Jonathan Borwein
Laureate Professor
and Director CARMA
www.carma.newcastle.edu.au

Why Convex: Some of My Favourite Convex Functions

PREPARED FOR
2009 ANNUAL CMS
WINTER MEETING
WINDSOR, ONT

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

"I never run for trains."
Nasim Nicholas Taleb (The Black Swan)

Revised 03-12-2009

Jonathan Borwein
Laureate Professor
and Director CARMA
www.carma.newcastle.edu.au

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

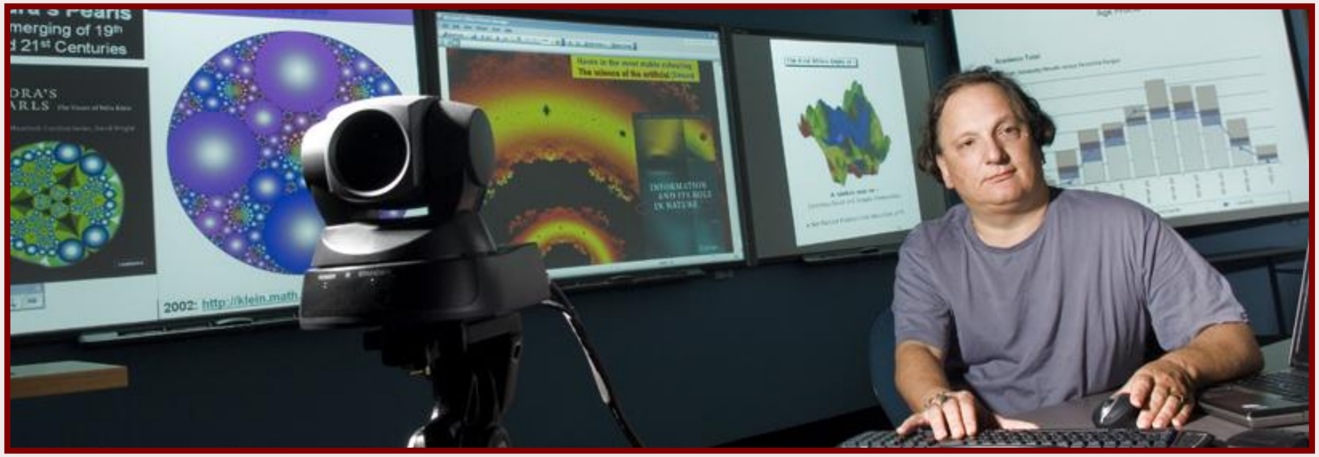
CMS-in-Oz

Adelaide
Finnur Larusson

Melbourne
Lynn Batten

Newcastle
Brian Alspach
Jon Borwein
Kathy Heinrich

Revised 03-12-2009



Newcastle
AMSI-AG
Room

Why Convex: Some of my Favourite Convex Functions

Jon Borwein, FRSC www.carma.newcastle.edu.au

Laureate Professor, Newcastle NSWb

"Harald Bohr is reported to have remarked

"Most analysts spend half their time hunting through the literature for inequalities they want to use, but cannot prove."

- D.J.H. Garling

Review of Michael Steele's ***The Cauchy Schwarz Master Class*** in MAA Monthly, June-July 2005, 575-579.

Also G.H. Hardy's ***A Proglemena to Inequalities***, Collected Works

Harald Bohr
1887-1951

CMS Plenary Lecture

December 5th 2009

Abstract of Convexity Talk, I

JONATHAN BORWEIN, University of Newcastle, NSW
Why Convex?

This lecture makes the case for the study of convex functions focussing on their structural properties. We highlight the centrality of convexity and give a selection of salient examples and applications.

It has been said that most of number theory devolves to the Cauchy-Schwarz inequality and the only problem is deciding ‘what to Cauchy with.’ In like fashion, much mathematics is tamed once one has found the right convex ‘Green's function.’

Why convex? Well, because ...

Abstract of Convexity Talk, II

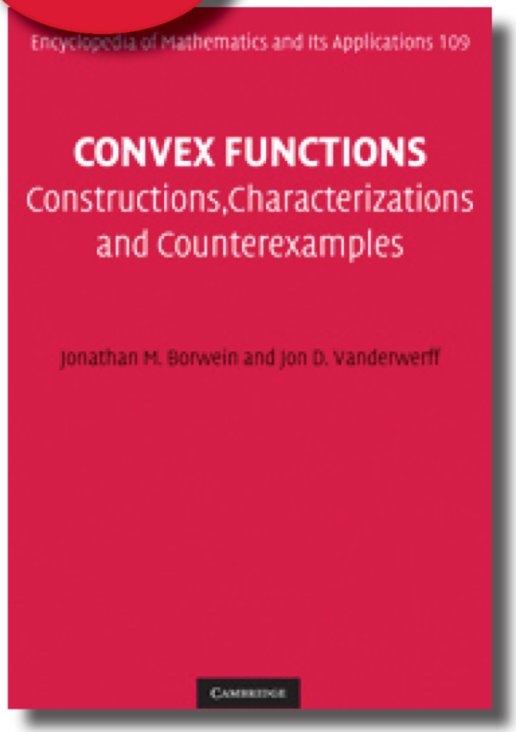
From Chapter 1 of *Convex Functions* (JMB and JDV, 2009)

The first modern formalization of the concept of convex function appears in J. L. W. V. Jensen, "Om konvexe funktioner og uligheder mellem midelvaerdier." Nyt Tidsskr. Math. B ***16*** *(1905), pp. 49–69. Since then, at first referring to "Jensen's convex functions," then more openly, without needing any explicit reference, the definition of convex function becomes a standard element in calculus handbooks.* (A. Guerraggio and E. Molho) Historia Mathematica 2004

Convexity theory ... reaches out in all directions with useful vigor. Why is this so? Surely any answer must take account of the tremendous impetus the subject has received from outside of mathematics, from such diverse fields as economics, agriculture, military planning, and flows in networks. With the invention of high-speed computers, large-scale problems from these fields became at least potentially solvable. Whole new areas of mathematics (game theory, linear and nonlinear programming, control theory) aimed at solving these problems appeared almost overnight. And in each of them, convexity theory turned out to be at the core. The result has been a tremendous spurt in interest in convexity theory and a host of new results. (A. Wayne Roberts and Dale E. Varberg, 1973) Klee **SIAM Rev 1976**

The Sum of What I know

November
2009



Key Features

- Unique focus on the functions themselves, rather than convex analysis
- Contains over 600 exercises showing theory and applications
- All material has been class-tested

Contents

November 2009
555 pages
10 tones / 640 exercises
50 worked examples / 45 figures
Hardback / 9780521850056
Price Is Not Yet Set

