Fifty Years of Maximal Monotonicity:

Recent Progress on Maximal Monotonicity



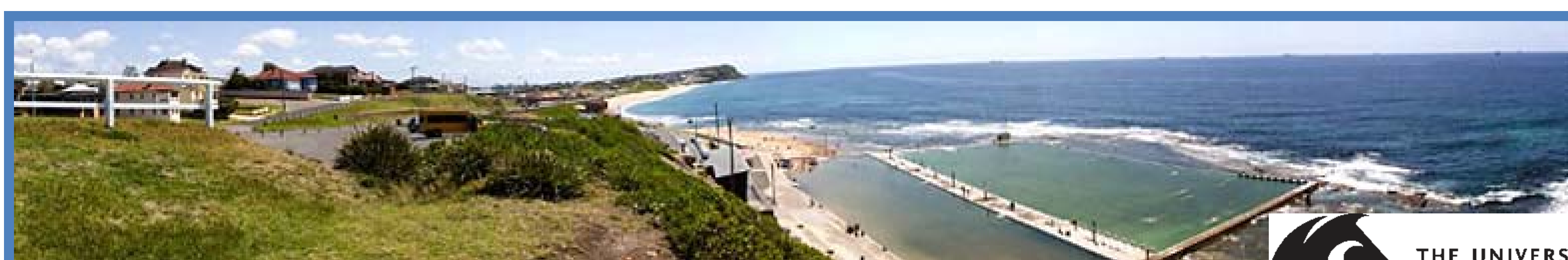
January 29 - Feb 2, 2012 Warrnambool, Victoria.

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

Laureate Professor University of Newcastle, NSW

Director, Centre for **Computer Assisted Research Mathematics and Applications**

CARMA



THE UNIVERSITY OF NEWCASTLE AUSTRALIA

Revised 15-01-12

CARMA

ABSTRACT

Jonathan M. Borwein

Director

Newcastle Centre for

Computer Assisted Research Mathematics and its Applications (CARMA)

Maximal(ly) monotone operator theory has just has turned fifty. I intend to briefly survey the history of the subject.

I shall try to explain why maximal monotone operators are both interesting and important--- culminating with a description of the remarkable progress made during the past decade.

See "Fifty years of maximal monotonicity," *Optimization Letters*, **4** (2010), 473-490.

"*The object of mathematical rigor is to sanction and legitimize the conquests of intuition, and there was never any other object for it*." – Jacques Hadamard

OUTLINE

- Background
- Five + decades of monotonicity (my highlights)
  - 1959-1970
  - 1970-1980
  - 1980-1990
  - 1990-2000
  - 2000-2012
- What we now know
- And still don't know
- Stories first
- Mathematics last

Monotone networks*

By G. J. Minty† Proc R. Soc. Lond. 1960

Department of Mathematics, Duke University, Durham, N.C., U.S.A.

(*Communicated by J. L. Synge, F.R.S.—Received* 12 *June* 1959—*Revised* 4 *February* 1960)

Fundamental existence and uniqueness theorems for electrical networks of non-linear resistors are proved in an abstract form, as theorems of pure mathematics. The two groups from which the 'currents' and 'voltage drops' are drawn are permitted to be either the real numbers, or discrete subgroups of the reals. It is found that the uniqueness theory is derivable from extremum principles for certain convex functions associated with the networks, and that the existence theory is derivable from a single new theorem of graph theory.

The abstract approach, besides revealing the logical structure of the subject more clearly than the 'concrete' approach, also (1) reveals the mathematical problem of solving a non-linear network to be identical with certain extremum problems arising in non-electrical applications, (2) contributes a numerical method, since the constructions for the discrete case are algorithmic, and (3) permits the application of the theorems to problems of pure mathematics.

Applications are not fully discussed; they will be treated at greater length in the appropriate technical journals.

1. Introduction

A well-known theorem, proved incorrectly by Poincaré (1901) and correctly by

"*in after years I have deeply regretted that I did not proceed far enough at least to understand something of the great leading principles of mathematics, **for men thus endowed seem to have an extra sense.***" - Charles Darwin

BACKGROUND

- First introduced by George Minty for networks
  - then in a PDE setting by Felix Browder and his school
  - and then for subgradients, variational inequalities, algorithms, mathematical economics ...
  - By **1975**, the main ideas were clear – if not easy – in Hilbert (reflexive Banach) space

A **monotone operator** from a Hausdorff locally convex space E to its dual E^* is a subset T of $E \times E^*$ such that $\langle x^* - y^*, x - y \rangle \geq 0$ for all (x^*, x) and (y^*, y) from T, where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between E and E^*.

