Decompositions of Monotone Operators

Jonathan M. Borwein, FRSC

Research Chair in IT
Dalhousie University

Halifax, Nova Scotia, Canada

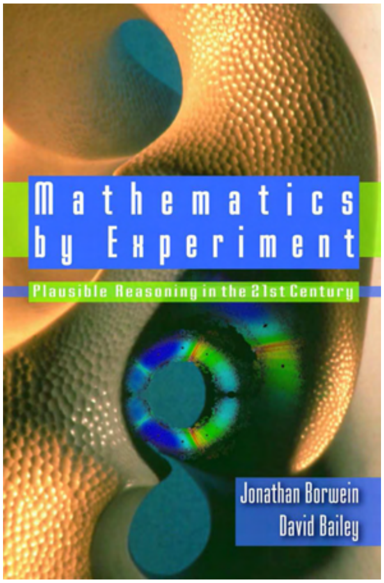
Control, Set-Valued Analysis & Applications

CSVAA '04, Guadeloupe, 5–8 April 2004

I'll be glad if I have succeeded in impressing the idea that it is not only pleasant to read at times the works of the old mathematical authors, but this may occasionally be of use for the actual advancement of science.

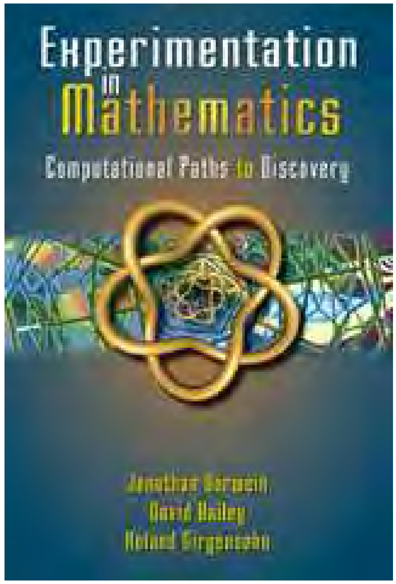
Constantin Carathéodory
in a 1936 address to the MAA

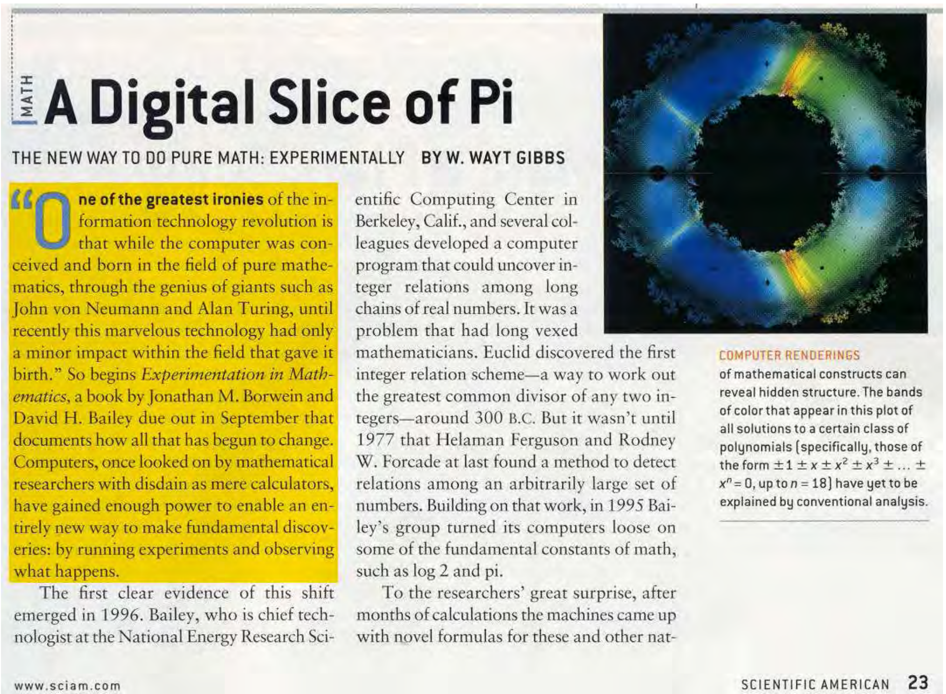
SOME SELF PROMOTION



Two fine new
AK PETERS
Monographs
(2003, 2004)

you should
certainly buy!