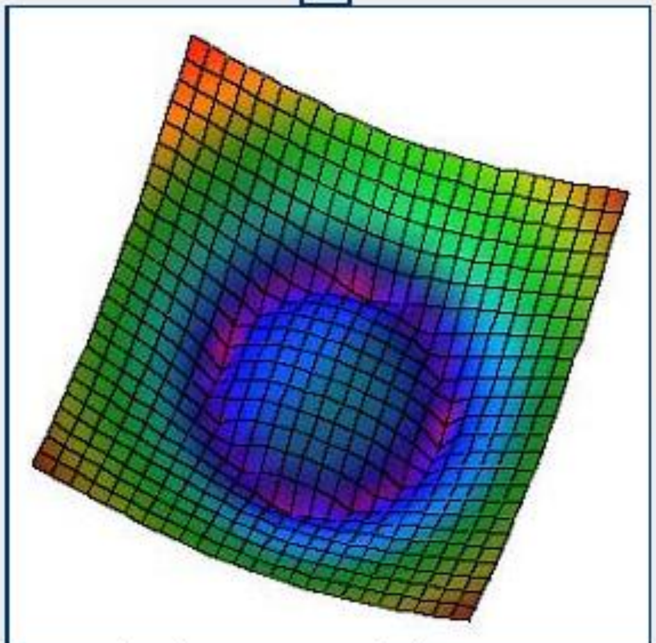
The Sum of What I know

Convex Functions

Web site

Sample | Order | Reviews | Addenda | Links | Contact | CUP

Jonathan Borwein, FRSC
University of Newcastle
and Dalhousie University

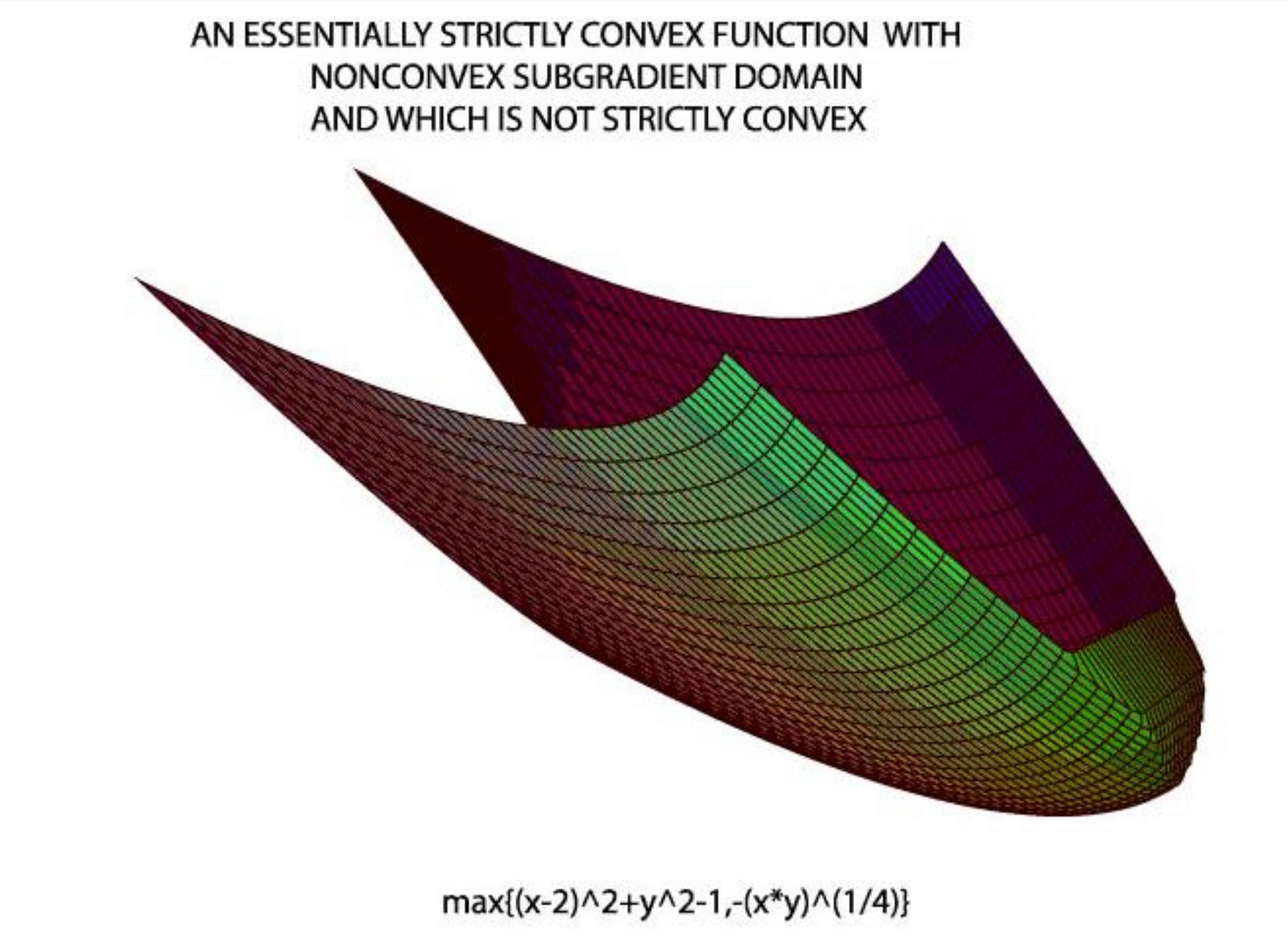


Coming in **2009** to a website near you.
Cambridge University Press *Encyclopedia of Mathematics and Applications* volume **109**, entitled
Convex Functions: Constructions, Characterizations and Counterexamples.
This book is intended for:

- Researchers, practitioners and students
- In computation, optimization, analysis
- With applications throughout the mathematical sciences.

Jon Vanderwerff
La Sierra University

Even Three Dimensions is Subtle



Abstract of Convexity Talk, III

I now offer a variety of examples of convexity appearing (often unexpectedly) in my research.

(Log) convex functions are not denatured. They are everywhere.

Each illustrates either the power of convexity, or of modern symbolic computation, or of both …

Principle of Uniform Boundedness

$$f_{\mathcal{A}}(x) := \sup_{A \in \mathcal{A}} \|A(x)\|$$

Proof. (i) f_A is convex and lower-semicontinuous as a supremum of such functions;

(ii) a **pointwise bounded** collection forces finiteness;

(iii) by Baire, f is continuous and so the linear operators are **uniformly bounded.** **QED**

Outline of Convexity Talk

A. Generalized Convexity of Volumes (Bohr-Mollerup, 1922).

B. Coupon Collecting and Convexity.

C. Convexity of Spectral Functions.

D. Characterizations of Banach space.

CARIMA

Volumes

The talk ends when I do
There are three bonus tracks!

Drive

Full details are in the reference texts and at http://projects.cs.dal.ca/ddrive/ConvexFunctions/ with some software

The Brothers Bohr

- One Nobel Prize
  - Nils (1885-1962)
  - Physics (1922)
- One Olympic Medal
  - Harald (**1887**-1951)
  - Soccer (1908)

1887-1920, 1887-1951, 1887-1985

Generalized Convexity of Volumes

A. Generalized Convexity of Gamma (Bohr-Mollerup, 1922).

Γ is usually defined for $\mathrm{Re}(x) > 0$ as

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1}\, dt. \qquad (1)$$

Theorem 1 (Bohr-Mollerup) *Γ is the unique function $f : (0,\infty) \to (0,\infty)$ such that:*
(a) $f(1) = 1$; (b) $f(x+1) = xf(x)$;
(c) f is log-convex ($x \to \log f(x)$ is convex).

- Application is often *automatable* in a computer algebra system, as I now illustrate:

Generalized Convexity of Volumes

A. Generalized Convexity of Gamma (Beta function).

The $\beta-$*function* is defined by

$$\beta(x,y) \;:=\; \int_0^1 t^{x-1}(1-t)^{y-1}\,dt \qquad (1)$$

for $\mathrm{Re}(x), \mathrm{Re}(y) > 0$. As is often established using polar coordinates and double integrals

$$\boxed{\beta(x,y) \;=\; \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}.} \qquad (2)$$

Proof (2) Use $f := x \to \beta(x,y)\,\Gamma(x+y)/\Gamma(y)$. (a) and (b) are easy. For (c) we show f is log-convex via Hölder's inequality. Thus $f = \Gamma$ as required. **QED**

- Γ is *hyper-transcendental* as is ζ.

Generalized Convexity of Volumes

A. Convexity of Volumes (Blaschke-Santalo inequality) (p-ball duality in Cinderella)

For a convex body C in R^n its *polar* is

$$C^\circ := \{y \in R^n : \langle y, x\rangle \leq 1 \text{ for all } x \in C\}.$$

Denoting n-dimensional Euclidean volume of $S \subseteq R^n$ by $V_n(S)$, **Blaschke-Santalo** says

$$V_n(C)\, V_n(C^\circ) \leq V_n(E)\, V_n(E^\circ) = V_n^2(B_n(2)) \tag{1}$$

where maximality holds (only) for *any* ellipsoid E and $B_n(2)$ is the Euclidean unit ball.

Question Explain cases of (1) as convexity estimates? Noting $B_p^\circ = B_q$ if $1/p + 1/q = 1$.

Generalized Convexity of Volumes

A. Convexity of Volumes (Dirichlet Formulae).

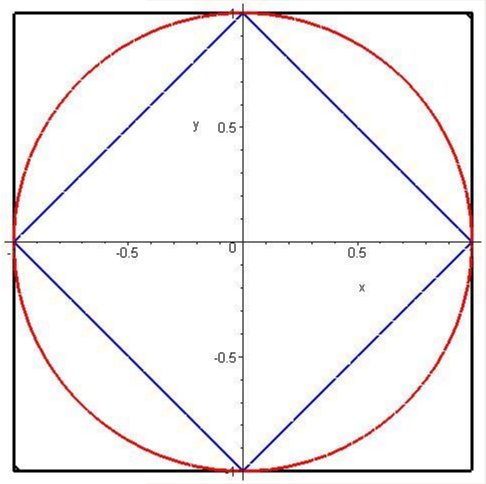
The volume of the ball in the $\|\cdot\|_p$-norm, $V_n(p)$, was first determined by Dirichlet

$$V_n(p) = 2^n \frac{\Gamma(1+\frac{1}{p})^n}{\Gamma(1+\frac{n}{p})}.$$

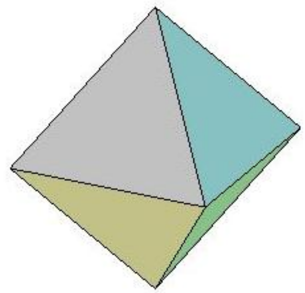
When $p = 2$,

$$V_n = 2^n \frac{\Gamma(\frac{3}{2})^n}{\Gamma(1+\frac{n}{2})} = \frac{\Gamma(\frac{1}{2})^n}{\Gamma(1+\frac{n}{2})},$$

is more concise than that usually recorded. *Maple* code derives this formula as an iterated integral for arbitrary p and fixed n.