"*He [Gauss] is like the fox, who effaces his tracks in the sand with his tail.*" - Niels Abel, 1802-1829

THREE CORE EXAMPLES

- SUBGRADIENTS (captures Banach space geometry)

$$\partial f(x) = \{x^* \in X^* \colon \langle x^*, y - x\rangle \le f(y) - f(x), \forall y \in X\}$$

$$J_X(x) := \frac{1}{2}\partial\|x\|^2 = \{x^* \in X^* \colon \|x\|^2 = \|x\|^2 = \langle x, x^*\rangle\}.$$

- SKEW LINEAR OPERATORS (have no ∂ part: "acyclic")

$$\langle Sx, y\rangle = -\langle Sy, x\rangle$$

$$\begin{bmatrix} 0 & A \\ -A^T & 0 \end{bmatrix}$$

All LP pairs become skew VIs

The non-cyclic rotation

$$S_\theta := \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

is only n-monotone for $\theta := \pi/n$.

  – these are in some sense the *extreme* cases (Asplund)

- LAPLACIANS (p-Laplacians, Elliptic PDEs)
  – *weak solutions* in appropriate Sobolev space, say, to Dirichlet's equation:

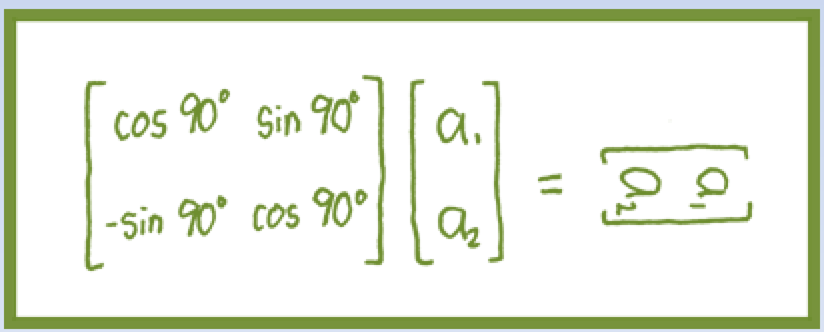
A NONLINEAR ROTATION OPERATOR

Theorem [Wiersma-JMB, 07] Define $S : \mathbf{R}^2 \to \mathbf{R}^2$ by $S(x, y) = (-y, x)$ for $x^2 + y^2 \leq 1$.
Then the unique maximal monotone extension $\widehat{S}$ of S with range restricted to the unit disc is **acyclic** and has:

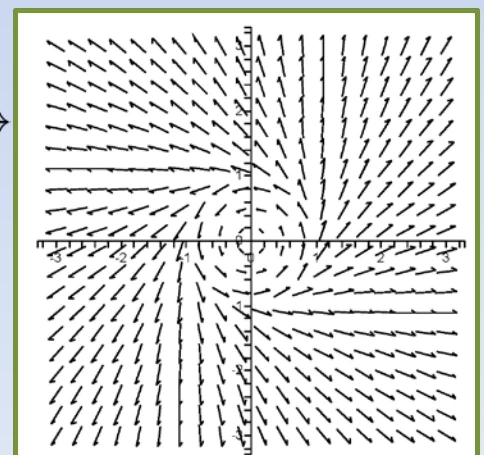
$$\widehat{S}(x) = \sqrt{1 - \frac{1}{\|x\|^2}}\frac{x}{\|x\|} + \frac{1}{\|x\|}S\left(\frac{x}{\|x\|}\right)$$

for $\|x\| \geq 1$.

Asplund's *Pareto* result:
can decompose $T = \partial f + \mathcal{A}$
where $\mathcal{A}$ has no more cyclic part.



S and $\widehat{S}$ $\rightarrow$



2009 all minimal monotone radially symmetric mappings are acyclic (Musev-Ribarska)

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 $|a_i(x,u)| \leq \nu\ (1 \leq i \leq n)$ and $|a_0(x,u)| \leq \nu|u| + a(x)$ 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

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

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 $0 < \lambda \leq 1$ 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 $2 \leq p < \infty$ the *p-Laplacian* $\triangle_p$ is maximal monotone: $\triangle_p u$ is given by

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

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

maximality

monotonicity

Key Advances: 1959-1970

- **Minty, Browder-Brezis-Hess, Asplund, Rockafellar, Zarantonello**, etc., lay down foundations
  - *Minty surjectivity theorem* R(T+J_X=X^*) if X is reflexive (**1962** Hilbert) [and J_X is 1-1 with 1-1 inverse] (proof at end)
  - application to **variational inequalities** (VI)

$$0 \in T(x) + \partial\delta_C(x) \Leftrightarrow \exists x^* \in T(x), x \in C, \sup_{c \in C}\langle x^*, c - x\rangle \leq 0.$$

  - sol'n techniques for (VI) or elliptic PDEs (Galerkin approximates) demand idea of maximality or at least "demiclosedness"

$$x_n^* \in T(x_n), x_n^* \to_s 0, x_n \to_w x \Rightarrow 0 \in T(x)$$

Minty, G., On some aspects of the theory of monotone operators. 1969 *Theory and Appl. of Monotone Operators* (Proc. NATO Adv. Study Inst., Venice, 1968) 67-82. "This is mainly a review article on monotone operators: the author points out their relation to Kirszbraun's theorem and to convexity theory, and sketches applications to Hammerstein's equation and to variational inequalities."

