J Optim Theory Appl
https://doi.org/10.1007/s10957-018-1254-0

Epi-convergence: The Moreau Envelope and Generalized Linear-Quadratic Functions

Chayne Planiden[1] · **Xianfu Wang**[2]

Received: 1 August 2017 / Accepted: 19 February 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract This work explores the class of generalized linear-quadratic functions, constructed using maximally monotone symmetric linear relations. Calculus rules and properties of the Moreau envelope for this class of functions are developed. In finite dimensions, on a metric space defined by Moreau envelopes, we consider the epigraphical limit of a sequence of quadratic functions and categorize the results. We examine the question of when a quadratic function is a Moreau envelope of a generalized linear-quadratic function; characterizations involving nonexpansiveness and Lipschitz continuity are established. This work generalizes some results by Hiriart-Urruty and by Rockafellar and Wets.

Keywords Attouch–Wets metric · Complete metric space · Epi-convergence · Extended seminorm · Fenchel conjugate · Firmly nonexpansive · Generalized linear-quadratic function · Linear relation · Lipschitz continuous · Maximally monotone · Nonexpansive · Moreau envelope · Proximal mapping

Dedicated to the pioneers of epi-convergence: Attouch, Beer, Wets, …

Communicated by Giuseppe Buttazzo.

✉ Xianfu Wang
shawn.wang@ubc.ca

Chayne Planiden
chayne@uow.edu.au

1 School of Mathematics & Applied Statistics, University of Wollongong, Northfields Avenue, Wollongong, NSW 2500, Australia

2 Department of Mathematics, University of British Columbia Okanagan, Kelowna, BC V1V 1V7, Canada

Published online: 05 March 2018

Mathematics Subject Classification 47A06 · 52A41 · 47H05 · 90C31

1 Introduction

The Moreau envelope is a well-established and extensively researched function that emerged in the 1960s [1]. It is of great use in optimization due to its regularizing properties, differentiability and coincidence of the minimizers of the objective function in the convex setting. This work continues the investigation into Moreau envelopes in finite dimensions, from the perspective of the generalized linear-quadratic objective function. The long-range reason for studying Moreau envelopes in general is, as alluded to in [2], that if we had a sufficient level of understanding about their properties, it would likely facilitate the development of minimization methods for Moreau envelopes, and therefore for their associated (nonsmooth) objective functions as well. In this work, we focus on generalized linear-quadratic functions, because it is a class of functions that has enough structure to secure solid results that do not require overly restrictive conditions, but allows us to obtain results that are useful for a wide range of functions. We define a metric space whose distance function is constructed using Moreau envelopes, with the intention of exploring the epi-convergence of a sequence of quadratic functions. (See [3] for more on epi-distance between functions.) The idea of studying epi-convergence of convex functions via the Moreau envelope is due to Attouch–Wets [3] and Attouch [4]. Several classes of functions can arise at the limit; these results are classified and illustrated. Then we approach the relationship between Moreau envelopes and quadratics from the opposite direction, asking under what conditions a given quadratic function is a Moreau envelope of another function, and whether said other function can be determined explicitly. A very nice characterization of convex Moreau envelopes arises from this study, one that involves Lipschitz gradients. This is one of the main results of the paper.

The linear relation is also a useful tool in functional analysis, notably documented and developed in [5], with more recent expansion in [6–9]. This paper continues to develop the theory of monotone linear relations, in particular for the class of generalized linear-quadratic functions. Such functions arise, for example, in the determination of the existence of a Hessian for the Moreau envelope [2,10]. In [2, Theorem 3.9], Rockafellar and Poliquin showed that a function does not have to be finite in order for its Moreau envelope to have a Hessian; it suffices that the second-order epi-derivative of the function be a generalized linear-quadratic function. The existence of a Hessian is of interest, since it is needed in order to do a second-order expansion of the Moreau envelope function, which leads to a second-order approximation of its objective function. Several properties and characterizations for the class of generalized linear-quadratic functions are provided in this work, and we demonstrate that it is useful and convenient to work in the setting of generalized linear-quadratic functions when considering matters of epi-convergence. In this paper, we show

(i) that monotone linear relations provide a unified framework for generalized linear-quadratic functions;
(ii) that the Fenchel conjugate of every generalized linear-quadratic convex function can be written in terms of the set-valued inverse of a monotone linear relation;

(iii) that a function is convex generalized linear-quadratic, if and only if its Moreau envelope is convex quadratic;
(iv) the characterization that a function is a Moreau envelope with parameter r of a proper, convex and lsc function, if and only if its gradient is r-Lipschitz.

We also establish calculus rules for the set of generalized linear-quadratic functions, and we generalize the result of Rockafellar [11, p. 136] and that of Hiriart-Urruty deconvolution [12, Example 2.7], from positive definite matrices to maximally monotone symmetric linear relations.

The rest of this paper is organized as follows. Section 2 contains notation, definitions and basic properties of Moreau envelopes, epi-convergence and monotone operators. In Sect. 3, we discuss epigraphical limits of linear-quadratic functions in one dimension. Several illustrative examples are presented, with graphs showing the limiting behavior of the Moreau envelope for sequences of quadratic functions. Section 4 contains the principal matter of this work: properties, characteristics and results on epi-convergence of generalized linear-quadratic functions on finite-dimensional space. Topics include symmetry, maximal monotonicity, nonexpansiveness, the subdifferential, sum, difference and infimal convolution rules, the adjoint, the set-valued and Moore–Penrose inverses and the Fenchel conjugate. This section includes characterizations of Moreau envelopes of generalized linear-quadratic functions. In Sect. 5, we give applications of these results and develop a calculus of the class of generalized linear-quadratic functions. Applications are to the seminorm function, the least squares problem and the limit of a sequence of generalized linear-quadratic functions. Section 6 makes concluding remarks.

2 Preliminaries

This section collects several definitions and facts from previous works, that we will use later in proving our main results. For proof of the facts, we refer the reader to the corresponding citations.

2.1 Notation

All functions in this work are defined on $\mathbb{R}^n$, Euclidean space equipped with inner product defined $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ and induced norm $\|x\| = \sqrt{\langle x, x\rangle}$. The extended real line $\mathbb{R} \cup \{\infty\}$ is denoted $\overline{\mathbb{R}}$. We use $\Gamma_0(\mathbb{R}^n)$ to represent the set of proper, convex and lower semicontinuous (lsc) functions on $\mathbb{R}^n$. The identity operator is denoted Id. Pointwise convergence is denoted $\xrightarrow{p}$, graphical convergence $\xrightarrow{g}$ and epi-convergence $\xrightarrow{e}$. The function $\frac{1}{2}\|\cdot\|^2$ is denoted q. We use $N_C(x)$ to represent the normal cone to C at x, as defined in [13]. The relative interior of a set A is denoted $\operatorname{ri} A$. The domain and range of an operator A are denoted $\operatorname{dom} A$ and $\operatorname{ran} A$, respectively. On $\overline{\mathbb{R}}$, where necessary we use inf-addition and accept the convention $\infty - \infty = \infty$ (see [13, p. 15]). We use S^n, S^n_+ and S^n_{++} to denote the sets of symmetric, positive semidefinite and positive definite matrices, respectively. The *graph* of an operator $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is defined

$$\text{gra}\, A = \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n : x^* \in Ax\}.$$

Its set-valued inverse $A^{-1} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is defined by the graph

$$\text{gra}\, A^{-1} = \{(x^*, x) \in \mathbb{R}^n \times \mathbb{R}^n : x^* \in Ax\}.$$

For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the function $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^n} \{\langle x^*, x\rangle - f(x)\}$$

is called the *Fenchel conjugate* of f. For $x \in \text{dom}\, f$, the *Fenchel subdifferential* of f is the set

$$\partial f(x) = \{x^* \in \mathbb{R}^n : f(y) \geq f(x) + \langle x^*, y - x\rangle \ \forall y \in \mathbb{R}^n\}.$$

For any $x \notin \text{dom}\, f$, $\partial f(x) = \varnothing$.

2.2 Moreau Envelopes, Proximal Mappings and Their Properties

We work with Moreau envelopes of functions throughout this paper, as defined below.

Definition 2.1 For a proper and lsc function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the *Moreau envelope* of f with prox-parameter $r > 0$ is denoted by $e_r f$ and defined

$$e_r f(x) = \inf_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{r}{2}\|y - x\|^2 \right\}.$$

The associated *proximal mapping* is the set of all points at which the above infimum is attained, denoted by Prox_f^r:

$$\text{Prox}_f^r(x) = \underset{y \in \mathbb{R}^n}{\text{argmin}} \left\{ f(y) + \frac{r}{2}\|y - x\|^2 \right\}.$$

Lemma 2.1 *For any proper function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$,

$$e_r f(x) = \frac{r}{2}\|x\|^2 - g^*(rx), \tag{1}$$

where $g(x) = f(x) + \frac{r}{2}\|x\|^2$.

Proof We have

$$\begin{aligned} e_r f(x) &= \inf_y \left\{ f(y) + \frac{r}{2}\|y-x\|^2 \right\} \\ &= -\sup_y \left\{ -f(y) - \frac{r}{2}(\|y\|^2 - 2\langle x, y\rangle + \|x\|^2) \right\} \\ &= \frac{r}{2}\|x\|^2 - \sup_y \left\{ \langle rx, y\rangle - \left(f(y) + \frac{r}{2}\|y\|^2 \right) \right\} \\ &= \frac{r}{2}\|x\|^2 - g^*(rx). \end{aligned}$$

□

Fact 2.1 [14, Example 23.3] *In the case of a convex function f, an alternate representation of the proximal mapping makes use of the* resolvent *of the subdifferential of f, which also provides a conversion to the proximal mapping with prox-parameter 1:*

$$\mathrm{Prox}_f^r = \left(\mathrm{Id} + \frac{1}{r}\partial f \right)^{-1} = \mathrm{Prox}_{\frac{1}{r}f}^1 .$$

An alternate expression for the Moreau envelope is reached through infimal convolution:

$$e_r f = f \,\square\, \frac{r}{2}\|\cdot\|^2 = f \,\square\, (rq). \tag{2}$$

Fact 2.2 [14, Theorem 16.2] *For any proper and lsc function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, and any $r > 0$, we have*

$$\begin{aligned} p \in \mathrm{Prox}_f^r(x) &\Rightarrow 0 \in \partial f(p) + r(p-x) \\ &\Leftrightarrow 0 \in \frac{1}{r}\partial f(p) + p - x. \end{aligned}$$

If, in addition, f is convex, then the first implication above becomes a two-way implication:

$$p \in \mathrm{Prox}_f^r(x) \Leftrightarrow 0 \in \partial f(p) + r(p-x).$$

Proposition 2.1 (Calculus of Moreau envelopes) *For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $r > 0$, $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$, the following hold:*

(i) $e_r(f+c) = e_r f + c$;
(ii) $e_r f = r e_1(f/r)$;
(iii) $e_r(f(\cdot - c)) = (e_r f)(\cdot - c)$;
(iv) $e_1 f = q - (f+q)^*$;
(v) $e_1(f + \langle \cdot, v\rangle) = e_1 f(\cdot - v) + \langle \cdot, v\rangle - q(v)$;
(vi) $(e_r f)^* = f^* + q/r$.

Proof (i) This is seen directly as a property of the infimum: for any function g and any $c \in \mathbb{R}$, $\inf\{g(x)+c\} = \inf\{g(x)\}+c$.

(ii) See [14, Proposition 12.22].

(iii) Let $z = y - c$. Then

$$\begin{aligned} e_r(f(\cdot - c))(x) &= \inf_{y\in\mathbb{R}^n}\left\{f(y-c)+\frac{r}{2}\|y-x\|^2\right\} \\ &= \inf_{z\in\mathbb{R}^n}\left\{f(z)+\frac{r}{2}\|z-(x-c)\|^2\right\} \\ &= (e_r f)(x-c). \end{aligned}$$

(iv) This is Lemma 2.1 with $r = 1$.

(v) Consider the left-hand side of statement (v) first. Applying statement (iv) to $f + \langle\cdot, v\rangle$, we have

$$e_1(f+\langle\cdot, v\rangle) = q - (f + \langle\cdot, v\rangle + q)^*.$$

Applying [14, Proposition 13.20(iii)] to the function $f+q$ with $y = 0$ and $\alpha = 0$, we have

$$e_1(f+\langle\cdot, v\rangle) = q - [f(\cdot - v) + q(\cdot - v)]^*. \tag{3}$$

Now consider the right-hand side of statement (v). Applying statement (iv) to $f(\cdot - v)$, we have

$$\begin{aligned} e_1(f(\cdot - v)) &= q(\cdot - v) - [f(\cdot - v) + q(\cdot - v)]^*, \\ &= q - [f(\cdot - v) + q(\cdot - v)]^* - \langle\cdot, v\rangle + q(v), \\ e_1(f(\cdot - v)) + \langle\cdot, v\rangle - q(v) &= q - [f(\cdot - v) + q(\cdot - v)]^*, \end{aligned}$$

which is the same as (3).

(vi) By [14, Proposition 13.21(iii)] with $g = rq$, we have $(f \,\square\, rq)^* = f^* + (rq)^*$. By (2), we have $(e_r f)^* = (f \,\square\, rq)^*$. Finally, [14, Example 13.4] gives us that $(rq)^* = q/r$. □

Definition 2.2 Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper and lsc. If there exists $r \geq 0$ such that $e_r f(x) > -\infty$ for some x, then f is said to be *prox-bounded*. The infimum of all such r is called the *threshold of prox-boundedness* of f.

Proposition 2.2 *Let $f \in \Gamma_0(\mathbb{R}^n)$. Then f is prox-bounded with threshold 0,* Prox_f^r *is single-valued and continuous, and $e_r f$ is convex and continuously differentiable. Moreover, the following properties hold:*

(i) $e_r f(x) + e_{\frac{1}{r}} f^*(rx) = \frac{r}{2}\|x\|^2$;

(ii) $\nabla e_r f(x) = r[x - \mathrm{Prox}_f^r(x)]$;

(iii) $\nabla e_r f^*(x) = \mathrm{Prox}_f^{\frac{1}{r}}(rx)$;

(iv) $\mathrm{Prox}_f^r(x) = \nabla g(x)$ *where* $g(x) = \frac{1}{r}\left[e_{\frac{1}{r}} f^*(rx)\right]$;

(v) $\mathrm{Prox}_{f^*}^r(x) = x - \frac{1}{r}\,\mathrm{Prox}_f^{\frac{1}{r}}(rx)$.

Proof The proof that f has threshold 0, Prox_f^r is single-valued and continuous, and $e_r f$ is convex and continuously differentiable is found in [13, Theorem 2.26].

(i) See [13, Example 11.26].

(ii) See [13, Theorem 2.26].

