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 “acyclic” in the sense that it contains no nontrivial subdifferential component. This is the nonlinear analogue of a skew linear operator. Examples of indecomposable and acyclic operators are given. In particular, we present an explicit nonlinear acyclic operator.

Key words. monotone operators, cyclic monotonicity, decompositions, convex subgradients, acyclic operators

AMS subject classifications. 47H04, 52A41

DOI. 10.1137/060658357

1. Introduction. Let X be a Banach space and X^* its topological dual. We denote the closed unit ball in X by B_X or B . 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 $y^* \in T(y)$.

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

In general, T could be a multivalued 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^* both are $\mathbb{R}^n$; in the following, the notation $T : \text{dom } T \subset X \rightarrow X^*$ (single arrow) always denotes a single valued operator. This is not an unreasonable restriction, since the results that hold in $\mathbb{R}^n$, such as continuity or differentiability theorems, usually have a reasonable extension at least to separable Asplund spaces [4]. Moreover, in $\mathbb{R}^n$, T is almost everywhere single valued on $\text{int dom } T$, from which most of our results naturally extend to the multivalued case. Further background and references may be found in [2, 3, 4, 5].

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* $\partial f : \text{dom } \partial f \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the monotone mapping

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

*Received by the editors April 27, 2006; accepted for publication (in revised form) April 4, 2007; published electronically October 4, 2007.

<http://www.siam.org/journals/siopt/18-3/65835.html>

[†]Faculty of Computer Science, Dalhousie University, 6050 University Avenue, Halifax, NS B3H 1W5, Canada (jborwein@cs.dal.ca).

[‡]Department of Mathematics and Statistics, Dalhousie University, 6050 University Avenue, Halifax, NS B3H 3J5, Canada (hwiersma@cs.dal.ca).

Subdifferential mappings enjoy a variety of nice properties: they are single valued on large sets and automatically maximal monotone [11, 4, 14] and seemingly belong to all classes of well-behaved maximal monotone operators in nonreflexive spaces (see [8, 9, 12, 13, 14]). 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 (see [1, 4]) shows that we can find a decomposition with R “acyclic,” as described 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 $\hat{S} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ in section 5. We further explore this mapping in section 6 and conclude with some open questions.

Since our central goal is to better understand acyclicity, little will be lost if the reader assumes throughout that every monotone operator is everywhere defined and single valued.

2. Skew decompositions. In this section we introduce various weakenings of the notion of a skew symmetric linear mapping and then link them to the properties of an associated function f_T due to Fitzpatrick as defined later in this paper. A mapping $SL : \text{dom } SL \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be *skew-like* if $\langle x^*, x \rangle = 0$ for all $(x, x^*) \in \text{gr } SL$, and $S : \text{dom } S \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is *skew* if it is linear and $\langle Sx, x \rangle = 0$ for all $x \in \text{dom } S$. We allow that $\text{dom } S \neq \mathbb{R}^n$; in this case we require that $S = \hat{S}|_{\text{dom } S}$ for some skew linear $\hat{S} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Thus, skew mappings are ab initio restrictions of skew and linear mappings.

FACT 1. Let $0 \in \text{int dom } S$.

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

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

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

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

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 $y \in \varepsilon B$.

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

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

so $\langle x^*, y \rangle = -\langle S(y), x \rangle$. Suppose $(x_1, x_1^*), (x_2, x_2^*), (\alpha x_1 + \beta x_2, w^*) \in \text{gr } S$. 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 $y \in \varepsilon B$, 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 $\text{dom } S$. That is, $S(\alpha x + \beta y) = \alpha S(x) + \beta S(y)$ whenever $x, y, \alpha x + \beta y \in \text{dom } S$.

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

(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 skew-like and monotone, and we can apply (1) to see that S is skew linear on $\text{dom } S$. $\square$

We will say that a monotone operator $T : \text{dom } T \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 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*. For example, the addition of a skew mapping to the subgradient of any norm produces a multivalued decomposable maximal monotone mapping.

Note that a skew-like operator need not be monotone. Note also that if $T(x)$ is single valued and nonempty, so necessarily is $S(x)$ and $\partial f(x)$.

For the following, we use the notation $DT(x)$ for the Jacobian matrix of $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ at x , and we note that T is $\mathcal{C}^1$ (Fréchet or, equivalently in finite dimensions, Gâteaux) on an open set C if and only if the mapping $x \rightarrow DT(x)$ is continuous on C .

