3B2v7.51c GML4.3.1 YJFAN : 4071 Prod.Type:com pp.1–22(col.fig.:NIL) ED:Susan PAGN: dinesh SCAN: NIL

ARTICLE IN PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

ACADEMIC PRESS

Journal of Functional Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

JOURNAL OF Functional Analysis

http://www.elsevier.com/locate/jfa

Variational composition of a monotone operator with a linear mapping with applications to elliptic PDEs with singular coefficients

Teemu Pennanen,[a,1] Julian P. Revalski,[b,2] and Michel Théra[c,*,3]

[a] *Department of Management Science, Helsinki School of Economics, PL 1210, 00101 Helsinki, Finland*
[b] *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Street, block 8, 1113 Sofia, Bulgaria*
[c] *LACO UMR 6090, Université de Limoges 12 Avenue A Thomas, 87060 Limoges, France*

Received 30 October 2001; revised 8 July 2002; accepted 12 October 2002

Abstract

This paper proposes a regularized notion of a composition of a monotone operator with a linear mapping. This new concept, called variational composition, can be shown to be maximal monotone in many cases where the usual composition is not. The two notions coincide, however, whenever the latter is maximal monotone. The utility of the variational composition is demonstrated by applications to subdifferential calculus, theory of measurable multifunctions, and elliptic PDEs with singular coefficients.

Keywords: Maximal monotone operator; Composition; Graphical convergence; Subdifferential; Measurable multifunction; Elliptic PDE

*Corresponding author.
E-mail addresses: pennanen@hkkk.fr (T. Pennanen), revalski@math.bas.bg (J.P. Revalski), michel.thera@unilim.fr (M. Théra).

[1] Partially supported by LACO (Laboratoire d'Arithmétique, Calcul Formel et Optimisation), UMR-CNRS 6090 of the University of Limoges, as well as by the Région Limousin under a Research Grant.
[2] Partially supported by the LACO (Laboratoire d'Arithmétique, Calcul Formel et Optimisation), UMR-CNRS 6090 of the University of Limoges, by Bulgarian NFSR under Grant MM-1105/01 and by NATO-CLG 978488.
[3] Research of Michel Théra has been partially supported by the French Chilean Scientific Cooperation Programme ECOS-SUD under Grant C00E05 and by NATO-CLG 978488.

PII: S0022-1236(02)00104-0

1. Introduction

Throughout this paper, U and X will be real reflexive Banach spaces and U^* and X^* their duals, unless otherwise specified. Recall that a set-valued mapping $T : U \rightrightarrows U^*$ is called *monotone* if

$$u_1^* \in T(u_1),\ u_2^* \in T(u_2) \Rightarrow \langle u_1 - u_2, u_1^* - u_2^* \rangle \geqslant 0,$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing between U and U^*. If a monotone mapping cannot be properly extended to another monotone mapping from U to U^*, it is called *maximal monotone*. An important example is the subdifferential

$$\partial f(u) = \{u^* \in U^* \mid f(v) \geqslant f(u) + \langle v - u, u^* \rangle \ \forall v \in U\}, \quad u \in U,$$

of a convex function $f : U \to \mathbb{R} \cup \{+\infty\}$. It has been shown by Rockafellar [30], in the Banach space setting, that $\partial f : U \rightrightarrows U^*$ is maximal monotone, whenever f is proper and lower semicontinuous, that is, when the *epigraph* $\operatorname{epi} f = \{(u, \alpha) \in U \times \mathbb{R} \mid f(u) \leqslant \alpha\}$ is nonempty and closed.

Let $A : X \to U$ be linear and continuous with adjoint $A^* : U^* \to X^*$. It is easily checked that the composite mapping $A^*TA : X \rightrightarrows X^*$, given by $A^*TA(x) := \cup\{A^*u^* \mid u^* \in T(Ax)\}$, is monotone. This kind of operators appear, for example, in partial differential equations in divergence form, and they also contain the pointwise sum of two or more operators as a special case. Without further conditions, however, A^*TA may fail to be maximal monotone; see [22,28,35] for sufficient conditions. It is then a natural idea to try to approximate A^*TA by a mapping which is guaranteed to be maximal monotone. A good candidate is $A^*T_\lambda A$, where T_λ is the Yosida regularization of T with parameter $\lambda > 0$. Indeed, since (after renorming of the space, if necessary) T_λ is a monotone continuous mapping, the same is then true of $A^*T_\lambda A$, which guarantees the maximality. If one now takes the limit of $A^*T_\lambda A$ as $\lambda \searrow 0$, in the sense of graphical convergence, it turns out that one obtains a mapping that is more likely to be maximal monotone than the pointwise composition A^*TA. This limit mapping, denoted here $(A^*TA)_v$ (to be given a precise definition in the next section) is what we call the *variational composition* of A and T. The purpose of this paper is to study the relation between A^*TA and $(A^*TA)_v$, to give sufficient conditions for maximality of $(A^*TA)_v$, and to give applications of this new concept.

Variational composition is a natural extension of the idea presented in Attouch et al. [3], where the notion of a *variational sum* was introduced. Their motivation was to define a new notion of a sum of two mappings, that is more likely to be maximal monotone than the usual pointwise sum. They studied the general properties of variational sums and showed how they arise quite naturally in practice. More applications and further study of this concept can be found in [4,14,26]. Note that if T_1 and T_2 are set-valued mappings from X to X^*, their pointwise sum can be expressed in the composite form A^*TA, by defining $U = X \times X$, $Ax = (x, x)$, and $T(x_1, x_2) = T_1(x_1) \times T_2(x_2)$. Indeed, then $A^*(x_1^*, x_2^*) = x_1^* + x_2^*$, and so $A^*TA(x) =$

$T_1(x) + T_2(x)$. This fact will allow us to draw connections between the variational composition and the variational sum.

The applications of the variational composition are similar to those of the variational sum. Whereas the variational sum gave an expression for the subdifferential of the sum of two convex functions [3, Theorem 7.2], the variational composition gives us a formula for the subdifferential of the composition $f \circ A$ of a linear continuous mapping $A : X \to U$ and a convex function f on U (Theorem 4.1). Much as [3, Theorem 7.2] was used to study the Schrödinger equation with singular potentials, we use our Theorem 4.1 to study elliptic PDEs with singular coefficients. In particular, we obtain an existence result for linear elliptic PDEs in divergence form in the case of locally integrable (instead of the usual essentially bounded) coefficients.

The precise definition of the variational composition will be given in the next section, after recalling some basic facts about monotone operators and their graphical convergence. In Section 3, we will study the relation between the pointwise and the variational compositions. Section 4 studies the special case of subdifferential mappings, and we obtain a new expression for the subdifferential of the composition of a convex function with a linear mapping. The last two sections are devoted to applications. In Section 5, we derive conditions for measurability of a family of composite mappings, and in Section 6, we use the variational composition to compute the subdifferential of an energy function associated with a partial differential equation with singular coefficients.

2. Preliminaries

We begin with some notations and basic facts about monotone operators. For more comprehensive introduction to the subject, see for example [11,24,38] or [35, Chapter 12]. The inverse $T^{-1} : U^* \rightrightarrows U$ of a set-valued mapping $T : U \rightrightarrows U^*$ is given by $T^{-1}(u^*) = \{u \in U \mid u^* \in T(u)\}$. The *graph* of T is the set $\operatorname{gph} T = \{(u, u^*) \in U \times U^* \mid u^* \in T(u)\}$, and the *domain* $\operatorname{dom} T$ and the *range* $\operatorname{rge} T$ of T are defined as the projections of $\operatorname{gph} T$ to U and U^*, respectively.