(iii) Replacing f with f^* in part (i) and using the fact that $f^{**} = f$, we have

$$e_r f^*(x) + e_{\frac{1}{r}} f(rx) = \frac{r}{2}\|x\|^2.$$

Differentiating both sides and rearranging yields

$$\nabla e_r f^*(x) = rx - \nabla e_{\frac{1}{r}} f(rx).$$

We substitute $z = rx$, then use part (ii) and the chain rule to get

$$\begin{aligned}
\nabla e_r f^*(x) &= rx - \nabla_x e_{\frac{1}{r}} f(z) \\
&= rx - \nabla_z e_{\frac{1}{r}} f(z) \nabla_x z \\
&= rx - \frac{1}{r}\left[z - \mathrm{Prox}_f^{\frac{1}{r}}(z)\right] r \\
&= \mathrm{Prox}_f^{\frac{1}{r}}(rx).
\end{aligned}$$

(iv) See [13, Exercise 11.27].

(v) Replacing f with f^* in part (iv), we have

$$\mathrm{Prox}_{f^*}^r(x) = \nabla g(x) \text{ where } g(x) = \frac{1}{r}\left[e_{\frac{1}{r}} f(rx)\right].$$

Substituting $z = rx$, then applying part (ii) and the chain rule yields

$$\begin{aligned}
\nabla g(x) &= \frac{1}{r}\nabla_z \left(e_{\frac{1}{r}} f(z)\right) \nabla_x z \\
&= \frac{1}{r}\left[\frac{1}{r}\left(z - \mathrm{Prox}_f^{\frac{1}{r}}(z)\right)\right] r \\
&= \frac{1}{r}\left(rx - \mathrm{Prox}_f^{\frac{1}{r}}(rx)\right) \\
&= x - \frac{1}{r}\,\mathrm{Prox}_f^{\frac{1}{r}}(rx).
\end{aligned}$$

□

2.3 Epi-convergence and the Attouch–Wets Metric

Epi-convergence plays a fundamental role in optimization and variational analysis (see [4,13,15–18]). In this section, we remind the reader of some definitions and facts related to epi-convergence, and we define a complete metric space that makes use of a variant of the Attouch–Wets metric.

Definition 2.3 For any sequence $\{f_k\}_{k\in\mathbb{N}}$ of functions on $\mathbb{R}^n$, the *lower epi-limit* $\operatorname{eliminf}_k f_k$ is the function having as its epigraph the outer limit of the sequence of sets $\operatorname{epi} f_k$:

$$\operatorname{epi}(\operatorname*{eliminf}_k f_k) = \limsup_k (\operatorname{epi} f_k).$$

The *upper epi-limit* $\operatorname{elimsup}_k f_k$ is the function having as its epigraph the inner limit of the sets $\operatorname{epi} f_k$:

$$\operatorname{epi}(\operatorname*{elimsup}_k f_k) = \liminf_k (\operatorname{epi} f_k).$$

When these two functions coincide, the *epi-limit* $\operatorname{elim}_k f_k$ is said to exist:

$$\operatorname*{elim}_k f_k = \operatorname*{eliminf}_k f_k = \operatorname*{elimsup}_k f_k.$$

In this event, the functions are said to *epi-converge* to f, symbolized by $f_k \xrightarrow{e} f$. Thus,

$$f_k \xrightarrow{e} f \Leftrightarrow \operatorname{epi} f_k \to \operatorname{epi} f.$$

There is a classical and ground-breaking epi-convergence result by Wijsman; a characterization that involves the Fenchel conjugate. It is the following.

Fact 2.3 [19, Theorem 6.2] *A sequence of functions $\{f_k\}_{k\in\mathbb{N}}$ in $\Gamma_0(\mathbb{R}^n)$ is epi-convergent to $f \in \Gamma_0(\mathbb{R}^n)$ iff the sequence $\{f_k^*\}_{k\in\mathbb{N}}$ is epi-convergent to f^*.*

Definition 2.4 A sequence of functions $\{f_k\}$ on $\mathbb{R}^n$ is *eventually prox-bounded* if there exists $r \geq 0$ such that $\liminf_{k\to\infty} e_r f_k(x) > -\infty$ for some x. The infimum of all such r is the *threshold of eventual prox-boundedness* of the sequence.

There is an important relationship among epi-convergence of proper and lsc functions, and pointwise and uniform convergence of their Moreau envelopes, as the following fact outlines.

Fact 2.4 [13, Theorem 7.37] *For proper, lsc functions f_k and f, the following are equivalent:*

(i) the sequence $\{f_k\}_{k\in\mathbb{N}}$ is eventually prox-bounded and $f_k \xrightarrow{e} f$;
(ii) f is prox-bounded and $e_r f_k \xrightarrow{p} e_r f$ for all $r \in]\varepsilon, \infty[$, $\varepsilon > 0$.

Then the pointwise convergence of $e_r f_k$ to $e_r f$ for $r > 0$ sufficiently large is uniform on all bounded subsets of $\mathbb{R}^n$, hence yields continuous convergence and epi-convergence as well. Indeed, $e_{r_k} f_k$ converges in all these ways to $e_r f$ whenever $r_k \to r \in]\bar{r}, \infty[$, where $\bar{r}$ is the threshold of eventual prox-boundedness. If f_k and f are convex, then $\bar{r} = 0$ and condition (ii) can be replaced by

(ii) $e_r f_k \xrightarrow{p} e_r f$ *for some* $r > 0$.

Since we are working in finite-dimensional space, epi-convergence is equivalent to the pointwise convergence of $\{d(\cdot, \operatorname{epi} f_k)\}$ to $d(\cdot, \operatorname{epi} f)$, which is in turn equivalent to uniform convergence of distance functions on bounded subsets of $\mathbb{R}^{n+1}$. Thus, the space is completely metrizable [20, pp. 78–80]. However, we are going to display a different metric for convex functions based on Fact 2.4.

Definition 2.5 *(Attouch–Wets metric)* Let $r > 0$. For $f, g \in \Gamma_0(\mathbb{R}^n)$, define the distance function

$$d(f, g) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\sup_{\|x\| \le i} |e_r f(x) - e_r g(x)|}{1 + \sup_{\|x\| \le i} |e_r f(x) - e_r g(x)|}.$$

Fact 2.5 [21, Proposition 3.5] *The space $(\Gamma_0(\mathbb{R}^n), d)$ is a complete metric space.*

Fact 2.6 [11, Theorem 25.7] *Let C be an open and convex set in $\mathbb{R}^n$. Let f be a convex function that is finite and differentiable on C. Let $\{f_k\}_{k \in \mathbb{N}}$ be a sequence of convex functions that are finite and differentiable on C, such that $\lim_{k \to \infty} f_k(x) = f(x)$ for every $x \in C$. Then*

$$\lim_{k \to \infty} \nabla f_k(x) = \nabla f(x) \ \forall x \in C.$$

In fact, the mappings ∇f_k converge to ∇f uniformly on every closed and bounded subset of C.

2.4 Monotone Operators and Resolvents

In this section, we list a number of facts involving monotonicity, maximal monotonicity and cyclic monotonicity.

Definition 2.6 An operator $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *monotone* if

$$\langle x^* - y^*, x - y \rangle \ge 0 \ \forall (x, x^*), (y, y^*) \in \operatorname{gra} A.$$

The monotone operator A is *maximally monotone* if there does not exist a monotone operator that contains A.

Definition 2.7 The *resolvent* J_A of a mapping $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is defined

$$J_A = (\operatorname{Id} + A)^{-1}.$$

Fact 2.7 [13, Lemma 12.14] *Every mapping* $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *obeys the identity*

$$\operatorname{Id} - (\operatorname{Id} + A)^{-1} = (\operatorname{Id} + A^{-1})^{-1}.$$

Fact 2.8 [13, Lemma 12.12] *Let* $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be monotone and* $\lambda > 0$. *Then* $(\operatorname{Id} + \lambda A)^{-1}$ *is monotone and nonexpansive. Moreover,* A *is maximally monotone iff the domain of* $(\operatorname{Id} + \lambda A)^{-1}$ *is* $\mathbb{R}^n$. *In that case,* $(\operatorname{Id} + \lambda A)^{-1}$ *is maximally monotone as well, and it is a single-valued mapping from all of* $\mathbb{R}^n$ *into itself.*

Fact 2.9 [14, Proposition 23.7] *Let* $D \neq \varnothing$ *be a subset of* $\mathbb{R}^n$, $T : D \to \mathbb{R}^n$, $A = T^{-1} - \operatorname{Id}$. *Then* T *is firmly nonexpansive iff* A *is monotone.*

Fact 2.10 [22, Theorem 6.6] *Let* $T : \mathbb{R}^n \to \mathbb{R}^n$. *Then* T *is the resolvent of the maximally cyclically monotone operator* $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *iff* T *has full domain,* T *is firmly nonexpansive, and for every set of points* $\{x_1, \ldots, x_m\}$ *where the integer* $m \geq 2$ *and* $x_{m+1} = x_1$, *one has*

$$\sum_{i=1}^{m} \langle x_i - T x_i, T x_i - T x_{i+1} \rangle \geq 0.$$

Fact 2.11 [14, Theorem 22.14] *Let* $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. *Then* A *is maximally cyclically monotone iff there exists* $f \in \Gamma_0(\mathbb{R}^n)$ *such that* $A = \partial f$.

Fact 2.12 (Baillon–Haddad Theorem) [23, Corollary 10] *Let* φ *be a convex* $\mathcal{C}^1$ *function on* $\mathbb{R}^n$. *Let* $A = \nabla \varphi$. *If* A *is* L*-Lipschitz, then*

$$\langle Au - Av, u - v \rangle \geq \frac{1}{L} \|Au - Av\|^2 \qquad \forall u, v \in \mathbb{R}^n.$$

Hence, $\frac{A}{L} = \nabla \left(\frac{\phi}{L} \right)$ *is firmly nonexpansive and 1-Lipschitz. Consequently,* $\frac{A}{L}$ *is a proximal mapping:*

$$\frac{A}{L} = \operatorname{Prox}_g^1 \text{ for some } g \in \Gamma_0(\mathbb{R}^n).$$

3 Epigraphical Limits of Quadratic Functions on $\mathbb{R}$

One of the main objectives of this paper is to present epi-convergence properties of generalized linear-quadratic functions and their Moreau envelopes. For the first set of results, we focus on quadratic functions on $\mathbb{R}$. This serves to show the variety of situations that can arise at the epigraphical limit of a sequence of quadratic functions. Then in the next section, we concentrate on the expansion to $\mathbb{R}^n$.

Theorem 3.1 *For all* $k \in \mathbb{N}$, *let* $a_k, b_k, c_k \in \mathbb{R}$ *with* $a_k \geq 0$. *Consider the sequence of functions*

$$\{f_k(x) = a_k x^2 + b_k x + c_k\}_{k \in \mathbb{N}} \text{ in } \Gamma_0(\mathbb{R}).$$

Then for $r > 0$ we have

$$e_r f_k(x) = \frac{a_k r}{2a_k + r} x^2 + \frac{b_k r}{2a_k + r} x + c_k - \frac{b_k^2}{2(2a_k + r)}. \tag{4}$$

Moreover, letting $k \to \infty$ and $f_k \xrightarrow{e} f$, we have the following trichotomy.

(i) If $f \equiv \infty$, then $e_r f \equiv \infty$.
(ii) If $f(x) = -\infty$ for some x, then $e_r f \equiv -\infty$.
(iii) If f is proper, then $e_r f$ is of the form $arx^2 + bx + c$ with $a \geq 0$. This is true even in the case where $a_k \to \infty$ and $f(x) = \iota_{\{b\}}(x) + c$.

Proof The Moreau envelope is not defined for improper functions such as those of parts (i) and (ii), but if we consider the same definition valid for improper functions, it is clear that in part (i) we have $e_r f \equiv \infty$ and in part (ii) we have $e_r f \equiv -\infty$. For part (iii), we want to consider the Moreau envelope at the limit of the sequence

$$\begin{aligned} e_r f_k(x) &= \inf_{y \in \mathbb{R}} \left\{ f_k(y) + \frac{r}{2}(y - x)^2 \right\} \\ &= \inf_{y \in \mathbb{R}} \left\{ \left(a_k + \frac{r}{2}\right) y^2 + (b_k - rx)y + c_k + \frac{r}{2}x^2 \right\}. \end{aligned}$$

The infimand above is a strictly convex quadratic function, so its minimum can be found by setting the derivative equal to zero and finding critical points. This yields the minimizer $y = \frac{rx - b_k}{2a_k + r}$, which gives

$$\begin{aligned} e_r f_k(x) &= \left(a_k + \frac{r}{2}\right) \frac{(rx - b_k)^2}{(2a_k + r)^2} + (b_k - rx) \frac{rx - b_k}{2a_k + r} + c_k + \frac{r}{2}x^2 \\ &= \frac{a_k r}{2a_k + r} x^2 + \frac{b_k r}{2a_k + r} x + c_k - \frac{b_k^2}{2(2a_k + r)}. \end{aligned}$$

As expected, $e_r f_k(x) \in \Gamma_0(\mathbb{R})$ for all k, since the quadratic coefficient is nonnegative. Now consider the sequence $f_k \xrightarrow{e} f$. By Fact 2.4, we need only consider the pointwise convergence of the sequence $\{e_r f_k\}_{k \in \mathbb{N}}$. Since $e_r f(x)$ is finite for all x, evaluating (3.1) at $x = 0$ and taking the limit as $k \to \infty$ gives us that the constant coefficient $c_k - b_k^2/[2(2a_k + r)]$ converges to some $c \in \mathbb{R}$. We know that $e_r f_k$ is differentiable for all k by Proposition 2.2, so $\nabla e_r f_k \to \nabla e_r f$ by Fact 2.6. Thus, differentiating (3.1) and evaluating at $x = 0$, we take the limit to find that the linear coefficient $b_k r/(2a_k + r)$ also converges, to some $b \in \mathbb{R}$. Finally, evaluating the same derivative at $x = 1$ and taking the limit, we have that the coefficient $a_k r/(2a_k + r)$ (which is nonnegative for all k) converges to ar for some $a \geq 0$. □

Theorem 3.1 leads one to ask which convex functions have quadratic functions as their Moreau envelopes. This question is answered by Proposition 3.1 below.

Proposition 3.1 *On $\mathbb{R}$, any convex quadratic function $f : \Gamma_0(\mathbb{R}) \to \overline{\mathbb{R}}$ defined by*

$$f(x) = \alpha x^2 + \beta x + \gamma, \ \ \alpha \geq 0$$

is a Moreau envelope of some convex function g, where g is either a quadratic function $g(x) = ax^2 + bx + c$, $a \geq 0$, or an indicator function $g(x) = \iota_{\{b\}}(x) + c$. Specifically, there exists a prox-parameter $r > 0$ such that the following hold.

(i) If $0 \leq \alpha < r/2$, then $g(x) = ax^2 + bx + c$, where

$$a = \frac{\alpha r}{r - 2\alpha}, \ \ b = \frac{\beta r}{r - 2\alpha}, \ \ c = \gamma + \frac{\beta^2}{2(r - 2\alpha)}.$$

(ii) If $\alpha = r/2$, then $g(x) = \iota_{\{b\}}(x) + c$, where

$$b = -\frac{\beta}{r}, \ \ c = \gamma - \frac{\beta^2}{2r}.$$

(iii) If $\alpha > r/2$, then $\nexists g \in \Gamma_0(\mathbb{R})$ such that $f = e_r g$.

