ASPLUND DECOMPOSITION OF MONOTONE OPERATORS

JONATHAN BORWEIN AND HERRE WIERSMA

Abstract. We establish representations of a monotone mapping as the sum of a maximal subdifferential mapping and a 'remainder' monotone mapping, where the remainder is either skew linear, or more broadly 'acyclic', in the sense that it contains no nontrivial subdifferential component. Examples are given of indecomposable and acyclic operators. In particular, we present an explicit nonlinear acyclic operator.

1. Introduction

Let X be a Banach space, and X^* its topological dual. Recall that a *monotone operator* $T : X \rightrightarrows X^*$ is a mapping that satisfies

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

whenever $x^* \in T(x)$ and $y^* \in T(y)$. In general, T could be a multi-valued mapping on an infinite dimensional space; however, the phenomena we wish to discuss are poorly understood, even for single-valued mappings in $\mathbb{R}^n$. We will restrict ourselves largely to this setting where T is single-valued, and X and X^* are both $\mathbb{R}^n$; in the following, the notation $T : X \to X^*$ (single arrow) always denotes a single-valued operator. This is not an unreasonable restriction, since arguments that hold in $\mathbb{R}^n$ usually have a reasonable extension to Asplund spaces [4]. Moreover, in $\mathbb{R}^n$, T is almost everywhere single-valued on $\operatorname{int}\operatorname{dom} T$, from which most of our results naturally extend to the multi-valued case. Further background and references may be found in [2], [3] and [4].

The *domain* of T is $\operatorname{dom} T = \{x \in \mathbb{R}^n \mid T(x) \neq \emptyset\}$, and the *range* of T is $\operatorname{ran} T = \{x^* \in \mathbb{R}^n \mid x^* \in T(x) \text{ for some } x \in \operatorname{dom} T\}$. The *graph* of T is the set $\operatorname{gr} T := \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n \mid x^* \in T(x)\}$. Of particular interest are maximal monotone operators: T is said to be *maximal monotone* if $\operatorname{gr} T \subset \operatorname{gr} S$ and S monotone implies that $T = S$.

One important instance of a maximal monotone operator is the subdifferential of a convex function. Let f be a proper convex lower semicontinuous function on $\mathbb{R}^n$. Then the *subdifferential* $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the mapping

$$\partial f(x) = \{x^* \in \mathbb{R}^n \mid \langle x^*, y - x\rangle + f(x) \leq f(y) \text{ for all } y \in \mathbb{R}^n\}.$$

Subdifferential mappings enjoy a variety of nice properties: they are single valued on large sets, automatically maximal monotone, and seemingly belong to all classes of well-behaved maximal monotone operators in non-reflexive spaces (see [7], [8], [10], [11], [12]). Thus, it appears that if $T = \partial f + R$ possesses any pathology, it is contributed by R. For an arbitrary monotone mapping T, it is therefore appealing to consider decompositions of the form $T = \partial f + R$, where R is a 'remainder' to be made as small as possible in some sense. This is an extension of the decomposition of a linear operator into its symmetric and skew parts: $L = (L + L^*)/2 + (L - L^*)/2$.

The 'nicest' form for R to take is the zero mapping, in which case T is just a subdifferential map. Barring that, perhaps the next simplest form for R to take is a *skew* or *skew-like* mapping. We investigate in section 2 when such a decomposition is possible. Examples of operators for which this decomposition is not possible are given in section 3. Even if R does not take such a simple form, a modernized version of a 1970 result of Asplund ([1], [4]) shows that we can find a decomposition with R 'acyclic', as we describe in section 4. Little is known about the properties of such acyclic mappings, however. We give the first explicit example of a nonlinear acyclic operator $\widehat{S} : \mathbb{R}^2 \to \mathbb{R}^2$ in section 5.

2. Skew Decompositions

A mapping $SL : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be *skew-like* if $\langle x^*, x\rangle = 0$ for all $(x, x^*) \in \operatorname{gr} SL$, and $S : \mathbb{R}^n \to \mathbb{R}^n$ is *skew* if it is linear on $\operatorname{dom} S$ and $\langle Sx, x\rangle = 0$ for all $x \in \operatorname{dom} S$. We allow that $\operatorname{dom} S \neq \mathbb{R}^n$; in this case we require that $S = \widehat{S}|_{\operatorname{dom} S}$ for some skew linear $\widehat{S} : \mathbb{R}^n \to \mathbb{R}^n$.

Fact 1. *Let* $0 \in \operatorname{int}\operatorname{dom} S$.

(1) *If* $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is monotone and skew-like then it is skew linear on* $\operatorname{dom} S$.
(2) *If* $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is monotone, and* $-S$ *is monotone with* $0 \in S(0)$, *then* S *is skew linear on* $\operatorname{dom} S$.

Proof. (1) Using monotonicity and the fact that $\langle x, x^*\rangle = 0$ when $x^* \in S(x)$, we have $\langle x^*, y\rangle \le -\langle y^*, x\rangle$ for all $(x, x^*), (y, y^*) \in \operatorname{gr} S$.

Choose $\epsilon > 0$ so that $\epsilon B \subset \operatorname{int}\operatorname{dom} S$, where B is the closed unit ball in $\mathbb{R}^n$. For $y, z \in \epsilon B$ choose $y_1^* \in S(y)$, $y_2^* \in S(-y)$ and $z^* \in S(z)$. Then $\langle y_1^*, z\rangle \le -\langle z^*, y\rangle$ and $\langle z^*, -y\rangle \le -\langle y_2^*, z\rangle$, which combine to give

$$\langle y_1^* + y_2^*, z\rangle \le 0 \text{ for all } z \in \epsilon B.$$

Hence $y_1^* = -y_2^*$ for all $y_1^* \in S(y)$ and $y_2^* \in S(-y)$, so $S(y)$ is singleton with $S(y) = -S(-y)$ whenever $y \in \epsilon B$.