For simplicity of notation, the norms on U and U^* will both be denoted by $\|\cdot\|$. It will be clear from the context which norm is meant. The *duality mapping* is $J_U : U \rightrightarrows U^*$, defined by $J_U = \partial\phi$, where $\phi(u) = \frac{1}{2}\|u\|^2$, $u \in U$. This is a maximal monotone mapping with $\operatorname{dom} J_U = U$, and it can be expressed as

$$J_U(u) = \{u^* \in U^* \mid \langle u, u^* \rangle = \|u\|^2 = \|u^*\|^2\}, \quad u \in U.$$

Furthermore, we have $J_U^{-1} = J_{U^*}$, the duality mapping on U^* associated with the dual norm. Due to a well-known renorming result of Troyanski (see e.g. [15]) we can (and will) assume that the norms on U and U^* are locally uniformly rotund. This implies that these norms satisfy the *Kadec–Klee property*:

$$u_n \to u \text{ weakly and } ||u_n|| \to ||u|| \text{ imply } u_n \to u \text{ strongly},$$

and then the duality mappings J_U and J_U^{-1} are single valued and norm-to-norm continuous.

As usual, the following Minty–Rockafellar criterion for maximal monotonicity will be crucial; see [29, Proposition 1].

Theorem 2.1. *A monotone mapping $T : U \rightrightarrows U^*$ is maximal if and only if for every $\lambda > 0$, $\operatorname{rge}(T + \lambda J_U) = U^*$. In this case, the inverse $(T + \lambda J_U)^{-1}$ is a single-valued maximal monotone operator which is norm to weak continuous.*

It follows from this and the properties of the chosen norms that, if T is maximal monotone, then for any $\lambda > 0$, the *Yosida regularization*

$$T_\lambda = (T^{-1} + \lambda J_U^{-1})^{-1}$$

of T is single valued, strongly continuous and maximal monotone with $\operatorname{dom} T_\lambda = U$; see for example [2, Proposition 3.56]. The following is well known (see for example [5, p. 63]), but for the convenience of the reader, we provide the simple proof.

Corollary 2.1. *Let T be maximal monotone.*

(a) *We have $u^* \in \operatorname{rge} T$ if and only if the family $\{u_\lambda \mid \lambda > 0\}$ of solutions to*

$$T(u) + \lambda J_U(u) \ni u^*$$

remains bounded as $\lambda \searrow 0$. When this happens, $||u_\lambda|| \leqslant ||\bar{u}||$ for all $\lambda > 0$, where $\bar{u}$ is the minimum norm solution of $T(u) \ni u^*$, and as $\lambda \searrow 0$, u_λ converges strongly to $\bar{u}$.

(b) *We have $u \in \operatorname{dom} T$ if and only if the family $\{T_\lambda(u) \mid \lambda > 0\}$ remains bounded as $\lambda \searrow 0$. When this happens, $||T_\lambda(u)|| \leqslant ||\bar{u}^*||$ for all $\lambda > 0$, where $\bar{u}^*$ is the minimum norm solution of $T(u) \ni u^*$, and as $\lambda \searrow 0$, $T_\lambda(u)$ converges strongly to $\bar{u}^*$.*

Proof. Part (a): By Theorem 2.1, the point u_λ is uniquely defined for every $\lambda > 0$. If $\{u_\lambda \mid \lambda > 0\}$ is bounded, it has a weak cluster point $\bar{u}$, and $\lambda J_U(u_\lambda) \to 0$ strongly. Since $u^* - \lambda J_U(u_\lambda) \in T(u_\lambda)$, we must have $u^* \in T(\bar{u})$, by the maximal monotonicity of T. This proves the "if" part. Now let $\bar{u}$ be the minimum norm element of $T^{-1}(u^*)$, which exists and is unique, since $T^{-1}(u^*)$ is closed and convex by the maximal monotonicity of T. Then, by monotonicity of T,

$$0 \leqslant \langle u_\lambda - \bar{u}, u^* - J_U(u_\lambda) - u^* \rangle \leqslant -||u_\lambda||^2 + ||\bar{u}||\,||u_\lambda||,$$

which implies $||u_\lambda|| \leqslant ||\bar{u}||$, proving the "only if" part. Combining the above arguments, we see that, whenever $\{u_\lambda\}$ is bounded, it satisfies $||u_\lambda|| \leqslant ||\bar{u}||$ and all

its weak cluster points are in $T^{-1}(u^*)$. From this it follows that the whole sequence $\{u_\lambda\}$ must converge weakly to $\bar{u}$, and by the Kadec–Klee property, the convergence is strong. Part (b) follows by applying (a) to T^{-1}. $\square$

In order to give the precise definition of variational composition, we need to recall the notion of *graphical convergence* of a family $\{C_\lambda : U \rightrightarrows U^*\}_{\lambda>0}$ of operators. The idea of graphical convergence is to identify the operators with their graphs and to consider Painlevé–Kuratowski convergence on them. We will denote by $\text{g-lim inf}_{\lambda\searrow 0} C_\lambda$ the mapping whose graph is the set of points (u, u^*) such that for every sequence $\lambda_n \searrow 0$ there is a sequence $(u_n, u_n^*) \to (u, u^*)$ with $u_n^* \in C_{\lambda_n}(u_n)$. Similarly, $\text{g-lim sup}_{\lambda\searrow 0} C_\lambda$ is the mapping whose graph is the set of points (u, u^*) such that there exist sequences $\lambda_n \searrow 0$ and $(u_n, u_n^*) \to (u, u^*)$ with $u_n^* \in C_{\lambda_n}(u_n)$. If $\text{g-lim inf}_{\lambda\searrow 0} C_\lambda = \text{g-lim sup}_{\lambda\searrow 0} C_\lambda$, one says that the family $\{C_\lambda\}_{\lambda>0}$ *graph-converges* to the common limit which is denoted by $\text{g-lim}_{\lambda\searrow 0} C_\lambda$. For reference on the general theory of convergence of sets and graph-convergence of operators, see for example [2,9,35].

We will need the following facts from Attouch [2, Chapter 3].

Theorem 2.2. *Let $\{C_\lambda\}_{\lambda>0}$ and C be maximal monotone mappings. Then*

(a) $\text{g-lim inf}_{\lambda\searrow 0} C_\lambda$ *is monotone*;
(b) $\text{g-lim}_{\lambda\searrow 0} C_\lambda = C$ *if and only if* $\text{g-lim inf}_{\lambda\searrow 0} C_\lambda \supset C$;
(c) $\text{g-lim}_{\lambda\searrow 0} C_\lambda = C$ *if and only if*

$$\lim_{\lambda\searrow 0} (C_\lambda + J_U)^{-1}(u^*) = (C + J_U)^{-1}(u^*) \quad \forall u^* \in U^*.$$

Now, let X, X^* be another dual pair of reflexive Banach spaces, endowed with locally uniformly rotund norms, and let $A : X \to U$ be linear and continuous. Since the Yosida regularization $T_\lambda : U \rightrightarrows U^*$ of a maximal monotone T is single valued and continuous for every $\lambda > 0$, so is the composition $A^*T_\lambda A : X \rightrightarrows X^*$. From the monotonicity of T_λ it then follows that $A^*T_\lambda A$ is maximal monotone for every $\lambda > 0$; see for example [38]. This, and the fact that $\text{g-lim}_{\lambda\searrow 0} T_\lambda = T$, suggest the following.

Definition 2.1. Let $A : X \to U$ be continuous and linear, and let $T : U \rightrightarrows U^*$ be maximal monotone. The *variational composition* $(A^*TA)_v : X \rightrightarrows X^*$ of A and T is the set-valued mapping

$$(A^*TA)_v = \text{g-}\liminf_{\lambda\searrow 0} A^*T_\lambda A.$$

By Theorem 2.2(a), $(A^*TA)_v$ is monotone, and by (b),

$$(A^*TA)_v = \text{g-}\lim_{\lambda \searrow 0} A^*T_\lambda A,$$

whenever $(A^*TA)_v$ is maximal monotone.

The idea of replacing A^*TA by $A^*T_\lambda A$, and taking the limit as $\lambda \searrow 0$, has been already used (in the finite-dimensional setting) in the proof of [35, Theorem 12.43], where a sufficient constraint qualification condition was found in order to assure that the family $\{A^*T_\lambda A\}$ graph-converges to A^*TA (which in this case guarantees the maximality of A^*TA).