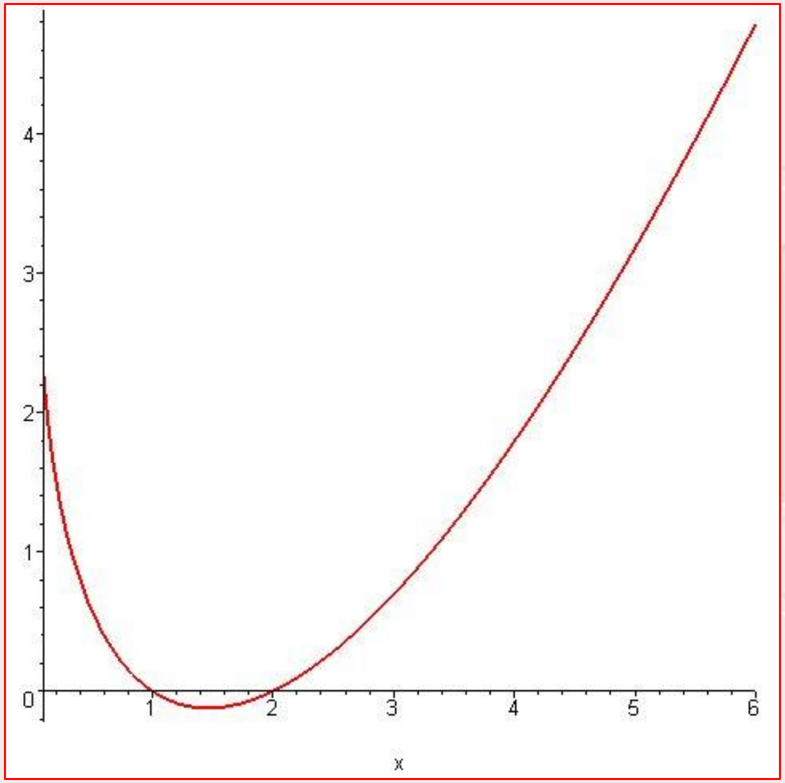
$1, 2, \infty$-balls in $\mathbb{R}^2$

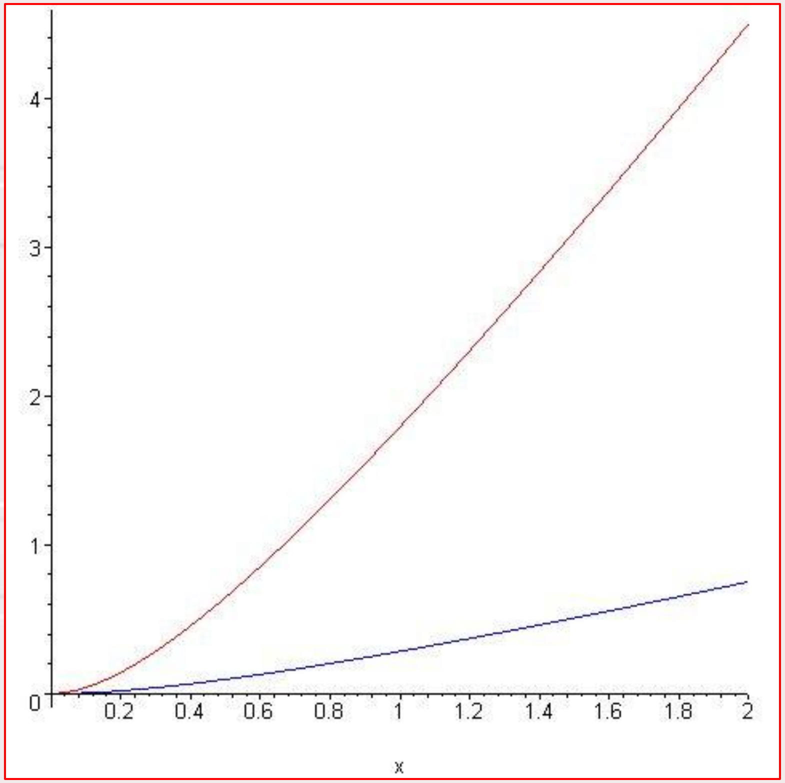


1-ball in R^3

Generalized Convexity of Volumes

A. Convexity of Volumes (Ease of Drawing Pictures).





$\log \Gamma(x)$ $\log V_a(1/x)$ for $a = 4/3, 3$

Discover the formula for $\sum_{n\geq 1} V_n(2)$

Generalized Convexity of Volumes

A. Convexity of Volumes ('mean' log-convexity). 2002

Theorem 2 [(H,A) log-concavity] *The function* $V_\alpha(p) := 2^\alpha \Gamma(1+\frac{1}{p})^\alpha / \Gamma(1+\frac{\alpha}{p})$ *satisfies*

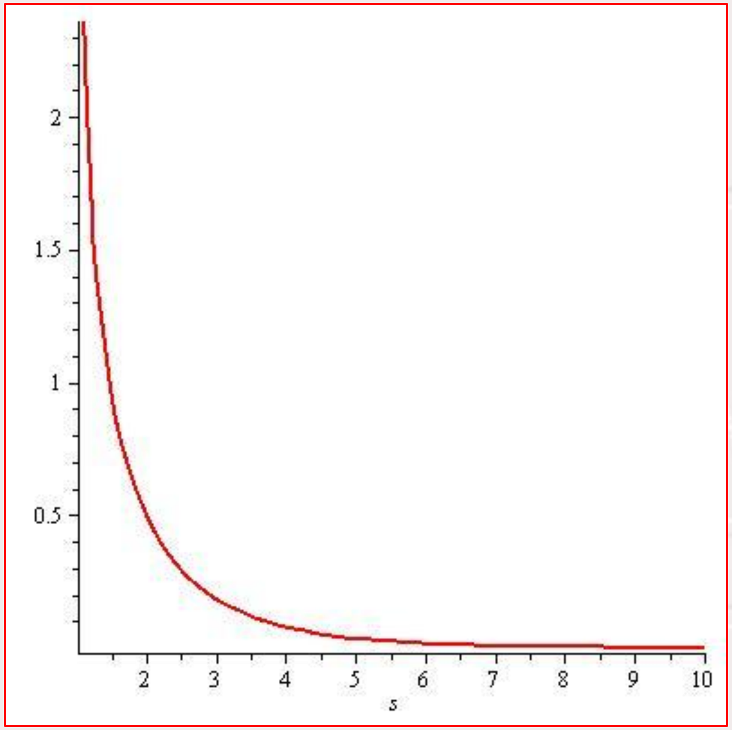
$$V_\alpha(p)^\lambda V_\alpha(q)^{1-\lambda} < V_\alpha\left(\frac{pq}{\lambda q + (1-\lambda)p}\right) \quad (1)$$

for all $\alpha > 1$*, if* $p, q > 1$*,* $p \neq q$*, and* $\lambda \in (0,1)$.

In (1) $\alpha = n, \frac{1}{p} + \frac{1}{q} = 1$ with $\lambda = 1 - \lambda = 1/2$ recovers the $p-$norm case of *Blaschke-Santalo*; and the *lower bound.* This extends to substitution norms. **Q.** How far can one take this?

Generalized Convexity of Zeta

(Ease of Drawing Pictures).



