Continuous Linear Monotone Operators on Banach Spaces

Heinz H. Bauschke* and Jonathan M. Borwein†
Centre for Experimental & Constructive Mathematics,
Simon Fraser University, Burnaby, B.C. V5A 1S6, Canada

August 29, 1995

Abstract

The concept of a monotone operator — which covers both linear positive semi-definite operators and subdifferentials of convex functions — has turned out to be very powerful in various branches of mathematics. Over the last few decades, several new notions of monotonicity have been introduced: Gossez’ maximal monotone of type (D), Simons’ monotone of type (WD) and of type (N1), Fitzpatrick and Phelps’ locally maximal monotone. While these monotonicities are automatic for maximal monotone operators in reflexive Banach spaces and for subdifferentials of convex functions, their precise relationship is largely unknown. In view of the origin of the theory of monotone operators, it is very natural to investigate linear monotone (i.e. positive semi-definite) operators. Here, it is shown — within the beautiful framework of Convex Analysis — that for continuous linear monotone operators, all these notions coincide and are equivalent to the monotonicity of the conjugate operator. The latter condition is analyzed and illustrated by several examples. Some nonlinear results on regularizations conclude the paper.

1991 M.R. Subject Classification.
Primary 47H05; Secondary 26A51, 46B10, 47B25.

Key words and phrases. positive semi-definite operator, antisymmetric operator, skew-symmetric operator, symmetric operator, monotone operator, monotone operator of type (D), locally maximal monotone operator, weakly compact operator, property (w), duality mapping, regularization.

*Electronic mail: bauschke@cem.sfu.ca.
†Electronic mail: jborwein@cem.sfu.ca. Research supported by NSERC and by the Shrum Endowment.
1 Introduction

Motivation
A monotone operator is a (possibly set-valued) map from a Banach space to its dual satisfying a certain relation. In the simplest case, when the space is just the real line, this relation corresponds to monotonically increasing functions, hence the name. Monotone operators appear in quite diverse areas such as Operator Theory, Numerical Analysis, Differentiability Theory of Convex Functions, and Partial Differential Equations, because the notion of a monotone operator is broad enough to cover two fundamental mathematical objects: linear positive semi-definite operators and subdifferentials of convex functions. Although the former object gave rise to the field, it is the latter that has been receiving most of the recent attention.

The urge to extract and study the quite strong monotonicity properties of subdifferentials of convex functions has led to the introduction of several new (and stronger) notions of monotonicity. While these notions are automatic for maximal monotone operators on reflexive Banach spaces, the situation in nonreflexive Banach spaces is far less understood. As a consequence, almost all applications of Monotone Operator Theory (for instance, in Partial Differential Equations) are restricted to the reflexive case — albeit the natural setting of these applications is often nonreflexive.

Quite surprisingly to us, these notions of monotonicity were largely untested even for the most natural candidates: continuous linear positive semi-definite operators. Thus:

The aim of this paper is to study the various notions of monotonicity for continuous linear positive semi-definite operators.

Using elegant and potent tools of Convex Analysis, we show that these notions all coincide with the monotonicity of the conjugate operator. Structure theorems from Banach Space Theory then imply that monotonicity of the conjugate operator seems to be the rule — with the notable exception of spaces that contain a complemented copy of the sequence space $\ell_1$. Some nonlinear results conclude the paper indicating that even for (nonlinear multi-valued) regularizations, the notions of monotonicity may well all coincide.

Monotone Operator Theory has attracted many eminent mathematicians who have contributed numerous important results; as we do not even attempt to give a historical account here (but see [31]), we want at least refer the reader to some basic literature: the conference proceedings [17, 5, 52] and the books [34, 53, 54]; for applications, see [10, 16, 33].

Outline
The first section continues with subsections on notation, basic facts, and Convex Analysis tools. The various notions of monotonicity by Gossez (dense type or
type (D) \[19\], by Phelps and Fitzpatrick (locally maximal monotone \[14\]), and by Simons (range-dense type or type (WD) and type (NI) \[45\]) are introduced and some of their basic relationships are reviewed. Fenchel's Duality Theorem and examples of conjugate functions are provided; for completeness, a self-contained proof of a useful result by Gossez on the weak* closure of a closed convex set (see Proposition 1.29) is included.

From Section 2 on, we focus on the case when the monotone operator is continuous and linear. Preliminary results are collected here: the key concepts are: Gossez' explicit computation of the graph of a certain extension (Theorem 2.4) and the simple yet crucial decomposition of the operator into a symmetric and a skew part (Proposition 2.14). The symmetric part is \( - \) as the gradient of a quadratic function \( - \) extremely well-behaved. The skew part nonetheless allows fairly explicit descriptions of the various extension types.

Section 3 contains the main result: given a continuous linear monotone operator, all notions of monotonicity are equivalent to the monotonicity of the conjugate operator or to the monotonicity of the conjugate of the skew part (Theorem 3.3). Therefore, the monotonicity behaviour of a continuous linear monotone operator is determined by its skew part. As a consequence, we give a \( ( \text{partial} ) \) affirmative answer to a question posed by Gossez more than two decades ago. It also follows that skewness of the conjugate of the skew part of the operator is equivalent to weak compactness (Theorem 3.5) \( - \) thus, in some sense, such monotonicity is \( \text{“one half of weak compactness”} \).

Examples are given in Section 4. Classical counter-examples by Gossez and by Fitzpatrick and Phelps (Example 4.2, Example 4.5) are presented as particularizations of Theorem 4.1 which is an \( “\text{instruction manual”} \) for constructing interesting linear monotone operators whose skew parts have nonmonotone conjugates.

In Section 5, we coin the notion of a convex monotone space \( (\text{cms}) \); these are spaces on which every continuous linear skew operator is weakly compact. This notion is compared to the stronger property \( (\text{w}) \), introduced by Saab and Saab. They showed that a Banach lattice is \( (\text{w}) \) if and only if it does not contain a complemented copy of \( \ell_1 \). Thus many classical Banach spaces are \( (\text{w}) \) and hence \( (\text{cms}) \) \( - \) with the exception of the spaces \( \ell_1 \) and \( L_1[0,1] \) which host \( “\text{bad”} \) operators given by Gossez and by Fitzpatrick and Phelps, respectively. Regularizations, i.e. \( (\text{potentially nonlinear multivalued}) \) perturbations by positive multiples of the duality map, are investigated in the final Section 6. The results demonstrate the close relationship between local maximal monotonicity and monotonicity of range-dense type even in this nonlinear context. The analysis is made very transparent and elegant by extracting a certain property of the duality map which we call ruggedness \( (\text{see, for instance, the alternative proof of a result by Gossez} 6.33) \).
Notation

The notation we employ here is fairly standard. Throughout the paper, we assume that

| X is a real Banach space with norm \( \| \cdot \| \) and dual space \( X^* \). |
| If \( x^* \in X^* \) and \( x \in X \), then \( \langle x^*, x \rangle \) is the value of \( x^* \) at \( x \). |

We often view \( X \) as a subspace in its bidual \( X^{**} := (X^*)^* \). The unit ball \( \{ x \in X : \| x \| \leq 1 \} \) is denoted \( B_X \). If \( (x^n) \) is a net in some dual space, then we write \( x^n \rightharpoonup x^* \) (resp. \( x^n \to x^* \)) to indicate convergence in the weak* (resp. norm) topology with limit point \( x^* \). If \( x, y \in X \), then \( [x, y] \) stands for the line segment \( \{ \lambda x + (1 - \lambda)y : \lambda \in [0, 1] \} \). If \( U \) (resp. \( V \)) is a subset of \( X \) (resp. \( X^* \)), then \( U^\perp \) (resp. \( V^\perp \)) stands for the annihilator \( \{ x^* \in X^*: \langle x^*, u \rangle = 0, \forall u \in U \} \) (resp. \( \{ x \in X : \langle x, v \rangle = 0, \forall v \in V \} = X \cap V^\perp \)). If \( T \) is a continuous linear operator from \( X \) to some other Banach space, then the conjugate (or adjoint, transpose) is denoted \( T^* \), the restriction of \( T \) to some subset \( U \) of \( X \) is written as \( T|_U \), and \( \ker T \) is the kernel (or null space) of \( T \) : \( \ker T := \{ x \in X : Tx = 0 \} \). Suppose \( C \) is a subset of \( X \). Then \( \text{span} C \) is the span of \( C \) (i.e. the set of all linear combinations of elements of \( C \)). Also, \( \text{cl} C \) (resp. int \( C \)) stands for the closure (resp. interior) of \( C \); here, the norm topology is the “default topology”.

If these operations are meant with respect to some other topology \( T \), then we indicate this by subscripts; for instance, \( \text{cl}_T C \) would be the closure of \( C \) with respect to the topology \( T \).

Notation from Convex Analysis appear throughout the paper. The indicator function of \( C \) is denoted \( 1_C \). If \( C \) is a convex cone, then \( \text{lin} C := C \cap (-C) \) is the linearity space of \( C \). Suppose \( f \) is a convex function on \( X \). Then \( \text{dom} f \) (resp. \( f^*, \partial f, \nabla f \)) stands for the essential domain (resp. conjugate, subdifferential, (Gâteaux-) gradient) of \( f \). Note that \( f^* \) is defined on \( X^* \) and hence \( f^{**} := (f^*)^* \) is defined on \( X^{**} \). If \( g \) is another convex function on \( X \), then \( f \oplus g \) is the infimal convolution of \( f \) and \( g \). Finally, the reals (resp. strictly positive integers \( \{ 1, 2, 3, \ldots \} \)) are abbreviated \( \mathbb{R} \) (resp. \( \mathbb{N} \)) and we used already \( \forall \) (resp. \( \exists \)) as a short form for “for all” (resp. “there exists”).

As general references on Functional Analysis, we recommend [26, 27, 51]; for more on Convex Analysis, see [37, 2, 11, 24, 25].

Basic Facts

**Reminder 1.1** Suppose \( Y, Z \) are sets and \( T \) is a set-valued map from \( Y \) to \( Z \), i.e. \( T \) is a map from \( Y \) to \( 2^Z \). Then the graph of \( T \) is denoted \( \text{gra} T \); so \( z \in Ty \) if and only if \( (y, z) \in \text{gra} T \), \( \forall y \in Y, z \in Z \). The domain (resp. range) of \( T \) is given by \( \text{dom} T := \{ y \in Y : Ty \neq \emptyset \} \) (resp. \( \text{ran} T := \{ z \in Z : z \in Ty, \text{ for some } y \in Y \} \)). The inverse of \( T \), denoted \( T^{-1} \), is the set-valued map from \( Z \) to \( Y \) defined by \( y \in T^{-1} z \) if and only if \( z \in Ty, \forall y \in Y, z \in Z \). If \( U \) is a subset of \( Y \), then we write \( T(U) \) for \( \bigcup_{u \in U} Tu \).
**Reminder 1.2** A set-valued map $T$ from $X$ to $X^*$ is a *monotone operator*, if

$$\langle x^* - y^*, x - y \rangle \geq 0, \quad \forall (x, x^*), (y, y^*) \in \text{gra} T.$$

If $T$ is a monotone operator from $X$ to $X^*$ for which $\text{gra} T$ is a maximal subset of $X \times X^*$ with respect to set-inclusion, then $T$ is called *maximal monotone*.

**Remark 1.3** A straight-forward application of Zorn’s Lemma guarantees maximal monotone extensions for any given monotone operator. We can analogously speak of (maximal) monotone operators from $X^*$ to $X$ or from $X^{**}$ to $X^*$ or of monotone operators whose graphs are maximal monotone with respect to some subsets and so forth.

The following extensions have been turned out to be useful when studying the non-reflexive case.

**Definition 1.4** Suppose $T$ is a set-valued map from $X$ to $X^*$. Define set-valued maps from $X^{**}$ to $X^*$ via their graphs as follows:

1. [19, Section 2] $(x^{**}, x^*) \in \text{gra} T_1$, if there exists a bounded net $(x_\alpha, x_\alpha^*)$ in $\text{gra} T$ with $x_\alpha \rightharpoonup x^{**}$ and $x_\alpha^* \rightharpoonup x^*$.
2. $(x^{**}, x^*) \in \text{gra} T_0$, if $\inf_{(y, y^*) \in \text{gra} T} \langle y^* - x^*, y - x^{**} \rangle = 0$.
3. [35, Section 3] $(x^{**}, x^*) \in \text{gra} \overline{T}$, if $\inf_{(y, y^*) \in \text{gra} T} \langle y^* - x^*, y - x^{**} \rangle \geq 0$; in other words: $\text{gra} \overline{T}$ consists of all pairs in $X^{**} \times X^*$ that are *monotonically related* to $\text{gra} T$.

**Remarks 1.5** Suppose $T$ is a set-valued map from $X$ to $X^*$.

1. $T$ is monotone if and only if $\text{gra} T \subseteq \text{gra} \overline{T}$.
2. If $T$ is monotone, then so is $T_1$.
3. $T$ is maximal monotone if and only if $\text{gra} T = (\text{gra} \overline{T}) \cap (X \times X^*)$.
4. If $T$ is maximal monotone, then $\overline{T}$ need not be monotone; see Remark 4.8.

**Proposition 1.6** Suppose $T$ is a monotone operator from $X$ to $X^*$. Then the following inclusions hold in $X^{**} \times X^*$:

$$\text{gra} T \subseteq \text{gra} T_1 \subseteq \text{gra} T_0 \subseteq \text{gra} \overline{T} \subseteq \text{gra} \overline{T_1}.$$ 

**Proof.** $\text{gra} T \subseteq \text{gra} T_1$ and $\text{gra} T_0 \subseteq \text{gra} \overline{T} \subseteq \text{gra} \overline{T_1}$ are obvious (even without monotonicity). Fix an arbitrary $(x^{**}, x^*) \in \text{gra} T_1$ and obtain a bounded net $(x_\alpha, x_\alpha^*)$ in $\text{gra} T$ with $x_\alpha \rightharpoonup x^{**}$ and $x_\alpha^* \rightharpoonup x^*$. Then $\langle x_\alpha - y, x_\alpha^* - y^* \rangle \geq 0, \forall \alpha \forall (y, y^*) \in \text{gra} T_1$; taking limits yields $\langle x^{**} - y, x^* - y^* \rangle \geq 0$. On the other hand, $(x^{**} - x_\alpha, x^* - x_\alpha^*) \rightharpoonup 0$; altogether, $0 \geq \inf_{(y, y^*) \in \text{gra} T} \langle x^{**} - y, x^* - y^* \rangle$, i.e. $(x^{**}, x^*) \in \text{gra} T_0$. Hence $\text{gra} T_1 \subseteq \text{gra} T_0$. Finally, pick $(z^{**}, z^*) \in \text{gra} \overline{T}$. Then $0 \leq \inf_{(y, y^*) \in \text{gra} T} \langle y^* - z^*, y - z^* \rangle \leq \lim_{\alpha} \langle x_\alpha - z^*, x_\alpha^* - z^{**} \rangle = \langle x^{**} - z^*, x^* - z^{**} \rangle$; so $(z^{**}, z^*)$ is monotonically related to $\text{gra} T_1$, hence $\text{gra} \overline{T} \subseteq \text{gra} \overline{T_1}$. 

These basic inclusions allow a nice motivation of the various types of monotonicity:
Definition 1.7 Suppose $T$ is a monotone operator from $X$ to $X^*$. Then:

(i) (Gossez’ [20]) $T$ is of dense type or of type (D), if $T_1 = T$.  
(ii) (Simons’ [45, Definition 14]) $T$ is of range-dense type or of type (WD), if for every $x^* \in \text{ran } T$, there exists a bounded net $(x_n, x_n^*) \in \text{gra } T$ with $x_n^* \to x^*$.  
(iii) (Simons’ [45, Definition 10]) $T$ is of type (NI), if $\inf_{(y, y^*) \in \text{gra } T} \langle y^* - x^*, y - x^* \rangle \leq 0$, for all $(x^{**}, x^*) \in X^{**} \times X^*$. If this holds only on some subset of $X^{**} \times X^*$, then we say that $T$ is of type (NI) with respect to this subset.  
(iv) (Fitzpatrick and Phelps’ [14, Section 3]) $T$ is locally maximal monotone, if $(\text{gra } T^{-1}) \cap (V \times X)$ is maximal monotone in $V \times X$, for every convex open set $V$ in $X^*$ with $V \cap \text{ran } T \neq \emptyset$.  
(v) (Fitzpatrick and Phelps’ [15, Section 3]) $T$ is maximal monotone locally, if $(\text{gra } T) \cap (U \times X^*)$ is maximal monotone in $U \times X^*$, for every convex open set $U$ in $X$ with $U \cap \text{dom } T \neq \emptyset$.  
(vi) $T$ is unique, if all maximal monotone extensions of $T$ in $X^{**} \times X^*$ coincide.

Fact 1.8 (see, for instance, [35, Example 3.2. (b) and Proposition 4.4]) Suppose $X$ is reflexive and $T$ is a monotone operator from $X$ to $X^*$. Then TFAE: (i) $T$ is maximal monotone; (ii) $T$ is maximal monotone and of dense type; (iii) $T$ is locally maximal monotone.

Fact 1.9 (Simons’ [45, Theorem 19]) Suppose $T$ is a monotone operator from $X$ to $X^*$. Then TFAE:

(i) $T$ is unique.  
(ii) $\overline{T}$ is the unique maximal monotone extension of $T$ in $X^{**} \times X^*$.  
(iii) $\overline{T}$ is maximal monotone.  
(iv) $\overline{T}$ is monotone.

Moreover: If $T$ is of type (NI), then (i)–(iv) hold.

Remark 1.10 A unique maximal monotone operator need not be of type (NI): consider the operator $G$ in Example 4.2. This example also shows that Simons’ [45, Theorem 19] cannot be improved (meaning that we cannot add the condition “$T$ is of type (NI)” to the list of characterizations of uniqueness in Fact 1.9).

The following characterizations are sometimes more handy to work with:

Proposition 1.11 Suppose $T$ is a monotone operator from $X$ to $X^*$. Then:

(i) $T$ is of dense type if and only if $T_1$ is maximal monotone.  
(ii) $T$ is of range-dense type if and only if $\text{ran } T_1 = \text{ran } \overline{T}$.  
(iii) $T$ is of type (NI) if and only if $T_0 = \overline{T}$.
(iv) (Phelps’ [35, Proposition 4.3]) \( T \) is locally maximal monotone if and only if for every weak* closed convex bounded subset \( C \) of \( X^* \) with \( \text{ran} T \cap \text{int} C \neq \emptyset \), and for every \( x_0 \in X \), \( x_0^* \in (\text{int} C) \setminus T x_0 \), there exist \( (z, z^*) \in \text{gra} T \cap (X \times C) \) with \( \langle z^*, z - x_0 \rangle < 0 \).

(v) \( T \) is maximal monotone locally if and only if for every bounded closed convex subset \( C \) of \( X \) with \( \text{dom} T \cap \text{int} C \neq \emptyset \), and for every \( x_0 \in \text{int} C \), \( x_0^* \in X^* \setminus T x_0 \), there exists \( (z, z^*) \in \text{gra} T \cap (C \times X^*) \) with \( \langle z^*, z - x_0 \rangle < 0 \).

**Proof.** (i): \( \Rightarrow \): \( T \) is of dense type \( \iff T_1 = \overline{T} \). Now \( T_1 \) is monotone (Remarks 1.5.(ii)), hence so is \( T \). By Fact 1.9, \( T = T_1 \) is maximal monotone. \( \Leftarrow \): Pick \( (x^{**}, x^*) \in \text{gra} \overline{T} \). Then (by Proposition 1.6) \( (x^{**, x^*}) \in \text{gra} T_1 \), i.e. this point is monotonically related to \( \text{gra} T_1 \). Now \( T_1 \) is maximal monotone, hence \( (x^{**}, x^*) \in \text{gra} T_1 \). (ii): \( \Rightarrow \): Pick \( x^* \in \text{ran} \overline{T} \). By assumption, there exists a bounded net \( (x_0, x_0^*) \) in \( \text{gra} \overline{T} \) such that \( x_0^* \to x^* \). Without loss, we can assume that \( x_0^* \to x^{**} \). Then \( (x^{**, x^*}) \in \text{gra} T_1 \) and in particular \( x^* \in \text{ran} T_1 \). \( \Leftarrow \) is even simpler. (iii): Let us abbreviate \( \inf_{(y, y^*) \in \text{gra} T \setminus x^{**} = y, x^* = y^*} I \) by \( I \).

