Journal of Convex Analysis
Volume 6 (1999), No. 2, 235–266

Graph-Convex Mappings and K-Convex Functions

Teemu Pennanen
Department of Mathematics, University of Washington,
Box 354350, Seattle, Washington 98195-4350, USA.
e-mail: pennanen@math.washington.edu

Received April 28, 1998
Revised manuscript received February 24, 1999

This paper studies global and local behavior of graph-convex set-valued mappings in finite-dimensional vector spaces. This is done in terms of recession mappings and graphical derivatives which are set-valued mappings whose graphs are convex cones. The main results are chain rules for computing the recession mapping and the graphical derivative of a composition of two set-valued mappings. The results on graph-convex mappings are applied to K-convex functions which are vector-valued generalizations of extended-real-valued proper convex functions. Many generalizations of classical results in convex analysis are obtained, along with a generalization of subdifferential calculus, in which the differential behavior of a function is described by a sublinear mapping that resembles the classical Jacobian. A particular advantage of this approach is that it leads to simple chain rules for compositions of vector-valued convex functions. The generality is reflected in the fact that most of the classical rules for computing recession functions and subdifferentials are obtained as special cases of the given chain rules. Some applications to mathematical programming and matrix analysis are given.

1. Introduction

Because of their flexibility in modeling various situations both in theory and practice, set-valued mappings provide a convenient framework for studying problems in optimization and variational analysis. Since a set-valued mapping $S : X \rightrightarrows U$ may be identified with its graph $\operatorname{gph} S = \{(x, u) \mid u \in S(x)\}$, its analysis can be reduced to the study of sets in the product space $X \times U$. A particularly convenient case occurs when the graph of S is convex. Then the analysis falls into the realm of convex analysis, and the results are stronger and more elegant than could be expected in the general case. This can be seen for example in Robinson [33], Borwein [4], Aubin and Ekeland [2], Aubin and Frankowska [3], and Rockafellar and Wets [37].

The purpose of this paper is to exploit convexity further, in the study of global and local behavior of set-valued mappings. We derive simple rules for describing such properties for mappings that have been constructed from other mappings whose corresponding properties are known. Specifically, we develop a recession calculus for graph-convex mappings, and we simplify and strengthen some of the already existing results on graphical differentiation of graph-convex mappings. This is based on the well known fact that a wide range of mappings and sets with special structure can be expressed in terms of compositions. Any property of a general composition yields information about these more special cases.

Two major themes in convex analysis are the study of recession and differential behavior of convex sets and convex functions. We develop a corresponding theory for graph convex mappings, and derive many extensions of the classical results as special cases. All the proofs follow one simple idea, and we only use the standard results in Rockafellar [35]. To emphasize the simplicity we consider only the finite-dimensional case. A major part

ISSN 0944-6532 / \$ 2.50

of our study is devoted to K-convex functions which can be identified with a special class of graph-convex set-valued mappings. The special properties of this class have consequences that show how K-convex functions share many important properties of extended-real-valued convex functions. The study of K-convex functions is also motivated by their increasing popularity in modeling various problems in optimization, especially in semidefinite programming, eigenvalue optimization and vector minimization; see for example Borwein [4], Craven [9], Shapiro [39, 40], and Lewis [24].

The global behavior of a graph-convex mapping will be described in terms of the "recession mapping" which is the sublinear mapping whose graph is the recession cone of the original graph. When the mapping is closed, the recession mapping coincides with the "horizon mapping" introduced for general set-valued mappings in [37]. Our main result on recession mappings is a new chain rule which states that, under mild conditions, the recession mapping of a composition is the composition of the recession mappings.

The local behavior of a graph-convex mapping will be described in terms of the "graphical derivative" which is the sublinear mapping whose graph is the tangent cone of the original graph at some of its points; see Aubin [1], as well as [2, 3, 37]. In the presence of convexity this is adjoint to the coderivative of Mordukhovich [26, 27]. In the general case, without convexity, chain rules have been given in inclusion form in [27, 37]. For graph-convex mappings strict chain rules are given in [37], and in [2] for the case where one of the mappings is linear. Our chain rule applies under slightly more general conditions than those in [37, Theorem 10.37], and in the finite-dimensional case, it implies those in [2]. Also, since we consider only the finite-dimensional convex case, our proofs are very simple, and the "constraint qualifications" are more in the spirit of convex analysis. For reference on generalized differentiation of set-valued mappings see [3, 37].

Let $K \subset U$ be a convex cone, that is, a nonempty convex set such that $\alpha x \in K$ whenever $x \in K$ and $\alpha \geq 0$. A function f from X to U defined on a set $\operatorname{dom} f$ is said to be *K-convex* if $\operatorname{dom} f$ is convex, and for every $x_1, x_2 \in \operatorname{dom} f$,

$$f(\alpha_1 x_1 + \alpha_2 x_2) - \alpha_1 f(x_1) - \alpha_2 f(x_2) \in K,$$

whenever $\alpha_1, \alpha_2 \in [0,1]$, and $\alpha_1 + \alpha_2 = 1$. The convexity of a real-valued function may be stated as $\mathbb{R}_-$-convexity, and concavity as $\mathbb{R}_+$-convexity. In accordance with the real-valued case, a function f from X to U with $\operatorname{dom} f = X$ will be called *affine* if it is $\{0\}$-convex. Note that, rather than extending the domain of a function to all of X, by defining something corresponding to extended-real-valued functions, we explicitly state the set $\operatorname{dom} f$ in which f has well defined values in U. Extended-real-valued functions that have values in $\mathbb{R} \cup (+\infty)$ fit this framework by defining $\operatorname{dom} f = \{x \mid f(x) < \infty\}$, as usual.

Just as in the real-valued case, it is easily shown that f is K-convex if and only if the set

$$\operatorname{epi}_K f = \{(x,u) \mid x \in \operatorname{dom} f,\ u \in f(x) - K\},$$

called the *K-epigraph* of f, is a convex set in $X \times U$. The K-epigraph of f is the graph of the *K-profile mapping* $S_{f,K} : X \rightrightarrows U$ of f defined by

$$S_{f,K}(x) = \begin{cases} f(x) - K & \text{if } x \in \operatorname{dom} f, \\ \emptyset & \text{otherwise.} \end{cases}$$

Through this relation all the results on graph-convex mappings can be translated into facts about K-convex functions.

An interesting example of a nonsmooth K-convex function may be found in matrix analysis. Choosing X to be the space of $n \times n$ Hermitian matrices and $U = \mathbb{R}^n$, one can show that the function $\lambda : X \to U$ giving the eigenvalues of a matrix in nonincreasing order is K-convex; see Friedland [11], Lewis [21, 23, 24] or Section 9. As an example, we will use K-convex functions to derive some recent results about λ, in Overton and Womersley [30], Hiriart-Urruty and Ye [16], Seeger [38], as well as in [11], which have important applications in eigenvalue optimization and semidefinite programming; see for example Overton [29], Vandenberghe and Boyd [43] and the references there in.

In [24], Lewis introduced the concept of a "normal decomposition system". A normal decomposition system involves a function γ which generalizes the function λ in the above example [24, Section 7]. By [24, Theorem 2.4] γ is K-convex, where K is the polar cone of the range of γ. Consequently, our results apply to normal decomposition systems yielding, for example, some of the closedness criteria and subgradient formulas in [24], as well as a generalization of the recession function formula given in [38] for spectrally defined matrix functions.

If f is a K-convex function from X to U, and g is an L-convex function from U to V, the classical criterion for the convexity of their composition $g \circ f$ is that g is nonincreasing, with respect to the partial order induced by L, in the directions of K. Because of its simplicity, this condition is often easy to verify, but for some applications it is too restrictive. In Section 6, we give a more general condition that also guarantees that the profile mapping of the composition coincides with the composition of the profile mappings of f and g. This will be used to translate facts about compositions of general graph-convex mappings into facts about compositions of K-convex functions.