The variational composition is closely related to the *variational sum* of two monotone mappings T^1 and T^2 from X to X^* defined in [3]:

$$(T^1 \underset{v}{+} T^2) := \text{g-}\liminf_{\lambda,\mu \searrow 0,\ \lambda\mu \neq 0} (T^1_\lambda + T^2_\mu). \tag{1}$$

If in Definition 2.1, we let $U = X \times X$, $Ax = (x, x)$, and $T(x_1, x_2) = T_1(x_1) \times T_2(x_2)$, we obtain

$$(A^*TA)_v = \text{g-}\liminf_{\lambda \searrow 0} (T^1_\lambda + T^2_\lambda),$$

so that $\text{gph}(T^1 +_v T^2) \subset \text{gph}(A^*TA)_v$. Thus, $(A^*TA)_v$ equals $T^1 +_v T^2$, whenever the latter is maximal monotone (which is the interesting case).

3. Comparison of the pointwise and the variational composition

The following simple inequality turns out to be useful in comparing A^*TA and $(A^*TA)_v$.

Lemma 3.1. *If T is monotone, then $u^* \in T(u)$ and $v^* = T_\lambda(v)$ imply*

$$\langle u - v, u^* - v^* \rangle \geqslant -\frac{\lambda}{4} ||u^*||^2.$$

Proof. Since $v^* = T_\lambda(v)$ means that $v - \lambda J_U^{-1}(v^*) \in T^{-1}(v^*)$, the monotonicity of T implies $\langle u - v + \lambda J_U^{-1}(v^*), u^* - v^* \rangle \geqslant 0$, and so,

$$\begin{aligned}\langle u-v, u^*-v^*\rangle &\geqslant \lambda\langle J_U^{-1}(v^*), v^*-u^*\rangle \\ &\geqslant \lambda(\|v^*\|^2 - \|u^*\|\,\|v^*\|) \\ &\geqslant \lambda \min_{\alpha\in\mathbb{R}} \{\alpha^2 - \|u^*\|\alpha\} = -\lambda\frac{\|u^*\|^2}{4}. \qquad \square\end{aligned}$$

In general, we cannot guarantee that $\operatorname{gph}(A^*TA)\subset\operatorname{gph}(A^*TA)_v$, but the following is true.

Proposition 3.1. *Let $A: X\to U$ be continuous and linear, and let $T: U\rightrightarrows U^*$ be maximal monotone. Then* $\operatorname{dom}(A^*TA)\subset\operatorname{dom}(A^*TA)_v$, *and if $(A^*TA)_v$ is maximal monotone, then* $\operatorname{gph}(A^*TA)\subset\operatorname{gph}(A^*TA)_v$.

Proof. If $x_0\in\operatorname{dom}(A^*TA)$, then $Ax_0\in\operatorname{dom}T$, so by Corollary 2.1(b), $T_\lambda(Ax_0)$ converges strongly to the minimum norm element of $T(Ax_0)$, say u_0^*. Thus, by continuity of A^*, $(A^*T_\lambda A)(x_0)$ converges strongly to $A^*u_0^*$. Then, by definition, $A^*u_0^*\in(A^*TA)_v(x_0)$, so $x_0\in\operatorname{dom}(A^*TA)_v$.

To prove the second part, let $\lambda>0$, $(x,x^*)\in\operatorname{gph}(A^*TA)$ and $(x_\lambda,x_\lambda^*)\in\operatorname{gph}(A^*T_\lambda A)$ be arbitrary, and let $u^*\in T(Ax)$ and $u_\lambda^*\in T_\lambda(Ax_\lambda)$ be such that $x^*=A^*u^*$ and $x_\lambda^*=A^*u_\lambda^*$. Then by Lemma 3.1,

$$\begin{aligned}\langle x-x_\lambda, x^*-x_\lambda^*\rangle &= \langle x-x_\lambda, A^*u^*-A^*u_\lambda^*\rangle \\ &= \langle Ax-Ax_\lambda, u^*-u_\lambda^*\rangle \geqslant -\frac{\lambda}{4}\|u^*\|^2.\end{aligned}$$

Since any point $(\tilde{x},\tilde{x}^*)\in\operatorname{gph}(A^*TA)_v$ can be written as a limit of (x_λ,x_λ^*) as $\lambda\searrow 0$, we must have

$$\langle x-\tilde{x}, x^*-\tilde{x}^*\rangle\geqslant 0 \quad \forall(\tilde{x},\tilde{x}^*)\in\operatorname{gph}(A^*TA)_v.$$

Since $(x,x^*)\in\operatorname{gph}(A^*TA)$ was arbitrary, this implies

$$\operatorname{gph}(A^*TA)\subset\operatorname{gph}(A^*TA)_v,$$

when $(A^*TA)_v$ is maximal monotone. $\square$

We next consider a particular case where $\operatorname{gph}(A^*TA)\subset\operatorname{gph}(A^*TA)_v$ does hold, and the variational composition can be seen to have a regularizing property. Our approach is obtained by modifying the one used in [3,26].

Given a set-valued mapping $S: U\rightrightarrows U^*$, we define $\bar{S}: U\rightrightarrows U^*$ by $\operatorname{gph}\bar{S}=\operatorname{cl}(\operatorname{gph}S)$. Obviously, if S is monotone, the same is true of $\bar{S}$.

Theorem 3.1. *If the mapping $\overline{A^*TA}$ is maximal monotone, then*

$$(A^*TA)_v = \overline{A^*TA}.$$

Proof. Let $y^* \in X^*$ be arbitrary. For $\lambda > 0$ denote by x_λ the unique solution of the equation

$$J_X(x) + (A^*T_\lambda A)(x) = y^*.$$

By Theorem 2.2(c), it suffices to show that as $\lambda \searrow 0$, x_λ converges strongly to the unique solution of

$$J_X(x) + (\overline{A^*TA})(x) \ni y^*.$$

Let $(x, x^*) \in \mathrm{gph}(A^*TA)$ be arbitrary, and let $u^* \in T(Ax)$ be such that $x^* = A^*u^*$. By Lemma 3.1,

$$-\frac{\lambda}{4}||u^*||^2 \leqslant \langle Ax - Ax_\lambda, u^* - T_\lambda(Ax_\lambda)\rangle = \langle x - x_\lambda, A^*u^* - (A^*T_\lambda A)(x_\lambda)\rangle,$$

so by the definition of x_λ,

$$\langle x - x_\lambda, x^* + J_X(x_\lambda) - y^*\rangle \geqslant -\frac{\lambda}{4}||u^*||^2. \tag{2}$$

This implies in particular that

$$-||x_\lambda||^2 - (||x|| + ||x^* - y^*||)||x_\lambda|| \geqslant -\langle x, x^* - y^*\rangle - \frac{\lambda}{4}||u^*||^2,$$

so $\{x_\lambda\}$ must be bounded, and it has a weak cluster point $\bar{x}$.

By monotonicity of J_X, $\langle x - x_\lambda, J_X(x)\rangle \geqslant \langle x - x_\lambda, J_X(x_\lambda)\rangle$, so (2) gives

$$\langle x - x_\lambda, x^* + J_X(x) - y^*\rangle \geqslant -\frac{\lambda}{4}||u^*||^2.$$

Passing to the limit,

$$\langle x - \bar{x}, x^* + J_X(x) - y^*\rangle \geqslant 0.$$

Since $(x, x^*) \in \mathrm{gph}(A^*TA)$ was arbitrary, this implies

$$\langle x - \bar{x}, x^* + J_X(x) - y^*\rangle \geqslant 0 \quad \forall (x, x^*) \in \mathrm{gph}\,\overline{(A^*TA)}.$$

Because $\overline{(A^*TA)}$ is maximal monotone, the same is true of $J_X + \overline{(A^*TA)}$, so we must have $(\bar{x}, y^*) \in \mathrm{gph}(J_X + \overline{(A^*TA)})$, or in other words,

$$J_X(\bar{x}) + \overline{(A^*TA)}(\bar{x}) \ni y^*. \tag{3}$$

Since the latter inclusion determines the point $\bar{x}$ uniquely, the whole family $\{x_\lambda\}$ must converge weakly to $\bar{x}$.