Let $(x, x^*) \in \operatorname{gr} S$, $y \in \epsilon B$. Then

$$\langle x^*, y\rangle \le -\langle S(y), x\rangle = \langle S(-y), x\rangle \le -\langle x^*, -y\rangle = \langle x^*, y\rangle,$$

so $\langle x^*, y\rangle = \langle S(y), x\rangle$. Suppose $(x_1, x_1^*), (x_2, x_2^*), (\alpha x_1 + \beta x_2, w^*) \in \operatorname{gr} S$. Then

$$\begin{aligned}\langle w^*, y\rangle &= -\langle S(y), \alpha x_1 + \beta x_2\rangle = -\alpha\langle S(y), x_1\rangle - \beta\langle S(y), x_2\rangle \\ &= \alpha\langle x_1^*, y\rangle + \beta\langle x_2^*, y\rangle = \langle \alpha x_1^* + \beta x_2^*, y\rangle\end{aligned}$$

for all $y \in \epsilon B$, so that $w^* = \alpha x_1^* + \beta x_2^*$. Choosing $x_2 = x_1$ and $\alpha + \beta = 1$ shows that S is single valued on $\operatorname{dom} S$. That is, $S(\alpha x+\beta y) = \alpha S(x)+\beta S(y)$ whenever $x, y, \alpha x + \beta y \in \operatorname{dom} S$.

Since $\epsilon B \subset \operatorname{dom} S$, it is clear that there is a unique skew linear extension $\widehat{S}$ of S to the whole space: $\widehat{S}(x) = (\|x\|/\epsilon)S(\epsilon x/\|x\|)$.

(2) If $x^* \in S(x)$ then

$$\langle x^*, x\rangle = \langle x^* - 0, x - 0\rangle = 0$$

since $0 \in S(0)$ and both S and $-S$ are monotone. So S is skewlike and monotone, and we can apply (1) to see that S is skew linear on $\operatorname{dom} S$.

□

We will say that a monotone operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is *weakly decomposable* if it can be written as the sum of a (possibly zero) skew-like operator and the subgradient of a proper lower semicontinuous convex function: $T = S + \partial f$; and *decomposable* if the skew-like part is actually skew. If T is not decomposable, we say that it is *indecomposable.* Note that a skew-like operator need not be monotone. For the following, we use the notation $DT(x)$ for the Jacobian matrix of $T : \mathbb{R}^n \to \mathbb{R}^n$ at x, and we say T is $\mathcal{C}^1$ on an open set C if the mapping $x \to DT(x)$ is continuous on C.

Fact 2. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a $\mathcal{C}^1$ maximal monotone mapping. Then the decomposition $T = S + \nabla f$ into a skew component S and a subdifferential component ∇f is unique when it exists.*

Proof. Suppose $T = S+\nabla f = S_1+\nabla g$. Then $S(x)-S_1(x) = \nabla g(x)-\nabla f(x)$. Differentiating gives

$$S - S_1 = \nabla^2(g - f)(x);$$

the left hand side is a skew matrix, and the right hand side is symmetric, so both must be zero matrices. □

Fact 3. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be $\mathcal{C}^1$ on an open set C with $0 \in C$ and $T(0) = 0$. Then T is monotone (resp. skew) if and only if $DT(z)$ is positive semidefinite (resp. skew) throughout C.*

Proof. We prove only the skew case; the monotone case is similar. Let $DT(z)$ be skew for each z in the interior of $\operatorname{dom} T$, and take $x, y \in \operatorname{dom} T$. The Mean-value theorem then provides $z \in [x, y]$ with

$$\langle T(x) - T(y), x - y\rangle = \langle DT(z)(x - y), x - y\rangle = 0,$$

so T and $-T$ are monotone. Fact 1 shows that T is skew linear. On the other hand, suppose T is skew, with $x \in \operatorname{int}\operatorname{dom} T$. Fixing h, we see that

$$\langle th, DT(x+sh)\ th\rangle = \langle T(x+th) - T(x), th\rangle = 0$$

for some $0 < s < t$. Thus

$$\langle h, DT(x+sh)\ h\rangle = 0;$$

letting $t \to 0$ shows that $DT(x)$ is skew. □

Define *Fitzpatrick's last function* f_T relative to a point $a \in \operatorname{int}\operatorname{dom} T$ by

$$f_T(x;a) := \int_0^1 \langle T\big(a + t(x-a)\big), x-a\rangle\, dt.$$

This construction was suggested to the authors by Simon Fitzpatrick in 2004. We use the notation $f_T(x) := f_T(x;0)$, where $0 \in \operatorname{int}\operatorname{dom} T$.

Lemma 1. *For any monotone $\mathcal{C}^1$ operator $T : \mathbb{R}^n \to \mathbb{R}^n$ with $0 \in \operatorname{int}\operatorname{dom} T$, it is always the case that $S := T - \nabla f_T$ is skew-like.*

Proof. Let f_T be convex. Fix $x, y \in \operatorname{int}\operatorname{dom} T$, and define

$$h(t) := \langle T(tx), ty\rangle.$$

We check that

$$(1)\ \langle T(x), y\rangle = h(1) - h(0) = \int_0^1 t\,\langle DT(tx)x, y\rangle\, dt + \int_0^1 \langle T(tx), y\rangle\, dt$$

and

$$\langle \nabla f_T(x), y\rangle = \int_0^1 t\,\langle DT(tx)^T x, y\,\rangle\, dt + \int_0^1 \langle T(tx), y\rangle\, dt; \tag{2}$$

we can switch the order of integration and differentiation since $(x,t) \to \langle T(tx), x\rangle$ is continuous. Then $S := T - \nabla f_T$ is skewlike, since $\langle T(x), x\rangle = \langle \nabla f_T(x), x\rangle$. □

Theorem 2. *Suppose we have a $\mathcal{C}^1$ maximal monotone operator $T : \mathbb{R}^n \to \mathbb{R}^n$ for which $0 \in \operatorname{int}\operatorname{dom} T$. Then the following are equivalent:*

1. *T is weakly decomposable.*
2. *f_T is convex.*

Proof. Letting $S := T - \nabla f_T$, Lemma 1 shows that S is skew-like. Hence if f_T is convex, T is weakly decomposable.

Conversely, suppose that $T = \nabla g + S$ with g convex and S skew-like. Then $f_{\nabla g} = f_T$ as is seen by writing $h(1) - h(0) = \int_0^1 h'(t)\, dt$ with $h := t \mapsto g(xt)$, which implies that $g = f_T$ and we are done. □

Theorem 3. *Suppose we have a $\mathcal{C}^1$ maximal monotone operator $T : \mathbb{R}^n \to \mathbb{R}^n$ for which $0 \in \operatorname{int}\operatorname{dom} T$. Then T is decomposable if and only if $T - \nabla f_T$ is skew.*

Proof. Without loss of generality we may assume $T(0) = 0$.