1959-1970

What reflexivity buys you is coercivity

- characterizes **Minty surjectivity theorem** (James theorem)
- ensures **resolvent** and **Yosida approximant** everywhere defined

$$R_\lambda := (T + \lambda J_X)^{-1}, \qquad T_\lambda := \left(T^{-1} + \lambda J_X^{-1}\right)^{-1}.$$

What Hilbert space (CAT(0)) buys you is nonexpansivity

- $\mathbf{J_H = I}$ (J_X and $(J_X)^{-1}$ are both smooth only on H)
- $\mathbf{T_\lambda}$ is non-expansive; indeed $$(x,y) \in P \Leftrightarrow \left(\frac{x-y}{2}, \frac{x+y}{2}\right) \in T$$
- T is **monotone** iff P is **non-expansive** (singleton)
- **proximal point** (below) and **Krasnoselskii algorithms** equivalent in H
- **Valentine-Kirzbraun theorem.** (1945-1932) *Every non expansive* P on A ⊂ H *extends to a nonexpansive* P^* *with* $D(P^*) = H$.

Quick Proof: $P \mapsto T \mapsto T^* \mapsto P^*$ (T^* any maximal extension of T).
Now $D(P^*) = R(T^*+I) = H$ by Minty's theorem.

Key Advances: 1970-1980

- **Rockafellar**, R. T., Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization* **14** (1976), 877-898.
  - On the maximality of sums of nonlinear monotone operators. *Trans. Amer. Math. Soc.* **149** (1970), 75-88. (highly technical: renorming, Brouwer, local bdd)
  - On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.* **33** (1970), 209-216. (far from easy—until recently)
  - (Maximal) monotonicity is hardly touched in *Convex Analysis* (1970)
- **Kenderov**, P., Semi-continuity of set-valued monotone mappings. *Fund. Math.* **88** (1975), 61--69. (early use for convex differentiability)
  - "*So, under some conditions, every maximal set-valued monotone mapping is single-valued almost everywhere*."
- **Gossez**, J.-P., Opérateurs monotones non linéaires dans les espaces de Banach non réflexifs. *J. Math. Anal. Appl.* **34** (1971), 371--376.
  - introduced **dense type** (all ∂f and all reflexive max. mon.) and lifted part of theory from reflexive space (despite complete failure of $R(T+J)=X^*$)

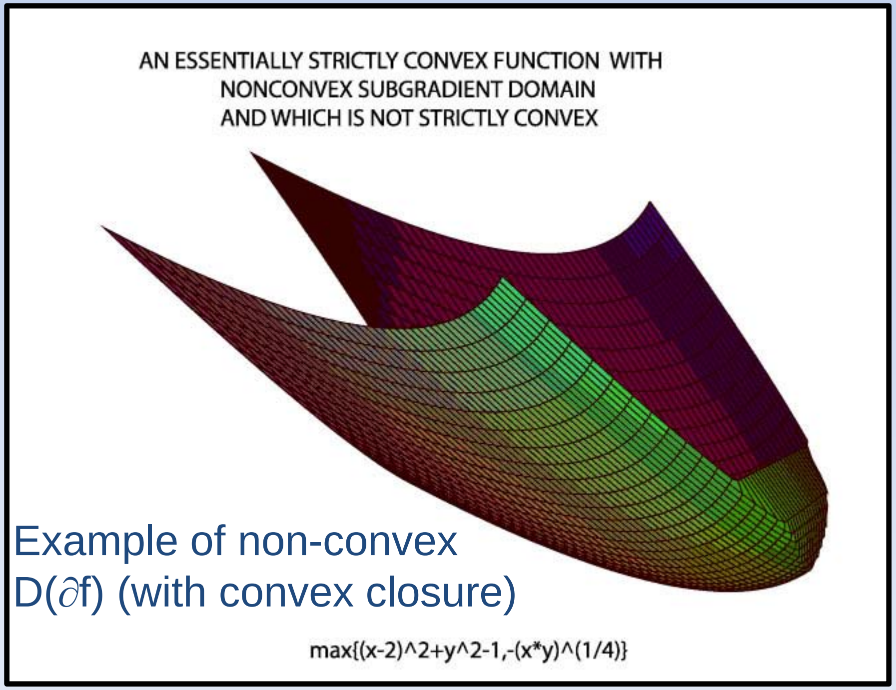
Examples : 1974 Gossez operator (a *non-dense type* linear max.mon.)
1976 *non-uniqueness* of extensions **1977** *non-convexity* of range of a max mon