The "K-recession mapping" of a K-convex function is defined as the recession mapping of its profile mapping. The special form of K-profile mappings implies interesting properties for the K-recession mappings. In particular, the adjoint sublinear mapping of the K-recession mapping can be expressed in terms of domains of convex conjugates of "scalarized" functions. This provides a useful link to the recession theory of extended-real-valued functions [35]. The graphical approach to the recession properties of set-valued mappings also yields closedness criteria for compositions of set-valued mappings and K-convex functions.

When applied to K-convex functions, graphical differentiation leads to a vector-valued generalization of the subdifferential calculus in convex analysis. The subdifferential is replaced by the "K-Jacobian" which is the graphical derivative of the K-profile mapping at a point $(x, f(x))$; see also Pennanen and Eckstein [32]. Subdifferentials and K-Jacobians are related in a way that allows us to interpret facts about either object in terms of the other (Section 6), and for extended-real-valued functions, we have a one-to-one correspondence between the two. In smooth analysis, this corresponds to identifying the space of real-valued linear mappings with a dual space X^* of X, and the real-valued Jacobians with points (gradients) in X^*. The relation between subdifferentials and K-Jacobians extends this correspondence in a natural way to the setting of convex analysis where points and linear mappings correspond to the more general concepts of convex sets and sublinear mappings, respectively, as pointed out in Rockafellar [34]. In general, this kind of representation is possible only when the range-space of the mapping is the real line. In order to obtain a general differential theory applicable to vector-valued functions as well, it seems necessary to pass from subdifferentials to the corresponding sublinear mappings. This differs from the usual approach where the differential behavior of a K-convex function is described by a set in the space of linear mappings, as in Levin [20], Kutateladze [19], Thera [41] and Borwein [6], to name just a few. Also, in view of classical analysis, it seems more natural to have a Jacobian which is a mapping instead of a set. Most importantly, since the K-Jacobian is a graphical derivative of the corresponding K-profile mapping, we can use the general calculus rules for graphical derivatives to compute K-Jacobians for composite functions.

This approach to nonsmooth analysis of vector-valued functions is very similar to that in [27] and [37]. However, our analysis replaces the local Lipschitz continuity by K-convexity, and we obtain exact chain rules in the form of equalities instead of inclusions. Relaxing the Lipschitz requirement, allows us to treat vector-valued functions which are defined only on a subset of the domain space. This corresponds to the distinction between real-valued and extended-real-valued functions.

The next section gives the basic definitions and some background on sublinear mappings. In Section 3, we define the recession mapping for a general graph-convex set-valued mapping, and derive a chain rule for the recession mapping of a composition. In Sections 4 and 5, we do the same for the graphical derivative and the "relative interior" of a set-valued mapping, which is the set-valued mapping whose graph is the relative interior of the original graph. In Sections 6 and 7, we apply the results to K-convex functions. The last two sections are devoted to applications. In Section 8, we study the composite model of convex programming, variational inequalities, and vector minimization problems. In Section 9, we show how the eigenvalue function in the space of Hermitian matrices can be treated under the framework of K-convex functions.

2. Preliminaries

We will use the notation of [37]. The *domain* and the *range* of a set-valued mapping S are defined as projections of $\operatorname{gph} S$ to X and U, respectively:

$$\operatorname{dom} S = \{x \in X \mid S(x) \neq \emptyset\}, \quad \operatorname{rge} S = \bigcup \{S(x) \mid x \in X\}.$$

The *inverse* of S is defined by $S^{-1}(u) = \{x \mid u \in S(x)\}$. The *closure* $\operatorname{cl} S$ of S is defined by $\operatorname{gph}(\operatorname{cl} S) = \operatorname{cl}(\operatorname{gph} S)$. If $S = \operatorname{cl} S$ it is said to be *closed.*

The main theme of this paper is composition of mappings. The *composition* of set-valued mappings $S : X \rightrightarrows U$ and $T : U \rightrightarrows V$ is defined by

$$(T \circ S)(x) = T(S(x)) = \bigcup \{T(u) \mid u \in S(x)\}.$$

The following graphical characterization of the composition was used in [37]. Because of its importance in all that follows, we provide the simple proof.

Lemma 2.1. *For any $S : X \rightrightarrows U$ and $T : U \rightrightarrows V$,*

$$\operatorname{gph}(T \circ S) = P_{X \times V}[(\operatorname{gph} S \times V) \cap (X \times \operatorname{gph} T)],$$

so that convexity of S and T implies the convexity of $T \circ S$.

Proof. Everything is clear from the expression

$$\begin{aligned}\operatorname{gph}(T\circ S) &= \{(x,v) \mid \exists u \in S(x):\ v \in T(u)\}\\ &= \{(x,v) \mid \exists u:\ (x,u,v) \in (\operatorname{gph} S \times V) \cap (X \times \operatorname{gph} T)\},\end{aligned}$$

and from the fact that intersections and linear mappings (e.g. projections) preserve convexity. □

Thus, to analyze "graphical properties" of compositions, it suffices that we are able to analyze intersections of convex sets and images of convex sets under linear mappings. The main reason to focus on compositions is their generality in the sense that other kinds of composite mappings are obtained by special choices of S or T.

Corollary 2.2.

(i) *Let $A : X \to U$ be linear, and let $T : U \rightrightarrows V$ be convex. Then $T\circ A : X \rightrightarrows V$ is convex.*

(ii) *Let $S : X \rightrightarrows U$ be convex, and let $B : U \to V$ be linear. Then $B\circ S : X \rightrightarrows V$ is convex.*

(iii) *Let $S_1, S_2 : X \rightrightarrows U$ be convex. Then $S_1 + S_2$ is convex.*

(iv) *Let $D \subset U$ and $T : U \rightrightarrows V$ be convex. Then $T(D)$ is convex.*

Proof. In (i) we have $S = A$, and in (ii), $T = B$. Part (iii), follows from (i) and (ii) by noting that $S_1 + S_2 = B\circ T\circ A$, where $A(x) = (x,x)$, $T(x_1,x_2) = S_1(x_1) \times S_2(x_2)$, and $B(u_1,u_2) = u_1 + u_2$. Part (iv) follows by choosing S to be the constant mapping $S(x) = D$, and using the fact that the range of the convex mapping $(T\circ S)(x) = T(D)$ is convex. □

It follows that various properties about general compositions of convex mappings can be translated into properties about the above cases. This will be the strategy in each of the subsequent sections: using the expression for the graph of a general composition in Lemma 2.1, we will first derive a result about a general composition, and then apply this to the special constructions in the proof of the above corollary. It should be noted that the above cases are just few special cases of composition. For example, mappings of the form $T\circ A^{-1}$, $B^{-1}\circ S$ or $(S_1 \cap S_2)(x) = S_1(x) \cap S_2(x)$ would be natural additions to the above list.

A set-valued mapping is called *sublinear* if its graph is a convex cone [37]. Sublinear mappings are set-valued generalizations of linear (single-valued) mappings. In particular, every linear mapping is sublinear. It is also easily checked that sums and compositions of sublinear mappings are again sublinear. We recall some basic facts about adjoints of sublinear mappings [35, 37].