MATH

A Digital Slice of Pi

THE NEW WAY TO DO PURE MATH: EXPERIMENTALLY BY W. WAYT GIBBS

"One of the greatest ironies of the information technology revolution is that while the computer was conceived and born in the field of pure mathematics, through the genius of giants such as John von Neumann and Alan Turing, until recently this marvelous technology had only a minor impact within the field that gave it birth." So begins *Experimentation in Mathematics*, a book by Jonathan M. Borwein and David H. Bailey due out in September that documents how all that has begun to change. Computers, once looked on by mathematical researchers with disdain as mere calculators, have gained enough power to enable an entirely new way to make fundamental discoveries: by running experiments and observing what happens.

The first clear evidence of this shift emerged in 1996. Bailey, who is chief technologist at the National Energy Research Scientific Computing Center in Berkeley, Calif., and several colleagues developed a computer program that could uncover integer relations among long chains of real numbers. It was a problem that had long vexed mathematicians. Euclid discovered the first integer relation scheme—a way to work out the greatest common divisor of any two integers—around 300 B.C. But it wasn't until 1977 that Helaman Ferguson and Rodney W. Forcade at last found a method to detect relations among an arbitrarily large set of numbers. Building on that work, in 1995 Bailey's group turned its computers loose on some of the fundamental constants of math, such as log 2 and pi.

To the researchers' great surprise, after months of calculations the machines came up with novel formulas for these and other nat-

COMPUTER RENDERINGS of mathematical constructs can reveal hidden structure. The bands of color that appear in this plot of all solutions to a certain class of polynomials (specifically, those of the form $\pm 1 \pm x \pm x^2 \pm x^3 \pm \ldots \pm x^n = 0$, up to $n = 18$) have yet to be explained by conventional analysis.