1970-1980

- **Mignot, F.,** Contrôle dans les inéquations variationelles elliptiques. *J. Functional Analysis* **22** (1976), 130--185.
  - shows in finite dimensions that a maximal monotone operator is a.e. *(Fréchet) differentiable*
  - this now provides the canonical proof of Alexandrov's theorem and is central to viscosity solution for PDEs
  - It is vaguely possible that it holds in separable Hilbert space

1972 JMB writes MSc on max mon operators by a happy mistake:

When his supervisor recommends wrong paper of Mosco's

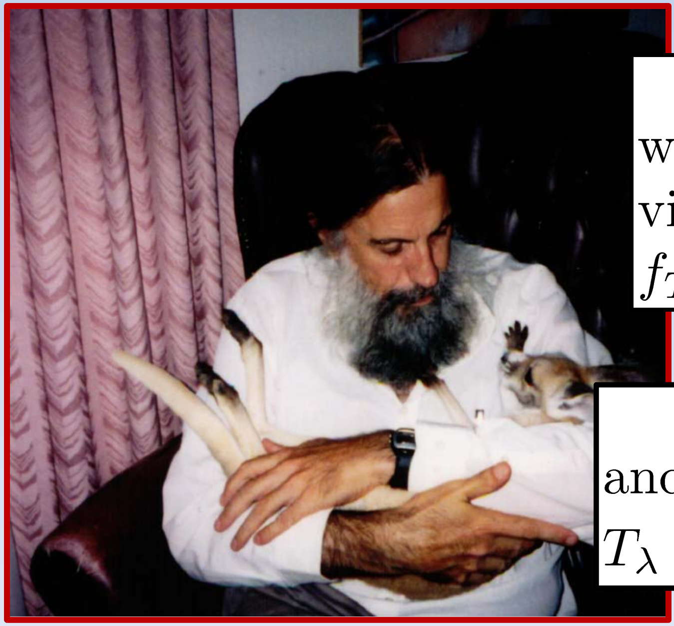


Key advances: 1980-1990

- **Fitzpatrick**, Simon, Representing monotone operators by convex functions. *Workshop/Miniconference on Functional Analysis and Optimization* (Canberra, 1988), 59-65.

"In an earlier work E. Krauss (1985) found a representation of monotone operators with the help of subdifferentials of saddle functions on $E \times E$. In the paper under review the author studies a monotone operator $T \subset (E \times E^*)$ by using the **convex function** $L_T\ E \times E^* \to \mathbf{R} \cup \{+\infty\}$ defined by

$$L_T(x, x^*) = \sup\{\langle x^*, y\rangle + \langle y^*, x - y\rangle : (y, y^*) \in T\}."$$



Brezis earlier used L_T without final term (oops) while JMB-SF proved local boundedness on int $D(T)$ via continuity of the *convex* function $f_T(x) := \sup\{\langle y^*, y - x\rangle : (y^*, y) \in T\}$.

Algorithmic advances. Spingarn's partial inverse and Lions-Mercier monotone splitting: all based on $T_\lambda := \left(T^{-1} + \lambda J^{-1}\right)^{-1}$.

Key Advances: 1990-2000

- **Preiss, David; Phelps, R. R.; Namioka**, I., Smooth Banach spaces, weak Asplund spaces and monotone or usco mappings. *Israel J. Math.* **72** (1990), no. 3, 257-279 (1991).
  - a maximal monotone operator on a space with a smooth norm is generically single-valued (also known in Asplund case)
  - precursor result (separable) **Zarantonello**, E.H., Dense single-valuedness of monotone operators. *Israel J. Math* **15** (1973), 158--166.
- **Simons, Stephen**, *Minimax and monotonicity*. Lecture Notes in Mathematics, **1693**. Springer-Verlag, Berlin, 1998.
  - first comprehensive treatment within (very subtle) convex analysis
  - 2nd edition: *From Hahn-Banach to Monotonicity* **2008** (much simpler)
- **1996 H.H. Bauschke** shows a *bounded* linear max. mon. T is dense type *iff* NI *iff* T^* is monotone
  - captured all known counter-examples (Gossez, Fitzpatrick-Phelps)
  - all of the form $T+\lambda J_X$ and can't exist on lattice X unless

Key Advances: 2000-2012

- Alvez-Marques, Bot, Burachik, Eberhard, Iusem, Martinez-Legas, Simons , Svaiter, Thera, Penot, Vosei, Zalinescu, et al have contributed mightily ...
  - details and accurate citations in Ch. Nine of Convex Functions (CUP)