If $T - \nabla f_T$ is skew, then

$$\langle \nabla f_T(x) - \nabla f_T(y), x - y\rangle = \langle T(x) - T(y), x - y\rangle \geq 0,$$

so ∇f_T is monotone. By Theorem 12.17 in [10] f_T is convex, so T is decomposable. On the other hand, suppose $T = \nabla g + S$ for some convex g and skew S. Then

$$\begin{aligned} f_T(x) &= \int_0^1 \langle \nabla g(tx) + S(tx), x\rangle dt \\ &= \int_0^1 \langle \nabla g(tx), x\rangle dt = g(x) - g(0), \end{aligned}$$

so $T - \nabla f_T = T - \nabla g = S$ is skew. □

3. Indecomposable Examples

The next example specifies an entire class of indecomposable operators. We require the following lemma:

Lemma 4. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be $\mathcal{C}^1$ and monotone. If there exist $x, y \in \mathbb{R}^n$ and $1 \leq i < j \leq n$ such that $DT(x)_{ij} - DT(x)_{ji} \neq DT(y)_{ij} - DT(y)_{ji}$, then T is indecomposable.*

Proof. Suppose that $T = \nabla f + S$ with f convex and S skew. Then the Hessian matrix $\nabla^2 f(z) = DT(z) - S$ is symmetric for each $z \in \mathbb{R}^n$. Setting $\Delta_{ij} = S_{ij} - S_{ji}$, we have:

$$DT(x)_{ij} = DT(x)_{ji} + \Delta_{ij} \text{ and } DT(y)_{ij} = DT(y)_{ji} + \Delta_{ij},$$

which implies $DT(x)_{ij} - DT(x)_{ji} = DT(y)_{ij} - DT(y)_{ji}$, a contradiction. □

Proposition 5. *Let $g \geq 0$ be a non-constant integrable real function such that with $g(x) \geq 1 = g(0)$ or $g(x) \leq 1 = g(0)$. Let*

$$G(x) := \int_0^x g \quad \textit{and} \quad K(x) := \int_0^x \{(1+g)/2\}^2.$$

Then

(1) *$T(x, y) := (K(x) - G(y), K(y) - G(x))$ is $\mathcal{C}^1$ and maximal monotone.*
(2) *T is indecomposable.*

Proof. To check that T is monotone, we check that the symmetric part of the Jacobian DT of T is positive semidefinite as required by Fact 3. First we compute:

$$DT = \begin{pmatrix} \left(\frac{1+g(x)}{2}\right)^2 & -g(x) \\ -g(y) & \left(\frac{1+g(y)}{2}\right)^2 \end{pmatrix},$$

so

$$DT_{sym} = (DT + DT^T)/2 = \begin{pmatrix} \left(\frac{1+g(x)}{2}\right)^2 & -\frac{g(x)+g(y)}{2} \\ -\frac{g(x)+g(y)}{2} & \left(\frac{1+g(y)}{2}\right)^2 \end{pmatrix}.$$

Since $\left(\frac{1+g(x)}{2}\right)^2 \geq 0$, we only need to check that $\operatorname{Det} DT_{sym} \geq 0$:

$$\begin{aligned}16 \operatorname{Det} DT_{sym} &= (1+g(x))^2 (1+g(y))^2 - 4(g(x)+g(y))^2 \\ &= (g(x)-1)(g(y)-1)\big((g(x)+1)(g(y)+1) + 2(g(x)+g(y))\big) \\ &\geq 0.\end{aligned}$$

The maximality of T is a consequence of example 12.7 in [10]. Lemma 4 with $i=1, j=2$ shows that T is indecomposable, since g is nonconstant. $\square$

Example 6. *If $g := x^2+1$ and T is constructed as in Proposition 5, then $T(x,y) = (x + 1/20\,x^5 + 1/3\,x^3 - 1/3\,y^3 - y, y + 1/20\,y^5 + 1/3\,y^3 - 1/3\,x^3 - x)$ is indecomposable. We have*

$$f_T(x,y) = \frac{1}{120}\,x^6 + \frac{1}{120}\,y^6 + \frac{1}{12}\,x^4 + \frac{1}{12}\,y^4 - \frac{1}{12}\,xy^3 - \frac{1}{12}\,yx^3 + \frac{1}{2}\,x^2 - xy + \frac{1}{2}\,y^2$$

and the Hessian of f_T is

$$\nabla^2 f_T(x,y) = \left[\begin{array}{cc} 1/4\,x^4 + x^2 - 1/2\,xy + 1 & -1/4\,x^2 - 1/4\,y^2 - 1 \\ -1/4\,x^2 - 1/4\,y^2 - 1 & 1/4\,y^4 + y^2 - 1/2\,xy + 1 \end{array}\right];$$

since $\nabla^2 f_T(x,y)_{11} < 0$ for large y and small positive x, f_T is not convex. By Theorem 2, T is also not weakly decomposable.

Example 7. *Consider the mapping*

$$T(x,y) := \left(\sinh(x) - \alpha\, y^2/2, \sinh(y) - \alpha\, x^2/2)\right).$$

Then

$$DT = \begin{pmatrix} \cosh(x) & -\alpha y \\ -\alpha x & \cosh(y) \end{pmatrix}$$

which is monotone iff

$$\alpha^2 \leq \frac{\cosh(x)}{x}\frac{\cosh(y)}{y}$$

for all $x, y > 0$. The right hand side is a separable convex function, and is minimized at $x = y = x_0 = \coth(x_0) = 1.199678\ldots$. So T is monotone iff $\alpha^2 \leq \sinh^2(x_0) = 2.276717\ldots$.

As before, since the difference between the off-diagonal entries of DT is nonconstant, T is indecomposable by Lemma 4.

4. Acyclic Decompositions

In this section, we reconstruct a modern version of a decomposition result found in [1]. We first need to recall some additional monotonicity notions. A mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *N-monotone* for $N \geq 2$ if for every $x_1, x_2, \ldots, x_n \in \operatorname{dom} T$ we have:

$$\sum_{i=1}^{N} \langle T(x_i), x_i - x_{i-1} \rangle \geq 0. \tag{3}$$

where $x_0 := x_N$. Note that 2-monotonicity is just monotonicity. We write $S \leq_N T$ to indicate that $T = S + R$ for some N-monotone R. In particular, this means that $\operatorname{dom} T \subset \operatorname{dom} S$.