\( \Rightarrow \): If \( (x^{**, x^*}) \in \text{gra} T \), then \( I \geq 0 \). Now \( T \) is of type (NI), hence \( I \leq 0 \). Thus \( I = 0 \). \( \Leftarrow \): Fix \( (x^{**, x^*}) \in X^{**} \times X^* \). If \( (x^{**, x^*}) \notin \text{gra} T \), then \( I < 0 \). Otherwise, \( (x^{**, x^*}) \in \overline{T} = T_0 \) and hence \( I = 0 \).

Finally, (v) is proved analogously to (iv). \( \square \)

**Proposition 1.12** Suppose \( T \) is a monotone operator from \( X \) to \( X^* \). Then \( \text{ran} \overline{(T \setminus T_0)} \subseteq \text{ran} \overline{T} \setminus \text{ran} T_1 \). (Here \( T \setminus T_0 \) is defined via its graph: \( \text{gra} (T \setminus T_0) := (\text{gra} \overline{T}) \setminus (\text{gra} T_0) \).

**Proof.** Let \( x^* \in \text{ran} \overline{(T \setminus T_0)} \), i.e. there is some \( x^{**} \in X^{**} \) such that \( (x^{**, x^*}) \in (\text{gra} \overline{T}) \setminus (\text{gra} T_0) \). Let \( \inf_{(y, y^*) \in \text{gra} T \setminus x^{**} = y, x^* = y^*} I \) be the contrary that \( x^* \in \text{ran} T_1 \), say \( (y^{**, x^*}) \in \text{gra} T_1 \), for some \( y^{**} \in X^{**} \). Then we would obtain a bounded net \( (y_0, y_0^*) \) in \( \text{gra} T \) with \( y_0^* \to y^{**} \) and \( y_0^* \to x^* \). It follows that \( \langle x^{**} = y_0, x^* = y_0^* \rangle \to \langle x^{**} = y^{**}, x^* = x^* \rangle = 0 \) implying the contradiction: \( \inf_{(y, y^*) \in \text{gra} T \setminus x^{**} = y, x^* = y^*} I \leq 0 \). \( \square \)

The following implications are due to Simons; to illustrate Proposition 1.11, we prove some of them.

**Fact 1.13** (Simons’ [45, Lemma 15 and Theorem 19]) For any monotone operator from \( X \) to \( X^* \), the following implications hold:

\[
\text{dense type } \Rightarrow \text{range-dense type } \Rightarrow \text{type (NI)} \Rightarrow \text{unique.}
\]

**Proof.** Suppose \( T \) is a monotone operator from \( X \) to \( X^* \). Using Definition 1.7, Proposition 1.11, and Proposition 1.12, we obtain the following cascade of implications: \( T \) is of dense type \( \iff T_1 = \overline{T} \iff \text{ran} T_1 = \text{ran} \overline{T} \iff T \) is of range-dense
type \Rightarrow \text{ran}(T \setminus T_0) = \emptyset \iff T_0 = T \iff T \text{ is of type (NI).} \text{ The remaining implication is contained in Fact 1.9.} \square

All these notions were introduced to generalize the properties of the subdifferential operator; so the following does not come as a surprise:

**Fact 1.14** Suppose \( f \) is convex lower semi-continuous proper function on \( X \). Then:

(i) (Gossez’ [18, Théorème 3.1]) \( \partial f \) is maximal monotone and of dense type; in fact, \( (\partial f)_1 = (\partial f^*)^{-1} \).

(ii) (Simons’ [46]) \( \partial f \) is locally maximal monotone.

(iii) (Fitzpatrick and Phelps’ [15, Corollary 3.4]) \( \partial f \) is maximal monotone locally.

**Fact 1.15** (Simons’ [45, Theorem 17]; Fitzpatrick and Phelps’ [14, Theorem 3.5]) Suppose \( T \) is a monotone operator from \( X \) to \( X^* \). If \( T \) is of range-dense type or locally maximal monotone, then \( \text{cl ran } T \) is convex.

**Fact 1.16** (Heisler; see [35, Remark after Problem 2.20]) Suppose \( T \) and \( T' \) are maximal monotone operators from \( X \) to \( X^* \) with \( \text{dom } T = \text{dom } T' = X \). Then \( T + T' \) is maximal monotone.

**Definition 1.17** Suppose \( T \) is a set-valued map from \( X \) to \( X^* \). Then \( T \) is called **coercive** (resp. **bounded**), if

\[
\lim_{\|x\| \to +\infty} \inf \frac{1}{\|x\|} \langle Tx, x \rangle = +\infty \text{ (resp. } T(B) \text{ is bounded, } \forall \text{ bounded } B \in 2^X) \text{.}
\]

It is easy to see that the following implications hold: \( \text{dom } T \) is bounded \( \Rightarrow \) \( T \) is coercive \( \Rightarrow T^{-1} \) is bounded. Neither implication can be reversed in general: indeed, let \( X := \mathbb{R}^2 \) and consider first the monotone operator \( T(x_1, x_2) := (-x_2, x_1) \). Then \( T^{-1} \) is bounded but \( T \) is not coercive. Also, the identity is coercive but \( \text{dom } T \) is unbounded.

**Fact 1.18** (Gossez’ [18, Théorème 8.1]) Suppose \( T \) is a coercive monotone operator from \( X \) to \( X^* \). If \( T \) is of dense type, then \( \text{ran } T = \text{cl ran } T = X^* \).

**Fact 1.19** (Phelps’ [35, Theorem 4.8.(ii),(iii)]) Suppose \( T \) is a coercive maximal monotone operator from \( X \) to \( X^* \). If \( \text{cl ran } T = X^* \), then \( T \) is locally maximal monotone.

As Phelps pointed out [35, Corollary 4.9], combining Fact 1.18 and Fact 1.19 shows that every coercive maximal monotone operator of dense type is locally maximal monotone. We conclude this subsection with a generalization of this observation.
Theorem 1.20 Suppose $T$ is a maximal monotone operator from $X$ to $X^\ast$. Consider the following conditions: (i) $T$ is coercive and of dense type; (ii) $(\mathcal{T})^{-1}$ is bounded and $T$ is of range-dense type; (iii) $T$ is locally maximal monotone. Then: (i)\(\Rightarrow\)(ii)\(\Rightarrow\)(iii).

Proof. "(i)\(\Rightarrow\)(ii)": using the definition of $T_1$ and weak* lower semicontinuity of the norm in $X^{**}$, it is easy to prove that $T_1$ is coercive whenever $T$ is. Now $T$ is of dense type, thus $T_1 = \mathcal{T}$ and $T$ is of range-dense type; so $(\mathcal{T})^{-1}$ is bounded.

"(ii)\(\Rightarrow\)(iii)": a closer look at Phelps' proof of [35, Theorem 4.8.(ii)] reveals that already $\text{cl ran } T = X^\ast$ and $T^{-1}$ bounded imply local maximal monotonicity of $T$. Since $\text{cl ran } T \supseteq \text{ran } T_1 = \text{ran } \mathcal{T}$ ($T$ is of range-dense type) and since $T^{-1}$ is bounded (because $(\mathcal{T})^{-1}$ is), it suffices to establish the following

Keystep: $\text{cl ran } \mathcal{T} = X^\ast$.

Fix $x^* \in X^\ast$ and $(u, u^*) \in \text{gra } T$ and let $(\lambda_n)$ be any sequence of strictly positive reals tending to 0. Simons' [45, Theorem 12.(a)] yields the existence of sequences $(x_n^*)$ in $X^{**}$ and of $(y_n^*)$, $(x_n^*)$ in $X^\ast$ with: $x^* = y_n^* + \lambda_n x_n^*$, $y_n^* \in \mathcal{T}x_n^{**}$, $\|x_n^*\|^2 = \|x_n^{**}\|^2 = \langle x_n^*, x_n^* \rangle$, $\forall n \in \mathbb{N}$. Also, by definition of $\mathcal{T}$, $\langle y_n^* - u^*, x_n^* - u \rangle \geq 0$, $\forall n \in \mathbb{N}$. Altogether,

$$\lambda_n\|x_n^{**}\|^2 = \langle \lambda_n x_n^*, x_n^* \rangle = \langle x^* - y_n^*, x_n^* \rangle \leq \langle x^*, x_n^* \rangle + \langle u^*, u^* - y_n^* \rangle - \langle u^*, x_n^* \rangle \leq (\|x^*\| + \|u^*\|)\|x_n^{**}\| + \|u^*\|\|u^*\| + \|y_n^*\|).$$

Claim: $(\lambda_n x_n^{**})$ is bounded.

Otherwise, WLOG $\lambda_n\|x_n^{**}\| \to +\infty$ (rename an appropriate subsequence). Hence $\|x_n^{**}\| \to +\infty$ and dividing the displayed chain of inequalities by $\|x_n^{**}\|$ yields $+\infty \leftarrow \|y_n^*\|/\|x_n^{**}\|$. On the other hand, $\|y_n^*\|/\|x_n^{**}\| = \|x^* - \lambda_n x_n^*\|/\|x_n^{**}\| \to 0$, an impossibility. The claim thus holds.

It follows that $(\lambda_n x_n^{**})$ is bounded and so is $(y_n^*)$. Since $(\mathcal{T})^{-1}$ is bounded, we conclude that the sequence $(x_n^*)$ is actually bounded. Consequently, $\lambda_n x_n^* \to 0$ and hence $x^* - y_n^* \in \text{ran } \mathcal{T}$, as desired.

\[\Box\]

Remark 1.21 The operator $(x_1, x_2) \mapsto (-x_2, x_1)$ on $\mathbb{R}^2$ shows that condition (ii) of Theorem 1.20 is really more general than condition (i). However, since $\mathbb{R}^2$ is very reflexive, we know already by Fact 1.8 that $T$ is locally maximal monotone. A much more sophisticated example (in not necessarily reflexive spaces) can be obtained as follows:

Suppose $J$ is the duality map (see Reminder 6.7) on $X$ and $S$ is a continuous linear skew operator (see Reminder 2.12) on some other real Banach space $Y$. Suppose further $S^\ast$ is skew, $S^{-1}$ is bounded, but $S$ is not coercive. Let $T$ be the map from $X \times Y$ to $X^\ast \times Y^\ast$
defined by \( T(x,y) = (Jx, Sy), \forall (x,y) \in X \times Y \). Then: \( T \) is maximal monotone of dense type, \((T)^{-1}\) is bounded, but \( T \) is not coercive.

We omit the verification – which relies on results presented later – since this construction is not crucial. More concretely, we could let \( X := \ell_1, Y := \mathbb{R}^2 \), and \( S \) be as above to obtain an example on \( \ell_1 \times \mathbb{R}^2 \). Since the space \( \ell_1 \times \mathbb{R}^2 \) is isometrically isomorphic to \( \ell_1 \), the construction results altogether in an example on the nonreflexive space \( \ell_1 \).

**Convex Analysis Tools**

**Fact 1.22** [Fenchel’s Duality Theorem; [2, Theorem 4.6.1.(b)], [37, Section 31]] Suppose \( Y \) is another real Banach space with dual space \( Y^* \), \( T \) is a continuous linear operator from \( X \) to \( Y \), \( f \) is convex lower semi-continuous proper function on \( X \) as is \( g \) on \( Y \). Define

\[
p := \inf_{x \in X} f(x) + g(Tx) \quad \text{and} \quad d := -\inf_{y^* \in Y^*} \{ f^*(-T^*y^*) + g^*(y^*) \}.
\]

Then \( p \geq d \). If \( T(\text{dom} f) \cap \text{int} \text{dom} g \neq \emptyset \) and \( p \) is finite, then \( p = d \) and \( d \) is attained.

All we need later is collected in the following:

**Proposition 1.23** Suppose \( f \) is a convex lower semi-continuous proper function on \( X \), \( z^* \in X^* \), and \( r \in \mathbb{R} \). Let \( h(x) := f(x) + \langle z^*, x \rangle + r, \forall x \in X \). Then \( h^*(x^*) = -r + f^*(x^* - z^*), \forall x^* \in X^* \).

**Example 1.24** Suppose \( z^* \in X^* \) and \( r \in \mathbb{R} \). Let \( h(x) := \langle z^*, x \rangle + r, \forall x \in X \). Then \( h^*(x^*) = -r + \langle z^*, x^* \rangle, \forall x^* \in X^* \).

**Example 1.25** Suppose \( z_0, z_1 \in X \) and \( \epsilon > 0 \). Let \( h := \iota_C \), where \( C := [z_0, z_1] + \epsilon B_X \). Then \( h^*(x^*) = \epsilon \|x^*\| + \max \{\{x^*, z_0\}, \{x^*, z_1\}\}, \forall x^* \in X^* \).

**Example 1.26** (see [2, Proposition 4.4.8]) Let \( h(x) := \frac{1}{2}\|x\|^2, \forall x \in X \). Then \( h^*(x^*) = \frac{1}{2}\|x^*\|^2, \forall x^* \in X^* \).

**Example 1.27** Suppose \( C \) is a closed convex nonempty subset of \( X \) (viewed in \( X^{**} \)). Then \( \iota_{C^*} = \iota_{\text{cl} \text{weak}^* C} \).

**Proof.** By assumption, \( \iota_C \) is closed convex lower semi-continuous proper, hence \( \iota_{C^*}|_X = \iota_C \) (see, for instance, [2, Theorem 4.2.2]). Also, since \( 0 \leq \iota_C \), we have \( 0 \leq \iota_{C^*} \). Fix an arbitrary \( x^{**} \in X^{**} \).

**Case 1:** \( x^{**} \in \text{cl}_{\text{weak}^*} C \). Then there exists a net \((x_\alpha)\) in \( C \) with \( x_\alpha \rightharpoonup x^{**} \).

Thus \( 0 \leq \iota_{C^*}(x^{**}) \leq \liminf \alpha \iota_{C^*}(x_\alpha) = \lim \alpha \iota_C(x_\alpha) = 0 \).

**Case 2:** \( x^{**} \notin \text{cl}_{\text{weak}^*} C \). Since \( \text{cl}_{\text{weak}^*} C \) is convex, we can separate \( x^{**} \) from it (viewed in \( X^{**} \) equipped with the weak* topology): there exists \( x^* \in X^* \) with \( \langle x^{**}, x^* \rangle > \iota_{C^*}(x^*) \). It follows that \( \iota_{C^*} \geq \sup_{n \in \mathbb{N}} \{\langle x^{**}, nx^* \rangle - \iota_{C^*}(nx^*)\} = +\infty. \square
Remark 1.28 The reader should be warned that in various textbooks, statements like “if $C$ is a closed convex nonempty subset of $X$, then $i_C^* = i_C^*$” appear. This does not contradict Example 1.27 – the ambiguity stems from the fact that most authors require that the domain of biconjugates be just the original space $X$ rather than $X^{**}$.

Proposition 1.29 (see also Gossez’ [18, Corollaire 3.2]) Suppose $C$ is a closed convex nonempty subset of $X$ (viewed in $X^{**}$) and $x^{**} \in \text{cl}_{\text{weak}^*} C$. Then there exists a net $(x_\alpha)$ in $C$ with $x_\alpha \rightharpoonup x^{**}$ and $\|x_\alpha\| \to \|x^{**}\|$.

Proof. Fix an arbitrary $r > \|x^{**}\|$.

Step 1: $C \cap r B_X \neq \emptyset$.

Assume to the contrary that $\|x\| > r$, $\forall x \in C$. By Fact 1.22 and Example 1.26, there exists some $x^* \in X^*$ such that

$$0 > \frac{1}{2} \|x\|^2 + \frac{1}{2} \|x^{**}\|^2 + i_C(x^*)$$

$$\geq \frac{1}{2} \|x\|^2 + \frac{1}{2} \|x^{**}\|^2 + \langle x^*, x^{**} \rangle$$

$$\geq 0,$$

an impossibility. Step 1 is thus verified.

Step 2: $(i_{rB_X} + i_C)^* = i_{rB_X}^* \square i_C^*$.

Fix $z^* \in X^*$. Then, using Fact 1.22 and Proposition 1.23,

$$(i_{rB_X} + i_C)^*(z^*) = \sup_{x \in X} (z^*, x) - (i_{rB_X} + i_C)(x)$$

$$= - \inf [(i_{rB_X}(x) - (z^*, x)) + i_C(x)]$$

$$= \min_{x^* \in X^*} (z^*, (i_{rB_X}^* + i_C^*)(x^*)$$

$$= (i_{rB_X}^* \square i_C^*)(z^*).$$

Step 3: $\text{cl}_{\text{weak}^*}(rB_X \cap C) = rB_{X^{**}} \cap \text{cl}_{\text{weak}^*} C$.

Then (remember that Step 1 holds for any $r > \|x^{**}\|$) $C \cap \text{int} r B_X \neq \emptyset$. Recall that by Goldstine’s Theorem (see, for instance, [9, Chapter II]) $\text{cl}_{\text{weak}^*} r B_X = r B_{X^{**}}$. Hence:

$$i_{\text{cl}_{\text{weak}^*}(rB_X \cap C)} = i_{rB_{X^{**}} \cap C}$$

(by Example 1.27)

$$= (i_{rB_X} + i_C)^{**}$$

$$= (i_{rB_X}^* \square i_C^*)^*$$

(by Step 2)

$$= i_{rB_X}^{**} + i_C^*$$

(by Example 1.27)

$$= rB_{X^{**}} + i_{\text{cl}_{\text{weak}^*} C}$$

(by Goldstine’s Theorem)

$$= i_{rB_{X^{**}} \cap \text{cl}_{\text{weak}^*} C}.$$

Thus $\text{cl}_{\text{weak}^*}(rB_X \cap C) = rB_{X^{**}} \cap \text{cl}_{\text{weak}^*} C$, as desired.
Final Step: By Step 3, there exists a net \((x_\alpha)\) in \(C\) with \(\|x_\alpha\| \leq r\) and \(x_\alpha \overset{w^*}{\rightharpoonup} x^{**}\). The weak* lower semicontinuity of the norm yields \(\limsup \|x_\alpha\| \geq \|x^{**}\|\). Now let \(T\) be the weak topology on \(\text{cl}_{\text{weak}^*} C\) which makes the following functions continuous (see [50, Section 6.3] or [49, Section 9.3]): \(\langle x^*, \cdot \rangle\), \(\forall x^* \in X^*\) and \(\|\cdot\|\). Since \(r > \|x^{**}\|\) was chosen arbitrarily, every \(T\)-neighborhood \(V\) of \(x^{**}\) contains a point in \(C\), say \(x_V\). The net \((x_V)\) does the job.

\[\Box\]

**Remark 1.30** Suppose \(X\) is in addition separable and does not contain a copy of \(l_1\) (see Reminder 5.6). Then every bounded subset of \(X^{**}\) is weak* sequentially dense in its weak* closure (Rosenthal's [40, Theorem 2]). Thus, using in addition Step 3 in the proof of Proposition 1.29, one obtains the following interesting sequential variant of Proposition 1.29: if \(C\) is a closed convex nonempty subset of \(X \subseteq X^{**}\), and \(x^{**} \in \text{cl}_{\text{weak}^*} C\), then there exists a bounded sequence \((x_n)\) in \(C\) with \(x_n \overset{w^*}{\rightharpoonup} x^{**}\).

## 2 Continuous Linear Operator Tools

In this section, we study continuous linear (single-valued) operators in some detail and lay the ground work for the results to follow.

**Reminder 2.1** [51, Section 11-4] Suppose \(T\) is a continuous linear operator from \(X\) to another real Banach space \(Y\). Then \(T\) is weakly compact (resp. tauberian), if ran \(T^{**}\mid_{X^{**}\setminus X} \subseteq Y\) (resp. ran \(T^{**}\mid_{X^{**}\setminus X} \subseteq Y^{**}\setminus Y\)).