Going back to (2), and using the inequality $\langle J_X(x_\lambda), x\rangle \leqslant \frac{1}{2}||x||^2 + \frac{1}{2}||x_\lambda||^2$, we get

$$\frac{1}{2}||x_\lambda||^2 \leqslant \frac{1}{2}||x||^2 + \langle x - x_\lambda, x^* - y^*\rangle + \frac{\lambda}{4}||u^*||^2,$$

from which

$$\limsup_{\lambda \searrow 0} ||x_\lambda||^2 \leqslant ||x||^2 + 2\langle x - \bar{x}, x^* - y^*\rangle.$$

Since $(x, x^*) \in \operatorname{gph}(A^*TA)$ was arbitrary, and since by (3), $(\bar{x}, y^* - J_X(\bar{x})) \in \operatorname{cl}\operatorname{gph}(A^*TA)$, we must have

$$\limsup_{\lambda \searrow 0} ||x_\lambda||^2 \leqslant ||\bar{x}||^2,$$

so by the Kadec–Klee property, $x_\lambda \to \bar{x}$ strongly. □

The following immediate consequence can be viewed as a consistency result for $(A^*TA)_v$.

Corollary 3.1. *If the mapping A^*TA is maximal monotone, then*

$$(A^*TA)_v = A^*TA.$$

Sufficient conditions for maximal monotonicity of the pointwise composition A^*TA have been given in [22,28,33]. In particular, A^*TA is maximal monotone whenever $0 \in \operatorname{ri}(\operatorname{rge} A - \operatorname{dom} T)$ [22, Corollary 4.4]. Here "ri" means the *relative interior* of a set.

For any number m of monotone mappings $T^1, \ldots, T^m$ from X to X^*, one could define a "variational sum" by $\text{g-lim inf}_{\lambda \searrow 0}\,(T^1_\lambda + \cdots + T^m_\lambda)$.

Corollary 3.2. *Let $T^1, \ldots, T^m$ be maximal monotone mappings from X to X^*. If the mapping $\overline{T^1 + \cdots + T^m}$ is maximal monotone, then*

$$\text{g-}\lim_{\lambda \searrow 0}\,(T^1_\lambda + \cdots + T^m_\lambda) = \overline{T^1 + \cdots + T^m}.$$

Proof. Let U be the space $X \times \cdots \times X$ equipped with the norm $||(x_1, \ldots, x_m)||_U^2 = ||x_1||_X^2 + \cdots + ||x_m||_X^2$, so that $J_U = (J_X, \ldots, J_X)$. If we define $T(x_1, \ldots, x_m) = T^1(x_1) \times \cdots \times T^m(x_m)$, and $Ax = (x, \ldots, x)$, then $T_\lambda = T^1_\lambda \times \cdots \times T^m_\lambda$, and

$$A^*TA = T^1 + \cdots + T^m,$$

$$A^*T_\lambda A = T^1_\lambda + \cdots + T^m_\lambda.$$

The result thus follows from Theorem 3.1. □

Corollary 3.2 is reminiscent of [3, Theorem 6.1], which states that the variational sum of two maximal monotone mappings equals the closure of their pointwise sum, whenever the latter is maximal monotone.

4. A subdifferential chain rule without constraint qualifications

If $f : U \to \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, and $A : X \to U$ is continuous and linear, then the composition $f \circ A$ is also convex and lower semicontinuous. Furthermore, by the chain rule of convex analysis,

$$\partial(f \circ A) \supset A^* \partial f A,$$

where equality holds whenever the "constraint qualification" $0 \in \operatorname{int}(\operatorname{rge} A - \operatorname{dom} f)$ is satisfied [32] (here $\operatorname{dom} f = \{u \in U \mid f(u) < +\infty\}$ as usual). Without the constraint qualification, however, the inclusion may be strict. The purpose of this section is to give a more general formula for $\partial(f \circ A)$ in terms of the variational composition.

For $\lambda > 0$, the *Moreau–Yosida regularization* f_λ of f is the function defined by

$$f_\lambda(u) = \inf_{v \in U} \left\{ f(v) + \frac{1}{2\lambda} ||v - u||^2 \right\}.$$

It is well known (see for example [8]) that f_λ is a convex C^1-function on U, with

$$\nabla f_\lambda = (\partial f)_\lambda.$$

Recall that a sequence $\{f_n\}_{n=1}^\infty$ of proper lower semicontinuous convex functions is said to *Mosco-converge* [19] to f, denoted by $f_n \to^M f$, if for every $u \in U$ the following two conditions are fulfilled:

(i) if $u_n \to u$ weakly, then $f(u) \leqslant \liminf_{n \to \infty} f_n(u_n)$;
(ii) there is a strongly converging sequence $u_n \to u$, with

$$\limsup_{n \to \infty} f_n(u_n) \leqslant f(u).$$

By the well-known result of Attouch [2, Theorem 3.66], we have $f_n \to^M f$ if and only if g-lim $\partial f_n = \partial f$ and a certain normalization condition holds.

Theorem 4.1. *Let $A : X \to U$ be linear and continuous, and let f be a proper lower semicontinuous convex function on U. If* $\operatorname{dom}(f \circ A) \neq \emptyset$, *then*

$$\partial(f \circ A) = (A^* \partial f A)_v.$$

Proof. By definition,

$$\begin{aligned}(A^* \partial f A)_v &= \text{g-lim inf}(A^*(\partial f)_\lambda A) \\ &= \text{g-lim inf}(A^* \nabla f_\lambda A) = \text{g-lim inf}(\nabla(f_\lambda \circ A)),\end{aligned}$$

where the last equality follows from the chain rule which applies by continuity of f_λ. As $\lambda \searrow 0$, the functions $f_\lambda \circ A$ monotonically increase to $f \circ A$, which by Attouch [2, Theorem 2.40] implies $f_\lambda \circ A \to^M f \circ A$. Then by the Attouch criterion,

$$\text{g-lim } \nabla(f_\lambda \circ A) = \partial(f \circ A).$$

This completes the proof. □

Theorem 4.1 gives an exact expression for $\partial(f \circ A)$, but it may be harder to evaluate than the pointwise composition $A^* \partial f A$. In Section 6, we give an example of a problem for which the constraint qualification $0 \in \operatorname{int}(\operatorname{rge} A - \operatorname{dom} T)$ fails, but where the variational composition can be computed. Theorem 4.1 resembles the results in [16–18,23,26,27,36,37], where subdifferential rules without constraint qualifications were given, e.g. in terms of limits of epsilon-subdifferentials and epsilon enlargements of subdifferentials.

5. Measurability of composite mappings

In this section, we will use the variational composition to study measurability properties of parameterized families of composite mappings. Throughout the section, Ω denotes a measurable space, and all the other spaces are separable Hilbert spaces.

Given a family of set-valued mappings $\{T(\omega) : H \rightrightarrows H\}_{\omega \in \Omega}$, define the mapping $\mathscr{L}_2[T] : L^2(\Omega; H) \rightrightarrows L^2(\Omega; H)$ (the canonical extension of T) by

$$\mathscr{L}_2[T](v) = \{v^* \in L^2(\Omega; H) \mid v^*(\omega) \in T(\omega)(v(\omega)) \text{ a.e. on } \Omega\}.$$

In this context, the measurability properties of the mapping $\omega \mapsto \operatorname{gph} T(\omega)$ are crucial.

Definition 5.1. A set-valued mapping $S : \Omega \rightrightarrows H$ is *measurable* if for any open $C \subset H$, the set

$$S^{-1}(C) = \{\omega \in \Omega \mid S(\omega) \cap C \neq \emptyset\}$$

is measurable. A family of set-valued mappings $\{T(\omega) : H \rightrightarrows H\}_{\omega \in \Omega}$ is *measurable* if the set-valued mapping $\omega \mapsto \operatorname{gph} T(\omega)$ is measurable.

Measurability of set-valued mappings has been studied extensively by many authors; see for example [1,12,33,35, Chapter 14]. It is particularly important when studying monotone mappings. If $T(\omega)$ is monotone a.e. on Ω, $\mathscr{L}_2[T]$ is monotone. The following result (see for example [11, Example 2.3.3]) gives a simple condition for maximality.