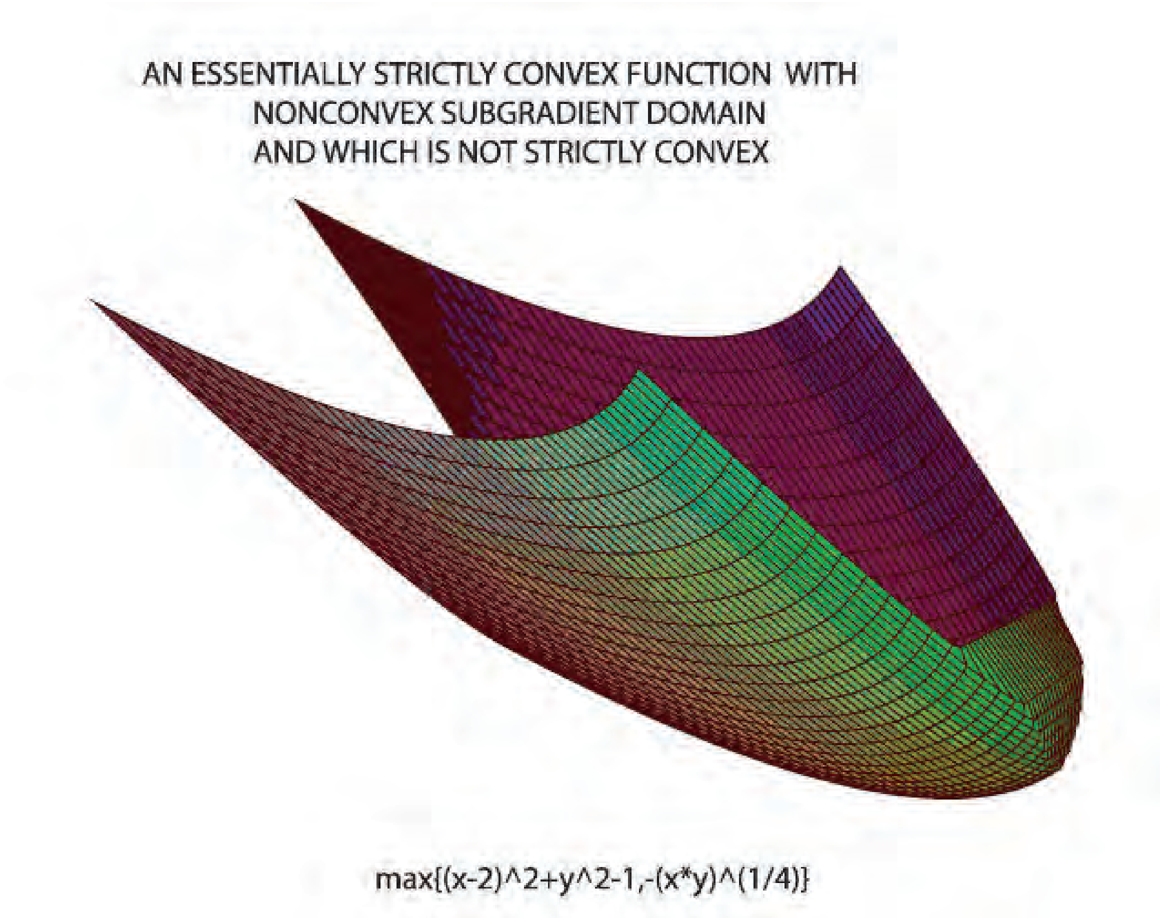
www.sciam.com SCIENTIFIC AMERICAN 23

http://www.expmath.info

www.cs.dal.ca/~jborwein

INTRODUCTION.

Details will appear in: J. Borwein and H. Wiersma, **Asplund Decompositions of Monotone Operators (on Asplund Spaces)**, in preparation.



★ "Even convex functions are hard ... !"

In a largely forgotten* 1968 or 1970 paper:

A MONOTONE CONVERGENCE THEOREM FOR SEQUENCES OF NONLINEAR MAPPINGS

Edgar Asplund

In this paper we prove a theorem generalizing the elementary theorem on convergence of bounded, monotone sequences of real numbers, and also the theorem of Vigier and Nagy, cf. [**2**, Appendice II] on the convergence of certain sequences of symmetric linear operators on Hilbert space.

The paper consists of two sections. In the first we prove the main monotone convergence theorem (Theorem 1) and apply it to prove a decomposition for monotone operators which generalizes the decomposition of a linear operator into symmetric and antisymmetric parts. In the second section we apply Theorem 1 to linear operators. Some rather messy linear algebra computations have to be performed in order to get Theorem 4, which is a natural generalization of the above mentioned theorem of Vigier and Nagy. In the process, we arrive at a characterization of n-monotone linear operators by their numerical ranges (Theorem 3) which shows that even for linear mappings on a space of two real dimensions, the classes of n-monotone operators are all distinct. This settles a question of Rockafellar [**3**, p. 500].

*In part, this is Asplund's fault. Titles really do matter!

Edgar Asplund (1931–1974), *inter alia*, provided a very provocative decomposition ... of a maximal monotone operator as the sum of a subgradient and an **acyclic** (“**skew**”) part.

Asplund’s other seminal contributions include

- Generic existence of **nearest and farthest points** to closed sets in Banach space

- Asplund **averaging** of good (re)norms on a Banach space

- **Generic differentiability** of convex functions (‘SDS spaces’ now called *Asplund spaces*)

- **Duality** between smoothness and roundness (exposedness) properties

---

43 #997

Asplund, Edgar

A monotone convergence theorem for sequences of nonlinear mappings. 1970

Nonlinear Functional Analysis (*Proc. Sympos. Pure Math., Vol. XVIII, Part* 1, *Chicago, Ill.*, 1968) *pp.* 1–9 *Amer. Math. Soc., Providence, R.I.*
47.80

Let X be a real Banach space and let D be a subset of X^* such that $0 \in D$ and D is weak* dense in some norm neighborhood of 0. For $f \in X^D$ one defines $f \geq_n 0$ if $\langle f(x_1), x_1 - x_n \rangle + \sum_{k=2}^n \langle f(x_k), x_k - x_{k-1} \rangle \geq 0$ for all n-tuples $(x_1, \cdots, x_n)$ in D^n, and $f \geq_c 0$ if $f \geq_n 0$ for all $n \in \mathbf{N}$. These "monotonicity" relations were introduced by R. T. Rockafellar [Pacific J. Math. **17** (1966), 497–510; MR **33** #1769]. Theorem 1: Let $\{f_\alpha\} \subset X^D$ be a net with $f_\alpha(0) = 0$ for all α, let $g \in X^D$, let $n \in \mathbf{N}$, suppose that g is norm-to-norm continuous at 0 in D, that $g(0) = 0$ and that $0 \leq_n f_\alpha \leq_n f_\beta \leq_2 g$ whenever $\alpha \leq \beta$; then there exists an $f \in X^D$ such that, in the norm of X, $\lim_\alpha f_\alpha(x) = f(x)$ for all $x \in D$. This monotone convergence theorem is applied to prove a decomposition for monotone operators that generalizes the decomposition of a linear operator into symmetric and antisymmetric parts. The author also gives, in the linear case, a characterization of the n-monotone (i.e., $f \geq_n 0$) operators in terms of their numerical ranges. This characterization shows that the classes of n-monotone operators are all distinct, which settles a question of Rockafellar [loc. cit.].