By duplicating entries in (3), it is easy to see that an N-monotone mapping is also M-monotone for $M \leq N$; in particular, an N-monotone mapping is monotone. Asplund [1] showed that these classes are distinct via the following example:

Example 8. *For $N \geq 2$ define a 2×2 matrix T_N by*

$$T_N = \begin{pmatrix} \cos(\pi/N) & -\sin(\pi/N) \\ \sin(\pi/N) & \cos(\pi/N) \end{pmatrix}$$

Then $x \to T_N(x)$ is N-monotone, but not $N+1$-monotone.

A more explicit proof to this surprisingly difficult proposition can be found in [2] and [3].

An operator that is N-monotone for every $N \geq 2$ is called *cyclically monotone* or *ω_0-monotone*. It is easy to see that subdifferential mappings are cyclically monotone; in fact, a classical result by Rockafellar [9] shows that subdifferential mappings are the only cyclically monotone mappings:

Theorem 9 ([9], Theorem 1 and Corollary 2)**.** *Suppose $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is cyclically monotone. Then C has a maximal cyclically monotone extension $\widehat{C}$ of the form $\widehat{C} = \partial f$ for some proper lower semicontinuous convex function f. Furthermore,* $\operatorname{ran} \widehat{C} \subset \overline{\operatorname{conv}} \operatorname{ran} C$.

The fact that $\widehat{C}$ preserves the range of C is implicit in the proof of Theorem 1 in [9], where the convex function f is of the form $f(x) = \sup\{\langle x^*_\alpha, x\rangle + r_\alpha \mid x^*_\alpha \in \operatorname{ran} C\}$. For clarity, we prove the following lemma:

Lemma 10. *Let $g(x) = \sup\{\langle x^*_\alpha, x\rangle + r_\alpha \mid \alpha \in A\}$. Then* $\operatorname{ran} \partial g \subset \overline{\operatorname{conv}} \{x^*_\alpha \mid \alpha \in A\}$.

Proof. If $x^* \in \partial g(x)$ and $x^* \notin \overline{\operatorname{conv}} \{x^*_\alpha \mid \alpha \in A\}$, then there is a $y \in \mathbb{R}^n$ such that $\langle x^*, y\rangle > \sup\{\langle x^*_\alpha, y\rangle \mid \alpha \in A\}$.

But $x^* \in \partial g(x)$ implies

$$\langle x^*, \lambda y - x\rangle \leq g(\lambda y) - g(x) \text{ for all } \lambda > 0,$$

which in turn implies that there exists α such that

$$\langle x^*, \lambda y\rangle \leq \langle x^*_\alpha, \lambda y\rangle + r_\alpha - g(x) + \langle x^*, x\rangle \text{ for all } \lambda > 0.$$

Dividing by λ and taking the limit as $\lambda \to \infty$ shows that $\langle x^*, y\rangle \leq \langle x^*_\alpha, y\rangle$, a contradiction.

□

Another range-preserving extension theorem we shall require is the following central case of the Debrunner-Flor theorem:

Theorem 11 ([4], [5])**.** *Suppose* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is monotone with range in* MB *for some* $M > 0$*. Then* T *has a bounded monotone extension* $\widehat{T}$ *with* $\operatorname{dom} \widehat{T} = \mathbb{R}^n$ *and* $\operatorname{ran} \widehat{T} \subset \overline{\operatorname{conv}} \operatorname{ran} T$*.*

The proof of the decomposition below hinges on a kind of monotone convergence theorem. We require the following definition: a monotone operator T is 3^-*-monotone* if

$$\langle T(x), y\rangle \leq \langle T(x), x\rangle + \langle T(y), y\rangle$$

for all $x, y \in \operatorname{dom} T$. In particular, this holds if T is N-monotone for $N \geq 3$, and $0 \in T(0)$.

Theorem 12 ([1], [4])**.** *Let* N *be one of* $3^-, 3, 4, \ldots,$ *or* ω_0*. Consider an increasing net of monotone operators* $T_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ *satisfying*

$$0 \leq_N T_\alpha \leq_N T_\beta \leq_2 T,$$

whenever $\alpha < \beta \in \mathcal{A}$*, for some monotone* $T : \mathbb{R}^n \to \mathbb{R}^n$*. Suppose that* $T(0) = 0$*,* $T_\alpha(0) = 0$ *for all* α*, and that* $0 \in \operatorname{int} \operatorname{dom} T$*. Then*

(i) *There is an* N*-monotone operator* $T_{\mathcal{A}}$ *with*

$$T_\alpha \leq_N T_{\mathcal{A}} \leq_2 T$$

for all $\alpha \in \mathcal{A}$*.*

(ii) *If* T *is maximal monotone and* $\operatorname{ran} T \subset MB$ *for some* $M > 0$ *then one may assume* $\operatorname{ran} T_{\mathcal{A}} \subset MB$*.*

Proof. **(i)** Let $\alpha < \beta$. Since $T(0) = 0$ and $0 \leq_2 T_\alpha \leq_2 T_\beta \leq_2 T$, we have

$$0 \leq \langle x, T_\alpha(x)\rangle \leq \langle x, T_\beta(x)\rangle \leq \langle x, T(x)\rangle, \tag{4}$$

for $x \in \operatorname{dom} T$. So $\lim_{\alpha\to\infty}\langle x, T_\alpha(x)\rangle$ exists.

Writing $T_{\beta\alpha} = T_\beta - T_\alpha$ and using $T_{\beta\alpha} \geq_{3^-} 0$ we get

$$\langle y, T_{\beta\alpha}(x)\rangle \leq \langle x, T_{\beta\alpha}(x)\rangle + \langle y, T_{\beta\alpha}(y)\rangle \tag{5}$$

for $x, y \in \operatorname{dom} T$. A monotone operator is locally bounded on the interior of its domain, see [4], and $0 \in \operatorname{int} \operatorname{dom} T$, so there exist $\epsilon > 0$ and $M > 0$ with $T(\epsilon B) \subset MB$ and $\epsilon B \subset \operatorname{dom} T$. Then

$$0 \leq \langle y, T_{\beta\alpha}(y)\rangle \leq \langle y, T(y)\rangle \leq \epsilon M \tag{6}$$

when $\|y\| \leq \epsilon$.