T is of **dense type** (D): $\inf_{(x,x^*)\in T}\langle x^*-z, x-z^{**}\rangle \geq 0$ implies some bounded net $(x_a, x_a^*) \in T \to_{w^*\times s} (z^{**}, z^*)$

1996 FP operator
$T\colon L^1[0,1] \to L^\infty[0,1]$ given by
$$T(x)(t) := \int_0^t x(s)\,ds - \int_t^1 x(s)\,ds$$
is skew and $\pm T$ not in (D).
(BWY: T in $AC^2[0,1]$ non maximal)

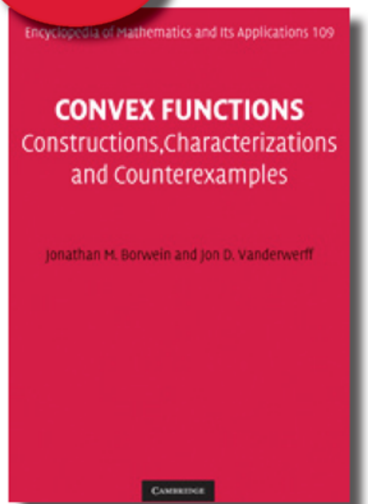
Key Features

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

Choice "Outstanding Academic Title" for 2011

Contents

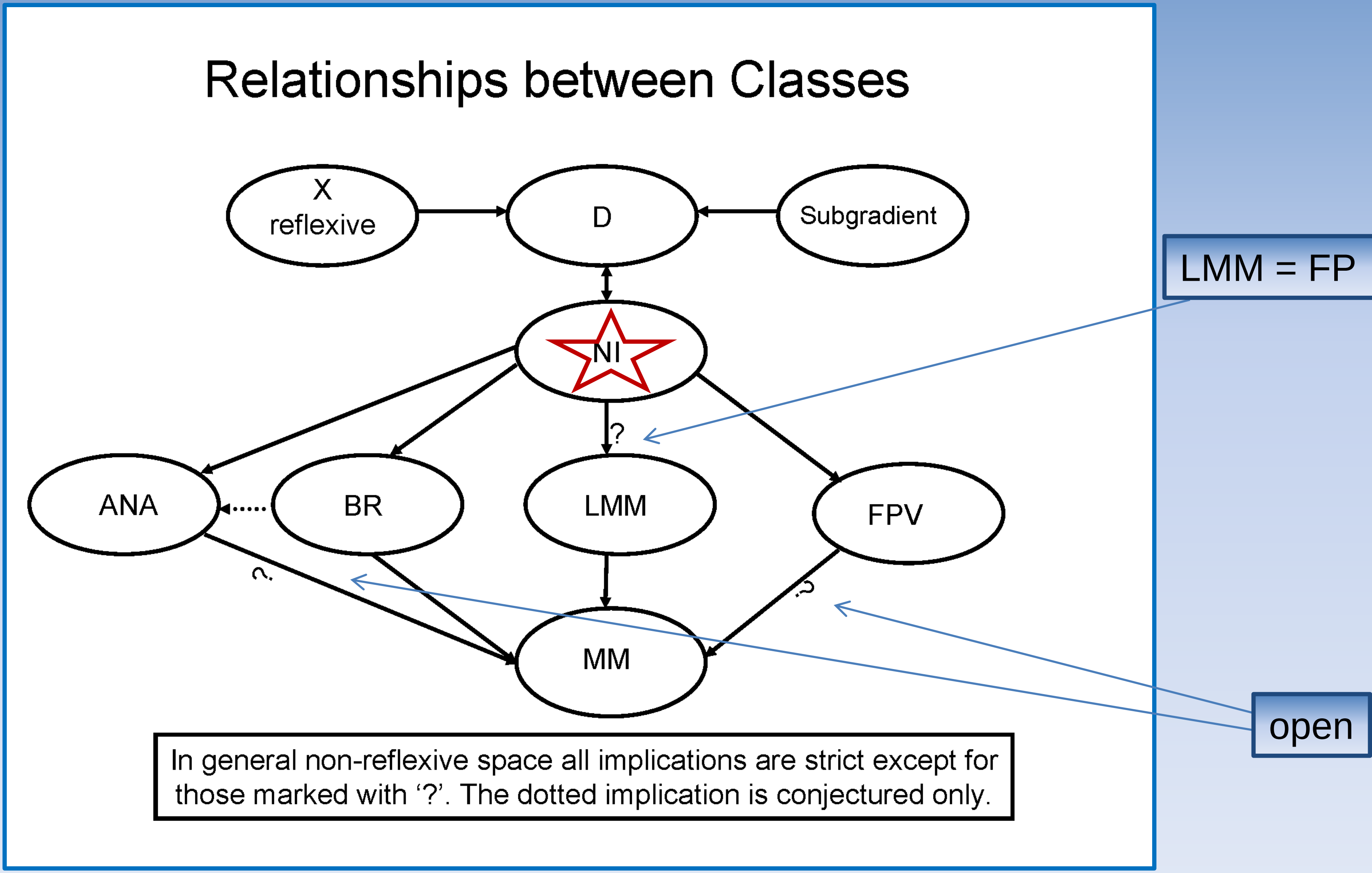
November 2009



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

2000-2010

Here was the current status (NI iff ED iff D)