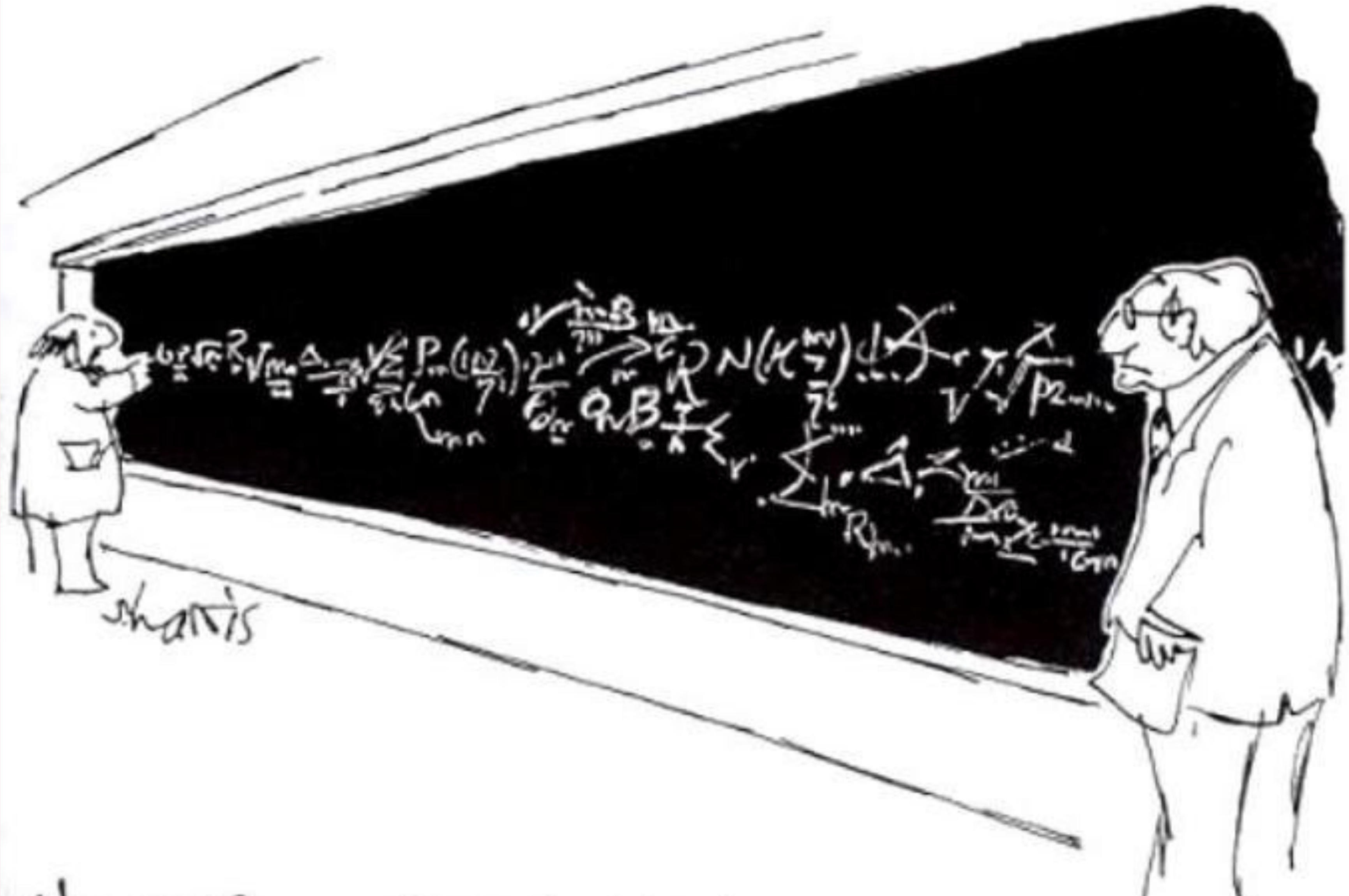
The **Euler product** shows

$$\log \zeta(x) = \sum_p \log(1 - e^{-x \log p})$$

is convex for $x > 1$
(p ranges over primes)

$\log \zeta(x)$ is convex

$\log(1 - e^{-x})$ has a nice Fenchel conjugate
$y \log y + (1-y)\log(1-y)$ (Fermi-Dirac entropy)



"HERE'S WHERE YOU MADE YOUR MISTAKE."

Outline of Convexity Talk

A. Generalized Convexity of Volumes (Bohr-Mollerup, 1922).

B. Coupon Collecting and Convexity.

C. Convexity of Spectral Functions.

D. Characterizations of Banach space.

Coupons

The talk ends when I do

Drive

CARMA

Coupon Collecting and Convexity

B. The origin of the problem.

Consider a network *objective function* p_N:

$$p_N(q) := \sum_{\sigma \in S_N} \left(\prod_{i=1}^{N} \frac{q_{\sigma(i)}}{\sum_{j=i}^{N} q_{\sigma(j)}} \right) \left(\sum_{i=1}^{N} \frac{1}{\sum_{j=i}^{N} q_{\sigma(j)}} \right),$$

summed over *all* $N!$ permutations; so a typical term is

$$\left(\prod_{i=1}^{N} \frac{q_i}{\sum_{j=i}^{N} q_j} \right) \left(\sum_{i=1}^{N} \frac{1}{\sum_{j=i}^{n} q_j} \right).$$

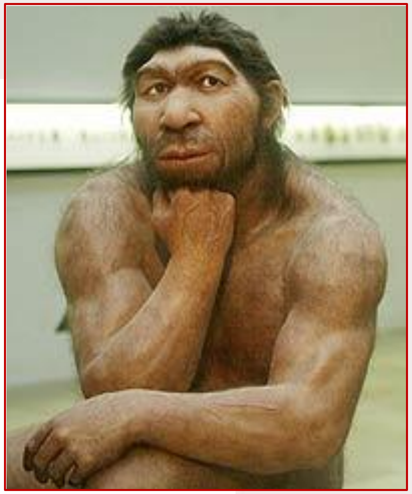
For example, with $N = 3$ this is

$$q_1 q_2 q_3 \left(\frac{1}{q_1 + q_2 + q_3} \right) \left(\frac{1}{q_2 + q_3} \right) \left(\frac{1}{q_3} \right) \left(\frac{1}{q_1 + q_2 + q_3} + \frac{1}{q_2 + q_3} + \frac{1}{q_3} \right).$$

This arose as the cost function in a **1999** PhD thesis on coupon collection. Ian Affleck wished to show p_N was convex on the positive orthant. I hoped not!

Coupon Collecting and Convexity

B. Doing What is Easy.



First, we try to simplify the expression for p_N. The *partial fraction decomposition* gives:

$$
\begin{aligned}
p_1(x_1) &= \frac{1}{x_1}, \\
p_2(x_1,x_2) &= \frac{1}{x_1}+\frac{1}{x_2}-\frac{1}{x_1+x_2}, \\
p_3(x_1,x_2,x_3) &= \frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}-\frac{1}{x_1+x_2}-\frac{1}{x_2+x_3}-\frac{1}{x_1+x_3} \\
&+ \frac{1}{x_1+x_2+x_3}. \qquad (1)
\end{aligned}
$$

Partial fractions are an arena in which computer algebra is hugely useful. Try performing the third case in (1) by hand. It is tempting to predict the "same" pattern will hold for $N = 4$. This is easy to confirm (by computer) and so we are led to:

A facet of Coxeter's favourite polyhedron

Coupon Collecting and Convexity

B. A Non-convex Integrand.

CONJECTURE. For each N, the function p_N given by

$$x \mapsto \int_0^1 \left\{ 1 - \prod_{k=1}^{N} \left(1 - t^{x_k}\right) \right\} \mathrm{d}t$$

is convex. Indeed $1/p_N$ is concave.

- Randomized numeric checks were run up to $N = 20$.
- $(N > 6)$ Computing the Hessian symbolically is impossible:
- Even just the diagonal will not fit on the largest *Maple*.