Proof We need to show the form of g such that $f(x) = e_r g(x)\ \forall x \in \mathbb{R}$ for any choice of $\alpha \geq 0$, $\beta, \gamma \in \mathbb{R}$. By Theorem 3.1, we have that

$$e_r g(x) = \frac{ar}{2a + r} x^2 + \frac{br}{2a + r} x + c - \frac{b^2}{2(2a + r)}.$$

We equate the coefficients of f accordingly:

$$\alpha = \frac{ar}{2a + r}, \ \ \beta = \frac{br}{2a + r}, \ \ \gamma = c - \frac{b^2}{2(2a + r)}. \tag{5}$$

Solving the first of these expressions for a, we find $a = \alpha r/(r - 2\alpha)$. Notice that $\alpha = r/2$ is a point of interest.

(i) If $\alpha \in [0, r/2[$, there is a one-to-one correspondence with $a \in [0, \infty[$. Then b and c are found by solving the equations in (5).
(ii) If $\alpha = r/2$, this corresponds to $g(x) = \iota_{\{b\}}(x) + c$:

$$\begin{aligned}
g(x) &= \begin{cases} c, & x = b, \\ \infty, & x \neq b, \end{cases} \\
e_r g(x) &= \inf_y \left\{ g(y) + \frac{r}{2}(y - x)^2 \right\}, \\
&= g(b) + \frac{r}{2}(b - x)^2, \\
&= \frac{r}{2}x^2 - brx + \frac{r}{2}b^2 + c.
\end{aligned}$$

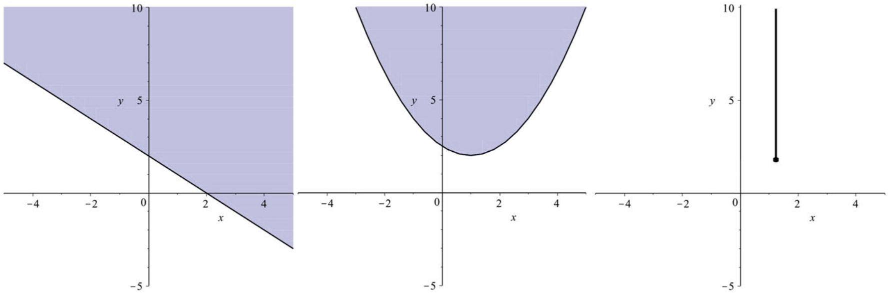

Fig. 1 Three possible general forms of the epigraph of $f(x)$

Equating $\beta = -br$ and $\gamma = rb^2/2 + c$, we find that $b = -\beta/r$, and that $c = \gamma - \beta^2/(2r)$. Then $f(x) = e_r g(x)$, where $g(x) = \iota_{\{b\}}(x) + c$.

(iii) Let $\alpha > r/2$. Suppose that $\exists g \in \Gamma_0(\mathbb{R})$ such that $f = e_r g$. By Proposition 2.2 and Fact 2.1, we have

$$\nabla e_r g(x) = r(\mathrm{Id} - J_{\partial g/r}).$$

Since $(\mathrm{Id} - J_{\partial g/r}) = J_{(\partial g/r)^{-1}}$ is nonexpansive, $\nabla e_r g$ is r-Lipschitz (see also Proposition 4.14). On the other hand, we have

$$\nabla e_r g(x) = \nabla f(x) = 2\alpha x + \beta,$$

which is L-Lipschitz only if $L \geq 2\alpha$. Hence, $r \geq 2\alpha$, which contradicts the condition that $\alpha > r/2$. Therefore, there does not exist $g \in \Gamma_0(\mathbb{R})$ such that $f = e_r g$. □

There are three possible epigraphical limits for the sequence defined in Theorem 3.1 (see Fig. 1). The first is $\operatorname{epi}(bx + c)$, the case where $a_k \to 0$. The second is $\operatorname{epi}(ax^2 + bx + c)$, the case where $a_k \to a > 0$. The third is $\operatorname{epi}(\iota_{\{b\}}(x) + c)$, the case where $a_k \to \infty$.

We present three examples here, to illustrate the three possibilities. In all three examples, we set $r = 1$.

Example 3.1 Define $f_k(x) = \left(1 + \frac{1}{k}\right) x^2 + \left(2 + \frac{1}{k}\right) x + \left(1 + \frac{1}{k}\right)$. Then

$$e_1 f_k(x) = \frac{k+1}{3k+2} x^2 + \frac{2k+1}{3k+2} x + \frac{2k^2 + 6k + 3}{k(6k+4)}.$$

Letting $k \to \infty$, we have $f_k \xrightarrow{e} f$ with

$$f(x) = (x+1)^2, \text{ and}$$
$$e_1 f(x) = \frac{1}{3}(x+1)^2.$$

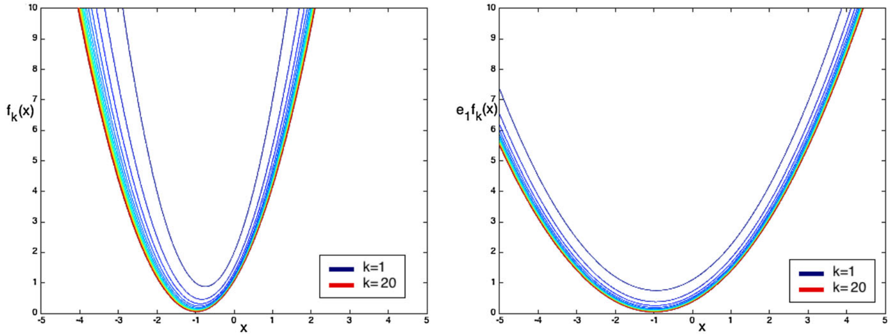

Fig. 2 Left: $f_k(x)$. Right: $e_1 f_k(x)$

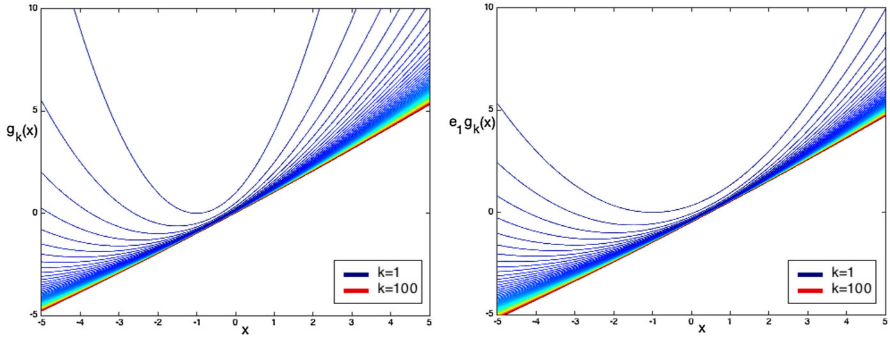

Fig. 3 Left: $g_k(x)$. Right: $e_1 g_k(x)$

Figure 2 shows the behavior of the graphs as a function of k.

Example 3.2 Define $g_k(x) = \frac{1}{k}x^2 + \left(1 + \frac{1}{k}\right) x + \frac{1}{k}$. Then

$$e_1 g_k(x) = \frac{1}{k+2}x^2 + \frac{k+1}{k+2}x + \frac{-k^2+3}{2k(k+2)}.$$

Letting $k \to \infty$, we have $g_k \xrightarrow{e} g$ with

$$g(x) = x, \text{ and}$$
$$e_1 g(x) = x - \frac{1}{2}.$$

Figure 3 shows the behavior of the graphs as a function of k.

Example 3.3 Define $h_k(x) = kx^2 + \frac{1}{k}x + \frac{1}{k}$. Then

$$e_1 h_k(x) = \frac{k}{2k+1}x^2 + \frac{1}{k(2k+1)}x + \frac{4k^2+2k-1}{2k^2(2k+1)}.$$

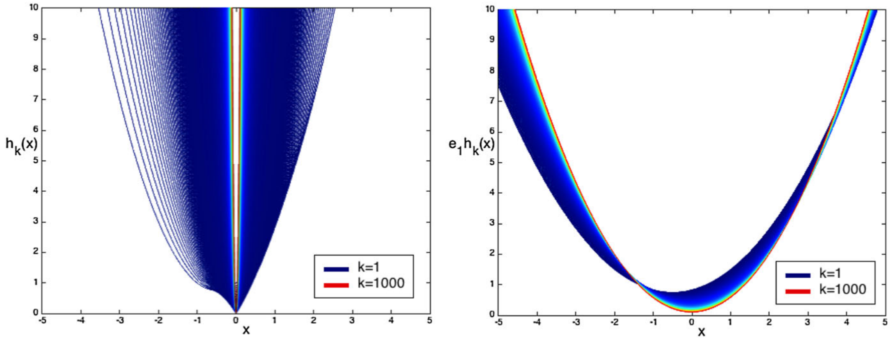

Fig. 4 Left: $h_k(x)$. Right: $e_1 h_k(x)$

Letting $k \to \infty$, we have $h_k \xrightarrow{e} h$ with

$$h(x) = \iota_{\{0\}}(x), \text{ and}$$
$$e_1 h(x) = \frac{1}{2}x^2.$$

Figure 4 shows the behavior of the graphs as a function of k.

4 Generalized Linear-Quadratic Functions on $\mathbb{R}^n$

Now we move on to finite-dimensional space. One natural goal that arises is that of unifying $f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle + c$ and $f(x) = \iota_{\{b\}}(x) + c$ in the more general setting of $\mathbb{R}^n$. To do so, we first need to establish several properties of monotone linear relations and generalized linear-quadratic functions.

4.1 Linear Relations and Generalized Linear-Quadratic Functions

Definition 4.1 An operator $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a *linear relation* if the graph of A is a linear subspace of $\mathbb{R}^n \times \mathbb{R}^n$.

Example 4.1 On $\mathbb{R}$, 'monotone' is equivalent to 'increasing.' Thus, a maximally monotone linear relation $A : \mathbb{R} \rightrightarrows \mathbb{R}$ is a straight line with nonnegative slope, and since it is a subspace, it must pass through the origin. There are three possibilities then: the x-axis, a line through the origin with positive slope, and the y-axis (see [13, Theorem 12.15] for details):

(i) $\operatorname{gra} A = \mathbb{R} \times \{0\} \Rightarrow A \equiv 0$,
(ii) $\operatorname{gra} A = \operatorname{span}\{(a, b)\},\ a, b \in \mathbb{R}\backslash\{0\} \Rightarrow A(x) = kx,\ k > 0$,
(iii) $\operatorname{gra} A = \{0\} \times \mathbb{R} \Rightarrow A = N_{\{0\}}$ (see Fig. 5).

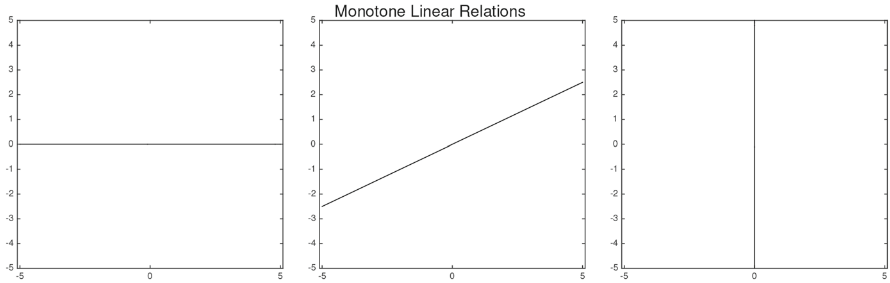

Fig. 5 Three possible forms of a monotone linear relation on $\mathbb{R}$

Definition 4.2 A *generalized linear-quadratic function* $p : \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined by a formula of the form

$$p(x) = \frac{1}{2}\langle x - a, A(x - a)\rangle + \langle b, x\rangle + c \quad \forall x \in \mathbb{R}^n,$$

where A is a linear relation, $a, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Our first question to consider here is: must the function p be well-defined?

Example 4.2 Define

$$A(x_1, x_2) = \{t(1, 1) : t \in \mathbb{R}\} \subseteq \mathbb{R}^2, \ \forall(x_1, x_2) \in \mathbb{R}^2.$$

Then A is a linear relation but not monotone, and $\langle x, Ax\rangle$ is not single-valued.

Proof It is elementary to show that A is a linear relation. Let $x_1 + x_2 \neq 0$. Then

$$\begin{aligned}\langle (x_1, x_2), A(x_1, x_2)\rangle &= \{\langle (x_1, x_2), t(1, 1)\rangle : t \in \mathbb{R}\} \\ &= \{t(x_1 + x_2) : t \in \mathbb{R}\} = \mathbb{R}.\end{aligned}$$

Therefore, $\langle x, Ax\rangle$ is not single-valued. Observe that A is not monotone. Indeed, set $t > 0$, and choose x_1, x_2 such that $x_1 + x_2 < 0$ and $t(1, 1) \in A(x_1, x_2)$. Note that $(0, 0) \in A(0, 0)$. Then

$$\begin{aligned}\langle (x_1, x_2) - (0, 0), A[(x_1, x_2) - (0, 0)]\rangle &= \langle (x_1, x_2), t(1, 1)\rangle \\ &= t(x_1 + x_2) < 0.\end{aligned}$$ □

The following fact says that when A is a monotone linear relation, p is well-defined.

Fact 4.1 [9, Proposition 3.2.1] *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a linear relation. Then A is monotone iff $\langle x, Ax\rangle \geq 0$ and $\langle x, Ax\rangle$ is single-valued for all $x \in$* dom A.

Our next question is: why do we write $\langle x - a, A(x - a)\rangle$, rather than the expanded form? It is because the two forms are not necessarily equivalent, as the following example demonstrates.

Example 4.3 Consider the example on $\mathbb{R}$ of $A = N_{\{0\}}$:

$$N_{\{0\}}(1-1) = \mathbb{R} \neq N_{\{0\}}(1) + N_{\{0\}}(-1) = \varnothing + \varnothing.$$

Fact 4.2 [9, Proposition 3.1.3] *The operator A is a linear relation if $\forall \alpha, \beta \in \mathbb{R}$ and $\forall x, y \in \mathbb{R}^n$, we have*

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay + A0.$$

Proposition 4.1 *Assume that A is a monotone linear relation. If both x and a are in* dom A*, or if* dom $A = \mathbb{R}^n$*, then*

$$\langle x - a, A(x-a)\rangle = \langle x, Ax\rangle - \langle x, Aa\rangle - \langle a, Ax\rangle + \langle a, Aa\rangle.$$

Proof When A is a monotone linear relation, $A0 \subset \operatorname{dom} A^{\perp}$. If $x, a \in \operatorname{dom} A$, then $\langle x, Ax\rangle$, $\langle a, Aa\rangle$, $\langle x, Aa\rangle$ and $\langle a, Ax\rangle$ are single-valued. It suffices to apply Fact 4.2. □

Definition 4.3 The *adjoint* A^* of a linear relation A is defined in terms of its graph:

$$\begin{aligned} \operatorname{gra} A^* &= \{(x^{**}, x^*) \in \mathbb{R}^n \times \mathbb{R}^n : (x^*, -x^{**}) \in (\operatorname{gra} A)^{\perp}\}, \\ &= \{(x^{**}, x^*) \in \mathbb{R}^n \times \mathbb{R}^n : \langle a, x^*\rangle = \langle a^*, x^{**}\rangle \ \forall (a, a^*) \in \operatorname{gra} A\}. \end{aligned}$$

Remark 4.1 We use the same asterisk notation for the adjoint of a linear relation as for the Fenchel conjugate of a function. The context makes it clear to which operation we are referring in every case; we hope that this double use is not distracting for the reader.