FACT 2. *Let $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuously Fréchet differentiable maximal monotone mapping on an open domain. Then the decomposition $T = S + \nabla f$ into a skew component S and a subdifferential component $\partial f = \{\nabla f\}$ is unique when it exists.*

Proof. From now on we will identify $\{\nabla f\}$ and ∇f . 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. $\square$

A useful observation is the following.

FACT 3. *Let $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be continuously Fréchet differentiable on an open convex set $C \subset \text{dom } T$ with $0 \in C$ and $T(0) = 0$. Then T is monotone (resp., skew) on C 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 C , and take $x, y \in C \subset \text{dom } T$. 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 monotones. Fact 1 shows that T is skew linear. On the other hand, suppose T is skew, with $x \in \text{int dom } T$. 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 $t \rightarrow 0$ shows that $DT(x)$ is skew. $\square$

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

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

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

We may use f_T to characterize both weak decomposability and decomposability. We start with a technical lemma.

LEMMA 1. *For any continuously Fréchet differentiable monotone operator $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ with $0 \in \text{int dom } T$, it is always the case that $S := T - \nabla f_T$ is skew-like on $\text{int dom } T$.*

Proof. Fix $x, y \in \text{int dom } T$, and define

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

We check that

$$(2.1) \quad \langle T(x), y \rangle = h(1) - h(0) = \int_0^1 t \langle DT(tx)x, y \rangle dt + \int_0^1 \langle T(tx), y \rangle dt$$

and

$$(2.2) \quad \langle \nabla f_T(x), y \rangle = \int_0^1 t \langle DT(tx)^T x, y \rangle dt + \int_0^1 \langle T(tx), y \rangle dt;$$

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

Throughout the rest of this section we assume $\text{dom } T$ is open so as to avoid technical complications at boundary points.

THEOREM 2 (weak decomposability). *Suppose T is a continuously Fréchet differentiable maximal monotone operator $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ for which $0 \in \text{int dom } T = \text{dom } T$. Then the following are equivalent:*

- (1) T is weakly decomposable on $\text{dom } T$,
- (2) f_T is convex on $\text{dom } T$.

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 $g - g(0) = f_T$ and we are done. $\square$

THEOREM 3 (decomposability). *Suppose we have a continuously differentiable maximal monotone operator $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ for which $0 \in \text{int dom } T = \text{dom } T$. Then T is decomposable on $\text{dom } T$ if and only if $T - \nabla f_T$ is skew on $\text{dom } T$.*

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 [12] 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. $\square$

So far we have not explicitly established that (weakly) indecomposable monotone operators actually exist. We address this in the next section.

3. Indecomposable examples. The next example specifies an entire class of everywhere-defined indecomposable operators. We require the following lemma.

LEMMA 4. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 on $\mathbb{R}^n$.*

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

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

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

PROPOSITION 5. *Let $g \geq 0$ be a nonconstant and continuous real function such that either $g(x) \geq 1 = g(0)$ or $g(x) \leq 1 = g(0)$. Let*

$$G(x) := \int_0^x g \quad \text{and} \quad K(x) := \int_0^x \left\{ \frac{(1+g)}{2} \right\}^2.$$

Then

- (1) $T(x, y) := (K(x) - G(y), K(y) - G(x))$ is both continuously differentiable and maximal monotone $\mathbb{R}^2$;
- (2) T is indecomposable on $\mathbb{R}^2$.

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

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

so

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

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

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

The maximality of T is a consequence of Example 12.7 in [12]. 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

$$\nabla^2 f_T(x, y) = \begin{bmatrix} 1/4 x^4 + x^2 - 1/2 xy + 1 & -1/4 x^2 - 1/4 y^2 - 1 \\ -1/4 x^2 - 1/4 y^2 - 1 & 1/4 y^4 + y^2 - 1/2 xy + 1 \end{bmatrix};$$

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

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

Then

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

which is monotone if and only if

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

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

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

We may now turn to the more general notion of an acyclic decomposition.

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 $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is said to be N -monotone for $N \geq 2$ if for every $x_1, x_2, \dots, x_N \in \text{dom } T$ we have

$$(4.1) \quad \sum_{i=1}^N \langle T(x_i), x_i - x_{i-1} \rangle \geq 0,$$

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 $\text{dom } T \subset \text{dom } S$.