Theorem 5.1. *Let $\{T(\omega)\}_{\omega \in \Omega}$ be a measurable family of maximal monotone mappings on H. If* $\operatorname{dom} \mathscr{L}_2[T] \neq \emptyset$ *then $\mathscr{L}_2[T]$ is maximal monotone.*

The above result is closely related to the theory of convex normal integrands [31]. A function f on $\Omega \times H$ is said to be a *convex normal integrand* if the mapping $\omega \mapsto \operatorname{epi} f(\omega, \cdot)$ is measurable with closed and convex values. If f is a convex normal integrand, then the integral functional

$$I_f(u) = \begin{cases} \int_\Omega f(\omega, u(\omega))\, d\omega & \text{if } f(\cdot, u(\cdot)) \in L^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

is a convex and lower semicontinuous function on $L^2(\Omega; H)$. By Attouch [1, Theorem 2.3], f is a convex normal integrand if and only if $\{\partial f(\omega, \cdot)\}_{\omega \in \Omega}$ is a measurable family of maximal monotone mappings on H, and there is a measurable function $u : \Omega \mapsto H$ such that $f(\cdot, u(\cdot))$ is measurable. The formula

$$\partial I_f = \mathscr{L}_2[\partial f]$$

is valid for any convex normal integrand provided $\operatorname{dom} \mathscr{L}_2[\partial f] \neq \emptyset$ [33]. This can also be seen as a consequence of Theorem 5.1 and the easily verified fact that $\partial I_f \supset \mathscr{L}_2[\partial f]$.

It is clear that $\mathscr{L}_2[T]^{-1} = \mathscr{L}_2[T^{-1}]$, where $T^{-1}(\omega) = T(\omega)^{-1}$, and that

$$\mathscr{L}_2[S] + \mathscr{L}_2[T] \subset \mathscr{L}_2[S + T], \tag{4}$$

where $(S + T)(\omega) = S(\omega) + T(\omega)$. Equality holds in (4), if $T(\omega)$ and $S(\omega)$ are monotone and $\mathscr{L}_2[S] + \mathscr{L}_2[T]$ is maximal monotone. Since the identity mapping on $L^2(\Omega; H)$ can be written as $J_{L^2(\Omega;H)} = \mathscr{L}_2[J_H]$, we have in particular that

$$\mathscr{L}_2[T]_\lambda = \mathscr{L}_2[T_\lambda],$$

where $T_\lambda(\omega) = T(\omega)_\lambda$, provided $\mathscr{L}_2[T]$ is maximal monotone (see Theorem 5.1).

The following can be found in [1].

Theorem 5.2. *Let (Ω, μ) be a positive σ-finite complete measure space, and consider a family $\{T(\omega)\}_{\omega\in\Omega}$ of maximal monotone mappings in H. The following are equivalent:*

(a) *$\{T(\omega)\}_{\omega\in\Omega}$ is measurable;*
(b) *$T(\cdot)_\lambda(v)$ is measurable for every $v\in H$ and $\lambda > 0$;*
(c) *There are measurable families $\{T_n(\omega)\}_{\omega\in\Omega}$, $n = 1, 2, \ldots$, of maximal monotone mappings such that*

$$T(\omega) = \operatorname*{g-lim}_{n\to\infty} T_n(\omega) \quad a.e.\ on\ \Omega.$$

These conditions hold if, in particular,

(d) *$T(\cdot)(v)$ is measurable for each $v\in H$, and $T(\omega)(\cdot)$ is continuous for each $\omega\in\Omega$ (such a T is called a Carathéodory mapping).*

Combining the measurability criteria of Theorem 5.2 with the general properties of variational composition, we obtain the following.

Theorem 5.3. *Let (Ω, μ) be a positive σ-finite complete measure space, let $\{T(\omega)\}_{\omega\in\Omega}$ be a measurable family of maximal monotone mappings on U, and let A be a Carathéodory mapping with $A(\omega)(\cdot) : X \to U$ linear for every $\omega\in\Omega$.*

(a) *If the mapping $\overline{A(\omega)^* T(\omega) A(\omega)}$ is maximal monotone a.e. on Ω, then*

$$\{\overline{A(\omega)^* T(\omega) A(\omega)}\}_{\omega\in\Omega}$$

is a measurable family.

(b) *If $T(\omega) = \partial f(\omega, \cdot)$ for a convex normal integrand f, that satisfies $\operatorname{dom} f(\omega, \cdot)\circ A(\omega) \neq \emptyset$ a.e. on Ω, then*

$$\{\partial(f(\omega, \cdot)\circ A(\omega))\}_{\omega\in\Omega}$$

is a measurable family, and $f(\omega, \cdot)\circ A(\omega)$ is a convex normal integrand.

Proof. By Theorem 5.2, the mapping T_λ is Carathéodory for every $\lambda > 0$. Since A is Carathéodory, the same is then true of A^*, and hence, of the family $\{A(\omega)^* T(\omega)_\lambda A(\omega)\}_{\omega\in\Omega}$. Part (a) now follows by using criteria (d) and (c) of Theorem 5.2 with Theorem 3.1. The first half of part (b) follows similarly through Theorem 4.1. The claim that $f(\omega, \cdot)\circ A(\omega)$ is a normal convex integrand, follows from [1, Theorem 2.3], since $\omega \mapsto f(\omega, A(\omega)x(\omega))$ is measurable for any measurable x, by the fact that f is a convex normal integrand. □

In the case $U = X \times X$, $T(\omega)(x_1, x_2) = T_1(\omega)(x_1) \times T_2(\omega)(x_2)$, and $A(\omega)(x) = (x, x)$, we recover the following result due to Attouch.

Corollary 5.1 (Attouch [1, Theorem 2.4]). *Let (Ω, μ) be a positive σ-finite complete measure space, and let $T_1(\omega)$ and $T_2(\omega)$ be measurable families of maximal monotone mappings on U. If for every ω the mapping $T_1(\omega) + T_2(\omega)$ is maximal monotone, then the family $\{T_1(\omega) + T_2(\omega)\}_{\omega \in \Omega}$ is measurable.*

It is interesting to note that the original proof of [1, Theorem 2.4] was based on properties that are characteristic of the variational sum introduced later in [3]. The proof of Theorem 5.3 is a natural extension of this approach.

6. Elliptic PDEs with singular coefficients

It was demonstrated in [3], with an example from quantum mechanics, how the variational sum can be useful in identifying the subdifferential of the sum of two discontinuous convex functions. Similarly, the expression

$$\partial(f \circ A) = (A^* \partial f A)_v$$

from Theorem 4.1 can be used to find $\partial(f \circ A)$, in cases where the chain rule $\partial(f \circ A) \supset A^* \partial f A$ fails to hold as an equality. The purpose of this section is to derive an expression for the subdifferential of a discontinuous "energy functional" through the computation of a variational composition.

Let $\Omega \subset \mathbb{R}^N$ be open and let $Q : \Omega \mapsto \mathbb{R}^{N \times N}$ be measurable with $Q(x)$ symmetric and positive semidefinite a.e. on Ω. Consider the function $g : H_0^1(\Omega) \to \mathbb{R} \cup \{+\infty\}$ defined by

$$g(u) = \begin{cases} \frac{1}{2} \int_\Omega \nabla u(x) \cdot Q(x) \nabla u(x)\, dx & \text{if } \nabla u \cdot Q \nabla u \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

Such functions arise frequently, e.g. in physics, and the fundamental problem is to minimize $g - \langle \cdot, u^* \rangle$ for some $u^* \in H_0^1(\Omega)^*$ over $u \in H_0^1(\Omega)$, or equivalently, to solve the inclusion $\partial g(u) \ni u^*$. It is often useful to have an explicit expression for ∂g.

Note that we can express g in the composite form $g = I_f \circ \nabla$, with the continuous linear map $\nabla : H_0^1(\Omega) \to L^2(\Omega; \mathbb{R}^N)$, $\nabla u = \{\frac{\partial u}{\partial x_i}\}_{i=1}^N$, and the convex normal integrand $f(x, v) = \frac{1}{2} v \cdot Q(x) v$. It follows that g is convex and lower semicontinuous since I_f is such. Also, since g is quadratic, $\operatorname{dom} g$ is a linear space. Because $\partial f(\omega, \cdot) = Q(\omega)$, we have $0 \in \mathscr{L}_2[\partial f](0)$ and $\partial I_f = \mathscr{L}_2[\partial f]$.

