol{R}^m \longrightarrow \boldsymbol{R}^m$ *satisfies,*

$$\lim_{x \to 0} \frac{\|f(n,x)\|}{\|x\|} = 0,$$

uniformly in n *and* M *is a constant* $n \times n$ *matrix all of whose eigenvalues lie inside the unit disk, then the zero solution [provided it is an isolated solution] of the difference equation,*

$$x_{n+1} = Mx_n + f(n, x_n),$$

is exponentially asymptotically stable; *that is, there exists* $\delta > 0$, $K > 0$ *and* $\zeta \in (0,1)$ *such that if* $\|x_0\| < \delta$ *then* $\|x_n\| \leq K\|x_0\|\zeta^n$.

Sims and I apply this to our iteration at the intersection point: The Hessian is fine for |h| <1

$$\begin{bmatrix} h^2 & -h\sqrt{1-h^2} \\ h\sqrt{1-h^2} & h^2 \end{bmatrix}$$

Summary

Theorem Let H be a hyperplane of height h.
Let B the unit sphere in N-space.
Except for intial values in a lower dimensional subspace we have:
(a) For $h = 0$ reflect-reflect and average converges to a point in $H \cap B$.
(b) For $-1 < h < 1$ reflect-reflect and average converges locally.
(c) For $|h| = 1$ reflect-reflect and average converges to an infeasible point.
(d) For $|h| > 1$ reflect-reflect and average diverges to infinity.

- For $0< |h|<1$, why is convergence global?
- Does this analysis lift to a general convex set ?
- To a p-ball ($0<p<\infty$), ellipse, ...?

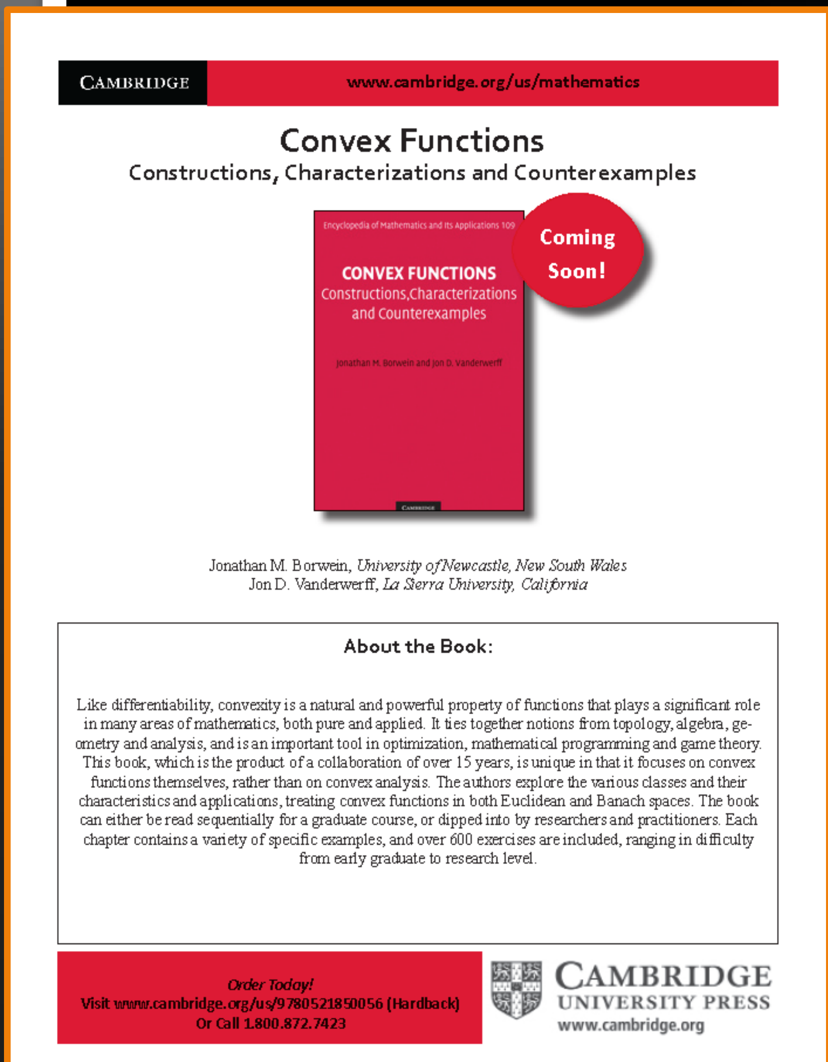
(back)

How to Maximize Surprise

(back)

Other General References

2009 **2005**


CAMBRIDGE www.cambridge.org/us/mathematics

Convex Functions
Constructions, Characterizations and Counterexamples

Encyclopedia of Mathematics and its Applications 109
CONVEX FUNCTIONS
Constructions, Characterizations and Counterexamples
Jonathan M. Borwein and Jon D. Vanderwerff

Coming Soon!

Jonathan M. Borwein, *University of Newcastle, New South Wales*
Jon D. Vanderwerff, *La Sierra University, California*

About the Book:

Like differentiability, convexity is a natural and powerful property of functions that plays a significant role in many areas of mathematics, both pure and applied. It ties together notions from topology, algebra, geometry and analysis, and is an important tool in optimization, mathematical programming and game theory. This book, which is the product of a collaboration of over 15 years, is unique in that it focuses on convex functions themselves, rather than on convex analysis. The authors explore the various classes and their characteristics and applications, treating convex functions in both Euclidean and Banach spaces. The book can either be read sequentially for a graduate course, or dipped into by researchers and practitioners. Each chapter contains a variety of specific examples, and over 600 exercises are included, ranging in difficulty from early graduate to research level.

Order Today!
Visit www.cambridge.org/us/9780521850056 (Hardback)
Or Call 1.800.872.7423

CAMBRIDGE UNIVERSITY PRESS
www.cambridge.org