By duplicating entries in (4.1), 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 $x \rightarrow T_N(x)$ is N -monotone, but not $(N+1)$ -monotone. A more explicit proof to this surprisingly difficult proposition can be found in [2, 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 [10] shows that subdifferential mappings are the only cyclically monotone mappings.

THEOREM 9 (maximal cyclic monotonicity [4, 10]). *Suppose $C : \text{dom } C \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 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, $\text{ran } \widehat{C} \subset \overline{\text{conv}} \text{ran } C$.*

The fact that $\widehat{C}$ preserves the range of C is implicit in the proof of Theorem 1 in [10], where the convex function f is of the form $f(x) = \sup\{\langle x_\alpha^*, x \rangle - r_\alpha \mid x_\alpha^* \in \text{ran } C\}$. For clarity, we prove the following lemma.

LEMMA 10. *Let X be a Banach space and let $x_\alpha^* \in X^*$ for $\alpha \in A$ and with each r_α real. Let $f(x) = \sup\{\langle x_\alpha^*, x \rangle - r_\alpha \mid \alpha \in A\}$. Then $\text{ran } \partial f \subset \overline{\text{conv}}^* \{x_\alpha^* \mid \alpha \in A\}$.*

Proof. Consider the convex function g defined on X^* by

$$g(x^*) := \inf \left\{ \sum \lambda_{\alpha_i} r_{\alpha_i} : \sum \lambda_{\alpha_i} x_{\alpha_i} = x^*, \sum \lambda_{\alpha_i} = 1, \lambda_{\alpha_i} > 0 \right\},$$

as we range over all finite subsets of $\{(x_\alpha^*, r_\alpha) \mid \alpha \in A\}$. It is easy to check that $g^*|_X = f$, and $f^* = g^{**}$ viewed in $\sigma(X^*, X)$. Now when $x^* \in \partial f(x)$ we have $f(x) + f^*(x^*) = \langle x^*, x \rangle$. Since $x \in \text{dom } f$, we see that $g^{**}(x^*)$ is finite and we are done since $\text{dom } g^{**} \subset \overline{\text{dom } g}^*$.

Alternative proof. If the conclusion fails we may find $x^* \in \partial f(x)$, $\varepsilon > 0$, and $h \in X$ such that

$$(4.2) \quad \langle x^*, h \rangle > \varepsilon + \sup_{\alpha \in A} \langle x_\alpha^*, h \rangle,$$

by the Hahn–Banach theorem. Thus, for each $\alpha \in A$ we have

$$\begin{aligned} \langle x^*, h \rangle &\geq \varepsilon + \langle x_\alpha^*, x \rangle = \varepsilon + (\langle x_\alpha^*, x+h \rangle - r_\alpha) - (\langle x_\alpha^*, x \rangle - r_\alpha) \\ &\geq \varepsilon + \langle x_\alpha^*, x+h \rangle - r_\alpha - f(x). \end{aligned}$$

Now $f(x)$ is finite and so supremizing over $\alpha \in A$ yields

$$\langle x^*, h \rangle \geq \varepsilon + f(x+h) - f(x),$$

in contradiction to $x^* \in \partial f(x)$. $\square$

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

THEOREM 11 (Debrunner–Flor extension [4, 6]). *Suppose $T : \text{dom } T \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is monotone with range in MB_{X^*} for some $M > 0$. Then T has a bounded monotone extension $\widehat{T}$ with $\text{dom } \widehat{T} = \mathbb{R}^n$ and $\text{ran } \widehat{T} \subset \overline{\text{conv}} \text{ran } T$.*

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

THEOREM 12 (monotone convergence [1, 4]). *Let N be one of 3^- , 3 , 4 , $\dots$, or ω_0 . Consider an increasing net of monotone operators $T_\alpha : \text{dom } T_\alpha \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfying*

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

whenever $\alpha < \beta \in \mathcal{A}$, for some monotone $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$. Suppose that $T(0) = 0$, $T_\alpha(0) = 0$ for all α , and that $0 \in \text{int dom } T$. 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 $\text{ran } T \subset MB$ for some $M > 0$, then one may assume $\text{ran } T_{\mathcal{A}} \subset MB$.*

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

$$(4.3) \quad 0 \leq \langle x, T_\alpha(x) \rangle \leq \langle x, T_\beta(x) \rangle \leq \langle x, T(x) \rangle$$

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

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

$$(4.4) \quad \langle y, T_{\beta\alpha}(x) \rangle \leq \langle x, T_{\beta\alpha}(x) \rangle + \langle y, T_{\beta\alpha}(y) \rangle$$

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

$$(4.5) \quad 0 \leq \langle y, T_{\beta\alpha}(y) \rangle \leq \langle y, T(y) \rangle \leq \varepsilon M$$

when $\|y\| \leq \varepsilon$.