If \(X\) is reflexive, then every continuous linear operator from \(X\) to another real Banach space \(Y\) is (trivially) weakly compact and tauberian.

**Fact 2.2** (Gantmacher; see, for instance, [51, Theorem 11-4-2 and Theorem 11-4-4]) Suppose \(T\) is a continuous linear operator from \(X\) to \(X^*\). Then TFAE:

1. \(T\) is weakly compact;
2. \(\text{cl} T(B_X)\) is weakly compact;
3. \(T^*\) is weakly compact.

**Fact 2.3** (Wilansky's [51, Theorem 11-4-5]) Suppose \(T\) is a continuous linear operator from \(X\) to \(Y\). Consider the following conditions: (i) \(T\) is tauberian;
(ii) \(\ker T^{**} = \ker T\); (iii) \(\ker T\) is reflexive. Then: (i) \(\Rightarrow\) (ii) \(\Rightarrow\) (iii). Moreover, if ran \(T\) is closed, then (i), (ii), and (iii) are equivalent.

**Theorem 2.4** (see also Gossez' [19, End of Section 2]) Suppose \(T\) is a continuous linear operator from \(X\) to \(X^*\). Then:

1. \(\text{gra} T \subseteq (\text{gra} T^{**}) \cap (\text{gra} (T^*|_X)^*) \subseteq (\text{gra} T^{**}) \cap (X^{**} \times X^*) \subseteq \text{gra} T^{**}\).
2. \(\text{gra} T_1 = \text{cl}_{\text{weak}^*\times \|\cdot\|} \text{gra} T = (\text{gra} T^{**}) \cap (X^{**} \times X^*)\).
Proof. (i): is straightforward and thus omitted. (ii): we start with the second equality. Consider \( Y := (X^*, \text{weak}^*) \times (X^*, \| \cdot \|) \). Then \( Y^* = (X^*, \| \cdot \|) \times (X^*, \| \cdot \|) \) (\cite{26, Theorem 18.4}). We can compute the closure of \( \text{gra} T \in Y \) using the Bipolar-Theorem (\cite{6, Theorem V.1.8}): one first checks that \( (\text{gra} T)^\perp = \{(T^* x^*, x^*) : x^* \in X^*\} \) and then that \( \text{cl}_{\text{weak}^* \times \| \cdot \|} \text{gra} T = \perp (\text{gra} T)^\perp \) = \( (\text{gra} T)^\perp \cap (X^* \times X^*) \).

Now we turn our attention to the first equality. It is clear that \( \text{gra} T_i \subseteq \text{cl}_{\text{weak}^* \times \| \cdot \|} \text{gra} T \). Conversely, pick \( (x^*, x^*) \in \text{cl}_{\text{weak}^* \times \| \cdot \|} \text{gra} T \), i.e. \( x^* = T^* x^* \in X^* \) (by the second equality). There exists a net \( (x_n) \) in \( X \) with \( x_n \xrightarrow{\text{weak}^*} x^* \) and \( T x_n \to x^* \). Now fix an arbitrary positive integer \( n \) and an arbitrary weak* neighborhood \( V \) of 0 (in \( X^* \)). Let \( C_n := T^{-1}(x^* + \frac{1}{n} B_{X^*}) \). Then \( C_n \) is closed and convex. Moreover, \( C_n \) contains \( x_n \) for all large \( n \); hence \( x^* \in \text{cl}_{\text{weak}^*} C_n \). Proposition 1.29 (applied to \( C_n \)) guarantees the existence of a point \( x_{(n, V)} \) with:

\[
x_{(n, V)} - x^* \in V, \quad \|x_{(n, V)}\| = \|x^*\| \leq \frac{1}{n}, \quad \text{and} \quad \|T x_{(n, V)} - x^*\| \leq \frac{1}{n} .
\]

Let \( \mathcal{V} \) be the set of all weak* neighborhoods of 0 and consider the following binary relation on \( \mathbb{N} \times \mathcal{V} \): \( (n_1, V_1) \geq (n_2, V_2) \) if \( n_1 \geq n_2 \) and \( V_1 \subseteq V_2 \). Then \( (\mathbb{N}, \mathcal{V}) \) is a directed set. The net \( x_{(n, V)} \) is bounded with \( x_{(n, V)} \xrightarrow{\text{weak}^*} x^* \) and \( T x_{(n, V)} \to x^* \). Consequently, \( (x^*, x^*) \in \text{gra} T_i \) and the proof is complete. \( \square \)

Corollary 2.5 Suppose \( T \) is a continuous linear operator from \( X \) to \( X^* \). Then \( T \) is weakly compact (resp. tauberian) if and only if \( T_1 = T^* \) (resp. \( T_1 = T \)).

Let us now impose some monotonicity.

Remark 2.6 Suppose \( T \) is a continuous linear operator from \( X \) to \( X^* \). Then \( T \) is called positive or positive semi-definite, if \( (T x, x) \geq 0, \forall x \in X \).

The following result is part of the folklore:

Proposition 2.7 Suppose \( T \) is a continuous linear operator from \( X \) to \( X^* \). Then TFAE: (i) \( T \) is positive; (ii) \( T \) is monotone; (iii) \( T \) is maximal monotone.

Proof. For “(ii)\( \Rightarrow \) (iii)”, see \cite[Example 1.5.(b)]: the rest is obvious. \( \square \)

Remark 2.8 If \( T \) is a linear monotone operator from \( X \) to \( X^* \) with \( \text{dom} \ T = X \), then \( T \) is necessarily continuous (see, for instance, \cite[Proposition 26.4.(b)]{54}). Thus in subsequent sections, we could have written “T is a linear monotone operator from \( X \) to \( X^* \) with \( \text{dom} \ T = X \)” instead of “\( T \) is a continuous linear monotone operator from \( X \) to \( X^* \)”; however, we prefer the latter formulation.

Corollary 2.9 (Gossez’ \cite[End of Section 2]{19}) Suppose \( T \) is a continuous linear monotone operator from \( X \) to \( X^* \). If \( T \) is weakly compact, then \( T \) is maximal monotone and of dense type.
Proof. $T$ weakly compact means $T_1 = T^{**}$ (Corollary 2.5.(i)). On the one hand, $T_1$ is monotone (see Remarks 1.5.(ii)). On the other hand, $T^{**}$ is linear. Now $T_1 = T^{**}$ is continuous linear monotone, hence maximal monotone (Proposition 2.7). Consequently, $T$ is maximal monotone and of dense type (Proposition 1.11.(i)). \hfill \Box

Remark 2.10 There exists a continuous linear maximal monotone operator of dense type that is not weakly compact; thus, the reverse implication in Corollary 2.9 is false in general: consider $\neg G$ from Example 4.2. Nonetheless, maximal monotonicity of dense type and weak compactness are closely related; see Section 3.

Proposition 2.11 Suppose $T$ is a continuous linear operator from $X$ to $X^*$. Then $T$ is monotone and of type (NI) with respect to $\text{gra}(-T^*)$ if and only if $T^*$ is monotone.

Proof. Suppose momentarily that $T$ is monotone. Fix $x^{**} \in X^{**}$ and $x \in X$. Then $\langle Tx + T^*x^{**}, x - x^{**} \rangle = \langle Tx, x \rangle + \langle T^*x^{**}, x \rangle - \langle T^*x^{**}, x^{**} \rangle \geq -\langle T^*x^{**}, x^{**} \rangle$. Hence $-\langle T^*x^{**}, x^{**} \rangle \leq \inf_{x \in X} \langle Tx + T^*x^{**}, x - x^{**} \rangle \leq \langle T0 + T^*x^{**}, 0 - x^{**} \rangle = -\langle T^*x^{**}, x^{**} \rangle$ and thus:

$$\inf_{(y, y^*) \in \text{gra} T} \langle y^* - (-T^*x^{**}), y - x^{**} \rangle = \inf_{x \in X} \langle Tx + T^*x^{**}, x - x^{**} \rangle = -\langle T^*x^{**}, x^{**} \rangle.$$ 

The result follows readily. \hfill \Box

Symmetric and Skew Operators

Some of the results in this subsection are related to unpublished work of Dr. R. Phelps [36] on linear monotone operators that are not necessarily continuous.

Reminder 2.12 Suppose $T$ is a continuous linear operator from $X$ to $X^*$. Then $T$ is symmetric (resp. skew), if $T^*|_X = T$ (resp. $T^*|_X = -T$).

It should be noted that skew operators are sometimes also called “skew-symmetric” or “antisymmetric”. The following result is easily established:

Proposition 2.13 Suppose $T$ is a continuous linear operator from $X$ to $X^*$. Then:

(i) $T$ is symmetric $\iff \langle Tx, y \rangle = \langle Ty, x \rangle, \forall x, y \in X.$

(ii) $T$ is skew $\iff \langle Tx, y \rangle = -\langle Ty, x \rangle, \forall x, y \in X \iff T$ and $-T$ are monotone.

Symmetric and skew operators are building blocks in a sense made precise in the following easy-to-verify proposition (for a more general version; see [36]).
**Proposition 2.14** Suppose $T$ is a continuous linear operator from $X$ to $X^*$. Then $T$ can be written as the sum of two continuous linear operators, $T = P + S$, where $P$ is symmetric and $S$ is skew. This decomposition is unique; in fact:

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

The key to a further analysis lies in the decomposition $T = P + S$. From an monotone operator theory point of view, the symmetric part $P$ is very nice (equal to the subdifferential of a convex function, provided that $T$ is monotone; see the following subsection). Although the skew part $S$ is very monotone (see Proposition 2.13.(ii)), it is far away from the subdifferential of a convex function (recall that Hessians of convex functions are symmetric – not skew!). So it is not unexpected that the skew part $S$ plays a major role (see Section 3).

Let us now investigate the symmetric and the skew part in more detail.

**Continuous Linear Monotone Symmetric Operators**

We now exploit the fact that the symmetric part of a monotone operator is the gradient of a quadratic convex function. (See [36] for related generalizations.)

**Proposition 2.15** Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with symmetric part $P$. Let $q(x) := \frac{1}{2}\langle x, Tx \rangle, \forall x \in X$. Then: $q$ is convex $\Leftrightarrow T$ is monotone $\Leftrightarrow P$ is monotone. In this case, we have furthermore:

(i) $\nabla q = P$.

(ii) $q^* \circ P = q$.

(iii) ran $P \subseteq \text{dom} \; q^* \subseteq \text{cl} \text{ran} \; P$.

(iv) $q^*$ is strictly convex on ran $P$.

(v) $q^*$ is nonnegative and quadratic-homogeneous, i.e. $q^*(tx^*) = t^2q^*(x^*), \forall x^* \in X^*, \forall t \in \mathbb{R}$.

**Proof.** Since $q$ is continuous, it suffices to check midpoint convexity; fixing two arbitrary points $x, y \in X$, we have

$$ q\left(\frac{x+y}{2}\right) \leq \frac{1}{2}q(x) + \frac{1}{2}q(y) \Leftrightarrow 0 \leq \langle x-y, T(x-y) \rangle \Leftrightarrow 0 \leq \langle x-y, P(x-y) \rangle. $$

(i): Pick $(x, x^*) \in \partial q$ (possible, because $q$ is continuous and convex on $X$). Then $t\langle x^*, h \rangle \leq q(x + th) - q(x), \forall h \in X, t > 0$; this simplifies to $\langle x^*, h \rangle \leq \langle \frac{1}{2}Tx + \frac{1}{2}T^*x, h \rangle + \frac{1}{2}t\langle h, Th \rangle$. Letting tend $t$ to 0 yields $x^* = \frac{1}{2}Tx + \frac{1}{2}T^*x = Px$.

(ii): Fix $x_0 \in X$. Then

$$ q^*(Px_0) = \sup_{x \in X} \langle Px_0, x \rangle - q(x) = - \inf_{x \in X} \{q(x) + \langle -Px_0, x \rangle\}; $$

this last infimum can be viewed as a little optimization problem that is easy to solve: indeed, after taking gradients, we learn that the set of minimizers equals
$x_0 + \ker P$. It follows that $q^*(Px_0) = q(x_0)$.

(iii): The first inclusion follows from (ii). Now fix an arbitrary $x^* \in \text{dom } q^*$. Again, viewing $q^*$ as an optimization problem turns out to be useful: let $f(x) := q(x) - \langle x^*, x \rangle$, then

$$q^*(x^*) = \inf_{x \in X} \{q(x) - \langle x^*, x \rangle\} = \inf_{x \in X} f(x).$$

Fix $\epsilon > 0$. By [34, Lemma 3.22], there exists some $x \in X$ such that $\|\nabla f(x)\| = \|Px - x^*\| < \epsilon$. As $\epsilon$ was chosen arbitrarily, it follows that $x^* \in \text{cl ran } P$.

(iv): follows precisely as in [25, Theorem X.4.1.3].

(v): $q^*(x^*) \geq \langle x^*, 0 \rangle - q(0) = 0$, $\forall x^* \in X^*$. $q^*$ is quadratic-homogeneous, since $q$ is. $\square$

The next proposition complements Proposition 2.7. It also follows from Fitzpatrick and Phelps’ [15, Theorem 3.10]; however, here we provide a simpler Convex Analysis proof.

**Proposition 2.16** Suppose $T$ is a continuous linear operator from $X$ to $X^*$. Then $T$ is maximal monotone locally if and only if $T$ is monotone.

**Proof.** “$\Rightarrow$”: choose $U := X$ in Definition 1.7.(v). “$\Leftarrow$”: in view of Proposition 1.11.(v), let’s fix a bounded closed convex subset $C$ in $X$, $x_0 \in \text{int } C$, and $x_0^* \in X^* \setminus Tx_0$. Let $p := \inf_{x \in C} \frac{1}{2}(Tx - x_0^*, x - x_0)$. Our aim is $p < 0$. Since $C$ is bounded, the infimum $p$ is finite. Define $f(x) := q(x) - \langle \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0, x \rangle + \frac{1}{2}\langle x_0^*, x_0 \rangle$ and $g := \iota_C$, where $g(x) := \frac{1}{2}\langle x, Tx \rangle$, $\forall x \in X$. Then $q$ is convex (Proposition 2.15) and $p = \inf_{x \in X} f(x) + g(x)$. Now Fact 1.22 and Proposition 1.23 yield

$$p = -\inf_{x^* \in X^*} f^*(-x^*) + g^*(x^*) = \frac{1}{2}(x_0^*, x_0) - \inf_{x^* \in X^*} \{q^*(-x^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) + \iota_C^*(x^*)\};$$

moreover, the last infimum is attained, say at $y^* \in X^*$. Then, since $x_0$ is in the interior of $C$,

$$p = \frac{1}{2}(x_0^*, x_0) - q^*(-y^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) - \iota_C^*(y^*)$$

$$< \frac{1}{2}(x_0^*, x_0) - q^*(-y^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) - \langle y^*, x_0 \rangle$$

$$= \langle \langle -y^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0, x_0 \rangle - q^*(-y^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) - \frac{1}{2}(x_0, Tx_0) \rangle$$

$$= \langle (y^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0, x_0) - q^*(-y^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) - q(x_0) \rangle$$

$$\leq 0.$$

The proof is complete. $\square$

**Theorem 2.17** Suppose $P$ is a continuous linear monotone symmetric operator from $X$ to $X^*$. Then $P_1 = P_0 = P = P^* = P^{**}$. Consequently: $P$ is maximal monotone of dense type, weakly compact, locally maximal monotone, and maximal monotone locally; $P^*$ is monotone and symmetric.
**Proof.** By Proposition 2.15, $P$ is the subdifferential of the continuous convex function $\frac{1}{2}(x, Px)$. Hence (Fact 1.14 and Proposition 2.16) $P$ is maximal monotone and of dense type, locally maximal monotone, and maximal monotone locally; in particular, $P_1 = P_0 = \overline{P}$ and $P$ is of type (NI). It follows that on the one hand, $P^*$ is a maximal monotone extension of $P$ (Proposition 2.11 and Proposition 2.7); on the other hand, $\overline{P}$ is the unique maximal monotone extension of $P$ in $X^* \times X^*$ (Fact 1.9). Altogether, $\overline{P} = P^*$. Now $P^* = P^{**}$, because $P_1 = P^*$ and $\text{gra } P_1 \subseteq \text{gra } P^{**}$ (Theorem 2.4). Finally, the weak compactness of $P$ follows from Corollary 2.5.(i). □

The following corollary is used repeatedly:

**Corollary 2.18** Suppose $P$ is a continuous linear monotone symmetric operator from $X$ to $X^*$. Then for every $x^{**} \in X^{**}$, there exists a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \rightharpoonup x^{**}$ and $Px_\alpha \rightharpoonup P^*x^{**} = P^{**}x^{**}$.

**Proof.** $P_1 = P^* = P^{**}$. □

Theorem 2.17 allows us to strengthen parts of Proposition 2.15:

**Proposition 2.19** Suppose $P$ is a continuous linear monotone symmetric operator from $X$ to $X^*$. Let $q(x) := \frac{1}{2}(x, Px), \forall x \in X$. Then

$$q^*(P^*x^{**}) = \frac{1}{2}(x^{**}, P^{**}x^{**}) = q^{**}(x^{**}), \ \forall x^{**} \in X^{**}.$$  

Also, $\text{dom } \partial q^* = \text{ran } P^*$ and $\nabla q^* = P^{**} = P^*$.

**Proof.** Fix $x^{**} \in X^{**}$ and define $g(x) := \langle -P^*x^{**}, x \rangle + \frac{1}{2}(x^{**}, P^{**}x^{**}), \forall x \in X$. Then $(x^{**}, P^{**}x^{**}) \in \text{gra } P_0$ (Theorem 2.17) and hence

$$0 = \frac{1}{2} \inf_{x \in X} (Px - P^*x^{**}, x - x^{**}) = \inf_{x \in X} q(x) + g(x).$$

The conjugate of $g$ is given by (see Example 1.24) $g^*(x^*) = -\frac{1}{2}(x^{**}, P^*x^{**}) + \text{int}_{\{-P^*x^{**}\}}(x^*), \forall x^* \in X^*$. Fact 1.22 yields

$$0 = - \inf_{x^* \in X^*} \{q^*(x^*) + g^*(-x^*)\} = \frac{1}{2}(x^{**}, P^{**}x^{**}) - q^*(P^*x^{**}),$$

which is the first equality. To prove the second equality, we first note that the first equality implies

$$q^{**}(x^{**}) = \sup_{x^{**} \in X^{**}} \langle x^{**}, x^{**} \rangle - q^*(x^*) \geq \langle x^{**}, P^{**}x^{**} \rangle - q^*(P^*x^{**}) = \frac{1}{2}(x^{**}, P^*x^{**}).$$

On the other hand, by Corollary 2.18, there is a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \rightharpoonup x^{**}$ and $Px_\alpha \rightharpoonup P^*x^{**}$. Then for every $x^* \in X^*$, we estimate

$$q^*(x^*) \geq \lim_{\alpha} \langle x^*, x_\alpha \rangle - \frac{1}{2}(x_\alpha, Px_\alpha) = \langle x^{**}, x^* \rangle - \frac{1}{2}(x^{**}, P^*x^{**}).$$
This in turn implies $\frac{1}{2}\langle x^{**}, P^* x^{**} \rangle \geq \sup_{x^* \in X^*} \langle x^{**}, x^* \rangle - q^*(x^*) = q^{**}(x^{**})$, which yields the second equality.

"Also" part: By Theorem 2.17 and Fact 1.14, $\text{dom} \partial q^* = \text{ran} (\partial q)_1 = \text{ran} P^*$.
Finally, $\nabla q^{**} = P^{**}$ (apply Proposition 2.15 to $P^{**} = P^*$).

\[ \square \]

**Remark 2.20** Borrowing the notation of the two preceding propositions and recalling that $\text{ran} P \subseteq \text{ran} P^*$, we thus precisely know what $q^*$ does except at points in $(\text{cl ran} P) \setminus \text{ran} P^*$. Therefore, if $\text{ran} P$ is closed, then $q^*$ is completely determined. This happens if $X$ is finite-dimensional, we then recover a well-known formula for $q^*$ (see, for instance, [25, Example X.1.14]).