Definition 4.4 An operator A is *symmetric* if $\operatorname{gra} A \subseteq \operatorname{gra} A^*$. Equivalently, A is symmetric if

$$\langle x, y^*\rangle = \langle y, x^*\rangle \ \forall (x, x^*), (y, y^*) \in \operatorname{gra} A.$$

Example 4.4 The following are maximally monotone and symmetric linear relations.

(i) A positive semidefinite matrix $A : \mathbb{R}^n \to \mathbb{R}^n$, and its set-valued inverse $A^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. This follows from

$$(A^{-1})^* = (A^*)^{-1} = A^{-1}.$$

(ii) The normal cone operator $N_L : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, where $L \subset \mathbb{R}^n$ is a subspace. This is because

$$\operatorname{gra} N_L = L \times L^{\perp}, \quad \operatorname{gra} (N_L)^* = L \times L^{\perp}.$$

Definition 4.5 For a monotone linear relation $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, we define

(i) $q_A(x) = \begin{cases} \frac{1}{2}\langle x, Ax\rangle, & \text{if } x \in \operatorname{dom} A, \\ \infty, & \text{if } x \notin \operatorname{dom} A, \end{cases}$

(ii) $A_+ = \frac{1}{2}(A + A^*)$.

Remark 4.2 The framework of a generalized linear-quadratic function provides alternative notation that is sometimes more convenient or compact. For instance, for $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$ the indicator function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$,

$$f(x) = \iota_{\{a\}}(x) + c = \begin{cases} c, & \text{if } x = a, \\ \infty, & \text{if } x \neq a, \end{cases}$$

can be expressed as a generalized linear-quadratic function:

$$f(x) = q_{N_{\{0\}}}(x - a) + c,$$

where

$$N_{\{0\}}(x) = \begin{cases} \mathbb{R}^n, & \text{if } x = 0, \\ \varnothing, & \text{if } x \neq 0, \end{cases}$$

is a maximally monotone linear relation.

Proposition 4.2 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a linear relation. Suppose that at least one of the following is true:*

(i) A is symmetric, or
(ii) A is monotone.

Then q_A is an extended-real-valued function.

Proof (i) Let A be symmetric. Then by Definition 4.4 with $y = x$, we have

$$\langle x, y^*\rangle = \langle x, x^*\rangle \ \forall (x, x^*), (x, y^*) \in \operatorname{gra} A.$$

That is, $q_A(x) = \langle x, Ax\rangle = \langle x, x^*\rangle$ is single-valued for all $x \in \operatorname{dom} A$.

(ii) This is direct from Fact 4.1. □

4.2 Properties and Calculus of q_A

The generalized linear-quadratic function q_A is instrumental in establishing our final main result. In this section, we collect a number of properties of q_A under conditions such as maximal monotonicity and symmetry.

Lemma 4.1 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be symmetric. Then A^{-1} is symmetric.*

Proof By definition, A is symmetric iff

$$\langle x, Ay\rangle = \langle Ax, y\rangle \qquad \forall x, y \in \operatorname{dom} A. \tag{6}$$

Let $u \in Ay$, $v \in Ax$. Then $u, v \in \operatorname{ran} A = \operatorname{dom} A^{-1}$, and $x \in A^{-1}v$, $y \in A^{-1}u$. Substituting into (6), we have

$$\langle A^{-1}v, u\rangle = \langle v, A^{-1}u\rangle \qquad \forall u, v \in \operatorname{dom} A^{-1},$$

which is the definition of symmetry of A^{-1}. □

Lemma 4.2 *Let $A_1, A_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be maximally monotone linear relations. Then $A_1 + A_2$ is a maximally monotone linear relation. If, in addition, A_1 and A_2 are symmetric, then $A_1 + A_2$ is symmetric.*

Proof Since $\operatorname{dom} A_1$ and $\operatorname{dom} A_2$ are linear subspaces of $\mathbb{R}^n$, $\operatorname{dom} A_1 - \operatorname{dom} A_2$ is a closed subspace. By [9, Theorem 7.2.2], $A_1 + A_2$ is maximally monotone. Since $\operatorname{gra} A_1$ and $\operatorname{gra} A_2$ are linear subspaces, $\operatorname{gra}(A_1 + A_2)$ is a linear subspace. Hence, $A_1 + A_2$ is a linear relation. It remains to prove that $A_1 + A_2$ is symmetric when A_1 and A_2 are. Let $(x, x^*), (y, y^*) \in \operatorname{gra}(A_1 + A_2)$ be arbitrary. Since $\operatorname{dom}(A_1 + A_2) = \operatorname{dom} A_1 \cap \operatorname{dom} A_2$, we have $x, y \in \operatorname{dom} A_1$ and $x, y \in \operatorname{dom} A_2$. Then there exist $x_1^*, y_1^* \in \operatorname{ran} A_1$ and $x_2^*, y_2^* \in \operatorname{ran} A_2$ such that

(i) $(x, x_1^*), (y, y_1^*) \in \operatorname{gra} A_1$ and $(x, x_2^*), (y, y_2^*) \in \operatorname{gra} A_2$, and
(ii) $x_1^* + x_2^* = x^*$ and $y_1^* + y_2^* = y^*$.

This gives us that

$$(x, x_1^* + x_2^*) = (x, x^*) \in \operatorname{gra}(A_1 + A_2) \text{ and } (y, y_1^* + y_2^*) = (y, y^*) \in \operatorname{gra}(A_1 + A_2).$$

Now consider $\langle x, y^*\rangle - \langle y, x^*\rangle$:

$$\begin{aligned}
\langle x, y^*\rangle - \langle y, x^*\rangle &= \langle x, y_1^*\rangle + \langle x, y_2^*\rangle - \langle y, x_1^*\rangle - \langle y, x_2^*\rangle \\
&= (\langle x, y_1^*\rangle - \langle y, x_1^*\rangle) + (\langle x, y_2^*\rangle - \langle y, x_2^*\rangle) \\
&= (\langle x, y_1^*\rangle - \langle x, y_1^*\rangle) + (\langle x, y_2^*\rangle - \langle x, y_2^*\rangle) \\
&\qquad (A_1 \text{ is symmetric}) \qquad (A_2 \text{ is symmetric}) \\
&= 0.
\end{aligned}$$

Thus, for any $(x, x^*), (y, y^*) \in \operatorname{gra}(A_1 + A_2)$, we have $\langle x, y^*\rangle = \langle y, x^*\rangle$. Therefore, $A_1 + A_2$ is symmetric. □

Proposition 4.3 *Let A_1, A_2 be maximally monotone and symmetric linear relations on $\mathbb{R}^n$. Then*

$$A_1^* + A_2^* = (A_1 + A_2)^*.$$

Proof (⇒) By definition of symmetry, we have

$$\text{gra } A_1 \subseteq \text{gra } A_1^*. \tag{7}$$

Since A_1 is maximally monotone, A_1^* is also maximally monotone by [24, Corollary 5.11]. Then (7) is actually an equality, and we have

$$A_1 = A_1^*, \text{ and similarly } A_2 = A_2^*. \tag{8}$$

Then by definition of adjoint,

$$\begin{aligned}
\text{gra}((A_1 + A_2)^*) &= \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n : (x^*, -x) \in (\text{gra}(A_1 + A_2))^\perp\} \\
&= \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n : (x^*, -x) \in (\text{gra}(A_1^* + A_2^*))^\perp\} \quad \text{by (8)} \\
&= \text{gra}((A_1^* + A_2^*)^*).
\end{aligned}$$

Once more by definition of symmetry, we have $\text{gra}(A_1^* + A_2^*) \subseteq \text{gra}((A_1^* + A_2^*)^*)$. Therefore, $\text{gra}(A_1^* + A_2^*) \subseteq \text{gra}((A_1 + A_2)^*)$.

(⇐) We have $\text{gra}((A_1 + A_2)^*) = \text{gra}((A_1^* + A_2^*)^*)$ from above, and invoking symmetry yields $\text{gra}((A_1^* + A_2^*)^*) \subseteq \text{gra}((A_1^* + A_2^*)^{**})$. Since we are in $\mathbb{R}^n$, $\text{gra}((A_1^* + A_2^*)^{**})$ is closed. Thus $\text{gra}((A_1^* + A_2^*)^{**}) = \text{gra}(A_1^* + A_2^*)$. Therefore, $\text{gra}((A_1 + A_2)^*) \subseteq \text{gra}(A_1^* + A_2^*)$.[1] □

Proposition 4.4 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone linear relation. Then*

(i) q_A is well-defined, i.e., $q_A : \mathbb{R}^n \to \overline{\mathbb{R}}$,
(ii) q_A is convex,
(iii) $q_A = q_{A_+}$, and
(iv) $\partial q_A = A_+$.

Proof (i) This is direct from Proposition 4.2.
(ii) See [7, Proposition 2.3].
(iii) Since A is maximally monotone, A^* is also maximally monotone [24, Corollary 5.11]. By definition of A^*, we have $\langle x, Ax\rangle = \langle A^*x, x\rangle = \langle x, A^*x\rangle$. Then

$$\begin{aligned}
q_A(x) = \frac{1}{2}\langle x, Ax\rangle &= \frac{1}{2}\left(\frac{\langle x, Ax\rangle + \langle x, A^*x\rangle}{2}\right) \\
&= \frac{1}{2}\left\langle x, \frac{Ax + A^*x}{2}\right\rangle \\
&= \frac{1}{2}\langle x, A_+x\rangle = q_{A_+}(x).
\end{aligned}$$

(iv) Since A is maximally monotone, A^* is as well, hence A_+ is as well. Then

$$\partial q_A = \partial q_{A_+} = A_+ = \frac{1}{2}(A + A^*).$$

□

[1] Thank you to Dr. Walaa Moursi for contributing to this Proof.

Lemma 4.3 *Let* $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be a maximally monotone and symmetric linear relation. Then* $\partial q_A = A$.

Proof Since A is symmetric, $A = A^*$. The result follows from Proposition 4.4(iv). □

Corollary 4.1 *Let* $A_1, A_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be maximally monotone and symmetric linear relations such that* $q_{A_1} = q_{A_2}$. *Then* $A_1 = A_2$.

Proof This follows from $\partial q_{A_1} = A_1$, $\partial q_{A_2} = A_2$. □

Remark 4.3 The maximal monotonicity condition of Corollary 4.1 is necessary. As a counterexample, consider a monotone selection S of A and set

$$A_1 = S, \qquad A_2 = S + A0.$$

Then $q_{A_1} = q_{A_2}$, but $A_1 \neq A_2$ unless $A0 = \{0\}$.

Proposition 4.5 *Let* $A_1, A_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be monotone linear relations. Then*

$$q_{A_1} + q_{A_2} = q_{A_1 + A_2}.$$

In addition, if $\operatorname{dom} A_1 \subseteq \operatorname{dom} A_2$ *and* $A_1 - A_2$ *is monotone, then*

$$q_{A_1} - q_{A_2} = q_{A_1 - A_2}.$$

Proof By definition, we have $q_{A_1}(x) = \frac{1}{2}\langle x, A_1 x\rangle$ if $x \in \operatorname{dom} A_1$, ∞ otherwise. Similarly, $q_{A_2}(x) = \frac{1}{2}\langle x, A_2 x\rangle$ if $x \in \operatorname{dom} A_2$, ∞ otherwise. Thus,

$$\begin{aligned}(q_{A_1} + q_{A_2})(x) &= \begin{cases} \frac{1}{2}\langle x, (A_1 + A_2)x\rangle, & \text{if } x \in \operatorname{dom} A_1 \cap \operatorname{dom} A_2, \\ \infty, & \text{if } x \notin \operatorname{dom} A_1 \cap \operatorname{dom} A_2, \end{cases} \\ &= q_{A_1 + A_2}(x).\end{aligned}$$

Now suppose that $\operatorname{dom} A_1 \subseteq \operatorname{dom} A_2$ and $A_1 - A_2$ is monotone. Then, for $x \in \operatorname{dom} A_2$ with $x \in \operatorname{dom} A_1$, we have that $q_{A_1} - q_{A_2}$ is single-valued, so that

$$q_{A_1}(x) - q_{A_2}(x) = q_{A_1 - A_2}(x).$$

When $x \notin \operatorname{dom} A_1$, we have

$$q_{A_1}(x) - q_{A_2}(x) = \infty - q_{A_2}(x) = \infty.$$

Now

$$\operatorname{dom} q_{A_1 - A_2} = \operatorname{dom}(A_1 - A_2) = \operatorname{dom} A_1 \cap \operatorname{dom} A_2 = \operatorname{dom} A_1,$$

so that

$$q_{A_1-A_2}(x)=\infty \text{ when } x\notin \operatorname{dom} A_1.$$

Therefore,

$$q_{A_1}-q_{A_2}=q_{A_1-A_2}.$$

$\square$

The condition $\operatorname{dom} A_1 \subseteq \operatorname{dom} A_2$ is necessary for $q_{A_1}-q_{A_2}=q_{A_1-A_2}$. The following example shows that Proposition 4.5 can fail if $\operatorname{dom} A_1 \nsubseteq \operatorname{dom} A_2$.

Example 4.5 Let $A_1, A_2 : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ be maximally monotone linear relations given by

$$A_1=\mathrm{Id},\quad A_2=N_{\mathbb{R}\times\{0\}},$$

where

$$N_{\mathbb{R}\times\{0\}}(x,y)=\begin{cases}\{0\}\times\mathbb{R}, & \text{if } y=0,\\ \varnothing, & \text{if } y\neq 0.\end{cases}$$

Then

$$(A_1-A_2)(x,y)=\begin{cases}(x,0)-\{0\}\times\mathbb{R}, & \text{if } y=0,\\ \varnothing, & \text{if } y\neq 0,\end{cases}$$

is a maximally monotone linear relation. We have

$$q_{A_1}(0,1)-q_{A_2}(0,1)=1/2-\infty=-\infty,$$

but

$$q_{A_1-A_2}(0,1)=\infty$$

because $(0,1)\notin \operatorname{dom}(A_1-A_2)$.

Proposition 4.6 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone and symmetric linear relation. Then the following are equivalent:*

(i) $q_A(x)=0$;
(ii) $x\in \operatorname{argmin} q_A$;
(iii) $0\in \partial q_A(x)$;
(iv) $0\in Ax$;
(v) $x\in A^{-1}0$.