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

$$(4.6) \quad 0 \leq \langle x, T_{\beta\alpha}(x) \rangle \leq \varepsilon^2$$

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

Combining (4.4), (4.5), and (4.6) gives

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

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

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

for all $y \in \mathbb{R}^n$, so $(T_\alpha(x))_\alpha$ 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 $\text{dom } T = \mathbb{R}^n$, since T is maximal. Fixing $x \in \mathbb{R}^n$, we know that

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

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

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 $\|y\| \rightarrow \infty$ 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 $T(x) := (\frac{T_N + T_N^T}{2})x = \cos(\pi/N)Ix$. Then $0 \leq_N T_\alpha \leq T_\beta \leq T$ for $\alpha < \beta$, and $T_\mathcal{A} = T_\alpha$, but

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

Now we are ready to present an updated version of a decomposition result provided in [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 $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 $C = \text{dom } A$ is implied.

THEOREM 13 (Asplund decomposition [1, 4]). *Suppose we are given a (single-valued) maximal monotone operator $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ with $\text{int dom } T \neq \emptyset$.*

(i) *T may be decomposed as*

$$T = \nabla f + A,$$

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

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

Proof. (i) First, shift the graph of T so that $0 \in \text{int dom } T$. 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 $\text{gr}(-\partial g|_{\text{dom } T}) \subset \text{gr } \partial h$ for some lower semicontinuous convex $h : \mathbb{R}^n \rightarrow \mathbb{R}$. Thus g is both convex and concave, hence affine, on $\text{dom } T$, and A is therefore acyclic with respect to $\text{dom } T$.

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

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

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 $\text{ran } \widehat{C} \subset MB$, so $\text{ran } \partial f \subset MB$ by Rockafellar's result. It is straightforward to show that this implies that f is M -Lipschitz. $\square$

An immediate corollary of this decomposition is the following.

COROLLARY 14 (nonlinear acyclicity). *Under the hypotheses of Theorem 13, if $T : \text{dom } T \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is maximal monotone with bounded range, then the acyclic part of the Asplund decomposition of T is either nonlinear or zero. In other words, A is nonlinear unless T is cyclically monotone.*

Proof. Since T is maximal monotone with bounded range, $\text{dom } T = \mathbb{R}^n$. The decomposition $T = \partial f + A$ shows that $\text{dom } \partial f = \text{dom } A = \mathbb{R}^n$, 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. $\square$

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

5. Explicit 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 $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 $\text{dom } f \supset \text{dom } S = \mathbb{R}^n$, 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 $F = 0$. $\square$

We leave it to the reader to check that the sum of an acyclic operator and a skew linear operator is still acyclic. It is not clear that the sum of two acyclic operators must be acyclic. 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.

We recall that Asplund was unable to find explicit examples of nonlinear acyclic mappings [1], 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 $\widehat{S} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$.

Precisely, $\widehat{S}$ 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 [7].

COROLLARY 16 (unique extension [4, 7]). *Suppose $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is maximal monotone and suppose that $\text{ran } T \cap \text{int } B \neq \emptyset$. Then there is a unique maximal monotone mapping $\widehat{T}$ such that $T(x) \cap B \subset \widehat{T}(x) \subset B$. Furthermore,*

$$(5.1) \quad \widehat{T}(x) = \{x^* \in B \mid \langle x^* - y^*, x - y \rangle \geq 0 \text{ for all } y^* \in T(y) \cap \text{int } B\}.$$

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 eventually show in Proposition 20 to be $\widehat{S}$ itself.

PROPOSITION 17. Define $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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, (5.1) 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 $t \rightarrow 0$ 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

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

for all $\|z\| \leq 1$. Let $z = (\frac{1}{a}, -\frac{\sqrt{a^2-1}}{a})$ so that $\widehat{S}(z) = S(z) = (\frac{\sqrt{a^2-1}}{a}, \frac{1}{a})$. 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 $z = (\frac{1}{a}, -\frac{\sqrt{a^2-1}}{a})$ also shows that $u = \sqrt{1 - 1/a^2}$.

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 commence to 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 $\|x\| = 1$.