Reviewed by *J.-P. Gossez*

OUTLINE. In this lecture I intend to:

1. **Motivate** revisiting Asplund's work:

   - How much do we know about **convex subgradients & monotone operators**?

   - Is **a (bounded) linear mapping monotone iff its adjoint is**??

   - Reviewing monotonicity theory in **non-reflexive** spaces

2. **Sketch** a modern version of E.A.'s result.

3. **Discuss** applications and extensions

4. **Pose** some **HARD CONJECTURES**: e.g.,

   All monotone pathologies are realizable with 'skew' mappings.

PRELIMINARIES

Definition 1 *A mapping $T : X \to X^*$ is* monotone *if for all $x, y \in D := \operatorname{dom} T$*

$$\langle y^* - x^*, y - x \rangle \geq 0$$

for all $x^* \in T(x), y^* \in T(y)$.

T is **skew** *if T and $-T$ are both monotone.*

Definition 2 *A (multi-)function $T : D \subset X \to X^*$ is n*-monotone *if*

$$\langle T(x_1), x_1 - x_n \rangle + \sum_{k=2}^{n} \langle T(x_k), x_k - x_{k-1} \rangle \geq 0$$

holds for all $(x_1, \ldots, x_n) \in D^n$, and T is **cyclically monotone** *if it is n-monotone for all n.*

Definition 3 *A monotone mapping $T : D \subset X \to X^*$ is* **acyclically-monotone** *if*

$$M + \partial f \subset T$$

and M monotone implies f is linear.

✓ These classes are distinct (Asplund)

- Our interest in cyclic monotonicity is largely motivated by the following

Theorem 4 (Rockafellar, 1966) *Given a relation ρ on $X \times X^*$, there exists a closed convex function f on X such that $\partial f \supset \rho$* if and only if *ρ is cyclically monotone.*

- Throughout, X is a Banach space and ∂f is the *convex subdifferential* familiar from convex analysis: $\partial f(x) =$

$$\{x^* \in X^* : \langle x^*, y - x\rangle + f(x) \leq f(y) \ \forall y \in X\}.$$

- ∂f is maximally cyclically monotone iff it is a subgradient of a closed convex function.

- Skew and monotone implies acyclic. The converse holds if T is linear. Skew and C^2 implies linear.

(CONVEX) SUBGRADIENTS

★ Take care, even in separable Hilbert space.

Example. *A proper lower semicontinuous convex function, f, on separable Hilbert space with the graph of ∂f not norm$\times$***bw** *closed.*

- Let $E := \ell_2(\mathbb{N})$ and define

  $$e_{p,m} := \frac{1}{p}\left(e_p + e_{p^m}\right), \qquad e^*_{p,m} := e^*_p + (p-1)e^*_{p^m}$$

  for $m, p, r, s \in \mathbb{N}$, $m \geq 2$ and p **prime**.

  We have $\langle e^*_{p,m}, e_{p',m'}\rangle$ **= 0** if $p \neq p'$, **=1/p** if $p = p', m \neq m'$ and **=1** if $p = p', m = m'$.

- For $x \in E$ define

  $$\mathbf{f(x)} := \max_{m>1,p}\left(\langle e^*_1, x\rangle + 1, \sup\{\langle e^*_{p,m}, x\rangle\}\right)$$

  so f is a proper lsc convex function on E.

1. Then $f(0) = f(e_{p,m}) = 1$, $f(-e_1) = 0$ and $f(x) \geq \langle e^*_{p,m}, x\rangle$ for $x \in E$, for $m \geq 2$ and p prime.

   - So, $e^*_{p,m} \in \partial f(e_{p,m})$.