• a notationally efficient representation of no help with a proof

Coupon Collecting and Convexity

B. A Very Convex Integrand. (Is there a direct proof?)

A year later, Omar Hijab suggested re-expressing p_N as the joint expectation of Poisson distributions. This leads to:

If $x = (x_1, \cdots, x_n)$ is a point in the positive orthant R^n_+, then

$$p_N(x) = \left(\prod_{i=1}^{n} x_i\right) \int_{R^n_+} e^{-\langle x,y\rangle} \max(y_1, \cdots, y_n)\, dy$$

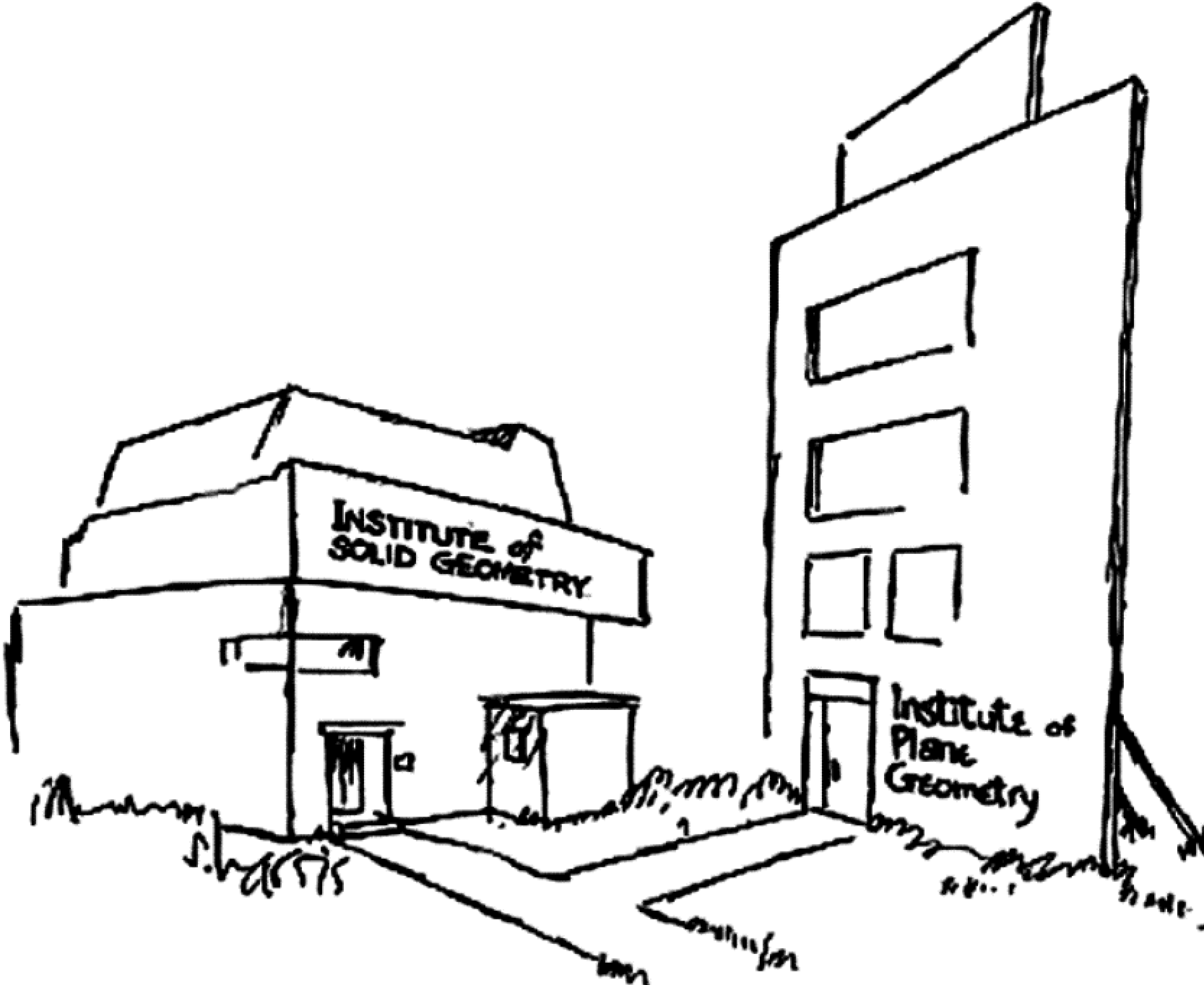
- $\langle x, y\rangle = x_1 y_1 + \cdots + x_n y_n$ is the inner product

Now $y_i \to x_i\, y_i$ and standard techniques show $1/p_N$ is concave, since the integrand is.[We can now ignore probability if we wish!]
Q "inclusion-exclusion" convexity? OK for $1/g(x) > 0$, g concave.

Goethe’s One Nice Comment About Us

“Mathematicians are a kind of Frenchmen:
whatever you say to them they translate into their own language, and right away it is something entirely different.”

(Johann Wolfgang von Goethe)
Maximen und Reflexionen, no. 1279


INSTITUTE of
SOLID GEOMETRY
Institute of
Plane
Geometry
S. HARRIS

Outline of Convexity Talk

A. Generalized Convexity of Volumes (Bohr-Mollerup).

B. Coupon Collecting and Convexity.

C. Convexity of Spectral Functions.

D. Characterizations of Banach space.

Spectra

The talk ends when I do

Convexity of Spectral Functions

C. Eigenvalues of symmetric matrices (Lewis (95) and Davis (59)).

λ(S) lists decreasingly the (real, resp. non-negative) eigenvalues of a (symmetric, resp. PSD) n-by-n matrix S. The Fenchel conjugate is the convex closed function given by

$$f^*(x) := \sup_y \langle y, x\rangle - f(y).$$

Theorem (Spectral conjugacy) If $f : R^n \mapsto (-\infty, \infty]$ is a symmetric function, it satisfies the formula $(f \circ \lambda)^* = f^* \circ \lambda$.

Fan $\mathrm{tr}(AB) \leq \lambda(A)^T \lambda(B)$

Corollary [Davis/Lewis] Suppose $f : R^n \mapsto (-\infty, \infty]$ is symmetric. The "spectral function" $f \circ \lambda$ is closed and convex (resp. differentiable) iff f is closed and convex (resp. differentiable). [Von Neumann for norms]

Also for trace class operators

Convexity of Spectral Functions

C. Three Amazing Examples (Lewis).

I. Log Determinant Let $\mathrm{lb}(x) := -\log(x_1 x_2 \cdots x_n)$ which is clearly symmetric and convex. The corresponding spectral function is $S \mapsto -\log\det(S)$.

II. Sum of Eigenvalues Ranging over permutations, let $f_k(x) := \max_\pi\{x_{\pi(1)} + x_{\pi(2)} + \cdots + x_{\pi(k)}\}$. This is clearly symmetric and convex. The corresponding spectral function is

$$\sigma_k(S) := \lambda_1(S) + \lambda_2(S) + \cdots \lambda_k(S).$$

In particular the largest eigenvalue, σ_1, is a continuous convex function of S and is differentiable if and only if the eigenvalue is simple.

Convexity of Spectral Functions

C. Three Amazing Examples (Lewis).

III. $k-$th Largest Eigenvalue The $k-$th largest eigenvalue may be written as

$$\mu_k(S) = \sigma_k(S) - \sigma_{k-1}(S).$$

In particular, this represents μ_k as the difference of two convex continuous, hence locally Lipschitz, functions of S and so *we discover the very difficult result* that for each k, $\mu_k(S)$ is a **locally Lipschitz function** of S.

$$\mathbf{N=3.}\ \lambda_2(A) = \mathrm{tr}(A) - \lambda_{\max}(A) - \lambda_{\min}(A)$$

- Hard analogues exist for *singular values, hyperbolic polynomials, Lie algebras*, etc.