For $x \in \operatorname{dom} T$, we may choose $\gamma(x)$ so that

$$0 \leq \langle x, T_{\beta\alpha}(x)\rangle \leq \epsilon^2 \tag{7}$$

whenever $\beta > \alpha > \gamma(x)$, since $\langle x, T_\alpha(x)\rangle$ is convergent.

Combining equations (5), (6) and (7) gives

$$\langle y, T_{\beta\alpha}(x)\rangle \leq \langle x, T_{\beta\alpha}(x)\rangle + \langle y, T_{\beta\alpha}(y)\rangle \leq (M + \epsilon)\epsilon$$

for all $\|y\| \leq \epsilon$ and $\beta > \alpha > \gamma(x)$. This shows

$$\langle y, T_{\beta\alpha}(x)\rangle \to 0$$

for all $y \in \mathbb{R}^n$, so $\big(T_\alpha(x)\big)_\alpha$ is Cauchy, and thus has a limit. Setting $T_{\mathcal{A}}(x)$ to this limit, it is clear from the definitions that $T_{\mathcal{A}}$ is N-monotone. It is straightforward to check $T_\alpha \leq_N T_{\mathcal{A}} \leq_2 T$.

(ii) The Debrunner-Flor result shows that $\operatorname{dom} T = \mathbb{R}^n$, since T is maximal. Fixing $x \in \mathbb{R}^n$, we know

$$\begin{aligned} \langle T_\alpha(x), y\rangle &\leq \langle T_\alpha(x), x\rangle + \langle T_\alpha(y), y\rangle \\ &\leq \langle T(x), x\rangle + \langle T(y), y\rangle \end{aligned}$$

for all $y \in \operatorname{dom} T = \mathbb{R}^n$.

From $\|T(y)\| \leq M$ we get

$$\|T_\alpha(x)\|\|y\| \leq \langle T(x), x\rangle + M\|y\|$$

for all $y \in \mathbb{R}^n$. Letting $\|y\| \to \infty$ in this expression gives $\|T_\alpha(x)\| \leq M$. $\square$

The maximality condition in part (ii) of Theorem 12 cannot be removed for $N \neq \omega_0$. Indeed, for a fixed $N \geq 3$ and T_N as in Example 8, define maps T_α and T on the unit ball B by $T_\alpha(x) := T_N(x)$ for each α in some directed set $\mathcal{A}$ and $T(x) := \left(\frac{T_N + T_N^T}{2}\right) x = \cos(\pi/N) I x$. Then $0 \leq_N T_\alpha \leq T_\beta \leq T$ for $\alpha < \beta$, and $T_{\mathcal{A}} = T_\alpha$, but

$$\operatorname{ran} T_{\mathcal{A}} = T_{\mathcal{A}}(B) = B \not\subseteq \cos(\pi/N) B = \operatorname{ran} T.$$

Next we present an updated version of a decomposition result provided by [1]. In this case, the decomposition takes the form of a subdifferential component, as before, and an *acyclic* (termed *irreducible* in [1]) remainder A. Given a set $C \subset \mathbb{R}^n$, a monotone operator $A : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *acyclic* with respect to C if $A = \partial f + R$ with R monotone implies that ∂f is constant on C (i.e. f is affine on C). That is, $A|_C$ has no nontrivial subdifferential component. If no set C is given, then $C = \operatorname{dom} A$ is implied.

Theorem 13 ([1], [4])**.** *Suppose that $T : \mathbb{R}^n \to \mathbb{R}^n$ is a maximal monotone operator with* $\operatorname{int} \operatorname{dom} T \neq \emptyset$.

(i) *T may be decomposed as*

$$T = \partial f + A$$

where f is lower semicontinuous and convex, while A is acyclic with respect to $\operatorname{dom} T$.

(ii) *If* $\operatorname{ran} T \subset MB$, *we may assume that f is M-Lipschitz.*

Proof. **(i)** First, shift the graph of T so that $0 \in \operatorname{int} \operatorname{dom} T$. Consider the set

$$\mathcal{C} := \{C \mid 0 \leq_{\omega_0} C \leq_2 T,\ C(0) = 0\},$$

ordered by the partial order $\leq_{\omega_0}$. Every chain in $\mathcal{C}$ has an upper bound $T_{\mathcal{A}}$ by Theorem 12, and $\mathcal{C}$ is nonempty since it contains the zero mapping, so Zorn's lemma provides a $\leq_{\omega_0}$-maximal $\widehat{C}$ in $\mathcal{C}$ with

$$0 \leq_{\omega_0} \widehat{C} \leq_2 T.$$

So $T = \widehat{C} + A$ for some monotone A. To show that A is acyclic, suppose $A = \partial g + M$. Then

$$T = (\widehat{C} + \partial g) + M,$$

so, by adding a constant to ∂g and subtracting it from M if necessary, we have $\partial g + \widehat{C} \in \mathcal{C}$. Since $\widehat{C}$ is $\leq_{\omega_0}$-maximal, we have $\widehat{C} + \partial g \leq_{\omega_0} \widehat{C}$, so $\operatorname{gr}(-\partial g|_{\operatorname{dom} T}) \subset \operatorname{gr} \partial h$ for some lower semicontinuous convex $h : \mathbb{R}^n \to \mathbb{R}$. Thus g is both convex and concave, hence affine, on $\operatorname{dom} T$, and A is therefore acyclic with respect to $\operatorname{dom} T$.

Now, $\widehat{C}$ is cyclically monotone, so Rockafellar's result shows that $\operatorname{gr} \widehat{C} \subset \operatorname{gr} \partial f$ for some proper convex lower semicontinuous f. This gives:

$$\operatorname{gr} T = \operatorname{gr}(\widehat{C} + A) \subset \operatorname{gr}(\partial f + A),$$

but $\partial f + A$ is monotone, and T is maximal monotone, so $T = \partial f + A$, as required.
(ii) Part (ii) of Theorem 12 shows that one may assume that $\operatorname{ran} \widehat{C} \subset MB$, so $\operatorname{ran} \partial f \subset MB$ by Rockafellar's result. It is straightforward to show that this implies that f is $M-$Lipschitz. □

An immediate corollary of this decomposition is:

Corollary 14. *Under the hypotheses of Theorem 13, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is maximal monotone with bounded range then the acyclic part of the Asplund decomposition of T is nonlinear or zero.*

Proof. Since T is maximal monotone with bounded range, $\operatorname{dom} T = \mathbb{R}^n$. The decomposition $T = \partial f + A$ shows that $\operatorname{dom} \partial f = \operatorname{dom} A = \mathbb{R}^n$, and we know that the range of ∂f is bounded as well. If A is nonzero and linear, then the range of A is unbounded, which is impossible. □

While Corollary 14 implies the existence of many nonlinear acyclic operators, it does not exhibit any explicitly. We remedy this in the next and final section.