**Continuous Linear Skew Operators**
Parts of the next two propositions are implicit in Gossez' work [19, 21].

**Proposition 2.21** Suppose $S$ is a continuous linear skew operator from $X$ to $X^*$ and $(x^{**}, x^*) \in X^{**} \times X^*$. Then:

(i) $x^* \in S_1 x^{**} \Leftrightarrow x^* = S^{**} x^{**} = -S^* x^{**} \Leftrightarrow -x^* \in (-S)_1 x^{**}$.
(ii) $x^* \in S_0 x^{**} \Leftrightarrow x^* = -S^* x^{**}$ and $(S^* x^{**}, x^*) = 0 \Leftrightarrow -x^* \in (-S)_0 x^{**}$.
(iii) $x^* \in \overline{S} x^{**} \Leftrightarrow x^* = -S^* x^{**}$ and $(S^* x^{**}, x^*) \leq 0$.

Consequently: $\text{gra} S_1 = \text{gra} S^{**} \cap \text{gra} (-S^*) \subseteq \text{gra} \overline{S} \subseteq \text{gra} (-S^*)$, $(-S)_1 = -S_1$, and $(-S)_0 = -S_0$. Moreover:

(iv) $-S^*$ is monotone if and only if $\overline{S} = -S^*$ and $S$ is unique.
(v) $S^*$ is skew if and only if $S$ is weakly compact.

**Proof.** Fix an arbitrary $y \in X$. Then, using the skewness of $S$,

$$\langle x^{**} - y, x^* - S y \rangle = \langle x^{**}, x^* \rangle - \langle y, S^* x^{**} + x^* \rangle.$$

Hence

$$\inf_{(y, y^*) \in \text{gra} S} \langle x^{**} - y, x^* - y^* \rangle = \begin{cases} 
\langle x^{**}, x^* \rangle & \text{if } x^* = -S^* x^{**}; \\
-\infty & \text{otherwise}.
\end{cases}$$

(ii) and (iii) follow readily. For (i), observe that: $x^* \in S_1 x^{**} \Leftrightarrow (x^{**}, x^*) \in \text{gra} S_1 \cap \text{gra} S_0$ (Proposition 1.6) \Leftrightarrow $x^* = S^{**} x^{**} = -S^* x^{**} \in X^*$ ((ii) and Theorem 2.4) \Leftrightarrow $x^* = S^{**} x^{**} = -S^* x^{**}$ (ran $S^* \subseteq X^*$).

Now the “Consequently” part follows from Proposition 1.6, (i), and (ii).

“Moreover”: Recall that $S$ is unique if and only if $\overline{S}$ is monotone (Fact 1.9); hence (iv) follows. For (v), observe that $S^*$ skew $\Leftrightarrow S^{**} = -S^* \Leftrightarrow S_1 = S^{**}$ (since gra $S_1 = \text{gra} S^{**} \cap \text{gra} (-S^*)$) $\Leftrightarrow S$ is weakly compact (Corollary 2.5(i)).

\[ \square \]
Remark 2.22 Although \((-S)_{1} = -S_{1}\) and \((-S)_{0} = -S_{0}\) for every continuous linear skew operator \(S\), the formula \(\overline{\overline{-S}} = -\overline{S}\) is false in general: in fact, \(\overline{\overline{-S}} = -\overline{S}\) if \(S\) is of type (NI) — but there exist continuous linear skew operators that are not of type (NI): take, for instance, the operator \(G\) from Example 4.2.

Proposition 2.23 Suppose \(S\) is a continuous linear skew operator from \(X\) to \(X^{*}\). Suppose further \(S^{*}\) or \(-S^{*}\) is monotone. If \(x^{**}\) is a point in \(X^{**}\) with \(\langle x^{**}, S^{*}x^{**} \rangle = 0\), then \(S^{**}x^{**} = -S^{*}x^{**}\). Therefore, \(S_{1} = S_{0}\) and \((-S)_{1} = (-S)_{0}\).

Proof. Suppose \(x^{**} \in X^{**}\) with \(\langle S^{*}x^{**}, x^{**} \rangle = 0\).

Case 1: \(S^{*}\) is monotone.

Fix \(y^{**} \in X^{**}\) and \(\lambda > 0\). Then:

\[
0 = \langle S^{*}x^{**}, x^{**} \rangle = \langle S^{*}(x^{**} + \lambda y^{**}), x^{**} + \lambda y^{**} \rangle = \lambda \langle S^{*}x^{**}, y^{**} \rangle + \lambda \langle S^{*}y^{**}, x^{**} \rangle \geq -\lambda \langle S^{*}x^{**}, y^{**} \rangle - \lambda \langle S^{*}y^{**}, x^{**} \rangle - \lambda^{2} \langle S^{*}y^{**}, y^{**} \rangle.
\]

Now divide by \(\lambda\) and then let \(\lambda\) tend to 0 to conclude \(\langle S^{*}x^{**}, y^{**} \rangle \geq -\langle S^{*}y^{**}, x^{**} \rangle\). Replace \(y^{**}\) by \(-y^{**}\) and obtain \(\langle S^{*}x^{**}, y^{**} \rangle \leq -\langle S^{*}y^{**}, x^{**} \rangle\). Altogether:

\[
\langle -S^{*}x^{**}, y^{**} \rangle = \langle S^{*}y^{**}, x^{**} \rangle = \langle S^{**}x^{**}, y^{**} \rangle, \quad \forall y^{**} \in X^{**}.
\]

It follows that \(S^{**}x^{**} = -S^{*}x^{**}\).

Case 2: \(-S^{*}\) is monotone.

By the first case, \((-S)^{*}x^{**} = -(S)^{*}x^{**}\), as desired.

Using Proposition 2.21 repeatedly, we conclude as follows: Fix an arbitrary \((x^{**}, x^{*}) \in \text{gra} S_{0}\). Then \(x^{*} = -S^{*}x^{**}\) and \(\langle S^{*}x^{**}, x^{**} \rangle = 0\); hence, by what we just proved, \(S^{**}x^{**} = -S^{*}x^{**}\) and so \((x^{**}, x^{*}) \in \text{gra} S_{1}\). Thus \(S_{1} = S_{0}\) which yields \((-S)_{1} = -S_{1} = -S_{0} = (-S)_{0}\).

3 Characterizations

Proposition 3.1 Suppose \(T\) is a continuous linear operator from \(X\) to \(X^{*}\) and \((x^{**}, x^{*}) \in X^{**} \times X^{*}\). Let \(q(x) := \frac{1}{2} \langle x, Tx \rangle\), \(\forall x \in X\). Then

\[
\frac{1}{2} \inf_{x \in X} \langle Tx - x^{*}, x - x^{**} \rangle = \frac{1}{2} \langle x^{**}, x^{*} \rangle - q^{*}(\frac{1}{2} x^{*} + \frac{1}{2} T^{*} x^{**}).
\]

and hence:

(i) \(T\) is of type (NI) with respect to \((x^{**}, x^{*})\) \(\Leftrightarrow q^{*}(\frac{1}{2} x^{*} + \frac{1}{2} T^{*} x^{**}) \geq \frac{1}{2} \langle x^{**}, x^{*} \rangle\).

(ii) \((x^{**}, x^{*}) \in \text{gra} T_{0} \Leftrightarrow q^{*}(\frac{1}{2} x^{*} + \frac{1}{2} T^{*} x^{**}) = \frac{1}{2} \langle x^{**}, x^{*} \rangle\).
(iii) \((x^{**}, x^*) \in \text{gra} T \iff q^*(\frac{1}{2}x^{**} + \frac{1}{2}T^* x^{**}) \leq \frac{1}{2}\langle x^{**}, x^* \rangle \).

If \(x^* = (T^*|_X)^* x^{**}\), then:

(iv) \(T\) is of type (NI) with respect to \((x^{**}, x^*) \iff \langle S^* x^{**}, x^* \rangle \geq 0\).

(v) \((x^{**}, x^*) \in \text{gra} T_0 \iff \langle S^* x^{**}, x^* \rangle = 0\).

(vi) \((x^{**}, x^*) \in \text{gra} T \iff \langle S^* x^{**}, x^* \rangle \leq 0\).

Consequently:

(vii) \(S^*\) is monotone if and only if \(T\) is (NI) with respect to \(\text{gra} (T^*|_X)^*\).

(viii) \(S^*\) is skew if and only if \(T^{**} = (T^*|_X)^*\) if and only if \(\text{gra} (T^*|_X)^* \subseteq \text{gra} T_0\).

(ix) \(-S^*\) is monotone if and only if \(\text{gra} (T^*|_X)^* \subseteq \mathbb{T}\).

**Proof.** The displayed formula, (i), (ii), and (iii) are easy to check. Now suppose \((x^{**}, x^*) \in \text{gra} (T^*|_X)^*\), i.e. \(x^* = P^* x^{**} - S^* x^{**}\). Then, using Proposition 2.19,

\[
\frac{1}{2}\langle x^{**}, x^* \rangle - q^*(\frac{1}{2}x^{**} + \frac{1}{2}T^* x^{**}) = \frac{1}{2}\langle x^{**}, P^* x^{**} \rangle - \frac{1}{2}\langle x^{**}, S^* x^{**} \rangle - q^*(\frac{1}{2}P^* x^{**} - \frac{1}{2}S^* x^{**} + \frac{1}{2}S^* x^{**}) = -\frac{1}{2}\langle x^{**}, S^* x^{**} \rangle,
\]

which yields (iv), (v), and (vi). The “Consequently” part follows from (iv), (v), (vi), and the fact that \(\text{dom}(T^*|_X)^* = X^{**}\).

**Proposition 3.2** Suppose \(T\) is a continuous linear monotone operator from \(X\) to \(X^*\) with symmetric part \(P\) and skew part \(S\). If \(S^*\) or \(-S^*\) is monotone, then \(\text{gra} T^{**} \cap \text{gra} (T^*|_X)^* = \text{gra} T_1\). Consequently: if \(S^*\) is skew, then \(T_1 = T^{**} = (T^*|_X)^*\).

**Proof.** “\(\subseteq\)” is clear from Theorem 2.4.

“\(\supseteq\)” Fix \((x^{**}, x^*) \in \text{gra} T_1\), i.e. \(x^* = T^{**} x^{**} = P^* x^{**} + S^{**} x^{**} \in X^*\) (Theorem 2.4 and Theorem 2.17). Using Proposition 1.6, we thus have \((x^{**}, x^*) \in \text{gra} T_0 \cap \text{gra} T^{**}\). Now Proposition 3.1.(ii), the fact that \(\langle S^{**} x^{**}, x \rangle = -\langle S^* x^{**}, x \rangle\), \(\forall x \in X\), and Proposition 2.19 give

\[
\frac{1}{2}\langle x^{**}, P^* x^{**} + S^{**} x^{**} \rangle = q^*(\frac{1}{2}P^* x^{**} + \frac{1}{2}S^{**} + \frac{1}{2}P^* x^{**} + \frac{1}{2}S^{**}) = \sup_{x \in X}\langle P^* x^{**} + \frac{1}{2}S^{**}, x \rangle + \frac{1}{2}\langle x, P^* x^{**} \rangle = \sup_{x \in X}\langle x, P^* x^{**} \rangle - \frac{1}{2}\langle x, P^* x^{**} \rangle = q^*(P^* x^{**}) = \frac{1}{2}\langle x^{**}, P^* x^{**} \rangle.
\]

Hence \(\langle S^* x^{**}, x^* \rangle = 0\) and so (by Proposition 2.23) \(S^* x^{**} = -S^{**} x^{**}\). Thus \((x^{**}, x^*) = (x^{**}, P^* x^{**} + S^{**} x^{**}) = (x^{**}, P^* x^{**} - S^* x^{**}) \in \text{gra} (T^*|_X)^*\). The “Consequently” part follows, since \(S^*\) is skew if and only if \((T^*|_X)^* = T^{**}\).

\(\Box\)

We are now ready for the main result.
Theorem 3.3 Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with symmetric part $P$ and skew part $S$. Then TFAE:

(i) $T$ is monotone and of dense type.
(ii) $T$ is monotone and of range-dense type.
(iii) $T$ is monotone and of type (NI).
(iv) $T$ is monotone and of type (NI) with respect to gra ($-T^*$).
(v) $T$ is locally maximal monotone.
(vi) $T^*$ is monotone.
(vii) $P$ and $S^*$ are monotone.
(viii) $P$ is monotone and $S$ is of dense type.
(ix) $P$ is monotone and $S$ is of range-dense type.
(x) $P$ is monotone and $S$ is of type (NI).
(xi) $P$ is monotone and $S$ is of type (NI) with respect to $-S^*$.
(xii) $P$ is monotone and $S$ is locally maximal monotone.
(xiii) gra $T_1 = gra T_0 = gra T = gra T^* \cap gra (T^*|_X)^*$.

Proof. Throughout this proof, let $q(x) := \frac{1}{2} \langle Tx, x \rangle = \frac{1}{2} \langle Px, x \rangle, \forall x \in X$.

“(i)⇒(ii)⇒(iii)⇒(iv)⇒(vi)” follow from Fact 1.13 and Proposition 2.11.
“(vi)⇒(vii)”: $T$ and $P$ are monotone, because $T^*$ is. Fix an arbitrary $x^* \in X^*$.

By Corollary 2.18, obtain a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \stackrel{w^*}{\to} x^{**}$ and $Px_\alpha \to P^*x^{**}$. Now

$$0 \leq \langle T^*(x^{**} - x_\alpha), x^{**} - x_\alpha \rangle$$
$$= \langle T^*x^{**}, x^{**} - x_\alpha \rangle - \langle T^*x_\alpha, x^{**} - x_\alpha \rangle$$
$$= \langle T^*x^{**}, x^{**} - x_\alpha \rangle - \langle Px_\alpha - Sx_\alpha, x^{**} - x_\alpha \rangle$$
$$= \langle T^*x^{**} - Px_\alpha, x^{**} - x_\alpha \rangle + \langle x_\alpha, S^*x^{**} \rangle$$
$$\to \langle x^{**}, S^*x^{**} \rangle;$$

consequently, $S^*$ is monotone and (vii) holds.

“(vii)⇒(i)”: Fix an arbitrary $(x^{**}, x^*) \in$ gra $T$. By Corollary 2.18, obtain a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \stackrel{w^*}{\to} x^{**}$ and $Px_\alpha \to P^*x^{**}$. Now Proposition 3.1 and the monotonicity of $S^*$ yield

$$\frac{1}{2} \langle x^{**}, x^* \rangle \quad \geq \quad q^*(\frac{1}{2}x^{**} + \frac{1}{2}T^*x^{**})$$
$$\geq \quad \lim_{\alpha} \langle \frac{1}{2}x^{**} + \frac{1}{2}T^*x^{**}, x_\alpha \rangle - \frac{1}{2} \langle x_\alpha, Px_\alpha \rangle$$
$$= \quad \frac{1}{2} \langle x^{**}, x^{**} \rangle + \frac{1}{2} \langle S^*x^{**}, x^{**} \rangle$$
$$\geq \quad \frac{1}{2} \langle x^{**}, x^* \rangle,$$

hence $\langle S^*x^{**}, x^{**} \rangle = 0$ and $q^*(\frac{1}{2}x^{**} + \frac{1}{2}T^*x^{**}) = \frac{1}{2} \langle x^{**}, x^* \rangle$. This has two important consequences: firstly, by Proposition 2.23,

$$S^*x^{**} = -S^*x^{**} \in$ ran $S^* \subseteq X^*$. 

21
Secondly, using Proposition 2.19,

\[
\left(\frac{1}{2}x^* + \frac{1}{2}T^*x^*, x^*\right) = \frac{1}{2}\langle x^*, x^*\rangle + \frac{1}{2}\langle x^*, P^*x^*\rangle
\]

\[
= q^\left(\frac{1}{2}x^* + \frac{1}{2}T^*x^*\right) + q^\left(x^*\right); \\
\text{thus: } x^* \in \partial q^\left(\frac{1}{2}x^* + \frac{1}{2}T^*x^*\right) \Rightarrow \frac{1}{2}x^* + \frac{1}{2}T^*x^* \in \partial q^\left(x^*\right) = \{P^*x^*\} \Rightarrow \\
x^* = P^*x^* - S^*x^*. \text{ Altogether,}
\]

\[
x^* = P^*x^* - S^*x^* = P^*x^* + S^*x^* = T^*x^* \in X^*,
\]

so that (Theorem 2.4) \((x^**, x^*) \in \text{gra} T_1\), as desired.

“(v)⇒(vi)”: \(T\) is monotone (use \(V = X^*\) in Definition 1.7.(iv)) and so is \(P\); the function \(q\) is convex (Proposition 2.15). Fix \(x_0^* \in X^*\). We aim for \(\langle T^*x_0^*, x_0^*\rangle \geq 0\) and can thus assume WLOG that \(x_0^* := T^*x_0^* \neq 0\). Select \(x_1 \in X\) with \(\langle x_0^*, x_1\rangle < 0\) and let \(x_1^* := Tx_1\). Let \(x_0 := 0\), fix an arbitrary \(\varepsilon > 0\), and define

\[
C_\varepsilon := [x_0^*, x_1^*] + \varepsilon B_{X^*}.
\]

Then \(C_\varepsilon\) is \(w^*\) closed, convex, bounded with \(x_1^* \in \text{ran} T \cap \text{int} C_\varepsilon\). Also, \(x_0^* \in (\text{int} C_\varepsilon) \setminus Tx_0\). Local maximal monotonicity (via Proposition 1.11.(iv)), Proposition 1.23, and Fact 1.22 yield

\[
0 > \frac{1}{2} \inf_{x \in X : T \in C_\varepsilon} \langle Tx - x_0^*, x - x_0\rangle
\]

\[
= \inf_{x \in X} q(x) + \langle -\frac{1}{2}x_0^*, x\rangle + \iota_{C_\varepsilon}(Tx)
\]

\[
\geq \inf_{x^* \in X^*} \left\{ q^\left(-T^*x^* + \frac{1}{2}x_0^*\right) + \iota_{C_\varepsilon}(x^*) \right\}.
\]

Now pick \(x^* := \frac{1}{2}x_0^*\); then, by Example 1.25 and the fact that \(q^*(0) = 0\),

\[
0 < q^\left(-\frac{1}{2}T^*x_0^* + \frac{1}{2}x_0^*\right) + \iota_{C_\varepsilon}\left(\frac{1}{2}x_0^*\right)
\]

\[
= \iota_{C_\varepsilon}\left(\frac{1}{2}x_0^*\right)
\]

\[
= \varepsilon\|\frac{1}{2}x_0^*\| + \max\{\langle \frac{1}{2}x_0^*, x_0^\rangle, \langle \frac{1}{2}x_0^*, x_1^*\rangle\}.
\]

Multiply by 2 and then let \(\varepsilon\) tend to 0 to obtain

\[
0 \leq \max\{\langle x_0^*, T^*x_0^*\rangle, \langle x_0^*, Tx_1\rangle\}.
\]

Since \(\langle x_0^*, Tx_1\rangle = \langle T^*x_0^*, x_1\rangle = \langle x_0^*, x_1\rangle < 0\), we conclude \(\langle x_0^*, T^*x_0^*\rangle \geq 0\).

“(vii)⇒(v)”: In view of Proposition 1.11.(iv), let’s fix a \(w^*\) closed convex bounded subset \(C\) of \(X^*\) with \(\text{ran} T \cap \text{int} C \neq \emptyset\). Let \(x_0 \in X\), \(x_0^* \in (\text{int} C) \setminus Tx_0\). Let

\[
p := \inf_{x \in X : T \in C} \frac{1}{2} \langle Tx - x_0^*, x - x_0\rangle.
\]

Clearly, \(p < +\infty\) and our aim is \(p < 0\). We thus can assume WLOG that \(p > -\infty\), hence \(p\) is finite. Let \(f(x) := q(x) + \frac{1}{2}\langle -x_0^* - T^*x_0^*, x\rangle + \frac{1}{2}\langle x_0^*, x_0\rangle\),
\(\forall x \in X\) and let \(g := \iota C\). Then, using Fact 1.22 and Proposition 1.23,

\[
\begin{align*}
p &= \inf_{x \in X} f(x) + g(Tx) \\
&= d \\
&:= -\inf_{x^* \in X^*} \{ f^*(-T^*x^*) + g^*(x^*) \} \\
&= \frac{1}{2}(x_0^*, x_0) - \inf_{x^* \in X^*} \{ q^*(-T^*x^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) + \iota_C^*(x^*) \}.
\end{align*}
\]

Moreover: the last infimum is attained (by Fact 1.22 and ran \(T \cap \text{int} C \neq \emptyset\), say at some \(x_0^* \in X^*\). Thus the proof of "(vii)\(\Rightarrow\)(v)" would be complete after reaching the following

Aim: \(\frac{1}{2}(x_0^*, x_0) < q^*(-T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) + \iota_C^*(x_0^*)\).