Trace class operators

Convexity of Barrier Functions

C. A Fourth Amazing Example (Nesterov & Nemirovskii, 1993).

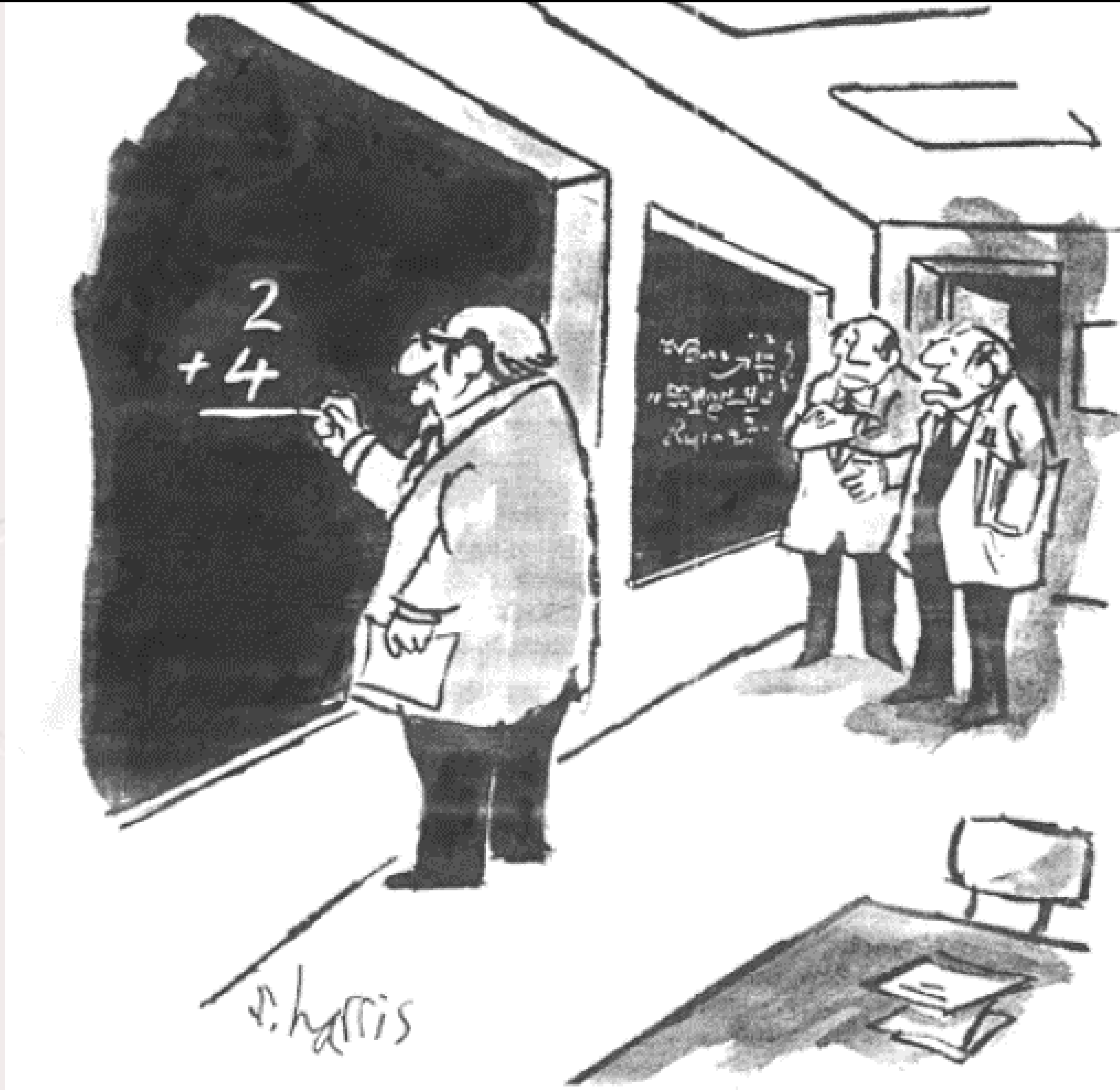
IV Self-concordant Barrier Functions Let A be a nonempty open convex set in R^N. Define, for $x \in A$,

$F_1(x)=|1/x-0|=1/x$

$$F_N(x) := \lambda_N((A-x)^\circ),$$

where λ_N is N-dimensional Lebesque measure and $(A-x)^o$ is the polar set. Then F_N is an essentially Fréchet smooth, log-convex, barrier function for A.

- Central to modern *interior point methods*.
- The orthant yields $\text{lb}(x) := -\sum_{k=1}^{N} \log x_k$.
- Hilbert space analog? (JB-JV, CUP, 2009)



"He was very big in Vienna."

Outline of Convexity Talk

A. Generalized Convexity of Volumes (Bohr-Mollerup).

B. Coupon Collecting and Convexity.

C. Convexity of Spectral Functions.

D. Characterizations of Banach Spaces

Characterizations

The talk ends when I do

Full details are in the three reference texts

D. Is not Madelung's Constant:
David Borwein CMS Career Award

$$= \sum_{n,m,p}{}' \frac{(-1)^{n+m+p}}{\sqrt{n^2+m^2+p^2}}$$

This polished solid silicon bronze sculpture is inspired by the work of David Borwein, his sons and colleagues, on the conditional series above for salt, Madelung's constant. This series can be summed to uncountably many constants; one is Madelung's constant for **electro-chemical stability of sodium chloride**. (**Convexity is hidden here too!**)

This constant is a period of an elliptic curve, a real surface in four dimensions. There are uncountably many ways to imagine that surface in three dimensions; one has negative gaussian curvature and is the tangible form of this sculpture. (As described by the artist.)

D. Characterizations

8

Convex functions and classifications of Banach spaces

A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between theories. (Stefan Banach)[1]

8.1 Canonical examples of convex functions

The first part of this chapter connects differentiability and boundedness properties of convex functions with respect to a *bornology* β (see p. 149 for the definition) with sequential convergence in the dual space in the topology of uniform convergence on the sets from the bornology. In some sense, many of the results in this chapter illustrate the degree to which linear topological properties carry over to convex functions. This chapter also examines extensions of convex functions that preserve continuity, as well as some related results.

Exemplars

Proposition 8.1.2. *Let X be a Banach space. Then the following are equivalent.*

(a) Mackey and norm convergence coincide sequentially in X^*.
(b) Every sequence of lsc convex functions that converges to a continuous affine function uniformly on weakly compact sets converges uniformly on bounded sets to the affine function.
(c) Every continuous convex function that is bounded on weakly compact subsets of X is bounded on bounded subsets of X.
(d) Weak Hadamard and Fréchet differentiability agree for continuous convex functions.

8.2 Characterizations of various classes of spaces

In this section we provide a listing of various classifications of Banach spaces in terms of properties of convex functions. Many of the implications follow from Theorem 8.1.3 or variants of the arguments upon which it is based. We will organize these results based upon when two of the following notions (Gâteaux, weak Hadamard or Fréchet) differentiability coincide for continuous convex functions on a space, and then for continuous weak*-lsc functions on the dual space. First we state the Josefson–Nissenzweig theorem proved independently by the two authors.

Theorem 8.2.1 (Josefson–Nissenzweig [271, 333]). *Suppose X is an infinite-dimensional Banach space, then there is a sequence $(x_n^*) \subset S_{X^*}$ that converges weak* to 0.*

Exemplars

First, we consider when Gâteaux and Fréchet differentiability coincide for continuous convex functions.

Theorem 8.2.2. *For a Banach space X, the following are equivalent.*

(a) X is finite-dimensional.
(b) Weak and norm convergence coincide sequentially in X^*.*
(c) Every continuous convex function on X is bounded on bounded subsets of X.
(d) Gâteaux and Fréchet differentiability coincide for continuous convex functions on X.

Basic idea: the convex $f(x) := \lim_{n\to\infty}\langle x_n^*, x\rangle$ captures the sequence (x_n^*).

A Banach space is said to have the *Dunford–Pettis property* if $\langle x_n^*, x_n\rangle \to 0$ whenever $x_n \to_w 0$ and $x_n^* \to_w 0$. The term DP*-property derives from the fact that weak convergence is replaced with weak* convergence in the dual sequence in the Dunford–Pettis property. Therefore, it follows immediately that a Banach space with the Grothendieck and Dunford–Pettis properties has the DP* property (but not conversely, e.g. ℓ_1). Consequently, the spaces $\ell_\infty(\Gamma)$ for any index set Γ have the DP*-property (see [184]).

Theorem 8.2.3. *For a Banach space X, the following are equivalent.*

(a) X has the DP-property.*
(b) Gâteaux and weak Hadamard differentiability coincide for all continuous convex functions on X.
(c) Every continuous convex function on X is bounded on weakly compact subsets of X.


THEORIES
PROOFS

Three Bonus Track Follows

A. Generalized Convexity of Volumes (Bohr-Mollerup).
B. Coupon Collecting and Convexity.
C. Convexity of Spectral Functions.
D. Characterizations of Banach space

E. Entropy and NMR.
F. Inequalities and the Maximum Principle.
G. Trefethen's 4th Digit-Challenge Problem.

References

Bonus

CARMA

Drive

REFERENCES

J.M. Borwein and D.H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century* A.K. Peters, 2003-2008.

J.M. Borwein, D.H. Bailey and R. Girgensohn, *Experimentation in Mathematics: Computational Paths to Discovery,* A.K. Peters, 2004. [Active CDs 2006]

J.M. Borwein and A.S. Lewis, *Convex Analysis and Nonlinear Optimization. Theory and Examples*, CMS-Springer, Second extended edition, 2005.

J.M. Borwein and J.D. Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, 2009.

Enigma

"The object of mathematical rigor is to sanction and legitimize the conquests of intuition, and there was never any other object for it."
- J. Hadamard quoted at length in E. Borel, *Lecons sur la theorie des fonctions*, 1928.

E. CONVEX CONJUGATES and NMR (MRI)

The *Hoch and Stern information measure* in complex N-space is $H(z) := \sum_{j=1}^{N} h(z_j/b)$ where h is convex and given (for scaling b) by

$$h(z) := |z| \ln\left(|z| + \sqrt{1+|z|^2}\right) - \sqrt{1+|z|^2}$$

for quantum theoretic (NMR) reasons. Recall the *Fenchel–Legendre conjugate*

$$f^*(y) = \sup_x \langle x, y\rangle - f(x).$$

Our symbolic convex analysis package produced

$$h^*(z) = \cosh(|z|).$$

Compare the *Shannon entropy* $z \ln(z) - z$ whose conjugate is $\exp(z)$.