In cases where the constraint qualification $0 \in \operatorname{int}(\operatorname{rge} \nabla - \operatorname{dom} I_f)$ holds, e.g. when $Q(x) \in L^\infty(\Omega; \mathbb{R}^{N \times N})$ so that $\operatorname{dom} I_f = L^2(\Omega; \mathbb{R}^N)$, the classical chain rule gives the simple formula $\partial g = \nabla^* \mathscr{L}_2[Q] \nabla$, where $\nabla^* = -\operatorname{div}$ (the divergence), that is,

$$\operatorname{dom} \partial g = \{u \in H_0^1(\Omega) \mid Q\nabla u \in L^2(\Omega; \mathbb{R}^N)\},$$

$$\partial g(u) = -\operatorname{div}(Q\nabla u).$$

We will next use Theorem 4.1 to derive a formula for ∂g in the case $Q \in L^1(\Omega; \mathbb{R}^{N\times N})$. In this situation, $\operatorname{dom} I_f \neq L^2(\Omega; \mathbb{R}^N)$, so the condition $0 \in \operatorname{int}(\operatorname{rge} \nabla - \operatorname{dom} I_f)$ may fail. We will need the following two lemmas.

Lemma 6.1. *Given any two norms $|\cdot|$ and $||\cdot||$ on $\mathbb{R}^N$ and $\mathbb{R}^{N\times N}$, respectively, there is a constant C such that for every symmetric positive semidefinite matrix $M \in \mathbb{R}^{N\times N}$*

1. $|Mv| \leqslant C(||M|| + v \cdot Mv)\ \forall v \in \mathbb{R}^N$,
2. $|Mv| \leqslant C\sqrt{||M||v \cdot Mv}\ \forall v \in \mathbb{R}^N$.

Proof. Let $M = Q^*\Lambda Q$, with $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_N)$, be the spectral decomposition of M, and denote the p-norm on $\mathbb{R}^N$ by $|\cdot|_p$. Since M is positive semidefinite $\lambda_i \geqslant 0$, and we get

$$\begin{aligned} |Mv|_2 = |\Lambda Qv|_2 &\leqslant c|\Lambda Qv|_1 \\ &= c\left(\sum_{i=1}^N \lambda_i |(Qv)_i|\right) \\ &\leqslant c\left(\sum_{i=1}^N \lambda_i + \sum_{i=1}^N \lambda_i (Qv)_i^2\right) \\ &\leqslant c\left(N\rho(M) + \sum_{i=1}^N \lambda_i (Qv)_i^2\right), \end{aligned}$$

where $\rho(M)$ is the spectral radius of M. Part 1 follows by noting that the spectral radius is a norm on symmetric matrices, and that $\sum_{i=1}^N \lambda_i (Qv)_i^2 = (Qv) \cdot \Lambda(Qv) = v \cdot Mv$. Applying part 1 to λv with an arbitrary $\lambda > 0$ we get

$$|Mv| \leqslant C \inf_{\lambda > 0}\left(\frac{||M||}{\lambda} + \lambda v \cdot Mv\right) = 2C\sqrt{||M||v \cdot Mv},$$

so the result follows by changing constants. □

Lemma 6.2. *Let $\varphi_n, \varphi : \Omega \to \bar{\mathbb{R}}$ be measurable functions such that $\varphi_n \to \varphi$ a.e. on Ω. If there exist a $\psi \in L^1(\Omega)$ and a continuous function $\rho : [0, \infty) \to [0, \infty)$ such that $\rho(0) = 0$ and*

$$\int_E |\varphi_n| < \rho\left(\int_E |\psi|\right) \quad \forall n, \tag{5}$$

for every measurable $E \subset \Omega$, then $\varphi_n, \varphi \in L^1(\Omega)$ *and* $||\varphi_n - \varphi||_{L^1(\Omega)} \to 0$.

Proof. It is clear that (5) implies $\varphi_n \in L^1(\Omega)$. It thus suffices by Vitali's theorem (see for example [25, Section 4.3, Theorem 11]) to check that

1. For each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\text{meas}(E) \leqslant \delta \Rightarrow \int_E |\varphi_n(x)|\, dx \leqslant \varepsilon \quad \forall n,$$

2. For each $\varepsilon > 0$ there is an $E_\varepsilon \subset \Omega$ such that $\text{meas}(E_\varepsilon) < \infty$ and

$$\int_{\Omega \setminus E_\varepsilon} |\varphi_n| < \varepsilon \quad \forall n.$$

Let $\varepsilon > 0$. By continuity of ρ, there is an $\eta > 0$, such that $\rho(\xi) \leqslant \varepsilon$ for all $\xi \in [0, \eta]$. On the other hand, by integrability of ψ, we can find $\delta > 0$, such that

$$\text{meas}(E) \leqslant \delta \Rightarrow \int_E |\psi| \leqslant \eta.$$

Condition 1 thus holds.

To show that the second condition holds, let $\{E_k\}_{k \in \mathbb{N}}$, be such that $E_k \subset E_{k+1}$, $\text{meas}(E_k) < \infty$, and $\cup E_k = \Omega$. Then, by the monotone convergence theorem,

$$\int_{E_k} |\psi| = \int_\Omega |\psi| \chi_{E_k} \nearrow \int_\Omega |\psi|$$

so that $\int_{\Omega \setminus E_k} |\psi| \to 0$. The second condition thus follows from (5) and the continuity of ρ. $\square$

Recall that the divergence is defined (in the distribution sense) for any vector-valued distribution, and that the dual $H^{-1}(\Omega)$ of $H_0^1(\Omega)$ is embedded in the space of distributions. In what follows, $C_c^\infty(\Omega)$ denotes the test functions on Ω, $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$, and J the duality mapping from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$.

Note also that $\nabla w \cdot Q \nabla u \in L^1(\Omega)$ whenever $w, u \in \text{dom}\, g$. Indeed, by Cauchy–Schwarz inequality,

$$|\nabla w \cdot Q \nabla u| \leq |\nabla w \cdot Q \nabla w|^{\frac{1}{2}} |\nabla u \cdot Q \nabla u|^{\frac{1}{2}},$$

and then by Hölder inequality,

$$\int_\Omega |\nabla w \cdot Q\nabla u| \leqslant \left[\int_\Omega \nabla w \cdot Q\nabla w\right]^{\frac{1}{2}} \left[\int_\Omega \nabla u \cdot Q\nabla u\right]^{\frac{1}{2}} = 2g(w)^{\frac{1}{2}} g(u)^{\frac{1}{2}}.$$

We are now ready to state the main result of this section.

Theorem 6.1. *If* $Q \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R}^{N\times N})$, *then* $C_c^\infty(\Omega) \subset \operatorname{dom} g$, $Q\nabla u \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R}^N)$ *for all* $u \in \operatorname{dom} g$, *and*

$$\operatorname{dom} \partial g = \Big\{ u \in H_0^1(\Omega) \mid u \in \operatorname{dom} g,\ \operatorname{div}(Q\nabla u) \in H^{-1}(\Omega),$$
$$\langle w, -\operatorname{div}(Q\nabla u)\rangle = \int_\Omega \nabla w \cdot Q\nabla u \quad \forall w \in \operatorname{dom} g \Big\}, \tag{6}$$

$$\partial g(u) = -\operatorname{div}(Q\nabla u). \tag{7}$$

Proof. If $w \in C_c^\infty(\Omega)$, then $\nabla w \in C_c^\infty(\Omega; \mathbb{R}^N)$ and so $\nabla w \cdot Q\nabla w \in L^1(\Omega)$. Thus, $C_c^\infty(\Omega) \subset \operatorname{dom} g$. By Lemma 6.1(1), we have for any $u \in H_0^1(\Omega)$,

$$|Q(x)\nabla u(x)| \leqslant C(||Q(x)|| + \nabla u(x) \cdot Q(x)\nabla u(x)),$$

so if $u \in \operatorname{dom} g$, we get $Q\nabla u \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R}^N)$.