WHAT WE KNEW IN 2005

- In reflexive spaces proofs are now very natural

Minty's theorem with smoothness hypotheses removed on X

Decoupling Lemma is just Fenchel or Hahn-Banach

Theorem 5.1.31 (Rockafellar) *Let X be a reflexive Banach space and let $T\colon X \to 2^{X^*}$ be a maximal monotone operator. Then* range$(T+J) = X^*$. *Here J is the* duality map *defined by $J(x) := \partial\|x\|^2/2$.*

Proof. The Cauchy inequality and (5.1.16) implies that for all x, x^*,

$$F_T(x,x^*) + \frac{\|x\|^2 + \|x^*\|^2}{2} \geq 0. \tag{5.1.17}$$

Applying the decoupling result of Lemma 4.3.1 to (5.1.17) we conclude that there exists a point $(w^*, w) \in X^* \times X$ such that

$$\begin{aligned} 0 \leq\ & F_T(x,x^*) - \langle w^*, x\rangle - \langle x^*, w\rangle \\ & + \frac{\|y\|^2 + \|y^*\|^2}{2} + \langle w^*, y\rangle + \langle y^*, w\rangle \end{aligned} \tag{5.1.18}$$

Choosing $y \in -Jw^*$ and $y^* \in -Jw$ in inequality (5.1.18) we have

$$F_T(x,x^*) - \langle w^*, x\rangle - \langle x^*, w\rangle \geq \frac{\|w\|^2 + \|w^*\|^2}{2}. \tag{5.1.19}$$

For any $x^* \in Tx$, adding $\langle w^*, w\rangle$ to both sides of the above inequality and noticing $F_T(x,x^*) = \langle x^*, x\rangle$ we obtain

$$\langle x^* - w^*, x - w\rangle \geq \frac{\|w\|^2 + \|w^*\|^2}{2} + \langle w^*, w\rangle \geq 0. \tag{5.1.20}$$

Since (5.1.20) holds for all $x^* \in Tx$ and T is maximal we must have $w^* \in Tw$. Now setting $x^* = w^*$ and $x = w$ in (5.1.20) yields

$$\frac{\|w\|^2 + \|w^*\|^2}{2} + \langle w^*, w\rangle = 0,$$

which implies $-w^* \in Jw$. Thus, $0 \in (T+J)w$. Since the argument applies equally well to all translations of T, we have range$(T+J) = X^*$ as required. ●

WHAT WE KNEW IN 2005

How maximality is assured

- The sum theorem follows similarly

***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].

*

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. (Maximality of ∂f for free)

In reflexive space: Fitzpatrick function yields the existence of maximal monotone extensions from **Hahn-Banach** in **ZF** without **Axiom of Choice** (Bauschke-Wang 2009)

"For many great theorems the necessity of a condition is trivial to prove, but the trick is to prove sufficiency. That's the hard part."
– George Minty (1929-86) (quoted by Andrew Lenard)

WHAT WE KNEW IN 2010

- (D) implies **FPV** implies **cl R(T)** is convex
- (D) iff **LMM** implies **cl D(T)** is convex
- S, T type (D) implies **S+T is maximal** when

  $$0 \in \text{core}[D(T)-D(S)]$$

  and S,T are maximal (recovers reflexive case)
- a nonlinear T with **unique extension** is in (D)
- much about (n-cyclic) Fitzpatrick functions

> *"The difficulty lies, not in the new ideas, but in escaping the old ones, which ramify, for those brought up as most of us have been, into every corner of our minds."* - John Maynard Keynes (1883-1946)

CURRENT REFERENCES

H.H. Bauschke, J.M. Borwein, S. Wang and L. Yao

1. Fitzpatrick-Phelps type coincides with dense type and negative-infimum type, *Optimization Letters*. E-published, Aug 2011.

2. Construction of pathological maximally monotone operators on non-reflexive Banach spaces. Available at http://arxiv.org/abs/1108.1463.