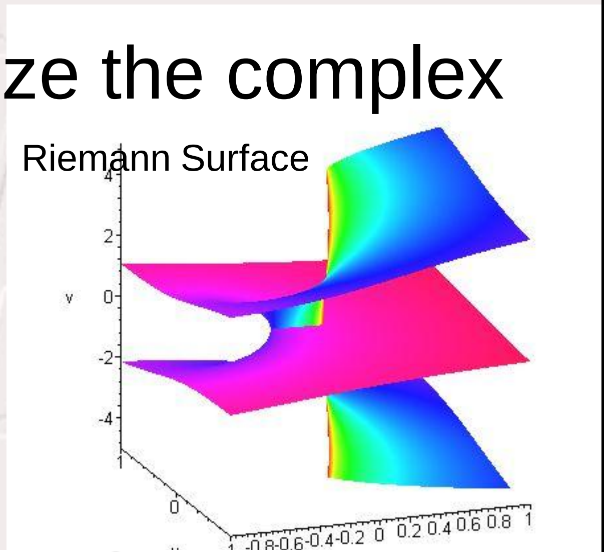
I'd never have tried by hand! Effective dual algorithms are now possible!

Knowing `Closed Forms' Helps

For example

$$(\exp\exp)^*(y) = y\ln(y) - y\{W(y) + W(y)^{-1}\}$$

where *Maple* or *Mathematica* recognize the complex *Lambert W function* given by



$$W(x)e^{W(x)} = x.$$

Thus, the conjugate's series is:

$$-1 + (\ln(y) - 1)\,y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \frac{3}{8}y^4 + \frac{8}{15}y^5 + O\left(y^6\right).$$

The literature is all in the last decade since W got a name!

WHAT is ENTROPY?

Despite the narrative force that the concept of entropy appears to evoke in everyday writing, in scientific writing entropy remains a thermodynamic quantity and a mathematical formula that numerically quantifies disorder. When the American scientist Claude Shannon found that the mathematical formula of Boltzmann defined a useful quantity in information theory, he hesitated to name this newly discovered quantity entropy because of its philosophical baggage. The mathematician John Von Neumann encouraged Shannon to go ahead with the name entropy, however, since “no one knows what entropy is, so in a debate you will always have the advantage."

The American Heritage Book of English Usage, p. 158

Information Theoretic Characterizations Abound

Theorem. Up to a positive scalar multiple

$$H(\overrightarrow{p}) = -\sum_{k=1}^{N} p_k \log p_k$$

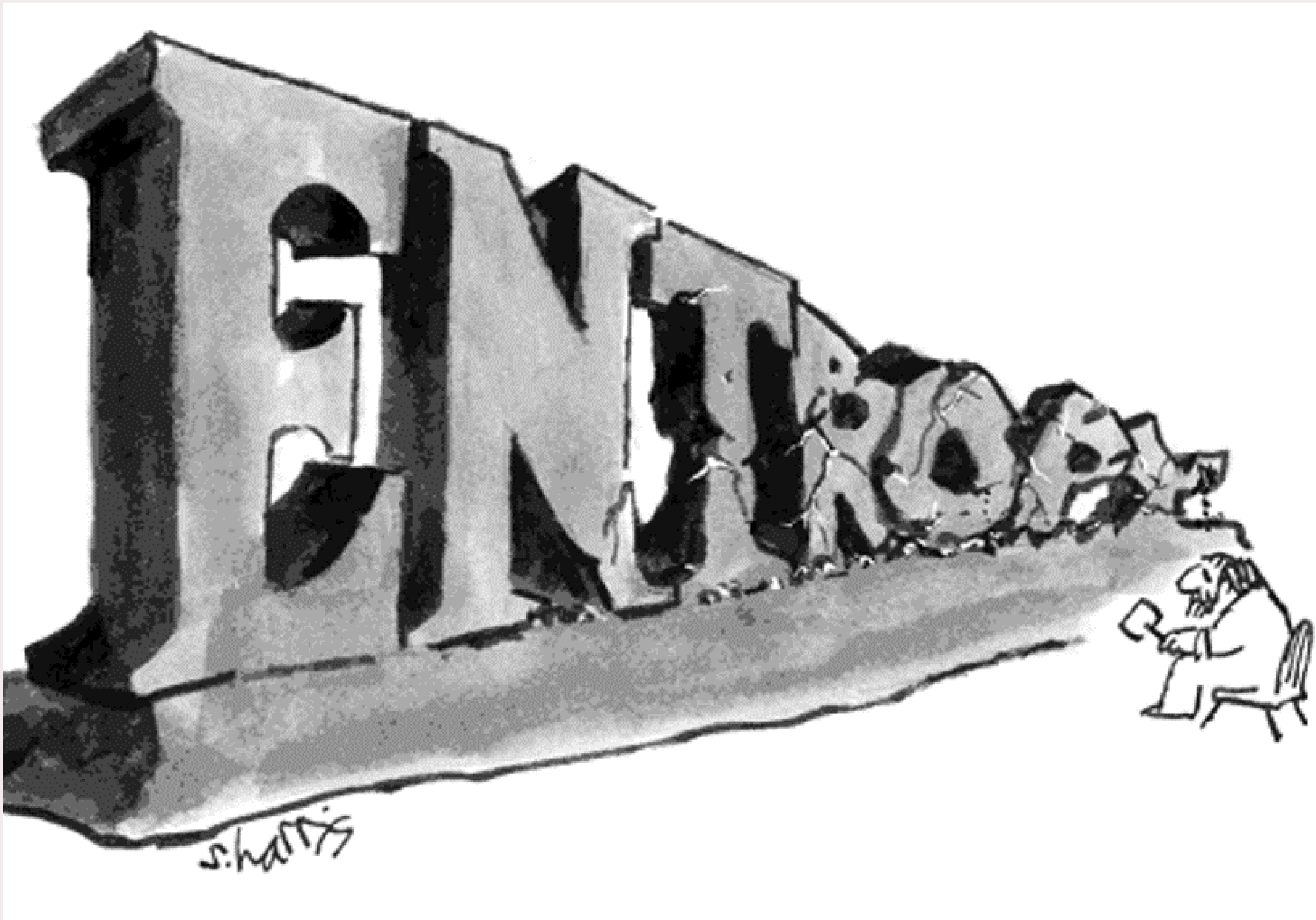
is the unique continuous function on finite probabilities such that [a.] **Uncertainly grows**:

$$H\left(\overbrace{\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}}^{n}\right)$$

increases with n.

[b.] **Subordinate choices are respected**: for distributions $\overrightarrow{p_1}$ and $\overrightarrow{p_2}$ and $0 < p < 1$,

$$H\left(p\,\overrightarrow{p_1}, (1-p)\,\overrightarrow{p_2}\right) = p\,H(\overrightarrow{p_1}) + (1-p)\,H(\overrightarrow{p_2}).$$


s.harris

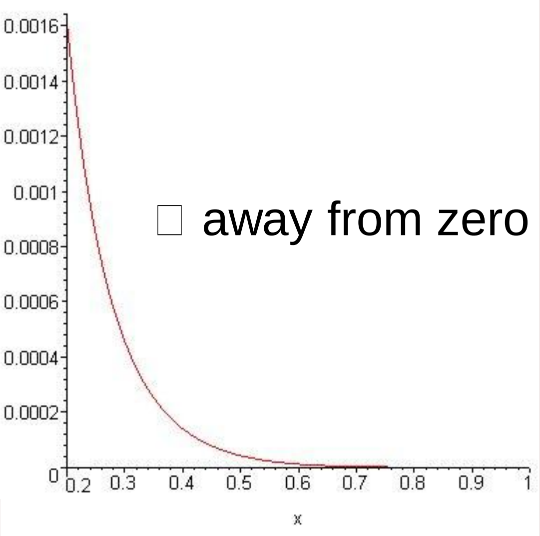
F. Inequalities and the Maximum Principle

- Consider the two *means*

$$\mathcal{L}^{-1}(x,y) := \frac{x-y}{\ln(x)-\ln(y)}$$

and

$$\mathcal{M}(x,y) := \sqrt[\frac{3}{2}]{\frac{x^{\frac{2}{3}}+y^{\frac{2}{3}}}{2}}$$