To derive the subdifferential formula, we will denote by G the mapping given by (6) and (7). Because $0 \in \operatorname{dom} g$, we have $\partial g = (\nabla^* \mathscr{L}_2[Q]\nabla)_v$ by Theorem 4.1. By maximal monotonicity of ∂g, it thus suffices to show that G is monotone and $(\nabla^* \mathscr{L}_2[Q]\nabla)_v \subset G$. Since for every $u_1, u_2 \in \operatorname{dom} G$

$$\begin{aligned}
\langle u_1 - u_2, G(u_1) - G(u_2)\rangle &= \langle u_1, -\operatorname{div}(Q\nabla u_1)\rangle - \langle u_1, -\operatorname{div}(Q\nabla u_2)\rangle \\
&\quad - \langle u_2, -\operatorname{div}(Q\nabla u_1)\rangle + \langle u_2, -\operatorname{div}(Q\nabla u_2)\rangle \\
&= \int_\Omega \nabla u_1 \cdot Q\nabla u_1 - \int_\Omega \nabla u_1 \cdot Q\nabla u_2 \\
&\quad - \int_\Omega \nabla u_2 \cdot Q\nabla u_1 + \int_\Omega \nabla u_2 \cdot Q\nabla u_2 \\
&= \int_\Omega \nabla(u_1 - u_2) \cdot Q\nabla(u_1 - u_2),
\end{aligned}$$

the monotonicity follows from the positive semidefiniteness of $Q(x)$. Note that $(\nabla^* \mathscr{L}_2[Q]\nabla)_v \subset G$ is equivalent to

$$[J + (\nabla^* \mathscr{L}_2[Q]\nabla)_v]^{-1}(u^*) \subset (J + G)^{-1}(u^*) \quad \forall u^* \in H^{-1}(\Omega)$$

where by Theorems 4.1 and 2.2

$$[J + (\nabla^* \mathscr{L}_2[Q]\nabla)_v]^{-1}(u^*) = \lim_{\lambda \searrow 0} [J + \nabla^* \mathscr{L}_2[Q]_\lambda \nabla]^{-1}(u^*).$$

To complete the proof it thus suffices to show that for each $u^* \in H^{-1}(\Omega)$, the limit $\bar{u}$ of the strongly convergent family $\{u_\lambda\} \subset H_0^1(\Omega)$ defined through

$$J(u_\lambda) + (\nabla^* \mathscr{L}_2[Q]_\lambda \nabla)(u_\lambda) = u^* \tag{8}$$

is a solution of

$$J(u) + G(u) = u^*. \tag{9}$$

First, since $\bar{u} \in \operatorname{dom}(\nabla^* \mathscr{L}_2[Q]_\lambda \nabla)_v = \operatorname{dom} \partial g$, by Theorem 4.1, we have $\bar{u} \in \operatorname{dom} g$. Because $\mathscr{L}_2[Q]_\lambda = \mathscr{L}_2[Q_\lambda]$, where $Q_\lambda \in L^\infty(\Omega; \mathbb{R}^{N \times N})$, (8) means that

$$\langle w, J(u_\lambda) \rangle + \int_\Omega \nabla w \cdot Q_\lambda \nabla u_\lambda = \langle w, u^* \rangle \tag{10}$$

for every $w \in H_0^1(\Omega)$. In particular, with $w = u_\lambda$ we get

$$||u_\lambda||^2_{H_0^1(\Omega)} + \int_\Omega \nabla u_\lambda \cdot Q_\lambda \nabla u_\lambda \leqslant ||u_\lambda||_{H_0^1(\Omega)} ||u^*||_{H^{-1}(\Omega)},$$

which implies

$$\int_\Omega \nabla u_\lambda \cdot Q_\lambda \nabla u_\lambda \leqslant \frac{1}{4} ||u^*||^2_{H^{-1}(\Omega)} \quad \forall \lambda > 0. \tag{11}$$

Now take $w \in \operatorname{dom} g$ in (10), and let

$$\varphi_\lambda = \nabla w \cdot Q_\lambda \nabla u_\lambda \quad \text{and} \quad \varphi = \nabla w \cdot Q \nabla \bar{u}.$$

We will show that the conditions of Lemma 6.2 are satisfied. Since $Q(x)_\lambda \to Q(x)$ for all $x \in \Omega$, and $\nabla u_\lambda \to \nabla \bar{u}$ strongly in $L^2(\Omega; \mathbb{R}^N)$, we have (by passing to a subsequence if necessary) that $\varphi_\lambda \to \varphi$ a.e. in Ω. By Cauchy–Schwarz inequality

$$|\varphi_\lambda(x)| \leqslant [\nabla w(x) \cdot Q(x)_\lambda \nabla w(x)]^{\frac{1}{2}} [\nabla u_\lambda(x) \cdot Q(x)_\lambda \nabla u_\lambda(x)]^{\frac{1}{2}},$$

where

$$\nabla w(x) \cdot Q(x)_\lambda \nabla w(x) \leqslant \nabla w(x) \cdot Q(x) \nabla w(x),$$

since for each x, the function $\phi_\lambda(v) := v \cdot Q(x)_\lambda v$ is the Moreau–Yosida regularization of $\phi(v) := v \cdot Q(x) v$. Thus by Hölder's inequality, we have for every measurable $E \subset \Omega$,

$$\int_E |\varphi_\lambda(x)|\,dx \leqslant \left[\int_E \nabla u_\lambda \cdot Q_\lambda \nabla u_\lambda\right]^{\frac{1}{2}} \left[\int_E \nabla w \cdot Q\nabla w\right]^{\frac{1}{2}}$$
$$\leqslant \frac{1}{2}\|u^*\|_{H^{-1}(\Omega)} \left[\int_E \nabla w \cdot Q\nabla w\right]^{\frac{1}{2}},$$

where the second inequality follows from (11). Since $w \in \operatorname{dom} g$, we have $\nabla w \cdot Q\nabla w \in L^1(\Omega)$, so the assumptions of Lemma 6.2 are in force. Thus, $\int_\Omega \varphi_\lambda \to \int_\Omega \varphi$, and passing to the limit in (10) gives

$$\langle w, J(\bar{u})\rangle + \int_\Omega \nabla w \cdot Q\nabla \bar{u} = \langle w, u^*\rangle \quad \forall w \in \operatorname{dom} g. \tag{12}$$

Because $C_c^\infty(\Omega) \subset \operatorname{dom} g$, by the first part of the theorem, (12) implies that

$$J(\bar{u}) - \operatorname{div}(Q\nabla \bar{u}) = u^* \tag{13}$$

in the sense of distributions. But since $J(\bar{u}), u^* \in H^{-1}(\Omega)$, (13) must hold also in $H^{-1}(\Omega)$, with $\operatorname{div}(Q\nabla\bar{u}) \in H^{-1}(\Omega)$. We thus have for every $w \in \operatorname{dom} g$,

$$\langle w, -\operatorname{div}(Q\nabla\bar{u})\rangle = \langle w, u^*\rangle - \langle w, J(\bar{u})\rangle$$
$$= \int_\Omega \nabla w \cdot Q\nabla \bar{u},$$

where the second equality follows from (12). In summary, $\bar{u}$ solves (9). □

Combining Theorem 6.1 with general results of convex analysis, one can derive existence criteria for PDEs associated with the operator ∂g. The following gives a simple example.

Corollary 6.1. *Let Ω be bounded, $\alpha > 0$, and let $Q \in L^1(\Omega; \mathbb{R}^{N\times N})$ be such that $v \cdot Q(x)v \geqslant \alpha|v|^2$ for all $v \in \mathbb{R}^m$ and for a.e. $x \in \Omega$. Then for each $u^* \in H^{-1}(\Omega)$, there exists a unique $u \in H_0^1(\Omega)$ such that*

$$-\operatorname{div}(Q\nabla u) = u^*,$$
$$\nabla u \cdot Q\nabla u \in L^1(\Omega), \qquad Q\nabla u \in L^1(\Omega; \mathbb{R}^N),$$

and

$$\langle w, -\operatorname{div}(Q\nabla u)\rangle = \int_\Omega \nabla w \cdot Q\nabla u$$

for all $w \in H_0^1(\Omega)$ such that $\nabla w \cdot Q\nabla w \in L^1(\Omega)$.