3. The Brezis-Browder theorem in an arbitrary Banach space. In revision for *J. Functional Analysis*, October 2011. Available at http://carma.newcastle.edu.au/jon/BrBr.pdf

2011-2012

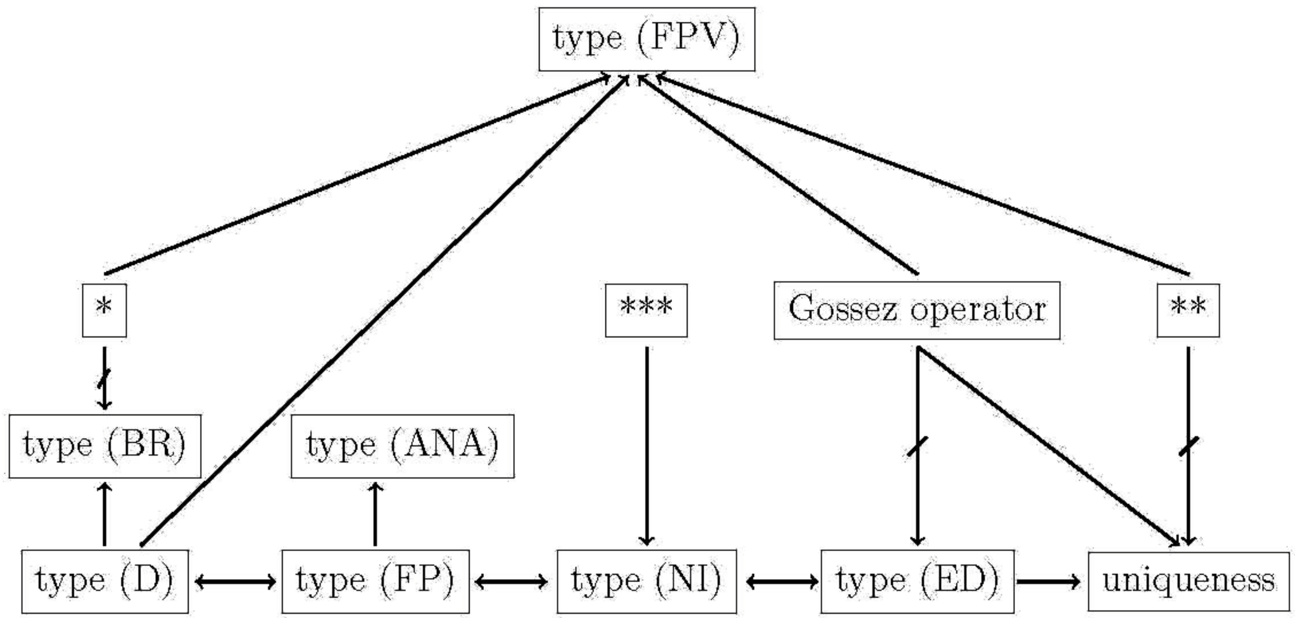
Here is the current status (FP iff NI iff D) (Refers to #2.)

"$*$" refers to skew operators such as T in Theorem 3.6, T_α in Example 4.1, A in Example 4.8, A in Corollary 4.11, and A in Corollary 4.12.

"$**$" refers to the operators such as $A\&T$ in Theorem 3.6, $A_\alpha\&T_\alpha$ in Example 4.1, A in Example 4.8, A in Corollary 4.11, A in Corollary 4.12, and $A + \langle\cdot, e\rangle e$ in Example 4.13.

"$***$" denotes maximally monotone and unique operators with non affine graphs.

We let (ANA), (FP) and (FPV) respectively denote the other monotone operator classes "almost negative alignment", "Fitzpatrick-Phelps" and "Fitzpatrick-Phelps-Veronas". Then by [35, 11, 9, 5, 33, 25, 36, 41], we have the following relationships.



WHAT WE KNOW NOW

Theorem [Brézis-Browder in general Banach space, 2010] Let $A\colon X \to X^*$ be a monotone linear relation such that graph A is closed. Then the following are equivalent.

1. A is maximally monotone of type (D).
2. A is maximally monotone of type (NI).
3. A is maximally monotone of type (FP).
4. A^* is monotone.

Hence in reflexive space A is maximal iff A^* is monotone (Brézis-Browder, 1975).

Recall that the operator *adjoint* of A, written as A^*, is defined by

$$\text{graph } A^* = \{(x^{**}, x^*) \in X^{**} \times X^* \colon (x^*, -x^{**}) \in (\text{graph } A)^\perp\}.$$

Conjecture. Every maximally monotone operator is BR iff X is reflexive. (We know non BR, hence non (D), examples exist in nearly all non-reflexive spaces, see #3.)

WHAT WE STILL DON'T KNOW

In an arbitrary Banach space X:

1. Does S, T maximal imply S+T is maximal if

   $0 \in$ core[D(T)-D(S)]?