The *polar* of a convex cone K is the closed convex cone $K^* = \{u^* \mid \langle u, u^*\rangle \le 0\ \forall u \in K\}$. The *upper adjoint* of a sublinear mapping $S : X \rightrightarrows U$ is the sublinear mapping $S^{*+} : U^* \rightrightarrows X^*$, defined by

$$S^{*+}(u^*) = \{x^* \mid (x^*, -u^*) \in (\operatorname{gph} S)^*\}.$$

The *lower adjoint* of S is defined by

$$S^{*-}(u^*) = \{x^* \mid (-x^*, u^*) \in (\operatorname{gph} S)^*\}.$$

If S is a linear single-valued mapping, then both S^{*+} and S^{*-} reduce to the usual adjoint linear mapping S^*. Since $K^{**} = \operatorname{cl} K$, for any convex cone K, it follows that $(S^{*+})^{*-} = (S^{*-})^{*+} = \operatorname{cl} S$. The inverse formulas $(S^{-1})^{*+} = (S^{*-})^{-1}$, and $(S^{-1})^{*-} = (S^{*+})^{-1}$ are also valid. The following generalizes the formula $(T{\circ}S)^* = S^*{\circ}T^*$ which holds when S and T are linear. For set-valued mappings S and T, $S \subset T$ means that $S(x) \subset T(x)$ for all x.

Lemma 2.3 ([35, Theorem 39.8]). *Let $S : X \rightrightarrows U$ and $T : U \rightrightarrows V$ be sublinear. Then*

$$(T{\circ}S)^* \supset S^*{\circ}T^*,$$

where $*$ stands for either $*+$ or $*-$. If $\operatorname{ri}\operatorname{rge} S \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset$, then equality holds. If S and T are closed, and $\operatorname{ri}\operatorname{rge} T^* \cap \operatorname{ri}\operatorname{dom} S^* \neq \emptyset$, then $T{\circ}S$ is closed, and $(T{\circ}S)^* = \operatorname{cl}(S^*{\circ}T^*)$.

One way to prove this is to use Lemma 2.1 and the rules for computing polar cones [37, Exercise 11.31]. Applying Lemma 2.3 to more special compositions, such as those in Corollary 2.2, one can derive expressions for adjoints of sums and other combinations of sublinear mappings. The following generalizes the fact that $\operatorname{rge} S^* = S^{-1}(0)^\perp$, when S is linear.

Lemma 2.4. *For any sublinear mapping $S : X \rightrightarrows U$,*

$$\begin{aligned}
(\operatorname{dom} S)^* &= S^{*+}(0) = -S^{*-}(0),\\
(\operatorname{rge} S)^* &= (S^{*-})^{-1}(0) = -(S^{*+})^{-1}(0),\\
(\operatorname{cl} S)(0) &= (\operatorname{dom} S^{*-})^* = -(\operatorname{dom} S^{*+})^*,\\
(\operatorname{cl} S)^{-1}(0) &= (\operatorname{rge} S^{*+})^* = -(\operatorname{rge} S^{*-})^*.
\end{aligned}$$

Proof. By definition

$$\begin{aligned}
(\operatorname{dom} S)^* &= (P_X(\operatorname{gph} S))^* = P_X^{*-1}[(\operatorname{gph} S)^*]\\
&= \{x^* \mid (x^*, 0) \in (\operatorname{gph} S)^*\} = S^{*+}(0) = -S^{*-}(0).
\end{aligned}$$

The second formula follows by applying the first to S^{-1}, and using the facts $(S^{-1})^{*+} = (S^{*-})^{-1}$ and $(S^{-1})^{*-} = (S^{*+})^{-1}$. The last two formulas follow by applying the first two to S^{*+}, and using the facts that $(S^{*+})^{*-} = (S^{*-})^{*+} = \operatorname{cl} S$, $\operatorname{dom} S^{*+} = -\operatorname{dom} S^{*-}$ and $\operatorname{rge} S^{*+} = -\operatorname{rge} S^{*-}$. □

3. Recession mappings

Let C be a nonempty convex set. The *recession cone* of C is defined by [35, Section 8]

$$\operatorname{rc} C = \{y \mid x + \tau y \in C,\ \forall x \in C,\ \forall \tau > 0\} = \bigcap_{x\in C}\bigcap_{\alpha>0} \alpha(C - x).$$

As an intersection of convex cones $\bigcap_{\alpha>0} \alpha(C - x)$, $\operatorname{rc} C$ is a convex cone, closed if C is closed. Also, if C is a cone then $\operatorname{rc} C = C$. If C is closed, we have by [35, Corollary 8.3.2]

$$\operatorname{rc} C = \bigcap_{\alpha>0} \alpha(C - x),$$

where $x \in C$ is arbitrary. Since the sets $\alpha(C-x)$ are monotonically decreasing as α goes to zero, we see that the recession cone of a closed convex set does not depend on the form of C on bounded sets. The recession cone describes a *global* behavior of a convex set.

Let $S : X \rightrightarrows U$ be graph-convex. The *recession mapping* of S is the sublinear mapping $RS : X \rightrightarrows U$, whose graph is the recession cone of the graph of S:

$$\operatorname{gph} RS = \operatorname{rc} \operatorname{gph} S.$$

The mapping RS is closed whenever S is. We always have $RS^{-1} = (RS)^{-1}$, and if S is sublinear (e.g. S or S^{-1} is linear), then $RS = S$. The upper adjoint of RS will be denoted by R^*S.

Chain rules for recession mappings are obtained by combining Lemma 2.1 with rules for computing recession cones. Since the conditions in the recession cone formulas also imply the preservation of closedness, we obtain closedness criteria for compositions as byproducts.

Lemma 3.1.
(i) ([35, Theorem 9.1]) *Let $C \subset X$ be closed and convex, and let $A : X \to U$ be linear. Then $\operatorname{rc} A(C) \supset A(\operatorname{rc} C)$, and if $\operatorname{rc} C \cap A^{-1}(0)$ is a subspace, then $A(C)$ is closed, and the inclusion holds as an equality.*
(ii) ([35, Corollary 8.3.3]) *Let C_1 and C_2 are closed and convex sets whose intersection is nonempty. Then $C_1 \cap C_2$ is closed, and*

$$\operatorname{rc}(C_1 \cap C_2) = \operatorname{rc} C_1 \cap \operatorname{rc} C_2.$$

The following lemma will be used to convert the chain rule into an adjoint form.

Lemma 3.2. *Let K_1 and K_2 be convex cones. Then*

$$\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \emptyset \quad \Longleftrightarrow \quad K_1^* \cap -K_2^* \text{ is a subspace.}$$

Proof. By [35, Theorem 6.5], $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \emptyset \Leftrightarrow 0 \in \operatorname{ri} K_1 - \operatorname{ri} K_2 \Leftrightarrow 0 \in \operatorname{ri}(K_1 - K_2)$. Since $K_1 - K_2$ is a cone, this means that $K_1 - K_2$ is a subspace, which is equivalent to $(K_1 - K_2)^* = K_1^* \cap -K_2^*$ being a subspace. □

Theorem 3.3. *Let $S : X \rightrightarrows U$ and $T : U \rightrightarrows V$ be closed and convex, such that $\operatorname{rge} S \cap \operatorname{dom} T \neq \emptyset$. Then*

$$R(T \circ S) \supset RT \circ RS,$$

and if $RS(0) \cap (RT)^{-1}(0)$ is a subspace, then $T \circ S$ is closed, and the inclusion holds as an equality, with

$$R^*(T \circ S) = \operatorname{cl}(R^*S \circ R^*T).$$

Proof. By Lemma 2.1, $\operatorname{gph}(T \circ S) = P_{X \times V}(C)$, where $C = (\operatorname{gph} S \times V) \cap (X \times \operatorname{gph} T)$. By Lemma 3.1(i),

$$\begin{aligned}\operatorname{gph} R(T \circ S) &= \operatorname{rc} \operatorname{gph}(T \circ S) \supset P_{X \times V}(\operatorname{rc} C) \\ &= P_{X \times V}[(\operatorname{rc} \operatorname{gph} S \times V) \cap (X \times \operatorname{rc} \operatorname{gph} T) = \operatorname{gph}(RT \circ RS),\end{aligned}$$

with equality if $\operatorname{rc} C \cap P^{-1}_{X\times V}(0)$ is a subspace, in which case $\operatorname{gph}(T\circ S)$ is closed. By Lemma 3.1(ii),

$$\begin{aligned}\operatorname{rc} C \cap P^{-1}_{X\times V}(0) &= (\operatorname{rc}\operatorname{gph} S \times V) \cap (X \times \operatorname{rc}\operatorname{gph} T) \cap (\{0\} \times U \times \{0\})\\ &= \{0\} \times (RS(0) \cap (RT)^{-1}(0)) \times \{0\},\end{aligned}$$

which is a subspace if and only if $RS(0) \cap (RT)^{-1}(0)$ is a subspace.