By assumption, \(0 \leq \langle S^*x_0^*, x_0^* \rangle + \langle S^*(x_0^* - x_0), x_0^* - x_0 \rangle\), which is (after some tedious yet elementary manipulations) equivalent to

\[
\begin{align*}
\langle -T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0, x_0 - 2x_0^* \rangle - \frac{1}{2}(x_0 - 2x_0^*), P^*(x_0 - 2x_0^*) \rangle \\
\geq \frac{1}{2}(x_0^*, x_0) - \langle x_0^*, x_0^* \rangle.
\end{align*}
\]

On the other hand, Corollary 2.18 gives a bounded net \((x_0)\) in \(X\) such that
\(x_0 \rightharpoonup x_0 - 2x_0^*\) and \(P x_0 \to P^*(x_0 - 2x_0^*)\); thus altogether

\[
\begin{align*}
q^*(-T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) \\
\geq \lim_{\to} \langle -T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0, x_0 \rangle - \frac{1}{2}(x_0, P x_0) \\
= \langle -T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0, x_0 - 2x_0^* \rangle - \frac{1}{2}(x_0 - 2x_0^*, P^*(x_0 - 2x_0^*)) \\
\geq \frac{1}{2}(x_0^*, x_0) - \langle x_0^*, x_0^* \rangle.
\end{align*}
\]

Consequently, since \(x_0^*\) is in the interior of \(C\),

\[
\begin{align*}
\frac{1}{2}(x_0^*, x_0) \leq q^*(-T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) + \langle x_0^*, x_0^* \rangle \\
< q^*(-T^*x_0^* + \frac{1}{2}x_0^* + \frac{1}{2}T^*x_0) + \iota_C^*(x_0^*),
\end{align*}
\]

which is what we aimed for!

We just proved that (i)-(viii) are equivalent for arbitrary continuous linear operators. So let’s apply this to \(\bar{T} := S\). The symmetric (resp. skew) part of \(\bar{T}\) is 0 (resp. \(S\)). Hence we obtain the equivalences:

\(S\) is of dense type \(\Leftrightarrow\) \(S\) is of range dense type \(\Leftrightarrow\) \((\text{NI})\Leftrightarrow\) \(S\) is of type (NI) with respect to \(\text{gra}(-S)\Leftrightarrow\) \(S\) is locally maximal monotone \(\Leftrightarrow\) \(S^*\) is monotone.

Therefore, (i)-(xi), (vii)\(\Rightarrow\)(xi): use Proposition 1.6 and Proposition 3.2.

“(xi)\(\Rightarrow\)(i)”: \(\bar{T}\) is monotone (since \(\text{gra} \bar{T} \subseteq \text{gra} \bar{T}_1 = \text{gra} \bar{T}\), see Remarks 1.5(i)) and of dense type.
Observation 3.4 Gossez [19, End of Section 2] found the following question interesting:

Suppose $T$ is a closed densely defined linear monotone operator from $X$ to $X^*$. Suppose further that $T^*$ is monotone. Is $T_1$ maximal monotone?

He then proved that the answer is “yes” if $T$ is continuous and skew. We are now able to give an affirmative answer to this question provided that $T$ is merely continuous: indeed, this follows from Theorem 3.3.\((vi)\Rightarrow(i)\) and Proposition 1.11.(i).

Theorem 3.5 Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with symmetric part $P$ and skew part $S$. Then TFAE:

(i) $T$ and $T^*|_X$ are monotone and of dense type.
(ii) $T$ and $T^*|_X$ are monotone and of type (NI).
(iii) $T$ and $T^*|_X$ are locally maximal monotone.
(iv) $T^*$ and $(T^*|_X)^*$ are monotone.
(v) $T$ is monotone and weakly compact.
(vi) $P$ is monotone and $S$ is weakly compact.
(vii) $P$ is monotone and $S^*$ is skew.
(viii) $P$ is monotone and $S$, $-S$ are of dense type.
(ix) $P$ is monotone and $S$, $-S$ are of type (NI).
(x) $P$ is monotone and $S$, $-S$ are locally maximal monotone.
(xi) $T_1 = T_0 = T = T^* = (T^*|_X)^*$.

Proof. Applying Theorem 3.3 to $T = P + S$ and $T^*|_X = T - S$ yields the equivalence of (i), (ii), (iii), (iv), (vii), (viii), (ix), and (x). Now (v) $\Leftrightarrow$ (vi) (by weak compactness of $P$; see Theorem 2.17) $\Leftrightarrow$ (vii) (by Proposition 2.21.(v)); so (i)-(x) are all equivalent. Finally, (vii) $\Rightarrow P$, $S^*$ are monotone and $S^*$ is skew $\Leftrightarrow$ gra$T_1 = gra T_0 = gra T = gra (T^*|_X)^*$ and $T^* = (T^*|_X)^*$ (Theorem 3.3 and Proposition 3.1.(viii)) $\Leftrightarrow$ (xi). \hfill $\Box$

Remark 3.6 Borrowing notation from the two theorems above, we see that monotonicity of $S^*$ can be interpreted as “one half of weak compactness” of $T$.

Proposition 3.7 Let $\mathcal{X}$ be the Banach space of all continuous linear operators from $X$ to $X^*$ (equipped with the usual operator norm). Let further $\mathcal{M} := \{T \in \mathcal{X} : T$ is monotone\}, $\mathcal{M}_* := \{T \in \mathcal{X} : T^*$ is monotone\}, and $\mathcal{M}_0 := \{T \in \mathcal{X} : T$ is monotone and weakly compact\}. Then $\mathcal{M}$, $\mathcal{M}_*$, and $\mathcal{M}_0$ are closed convex cones and $\mathcal{M} \subseteq \mathcal{M}_* \subseteq \mathcal{M}_0$. Moreover, $\text{lin} \mathcal{M} = \{T \in \mathcal{X} : T$ is skew\} and $\text{lin} \mathcal{M}_* = \text{lin} \mathcal{M}_0 = \{T \in \mathcal{X} : T^*$ is skew\}.  

24
Proof. Clearly, $\mathcal{M}$, $\mathcal{M}_+$, and $\mathcal{M}_0$ are convex cones; the announced inclusions follow from Theorem 3.3 and Theorem 3.5. To verify closedness of the cones, suppose $T_n \to T$ in $\mathcal{X}$. Recall that $T_n^* \to T^*$ in $\mathcal{X}^*$ and $T_n^{**} \to T^{**}$ in $\mathcal{X}^{**}$ (because $\|T_n - T\| = \|T_n^* - T^*\| = \|T_n^{**} - T^{**}\|$).

(i): Suppose the net $(T_n)$ is in $\mathcal{M}$. Then $(T_n x, x) \geq 0$, $\forall x$, $\forall x \in X$; hence $(Tx, x) \geq 0$. It follows that $\mathcal{M}$ is closed.

(ii): Suppose the net $(T_n)$ is in $\mathcal{M}_+$. Then $(T_n x^+, x^+) \geq 0$, $\forall x^+, \forall x^+ \in X^+$; hence $(T^* x^+, x^+) \geq 0$. It follows that $\mathcal{M}_+$ is closed.

(iii): Suppose the net $(T_n)$ is in $\mathcal{M}_0$. Then $T_n^{**} x^+ \in X^*$, $\forall x^+$, $\forall x^+ \in X^*$. Since $X^*$ is closed in $X^{**}$, we obtain $T^{**} x^+ \in X^*$ and we saw in (i) already that $(Tx, x) \geq 0$, $\forall x \in X$. Hence $\mathcal{M}_0$ is closed.

Now $T \in \text{lin} \mathcal{M} \Rightarrow T$ and $-T$ are monotone $\Rightarrow T$ is skew. Finally: $T \in \text{lin} \mathcal{M}_+$ $\Rightarrow T^*$ and $-T^*$ are monotone $\Rightarrow P$, $-P$, $S^*$, $-S^*$ are monotone (Theorem 3.3) $\Rightarrow P$, $-P$ are monotone and $S^*$, $-S^*$ are skew $\Rightarrow T$, $-T$ are weakly compact (Theorem 3.5). □

Remark 3.8 Borrowing notation from Proposition 3.7, we see with Theorem 3.3 that the closed convex cone $\mathcal{M}_+$ is equal to the set of of all continuous linear monotone operators that are of dense type (equivalently: of range-dense type, of type (NI), or locally maximal monotone). In particular, the latter set is closed under addition, nonnegative scalar-multiplication, and taking limits in $\mathcal{X}$.

4 Examples

Theorem 4.1 Suppose $T$ is a continuous linear operator from $X$ to $X^*$ and there exists some $e \in X^*$ such that

$$e \notin \text{cl} \text{ran} \ T \land \langle Tx, x \rangle = \langle e, x \rangle^2, \ \forall x \in X.$$ 

Then $T$ is monotone, its symmetric part $P$ is given by $Px := \langle e, x \rangle e$, $\forall x \in X$, its skew part $S := T - P$, and $P^* x^* := \langle x^*, e \rangle e$, $\forall x^* \in X^*$. Moreover: $S^*$ is not monotone; $S$ is neither of type (NI) nor locally maximal monotone; $S$ and $-S$ are not weakly compact.

If $\text{ran} T^* = \text{ran} T^*|_X$; equivalently, $(\text{ran} T)^* \subseteq X$ or

\begin{align*}
\forall x^* \in X^* \exists \tilde{x} \in X \subseteq X^* : x^*|_{\text{ran} T} = \tilde{x}|_{\text{ran} T},
\end{align*}

then $-S^*$ is monotone (in fact: $\langle -S^* x^*, x^* \rangle = \langle x^* - \tilde{x}, e \rangle^2, \forall x^* \in X^*$); $S$ is unique; $-S$ is of dense type and locally maximal monotone;

$\text{gra} S_1 = \text{gra} S^* \cap \text{gra} (-S^*) = \text{gra} S_0 \subsetneq \text{gra} S = \text{gra} (-S^*)$;

$\text{gra} (-S)_1 = \text{gra} (-S^*) \cap \text{gra} S^* = \text{gra} (-S)_0 = \text{gra} (-S) \subsetneq \text{gra} (-S)$

If $\ker T \subsetneq \ker T^*$; or, even more restrictive,

\begin{align*}
T \text{ is one-to-one and ran } T^* \text{ is not norm dense in } X^*,
\end{align*}

then $S$ and $-S$ are not tauberian; $\text{gra} S \subsetneq \text{gra} S_1$; $\text{gra} (-S) \subsetneq \text{gra} (-S)_1$.

25
Proof. $T$ is obviously monotone. Let $P x := \langle e, x \rangle e, \forall x \in X$; then $P^* x^* = \langle x^*, P x \rangle = \langle x^*, e \rangle \langle e, x \rangle$ and hence $P^* x^* = \langle x^*, e \rangle e, \forall x^* \in X^*$. So $P$ is symmetric. Consider now $S := T - \bar{P}$. Then $\langle S x, x \rangle = \langle T x, x \rangle \langle P x, x \rangle = \langle T x, e \rangle \langle e, x \rangle = 0, \forall x \in X$, thus $S$ is skew. Since $T = P + S$, the symmetric (resp. skew) part of $T$ is $P$ (resp. $S$) by Proposition 2.14.

Because $e \notin \text{cl ran } T = \{ (\ker T) \} [51, \text{Lemma } 11-1-7.(c)]$, there exists some $x_0^* \in \ker T^*$ with $\langle x_0^*, e \rangle \neq 0$. Hence

$$\langle S^* x_0^*, x_0^* \rangle = \langle T^* x_0^*, x_0^* \rangle - \langle P^* x_0^*, x_0^* \rangle = 0 - \langle x_0^*, e \rangle^2 < 0;$$

so $S^*$ is not monotone. Thus $S$ is neither of type (NI) nor locally maximal monotone (Theorem 3.3) and $S, -S$ are not weakly compact (Proposition 2.21.(v)). Since $\text{ran } T \subseteq X^*$, the Hahn/Banach Theorem allows us to identify $\text{ran } T^*$ with $\{ x^* | \text{ran } T : x^* \in X^* \}$. Hence, as announced:

$$(\text{ran } T)^* \subseteq X \iff \forall \hat{x}^* \in (\text{ran } T)^* \exists \hat{x} \in X \subseteq X^* \hat{z}^* = \hat{x}|_{\text{ran } T}$$

$\iff \forall \hat{x}^* \in X^* \exists \hat{x} \in X \hat{z}^*|_{\text{ran } T} = \hat{x}|_{\text{ran } T}$$

$\iff \forall \hat{x} \in X \hat{t} \in T^* \hat{x}^* = T^* \hat{x}$

$\iff \text{ran } T^* = \text{ran } T^*|_X.$

Suppose now (1) holds and fix an arbitrary $x^* \in X^*$. Pick $\hat{x} \in X \subseteq X^*$ with $x^*|_{\text{ran } T} = \hat{x}|_{\text{ran } T}$; equivalently, $T^* x^* = T^* \hat{x}$. Then we have $\langle S^* x^*, x \rangle = \langle T^* x^*, x \rangle - \langle P^* x^*, x \rangle = \langle T^* \hat{x}, x \rangle - \langle P^* x^*, x \rangle, \forall x \in X$; hence

$$\langle S^* x^*, x \rangle = \langle T^* \hat{x}, x \rangle - \langle P^* x^*, x \rangle$$

Because $T^*|_X = P - S = 2P - T$, we further obtain

$$\langle S^* x^*, x \rangle = \langle T^* \hat{x}, x \rangle - \langle x^*, e \rangle^2$$

$$= 2 \langle P \hat{x}, x^* \rangle - \langle T \hat{x}, x^* \rangle - \langle x^*, e \rangle^2$$

$$= 2 \langle P \hat{x}, x^* \rangle - \langle \hat{x}, T^* x^* \rangle - \langle x^*, e \rangle^2$$

$$= 2 \langle P \hat{x}, x^* \rangle - \langle \hat{x}, T^* \hat{x} \rangle - \langle x^*, e \rangle^2$$

$$= 2 \langle e, \hat{x} \rangle \langle x^*, e \rangle - \langle e, \hat{x} \rangle^2 - \langle x^*, e \rangle^2$$

$$= -\langle e, x^* - \hat{x} \rangle^2 \leq 0;$$

consequently, $-S^*$ is monotone. Then Proposition 2.21 and Proposition 2.23 imply: $S$ is unique, $S = -S^*$, $S_1 = S_0 = -(-S)_1 = (-S)_0, \text{gra } S_1 = \text{gra } S^* \cap \text{gra } S$ (Theorem 3.3 on $-S, S$); so all announced consequences of (1) are verified. Finally, (2) $\iff \ker T = \{0\}$ and $\text{cl ran } T^* \subseteq X^* \iff \ker T = \{0\}$ and $(\text{cl ran } T^*)^\perp = \ker T^* \supseteq \{0\} [51, \text{Lemma } 11-1-7.(b)] \iff \ker T^* \supseteq \ker T^* \Rightarrow T$ and $-T$ are not tauberian ([51, Theorem 11-4.5]) $\iff S$ and $-S$ are not tauberian (since $P$ is weakly compact) $\iff \text{gra } S \supseteq \text{gra } S_1$ and $\text{gra } (-S) \supseteq \text{gra } (-S)_1$ (by Theorem 2.4 and Corollary 2.5). 

$\square$
Example 4.2 (Gossez’ [19, Example in Section 2]) Define the mapping $G$ from $\ell_1$ to $\ell_\infty$ by
\[
(Gx)_n := -\sum_{k<n} x_k + \sum_{k>n} x_k, \quad \forall x = (x_k) \in \ell_1 \forall n \in \mathbb{N}.
\]
Then: $G$ and $-G$ are continuous linear skew operators from $\ell_1$ to $\ell_1^* = \ell_\infty$. $G^*$ is not monotone but $-G^*$ is. $G$ is neither of type (NI) nor locally maximal monotone. $-G$ is of dense type and locally maximal monotone. Both $G$ and $-G$ are unique, but neither is weakly compact nor tauterian.

Proof. Consider the mapping $T$ from $\ell_1$ to $\ell_\infty$ given by
\[
(Tx)_n := x_n + 2 \sum_{k>n} x_k, \quad \forall x = (x_k) \in \ell_1 \forall n \in \mathbb{N}.
\]
Then $T$ is linear, continuous (in fact $\|T\| = 2$), and $\text{ran} \, T \subseteq c_0 \subseteq \ell_\infty$. Let $e := (1,1,\ldots) \in \ell_\infty = \ell_1^*$. Then $e \notin \text{cl} \, \text{ran} \, T \subseteq \text{cl} \, c_0 = c_0$ and for every $x \in \ell_1$,
\[
\langle T x, x \rangle = \sum_n x_n (x_n + \sum_{k>n} 2x_k) = \sum_n x_n^2 + \sum_n \sum_{k>n} 2x_n x_k = \sum_n x_n^2 + \sum_{k\neq n} x_n x_k = \langle \sum_n x_n \cdot 1, (\sum_k x_k \cdot 1) \rangle = \langle e, x \rangle^2.
\]
By Theorem 4.1, the symmetric part $P$ of $T$ is given by $Px = \langle e, x \rangle e$, $\forall x \in \ell_1$ and the skew part of $T$ is $S := T - P$. Now for all $x \in \ell_1, n \in \mathbb{N}$:
\[
(Sx)_n = (Tx)_n - (Px)_n = x_n + 2 \sum_{k>n} x_k - \sum_k x_k = -\sum_{k<n} x_k + \sum_{k>n} x_k = (Gx)_n;
\]
hence $S = G$. It follows that $G^*$ is not monotone; $G$ is neither of type (NI) nor locally maximal monotone; $G$ and $-G$ are not weakly compact (Theorem 4.1). The Hahn/Banach Theorem yields $(\text{ran} \, T^*)_* \subseteq c_0^* = \ell_1$, so (1) in Theorem 4.1 holds and hence: $-G^*$ is monotone; $-G$ is of dense type, locally maximal monotone, and unique; $G$ is at least unique. It is easy to check that $T$ is one-to-one. Note also that by (1), $\text{cl} \, \text{ran} \, T^* = \text{cl} \, \text{ran} \, T^*|_X = \text{cl} \, \text{ran} \, (2P - T) \subseteq \text{cl} \, (\text{ran} \, T + \mathbb{R}e) \subseteq \text{cl} \, (c_0 + \mathbb{R}e) \subseteq c_{\ell_{\infty}} \subseteq \ell_\infty^*$. Hence (2) in Theorem 4.1 holds and $G$, $-G$ are not tauterian.

Remark 4.3 Let $G$ denote the Gossez operator from Example 4.2. Gossez [19] proved that $G$ is unique but not of dense type. Phelps [35, Example 4.5] showed
that $G$ is not locally maximal monotone. We observe that the discussion of the Gossez operator via Theorem 4.1 is conceptionally much simpler, does not require the Stone–Cech compactification, and gives a bit more insight.