   - ✓ if int D(T) $\cap$ int D(S) $\neq \emptyset$ then S+T is maximal (JMB 07)
   - ✓ the right generalization of $(T^{-1} + \lambda J^{-1})^{-1}$ might be useful

2. Are any non-reflexive spaces of **type (D)** ?

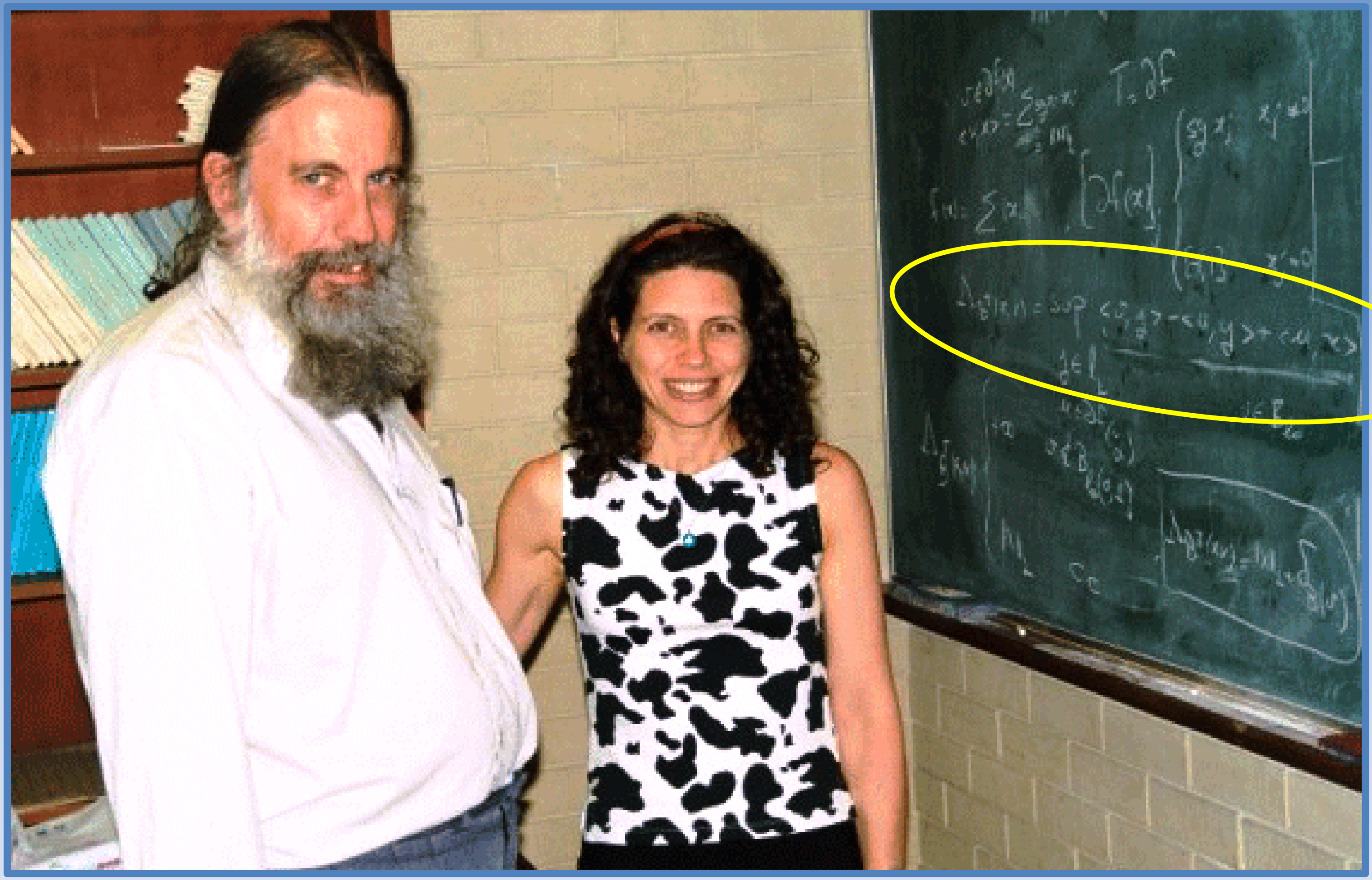
   - ✓ i.e., all maximal monotones on X are type (D)
   - ✓ this seems very unlikely (#3). It fails in all spaces containing c_0 (Svaiter) l^1, James quasi-reflexive space, ...

3. Does every max mon have cl D(T) convex?

   - ✓ cl R(T) convex characterizes reflexive space
   - ✓ I conjecture 'yes'. (It is implied by restricted sum rules)

Simon Fitzpatrick and Regina Burachik

(2004)



THAT'S ALL FOLKS !