The closedness of S and T implies that of RS and RT, which in turn implies the closedness of $RS(0)$ and $(RT)^{-1}(0)$. Thus, by Lemma 3.2, $RS(0) \cap (RT)^{-1}(0)$ is a subspace if and only if $\operatorname{ri}((RT)^{-1}(0))^* \cap \operatorname{ri}(-RS(0))^* \neq \emptyset$, which by Lemma 2.4 may be written as $\operatorname{ri}\operatorname{rge} R^*T \cap \operatorname{ri}\operatorname{dom} R^*S \neq \emptyset$. The expression for $R^*(T\circ S)$ now follows from Lemma 2.3. □

Applying Theorem 3.3 to the special compositions used to construct the mappings in Corollary 2.2, we obtain the following.

Corollary 3.4.

(i) *Let $A : X \to U$ be linear, and let $T : U \rightrightarrows V$ be closed and convex, such that* $\operatorname{rge} A \cap \operatorname{dom} T \neq \emptyset$. *Then $T\circ A$ is closed, and*

$$\begin{aligned}R(T\circ A) &= RT\circ A\\ R^*(T\circ A) &= \operatorname{cl}(A^*\circ R^*T).\end{aligned}$$

(ii) *Let $S : X \rightrightarrows U$ be closed and convex, and let $B : U \to V$ be linear. Then*

$$R(B\circ S) \supset B\circ RS,$$

and if $RS(0) \cap B^{-1}(0)$ is a subspace, then $B\circ S$ is closed, and the inclusion holds as an equality, with

$$R^*(B\circ S) = R^*S\circ B^*.$$

(iii) *Let $S_1, S_2 : X \rightrightarrows U$ be closed and convex, such that* $\operatorname{dom} S_1 \cap \operatorname{dom} S_2 \neq \emptyset$. *Then*

$$R(S_1 + S_2) \supset RS_1 + RS_2,$$

and if $RS_1(0) \cap -RS_2(0)$ is a subspace, then $S_2 + S_2$ is closed, and the inclusion holds as an equality, with

$$R^*(S_1 + S_2) = \operatorname{cl}(R^*S_1 + R^*S_2).$$

(iv) *Let $D \subset U$ and $T : U \rightrightarrows V$ be closed and convex, such that $D \cap \operatorname{dom} T \neq \emptyset$. Then*

$$\operatorname{rc} T(D) \supset RT(\operatorname{rc} D).$$

If $\operatorname{rc} D \cap (RT)^{-1}(0)$ *is a subspace, then $T(D)$ is closed, and the inclusion holds as an equality.*

Proof. In (i), we choose $S = A$, so that $RS(0) = A(0) = 0$, and $RS(0) \cap (RT)^{-1}(0)$ is trivially a subspace. In (ii), the closure operation in the adjoint relation is superfluous, since $\operatorname{dom} B = U$ implies that $R^*S\circ B^*$ is always closed by Lemma 2.3. To obtain (iii), define A, T and B as in the proof of Corollary 2.2(iii), and apply (i) to A and T, and then (ii) to $T\circ A$ and B. In this case, $R(T\circ A)(0) = RS_1(0) \times RS_2(0)$, and $B^{-1}(0) = \{(u_1, u_2) \mid u_2 = -u_1\}$, so that the condition in (ii) becomes $RS_1(0) \cap -RS_2(0)$. Part (iv) follows by defining $S(x) = D$. Then $RS(x) = \operatorname{rc} D$, and $R(T\circ S)(x) = \operatorname{rc} T(D)$, for all x. Also, a constant mapping is closed if and only if its value is closed. □

Note that, if S is linear in (iv), we recover Lemma 3.1(i), and if S is the inverse of a linear mapping, we obtain the closedness criterion in [35, Theorem 6.7], and the recession cone formula in [35, Corollary 8.3.4].

Expressions in (i) and (ii) in the following corollary may be convenient in checking the subspace condition in Theorem 3.3.

Corollary 3.5.
Let $S : X \rightrightarrows U$ be closed and convex, and let $x \in \operatorname{dom} S$ and $u \in \operatorname{rge} S$ be arbitrary. Then

(i) $RS(0) = \operatorname{rc} S(x)$,
(ii) $(RS)^{-1}(0) = \operatorname{rc} S^{-1}(u)$,
(iii) $\operatorname{rge} RS \subset \operatorname{rc} \operatorname{rge} S$, *with equality if* $\operatorname{rc} S^{-1}(u)$ *is a subspace, in which case* $\operatorname{rge} S$ *is closed,*
(iv) $\operatorname{dom} RS \subset \operatorname{rc} \operatorname{dom} S$, *with equality if* $\operatorname{rc} S(x)$ *is a subspace, in which case* $\operatorname{dom} S$ *is closed.*

Proof. Part (i) follows from Corollary 3.4(iv) by choosing $T = S$ and $D = \{x\}$, and (iii) follows by choosing $D = U$. Since $RS^{-1} = (RS)^{-1}$, parts (ii) and (iv) follow by applying (i) and (iii) to S^{-1}. □

4. Graphical derivatives

Let C be a nonempty convex set, and let $x \in C$. The *tangent cone* $T_C(x)$ of C at $x \in C$ is the closure of the *positive hull*

$$\operatorname{pos}(C - x) = \bigcup_{\alpha > 0} \alpha(C - x)$$

of $C - x$ [37]. Thus, $T_C(x)$ is a closed convex cone. If C is a subspace, then $T_C(x) = C$ for all $x \in C$. Since the sets $\alpha(C - x)$ are monotonically increasing with α, we see that $T_C(x)$ does not depend on the form of C outside any neighborhood of x. The tangent cone describes a *local* behavior of C at x.

Let $S : X \rightrightarrows U$ be convex, and let $(x, u) \in \operatorname{gph} S$. The *graphical derivative* [37] (or the *contingent derivative* [1]) of S at (x, u) is the sublinear mapping $DS(x\,|\,u) : X \rightrightarrows U$, whose graph is the tangent cone to $\operatorname{gph} S$ at (x, u):

$$\operatorname{gph} DS(x\,|\,u) = T_{\operatorname{gph} S}(x, u).$$

It is clear that $DS^{-1}(u\,|\,x) = DS(x\,|\,u)^{-1}$, and that if $\operatorname{gph} S$ is a subspace (e.g. S or S^{-1} is linear), then $DS(x\,|\,u) = S$ for all $(x, u) \in \operatorname{gph} S$. The upper adjoint of $DS(x\,|\,u)$ is denoted by $D^*S(x\,|\,u)$, and it is called the *coderivative* of S at (x, u) [26, 27, 37].

We will use the same strategy as in the previous section. The following well known rules for tangent cones correspond to Lemma 3.1. They can be derived for example from the formulas for subdifferentials and polar cones in [35].

Lemma 4.1.
(i) *Let $C \subset X$ be convex, and let $A : X \to U$ be linear. Then for any $u \in A(C)$*

$$T_{A(C)}(u) = \operatorname{cl} A(T_C(x)),$$

where $x \in C \cap A^{-1}(u)$ is arbitrary.