Proof. Let g be as in Theorem 6.1. Then by our assumptions on Q and by Poincaré's inequality, we can find $c > 0$ such that

$$g(u) \geqslant \frac{\alpha}{2} \int_{\Omega} |\nabla u(x)|^2 \, dx \geqslant c \|u\|^2_{H^1_0(\Omega)} \quad \forall u \in H^1_0(\Omega).$$

Thus, g is coercive and so is then $g - \langle \cdot, u^* \rangle$. This implies that $g - \langle \cdot, u^* \rangle$ has a unique minimizer, or equivalently, that the inclusion $\partial g(u) \ni u^*$ has a unique solution. The result thus follows from Theorem 6.1. □

Combining Theorem 6.1 with the results of [13], one can derive existence results for evolution equations associated with the operator ∂g. Similarly, a combination of Theorems 6.1 and 5.3 could be used to study time-dependent evolution equations. Using the above techniques, one could also study nonlinear PDEs, where the linear operator $Q(x)$ is replaced by $\partial f(x, \cdot)$ for a more general convex normal integrand f. Results in this direction have been obtained, e.g. in [6] in a slightly different setting. Our approach is also related to [20] where the convergence of sequences of Dirichlet forms related to composite media was studied. Interesting links can also be found with energy functionals with respect to measures [10], generalized quadratic forms and the second-order nonsmooth calculus of Rockafellar [34]; see also [21].

7. Uncited references

[7]

Acknowledgments

We are very much indebted to an anonymous referee for a number of important remarks, particularly concerning an earlier version of our Theorem 6.1 and its proof. We are also grateful to professors H. Attouch, P. Combettes and M. Willem for many helpful suggestions.

References

[1] H. Attouch, Familles d'opérateurs maximaux monotones et mesurabilité, Ann. Mat. Pura Appl. 120 (1979) 35–111.
[2] H. Attouch, Variational Convergence for Functions and Operators, Applicable Mathematics Series, Pitman, London, 1984.
[3] H. Attouch, J.-B. Baillon, M. Théra, Variational sum of monotone operators, J. Convex Anal. 1 (1994) 1–29.
[4] H. Attouch, J.-B. Baillon, M. Théra, Weak solutions of evolution equations and variational sum of maximal monotone operators, Southest Asian Bull. Math. 19 (2) (1995) 117–126.
[5] H. Attouch, Z. Chbani, A. Moudafi, Recession operators and solvability of variational problems in reflexive Banach spaces, in: Calculus of Variations, Homogenization and Continuum Mechanics

(Marseille, 1993), Series Advanced in Mathematical and Applied Sciences, Vol. 218, World Scientific Publishing, River Edge, 1994, pp. 51–67.

[6] H. Attouch, A. Damlamian, Application des méthodes de convexité et monotonie à l'étude de certaines équations quasi linéaires, Proc. Roy. Soc. Edinburgh Sect. A 79 (1–2) (1977/78) 107–129.

[7] H. Attouch, M. Théra, Convergences en analyse multivoque et unilatérale, MATAPLI, Bull. de liaison 36 (1993) 23–40.

[8] H. Attouch, R.J-B. Wets, Quantitative analysis of variational systems: I. The epigraphical distance, Trans. Amer. Math. Soc. 328 (1991) 695–729.

[9] G. Beer, Topologies on Closed and Closed Convex Sets, in: Mathematics and its Applications, Vol. 268, Kluwer Academic Publishers Group, Dordrecht, 1993.

[10] G. Bouchitté, G. Buttazzo, P. Seppecher, Energies with respect to a measure and applications to low-dimensional structures, Calc. Var. Partial Differential Equations 5 (1) (1997) 37–54.

[11] H. Brezis, Opérateurs Maximaux Monotones et Semi Groupes de Contractions Dans Les Espaces de Hilbert, North-Holland, Amsterdam, 1973.

[12] C. Castaing, M. Valadier, Convex analysis and measurable multifunctions, in: Lecture Notes in Mathematics, Vol. 580, Springer, Berlin, New York, 1977.

[13] M.G. Crandall, A. Pazy, Semi-groups of nonlinear contractions and dissipative sets, J. Funct. Anal. 3 (1969) 376–418.

[14] T. Diagana, Variational sum and Kato's conjecture, preprint, 2000.

[15] J. Diestel, Geometry of Banach spaces, Selected Topics, Lecture Notes in Mathematics, Springer, Berlin, 1975.

[16] J.-B. Hiriart-Urruty, M. Moussaoui, A. Seeger, M. Volle, Subdifferential calculus without qualification conditions, using approximate subdifferentials: a survey, Nonlinear Anal. TMA 24 (1995) 1727–1754.

[17] J.-B. Hiriart-Urruty, R.R. Phelps, Subdifferential calculus using ε-subdifferentials, J. Funct. Anal. 118 (1) (1993) 154–166.

[18] A. Jourani, Variational sum of subdifferentials of convex functions, in: C. Garcia, C. Olivé, M. Sanroma (Eds.), Proceedings of the Fourth Catalan Days on Applied Mathematics, Tarragona Press University, Tarragona, 1998, pp. 71–80.

[19] U. Mosco, Convergence of convex sets and of solutions of variational inequalities, Adv. Math. 3 (1969) 510–585.

[20] U. Mosco, Composite media and asymptotic Dirichlet forms, J. Funct. Anal. 123 (21) (1994) 368–421.

[21] J.L. Ndoutoune, M. Théra, Generalised second order derivatives of convex functions in reflexive Banach spaces, Bull. Austral. Math. Soc. 51 (1995) 55–72.

[22] T. Pennanen, Dualization of generalized equations of maximal monotone type, SIAM J. Optim. 10 (2000) 809–835.

[23] J.-P. Penot, Subdifferential calculus without qualification conditions, J. Convex Anal. 3 (1996) 1–13.

[24] R.R. Phelps, Convex Functions, Monotone Operators and Differentiability, in: Lecture Notes in Mathematics, Vol. 1364, Springer, Berlin, 1989.

[25] M.M. Rao, Measure Theory and Integration, Pure and Applied Mathematics, A Wiley-Interscience Publication, Wiley, New York, 1987.

[26] J.P. Revalski, M. Théra, Variational and extended sums of monotone operators, in: M. Théra, R. Tichatschke (Eds.), Ill-Posed Variational Problems and Regularization Techniques, Lecture Notes in Economics and Mathematical Systems, Vol. 477, Springer, Berlin, 1999, pp. 229–246.

[27] J.P. Revalski, M. Théra, Enlargements and sums of monotone operators, Nonlinear Anal. TMA 48 (2002) 505–519.

[28] S.M. Robinson, Composition duality and maximal monotonicity, Math. Programing Ser. A 85 (1999) 1–13.

[29] R.T. Rockafellar, On the maximality of sums of nonlinear monotone operators, Trans. Amer. Math. Soc. 149 (1970) 75–88.

[30] R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math. 33 (1970) 209–216.

[31] R.T. Rockafellar, Convex integral functionals and duality, in: Contributions to Nonlinear Functional Analysis, Proceedings of the Symposium, Mathematical Research Center, University of Wisconsin, Madison, WI, Academic Press, New York, 1971, pp. 215–236.
[32] R.T. Rockafellar, Conjugate Duality and Optimization, SIAM, 1974.
[33] R.T. Rockafellar, Integral functionals, normal integrands and measurable selections, in: Nonlinear Operators and the Calculus of Variations, Lecture Notes in Mathematics, Vol. 543, Springer, Berlin, 1976, pp. 157–207.
[34] R.T. Rockafellar, Generalized derivatives of convex functions and saddle functions, Trans. Amer. Math. Soc. 320 (1990) 810–822.
[35] R.T. Rockafellar, R.J-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[36] L. Thibault, A general sequential formula for subdifferentials of sums of convex functions defined on Banach spaces, in: R. Durier, C. Michelot (Eds.), Recent Developments in Optimization (Seventh French–German Conference held at Dijon in June 27–July 2, 1994), Lecture Notes in Economics and Mathematical Systems, Vol. 429, Springer, Berlin, 1995, pp. 340–345.
[37] L. Thibault, A short note on sequential convex subdifferential calculus, unpublished paper, 1995.
[38] E. Zeidler, Nonlinear Functional Analysis and its Applications II, Springer, Berlin, 1990.