The Gossez operator arises quite naturally from an operator-theoretic viewpoint:

**Remark 4.4** The universal nonweakly compact operator (see [9, Exercise VII.6]) is the sum operator $\sigma$ from $l_1$ to $l_\infty$:

$$(\sigma x)_n := \sum_{k \leq n} x_k, \quad \forall x = (x_k) \in l_1 \forall n \in \mathbb{N},$$

Then the symmetric part $P$ of $\sigma$ is given by $(Px)_n := x_n + \frac{1}{n} \sum_{k \not= n} x_k, \forall x \in l_1$; the skew part of $\sigma$ is $-\frac{1}{n} G$, where $G$ denotes Gossez’ operator from Example 4.2; and $\sigma$ is monotone: denoting $(1, 1, 1, \ldots) \in l_\infty$ by $e$, we have $(\sigma x, e) = \frac{1}{2} \|e\|^2 + \frac{1}{n} \|x\|^2_n$. Similarly, the tail operator $\tau$ from $l_1$ to $l_\infty$ is given by $(\tau x)_n := \sum_{k \geq n} x_k, \forall x \in l_1 \forall n \in \mathbb{N}$. Then $\sigma$ and $\tau$ possess the same symmetric part (so $\tau$ is also monotone), but the skew part of $\tau$ equals $\frac{1}{2} G$.

The next example is a “continuous” version of the (negative) Gossez operator; see Example 4.2.

**Example 4.5** (Fitzpatrick and Phelps’ [15, Example 3.2]) Define the mapping $F$ from $L_1[0, 1]$ to $L_\infty[0, 1]$ by

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

Then $F$, $-F$ are continuous linear skew operators from $L_1[0, 1]$ to $L_\infty^*[0, 1] = L_\infty[0, 1]$ that are not of type (NI), not locally maximal monotone, and not weakly compact. Neither $F^*$ nor $-F^*$ is monotone.

**Proof. Step 1:** Define the mapping $T$ from $L_1[0, 1]$ to $L_\infty[0, 1]$ by

$$(Tx)(t) := 2 \int_0^t x(s) ds, \quad \forall x \in L_1[0, 1] \forall t \in [0, 1].$$

Then $T$ is linear and continuous (with $\|T\| = 2$). The range of $T$ is contained in the subspace $C_{0,0}$ of $L_\infty[0, 1]$ that consists of all equivalence classes that contain a continuous function vanishing at $0$. Let $e$ denote the equivalence class in $L_\infty[0, 1]$ that contains the constant function $1$. Then the distance from $e$ to any member in $C_{0,0}$ is at least $1$; thus certainly $e \notin \text{cl ran } T$. Also, for every $x \in L_1[0, 1]$,

$$\langle Tx, x \rangle = 2 \int_0^1 (\int_0^t x(s) ds) x(t) \, dt$$
$$= 2 \int_0^1 x(s) x(t) \, ds \, dt$$
$$= \int_{[0,1] \times [0,1]} x(s) x(t) \, ds \, dt$$
$$= (\int_0^1 x(s) \cdot 1 \, ds)(\int_0^1 x(t) \cdot 1 \, dt)$$
$$= \langle x, e \rangle^2.$$
Then (Theorem 4.1) the positive part \( P \) of \( T \) is given by \( Px := \langle e, x \rangle e \), and the skew part \( S \) of \( T \) is given by

\[
(Sx)(t) := (Tx)(t) - (Px)(t)
\]

\[
= 2 \int_0^t x(s) \, ds - \langle e, x \rangle e(t)
\]

\[
= 2 \int_0^t x(s) \, ds - \int_0^1 x(s) \, ds
\]

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

consequently, \( S = F \). Now Theorem 4.1 implies that \( F^* \) is not monotone and \( F \) is neither of type (NI) nor locally maximal monotone.

**Step 2:** This time, we define the mapping \( T \) by

\[
(Tx)(t) := 2 \int_0^1 x(s) \, ds, \quad \forall x \in L_1[0,1] \ \forall t \in [0,1].
\]

We define \( e \) as in Step 1 and check analogously: \( T \) is continuous, linear, \( e \notin \text{ran } T \), and \( \langle Tx, x \rangle = \langle x, x \rangle^2, \forall x \in L_1[0,1] \). This time, however, the skew part \( S \) of \( T \) is \( -F \) ! We conclude as above that \( -F^* \) is not monotone and that \( -F \) is neither of type (NI) nor locally maximal monotone. \( \square \)

**Remarks 4.6** Consider the operator \( F \) from Example 4.5 and let \( X := L_1[0,1] \). Fitzpatrick and Phelps showed directly that the operator \( F \) is not locally maximal monotone; see [15, Example 3.2]. We now sketch a proof that \( F \) is not tauberian. Define the operator \( T \) from \( X \) to \( X^* \) by \( (Tx)(t) := \int_0^t x(s) \, ds \), \( \forall x \in X \ \forall t \in [0,1] \). Since the skew part of \( F \) is \( 2T \) (see the proof of Example 4.5), the operator \( F \) is tauberian if and only if \( T \) is. If \( T \) were tauberian, then \( T(B_X) \) would be closed (see [51, Problem 11-4-126] or [32, Theorem 2.1]).

**Keystep** (Erdélyi): \( T(B_X) \) is not closed.

To see this, denote Lebesgue’s singular function by \( \Lambda \) (see [23, Exercise 8.28] for the construction and a sketch). Let \( \Phi_1 \) (resp. \( \Phi_2 \)) be the piecewise linear continuous function on \([0,1]\) determined by the points \( \{(0,0), (\frac{1}{3}, \frac{1}{3}), (\frac{2}{3}, \frac{1}{3}), (1,1)\} \) (resp. by \( \{(0,0), (\frac{1}{5}, \frac{1}{5}), (\frac{2}{5}, \frac{2}{5}), (\frac{3}{5}, \frac{3}{5}), (\frac{4}{5}, \frac{4}{5}), (\frac{5}{5}, \frac{5}{5}), (1,1)\} \)). Continue in the spirit of the construction of \( \Lambda \) via Cantor’s ternary set – to obtain a sequence \( (\Phi_n) \) of continuous piecewise linear functions and denote their derivatives by \( (\varphi_n) \). Then \( (\varphi_n) \) lies in \( B_X \) (in fact: \( \|\varphi_n\|_1 = 1, \forall n \in \mathbb{N} \)) and the sequence \( (T\varphi_n) = (\Phi_n) \) converges uniformly to \( \Lambda \) and hence in \( L_\infty[0,1] \). However, \( \Lambda \) is not absolutely continuous ([47, Example 3.138]); therefore, \( \Lambda \notin \text{ran } T \) ([47, Theorem 6.8.4]).

Altogether, the operator \( F \) is not tauberian.

We conclude this section with an important example, also due to Gossez.

**Fact 4.7** (Gossez [21]) There exists a continuous linear skew operator from \( \ell_1 \) to \( \ell_\infty \) that is not unique.

**Remark 4.8** Denoting Gossez’ operator from Fact 4.7 by \( S \), we see that \( S \) is not monotone (by Fact 1.9).
5 Conjugate monotone spaces

Suppose $S$ is a continuous linear skew operator from $X$ to $X^*$. Then exactly one of the following alternatives occurs:

(i) $S^*$ and $-S^*$ are monotone.
(ii) Either $S^*$ or $-S^*$ is monotone but not both.
(iii) Neither $S^*$ nor $-S^*$ is monotone.

For instance: (i) always happens in reflexive Banach spaces (Theorem 3.3); (ii) occurs in $\ell_1$ (Example 4.2); (iii) takes place in $L_1[0,1]$ (Example 4.5).

Let us now single out spaces that do not allow “bad” continuous linear monotone operators:

**Definition 5.1** We say that $X$ is a conjugate monotone space (cms), if the conjugate of every continuous linear monotone operator from $X$ to $X^*$ is monotone.

**Proposition 5.2** TFAE:

(i) $X$ is (cms).
(ii) Every continuous linear skew operator from $X$ to $X^*$ has a skew conjugate.
(iii) Every continuous linear skew operator from $X$ to $X^*$ is weakly compact.
(iv) Every continuous linear monotone operator from $X$ to $X^*$ is weakly compact.

**Proof.** “(iv)⇒(iii)”: is trivial. “(ii)⇒(i)”: use Proposition 2.21.(v). “(ii)⇒(i)”: follows from Theorem 3.3. “(i)⇒(iv)”: Fix a continuous linear monotone operator $T$ from $X$ to $X^*$. Then $T^*|_X$ is continuous linear monotone as well. Since $X$ is (cms), both $T^*$ and $(T^*|_X)^*$ are monotone; equivalently, by Theorem 3.5, $T$ is weakly compact. \[\Box\]

**Remarks 5.3**

(i) Every reflexive Banach space is (cms); this follows from Theorem 3.5 and Fact 2.2. (See also Brezis’ [3, Theorem 1] for a related result.)
(ii) Neither $\ell_1$ nor $L_1[0,1]$ is (cms) (Example 4.2, Example 4.5).
(iii) The Banach space $\ell^*_\infty$ is not (cms) either: indeed, let $T := -G^*$, where $G$ is the Gousses operator from Example 4.2. Then $T^*$ is monotone, but $T$ is not weakly compact (Fact 2.2).
(iv) (Fitzpatrick) Suppose $X$ is isometrically isomorphic to another real Banach space $Y$ (see Reminder 5.4.(iv)). Using Proposition 5.5.(iii) below, it is easily verified that $X$ is (cms) if and only if $Y$ is. In particular: $L^*_\infty[0,1]$ is not (cms): this follows from (iii) together with the fact (due to Peckzynski) that $L^*_\infty[0,1]$ and $\ell^*_\infty$ are isometrically isomorphic (see [26, Theorem 24.4.1]). In contrast to (iii), here the operator $F$ from Example 4.5 cannot be used to obtain a counterexample (since neither $F^*$ nor $-F^*$ is monotone).

30
In the remainder of this section, we want to provide more examples of Banach spaces that are or are not (cms). To do so, we need numerous facts from the theory of Banach spaces. For the reader’s convenience, we give detailed definitions and pointers to the relevant literature.

**Reminder 5.4** Suppose $Y$ is another real Banach space. (i) A continuous linear one-to-one operator from $X$ onto $Y$ is called an *isomorphism*; the Banach spaces $X$ and $Y$ are said to be *isomorphic*. (ii) A continuous linear one-to-one operator from $X$ to $Y$ whose range is a closed subspace is called an *injection*. (iii) A continuous linear operator from $X$ onto $Y$ is called a *surjection*. (iv) Finally, an isomorphism from $X$ to $Y$ that preserves the norm of each element is called an *isometric isomorphism*; the Banach spaces $X$ and $Y$ are then *isometrically isomorphic*.

The following properties are quite useful:

**Proposition 5.5** Suppose $Q$ is a surjection from $X$ to another real Banach space $Y$. Then:

(i) $Q^*$ is a tauberian injection from $X$ to $Y$.
(ii) $Q^{**}$ is a surjection from $X^{**}$ to $Y^{**}$.
(iii) Suppose $T$ is a continuous linear operator from $Y$ to $Y^*$. Then $T$ is weakly compact if and only if $Q^*TQ$ is.

**Proof.** (i): $Q^*$ is an injection ([51, Theorem 11-3-4.(b)]), hence tauberian (Fact 2.3). (ii): $Q^{**}$ is onto by (i) and [51, Theorem 11-3-4.(a)]. (iii): “$\Rightarrow$”: fix an arbitrary $x^{**} \in X^{**}$. Then: $Q^{**}x^{**} \in Y^{**} \Rightarrow T^{**}Q^{**}x^{**} \in Y^{**} \Rightarrow Q^{**}T^{**}Q^{**}x^{**} \in X^*$ (since $Q^{**}|_{Y^*} = Q^*$). Hence ran$(Q^*TQ)^{**} \subseteq X^*$ and so $Q^*TQ$ is weakly compact. “$\Leftarrow$”: We prove the contra-positive, so suppose $T$ is not weakly compact. Then there is some $y^{**} \in Y^{**}$ such that $T^{**}y^{**} \in Y^{***} \setminus Y^*$. Since $Q^{**}$ is onto (by (ii)), we obtain $x^{**} \in X^{**}$ with $Q^{**}x^{**} = y^{**}$. However, $Q^*$ is tauberian (by (i)), hence $Q^{***}T^{**}y^{**} \in X^{***} \setminus X^*$. Altogether, $(Q^*TQ)^{**}x^{**} \in X^{***} \setminus X^*$; so $Q^*TQ$ is not weakly compact. \hfill \Box

**Reminder 5.6** Suppose $Y$ is another real Banach space. (i) $Y$ contains a complemented copy of $X$, if there exists an injection from $X$ to $Y$ whose range is a complemented subspace. Working definition: there exist a closed subspace $Z$ of $Y$, a projection $P_Z$ onto $Z$ (i.e. $P_Z$ is a continuous linear operator from $Y$ to $Y$ with ran $P_Z = Z$ and $P_ZP_Z = P_Z$), and an isomorphism from $Z$ to $X$. (The reader is referred to Jameson’s [27, Chapter 29] for a gentle and thorough introduction to the concept of a “complemented subspace”.) (ii) $Y$ contains a copy of $X$, if there exists an injection from $X$ to $Y$. (iii) $Y$ is a quotient of $X$, if there exists a surjection from $X$ to $Y$.

**Proposition 5.7** Suppose $Y$ is a quotient of $X$. If $X$ is (cms), then so is $Y$.  

31
**Proof.** Let $T$ be a surjection from $X$ to $Y$. Fix an arbitrary continuous linear monotone operator from $Y$ to $Y^*$. Then $Q^*TQ$ is a continuous linear monotone operator from $X$ to $X^*$, hence weakly compact. Proposition 5.5.(iii) implies that $T$ is weakly compact as well. \hfill $\blacksquare$

**Proposition 5.8** Suppose $X$ contains a complemented copy of $\ell_1$. Then $X$ is not (cms).

**Proof.** The hypothesis means (Reminder 5.6) that there exist a complemented closed subspace $Y$ of $X$, a continuous linear operator $P_Y$ from $X$ to $X$ with $\text{ran } P_Y = Y$ (and $P_YP_Y = P_Y$), and a continuous linear one-to-one and onto operator $I$ from $Y$ to $\ell_1$. Denote the Gossz operator from $\ell_1$ to $\ell_\infty$ of Example 4.2 by $G$. Consider $S := (IP_Y)^*G(IP_Y)$. Then $S$ is a continuous linear skew operator from $X$ to $X^*$. Since $G$ is not weakly compact, neither is $S$ (Proposition 5.5.(iii)). Hence $X$ is not (cms). \hfill $\blacksquare$

**Fact 5.9** (Bessaga and Pelczynski; see, for instance, [9, Theorem V.10 on page 48 and [27, 29.17]) TFAE: (i) $X$ contains a complemented copy of $\ell_1$; (ii) $X^*$ contains a copy of $c_0$; (iii) $\ell_1$ is a quotient of $X$.

**Fact 5.10** (see, for instance, [9, Theorem VII.6 on page 74 and Theorem on page 94])

(i) (Pelczynski) Every infinite-dimensional closed subspace of $\ell_1$ contains a complemented copy of $\ell_1$.

(ii) (Kadec and Pelczynski) Every nonreflexive closed subspace of $L_1[0,1]$ contains a complemented copy of $\ell_1$.

**Reminder 5.11** If $T$ is a continuous linear operator from $X$ to another real Banach space $Y$ with $\text{cl } T(B_X)$ compact, then $T$ is compact.

**Fact 5.12** (Schauder; see, for instance, [6, Theorem VI.3.4]) Suppose $T$ is a continuous linear operator from $X$ to another real Banach space $Y$. Then $T$ is compact if and only if $T^*$ is.

**Definition 5.13** We say that $X$ is (c) (resp. $X$ is (w)) if every continuous linear operator from $X$ to $X^*$ is compact (resp. weakly compact).

Property (w) was defined by Saab and Saab in [42]. Clearly, the following implications hold: (c) $\Rightarrow$ (w) $\Rightarrow$ (cms). It is obvious that finite-dimensional Banach spaces are (c) and that reflexive Banach spaces are (w).

**Reminder 5.14** A sequence $(x_n)$ in $X$ is called weakly Cauchy, if $(\langle x^*, x_n \rangle)$ is Cauchy (in $\mathbb{F}$), $\forall x^* \in X^*$. If every weakly Cauchy sequence in $X$ is actually weakly convergent, then $X$ is weakly sequentially complete.

32
For a proof of the following deep fact, see also [9, Chapter XI].

**Fact 5.15** (Rosenthal [39]) $X$ does not contain a copy of $\ell_1$ if and only if every bounded sequence in $X$ has a weakly Cauchy subsequence.

**Reminder 5.16** $X$ is Schur, if every weakly convergent sequence in $X$ is actually norm convergent.

**Remark 5.17** It is not hard to check that $X$ is Schur if and only if every weakly Cauchy sequence is norm convergent (Sketch: fix a weakly Cauchy sequence in a Schur space. Were it not norm convergent, then it would possess a subsequence, say $(x_n)$, with $\inf_{n \in \mathbb{N}} \|x_n - x_{n+1}\| > 0$. Now $x_n - x_{n+1}$ would converge weakly to 0, hence also in norm — contradiction!). Consequently, Schur spaces are weakly sequentially complete.

**Proposition 5.18** Suppose $X^*$ is Schur. Then $X$ is (c).

**Proof.** The space $X^*$ is Schur, thus every continuous linear operator from any other real Banach space to $X^*$ is “completely continuous” or “Dunford-Pettis” (i.e. it sends weakly Cauchy sequences to norm Cauchy sequences); it then follows from Emmanuele’s [12, Corollary 7] that $X$ does not contain a copy of $\ell_1$. Now let $T$ be an arbitrary continuous linear operator from $X$ to $X^*$. Fix an arbitrary sequence $(b_n)$ in $B_X$. By Fact 5.15, $(b_n)$ possesses a weakly Cauchy subsequence, say $(b_{k_n})$. Then $(Tb_{k_n})$ is weakly Cauchy, too. By Remark 5.17, the sequence $(Tb_{k_n})$ is norm convergent. It follows that $T(B_X)$ is relatively sequentially compact; consequently, $T$ is compact. \[\square\]

The space $c_0(\Gamma)$ from the next example and its dual $\ell_1(\Gamma)$ are discussed in Holmes’ [26, Section 16.I].

**Example 5.19** The space $c_0(\Gamma)$ is (c), for every nonempty set $\Gamma$.

**Proof.** In view of Proposition 5.18, it is enough to show that $c_0^*(\Gamma)$ is Schur. If $\Gamma$ is finite, then $c_0(\Gamma)$ is finite-dimensional and thus $c_0^*(\Gamma)$ is certainly Schur. Otherwise, use the well-known “Schurness” of $\ell_1 = c_0^*$ ([27, Proposition 27]). \[\square\]

**Fact 5.20** (Rosenthal/Pitt; see [38, Theorem A2]) Suppose $1 < p < +\infty$. Then $L_p[0,1]$ and $\ell_p$ are (w). Moreover: $\ell_p$ is (c) if and only if $2 < p$ while $L_p[0,1]$ is never (c).

**Proposition 5.21** (see also Saab and Saab’s [43, Remark following Corollary 24]) Suppose $X$ does not contain a copy of $\ell_1$ and $X^*$ is weakly sequentially complete. Then $X$ is (w).

**Proof.** As in the proof of Proposition 5.18, let $T$ be a continuous linear operator from $X$ to $X^*$ and fix an arbitrary sequence $(b_n)$ in $B_X$. Obtain a weakly
Cauchy subsequence \((g_{k_n})\) of \((b_n)\) by Fact 5.15. Then \((Tb_{k_n})\) is weakly Cauchy in \(X^*\) and hence weakly convergent. It follows that \(T(\text{cl} B_X)\) is relatively weakly sequentially compact and hence \(\text{cl} T(B_X)\) is weakly compact by Eberlein/Smulian [9, Theorem on page 18]. Thus \(T\) is weakly compact (Fact 2.2).

\(\Box\)

**Remark 5.22** “\(X^*\) is weakly sequentially complete” is not necessary for “\(X\) is (w)”: Leung [29, Section 2] constructed a James type space that is (c), does not contain a copy of \(\ell_1\), and its dual is not weakly sequentially complete.

The following proposition can be proved like Proposition 5.7.