Proof (i)⇒(ii) Let $q_A(x)=0$. Since A is monotone, by Fact 4.1 we have $q_A(y)\geq 0\ \forall y\in\mathbb{R}^n$. Hence,

$$\min_{y\in\mathbb{R}^n} q_A(y)=0=q_A(x)\Rightarrow x\in\operatorname{argmin} q_A.$$

(ii)⇒(iii) This is direct from Fermat's Theorem.
(iii)⇒(iv) Let $0\in\partial q_A(x)$. Since A is symmetric and maximally monotone, by Lemma 4.3 we have $\partial q_A(x)=A(x)$. Therefore, $0\in Ax$.
(iv)⇒(i) Let $0\in Ax$. Then, since $q_A(x)$ is single-valued by Fact 4.1, we have

$$q_A(x)=\frac{1}{2}\langle x,Ax\rangle=\frac{1}{2}\langle x,0\rangle=0.$$

(iv)⇔(v) We have $0\in Ax\Leftrightarrow x\in A^{-1}0$. □

4.3 The Fenchel Conjugate of q_A

Conjugacy plays a vital role in convex analysis [25, Chapter X]. One often finds it beneficial to work temporarily in a dual space in order to solve a problem, then return the answer to the primal space. In this section, we explore the Fenchel conjugate of q_A. We show that the set-valued inverse A^{-1} is convenient for computing q_A^*.

Proposition 4.7 *Let $A:\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ be a maximally monotone and symmetric linear relation. Then $q_A^*=q_{A^{-1}}$, that is,*

$$q_A^*(y)=\begin{cases}\frac{1}{2}\langle y,A^{-1}y\rangle, & \text{if } y\in\operatorname{ran}A,\\ \infty, & \text{if } y\notin\operatorname{ran}A.\end{cases}$$

Consequently,

$$q_A^{**}(x)=\begin{cases}\frac{1}{2}\langle x,Ax\rangle, & \text{if } x\in\operatorname{ran}A^{-1},\\ \infty, & \text{if } x\notin\operatorname{ran}A^{-1}.\end{cases}$$

Thus, $q_A^{**}=q_A$, so q_A is lsc and convex.

Proof By the definition of q_A^*, we have

$$\begin{aligned}q_A^*(y)&=\sup_x\{\langle y,x\rangle-q_A(x)\}\\&=\sup_{x\in\operatorname{dom}A}\left\{\langle y,x\rangle-\frac{1}{2}\langle x,Ax\rangle\right\}.\end{aligned}\tag{9}$$

Consider two cases.

(i) Let $y \in \operatorname{ran} A$. Then the solution to the supremum in (9) is $\bar{x}$ such that

$$0 \in \partial \left[\langle y, \bar{x} \rangle - \frac{1}{2} \langle \bar{x}, A\bar{x} \rangle \right].$$

This gives $y \in A\bar{x}$, hence, $\bar{x} \in A^{-1}y$. Then

$$q_A^*(y) = \langle y, A^{-1}y \rangle - \frac{1}{2} \langle A^{-1}y, y \rangle = \frac{1}{2} \langle y, A^{-1}y \rangle.$$

(ii) Let $y \notin \operatorname{ran} A$. Note that since $\operatorname{ran} A$ is closed and convex, by [11, Corollary 11.4.2] there exist $z \in \mathbb{R}^n$ and $r \in \mathbb{R}$ such that

$$\langle z, y \rangle > r \geq \sup_{x \in \operatorname{ran} A} \langle x, z \rangle.$$

Since $\operatorname{ran} A$ is a subspace, we have $0 \in \operatorname{ran} A$ and $r \geq 0$. Also since $\operatorname{ran} A$ is a subspace, we have $kx \in \operatorname{ran} A\ \forall x \in \operatorname{ran} A$, $\forall k \in \mathbb{R}$. Hence, $r \geq k\langle x, z \rangle$ for all $x \in \operatorname{ran} A$ and for all $k \in \mathbb{R}$. Thus, $\langle x, z \rangle = 0$ for all $x \in \operatorname{ran} A$ (otherwise, for $\langle x, z \rangle \neq 0$ one could choose k such that $k\langle x, z \rangle > r$). Then $\sup_{x \in \operatorname{ran} A} \langle x, z \rangle = 0$, hence $\langle z, y \rangle > 0$. Noting that

$$\sup_{k>0} \left\{ \langle y, kz \rangle - \frac{1}{2} \langle kz, A(kz) \rangle \right\} = \sup_{k>0} \langle y, kz \rangle = \infty, \tag{10}$$

and that the supremum of (9) is greater than or equal to that of (10), we have $q_A^*(y) = \infty$. □

Corollary 4.2 *Let A be positive definite and nonsingular. Then $q_A^* = q_{A^{-1}}$, where A^{-1} is the classical inverse.*

Corollary 4.3 *Let A be positive semidefinite. Then $q_A^* = q_{A^{-1}}$ where*

$$q_{A^{-1}}(x) = \begin{cases} \frac{1}{2}\langle x, A^{-1}x \rangle, & \text{if } x \in \operatorname{ran} A, \\ \infty, & \text{if } x \notin \operatorname{ran} A, \end{cases}$$

and A^{-1} is the set-valued inverse.

Proposition 4.8 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone and symmetric linear relation. Then $q_A^* = q_A$ iff $A = A^{-1}$, iff $A = \operatorname{Id}$.*

Proof The proof that $A = A^{-1}$ iff $A = \operatorname{Id}$ is found in [6, Proposition 2.8]. In the sequel, we prove that $q_A^* = q_A$ iff $A = A^{-1}$.

($\Leftarrow$) Suppose that $A = A^{-1}$. Then we see immediately by Corollary 4.3 that $q_A^* = q_A$.
($\Rightarrow$) Suppose that $q_A^* = q_A$. Then by Lemma 4.3, we have that $\partial q_A = A = \partial q_A^*$. By Corollary 4.3, we find $\partial q_A^* = A^{-1}$. Therefore, $A = A^{-1}$. □

Proposition 4.9 *Let $A_i : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone and symmetric linear relation for each $i \in \{1, \dots, m\}$. Then the infimal convolution defined $f = q_{A_1} \square \cdots \square q_{A_m}$ is a generalized linear-quadratic function, and*

$$\partial f = \left(\sum_{i=1}^{n} A_i^{-1} \right)^{-1},$$

which is the parallel sum *of A_i.*

Proof By Lemmas 4.1 and 4.2, we have that $A_1^{-1} + \cdots + A_m^{-1}$ is a maximally monotone and symmetric linear relation. By Corollary 4.3, $q^*_{A_i} = q_{A_i^{-1}}$, and since $0 \in \bigcap_{i=1}^m \operatorname{ri} \operatorname{ran} A_i$, [11, Theorem 16.4] gives

$$\begin{aligned} q_{(A_1^{-1}+\cdots+A_m^{-1})^{-1}} = q^*_{A_1^{-1}+\cdots+A_m^{-1}} &= \left(q_{A_1^{-1}} + \cdots + q_{A_m^{-1}} \right)^* \\ &= \left(q^*_{A_1} + \cdots + q^*_{A_m} \right)^* \\ &= q_{A_1} \square \cdots \square q_{A_m} = f. \end{aligned}$$

Therefore, by Lemma 4.3, we have the statement of the proposition. □

Proposition 4.10 *Let $A_1 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone and symmetric linear relation, and $A_2 : \mathbb{R}^n \to \mathbb{R}^n$ be positive definite. Define the function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ by*

$$q_{A_2} \square h = q_{A_1}. \tag{11}$$

Then for every $x \in \mathbb{R}^n$,

$$h(x) = \left(q^*_{A_1} - q^*_{A_2} \right)^* (x) = \sup_y \left\{ q_{A_1}(x+y) - q_{A_2}(y) \right\}. \tag{12}$$

Consequently, when $A_1^{-1} - A_2^{-1}$ is monotone, one has

$$\partial h = \left(A_1^{-1} - A_2^{-1} \right)^{-1}, \tag{13}$$

which is the star-difference *of A_1 and A_2.*

Proof Taking the Fenchel conjugate of (11) yields $h^* = q^*_{A_1} - q^*_{A_2}$. Then by the Toland–Singer duality, we have (12). Observe that $A_1^{-1} - A_2^{-1}$ is maximally monotone because of the following. We have

$$\begin{aligned} &\operatorname{dom}(A_1^{-1} - A_2^{-1}) = \operatorname{ran} A_1 \cap \operatorname{ran} A_2 = \operatorname{ran} A_1 = \operatorname{dom} A_1^{-1}, \text{ and} \\ &\quad (A_1^{-1} - A_2^{-1})(0) = A_1^{-1}(0). \end{aligned}$$

Because A_1^{-1} is maximally monotone, $(\operatorname{dom} A_1^{-1})^\perp = A_1^{-1}(0)$. Then by [8, Fact 2.4(v)], $A_1^{-1} - A_2^{-1}$ is maximally monotone. We have that $A_1^{-1} - A_2^{-1}$ is a maximally monotone and symmetric linear relation, and that $q^*_{A_i} = q_{A_i^{-1}}$ and $q_{A_1^{-1}} - q_{A_2^{-1}} = q_{A_1^{-1}-A_2^{-1}}$. Therefore, we have (13). □

Remark 4.4 This result generalizes that of Hiriart-Urruty [12], because the matrix $A_1^{-1} - A_2^{-1}$ need not be positive definite [12, Example 2.7].

4.4 Relating the Set-Valued Inverse and the Moore–Penrose Inverse

The set-valued inverse A^{-1} of a linear mapping and the Moore–Penrose inverse $A^\dagger$ both have their uses. For properties of $A^\dagger$, see [26, p. 423–428]. In this section, we show how the two inverses are closely related. We also include a description of the Moore–Penrose inverse for a particular mapping, the orthogonal projector.

Proposition 4.11 *The following hold.*

(i) Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a linear mapping. Then

$$A^{-1}x = \begin{cases} A^\dagger x + A^{-1}0, & \textit{if } x \in \operatorname{ran} A, \\ \varnothing, & \textit{if } x \notin \operatorname{ran} A. \end{cases}$$

(ii) Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be maximally monotone. Then

$$A^{-1}x = A^\dagger x + N_{\operatorname{dom} A^{-1}} = \begin{cases} A^\dagger x + (\operatorname{ran} A)^\perp, & \textit{if } x \in \operatorname{ran} A, \\ \varnothing, & \textit{if } x \notin \operatorname{ran} A. \end{cases}$$

(iii) Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be monotone, symmetric and linear. Then

$$A^{-1} = P_{\operatorname{ran} A} A^\dagger P_{\operatorname{ran} A} + N_{\operatorname{dom} A^{-1}}.$$

Proof (i) Since $AA^\dagger = P_{\operatorname{ran} A}$ $\forall x \in \operatorname{ran} A$, it holds that

$$AA^\dagger x = P_{\operatorname{ran} A} x = x \Rightarrow A^\dagger x \in A^{-1}x.$$

Since $A^{-1}x = x^* + A^{-1}0$ for every $x^* \in Ax$, we have

$$A^{-1} = A^\dagger + A^{-1}0 \text{ on } \operatorname{ran} A.$$

(ii) Since A is maximally monotone, A^{-1} is as well, and

$$(\operatorname{dom} A^{-1})^\perp = A^{-1}0.$$

Applying part (i) completes the proof.

(iii) If A is maximally monotone and linear, then

$$\begin{aligned} \operatorname{ran} A^\dagger = \operatorname{ran} A^\top = \operatorname{ran} A^* = \operatorname{ran} A, \text{ and} \\ N_{\operatorname{dom} A^{-1}}(x) = (\operatorname{dom} A^{-1})^\perp = A^{-1}0. \end{aligned}$$

This implies that on $\operatorname{ran} A = \operatorname{dom} A^{-1}$,

$$P_{\operatorname{ran} A} A^{\dagger} P_{\operatorname{ran} A} = A^{\dagger}.$$

Now we apply part (ii). Let $x \in \mathbb{R}^n$, $u = P_{\operatorname{ran} A} x$. Denote $P_{\operatorname{ran} A} A^{\dagger} P_{\operatorname{ran} A}$ by L. Using $AA^{\dagger}A = A$, L is monotone because

$$\begin{aligned}
\langle x, Lx\rangle &= \langle P_{\operatorname{ran} A} x, A^{\dagger} P_{\operatorname{ran} A} x\rangle \\
&= \langle Au, A^{\dagger} Au\rangle \\
&= \langle u, AA^{\dagger} Au\rangle \\
&= \langle u, Au\rangle \\
&\geq 0.
\end{aligned}$$

We have that L is symmetric, because

$$(A^{\dagger})^* = (A^*)^{\dagger} = A^{\dagger}.$$

□

In [14, Exercise 3.13], for a linear mapping A, one has

$$A^{\dagger} = P_{\operatorname{ran} A^*} A^{-1} P_{\operatorname{ran} A}. \tag{14}$$

For a set $\Omega \subset \mathbb{R}^n$, define the indicator mapping $\triangle\colon \mathbb{R}^n \to \mathbb{R}^n$ of Ω relative to $\mathbb{R}^n$ by

$$\triangle_{\Omega}(x) = \begin{cases} 0 \in \mathbb{R}^n, & \text{if } x \in \Omega, \\ \varnothing, & \text{if } x \notin \Omega. \end{cases}$$

(See, e.g., [27].) Combining (14) and Proposition 4.11, we obtain a complete relationship between A^{-1} and $A^{\dagger}$.

Corollary 4.4 *(i) When A is a linear mapping on $\mathbb{R}^n$,*

$$A^{-1} = A^{\dagger} + \triangle_{\operatorname{dom} A^{-1}} + A^{-1}0 \quad \text{and} \quad A^{\dagger} = P_{\operatorname{ran} A^*} A^{-1} P_{\operatorname{ran} A}.$$

(ii) If, in addition, A is maximally monotone, then

$$A^{-1} = A^{\dagger} + N_{\operatorname{dom} A^{-1}} \quad \text{and} \quad A^{\dagger} = P_{\operatorname{ran} A} A^{-1} P_{\operatorname{ran} A}.$$

Corollary 4.4(i) is a corollary of Proposition 4.11.

Corollary 4.5 *Let A be a maximally monotone and symmetric linear relation. Then*

$$(q_A)^* = \begin{cases} q_{A^{\dagger}}, & \text{if } x \in \operatorname{ran} A, \\ \infty, & \text{if } x \notin \operatorname{ran} A. \end{cases}$$

In the sequel, we present the Moore–Penrose inverse of the projector mapping. We remind the reader of the definition.

Definition 4.6 Let $C \subseteq \mathbb{R}^n$ be closed and convex. The *projection* of a point x onto C is defined

$$P_C x = \{y \in C : \|y - x\| = d_C(x)\},$$

where d_C is the *distance function:* $d_C(x) = \inf_{y \in C} \|y - x\|$. We call P_C the projection operator.