(ii) *Let C_1 and C_2 be convex, and let $x \in C_1 \cap C_2$. Then*

$$T_{C_1 \cap C_2}(x) \subset T_{C_1}(x) \cap T_{C_2}(x),$$

and if $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset$*, then equality holds.*

The following lemma will be used to obtain the adjoint form of the chain rule.

Lemma 4.2. *Let $C_1, C_2 \subset X$ be convex, and let $x \in C_1 \cap C_2$ be arbitrary. Then*

$$\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset \iff \operatorname{ri} T_{C_1}(x) \cap \operatorname{ri} T_{C_2}(x) \neq \emptyset.$$

Proof. We have $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset \Leftrightarrow 0 \in \operatorname{ri}(C_1 - C_2) \Leftrightarrow$ "$T_{C_1-C_2}(0)$ is a subspace". Defining $C = C_1 \times C_2$ and $A(x_1, x_2) = x_1 - x_2$ in Lemma 4.1(i), we have $T_{C_1-C_2}(0) = \operatorname{cl}(T_{C_1}(x) - T_{C_2}(x))$. Thus, $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset$ if and only if $\operatorname{cl}(T_{C_1}(x) - T_{C_2}(x)) = (T_{C_1}(x)^* \cap -T_{C_2}(x)^*)^*$ is a subspace, which by Lemma 3.2 holds if and only if $\operatorname{ri} T_{C_1}(x) \cap \operatorname{ri} T_{C_2}(x) \neq \emptyset$. □

Theorem 4.3. *Let $S : X \rightrightarrows U$ and $T : U \rightrightarrows V$ be convex, and let $(x, v) \in \operatorname{gph}(T \circ S)$. Then*

$$D(T\circ S)(x \,|\, v) \subset \operatorname{cl}(DT(u \,|\, v)\circ DS(x \,|\, u)),$$
$$D^*(T\circ S)(x \,|\, v) \supset D^*S(x \,|\, u)\circ D^*T(u \,|\, v),$$

where $u \in S(x) \cap T^{-1}(v)$ is arbitrary. If $\operatorname{ri} \operatorname{rge} S \cap \operatorname{ri} \operatorname{dom} T \neq \emptyset$*, then equalities hold.*

Proof. By Lemma 2.1, $\operatorname{gph}(T\circ S) = P_{X\times V}(C)$, where $C = (\operatorname{gph} S \times V) \cap (X \times \operatorname{gph} T)$. Thus, by Lemma 4.1(i),

$$\operatorname{gph} D(T\circ S)(x \,|\, v) = T((x, v) |\operatorname{gph} T\circ S) = \operatorname{cl} P_{X\times V} T_C(x, u, v),$$

where $(x, u, v) \in C \cap P_{X\times V}^{-1}(x, v)$ is arbitrary, that is, $u \in S(x) \cap T^{-1}(v)$ is arbitrary. By Lemma 4.1(ii),

$$T_C(x, u, v) \subset T((x, u, v) |\operatorname{gph} S \times V) \cap T((x, u, v) | X \times \operatorname{gph} T)$$
$$= (\operatorname{gph} DS(x \,|\, u) \times V) \cap (X \times \operatorname{gph} DT(u \,|\, v)),$$

with equality under the condition $\operatorname{ri}(\operatorname{gph} S \times V) \cap \operatorname{ri}(X \times \operatorname{gph} T) \neq \emptyset$, which by [35, Theorem 6.8], is equivalent to $\operatorname{ri} \operatorname{rge} S \cap \operatorname{ri} \operatorname{dom} T \neq \emptyset$. Combining,

$$\operatorname{gph} D(T\circ S)(x \,|\, v) \subset \operatorname{cl} P_{X\times V}(\operatorname{gph} DS(x \,|\, u) \times V) \cap (X \times \operatorname{gph} DT(u \,|\, v))$$
$$= \operatorname{cl} \operatorname{gph}(DT(u \,|\, v)\circ DS(x \,|\, u)),$$

with equality if $\operatorname{ri} \operatorname{rge} S \cap \operatorname{ri} \operatorname{dom} T \neq \emptyset$.

The first formula and Lemma 2.3 imply

$$D^*(T\circ S)(x \,|\, v) \supset (DT(u \,|\, v)\circ DS(x \,|\, u))^* \supset D^*S(x \,|\, u)\circ D^*T(u \,|\, v),$$

where the first inclusion holds as an equality if $\operatorname{ri} \operatorname{rge} S \cap \operatorname{ri} \operatorname{dom} T \neq \emptyset$, and then also $\operatorname{ri} T_{\operatorname{rge} S}(u) \cap \operatorname{ri} T_{\operatorname{dom} T}(u) \neq \emptyset$, by Lemma 4.2. Applying Lemma 4.1(i) to $\operatorname{rge} S = P_U(\operatorname{gph} S)$ and $\operatorname{dom} T = P_U(\operatorname{gph} T)$, we see that this is equivalent to $\operatorname{ri} \operatorname{rge} DS(x \,|\, u) \cap \operatorname{ri} \operatorname{dom} DT(u \,|\, v) \neq \emptyset$, so that Lemma 2.3 implies the equality in the second inclusion. □

The following corresponds to Corollary 3.4.

Corollary 4.4.

(i) *Let $A : X \to U$ be linear, and let $T : U \rightrightarrows V$ be convex. Then for any $(x, v) \in \operatorname{gph}(T \circ A)$,*

$$D(T\circ A)(x\,|\,v) \subset DT(Ax\,|\,v)\circ A,$$
$$D^*(T\circ A)(x\,|\,v) \supset A^*\circ D^*T(A(x)\,|\,v),$$

with equalities if $\operatorname{rge} A \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset$.

(ii) *Let $S : X \rightrightarrows U$ be convex, and let $B : U \to V$ be linear. Then*

$$D(B\circ S)(x\,|\,v) = \operatorname{cl}(B\circ DS(x\,|\,u)),$$
$$D^*(B\circ S)(x\,|\,v) = D^*S(x\,|\,u)\circ B^*,$$

where $u \in S(x) \cap B^{-1}(v)$ is arbitrary.

(iii) *Let $S_1, S_2 : X \rightrightarrows U$ be convex. Then*

$$D(S_1 + S_2)(x\,|\,u) \subset \operatorname{cl}(DS_1(x\,|\,u_1) + DS_2(x\,|\,u_2)),$$
$$D^*(S_1 + S_2)(x\,|\,u) \supset D^*S_1(x\,|\,u_1) + D^*S_2(x\,|\,u_2),$$

where $u_i \in S_i(x)$ are arbitrary such that $u_1 + u_2 = u$. If $\operatorname{ri}\operatorname{dom} S_1 \cap \operatorname{ri}\operatorname{dom} S_2 \neq \emptyset$, *then equalities hold.*

(iv) *Let $C \subset X$ and $S : U \rightrightarrows V$ be convex. Then for any $u \in S(C)$*

$$T_{S(C)}(u) \subset \operatorname{cl} DS(x\,|\,u)(T_C(x)),$$
$$N_{S(C)}(u) \supset -D^*S(x\,|\,u)^{-1}(-N_C(x)),$$

where $x \in C \cap S^{-1}(u)$ is arbitrary. If $\operatorname{ri} C \cap \operatorname{ri}\operatorname{dom} S \neq \emptyset$, *then equalities hold.*

Proof. Parts (i)–(iii) follow by applying Theorem 4.3 to the special cases of Corollary 2.2. Part (iv) follows by noting that for any constant mapping $S(x) = D$, we have $DS(x\,|\,u) = T_D(u)$, and $D^*S(x\,|\,u)^{-1} = -N_D(u)$. □

Parts (i), (ii) and (iii) are finite-dimensional versions of Theorem 6, Proposition 7 and Corollary 10, respectively, of [2, Section 4.2]. Note that, if S is linear in (iv), we recover Lemma 4.1(i), and if S is the inverse of a linear mapping A, we obtain the familiar formula $T_{A^{-1}(D)}(x) = \operatorname{cl} A^{-1}(T_D(Ax))$, for $x \in A^{-1}(D)$, where $D \subset U$ is such that $\operatorname{rge} A \cap \operatorname{ri} D \neq \emptyset$. Using the fact that

$$D^*S^{-1}(u\,|\,x) = -D^*S(x\,|\,u)^{-1}\circ(-I), \tag{4.1}$$

we could write the normal cone formula in (iv) as $N_{S(C)}(u) \supset D^*S^{-1}(u\,|\,x)(N_C(x))$.