**Proposition 5.23** Suppose \(Y\) is a quotient of \(X\). If \(X\) is (w), then so is \(Y\).

**Proposition 5.24** Suppose \(X\) is not (w) (resp. not (c)). Then \(X^{**}\) is not (w) (resp. not (c)).

**Proof.** Select a continuous linear operator from \(X\) to \(X^*\) that is not weakly compact (resp. not compact). By Fact 2.2 (resp. Fact 5.12), \(T^{**}\) is not weakly compact (resp. not compact), either. Hence \(X^{**}\) is not (w) (resp. not (c)). \(\Box\)

**Remark 5.25** It is not clear whether “\(X\) is not (cms)” implies “\(X^{**}\) is not (cms)”: the proof of Proposition 5.24 does not generalize, since conjugates of skew operators need not be skew (Example 4.2 and Example 4.5).

A complete characterization of (w) Banach lattices was provided by Saab and Saab. (For more on Banach lattices, the reader is referred to Schaefer’s [44] and Meyer-Nieberg’s [30].)

**Fact 5.26** (Saab and Saab) Suppose \(X\) is a Banach lattice. Then TFAE: (i) \(X\) is (w); (ii) \(X\) is (cms); (iii) \(X\) does not contain a complemented copy of \(\ell_1\); (iv) \(X^*\) is weakly sequentially complete.

**Proof.** “(i)\(\iff\)(iii)”: is [42, Corollaire 11]. “(iii)\(\iff\)(iv)”: (iii) \(\Leftrightarrow X^*\) does not contain a copy of \(c_0\) (Fact 5.9) \(\Leftrightarrow (iv)\) (by [30, Theorem 5.1.14.(i)]). “(i)\(\Rightarrow\)(ii)”: is trivial. “(ii)\(\Rightarrow\)(iii)”: follows from Proposition 5.8. \(\Box\)

**Examples 5.27** Every AM-space is (w). In particular: \(\ell_0\), \(\ell_1\), \(L_\infty[0,1]\), and \(C(\Omega)\) (for any compact Hausdorff space \(\Omega\)) are all (w).

**Proof.** The dual of an AM-space is an AL-space ([44, Proposition II.9.1]). And every AL-space is weakly sequentially complete ([44, Corollary to Proposition II.8.8]); so Fact 5.26 applies. For the examples, see [44, page 102]. \(\Box\)
Remarks 5.28  Note that, for instance, $c_0$ does not contain a copy of $\ell_1$ ([48, page 10]) whereas $\ell_\infty$ does (as $\ell_\infty$ is universal for all separable Banach spaces; see [26, Section 25]). In view of Examples 5.27, “not containing a copy of $\ell_1$” is not necessary for property (w). (For more on Banach spaces not containing a copy of $\ell_1$, see van Dulst’s [48].) Saab and Saab’s result (Fact 5.26) and Proposition 5.8 highlight the important role of the property “not containing a complemented copy of $\ell_1$”. We do not know whether or not there exists a Banach space that is (w) but does not contain a complemented copy of $\ell_1$.

Further facts on properties related to (w) can be found in [13] and [42].

Definition 5.29 We say that $X$ is (symmetric w), if every continuous linear symmetric operator from $X$ to $X^*$ is weakly compact.

The property (symmetric w) complements property (cms) nicely:

Proposition 5.30 $X$ is (w) if and only if $X$ is both (symmetric w) and (cms).


□

Remarks 5.31

(i) For more on (symmetric w) in complex Banach spaces, see [1].
(ii) According to Gutiérrez [22, page 151], it is an open problem whether or not a Banach space with property (symmetric w) is necessarily (w).
(iii) Similarly, we do not know whether or not a Banach space that is (cms) is necessarily (w).
(iv) In view of Proposition 5.30, the questions asked in (ii) and (iii) are really about the relationship between the properties (symmetric w) and (cms).
(v) Suppose $X$ is a Banach space that allows us to write every continuous linear symmetric operator from $X$ to $X^*$ as the difference of two continuous linear monotone symmetric operators. Then, using Theorem 2.17: $X$ is (w) if and only if $X$ is (cms).
(vi) Let us now briefly outline that the approach suggested in (v) works in Hilbert space: indeed, given a continuous linear operator $T$ from $X$ to $X^*$, define the positive (resp. negative) part of $T$ by $T_+ := \frac{1}{2}[T] + \frac{1}{2}T$ (resp. $T_- := \frac{1}{2}[T] - \frac{1}{2}T$), where $[T] := \sqrt{T^2}$ (Spectral Theorem1). Then $T_+, T_-$ are continuous linear monotone symmetric operators and $T = T_+ - T_-$. (Since the notions considered depend only on the topology rather than the given norm, the same is true for all renormed Hilbert spaces; in particular, finite-dimensional Banach spaces.)

35
(vi) There exists a continuous linear symmetric operator from \( t_1 \) to \( t_\infty \) that cannot be written as the difference of two continuous linear monotone symmetric operators: indeed, Aron et al. [1, page 83] provide a continuous linear symmetric operator that is not weakly compact (and necessarily not monotone; see Theorem 3.5). In particular, \( t_1 \) is not (symmetric w).
(vii) It would be interesting to know if the notions of the positive (resp. negative) part of an operator (see (vi)) can be made meaningful for a "reasonable" class of Banach spaces.
(viii) Proposition 5.7 and Proposition 5.8 remain valid, if we replace "(cms)" by "(symmetric w)" (for the proof of Proposition 5.8, use the operator mentioned in (vii) rather than the Gossez operator).
(ix) In view of (ix), we could add "\( X \) is (symmetric w)" to the list of items characterizing (w) Banach lattices in Fact 5.26.

Summary 5.32 Based on powerful result from the theory of Banach spaces, we saw in this section that the examples by Gossez on \( t_1 \) and by Fitzpatrick and Phelps on \( L_1[0,1] \) (see Example 4.2, Example 4.5) are in some sense "the only examples of bad continuous linear monotone operators": indeed, many classical Banach spaces like \( c_0, c, t_\infty, L_\infty[0,1], C[0,1] \) as well as all reflexive Banach spaces are conjugate monotone spaces and thus are unable to host any "bad" continuous linear monotone operators.

6 Nonlinear results

Sums

Proposition 6.1 Suppose \( M \) is a monotone operator from \( X \) to \( X^* \) and \( S \) is a continuous linear skew operator from \( X \) to \( X^* \). If \( M \) and \( S \) are of type (NI), then so is \( M + S \); moreover: \( (M + S)x^* = M_0x^* + S_0x^* \), \( \forall x^* \in X^* \).

Proof. Suppose \( M \) and \( S \) are of type (NI), i.e. \( M_0 = \mathcal{M} \) and \( S_0 = \mathcal{S} \) (Proposition 1.11.(iii)). Fix \( (x^*, x^*) \in X^{**} \times X^* \). It is clear that \( M + S \) is monotone and that \( (M + S)x^* \subseteq \mathcal{M} + Sx^* \) (Proposition 1.6). Then for every \( (u, u^*) \in \text{gra} M \):

\[
(x^{**} - u, x^* - (u^* + S u)) \leq (x^{**} - u, x^* - u^*) + (S^* x^{**}, -u) \leq (x^{**} - u, x^* - u^*) + (S^* x^{**}, x^{**} - u) - (S^* x^{**}, x^{**}) = (x^{**} - u, (x^* + S^* x^{**}) - u^*) - (S^* x^{**}, x^{**}).
\]

Since \( S^* \) is monotone (Theorem 3.3), we obtain:

\[
\inf_{(v, v^*) \in \text{gra} (M + S)} (x^{**} - v, x^* - v^*) \leq -\langle S^* x^{**}, x^{**} \rangle + \inf_{(u, u^*) \in \text{gra} M} (x^{**} - u, (x^* + S^* x^{**}) - u^*) \leq \inf_{(u, u^*) \in \text{gra} M} (x^{**} - u, (x^* + S^* x^{**}) - u^*).
\]
It follows that \( M + S \) is of type (NI): \((M + S)_0 = \overline{M + S} \). If \( x^* \in \overline{M + S} \), then (from the inequalities above) \( x^* + S^*x^* \in \overline{M}x^* = M_0x^* \) so that \( \langle S^*x^*, x^* \rangle = 0 \). Hence (Proposition 2.21(ii) and Theorem 3.3) \(-S^*x^* \in S_0x^* = \overline{S}x^* \) and thus altogether \( \overline{M + S} \subseteq \overline{M}x^* + \overline{S}x^* = M_0x^* + S_0x^* \). Finally, let \( y^* \in M_0x^* \) and \(-S^*x^* \in S_0x^* \), i.e. \( \langle S^*x^*, x^* \rangle = 0 \) (Proposition 2.21(ii)). Then

\[
0 = \inf_{(u, u^*) \in \text{gra} M} \{ x^* - u, y^* - u \} \\
= \inf_{(u, u^*) \in \text{gra} M} \{ x^* - u, (y^* - u) + (-S^*x^* - Su) \} \\
= \inf_{(v, v^*) \in \text{gra} (M + S)} \{ x^* - v, (y^* - S^*x^*) - v \};
\]

therefore, \( y^* - S^*x^* \in (M + S)_0x^* \) and the proof is complete. \( \square \)

**Corollary 6.2** Suppose \( f \) is a convex lower semi-continuous proper function on \( X \) and \( T \) is a continuous linear monotone operator from \( X \) to \( X^* \). If \( T \) is of type (NI), then so is \( \partial f + T \).

**Proof.** Decompose \( T \) into its symmetric part \( P \) and its skew part \( S \) (Proposition 2.14). Then \( P \) is a subdifferential (Proposition 2.15) and so is \( \partial f + P \); as such, \( \partial f + P \) is of dense type (Fact 1.14(i)) and hence of type (NI) (Fact 1.13).

If \( T \) is of type (NI), then so is \( S \) (Theorem 3.3) and, by Proposition 6.1, \( (\partial f + P) + S = \partial f + T \) is of type (NI) as well. \( \square \)

**Definition 6.3** We call a set-valued map \( K \) from \( X \) to another real Banach space \( Y \) *compact*, if \( \text{cl} K(B) \) is compact, for every bounded subset \( B \) of \( X \).

A continuous linear operator is compact in this sense if and only if it is compact in the sense of Reminder 5.11; hence Definition 6.3 is a reasonable generalization.

**Proposition 6.4** Suppose \( K, M \) are monotone operators from \( X \) to \( X^* \), \( K \) is compact, dom \( K \supseteq \text{dom} M \), and \( x^* \in X^* \). Then \( M_1x^* \subseteq (K + M)_1x^* = K_1x^* \). If \( K \) is at most single-valued, then \( (K + M)_1x^* = K_1x^* + M_1x^* \).

**Proof.** Fix \( (x^*, x^*) \in \text{gra} M_1 \) and obtain a bounded net \((x_\alpha, x_\alpha^*) \) in \( \text{gra} M \) with \( x_\alpha \rightharpoonup x^* \) and \( x_\alpha^* \rightharpoonup x^* \). By hypothesis, \( \text{cl} K(x_\alpha) \) is compact, hence there are subsequences \((x_\beta) \) of \((x_\alpha) \) and \( y_\beta^* \in Kx_\beta \) such that \( y_\beta^* \rightarrow y^* \in K_1x^* \). Now \((x_\beta, x_\beta^* + y_\beta^* \) \( \in \text{gra} (K + M) \), \( x_\beta \rightharpoonup x^* \), \( x_\beta^* + y_\beta^* \rightharpoonup x^* + y^* \); consequently, \( x^* + y^* \in (K + M)_1x^* \) and thus \( x^* \in (K + M)_1x^* = K_1x^* \).

“II” part: from the above, \( K_1x^* + M_1x^* \subseteq (K + M)_1x^* \). Let’s now pick \( x^* \in (K + M)_1x^* \). Then there is a bounded net \((x_\alpha, Kx_\alpha + z_\alpha^*) \) in \( \text{gra} (K + M) \) with \((x_\alpha, z_\alpha^*) \) \( \in \text{gra} M \), \( x_\alpha \rightharpoonup x^* \), and \( Kx_\alpha + z_\alpha^* \rightharpoonup x^* \). After passing to a subnet if necessary, we assume \( Kx_\alpha \rightarrow y^* \in K_1x^* \). Hence \( z_\alpha^* \rightarrow x^* - y^* \in M_1x^* \) and the entire proposition is proven. \( \square \)
Proposition 6.5 Suppose $M$ is a monotone operator from $X$ to $X^*$ and $S$ is a continuous linear skew operator from $X$ to $X^*$. If $M$ is of dense type and $S$ is compact, then $M+S$ is of dense type; in fact: $(M+S)x^{**} = Mx^{**} + Sx^{**} = M_1x^{**} + S^{**}x^{**}, \forall x^{**} \in X^{**}$.

Proof. Suppose $M$ is of dense type and $S$ is compact. Fix $x^{**} \in X^{**}$. Then $M_1 = \overline{M}$ and $S$ is of dense type (and of type (NI); see Theorem 3.5). Hence, by Proposition 6.1, $\overline{M+S}x^{**} = M_1x^{**} + Sx^{**} = Mx^{**} + S^{**}x^{**}$. On the other hand, since $S$ is compact and single-valued, Proposition 6.4 and Proposition 1.6 apply and give $M_1x^{**} + Sx^{**} = (M+S)x^{**} \subseteq \overline{M+S^{**}}$. Altogether, $(M+S)x^{**} = \overline{M+S}x^{**}$ and the proof is complete.

Corollary 6.6 Suppose $f$ is a convex lower semi-continuous proper function on $X$ and $T$ is a continuous linear monotone operator from $X$ to $X^*$. If the skew part of $T$ is compact or if $X$ is (c), then $\partial f + T$ is of dense type.

Proof. If $X$ is (c), then the skew part of $T$ is certainly compact. So, much as in the proof of Corollary 6.2, denote the symmetric (resp. skew) part of $T$ by $P$ (resp. $S$). Then $\partial f + P$ is of dense type and $S$ is compact; hence Proposition 6.5 applies.

The duality map and rugged Banach spaces

Reminder 6.7 The subdifferential map for the function $\frac{1}{2}\|\cdot\|^2$ on $X$ is called the (normalized) duality map and denoted by $J$: $Jx := \{x^* \in X^* : \langle x^*, x \rangle = \|x^*\|^2 = \|x\|^2\}, \forall x \in X$.

In Hilbert space, the duality map is just the identity; the duality map is a very successful attempt to mimic this outside Hilbert spaces. It is well-known that $J$ is homogeneous and that $Jx$ is bounded, convex, weak* closed, and nonempty, $\forall x \in X$; see, for instance, [34].

Some notation before we start: denote the ordinary signum function by sign, i.e. sign $r = +1$, if $r > 0$; sign $0 = 0$; sign $r = -1$, if $r < 0$. For some of the next examples, it will be convenient to abbreviate the subdifferential map of the absolute value on $\mathbb{R}$ by Sign; thus: Sign $r = +1$, if $r > 0$; Sign $0 = [-1, +1]$; Sign $r = -1$, if $r < 0$. Also, if $\Gamma$ is a nonempty set and $\gamma \in \Gamma$, then denote the unit vector corresponding to $\gamma$ in $c_0(\Gamma)$ etc. by $e_\gamma$.

Example 6.8 Suppose $\Gamma$ is a nonempty set. Then $Jx = \overline{\text{conv}} \{x_\gamma \cdot e_\gamma : |x_\gamma| = \|x\|, \forall x \in c_0(\Gamma)\}$.

Example 6.9 Suppose $\Gamma$ is a nonempty set. Then $(Jx)_\gamma = \|x\| \text{Sign} x_\gamma, \forall x \in \ell_1(\Gamma) \forall \gamma \in \Gamma$.
Example 6.10 Suppose $(A, \mathcal{A}, \mu)$ is a $\sigma$-finite complete measure space. Then 
$$(Jf)(a) = \|f\| \text{Sign} f(a), \forall f \in L_1(A, \mathcal{A}, \mu) \forall a \in A.$$ 

Example 6.11 (Deimling’s [7, Example 12.2]) Suppose $\Omega$ is a compact Hausdorff space. Recall that the dual of $C(\Omega)$ is $M(\Omega)$, the space of finite signed Baire measures on $\Omega$ (for details, see [23, 41]). The total variation (resp. support) of an arbitrary measure $\mu \in M(\Omega)$ is denoted $\|\mu\|_s(\Omega)$ (resp. $\text{supp}\mu$). Let $C^+(\Omega)$ be the set of all positive measures on $\Omega$. Fix an arbitrary $f \in C(\Omega)$ and abbreviate $\Omega_f := \{\omega \in \Omega : |f(\omega)| = \|f\|\}$. Then: 

$$Jf = \{\mu \in M(\Omega) : \|\mu\| = \|\mu\|_s(\Omega) = \|f\|, \text{ supp}\mu \subseteq \Omega_f, \text{ and } \text{sign} f \in C^+(\Omega)\}.$$ 

Remark 6.12 The above examples are part of the folklore. Unfortunately, we have not been able to find a rich or detailed collection of examples in the literature. The reader is referred to Deimling’s [7, Subsection 12.2] and Holmes’ [26, Section 20.7] for results that are helpful in verifying the above examples.

Definition 6.13 We say that $X$ is rugged, if $\text{cl span ran } (J - J) = X^*$. 

Rugged spaces have to have points where the norm is not Gâteaux differentiable; they are never smooth. In particular, none of the following spaces is rugged: Hilbert spaces, uniformly convex spaces, $L_p[0, 1]$ for $1 < p < +\infty$.

Remark 6.14 Leach and Whitfield [28, page 121] coined the notion of a rough norm, where rough is also meant in a “nonsmooth sense”. However, the properties rough and rugged are unrelated (the reader is referred to [8, 34] for further information and notation): on the one hand, the Euclidean plane $\mathbb{R}^2$ is an Asplund space and thus does not admit an equivalent rough norm ([8, Theorem 1.5.3]) but does admit an equivalent rugged norm (Proposition 6.22 below). Hence rugged $\not\implies$ rough. On the other hand, Phelps constructed an equivalent norm on $\ell_1$ that is Gâteaux differentiable but nowhere Fréchet differentiable ([34, Example following Theorem 5.12]). By Deville et al.’s [8, Theorem III.1.9], the space $\ell_1$ admits an equivalent rough Gâteaux differentiable norm; this norm cannot be rugged. Thus rough $\not\implies$ rugged.

Example 6.15 Suppose $\Gamma$ is a nonempty set. Then $c_0(\Gamma)$ is rugged if and only if $\Gamma$ contains at least two elements.

Proof. If $\Gamma$ is a singleton, then $c_0(\Gamma)$ is essentially $\mathbb{R}$ which is not rugged. Now suppose $\Gamma$ contains at least two elements. Since $c_0^0(\Gamma) = \ell_1(\Gamma)$ and $\text{cl span } \{e_\gamma : \gamma \in \Gamma\} = \ell_1(\Gamma)$, it suffices to show that $e_\gamma \in \text{span ran } (J - J)$. So fix $\gamma \in \Gamma$ and any other $\delta \in \Gamma$, where $\gamma \neq \delta$. Let $x := e_\gamma + e_\delta$ and $y := e_\gamma - e_\delta$. Then, by Example 6.8, $Jx = \text{conv } \{e_\gamma, e_\delta\}$ and $Jy = \text{conv } \{e_\gamma, -e_\delta\}$. Thus $e_\gamma - e_\delta \in Jx - Jx$ and $e_\gamma - (-e_\delta) \in Jy - Jy$. Therefore, $e_\gamma \in \frac{1}{2}[(Jx - Jx) + (Jy - Jy)]$ and we are done. \(\square\)
Example 6.16 Suppose $\Gamma$ is a nonempty set. Then $\ell_1(\Gamma)$ is rugged if and only if $\Gamma$ contains at least two elements.

Proof. Again we can assume that $\Gamma$ contains at least two different elements. We show that $\ell_1(\Gamma)$ is rugged for $\Gamma = \mathbb{N}$ since it is notationally much more convenient (the other cases are proved analogously). By Example 6.9,

$$\bigcup_{r \geq 0} r[(J e_1 - J e_1) + (J e_2 - J e_2)] \supseteq \bigcup_{r \geq 0} r([-1, +1], [-1, +1], \ldots) = \ell_\infty;$$

the proof is complete.