Proposition 4.12 *Let P_L be the orthogonal projector onto a subspace L of $\mathbb{R}^n$. Then the Moore–Penrose generalized inverse $P_L^\dagger = P_L$.*

Proof By [28, Theorem 10.5], P_L is idempotent ($P_L^2 = P_L$) and Hermitian ($P_L^* = P_L$). Since $P_L^\dagger$ is the unique operator that satisfies the four Moore–Penrose equations, it is a simple matter to verify that each of them is satisfied by $P_L^\dagger = P_L$:

(i) $AA^\dagger A = A : P_L P_L P_L = P_L$,
(ii) $A^\dagger A A^\dagger = A^\dagger : P_L P_L P_L = P_L$,
(iii) $(AA^\dagger)^* = AA^\dagger : (P_L P_L)^* = P_L P_L \Rightarrow P_L^* = P_L$,
(iv) $(A^\dagger A)^* = A^\dagger A : (P_L P_L)^* = P_L P_L \Rightarrow P_L^* = P_L$.

Therefore, $P_L^\dagger = P_L$. □

Corollary 4.6 *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}} : x \mapsto \frac{1}{2}\langle x, P_L x\rangle$, where L is a subspace. Then*

$$f^*(x^*) = \begin{cases} \frac{1}{2}\langle x^*, P_L x^*\rangle, & \textit{if } x^* \in L, \\ \infty, & \textit{if } x^* \notin L \end{cases} = \begin{cases} \frac{1}{2}\langle x^*, x^*\rangle, & \textit{if } x^* \in L, \\ \infty, & \textit{if } x^* \notin L. \end{cases}$$

Proof Using Proposition 4.12 and Corollary 4.5, the proof is immediate. □

We end this section with an extension of Rockafellar's and Wets' result [13, Example 11.10].

Proposition 4.13 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a symmetric and monotone linear relation, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Suppose that*

$$f(x) = q_A(x) + \langle b, x\rangle + c.$$

Then, for every $y \in \mathbb{R}^n$, the Fenchel conjugate of f is

$$f^*(y) = q_{A^{-1}}(y - b) - c = \begin{cases} \frac{1}{2}\langle y - b, A^\dagger(y - b)\rangle - c, & \textit{if } y - b \in \operatorname{ran} A, \\ \infty, & \textit{if } y - b \notin \operatorname{ran} A. \end{cases}$$

Proof Applying Theorem 4.7, we have $\forall y \in \mathbb{R}^n$,

$$f^*(y) = (q_A)^*(y-b) - c = q_{A^{-1}}(y-b) - c.$$

By Proposition 4.11(ii),

$$A^{-1} = A^\dagger + N_{\operatorname{ran} A}.$$

This gives

$$f^*(y) = \begin{cases} \frac{1}{2}\langle y-b, A^\dagger(y-b)\rangle - c, & \text{if } y-b \in \operatorname{ran} A, \\ \infty, & \text{if } y-b \notin \operatorname{ran} A. \end{cases}$$

□

Remark 4.5 In [13, Example 11.10], the authors assume that $A \in \mathbb{R}^{n\times n}$, i.e., A is a linear operator. In Example 4.13, A is a linear relation.

4.5 Characterizations of Moreau Envelopes

In this section, we present several useful properties of Moreau envelopes of convex functions. We identify the form of the Moreau envelope for quadratic functions, and provide a characterization of Moreau envelopes that involves Lipschitz continuity. This leads to a sum rule for Moreau envelopes of convex functions. Then we follow up with one of the main results of this paper: Theorem 4.2 is a characterization that relates generalized linear-quadratic functions to nonexpansive mappings.

Proposition 4.14 *Let $f \in \Gamma_0(\mathbb{R}^n)$. Then $f = e_r g$ for some $g \in \Gamma_0(\mathbb{R}^n)$ iff ∇f is r-Lipschitz.*

Proof ($\Rightarrow$) Let $f = e_r g$ for some $g \in \Gamma_0(\mathbb{R}^n)$. Then by Proposition 2.2(ii), we have

$$\nabla f = r(\mathrm{Id} - \mathrm{Prox}_g^r).$$

Let $x, y \in \mathbb{R}^n$. Then

$$\begin{aligned} \|\nabla f(x) - \nabla f(y)\| &= r\|x - \mathrm{Prox}_g^r(x) - y + \mathrm{Prox}_g^r(y)\| \\ &= r\left\|\frac{1}{r}\mathrm{Prox}_{g^*}^{\frac{1}{r}}(rx) - \frac{1}{r}\mathrm{Prox}_{g^*}^{\frac{1}{r}}(ry)\right\| \quad \text{(Proposition 2.2(}v\text{))} \\ &= \left\|\mathrm{Prox}_{g^*}^{\frac{1}{r}}(rx) - \mathrm{Prox}_{g^*}^{\frac{1}{r}}(ry)\right\| \\ &= \|J_{r\partial g^*}(rx) - J_{r\partial g^*}(ry)\| \quad \text{(Fact 2.1)} \\ &\le \|rx - ry\| = r\|x-y\|. \end{aligned}$$

Therefore, ∇f is r-Lipschitz.

($\Leftarrow$) Let ∇f be r-Lipschitz. Then by Fact 2.12, $\frac{1}{r}\nabla f$ is firmly nonexpansive. By Fact 2.9, we have that $\frac{1}{r}\nabla f = (\mathrm{Id} + A)^{-1}$ for some A monotone. Since $f \in \Gamma_0(\mathbb{R}^n)$, A is maximally cyclically monotone by Fact 2.10. Thus, $A = \partial g$ for some $g \in \Gamma_0(\mathbb{R}^n)$ by Fact 2.11. Hence,

$$\nabla f = r(\mathrm{Id} + \partial g)^{-1} = \left[(\mathrm{Id} + \partial g) \circ \left(\frac{\mathrm{Id}}{r}\right)\right]^{-1}.$$

Then we have

$$\begin{aligned}\partial f^* = (\nabla f)^{-1} &= (\mathrm{Id} + \partial g)\left(\frac{\mathrm{Id}}{r}\right)\\ &= \frac{\mathrm{Id}}{r} + \partial g\left(\frac{\mathrm{Id}}{r}\right),\end{aligned}$$

so that

$$f^* = \frac{q}{r} + rg\left(\frac{\cdot}{r}\right) + c,\ c \in \mathbb{R}.$$

Taking the conjugate of both sides yields

$$\begin{aligned}f &= \left[\frac{q}{r} + rg\left(\frac{\cdot}{r}\right) + c\right]^*\\ &= \left(\frac{q}{r}\right)^* \Box \left[rg\left(\frac{\cdot}{r}\right) + c\right]^*\\ &= (rq) \Box (rg^* - c)\\ &= e_r(rg^* - c),\end{aligned}$$

where $g^* \in \Gamma_0(\mathbb{R}^n)$. $\square$

Corollary 4.7 *Let* $r_1, r_2 > 0,\ g, h \in \Gamma_0(\mathbb{R}^n)$. *Then*

$$e_{r_1} g + e_{r_2} h = e_{r_1+r_2} f \tag{15}$$

for some $f \in \Gamma_0(\mathbb{R}^n)$. *Specifically,*

$$f(x) = \sup_{v \in \mathbb{R}^n} \left\{\left[e_{r_1} g(x+v) - r_1 q(v)\right] + \left[e_{r_2} h(x+v) - r_2 q(v)\right]\right\}.$$

Proof Denote $e_{r_1} g, e_{r_2} h$ by $\bar{g}, \bar{h}$, respectively. Then by Proposition 4.14, $\nabla \bar{g}$ is r_1-Lipschitz and $\nabla \bar{h}$ is r_2-Lipschitz. Hence, $\nabla \bar{f}$ is (r_1+r_2)-Lipschitz, where $\bar{f} = \bar{g} + \bar{h} = e_{r_1} g + e_{r_2} h$. Applying Proposition 4.14 again, we have that $\bar{f} = e_{r_1+r_2} f$ for some $f \in \Gamma_0(\mathbb{R}^n)$. Now to find f, we apply the Fenchel conjugate to (15):

$$\begin{aligned} f^* + \frac{q}{r_1 + r_2} &= (e_{r_1} g + e_{r_2} h)^* \\ f &= \left[(e_{r_1} g + e_{r_2} h)^* - \frac{q}{r_1 + r_2} \right]^* . \end{aligned}$$

By the Toland–Singer duality for the conjugate of a difference [14, Corollary 14.19], we obtain that for every $x \in \mathbb{R}^n$,

$$\begin{aligned} f(x) &= \sup_{v \in \mathbb{R}^n} \left\{ (e_{r_1} g + e_{r_2} h)^{**}(x+v) - \left(\frac{q}{r_1 + r_2} \right)^* (v) \right\} \\ &= \sup_{v \in \mathbb{R}^n} \left\{ (e_{r_1} g + e_{r_2} h)(x+v) - (r_1 + r_2) q(v) \right\} \\ &= \sup_{v \in \mathbb{R}^n} \left\{ \left[e_{r_1} g(x+v) - r_1 q(v) \right] + \left[e_{r_2} h(x+v) - r_2 q(v) \right] \right\} . \end{aligned}$$

□

Remark 4.6 Corollary 4.7 gives us that for $r > 0$ and $f \in \Gamma_0(\mathbb{R}^n)$,

$$f = \left[(e_r f)^* - \frac{q}{r} \right]^* .$$

Therefore, by the Toland–Singer duality, for every $x \in \mathbb{R}^n$ we have

$$f(x) = \sup_{v \in \mathbb{R}^n} \{ e_r f(x+v) - r q(v) \}.$$

This is the Hiriart-Urruty deconvolution [12].

Proposition 4.15 *Let $A \in S_+^n$. Then the following are equivalent:*

(i) A is nonexpansive, i.e., $\|Ax - Ay\| \le \|x - y\|$ for all $x, y \in \mathbb{R}^n$;
(ii) A is firmly nonexpansive, i.e., $\|Ax - Ay\|^2 \le \langle x - y, Ax - Ay \rangle$ for all $x, y \in \mathbb{R}^n$;
(iii) $A = (P + \mathrm{Id})^{-1}$ for some maximally monotone linear relation P.

Proof Denote the eigenvalues of A as $\lambda_1, \lambda_2, \ldots, \lambda_n$. Since $A \in S_+^n$, all eigenvalues are real and nonnegative (see [26] Section 7.6).

(i)⇔(ii) Suppose that statement (i) is true. Then, letting $z = x - y$ and squaring both sides, we have

$$\begin{aligned} &\|Az\|^2 \le \|z\|^2 \\ &\Leftrightarrow \langle z, A^\top A z \rangle \le \langle z, z \rangle \\ &\Leftrightarrow \langle z, A^2 z \rangle \le \langle z, z \rangle \\ &\Leftrightarrow \langle z, (\mathrm{Id} - A^2) z \rangle \ge 0 \text{ for all } z \in \mathbb{R}^n. \end{aligned}$$

The inequality above is equivalent to the statement $\mathrm{Id} - A^2 \in S_+^n$, so

$$1 - \lambda_i^2 \ge 0 \ \forall i \in \{1, 2, \ldots, n\}.$$

Since $A \in S^n_+$, we have $\lambda_i \geq 0$ for all i. Hence, statement (i) is equivalent to the following:

$$0 \leq \lambda_i \leq 1 \text{ for all } i \in \{1, 2, \dots, n\}. \tag{16}$$

Now suppose that statement (ii) is true. This gives

$$\begin{aligned}
&\langle z, A^\top Az\rangle \leq \langle z, Az\rangle \\
&\Leftrightarrow \langle z, A^2 z\rangle \leq \langle z, Az\rangle \\
&\Leftrightarrow \langle z, (A - A^2)z\rangle \geq 0 \text{ for all } z \in \mathbb{R}^n.
\end{aligned}$$

Then $(\lambda_i - \lambda_i^2) \geq 0 \Rightarrow \lambda_i(1 - \lambda_i) \geq 0$ for all $i \in \{1, 2, \dots, n\}$, so that $0 \leq \lambda_i \leq 1$. Hence, statement (ii) is equivalent to (16).

(ii)$\Leftrightarrow$(iii) Suppose that statement (ii) is true. Then Fact 2.9 gives us that $A = (\mathrm{Id} + P)^{-1}$ for some maximally monotone relation P. Since A is a matrix, we have that A is linear, so that A^{-1} is a linear relation. Note that the matrix inverse of A may not exist; here we are referring to the general set-valued inverse of A. Then we have $\mathrm{Id} + P = A^{-1} \Rightarrow P = A^{-1} - \mathrm{Id}$, so that P is a linear relation. Thus statement (ii) implies statement (iii). Conversely, supposing that statement (iii) is true and applying Fact 2.9, statement (ii) is immediately implied. $\square$

Proposition 4.15 will allow us to prove Theorem 4.2, one of the main results of this paper. Existence of a Moreau envelope is closely tied to nonexpansiveness, as the following proposition and example demonstrate, and Theorem 4.2 ultimately concludes.

Proposition 4.16 *If $A \in S^n_+$ is not nonexpansive, then the function*

$$f(x) = \frac{r}{2}\langle x, Ax\rangle + \langle b, x\rangle + c$$

is not the Moreau envelope with prox-parameter r of a proper, lsc, convex function.

Proof Suppose that A is not nonexpansive. Then

$$\exists\, x, y \in \mathbb{R}^n \text{ such that } \|Ax - Ay\| > \|x - y\|. \tag{17}$$

Suppose that f is the Moreau envelope with prox-parameter r of some $g \in \Gamma_0(\mathbb{R}^n)$. Then by Theorem 4.14, ∇f is r-Lipschitz. That is, for all $x, y \in \mathbb{R}^n$, we have

$$\begin{aligned}
&\|\nabla f(x) - \nabla f(y)\| \leq r\|x - y\| \\
&\|(rAx + b) - (rAy + b)\| \leq r\|x - y\| \\
&\|Ax - Ay\| \leq \|x - y\|.
\end{aligned}$$

This is a contradiction to (17). Therefore, f is not the Moreau envelope with prox-parameter r of any $g \in \Gamma_0(\mathbb{R}^n)$. $\square$

Example 4.6 Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$, $f(x) = \frac{1}{2}\langle x, Ax\rangle$. Then there does not exist $g \in \Gamma_0(\mathbb{R}^2)$ such that $f(x) = e_1 g(x)$. However, with $g(x) = \frac{1}{2}\langle x, x\rangle$ we have $f(x) = e_3 g(x)$.

Proof Setting $r = 1$, we know that there cannot exist g with $f = e_1 g$ as a direct consequence of Proposition 4.16, since A is not nonexpansive. However, rearranging the expression as $f(x) = \frac{3}{2}\langle x, \operatorname{Id} x\rangle$ gives a larger prox-parameter $\tilde{r} = 3$ and a nonexpansive matrix Id, so there does exist $g \in \Gamma_0(\mathbb{R}^2)$ such that $f(x) = e_3 g(x)$. □

4.6 A Characterization of Generalized Linear-Quadratic Functions

In this section, we present the remaining main result of the paper: a characterization of when a convex function is a generalized linear-quadratic. It has to do with convex Moreau envelopes, and we begin with a theorem that explicitly determines the Moreau envelope of a generalized linear-quadratic function.