Corollary 4.5. *Let $S : X \rightrightarrows U$ be convex, and let $(x, u) \in \operatorname{gph} S$ be arbitrary.*

(i) $DS(x\,|\,u)(0) \supset T_{S(x)}(u)$, *and* $\operatorname{dom} D^*S(x\,|\,u) \subset -N_{S(x)}(u)$, *with equality if* $x \in \operatorname{ri}\operatorname{dom} S$,

(ii) $DS(x\,|\,u)^{-1}(0) \supset T_{S^{-1}(u)}(x)$, *and* $\operatorname{rge} D^*S(x\,|\,u) \subset N_{S^{-1}(u)}(x)$, *with equality if* $u \in \operatorname{ri}\operatorname{rge} S$,

(iii) $\operatorname{cl}\operatorname{rge} DS(x\,|\,u) = T_{\operatorname{rge} S}(u)$, *and* $D^*S(x\,|\,u)^{-1}(0) = -N_{\operatorname{rge} S}(u)$,

(iv) $\operatorname{cl}\operatorname{dom} DS(x\,|\,u) = T_{\operatorname{dom} S}(x)$, *and* $D^*S(x\,|\,u)(0) = N_{\operatorname{dom} S}(x)$.

Proof. These are special cases of Corollary 4.4(iv). Part (i) follows by choosing $C = \{x\}$, and noting that the closure is superfluous by Corollary 3.4(iv). Part (ii) follows by applying (i) to S^{-1}, and using (4.1). To get (iii) we choose $D = X$, and (iv) follows from (iii) and (4.1). □

5. Relative interiors of graph-convex mappings

In the differential calculus of the previous section, the relative interiors of the domains and ranges of mappings have an important role. In calculating relative interiors associated with graph-convex mappings, the following concept turns out to be useful. The *relative interior* of a convex mapping $S : X \rightrightarrows U$ is the mapping $\operatorname{ri} S : X \rightrightarrows U$ with the graph

$$\operatorname{gph}\operatorname{ri} S = \operatorname{ri}\operatorname{gph} S.$$

By [35, Theorem 6.2], $\operatorname{ri} S$ is convex. It is also clear that $\operatorname{ri} S^{-1} = (\operatorname{ri} S)^{-1}$, and if $\operatorname{gph} S$ is affine, then $\operatorname{ri} S = S$.

In studying relative interiors of composite mappings we take the approach of the previous sections.

Lemma 5.1.

(i) ([35, Theorem 6.6]) *Let $C \subset X$ be convex, and let $A : X \to U$ be linear. Then* $\operatorname{ri} A(C) = A(\operatorname{ri} C)$.

(ii) ([35, Theorem 6.5]) *Let C_1 and C_2 be convex. If* $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset$, *then* $\operatorname{ri}(C_1 \cap C_2) = \operatorname{ri} C_1 \cap \operatorname{ri} C_2$.

Theorem 5.2. *Let $S : X \rightrightarrows U$ and $T : U \rightrightarrows V$ be convex. If* $\operatorname{ri}\operatorname{rge} S \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset$, *then*

$$\operatorname{ri}(T \circ S) = \operatorname{ri} T \circ \operatorname{ri} S.$$

If in addition, $RS(0) \cap (RT)^{-1}(0)$ is a subspace, then also

$$\operatorname{cl}(T \circ S) = \operatorname{cl} T \circ \operatorname{cl} S.$$

Proof. By Lemma 2.1, $\operatorname{gph}(T \circ S) = P_{X \times V}(C)$, where $C = (\operatorname{gph} S \times V) \cap (X \times \operatorname{gph} T)$. So by Lemma 5.1(i), $\operatorname{gph}\operatorname{ri}(T \circ S) = P_{X \times V}(\operatorname{ri} C)$. Since the condition $\operatorname{ri}\operatorname{rge} S \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset$ is equivalent to $\operatorname{ri}(\operatorname{gph} S \times V) \cap \operatorname{ri}(X \times \operatorname{gph} T) \neq \emptyset$, we have by Lemma 5.1(ii) that

$$\operatorname{gph}\operatorname{ri}(T \circ S) = P_{X \times V}[(\operatorname{ri}\operatorname{gph} S \times V) \cap (X \times \operatorname{ri}\operatorname{gph} T)]$$

which by Lemma 2.1 equals $\operatorname{gph}(\operatorname{ri} T \circ \operatorname{ri} S)$. The second assertion follows similarly by using the closure formulas in Theorems 6.5 and 9.1 of [35]. □

Corollary 5.3.

(i) *Let $A : X \to U$ be linear, and let $T : U \rightrightarrows V$ be convex. If* $\operatorname{rge} A \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset$, *then* $\operatorname{ri}(T \circ A) = \operatorname{ri} T \circ A$, *and* $\operatorname{cl}(T \circ A) = \operatorname{cl} T \circ A$.

(ii) *Let $S : X \rightrightarrows U$ be convex, and let $B : U \to V$ be linear. Then* $\operatorname{ri}(B \circ S) = B \circ \operatorname{ri} S$, *and if $RS(0) \cap B^{-1}(0)$ is a subspace, then also* $\operatorname{cl}(B \circ S) = B \circ \operatorname{cl} S$.

(iii) *Let* $S_1, S_2 : X \rightrightarrows U$ *be convex. If* $\operatorname{ri}\operatorname{dom} S_1 \cap \operatorname{ri}\operatorname{dom} S_2 \neq \emptyset$*, then* $\operatorname{ri}(S_1 + S_2) = \operatorname{ri} S_1 + \operatorname{ri} S_2$*. If in addition* $RS_1(0) \cap -RS_2(0)$ *is a subspace, then also* $\operatorname{cl}(S_1 + S_2) = \operatorname{cl} S_1 + \operatorname{cl} S_2$.

(iv) *Let* $C \subset X$ *and* $S : X \rightrightarrows U$ *be convex. If* $\operatorname{ri} C \cap \operatorname{ri}\operatorname{dom} S \neq \emptyset$*, then* $\operatorname{ri} S(C) = (\operatorname{ri} S)(\operatorname{ri} C)$*. If in addition* $\operatorname{rc} C \cap (RS)^{-1}(0)$ *is a subspace, then also* $\operatorname{cl} S(C) = (\operatorname{cl} S)(\operatorname{cl} C)$*. In particular,* $\operatorname{dom}(\operatorname{ri} S) = \operatorname{ri}\operatorname{dom} S$*, and* $\operatorname{rge}(\operatorname{ri} S) = \operatorname{ri}\operatorname{rge} S$.

(v) *For any graph-convex mapping* S*,*

$$(\operatorname{ri} S)(x) = \begin{cases} \operatorname{ri} S(x) & \text{for } x \in \operatorname{ri}\operatorname{dom} S, \\ \emptyset & \text{for } x \notin \operatorname{ri}\operatorname{dom} S, \end{cases}$$

and for any $x \in \operatorname{ri}\operatorname{dom} S$*,* $(\operatorname{cl} S)(x) = \operatorname{cl} S(x)$.