2. Also $0^* \notin \partial f(0)$, since $0^* \in \partial f(0)$ is equivalent to $f(x) - f(0) \geq 0$ for all $x \in E$, which fails for $x = -e_1$.

   - Thus $(0, 0^*)$ is not in the graph of ∂f.

3. The graph of ∂f is not norm×bw closed:

$(0, 0^*)$ is in the norm×bw closure of

$$\{(e_{p,m}, e^*_{p,m}) : \quad m \geq 2, p \text{ prime}\} \subseteq \text{graph}\, \partial f$$

   - Informally, this is true since $e_{p,m}$ tends in norm to 0 for large p, and also 0^* is a bw-cluster point of the $e^*_{p,m}$.

■

Remark. Less instructively use $E = \ell^2([0,1])$ (non–separable) and

$$f_1(x) := \max_{0<r\le 1}\left(\langle e_0, x\rangle + 1, \sup\{r^{-1}\langle e_r, x\rangle\}\right).$$

- Before we used an unbounded sequence with a w*-cluster point; here that $\{r^{-1} e_r : 0 < r \le 1\}$ has 0^* in its bw* closure.*

Theorem 5 *More generally, let E be a Banach space. TFAE:*

i. E is finite dimensional.

ii. The graph of ∂f is norm×bw closed for each closed proper convex f on E.*

iii. The graph of each maximal monotone T on E is norm×bw closed.*

Built, after Namioka noted the bw topology is nastier than we knew; idea originates with Von Neumann.

- Thus, *all* limiting constructions of generalized gradients, that capture the convex subdifferential, *must* fail to be closed for general lower semi-continuous mappings, unless they are locally bounded.

Question. Is Theorem 5 true if int $D(T)$ (int dom f) is required to be non-empty?

- We conjecture "**It is Not**."

That is, we think it possible, at least in reflexive space, that:

The graph of **every** *maximal monotone T with $D(T)$ having nonempty interior interior is norm $\times$ bw* closed.*

THE MONOTONE ‘ZOO’

Definition 1. Suppose T is a (monotone) set-valued map from X to X^*.

Define set-valued maps T_1, T_0, $\overline{T}$ from X^{**} to X^* via:

1. $(x^{**}, x^*) \in \mathrm{Gr}(T_1)$, if there is a bounded net (x_α, x^*_α) in $\mathrm{Gr}(T)$ with $x_\alpha \to_* x^{**}$ and $x^*_\alpha \to x^*$.

2. $(x^{**}, x^*) \in \mathrm{Gr}(T_0)$, if

$$\inf_{(y,y^*)\in \mathrm{Gr}(T)} \langle y^* - x^*, y - x^{**}\rangle = 0.$$

3. $(x^{**}, x^*) \in \mathrm{Gr}(\overline{T})$, if

$$\inf_{(y,y^*)\in \mathrm{Gr}(T)} \langle y^* - x^*, y - x^{**}\rangle \geq 0.$$

Then ...

Definition 2. We say

(i) T *is* **dense type (D)** if $T_1 = \overline{T}$. (Gossez '76)

(ii) T *is* **range-dense type (WD)** if for every $x^* \in \mathrm{R}(\overline{T})$, there is a bounded net $(x_\alpha, x^*_\alpha) \in \mathrm{Gr}(T)$ with $x^*_\alpha \to x^*$. (Simons '95)

(iii) T *is* **type (NI)**, if $\inf_{(y,y^*)\in \mathrm{Gr}(T)} \langle y^* - x^*, y - x^{**}\rangle \leq 0$, for all $(x^{**}, x^*) \in X^{**} \times X^*$. (Simons)

(iv) T *is* **locally maximal monotone (FP)**, if $(\mathrm{Gr}(\mathrm{T})T^{-1}) \cap (V \times X)$ is max. monotone in $V \times X$, for every convex open V in X^* with $V \cap \mathrm{R}(T) \neq \emptyset$. (Fitzpatrick and Phelps '92)

(v) T *is* **unique**, if all maximal monotone extensions of T in $X^{**} \times X^*$ coincide.

REFLEXIVE SPACES

. . . and there are other classes (Simons, Bauschke-Borwein)!

▷ Convex subgradients have all these properties.

▶ Maximal monotone and dense type, or locally maximal monotone implies maximal monotone.

▶ The converses hold in reflexive space, usually easily.