5. Acyclic Examples

Skew linear mappings are canonical examples of monotone mappings that are not subdifferential mappings. It is therefore reassuring to know that they are acyclic:

Proposition 15. *Suppose that $S : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous linear operator satisfying $\langle S(x), x\rangle = 0$ for all $x \in \mathbb{R}^n$. Then S is acyclic.*

Proof. Let $S = F+R$ where F is a subdifferential mapping and R is maximal monotone. Since S is single valued, F and R are single valued. In particular. $F = \nabla f$ for some convex differentiable f. Since R is monotone, we have

$$\begin{aligned} 0 &\leq \langle R(x) - R(y), x - y\rangle = \langle S(x) - S(y), x - y\rangle - \langle F(x) - F(y), x - y\rangle \\ &= -\langle F(x) - F(y), x - y\rangle = \langle \nabla(-f)(x) - \nabla(-f)(y), x - y\rangle \end{aligned}$$

This shows that $-f$ is convex, so f is convex and concave, hence linear on its domain. But $\operatorname{dom} f \supset \operatorname{dom} S = \mathbb{R}^n$, so $f \in \mathbb{R}^n$. So $F = \nabla f$ is constant. In fact, by subtracting from F and adding to R, we may assume that $F = 0$. $\square$

For continuous linear monotone operators, then, the usual decomposition into symmetric and skew parts is the same as the Asplund decomposition into subdifferential and acyclic parts.

Recall that Asplund was unable to find explicit examples of nonlinear acyclic mappings, and we have found this quite challenging as well. In particular, we wish to determine a useful characterization of acyclicity. We make some progress in this direction by providing an explicit and to our mind surprisingly simple example: we present a nonlinear acyclic monotone mapping $\widehat{S} : \mathbb{R}^2 \to \mathbb{R}^2$. $\widehat{S}$ is constructed by restricting the range of the skew mapping $S(x, y) = (-y, x)$ to the unit ball, and taking a range-preserving maximal monotone extension of the restriction. This extension is unique, as we see from the following corollary of Proposition 14 from [4], work that originates in [6]:

Corollary 16 ([4], [6]). *Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is maximal monotone and suppose that* $\operatorname{ran} T \cap \operatorname{int} B \neq \emptyset$. *Then there is a unique mapping $\widehat{T}$ such that $T(x) \cap B \subset \widehat{T}(x) \subset B$. Furthermore,*

$$\widehat{T}(x) = \{x^* \in B \mid \langle x^* - y^*, x - y\rangle \geq 0 \text{ for all } y^* \in T(y) \cap \operatorname{int} B\}. \tag{8}$$

Note that $\widehat{T}$ is either a Lipschitz subgradient or it has a nonlinear acyclic part: the acyclic part is bounded so it cannot be nontrivially linear. Hence in the construction of Proposition 17 we know that $\widehat{S}$ has nonlinear acyclic part, which we shall show to be $\widehat{S}$ itself.

Proposition 17. *Define $S : \mathbb{R}^2 \to \mathbb{R}^2$ 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:*

$$\widehat{S}(x) = \begin{cases} S(x) & \|x\| \leq 1 \\ \sqrt{1 - \frac{1}{\|x\|^2}}\frac{x}{\|x\|} + \frac{1}{\|x\|} S\left(\frac{x}{\|x\|}\right) & \|x\| > 1 \end{cases}$$

Proof. From Corollary 16, we know that $\widehat{S}$ exists and is uniquely defined. In the interior of the unit ball, equation (8) shows that $\widehat{S}(x) = S(x)$. Indeed, let $t > 0$ be so small that $z = x + ty \in B$ for all unit length y. Then

$$\langle S(x + ty) - \widehat{S}(x), y\rangle \geq 0$$

for all unit y. Letting $t \to 0$ shows that $\widehat{S}(x) = S(x)$. To determine $(u, v) = \widehat{S}(x)$ for $\|x\| \geq 1$, it suffices by rotational symmetry to consider points $x = (a, 0)$ with $a \geq 1$. Then monotonicity requires that

$$\langle \widehat{S}(x) - S(z), x - z\rangle \geq 0$$

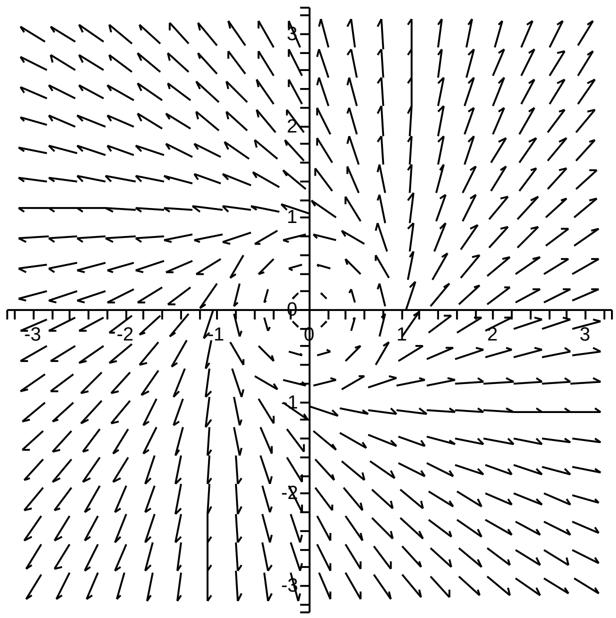

FIGURE 1. A field plot of $\widehat{S}$.

for all $\|z\| \leq 1$. Let $z = \left(\frac{1}{a}, -\frac{\sqrt{a^2-1}}{a}\right)$ so that $\widehat{S}(z) = S(z) = \left(\frac{\sqrt{a^2-1}}{a}, \frac{1}{a}\right)$. Then

$$\left\langle (u,v) - \left(\frac{\sqrt{a^2-1}}{a}, \frac{1}{a}\right), (a,0) - \left(\frac{1}{a}, -\frac{\sqrt{a^2-1}}{a}\right)\right\rangle \geq 0.$$

Expanding this gives

$$u\left(a - \frac{1}{a}\right) + \sqrt{1 - \frac{1}{a^2}}(v - a) \geq 0,$$

and noting that $u \leq \sqrt{1-v^2}$ gives

$$\sqrt{1-v^2}(a^2-1) + \sqrt{a^2-1}(v-a) \geq 0$$

which reduces to $(av-1)^2 \leq 0$, that is, $v = 1/a$. Similarly, setting $z = \left(\frac{1}{a}, -\frac{\sqrt{a^2-1}}{a}\right)$ also shows that $u = \sqrt{1 - \frac{1}{a^2}}$.