(vi) *For any graph-convex* $S : X \rightrightarrows U$ *and* $T : U \rightrightarrows V$ *such that* $\operatorname{ri}\operatorname{rge} S \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset$*, we have*

$$\operatorname{ri}\operatorname{dom}(T \circ S) = \{x \in \operatorname{ri}\operatorname{dom} S \mid \operatorname{ri} S(x) \cap \operatorname{ri}\operatorname{dom} T \neq \emptyset\},$$
$$\operatorname{ri}\operatorname{rge}(T \circ S) = \{v \in \operatorname{ri}\operatorname{rge} T \mid \operatorname{ri} T^{-1}(v) \cap \operatorname{ri}\operatorname{rge} S \neq \emptyset\},$$

Proof. Parts (i)–(iv) are obvious counterparts of (i)–(iv) in Corollary 2.2. Part (v) is obtained from (iv) by choosing C to be a singleton. The last part follows by noting that $\operatorname{dom}(T \circ S) = S^{-1}(\operatorname{dom} T)$, so that by (iv) $\operatorname{ri}\operatorname{dom}(T \circ S) = (\operatorname{ri} S)^{-1}(\operatorname{ri}\operatorname{dom} T)$. The final form for $\operatorname{ri}\operatorname{rge}(T \circ S)$ now follows from (v). The second formula follows from the first since $\operatorname{rge}(T \circ S) = \operatorname{dom}(T \circ S)^{-1} = \operatorname{dom}(S^{-1} \circ T^{-1})$. ☐

If S is linear in (iv), we recover Lemma 5.1(i), and if S is the inverse of a linear mapping, we obtain the rule in [35, Theorem 6.7]. The expression for $\operatorname{ri} S$ in (v) is just [35, Theorem 6.8] stated in terms of a set-valued mapping.

6. K-convex functions

Recall that a function f from X to U defined on a set $\operatorname{dom} f$ is said to be *K-convex* if for every $x_1, x_2 \in \operatorname{dom} f$,

$$f(\alpha_1 x_1 + \alpha_2 x_2) - \alpha_1 f(x_1) - \alpha_2 f(x_2) \in K,$$

whenever $\alpha_1, \alpha_2 \in [0, 1]$, and $\alpha_1 + \alpha_2 = 1$. By the definition of polar cone, this implies that the extended-real-valued functions $\langle u^*, f\rangle(x) := \langle u^*, f(x)\rangle$, with $\operatorname{dom}\langle u^*, f\rangle = \operatorname{dom} f$, are convex for every $u^* \in K^*$. Since $K^{**} = \operatorname{cl} K$, the latter condition is equivalent to $(\operatorname{cl} K)$-convexity of f. This provides an important link to the extensive theory of real-valued convex functions [35]. It allows us to use various facts about real-valued convex functions to deduce facts about K-convex functions and vice versa.

For any function f from X to U, let $\mathcal{K}_f$ be the collection of all convex cones with respect to which f is convex. As long as $\operatorname{dom} f$ is convex, we have $U \in \mathcal{K}_f$, so that $\mathcal{K}_f$ is nonempty. By definition, $K_1, K_2 \in \mathcal{K}_f$ implies $K_1 \cap K_2 \in \mathcal{K}_f$, and $K_1 \in \mathcal{K}_f$ and $K_2 \supset K_1$ imply $K_2 \in \mathcal{K}_f$. Thus, a small cone corresponds to a strong property. The following lemma characterizes the smallest closed member of $\mathcal{K}_f$.

Lemma 6.1. *Let f be a function from X to U be such that* $\operatorname{dom} f$ *is convex. Then*

$$K_f^* = \{u^* \in U^* \mid \langle u^*, f\rangle \text{ is convex}\},$$

is the polar of the smallest closed convex cone K_f in $\mathcal{K}_f$.

Proof. Denoting $K_f^\circ = \{u^* \mid \langle u^*, f\rangle \text{ is convex}\}$, we have

$$\begin{aligned} K_f^\circ &= \{u^* \mid \langle u^*, f\rangle (\alpha_1 x_1 + \alpha_2 x_1) \le \alpha_1 \langle u^*, f\rangle (x_1) + \alpha_2 \langle u^*, f\rangle (x_2), \\ &\qquad\qquad\qquad\qquad \forall x_i \in \operatorname{dom} f,\ \alpha_i \in [0,1] : \alpha_1 + \alpha_2 = 1\} \\ &= \{u^* \mid \langle u^*, f(\alpha_1 x_1 + \alpha_2 x_2) - \alpha_1 f(x_1) - \alpha_2 f(x_2)\rangle \le 0, \\ &\qquad\qquad\qquad\qquad \forall x_i \in \operatorname{dom} f,\ \alpha_i \in [0,1] : \alpha_1 + \alpha_2 = 1\}. \end{aligned}$$

Thus, K_f° is a closed convex cone, and it is the largest cone whose polar is in $\mathcal{K}_f$. Since, in terms of inclusion, the polarity operation is order reversing, $(K_f^\circ)^*$ has to be the smallest closed cone in $\mathcal{K}_f$. □

We say that a function f is K-*closed* if $S_{f,K}$ is closed, or equivalently, if $\operatorname{epi}_K f$ is closed. This property of K-convex functions has been studied e.g. in [33, 42]. Just as in the real-valued case, it is easily shown that K-closedness of f implies

(i) All the nonempty sets of the form $\{x \mid f(x) - u \in K\}$ are closed,
(ii) For any $x_k \xrightarrow[\operatorname{dom} f]{} x \in \operatorname{dom} f$, we have $f(x) - \lim f(x_k) \in K$, whenever the limit exists.

It is also easily verified that if K is pointed, as in the real-valued case $K = \mathbb{R}_-$, the converse implications hold. For a real-valued function, $\mathbb{R}_-$-closedness is equivalent to lower-semicontinuity, and $\mathbb{R}_+$-closedness is equivalent to upper-semicontinuity. For any f, $\{0\}$-continuity is equivalent to continuity relative to $\operatorname{dom} f$, and U-continuity is equivalent to closedness of $\operatorname{dom} f$. Note that since the closedness of $S_{f,K}$ implies the closedness of its values $S_{f,K}(x) = f(x) - K$, K-closedness of f implies the closedness of K.

Lemma 6.2. *If K is closed and $\langle u^*, f\rangle$ is lower-semicontinuous for every $u^* \in K^*$ then f is K-closed.*

Proof. Assume, for a contradiction, that there is a sequence $\{(x_i, u_i)\} \subset \operatorname{epi}_K f$ such that $(x_i, u_i) \to (x, u) \notin \operatorname{epi}_K f$. Then, since K is closed, there is a $u^* \in K^*$ such that $\langle u^*, f\rangle (x) > \langle u^*, u\rangle$. Since $\langle u^*, f\rangle (x_i) \le \langle u^*, u_i\rangle$, and $(x_i, \langle u^*, u_i\rangle) \to (x, \langle u^*, u\rangle)$, the function $\langle u^*, f\rangle$ cannot be lower-semicontinuous. □

The converse of Lemma 6.2 does not hold in general as can be seen by taking $u^* = 0$, and $f(x) = 1/x$ if $x > 0$ and $+\infty$ otherwise. Then f is $\mathbb{R}_-$-closed, but $\langle u^*, f\rangle = 0$ for $x > 0$ and 0 otherwise, so that $\langle u^*, f\rangle$ is not lower-semicontinuous. A sufficient condition for the closedness of $\langle u^*, f\rangle$ will be given in Corollary 7.4(ii).