▶ Linear examples show this may fails in some non-reflexive spaces (below).

- In reflexive space the theory is fairly good:* sum rules, domain behaviour (Simons).

- Simons' (1998) accounting of the non-reflexive case is detailed and subtle (domain and range behaviour).

- Generally, things are a mess with few counter-examples. In part, because we can, *unfortunately*, say a lot about the linear case:

*At least when some *core* condition is in force.

Proposition 6 *Suppose T is a continuous linear operator from X to X^*. Then T is weakly compact if and only if $T_1 = T^{**}$.*

Also TFAE: (i) T is positive; (ii) T is monotone; (iii) T is maximal monotone.

We rely also on the following easy-to-prove yet immensely useful decomposition principle.

Proposition 7 *Suppose T is a continuous linear operator from X to X^*. Then T can be written uniquely as the sum of two continuous linear operators, $T = P + S$, where P is* **symmetric** *and S is* **skew***:*

$$Px = \frac{1}{2}Tx + \frac{1}{2}T^*x, \quad Sx = \frac{1}{2}Tx - \frac{1}{2}T^*x, \qquad \forall x \in X.$$

- *P (resp. S) is the* symmetric part *(resp.* skew part*) of T.*

We are now ready for the main linear result.

Theorem 8 Suppose T is a continuous* linear operator from X to X^*. Then TFAE:

(i) T is monotone and of dense type or range-dense type or type (NI).

(ii) T is locally maximal monotone.

(iii T^* is monotone.

(iv) P and S^* are monotone.

(v) P is monotone and S is of dense type or range-dense type or type (NI) or locally maximal monotone.

*Closed and densely defined suffices (Phelps & Simons).

- ▶ This slightly mind-numbing result says ”Linear maps can not distinguish any of the classes”

- In particular, there is a **bad** positive map if and only if it is not the case that X **is such that every bounded map from** X **to** X^* **is weakly compact** (X is a **cms space**).

- Moreover, if X is a Banach lattice TFAE:

  – The adjoint of every positive (resp. skew) map is positive (skew) *iff* X contains no **isometric** copy of ℓ_1, as is the case for $C[0, 1]$.

- ▶ The—only—two fundamental examples are due to Gossez and Fitzpatrick & Phelps:

'SMOOTH' & SO-SO OPERATORS

♠ There are three mutually exclusive cases:

- T is "**good**": S^* and $-S^*$ are monotone.
- T is "**so-so**": one of S^* or $-S^*$ is monotone.
- T is "**bad**": neither S^* nor $-S^*$ is monotone.

Here is an example of a "so-so" operator.

Example 9 (Gossez) Consider $G : \ell_1 \to \ell_\infty$ with

$$(Gx)_n := -\sum_{k<n} x_k + \sum_{k>n} x_k, \quad \forall x = (x_k) \in \ell_1, n \in \mathbb{N}.$$

Then G **and** $-G$ **are skew** operators from ℓ_1 to ℓ_∞ and G^* **is not monotone but** $-G^*$ **is** and so both of dense type and locally maximal monotone.

A BAD OPERATOR

- Surprisingly, the "continuous" version of the (negative) *Gossez* operator is "bad".

Example 10 (Fitzpatrick and Phelps) Define $F : L_1[0,1] \to L_\infty[0,1]$ by

$$(Fx)(t) := \int_0^t x(s)ds - \int_t^1 x(s)ds, \; \forall x \in L_1[0,1],$$

for $t \in [0,1]$.
Then F**,** $-F$ **are skew** from $L_1[0,1]$ to $L_\infty[0,1]$ but **neither** F^* **nor** $-F^*$ **is monotone**.
Consequently, neither F or $-F$ is of type (NI) nor locally maximal monotone.

- Adding a regularizing term (duality map, subgradient) provably can not *worsen* things.

- We have exhausted all known counter-examples:

A conjecture is brewing

ASPLUND'S DECOMPOSITION

★ Studying a reworded version of Asplund's result to the stew reinforces this feeling:

Theorem 11 *(Asplund, Theorem 2) Suppose that T is a single-valued maximal monotone operator defined on a set D in X^* (resp. X) whose weak* (resp. weak) closure has norm interior, and is norm-to-norm continuous (on D) at a point in this interior.*

There is a convex subgradient operator $G = \partial f$ and an acyclically monotone S such that

$$T = \partial f + S.$$

- Linear mappings are acyclic *iff* skew.