Example 6.17 Suppose $(A, \mathcal{A}, \mu)$ is a $\sigma$-finite complete measure space. Then $L_1(A, \mathcal{A}, \mu)$ is rugged if and only if $\mathcal{A}$ contains at least two disjoint sets of finite strictly positive measure.

Proof. Once more, we can assume that there are $A_1, A_2 \in \mathcal{A}$ with $A_1 \cap A_2 = \emptyset$ and $0 < \mu(A_1), \mu(A_2) < +\infty$. Let $f_i := \frac{1}{\mu(A_i)} \chi_{A_i}$ for $i = 1, 2$ (where $\chi$ denotes the characteristic function), and proceed as in the proof of Example 6.16.

Example 6.18 Suppose $\Omega$ is a compact Hausdorff space. Then $C(\Omega)$ is rugged if and only if $\Omega$ contains at least two elements.

Proof. The reader guessed right that we can assume the existence of two distinct points $\omega_1, \omega_2 \in \Omega$. Pick two disjoint open neighborhoods $U_1$ of $\omega_1$, and $U_2$ of $\omega_2$. Define $E_1 := \text{cl} U_1$, $E_2 := \Omega \setminus U_1$, and $E_3 := E_1 \cap E_2$. Then $E_1, E_2, E_3$ are closed with $\omega_2 \notin E_1, \omega_1 \notin E_2, \omega_2 \notin E_2$, and $E_1 \cup E_2 \cup E_3 = \Omega$. Let $\mathcal{B}$ be the class of all Baire sets of $\Omega$. If $E$ is a closed subset of $\Omega$ and $\mu \in M(\Omega)$, then define $\mu_E := \mu(B \cap E), \forall B \in \mathcal{B}$. Now fix an arbitrary finite signed Baire measure $\mu \in M(\Omega)$. Then we can decompose $\mu$ into

$$\mu = \mu_{E_1} + \mu_{E_2} - \mu_{E_3}.$$ 

We want to show that $C(\Omega)$ is rugged: by the above decomposition, it is enough to show that $\mu_{E_i} \in \text{span ran } (J - J)$. Moreover, in view of the Jordan decomposition for signed measures and the homogeneity of $\text{span ran } (J - J)$, we can assume WLOG that $\mu$ is a probability measure (i.e. a positive measure with $|\mu| = \mu(\Omega) = 1$). Altogether, it suffices to establish the following

Key step: If $E$ is a closed convex subset of $\Omega$, $\mu$ is a probability measure on $\Omega$ with $\text{supp } \mu \subseteq E$, and $\omega_0 \in \Omega \setminus E$, then $\mu \in \text{span ran } (J - J)$.

Recall that every compact Hausdorff space is normal ([27, Corollary 9.15]). By Tietze’s Extension Theorem ([27, Theorem 12.4]), there exists $f \in C(\Omega)$ with $f|_E \equiv +1$ and $f(\omega_0) = -1$. Let $\delta_0 \in M(\Omega)$ be the Dirac measure at $\omega_0$: $\delta_0(B) = +1$, if $\omega \in B$; $\delta_0(B) = 0$, otherwise, $\forall B \in \mathcal{B}$. A direct check or Example 6.11 shows that

$$\mu, -\delta_0 \in J f.$$
If \( l \) denotes the function in \( C(\Omega) \) that is identically equal to 1, then similarly

\[
\mu + \delta_0 \in J_1.
\]

Consequently, \( \mu + \delta_0 \in (J - J)(f) \) and \( \mu - \delta_0 \in (J - J)(1) \) which yields the desired \( \mu \in \frac{1}{2}(J - J)(f) + (J - J)(1) \subseteq \text{span ran } (J - J). \)

\( \square \)

**Corollary 6.19** Every AM space with unit that is at least two-dimensional is rugged.

**Proof.** By a well-known result of Kakutani (see [44, Corollary II.7.1] or [30, Theorem 2.1.3]), every AM space with unit is isometrically isomorphic to some \( C(\Omega) \), where \( \Omega \) is a compact Hausdorff space.

\( \square \)

**Example 6.20** \( C[0, 1] \) is rugged.

**Example 6.21** Suppose \( \Gamma \) is a nonempty set. Then \( c(\Gamma) \) (resp. \( l_\infty(\Gamma) \)) is rugged if and only if \( \Gamma \) contains at least two elements.

**Proof.** These spaces are AM spaces with unit and they have dimensions greater or equal to 2. Alternatively, equip \( \Gamma \) with the discrete topology. Then \( c(\Gamma) \) (resp. \( l_\infty(\Gamma) \)) can be identified with \( C(\Omega) \), where \( \Omega \) is the one-point–compactification (resp. Stone–Cech compactification).

\( \square \)

It is amusing that (almost) all Banach spaces can be renormed to become rugged:

**Proposition 6.22** (Vanderwerff) \( X \) admits an equivalent rugged norm if and only if the dimension of \( X \) greater than 1.

**Proof.** “\( \Rightarrow \)”: neither \( \{0\} \) nor \( \mathbb{R} \) is rugged. “\( \Leftarrow \)”: fix an arbitrary \( u \in X \) with \( \|u\| = 1 \) and \( u^* \in J_u \). Then \( X = \ker u^* \oplus \mathbb{R}u \) which allows us to view \( X \) as \( X = Y \oplus \mathbb{R} \), where \( Y := \ker u^* \). Define a norm on \( X \) by \( \|y, r\| := \|y\| + |r|, \forall (y, r) \in X = Y \oplus \mathbb{R} \). Clearly, \( \|\cdot\| \leq \|\cdot\| \); thus, by the Inverse Mapping Theorem ([6, Theorem II.1.3]), these norms are equivalent. The norm in \( (X, \|\cdot\|) \) is given by \( \|\cdot\| := \max\{\|y\|, |r^*|\}, \forall y^* \in Y^* \forall r^* \in \mathbb{R}^* = \mathbb{R} \). For the remainder of this proof, \( J \) denotes the duality map of \( (X, \|\cdot\|) \). Then \( J(y, r) = \{(y^*, r^*) \in Y^* \times \mathbb{R} : \langle y^*, y \rangle + r^*r = \|y\| + |r| = \max\{\|y\|, |r^*|\}\} \). On the one hand, \( J(u, 0) \supseteq \{u^*\} \times \{-1, +1\} \) so that \( J(u, 0) - J(u, 0) \supseteq \{0\} \times [-2, +2] \).

On the other hand, \( J(0, 1) \supseteq 2B_{Y^*} \times \{1\}; \) thus \( J(0, 1) - J(0, 1) \supseteq 2B_{Y^*} \times \{0\} \).

Altogether, \( \bigcup_{n \in \mathbb{N}} \overline{(J(u, 0) - J(u, 0)) + (J(0, 1) - J(0, 1))} = Y^* \times \mathbb{R} = X^* \).

Consequently, \( (X, \|\cdot\|) \) is rugged.

\( \square \)
Regularization

**Definition 6.23** Suppose $T$ is a monotone operator from $X$ to $X^*$ and $\lambda > 0$. Then the operator $T + \lambda J$ is called a regularized of $T$.

The reader is referred to [54] for motivation on the regularization. We start with some basic properties.

**Proposition 6.24** Suppose $T$ is a continuous linear monotone operator from $X$ to $X^*$. Then the regularization $T + \lambda J$ is coercive and maximal monotone, $\forall \lambda > 0$.

**Proof.** Coercivity is obvious from $\frac{1}{\lambda^2} \langle (T + \lambda J)x, x \rangle \geq \lambda \|x\|$. The operators $T$ and $\lambda J$ are maximal monotone and their domain is the entire space; the same is true for $T + \lambda J$ (Fact 1.16). \hfill \Box

**Definition 6.25** Suppose $\epsilon \geq 0$. Then the $\epsilon$-subdifferential map of the function $\frac{1}{2} \| \cdot \|^2$ on $X$ is denoted by $J_\epsilon$; thus $J_\epsilon := \{ x^* \in X^* : \langle x^*, x \rangle \geq \frac{1}{2} \| x \|^2 + \frac{1}{2} \| x^* \|^2 - \epsilon \}, \forall x \in X$.

Approximate or $\epsilon$-subdifferentials have become a useful tool in optimization; for some motivation, see [25, Chapter XI]. (The reader should be warned that the notation of Definition 6.25 is not quite unambiguous: on the one hand, $J_\epsilon$ has a precise meaning for $\epsilon$ equal to 0 or 1 (for $\epsilon = 0$, we recover $J$). On the other hand, one could say that $J_0 = J_1$ is $(J^*)^{-1}$ (see Definition 1.4 and Fact 1.14(i)). However, for the remainder of this section, $J_\epsilon$ is always meant to be as in Definition 6.25.)

**Fact 6.26** (Gossez’ [18, Théorème 4.1]) Suppose $T$ is a monotone operator of dense type from $X$ to $X^*$. Then $\text{ran} \left( T + \lambda J_\epsilon \right) = X^*$, $\forall \lambda > 0 \forall \epsilon > 0$.

**Fact 6.27** (Gossez’ [19, Lemma 1]) Suppose $T$ is a monotone operator from $X$ to $X^*$ and $\lambda > 0$. If $\text{cl ran}(T + \lambda J) = X^*$, then $\text{ran}(T + \lambda J_\epsilon) = X^*$, $\forall \epsilon > 0$.

The converse holds true in the continuous linear setting. (The proof below works more generally for “uniformly continuous” operators, but there is no need to make this precise.)

**Proposition 6.28** Suppose $T$ is a continuous linear operator from $X$ to $X^*$ and $\lambda > 0$. If $\text{ran}(T + \lambda J_\epsilon) = X^* \forall \epsilon > 0$, then $\text{cl ran}(T + \lambda J) = X^*$.

**Proof.** Fix $z^* \in X^*$. For every $\epsilon > 0$, there exists $x_\epsilon \in X$ and $u^*_\epsilon \in J_\epsilon x_\epsilon$ such that $z^* = Tx_\epsilon + \lambda u^*_\epsilon$. By Brondsted and Rockafellar’s [1, Lemma], there exist $y_\epsilon \in X$ and $v^*_\epsilon \in J_\epsilon y_\epsilon$, such that $\| x_\epsilon - y_\epsilon \| \leq \sqrt{\epsilon}$ and $\| u^*_\epsilon - v^*_\epsilon \| \leq \sqrt{\epsilon}$. Then $Ty_\epsilon + \lambda v^*_\epsilon \in \text{ran}(T + \lambda J)$ and

$$\| z^* - (Ty_\epsilon + \lambda v^*_\epsilon) \| = \| (Tx_\epsilon + \lambda u^*_\epsilon) - (Ty_\epsilon + \lambda v^*_\epsilon) \| \leq (\| T \| + \lambda) \sqrt{\epsilon}.$$
Letting $\epsilon$ tend to 0 from above yields $z^* \in \text{cl ran} \ (T + \lambda J)$. \hfill $\square$

The next proposition shows that in [19, Lemma 2] Gossez proved more than actually stated:

**Proposition 6.29** Suppose $T$ is a monotone operator from $X$ to $X^*$ and $\epsilon \geq 0$. If there exists a sequence of positive reals $(\lambda_n)$ with $\lambda_n \to 0$ and ran $(T + \lambda_n J) = X^*$, then $T$ is of range-dense type.

**Proof.** In view of Proposition 1.6 and Proposition 1.11(ii), it suffices to show that ran $T \subseteq$ ran $T_1$. So fix $z^* \in \text{ran} \ T$, say $z^* \in T_{x^*}^*$ for some $z^* \in X_{x^*}$. Obtain for every $n \in \mathbb{N}$ a vector $x_n \in X$ with $z^* = y_n^* + \lambda_n u_n^*$, where $y_n^* \in Tx_n$ and $u_n^* \in Jx_n$. Then (by definition of $T$) $\langle z^*, x_n, z^* - y_n^* \rangle \geq 0$ and so $\langle z^*, u_n^* \rangle \geq \langle x_n, u_n^* \rangle$. Using the definition of $J$, we estimate $\epsilon \|x_n\|^2 \leq \epsilon^2 + \frac{1}{2} \|z^*\|^2$; the sequence $(x_n)$ stays bounded and so does the sequence $(u_n^*)$. It follows that $\lambda_n u_n^* \to 0$. Moreover, there is a subnet $(x_{n_k})$ of $(x_n)$ with $x_{n_k} \xrightarrow{w} x^{**}$, for some $x^{**} \in X^{**}$. Now $(x_{n_k}, y_{n_k}^*) \in \text{gra} \ T$, $x_{n_k} \xrightarrow{w} x^{**}$, and $y_{n_k}^* = z^* - \lambda_n u_{n_k}^* \to z^*$; hence $(x^{**}, z^*) \in \text{gra} \ T_1$. Consequently, $z^* \in \text{ran} \ T_1$, as desired. \hfill $\square$

**Theorem 6.30** Suppose $T$ is a continuous linear monotone operator from $X$ to $X^*$. Then TFAE:

(i) $T$ is of dense type.

(ii) ran $(T + \lambda J) = X^*$, $\forall \lambda > 0$ $\forall \epsilon > 0$.

(iii) cl ran $(T + \lambda J) = X^*$, $\forall \lambda > 0$.

(iv) $T + \lambda J$ is of range-dense type, $\forall \lambda \geq 0$.

(v) There exists a sequence $(\lambda_n)$ of positive reals tending to 0 with cl ran $(T + \lambda_n J) = X^*$, $\forall n \in \mathbb{N}$.

(vi) There exist $\epsilon > 0$ and a sequence $(\lambda_n)$ of positive reals tending to 0 with $\text{ran} \ (T + \lambda_n J) = X^*$, $\forall n \in \mathbb{N}$.

**Proof.** “(i)$\Rightarrow$(ii)”: Fact 6.26. “(ii)$\Leftrightarrow$(iii)”: Fact 6.27 and Proposition 6.28. “(iii)$\Rightarrow$(iv)”: fix $\lambda \geq 0$ and $\epsilon > 0$. Then cl ran $((T + \lambda J) + \frac{1}{\lambda} J) = X^*$; hence (Fact 6.27) on $T + \lambda J$ ran $((T + \lambda J) + \frac{1}{\lambda} J_n) = X^*$, $\forall n \in \mathbb{N}$. By Proposition 6.29, the regularization $T + \lambda J$ is of range-dense type. “(iii)$\Rightarrow$(v)”: is trivial.

“(v)$\Rightarrow$(vi)”: immediate from Fact 6.27. “(iv)$\Rightarrow$(i)”: $T + 0J = T$ is monotone and of range-dense type, hence of dense type (Theorem 3.3). “(vi)$\Rightarrow$(i)”: By Proposition 6.29, $T$ is of range-dense type, thus of dense type (Theorem 3.3). \hfill $\square$

**Remark 6.31** By contrast, Simons’ [45, Theorem 12(a)] shows that if $T$ is an arbitrary monotone operator of type (NI) from $X$ to $X^*$, then $T + \lambda J$ is onto, for every $\lambda > 0$. 

43
The next two results apply to the various rugged spaces identified in the previous subsection.

**Proposition 6.32** Suppose $X$ is rugged, $T$ is a continuous linear operator from $X$ to $X^*$, and $\lambda > 0$. Then TFAE: (i) $T + \lambda J$ is locally maximal monotone; (ii) $\text{cl ran} (T + \lambda J)$ is convex; (iii) $\text{cl ran} (T + \lambda J) = X^*$.

**Proof.** “(i)⇒(ii)”: follows from Fact 1.15. “(ii)⇒(iii)”: Otherwise, assume to the contrary that there exists $x^* \in X^* \setminus (\text{cl ran} (T + \lambda J))$ which we thus separate: $(x^{**}, x^*) > \sup_{x \in X} (x^{**}, (T + \lambda J)x)$, for some $x^{**} \in X^{**} \setminus \{0\}$. The homogeneity of $T$ and $J$ implies $(x^{**}, (T + \lambda J)x) \equiv 0$, $\forall x \in X$. Taking the difference ($T$ is single-valued!) and dividing by $\lambda$ yields $(x^{**}, Jx - Jx) = 0$, $\forall x \in X$. Hence $x^{**} \in (\text{ran} (J - J))^\perp = (\text{cl span ran} (J - J))^\perp = \{0\}$, an impossibility. “(iii)⇒(i)”: By Proposition 6.24, $T + \lambda J$ is maximal monotone and coercive, hence Fact 1.19 yields (i). □

**Remark 6.33** Gossez proved in [20] the following: “Let $G$ be the Gossez operator from Example 4.2 and $\lambda > 0$ such that $\text{cl ran} (G + \lambda J) \neq \ell_\infty$. Then $\text{cl ran} (G + \lambda J)$ is not convex.” His proof is quite involved (relying on the Stone-Čech compactification and the definition of $G$). However, his result also follows from Example 6.16 and Proposition 6.32 – in a very structured and almost effortless way!

In rugged spaces, some more items can be added to the list of characterizations of Theorem 6.30:

**Theorem 6.34** Suppose $X$ is rugged and $T$ is a continuous linear monotone operator from $X$ to $X^*$. Then TFAE:

(i) $T$ is of dense type.

(ii) $T + \lambda J$ is locally maximal monotone, $\forall \lambda \geq 0$.

(iii) There exists a sequence $(\lambda_n)$ of positive real numbers tending to 0 such that $\text{cl ran} (T + \lambda_n J)$ is convex, $\forall n \in \mathbb{N}$.

**Proof.** “(i)⇒(ii)”: $T = T + 0 J$ is locally maximal monotone by Theorem 3.3. Suppose now $\lambda > 0$. By Theorem 6.30, $\text{cl ran} (T + \lambda J) = X^*$. Hence (Proposition 6.32) $T + \lambda J$ is locally maximal monotone. “(ii)⇒(iii)”: clear from Fact 1.15. “(iii)⇒(i)”: by Proposition 6.32, $\text{cl ran} (T + \lambda_n J) = X^*$, $\forall n \in \mathbb{N}$.

Now apply Theorem 6.30. □

**Remarks 6.35** The operator $F$ from Example 4.5 is not of dense type and $L_1[0, 1]$ is rugged (Example 6.17); hence, by Theorem 6.34, we expect $\text{cl ran} F + \lambda J$ to be nonconvex, for all small $\lambda > 0$. In fact, Fitzpatrick and Phelps showed directly that $\text{cl ran} (F + 1 J)$ is nonconvex; see [15, Example 3.2]. It follows that
$F + J$ is neither of range-dense type nor locally maximal monotone (Fact 1.15). Suppose $X$ is rugged and $T$ is a continuous linear monotone operator from $X$ to $X^*$ that is not of dense type (see Example 4.2 or Example 4.5). Then $T + \lambda J$ is neither of range-dense type nor locally maximal monotone, for all small $\lambda > 0$: indeed, Theorem 6.34 yields nonconvex $\text{cl ran}(T + \lambda J)$, for all small $\lambda > 0$ and so Fact 1.15 applies.

We conclude with a theorem that applies in particular to $c_0(\Gamma)$, for every nonempty set $\Gamma$.

**Theorem 6.36** Suppose $X$ is (c) and $T$ is a continuous linear monotone operator from $X$ to $X^*$. Then $T + \lambda J$ is of dense type and locally maximal monotone, $\forall \lambda \geq 0$.

**Proof.** For $\lambda = 0$, this follows from Theorem 3.5. So suppose $\lambda > 0$. Corollary 6.6 implies that $T + \lambda J$ is of dense type. Then Theorem 6.30 yields $\text{cl ran}(T + \lambda J) = X^*$, $\forall \lambda > 0$. Since $T + \lambda J$ is maximal monotone and coercive (Proposition 6.24), we apply Fact 1.19 and conclude that $T + \lambda J$ is locally maximal monotone, $\forall \lambda > 0$. \hfill $\square$

**Remark 6.37** The conclusion of Theorem 6.36 remains true if we replace "X is (c)" by "X is reflexive" (Proposition 6.24 and Fact 1.8). Similar remarks apply to some of the other results in this subsection (although the known results are usually stronger).

**Acknowledgments**

It is our great pleasure to thank Tamás Erdélyi, Simon Fitzpatrick, and Jon Vanderwerff for various helpful comments and discussions. We also wish to thank Robert Phelps and Steven Simons for kindly sending us the manuscripts [36, 45].

**References**