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

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 $S^2 = -I$. $\square$

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

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

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

However, such mappings do exist for each even-dimensional $\mathbb{R}^{2n}$, and these can be embedded in $\mathbb{R}^{2n+1}$ in an obvious way. Thus, our construction provides an acyclic nonlinear mapping for each $\mathbb{R}^n$, $n > 1$.

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 we 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

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

and 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 . $\square$

PROPOSITION 20. *The extension mapping $\widehat{S}$ given explicitly in Proposition 17 is nonlinear and acyclic with bounded range and full domain.*

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 are 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 $y \rightarrow \widehat{x}(y)$ is continuous. We will determine $f(y)$ by integrating F along the ray $s \rightarrow x + sS(x)$. 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 $y \rightarrow c(\widehat{x}(y))$ 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

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

This means that

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

or

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

for all $\lambda \in (0, 1)$. Letting $\lambda \rightarrow 0$, we get $\widehat{x}((1+2\lambda)y) \rightarrow y$, so $c(\widehat{x}((1+\lambda)y)) \rightarrow c(y) > 0$. Since $\sqrt{1+2/\lambda} \rightarrow \infty$, 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. $\square$

It seems probable that the construction above applied to any nontrivial skew linear mapping always leads to an acyclic mapping—and that more ingenuity will allow some reader to prove this. We conclude the paper by exploring Fitzpatrick's last function for $\widehat{S}$ as above.

6. Computing $f_{\widehat{S}}$. We can also explicitly compute Fitzpatrick's last function $f_{\widehat{S}}$ as in the previous section. We have the following proposition.

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

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

Note that $f_{\widehat{S}}$ is convex, since it is a composition of the norm $x \rightarrow \|x\|$ with the increasing convex function $t \rightarrow \int_1^t \sqrt{1 - 1/s^2} ds$. 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 $SL = x \rightarrow h(\|x\|)S(x)$ is clearly skew-like. Note finally that SL is not monotone.

7. 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 all monotone operators defined on the space inherit these properties? Are such properties determined by the behavior of the acyclic part? 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? (2) Is there an effective characterization of acyclicity that allows one to easily determine whether a given operator is acyclic? (3) When is the sum of acyclic mappings acyclic? (4) Can one exhibit an acyclic mapping whose domain is not the whole space?

Acknowledgments. We would like to thank Heinz Bauschke and the referees whose very careful reading greatly improved the paper’s exposition.

REFERENCES

- [1] E. ASPLUND, *A monotone convergence theorem for sequences of nonlinear mappings*, Proc. Sympos. Pure Math., 18 (1970), pp. 1–9.
- [2] S. BARTZ, H. H. BAUSCHKE, J. M. BORWEIN, S. REICH, AND X. WANG, *Fitzpatrick functions, cyclic monotonicity, and Rockafellar’s antiderivative*, Nonlinear Anal., 66 (2007), pp. 1198–1223.
- [3] H. H. BAUSCHKE, J. M. BORWEIN, AND X. WANG, *Fitzpatrick functions and continuous linear monotone operators*, SIAM J. Optim., 18 (2007), pp. 789–809.
- [4] J. M. BORWEIN, *Maximal monotonicity via convex analysis*, J. Convex Anal., 13 (2006), pp. 561–586.
- [5] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley and Sons, New York, 1983.
- [6] H. DEBRUNNER AND P. FLOR, *Ein Erweiterungssatz für monotone Mengen*, Arch. Math., 15 (1964), pp. 445–447.
- [7] S. FITZPATRICK AND R. R. PHELPS, *Bounded approximants to monotone operators on Banach spaces*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 9 (1992), pp. 573–595.
- [8] S. FITZPATRICK AND R. R. PHELPS, *Some properties of maximal monotone operators on non-reflexive Banach spaces*, Set-Valued Anal., 3 (1995), pp. 51–69.
- [9] R. R. PHELPS, *Convex Functions, Monotone Operators, and Differentiability*, Lecture Notes in Math. 1364, Springer-Verlag, Berlin, 1989.
- [10] R. T. ROCKAFELLAR, *Characterization of the subdifferentials of convex functions*, Pacific J. Math., 17 (1966), pp. 497–510.
- [11] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [12] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
- [13] S. SIMONS, *Subdifferentials are locally maximal monotone*, Bull. Austral. Math. Soc., 47 (1993), pp. 465–471.
- [14] S. SIMONS, *Minimax and Monotonicity*, Lecture Notes in Math. 1693, Springer-Verlag, New York, 1998.