★ In finite dimensions, directly, a C^2 monotone is the sum of a convex gradient and a skew linear mapping.

★ The **basic reason why** we may decompose T as

$$T = \partial f + S$$

is $\cdots$

- A delicate 'Zornification' to obtain a maximally cyclic part.
  - The **interiority** condition enforces convergence of cyclically increasing nets

- Rockafellar's result makes this cyclic part a convex subgradient

- Maximality (in the cyclic order) forces the remainder to be acyclically monotone.

$\cdots$ AND EXTENSIONS

★ Put in a modern context we have:

Theorem 12 (Asplund, Theorem $2'$) *Suppose T is a multivalued-maximal monotone operator, $D(T)$ has non-empty interior and X is reflexive Banach space (or Asplund, or X is weak Asplund and T is bounded).*

Then the same decomposition result as in Theorem 11 holds.

In particular, whenever

$$\text{core}\, D(T) \neq 0,$$

locally T *decomposes as the sum of a convex subgradient, ∂f, and an acyclically monotone operator, S.*

Proof sketch. (In Asplund space with $D(T)$ open.)

- Replace $D(T)$ by a generic subset D where T is single-valued and norm continuous.

- Then, by Asplund's result

$$T_{|D} \subset (\partial f)_{|D} + S_{|D}.$$

- Now use the fact that T is a minimal cusco to deduce

$$T \subset \partial f + CUSCO(S_{|D}).$$

- Finally, $\overline{S} := CUSCO(S_{|D})$ is still acyclic.

♡♠♢♣ Maximal monotone operators' domains/ranges, local boundedness and outer continuity, etc. were not well understood when Asplund died 30 years ago (1974).

MY SUBGRADIENT CONJECTURES

- So after many years and failures I have the following **Structure Conjectures”**

1. All known “nice” monotone classes, $\mathcal{M}$, are closed under addition of a subgradient:

$$\mathcal{M} + \partial f \subset \mathcal{M}$$

2. All these “nice” classes, $\mathcal{M}$, coincide.

3. “Bad” operators can be realized by skew (linear).

4. “Nice” monotone operator are locally sums of subgradients and acyclic (often skew linear) maps.

® In short, subgradients and ’skews’ are ubiquitous.

CONCLUSIONS and QUESTIONS

Epiphany: Informal is not the same as sloppy:

(www.colab.sfu.ca)

1. 'de-Zornification': make Asplund's decomposition more explicit or constructive, at least for special spaces or operators?

2. Are there other interesting variants (e.g., conditions on T ('compact') or $D(T)$ ('small')?

GEORGES IFRAH

'A wealthy (15th Century) German merchant, seeking to provide his son with a good business education, consulted a learned man as to which European institution offered the best training. "*If you only want him to be able to cope with addition and subtraction*," the expert replied, "*then any French or German university will do. But if you are intent on your son going on to multiplication and division—assuming that he has sufficient gifts—then you will have to send him to Italy*." '

- From page 577 of *The Universal History of Numbers: From Prehistory to the Invention of the Computer*, Wiley, 2000.

- Emphasizing how great an advance positional notation was.

REFERENCES

- E. Asplund, *A monotone convergence theorem for sequences of nonlinear mappings*, Proceedings of Symposia in Pure Mathematics. Vol. **18**, Part 1, (1970), 1–9.

- H. H. Bauschke and J. M. Borwein, *Continuous Linear Monotone Operators on Banach Spaces*, Pacific J. Math. **189** (1989), no. 1, 1–20.

- Heinz H. Bauschke, Jonathan M. Borwein and Patrick L. Combettes, *Essential smoothness, essential strict convexity, and convex functions of Legendre type in Banach spaces*, Communications in Contemporary Mathematics, **3** (2001), 615–648. [CECM Preprint 00:156].

- J.M. Borwein, S.P. Fitzpatrick and R. Girgensohn, *Subdifferentials whose graphs are not bounded $\times$ weak-star closed,* Can. Math. Bull., **46** (2003), 538–545 [CoLab Preprint 01:175].

- R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. **17** (1966), 497–510.

- S. Simons, *Minimax and Monotonicity,* Lecture Notes in Math. 1693, Springer-Verlag, New York, 1998.