Theorem 4.1 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone and symmetric linear relation. Let $r > 0$, $a, b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Define the generalized linear-quadratic function*

$$f(x) = \frac{r}{2}\langle x - a, A(x - a)\rangle + \langle b, x\rangle + c.$$

Then for every $x \in \mathbb{R}^n$,

$$e_r f(x) = r q_{(\operatorname{Id} + A^{-1})^{-1}}\left(x - a - \frac{b}{r}\right) + \langle b, x\rangle - \frac{1}{r} q(b) + c.$$

Proof By Proposition 2.1, we have

$$\begin{aligned} e_r f &= r e_1(f/r) = r e_1(q_A(\cdot - a) + \langle \cdot, b/r\rangle + c/r) \\ &= r[e_1(q_A(\cdot - a))(\cdot - b/r) + \langle \cdot, b/r\rangle - q(b/r) + c/r] \\ &= r[q_{(\operatorname{Id} + A^{-1})^{-1}}(\cdot - b/r - a) + \langle \cdot, b/r\rangle - q(b/r) + c/r] \\ &= r q_{(\operatorname{Id} + A^{-1})^{-1}}(\cdot - a - b/r) + \langle \cdot, b\rangle - \frac{1}{r} q(b) + c. \end{aligned}$$

□

Corollary 4.8 *Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximally monotone and symmetric linear relation. Then*

(i) $e_1(q_A) = q_{(\operatorname{Id} + A^{-1})^{-1}}$,
(ii) $q_A = e_1 g$ *for some* $g \in \Gamma_0(\mathbb{R}^n)$ *iff* A *is nonexpansive.*

Proof (i) This follows from Theorem 4.1 with $a = b = c = 0$ and $r = 1$.
(ii) This follows from part (i) above and Proposition 4.14 with $r = 1$. □

Proposition 4.17 *Let $f \in \Gamma_0(\mathbb{R}^n)$ be a quadratic function:*

$$f(x) = \frac{r}{2}\langle x, Ax\rangle + \langle b, x\rangle + c, \quad A \in S^n_+,\ b \in \mathbb{R}^n,\ \ \in \mathbb{R},\ r > 0.$$

Then $e_r f \in \Gamma_0(\mathbb{R}^n)$ and $e_r f$ is quadratic. Specifically,

$$e_r f(x) = \frac{r}{2}\langle x, [\mathrm{Id} - (\mathrm{Id} + A)^{-1}]x\rangle + \langle b, (\mathrm{Id} + A)^{-1}x\rangle - \frac{1}{2r}\langle b, (\mathrm{Id} + A)^{-1}b\rangle + c,$$

where $\mathrm{Id} - (\mathrm{Id} + A)^{-1} \in S^n_+$.

Proof Applying Theorem 4.1 with $a = 0$ and denoting $(\mathrm{Id} + A)^{-1}$ as $\mathcal{B}$, we have

$$\begin{aligned}
e_r f(x) &= r q_{(\mathrm{Id} + A^{-1})^{-1}}\left(x - \frac{b}{r}\right) + \langle b, x\rangle - \frac{1}{r} q(b) + c,\\
&= \frac{r}{2}\left\langle x - \frac{b}{r}, (\mathrm{Id} + A^{-1})^{-1}\left(x - \frac{b}{r}\right)\right\rangle + \langle b, x\rangle - \frac{1}{2r}\langle b, b\rangle + c,\\
&= \frac{r}{2}\left\langle x - \frac{b}{r}, [\mathrm{Id} - (\mathrm{Id} + A)^{-1}]\left(x - \frac{b}{r}\right)\right\rangle + \langle b, x\rangle - \frac{1}{2r}\langle b, b\rangle + c,\\
&= \frac{r}{2}\langle x, (\mathrm{Id} - \mathcal{B})x\rangle + \langle x, \mathcal{B}b\rangle - \frac{1}{2r}\langle b, \mathcal{B}b\rangle + c,\\
&= \frac{r}{2}\langle x, [\mathrm{Id} - (\mathrm{Id} + A)^{-1}]x\rangle + \langle b, (\mathrm{Id} + A)^{-1}x\rangle - \frac{1}{2r}\langle b, (\mathrm{Id} + A)^{-1}b\rangle + c.
\end{aligned}$$

Since $\mathrm{Id} - (\mathrm{Id} + A)^{-1} = (\mathrm{Id} + A^{-1})^{-1}$ is monotone and symmetric, we have that $\mathrm{Id} - (\mathrm{Id} + A)^{-1} \in S^n_+$ and the proof is complete. □

Theorem 4.2 *Let f be a convex quadratic function:*

$$f(x) = \frac{r}{2}\langle x, Ax\rangle + \langle b, x\rangle + c, \quad A \in S^n_+,\ b \in \mathbb{R}^n,\ c \in \mathbb{R},\ r > 0.$$

Then, A is nonexpansive iff $f = e_r g$ where g is a generalized linear-quadratic function:

$$g(x) = \begin{cases} \frac{r}{2}\langle x, P^{-1}x\rangle + \langle t, x\rangle + s, & \text{if } x \in \operatorname{dom} P^{-1},\\ \infty, & \text{if } x \notin \operatorname{dom} P^{-1}, \end{cases}$$

with P^{-1} a monotone linear relation. This includes the case $g(x) = \iota_{\{t\}}(x) + s$. Specifically, g is as follows.

(i) If $A = \mathrm{Id}$, then $g(x) = \iota_{\{t\}}(x) + s = q_{N_{\{0\}}}(x - t) + s$ (see Remark 4.2)*, where*

$$t = -\frac{b}{r} \text{ and } s = c - \frac{r}{2}\langle b, b\rangle. \tag{18}$$

(ii) *The matrix* $A \in S^n_+ \setminus \mathrm{Id}$ *is nonexpansive iff* $g(x) = \frac{r}{2}\langle x, P^{-1}x\rangle + \langle t, x\rangle + s$, *where*

$$P^{-1} = (\mathrm{Id} - A)^{-1} - \mathrm{Id},\ t \in (\mathrm{Id} - A)^{-1}b,\ \textit{and}\ s = c + \frac{1}{2r}\langle b, (\mathrm{Id} - A)^{-1}b\rangle. \tag{19}$$

Proof (i) Let $g(x) = \iota_{\{t\}}(x) + s = \begin{cases} s, & x = t, \\ \infty, & x \neq t. \end{cases}$ Then

$$\begin{aligned} e_r g(x) &= \inf_y \left\{ g(y) + \frac{r}{2}\|y - x\|^2 \right\} \\ &= g(t) + \frac{r}{2}\|t - x\|^2 \\ &= s + \frac{r}{2}\langle t - x, t - x\rangle \\ &= s + \frac{r}{2}(\langle t, t\rangle - 2\langle t, x\rangle + \langle x, x\rangle) \\ &= \frac{r}{2}\langle x, \mathrm{Id}\, x\rangle - r\langle t, x\rangle + \frac{r}{2}\langle t, t\rangle + s. \end{aligned}$$

Equating

$$A = \mathrm{Id}, \quad b = -rt, \quad \text{and } c = \frac{r}{2}\langle t, q\rangle + s, \tag{20}$$

we have that for any choice of $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$, there exists $g(x) = \iota_{\{t\}}(x) + s$ such that

$$f(x) = \frac{r}{2}\langle x, x\rangle + \langle b, x\rangle + c = e_r g(x).$$

The equations in (18) are obtained by solving (20) for t and s.

(ii) By Proposition 4.16, if A is not nonexpansive, then there does not exist $g \in \Gamma_0(\mathbb{R}^n)$ such that $f(x) = e_r g(x)$. Thus, supposing that there does exist such a g, we have that A is nonexpansive. Then by Fact 2.12, $A = (\mathrm{Id} + P)^{-1}$ for some maximally monotone operator P. Since $A \in S^n_+$, P is a symmetric linear relation by Proposition 4.15. Now using the general set-valued inverse P^{-1}, we set

$$g(x) = \begin{cases} \frac{r}{2}\langle x, P^{-1}x\rangle + \langle t, x\rangle + s, & \text{if } x \in \mathrm{dom}\ P^{-1}, \\ \infty, & \text{if } x \notin \mathrm{dom}\ P^{-1}. \end{cases}$$

This function g is well-defined due to Fact 4.1. Since P is a monotone linear relation, the function

$$h(x) = \begin{cases} \frac{1}{2}\langle x, P^{-1}x\rangle, & \text{if } x \in \mathrm{dom}\ P, \\ \infty, & \text{if } x \notin \mathrm{dom}\ P \end{cases}$$

is single-valued. Then by Proposition 4.17, we have

$$\begin{aligned} e_r g(x) &= \frac{r}{2}\langle x, [\mathrm{Id} - (\mathrm{Id} + P^{-1})^{-1}]x\rangle + \langle t, (\mathrm{Id} + P^{-1})^{-1}x\rangle \\ &\quad - \frac{1}{2r}\langle t, (\mathrm{Id} + P^{-1})^{-1}t\rangle + s \\ &= \frac{r}{2}\langle x, (\mathrm{Id} + P)^{-1}x\rangle + \langle t, (\mathrm{Id} + P^{-1})^{-1}x\rangle \\ &\quad - \frac{1}{2r}\langle t, (\mathrm{Id} + P^{-1})^{-1}t\rangle + s. (\text{Fact 2.7}) \end{aligned}$$

Equating

$$A = (\mathrm{Id} + P)^{-1}, \quad b = (\mathrm{Id} + P^{-1})^{-1}t, \text{ and } c = s - \frac{1}{2r}\langle t, (\mathrm{Id} + P^{-1})^{-1}t\rangle, \tag{21}$$

we have that $f(x) = \frac{r}{2}\langle x, Ax\rangle + \langle b, x\rangle + c$ is the Moreau envelope of the function $g(x) = \frac{r}{2}\langle x, P^{-1}x\rangle + \langle t, x\rangle + s$. The equations in (19) are obtained by solving (21) for P^{-1}, t, and s. □

Theorem 4.3 *The function $f \in \Gamma_0(\mathbb{R}^n)$ is a generalized linear-quadratic function iff $e_r f \in \Gamma_0(\mathbb{R}^n)$ is a quadratic function. Specifically,*

$$e_r f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle + c \qquad \forall x \in \mathbb{R}^n$$

with $A \in S^n_+$, iff

$$f(x) = q_{(A^{-1} - \mathrm{Id}/r)^{-1}}\left(x + \frac{b}{r}\right) + \langle b, x\rangle + c + \frac{1}{2r}\|b\|^2 \qquad \forall x \in \mathbb{R}^n.$$

Proof (⇒) This is the statement of Theorem 4.1.
(⇐) Let $e_r f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle + c$, with A symmetric, linear and monotone, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then

$$(e_r f)^* = f^* + \frac{1}{r}q,$$

and

$$(q_A + \langle \cdot, b\rangle + c)^* = q_{A^{-1}}(\cdot - b) - c.$$

This gives us that

$$f^* = q_{A^{-1} - \mathrm{Id}/r}(\cdot - b) - \langle \cdot, b/r\rangle - c + q(b)/r.$$

It follows that

$$\begin{aligned} f &= (q_{A^{-1}-\mathrm{Id}/r}(\cdot - b))^*(\cdot + b/r) + c - q(b)/r \\ &= q_{(A^{-1}-\mathrm{Id}/r)^{-1}}(\cdot + b/r) + \langle \cdot + b/r, b\rangle + c - q(b)/r \\ &= q_{(A^{-1}-\mathrm{Id}/r)^{-1}}(\cdot + b/r) + \langle \cdot, b\rangle + c + q(b)/r. \end{aligned}$$

Thus, $f \in \Gamma_0(\mathbb{R}^n)$ is a generalized linear-quadratic function. □

5 Applications

This section presents a few applications of the theory seen thus far. We build on the idea of extended norms, give an application to the least squares problem and explore the limit of a sequence of generalized linear-quadratic functions.

5.1 A Seminorm with Infinite Values

In [29,30], Beer and Vanderwerff present the idea of norms that are allowed to take on infinite values. These so-called extended norms are functions on linear spaces that satisfy the properties of a norm when they are finite-valued, but can be infinite-valued as well. The authors extend many properties of norms to the setting of an extended norm space $(X, \|\cdot\|)$, where X is a vector space and $\|\cdot\|$ is an extended norm. In that spirit, we present here an extended seminorm.

Definition 5.1 A function $k : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a *gauge* if k is a nonnegative, positively homogeneous and convex function such that $k(0) = 0$. Thus, a gauge is a function k such that

$$k(x) = \inf\{\mu \geq 0 : \ x \in \mu C\}$$

for some nonempty convex set C.

Definition 5.2 The *polar* of a gauge k is the function k^o defined by

$$k^o(x^*) = \inf\{\mu^* \geq 0 : \ \langle x, x^*\rangle \leq \mu^* k(x) \ \forall x \in \mathbb{R}^n\}.$$

If k is finite everywhere and positive except at the origin, then the polar of k can be written as

$$k^o(x^*) = \sup_{x \neq 0} \frac{\langle x, x^*\rangle}{k(x)}.$$

Definition 5.3 A function $k : \mathbb{R}^n \to \overline{\mathbb{R}}$ is an *extended seminorm* if

(i) $k(x) \geq 0 \ \forall x \in \mathbb{R}^n$,
(ii) $k(\alpha x) = |\alpha| k(x) \ \forall x \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}$,

(iii) $k(x+y) \leq k(x) + k(y)$ $\forall x, y \in \mathbb{R}^n$,
(iv) $k(x) = \infty$ if $x \notin \operatorname{dom} k$.

Theorem 5.1 *Let A be a maximally monotone and symmetric linear relation. Then the following hold.*

(i) The function

$$k = (2q_A)^{1/2}$$

is an extended seminorm. Moreover,

$$k^{-1}(0) = A^{-1}0. \tag{22}$$

(ii) For all $x \in \operatorname{dom} A$ *and for all* $x^* \in \operatorname{ran} A$, *we have*

$$\langle x, x^* \rangle \leq \sqrt{\langle x, Ax \rangle}\sqrt{\langle x^*, A^{-1}x^* \rangle}.$$

(iii) The closed and convex sets

$$C = \{x : q_A(x) \leq 1\}, \qquad C^* = \{x^* : q_{A^{-1}}(x^*) \leq 1\}$$

are polar to each other.

Proof (i) Applying [11, Corollary 15.3.1] with $f = q_A$ and $p = 2$, We have that k is a gauge function. Thus, k is an extended seminorm. To see (22), we have that $k(x) = 0 \Leftrightarrow q_A(x) = 0$, so it suffices to apply Proposition 4.6.

(ii) By Proposition 4.7, $q_A^* = q_{A^{-1}}$. By [11, Corollary 15.3.1], we have that $k^o(x^*) = (2q_A^*(x^*))^{1/2}$, and that $\forall x \in \operatorname{dom} A$, $\forall x^* \in \operatorname{ran} A$,

$$\begin{aligned}\langle x, x\rangle^* &\leq k(x)k^o(x^*) \\ &= 2(q_A(x))^{1/2}(q_{A^{-1}}(x^*))^{1/2} \\ &= \sqrt{\langle x, Ax \rangle}\sqrt{\langle x^*, A^{-1}x^* \rangle}.\end{aligned}$$

(iii) By [11, Corollary 15.3.2], we have that the closed and convex sets

$$C = \{x : \langle x, Ax \rangle \leq 1\}, \qquad C^* = \{x^* : \langle x^*, A^{-1}x^* \rangle \leq 1\}$$

are polar to each other. □

Remark 5.1 The above result generalizes Rockafellar's result on [11, p. 136] with $Q = A$, from a positive definite matrix to a maximally monotone and symmetric linear relation.