So $\widehat{S}(x) = \widehat{S}(a,0) = \left(\sqrt{1-\frac{1}{a^2}}, \frac{1}{a}\right) = \sqrt{1 - \frac{1}{\|x\|^2}}\frac{x}{\|x\|} + \frac{1}{\|x\|}S\left(\frac{x}{\|x\|}\right)$. The same result holds for general $\|x\| \geq 1$ by considering the coordinate system given by the orthogonal basis $\{x, S(x)\}$. $\square$

Figure 1 shows the graph of the vector field $\widehat{S}$. Having computed $\widehat{S}$, we now show that it is acyclic, with the aid of two technical lemmas:

Lemma 18. *$\widehat{S}(x + tS(x)) = S(x)$ for all $t \geq 0$, and all $\|x\| = 1$.*

Proof.

$$\begin{aligned}\widehat{S}(x + tS(x)) &= \sqrt{1 - \frac{1}{1+t^2}}\,\frac{x + tS(x)}{\sqrt{1+t^2}} + \frac{1}{1+t^2}\,S(x + tS(x)) \\ &= \frac{t}{1+t^2}\,(x + tS(x)) + \frac{1}{1+t^2}\,(S(x) - tx) = S(x),\end{aligned}$$

since $S^2 = -I$. □

This construction does not extend immediately to all skew mappings, since it assumes that $S^2 = -I$, which can only occur in even dimensions:

Fact 4. *Skew orthogonal matrices exist only in even dimensions.*

Proof. $\operatorname{Det} S = \operatorname{Det}(S^\top) = \operatorname{Det}(-S) = (-1)^n \operatorname{Det} S$. □

However, such mappings do exist for each even-dimensional $\mathbb{R}^{2n}$, and these can be embedded in $\mathbb{R}^{2n+1}$ in the obvious way. Thus, our construction provides an acyclic nonlinear mapping for each $\mathbb{R}^n$, $n > 1$. In general, it seems probable that our construction of extending a restricted skew linear mapping always leads to an acyclic mapping—and that more ingenuity will allow some reader to prove this.

To show that $\widehat{S}$ is acyclic, we suppose that $\widehat{S} = F + R$, where $F = \partial f$ for some convex proper lower semicontinuous function f and R is maximal monotone, and show that F is constant.

Lemma 19. *Let $\|x\| = 1$, $t \geq 0$ and $y(t) = x + tS(x)$. Then $\langle F(y(t)), S(x)\rangle = c(x)$ for some constant $c(x)$.*

Proof. Suppose $t_1 \neq t_2$. Then $\widehat{S}(y(t_1)) = \widehat{S}(y(t_2))$, by Lemma 18, so

$$\begin{aligned}0 &\leq \langle R(y(t_1)) - R(y(t_2)),\, y(t_1) - y(t_2)\rangle \\ &= \langle \widehat{S}(y(t_1)) - \widehat{S}(y(t_2)),\, y(t_1) - y(t_2)\rangle - \langle F(y(t_1)) - F(y(t_2)),\, y(t_1) - y(t_2)\rangle \\ &= -\langle F(y(t_1)) - F(y(t_2)),\, y(t_1) - y(t_2)\rangle \leq 0,\end{aligned}$$

so

$$\langle F(y(t_1)) - F(y(t_2)), x + t_1 S(x) - (x + t_2 S(x))\rangle = 0,$$

that is

$$\langle F(y(t_1)), S(x)\rangle = \langle F(y(t_2)), S(x)\rangle$$

for any t_1, t_2. □

Proposition 20. *The extension mapping $\widehat{S}$ given explicitly in Proposition 17 is acyclic.*

Proof. First note that if $\widehat{S} = F + R$ with R monotone and $F = \partial f$, then both are single valued, so $F = \nabla f$. As in Proposition 15, we can assume that $f(x) = 0$ when $\|x\| \leq 1$.

Let $\|y\| > 1$. Then there is a unit vector x and a t such that $y = x + tS(x)$:

$$x = \widehat{x}(y) := \frac{y}{\|y\|^2} - \sqrt{\frac{1}{\|y\|^2} - \frac{1}{\|y\|^4}}S(y),$$

$$t = t(y) = \sqrt{\|y\|^2 - 1},$$

and we note that $y \to \widehat{x}(y)$ is continuous. We will determine $f(y)$ by integrating F along the ray $s \to x + sS(x)$. Using Lemma 19, we have:

$$\begin{aligned} f(y) - f(x) &= \int_0^t \langle \nabla f(x + sS(x)), S(x)\rangle ds \\ &= \int_0^t c(x) ds = c(x)t. \end{aligned}$$

Since f is continuous and convex, c is continuous and positive, so $y \to c(\widehat{x}(y))$ is continuous and positive.

Plugging in $t(y)$ gives $f(y) = c(\widehat{x}(y))\sqrt{\|y\|^2 - 1}$ when $\|y\| > 1$ and $f = 0$ for $\|y\| \leq 1$. Suppose $c(y) > 0$ for some $\|y\| = 1$. Then for f to be convex on the segment $[y, 2y]$ we require that:

$$(1 - \lambda)f(y) + \lambda f(2y) \geq f\big((1 + \lambda)y\big) \text{ for all } \lambda \in (0, 1).$$

This means

$$0 + \lambda c\big(\widehat{x}(2y)\big)\sqrt{3} \geq c\big(\widehat{x}\big((1 + \lambda)y\big)\big)\sqrt{\lambda^2 + 2\lambda},$$

or

$$c\big(\widehat{x}(2y)\big)\sqrt{3} \geq c\big(\widehat{x}\big((1 + \lambda)y\big)\big)\sqrt{1 + \frac{2}{\lambda}},$$

for all $\lambda \in (0, 1)$. Letting $\lambda \to 0$, we get $\widehat{x}\big((1 + 2\lambda)y\big) \to y$, so $c\big(\widehat{x}\big((1 + \lambda)y\big)\big) \to c(y) > 0$. Since $\sqrt{1 + \frac{2}{\lambda}} \to \infty$, the inequality does not hold for small λ unless $c(y) = 0$.