A conformal function estimated reduced to

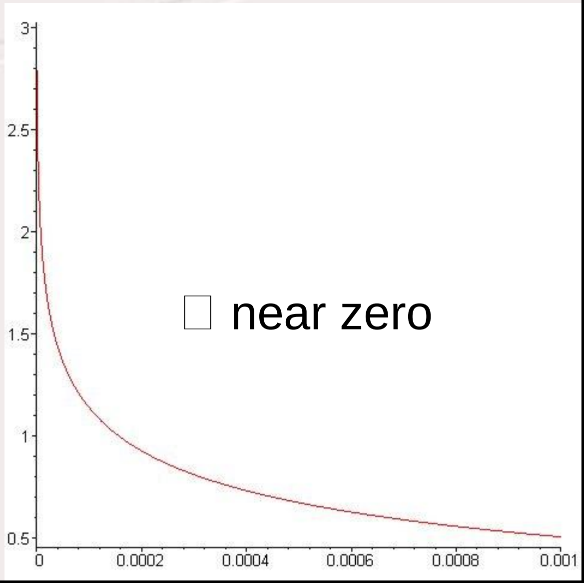
$$\mathcal{L}(\mathcal{M}(x,1),\sqrt{x}) > \mathcal{L}(x,1) > \mathcal{L}(\mathcal{M}(x,1),1)$$

for $0 < x < 1$.

tight

We first discuss showing

$$\mathcal{E}(x) := \mathcal{L}(\mathcal{M}(x,1),\sqrt{x}) - \mathcal{L}(x,1) > 0.$$



I. Numeric/Symbolic Methods

- $\lim_{x \to 0+} \mathcal{E}(x) = \infty$.

- *Newton-like iteration* shows that $\mathcal{E}(x) > 0$ on $[0.0, 0.9]$.

When we make each step effective.
This is hardest for the integral.

- *Taylor series* shows $\mathcal{E}(x)$ has 4 zeroes at 1.

$$= \frac{7}{51840}(x-1)^4 - \frac{7}{20736}(x-1)^5 + O\left((x-1)^6\right)$$

- *Maximum Principle* shows there are no more zeroes inside $C := \{z : |z-1| = \frac{1}{4}\}$:

$$\frac{1}{2\pi i}\int_C \frac{\mathcal{E}'}{\mathcal{E}} = \#(\mathcal{E}^{-1}(0); C)$$

Drive

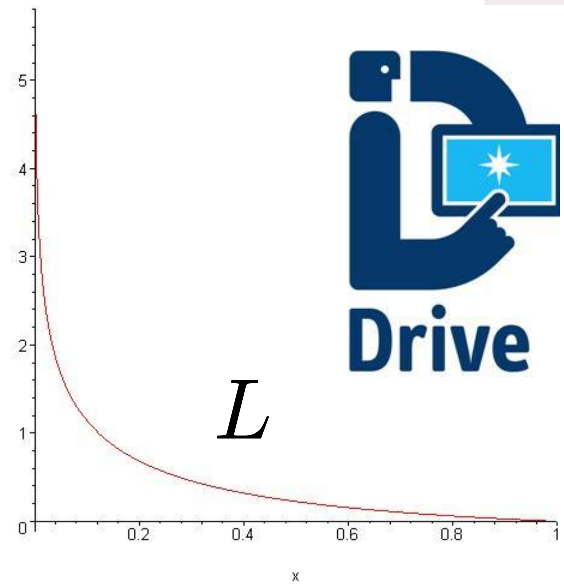
II. Graphic/Symbolic Methods

Consider the opposite (cruder) inequality

$$\Lambda := \mathcal{L}(x,1) - \mathcal{L}(\mathcal{M}(x,1),1) > 0.$$

We may observe that it holds since:

- $\mathcal{M}$ is a mean;
- $\mathcal{L}(x,1)$ decreases with x.
- There is an algorithm (Collins) for universal algebraic inequalities.

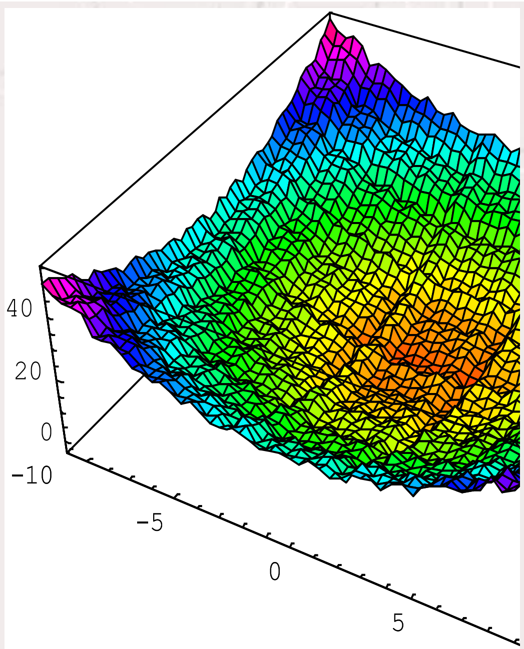


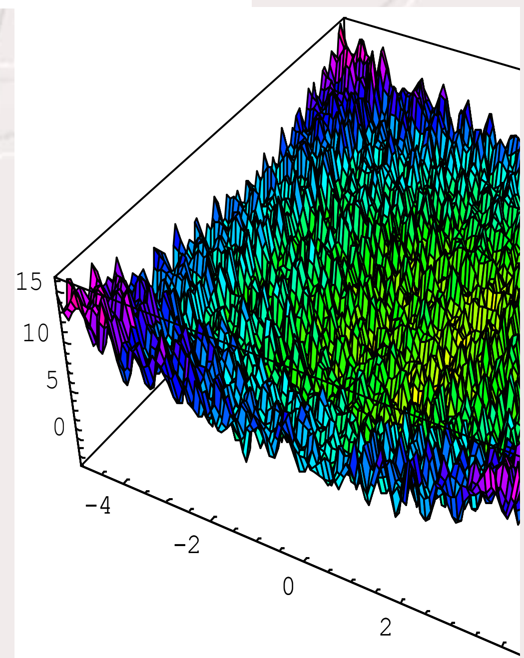
F. Nick Trefethen's 100 Digit/100 Dollar Challenge, Problem 4 (SIAM News, 2002)

 # 4. What is the global minimum of the function

$$\exp(\sin(50x)) + \sin(60e^y) + \sin(70\sin x)$$
$$+\sin(\sin(80y)) - \sin(10(x+y)) + (x^2+y^2)/4?$$

- no bounds are given.





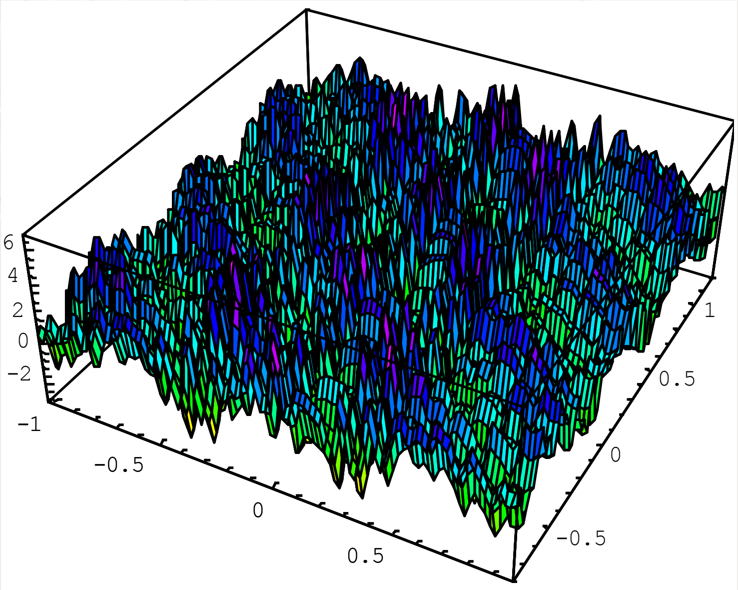
… HDHD Challenge, Problem 4

- This model has been numerically solved by LGO, MathOptimizer, MathOptimizer Pro, TOMLAB /LGO, and the *Maple GOT* (by Janos Pinter who provide the pictures).
- The solution found agrees to 10 places with the announced solution (the latter was originally based (provably) on a huge grid sampling effort, interval analyisis and local search).

$x^* \sim (-0.024627\ldots, 0.211789\ldots)$

$f^* \sim -3.30687\ldots$

Close-up picture near global solution: the problem still looks rather difficult … *Mathematica 6* can solve this by “zooming”!



See lovely SIAM solution book by Bornemann, Laurie, Wagon and Waldvogel and my Intelligencer Review at http://users.cs.dal.ca/~jborwein/digits.pdf