5.2 The Least Squares Problem

In this section, we show that generalized linear-quadratic functions can be used to study the least squares problem. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. The general least squares problem is to find a vector $x \in \mathbb{R}^n$ that minimizes

$$\ell : \mathbb{R}^n \to \mathbb{R} : x \mapsto \frac{1}{2}\|Ax - b\|^2 = q_{A^\top A}(x) - \langle x, A^\top b\rangle + q_{\mathrm{Id}}(b). \tag{23}$$

Theorem 5.2 *For the function ℓ given by (23), we have*

(i) $\ell^*(y) = q_{(A^\top A)^{-1}}(y + A^\top b) - q_{\mathrm{Id}}(b) \quad \forall y \in \mathbb{R}^n$;

(ii)

$$\partial\ell^*(y) = (A^\top A)^{-1}(y + A^\top b) \quad \forall y \in \mathbb{R}^n, \tag{24}$$

and

$$\mathrm{dom}\ \ell^* = \mathrm{ran}\ A^\top. \tag{25}$$

Proof (i) Apply Example 4.13.

(ii) Apply Proposition 4.4 to obtain (24). To see (25), using the facts that ran $A^\top A =$ ran $A^\top$ (c.f. [26, page 212]) and that ran $A^\top$ is a subspace, we have

$$\begin{aligned}\mathrm{dom}\ \ell^* &= \mathrm{dom}[(A^\top A)^{-1} - A^\top b] = \mathrm{ran}(A^\top A - A^\top b)\\ &= \mathrm{ran}(A^\top - A^\top b) = \mathrm{ran}\ A^\top.\end{aligned}$$

□

5.3 Sequences and Calculus Rules

We end this work with an application for sequences of q_{A_k} functions with A_k linear relations, and the development of a calculus for the generalized linear-quadratic functions.

Proposition 5.1 (Epi-convergence)

(i) For all $k \in \mathbb{N}$, let

$$f_k = q_{A_k}(\cdot - a_k) + \langle b_k, \cdot\rangle + c_k, \tag{26}$$

where A_k is a maximally monotone and symmetric linear relation, $a_k, b_k \in \mathbb{R}^n$ and $c_k \in \mathbb{R}$. Suppose that $f_k \xrightarrow{e} f$ and that f is proper. Then f is a generalized linear-quadratic function:

$$f = q_A(\cdot - a) + \langle b, \cdot\rangle + c, \tag{27}$$

where A is a maximally monotone symmetric linear relation, $a, b \in \mathbb{R}^n$, $c \in \mathbb{R}$.

(ii) For all $k \in \mathbb{N}$, let

$$f_k = q_{A_k} + c_k, \tag{28}$$

where A_k is a maximally monotone and symmetric linear relation, and $c_k \in \mathbb{R}$. Suppose that $f_k \xrightarrow{e} f$ and that f is proper. Then f is a generalized linear-quadratic function:

$$f = q_A + c, \tag{29}$$

where A is a maximally monotone and symmetric linear relation, and $c \in \mathbb{R}$.

Proof (i) As $f_k \xrightarrow{e} f$, we have $\partial f_k \xrightarrow{g} \partial f$. Differentiating (26), we find that $\partial f_k = A_k(\cdot - a_k) + b_k$, so that gra ∂f_k = gra $A_k + (a_k, b_k)$ is maximally monotone and affine. Thus, gra ∂f is maximally monotone and affine. By [8, Theorem 4.3], gra ∂f = gra $A + (a, b)$ for some maximally monotone and symmetric linear relation A. Then by Proposition 4.4, we have $A = \partial q_A$, so that (27) is true.

(ii) The proof is similar to that of part (i). Differentiating (28), we find that $\partial f_k = A_k \; \forall x \in \mathbb{R}^n$, so that gra ∂f_k = gra A_k is a linear subspace. Thus, gra ∂f is a linear subspace, gra ∂f = gra A for some maximally monotone and symmetric linear relation A, and Proposition 4.4 gives $A = \partial q_A$. Therefore, (29) is true. □

As a result of Proposition 5.1, we are now able to define calculus rules for generalized linear-quadratic functions. We do so in the form of Theorems 5.3 and 5.4, for which we define the following sets.

Definition 5.4 Denote by $\mathcal{A}$ the set of maximally monotone and symmetric linear relations on $\mathbb{R}^n$. We define S as the set of convex generalized linear-quadratic functions:

$$S = \left\{ f = q_A(\cdot - a) + \langle b, \cdot \rangle + c : A \in \mathcal{A}, a, b \in \mathbb{R}^n, c \in \mathbb{R}, f \in \Gamma_0(\mathbb{R}^n) \right\}.$$

We define T as the subset of S obtained by setting $a = 0$:

$$T = \left\{ f = q_A + \langle b, \cdot \rangle + c : A \in \mathcal{A}, b \in \mathbb{R}^n, c \in \mathbb{R}, f \in \Gamma_0(\mathbb{R}^n) \right\}.$$

We define U as the subset of S obtained by setting $a = b = 0$:

$$U = \left\{ f = q_A + c : A \in \mathcal{A}, c \in \mathbb{R}, f \in \Gamma_0(\mathbb{R}^n) \right\}.$$

We begin with calculus rules for the simpler case, the set U.

Theorem 5.3 *Let d be the Attouch–Wets metric* (see Definition 2.5). *The following hold.*

(i) The metric space (U, d) is complete.
(ii) If $f \in U$, then $f^* \in U$.

(iii) If $f \in U$ *and* $\lambda > 0$, *then* $\lambda f \in U$.
(iv) If $f_1, f_2 \in U$, *then* $f_1 + f_2 \in U$.
(v) If $f_1, f_2 \in U$, *then* $f_1 \square f_2 \in U$.

Proof (i) This follows from Proposition 5.1.
(ii) Let $f \in U$, $f = q_A + c$. By Proposition 4.7, we have

$$f^* = (q_A + c)^* = q_A^* - c = q_{A^{-1}} - c,$$

which is a convex generalized linear-quadratic function of the required form. Therefore, $f^* \in U$.
(iii) Clearly, $f = q_A + c \in U$ and $\lambda > 0$ yields $\lambda f = q_{\lambda A} + \lambda c \in U$.
(iv) Let $f_1, f_2 \in U$, $f_1 = q_{A_1} + c_1$, $f_2 = q_{A_2} + c_2$. By Proposition 4.5, we have that

$$(f_1 + f_2) = q_{A_1+A_2} + c_1 + c_2$$

is a convex generalized linear-quadratic function of the form found in U. Therefore, $f_1 + f_2 \in U$.
(v) Let $f_1, f_2 \in U$, $f_1 = q_{A_1} + c_1$, $f_2 = q_{A_2} + c_2$. By Propositions 4.9 and 4.7, we have

$$f_1 \square f_2 = q^*_{A_1^{-1}+A_2^{-1}} + c_1 + c_2 = q_{(A_1^{-1}+A_2^{-1})^{-1}} + c_1 + c_2 \in U.$$

□

For the more general setting of the sets S and T, the calculus rules are not so straightforward. More stringent conditions are necessary; the following theorem provides the obtainable results.

Theorem 5.4 *Let d be the Attouch–Wets metric* (see Definition 2.5). *The following hold.*

(i) The metric space (S, d) *is complete.*
(ii) If $f \in S$, *then* $f^* \in S$.
(iii) If $f \in S$ *with* $b = 0$, *then* $f^* \in T$.
(iv) If $f \in S$ $(f \in T)$ *and* $\lambda > 0$, *then* $\lambda f \in S$ $(\lambda f \in T)$.
(v) If $f_1, f_2 \in T$, *then* $f_1 + f_2 \in T$.

Proof (i) This follows from Proposition 5.1.
(ii) Let $f \in S$, $f = q_A(\cdot - a) + \langle b, \cdot\rangle + c$. Combining Proposition 4.7 and [14, Proposition 13.20], we have

$$\begin{aligned} f^* &= (q_A(\cdot - a) + \langle b, \cdot\rangle + c)^* \\ &= (q_A(\cdot - a))^*(\cdot - b) - c \\ &= (q_{A^{-1}} + \langle a, \cdot\rangle)(\cdot - b) - c \\ &= q_{A^{-1}}(\cdot - b) + \langle a, \cdot - b\rangle - c, \end{aligned}$$

which is a convex generalized linear-quadratic function. Therefore, $f^* \in S$.

(iii) Let $f \in S$, $f = q_A(\cdot - a) + c$. By the same procedure as in the proof of (ii), we have

$$f^* = q_{A^{-1}} + \langle a, x\rangle - c \in T.$$

(iv) Clearly, since $f = q_A(\cdot - a) + \langle b, \cdot\rangle + c \in S$ and $\lambda > 0$, we have that $\lambda f = q_{\lambda A}(\cdot - a) + \lambda\langle b, \cdot\rangle + \lambda c \in S$. Also clearly, $a = 0$ yields $\lambda f \in T$.

(v) Let $f_1, f_2 \in T$, $f_1 = q_{A_1} + \langle b_1, \cdot\rangle + c_1$, $f_2 = q_{A_2} + \langle b_2, \cdot\rangle + c_2$. By Proposition 4.5, we have that

$$(f_1 + f_2) = q_{A_1+A_2} + \langle b_1 + b_2, \cdot\rangle + c_1 + c_2$$

is a convex generalized linear-quadratic function. Therefore, setting

$$f = f_1 + f_2,\ A = A_1 + A_2,\ b = b_1 + b_2 \text{ and } c = c_1 + c_2,$$

we have that $f = q_A + \langle b, \cdot\rangle + c \in T$. □

6 Conclusions

On $\mathbb{R}^n$, the Moreau envelope of a generalized linear-quadratic function was explicitly identified. Conversely, it was determined under what conditions a quadratic function is a Moreau envelope of a generalized linear-quadratic. Characterizations of the existence of the Moreau envelope for convex functions involving Lipschitz continuity and nonexpansiveness were established, and we showed that a convex function is generalized linear-quadratic if and only if its Moreau envelope is convex linear-quadratic. The topic of generalized linear-quadratic functions was discussed at length; several useful characterizations and properties were established. We gave applications to generalized seminorms, the least squares problem, the epi-limit of a sequence and the calculus of generalized linear-quadratic functions.

Acknowledgements The authors would like to thank the editor and the anonymous referee for their very helpful comments and suggestions for improvement of this manuscript. Chayne Planiden was supported by UBC University Graduate Fellowship and by Natural Sciences and Engineering Research Council of Canada. Xianfu Wang was partially supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant.

References

1. Moreau, J.J.: Proximité et dualité dans un espace Hilbertien. Bull. Soc. Math. Fr. **93**, 273–299 (1965)
2. Poliquin, R., Rockafellar, R.: Generalized Hessian properties of regularized nonsmooth functions. SIAM J. Optim. **6**(4), 1121–1137 (1996)
3. Attouch, H., Wets, R.: Quantitative stability of variational systems: I. The epigraphical distance. Trans. Am. Math. Soc. **328**, 695–729 (1991)
4. Attouch, H.: Variational Convergence for Functions and Operators. Applicable Mathematics Series. Pitman, Boston (1984)

5. Cross, R.: Multivalued Linear Operators. Monographs and Textbooks in Pure and Applied Mathematics, vol. 213. Marcel Dekker Inc, New York (1998)
6. Bartz, S., Bauschke, H., Moffat, S., Wang, X.: The resolvent average of monotone operators: dominant and recessive properties. SIAM J. Optim. **26**(1), 602–634 (2016)
7. Bauschke, H., Wang, X., Yao, L.: Monotone linear relations: maximality and Fitzpatrick functions. J. Convex Anal. **16**(3–4), 673–686 (2009)
8. Bauschke, H., Wang, X., Yao, L.: On Borwein–Wiersma decompositions of monotone linear relations. SIAM J. Optim. **20**(5), 2636–2652 (2010)
9. Yao, L.: On Monotone Linear Relations and the Sum Problem in Banach Spaces. UBC Ph. D. thesis (2011)
10. Lemaréchal, C., Sagastizábal, C.: Practical aspects of the Moreau–Yosida regularization: theoretical preliminaries. SIAM J. Optim. **7**(2), 367–385 (1997)
11. Rockafellar, R.: Convex Analysis. Princeton Landmarks in Mathematics. Princeton University Press, Princeton (1997)
12. Hiriart-Urruty, J.B.: The deconvolution operation in convex analysis: an introduction. Cybern. Syst. Anal. **30**(4), 555–560 (1994)
13. Rockafellar, R., Wets, R.: Variational Analysis. Springer, Berlin (1998)
14. Bauschke, H., Combettes, P.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2011)
15. Borwein, J., Vanderwerff, J.: Convex Functions: Constructions, Characterizations and Counterexamples. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge (2010)
16. Attouch, H., Beer, G.: On the convergence of subdifferentials of convex functions. Arch. Math. (Basel) **60**(4), 389–400 (1993)
17. Burke, J., Hoheisel, T.: Epi-convergent smoothing with applications to convex composite functions. SIAM J. Optim. **23**(3), 1457–1479 (2013)
18. Rockafellar, R., Royset, J.: Random variables, monotone relations, and convex analysis. Math. Program. **148**(1–2, Ser. B), 297–331 (2014)
19. Wijsman, R.A.: Convergence of sequences of convex sets, cones and functions. II. Trans. Am. Math. Soc. **123**, 32–45 (1966)
20. Beer, G.: Topologies on Closed and Closed Convex Sets. Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht (1993)
21. Planiden, C., Wang, X.: Strongly convex functions, Moreau envelopes and the generic nature of convex functions with strong minimzers. SIAM J. Optim. **26**(2), 1341–1364 (2016)
22. Bartz, S., Bauschke, H., Borwein, J., Reich, S., Wang, X.: Fitzpatrick functions, cyclic monotonicity and Rockafellar's antiderivative. Nonlinear Anal. **66**(5), 1198–1223 (2007)
23. Baillon, J.B., Haddad, G.: Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. Isr. J. Math. **26**(2), 137–150 (1977)
24. Bauschke, H., Borwein, J., Wang, X., Yao, L.: The Brezis–Browder theorem in a general Banach space. J. Funct. Anal. **262**(12), 4948–4971 (2012)
25. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods, vol. 306. Springer, Berlin (2013)
26. Meyer, C.: Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2000)
27. Mordukhovich, B.: Variational Analysis and Generalized Differentiation I: Basic Theory, vol. 330. Springer, Berlin (2006)
28. Roman, S.: Advanced Linear Algebra, vol. 3. Springer, Berlin (2005)
29. Beer, G.: Norms with infinite values. J. Convex Anal. **22**(1), 35–58 (2015)
30. Beer, G., Vanderwerff, J.: Structural properties of extended normed spaces. Set-Valued Var. Anal. **23**(4), 613–630 (2015)