For f to be convex and everywhere defined, then, we require $c(y) = 0$ for all $\|y\| = 1$. That is, f is identically zero. □

5.1. **Computing $f_{\widehat{S}}$.** We can also explicitly compute Fitzpatrick's last function $f_{\widehat{S}}$ in this case:

Proposition 21. *With $\widehat{S}$ as before, we have:*

$$f_{\widehat{S}}(x) = \begin{cases} 0 & \|x\| \leq 1 \\ \sqrt{\|x\|^2 - 1} + \arctan\left(\frac{1}{\sqrt{\|x\|^2 - 1}}\right) - \frac{\pi}{2} & \|x\| > 1 \end{cases}.$$

Proof. It is immediate from the definition that $f_{\widehat{S}}(x) = 0$ when $\|x\| \leq 1$. For $\|x\| > 1$, we get:

$$
\begin{aligned}
f_{\widehat{S}}(x) &= \int_0^1 \langle x, \widehat{S}(tx)\rangle dt \\
&= \int_0^{\frac{1}{\|x\|}} t\langle x, S(x)\rangle dt + \int_{\frac{1}{\|x\|}}^1 \sqrt{1 - \frac{1}{t^2\|x\|^2}} \frac{1}{\|x\|}\langle x, x\rangle dt + \int_{\frac{1}{\|x\|}}^1 \frac{1}{t\|x\|^2}\langle S(x), x\rangle dt \\
&= \int_{\frac{1}{\|x\|}}^1 \sqrt{1 - \frac{1}{t^2\|x\|^2}}\|x\| dt \\
&= \int_1^{\|x\|} \sqrt{1 - \frac{1}{s^2}} ds \\
&= \sqrt{\|x\|^2 - 1} + \arctan\left(\frac{1}{\sqrt{\|x\|^2 - 1}}\right) - \frac{\pi}{2}
\end{aligned}
$$

□

Note that $f_{\widehat{S}}$ is convex, since it is a composition of the norm $x \to \|x\|$ with the increasing convex function $t \to \int_1^t \sqrt{1 - \frac{1}{s^2}} ds$. So $\widehat{S}$ is weakly decomposable as $\widehat{S} = \nabla f_{\widehat{S}} + SL$ where SL is skew-like. To determine SL, we compute:

$$
\nabla f_{\widehat{S}}(x) = \begin{cases} 0 & \|x\| < 1 \\ \sqrt{1 - \frac{1}{\|x\|^2}} \frac{x}{\|x\|} & \|x\| \geq 1 \end{cases}.
$$

So $\widehat{S}(x) = \nabla f_{\widehat{S}}(x) + h(\|x\|)S(x)$, where

$$
h(t) = \begin{cases} 1 & t \leq 1 \\ \frac{1}{t^2} & t \geq 1 \end{cases}.
$$

So $\widehat{S}$ is not decomposable, but is weakly decomposable, since $SL = x \to h(\|x\|)S(x)$ is clearly skew-like. Note finally that SL is not monotone.

6. Conclusion

In this paper, we have provided some tools for the decomposition of monotone operators. This was originally motivated by observing that the classical counterexamples in monotone operator theory (see section 6 of [4]) are built from skew operators; in some sense, subgradients ("symmetric" operators) and acyclic mappings ("skew" operators) represent the extreme points of the space of monotone operators. The results we have given in this paper make this more concrete.

We remain convinced that a better understanding of acyclic operators will shed light on a number of open questions. For instance, if a Banach space has good differentiability properties, do monotone operators defined on the space inherit these? In a more limited fashion it seems important to answer the following questions: (1) Is there an iterative construction to compute the acyclic part of a monotone operator in finite dimensional space? and (2) Is there an effective characterization of acyclicity that allows one to easily determine whether a given operator is acyclic?

References

1. Edgar Asplund, *A monotone convergence theorem for sequences of nonlinear mappings*, Proceedings of Symposia in Pure Mathematics **18** (1970), 1–9.
2. Sedi Bartz, Heinz H. Bauschke, Jonathan M. Borwein, Simeon Reich, and Shawn Wang, *Fitzpatrick functions, cyclic monotonicity and Rockafellar's antiderivative*, Nonlinear Analysis: Theory Methods and Applications (2005).
3. Heinz H. Bauschke, Jonathan M. Borwein, and Shawn Wang, *Fitzpatrick functions and continuous linear monotone operators*, Submitted to SIAM J. Opt. (2006).
4. J. M. Borwein, *Maximal montonicity via convex analysis*, Journal of Convex Analysis (2006).
5. H. Debrunner and P. Flor, *Ein Erweiterungssatz für monotone Mengen*, Arch. Math. **15** (1964), 445–447.
6. S. Fitzpatrick and R.R. Phelps, *Bounded approximants to monotone operators on banach spaces*, Annales de l'Institut Henri Poincaré, Analyse Non Linéaire **9** (1992), 573–595.
7. ______, *Some properties of maximal monotone operators on nonreflexive banach spaces*, Set-Valued Analysis **3** (1995), 51–69.
8. R.R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture Notes in Mathematics, vol. 1364, Springer-Verlag, Berlin, 1989.
9. R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. **17** (1966), 497–510.
10. R. T. Rockafellar and R. J-B. Wets, *Variational analysis*, Grundlehren der mathematischen Wissenschaften, vol. 317, Springer, Berlin; Heidelberg; New York, 1998.
11. Stephen Simons, *Subdifferentials are locally maximal monotone*, Bulletin of the Australian Mathematical Society **47** (1993), 465–471.
12. ______, *Minimax and monotonicity*, Lecture Notes in Mathematics, vol. 1693, Springer-Verlag, New York, 1998.