Other kinds of "semicontinuity" properties, which reduce to the familiar one in the case of extended-real-valued functions, have been studied for example in [42, 6]. However, the geometric nature of K-closedness is often more useful in applications. Since K-closedness is equivalent to the graph-closedness of the associated K-profile mapping, we are able to use the general closedness criteria of Section 3 to obtain criteria for K-closedness of composite functions. Just like the recession properties of set-valued mappings were used to guarantee the closedness of their compositions, the closedness of compositions of K-convex functions will depend on the recession properties of the functions involved.

Let f be a K-convex function from X to U. The *K-recession mapping* $R_K f : X \rightrightarrows U$ of f is defined as $R_K f = RS_{f,K}$, or directly by

$$\operatorname{gph} R_K f = \operatorname{rc} \operatorname{epi}_K f.$$

In other words, $u \in R_K f(x)$ means that

$$\begin{aligned} & (x' + \alpha x, u' + \alpha u) \in \operatorname{epi}_K f, \quad \forall (x', u') \in \operatorname{epi}_K f,\ \forall \alpha \geq 0 \\ \Leftrightarrow \quad & f(x' + \alpha x) - u' - \alpha u \in K, \quad \forall (x', u') \in \operatorname{epi}_K f,\ \forall \alpha \geq 0 \\ \Leftrightarrow \quad & f(x' + \alpha x) - f(x') - \alpha u \in K, \quad \forall x' \in \operatorname{dom} f,\ \forall \alpha \geq 0. \end{aligned}$$

The set

$$\operatorname{rc}_K f = (R_K f)^{-1}(0) = \{x \mid f(x' + \alpha x) - f(x') \in K, \quad \forall x' \in \operatorname{dom} f,\ \forall \alpha \geq 0\}$$

will be called the *K-recession cone* of f. It gives the set of directions in which f is nonincreasing with respect to the order induced by K.[1] For an extended-real-valued f, the $\mathbb{R}_-$-recession mapping is the epigraph of the *recession function* of f, and the $\mathbb{R}_-$-recession cone is the ordinary recession cone of f [35, Section 8]. Note that, if f is K-closed, then by [35, Theorem 8.3], $u \in R_K f(x)$ means that

$$f(x' + \alpha x) - f(x') - \alpha u \in K, \quad \forall \alpha \geq 0,$$

where $x' \in \operatorname{dom} f$ is arbitrary. Similarly, for a K-closed f

$$\operatorname{rc}_K f = \{x \mid f(x' + \alpha x) - f(x') \in K, \quad \forall \alpha \geq 0\}.$$

The *K-range* of f is the set

$$\operatorname{rge}_K f = \operatorname{rge} S_{f,K} = P_U(\operatorname{epi}_K f) = \bigcup_{x \in \operatorname{dom} f} f(x) - K = \operatorname{rge} f - K$$

Unlike the ordinary range $\operatorname{rge} f$, the K-range of a K-convex function is always convex.

Corollary 3.5 translates to the following.

Proposition 6.3. *Let f be a K-convex and K-closed function from X to U, and define $L(u) = S_{f,K}^{-1}(u) = \{x \mid f(x) - u \in K\}$. Then*

(i) $R_K f(0) = -K$,

(ii) $\operatorname{rc} L(u) = \operatorname{rc}_K f$, *for any* $u \in \operatorname{rge}_K f$

(iii) $\operatorname{rge} R_K f \subset \operatorname{rc} \operatorname{rge}_K f$, *with equality if* $\operatorname{rc}_K f$ *is a subspace, in which case* $\operatorname{rge}_K f$ *is closed,*

(iv) $\operatorname{dom} R_K f \subset \operatorname{rc} \operatorname{dom} f$, *with equality if K is a subspace, in which case* $\operatorname{dom} f$ *is closed.*

[1] We think of K as a generalization of $\mathbb{R}_-$, so that $u_1 \leq u_2 \Leftrightarrow u_1 - u_2 \in K$.

Part (ii) generalizes the recession cone formula for level sets in [35, Theorem 8.7]. Part (iii) will be useful in guaranteeing the existence of solutions in extremum problems.

As usual, f^* stands for the convex conjugate of an extended-real-valued convex function f:

$$f^*(x^*) = \sup\{\langle x, x^* \rangle - f(x)\}.$$

The upper adjoint of $R_K f$ will be denoted by $R_K^* f$.

Proposition 6.4. *Let f be a K-convex and K-closed function from X to U. Then $R_K^* f$ is the closure of the sublinear mapping $H : U^* \rightrightarrows X^*$ given by*

$$H(u^*) = \begin{cases} \operatorname{dom} \langle u^*, f \rangle^* & \text{for } u^* \in K^*, \\ \emptyset & \text{for } u^* \notin K^*, \end{cases}$$

and the $(\operatorname{cl} K)$-recession cone of f may be expressed as

$$\operatorname{rc}_{\operatorname{cl} K} f = \bigcap_{u^* \in K^*} \operatorname{rc}_{\mathbb{R}_-} \langle u^*, f \rangle .$$

Proof. By [35, Corollary 14.2.1],

$$\begin{aligned} \operatorname{gph} R_K^* f &= \{(u^*, x^*) \mid (x^*, -u^*) \in (\operatorname{gph} R_K f)^*\} \\ &= \{(u^*, x^*) \mid (x^*, -u^*) \in (\operatorname{rc} \operatorname{epi}_K f)^*\} = \operatorname{cl} \left\{ (u^*, x^*) \;\middle|\; \sigma_{\operatorname{epi}_K f}(x^*, -u^*) < \infty \right\}, \end{aligned}$$

where the support function $\sigma_{\operatorname{epi}_K f}$ may be expressed as

$$\begin{aligned} \sigma_{\operatorname{epi}_K f}(x^*, -u^*) &= \sup_{x,u} \{\langle x^*, x \rangle - \langle u^*, u \rangle \mid f(x) - u \in K\} \\ &= \sup_x \{\langle x^*, x \rangle - \langle u^*, f(x) \rangle\} + \sup_{u \in K} \langle u^*, u \rangle = \begin{cases} \langle u^*, f \rangle^* (x^*) & \text{for } u^* \in K^*, \\ +\infty & \text{for } u^* \notin K^*. \end{cases} \end{aligned}$$

This proves the first formula.

By the definition of polar cone,

$$\begin{aligned} \operatorname{rc}_{\operatorname{cl} K} f &= \{y \mid f(x + \alpha y) - f(x) \in \operatorname{cl} K,\ \forall \alpha \geq 0,\ \forall x \in \operatorname{dom} f\} \\ &= \{y \mid \langle u^*, f \rangle (x + \alpha y) \leq \langle u^*, f \rangle (x),\ \forall u^* \in K^*,\ \forall \alpha \geq 0,\ \forall x \in \operatorname{dom} f\} \\ &= \left\{ y \;\middle|\; y \in \operatorname{rc}_{\mathbb{R}_-} \langle u^*, f \rangle,\ \forall u^* \in K^* \right\} = \bigcap_{u^* \in K^*} \operatorname{rc}_{\mathbb{R}_-} \langle u^*, f \rangle . \end{aligned}$$

□

Since $\operatorname{dom} \langle u^*, f \rangle = \operatorname{dom} f$, the domain of $\langle u^*, f \rangle^*$ is nonempty for any $u^* \in K^*$ [35, Theorem 12.2], so that $\operatorname{dom} R_K^* f = K^*$. Hence, by Corollary 5.3(v)

$$R_K^* f(u^*) = \operatorname{cl} \operatorname{dom} \langle u^*, f \rangle^* \quad \forall u^* \in \operatorname{ri} K^*.$$

For extended-real-valued f, this gives the following one-to-one correspondence between the recession mapping of f and the closure of $\operatorname{dom} f^*$:

$$\begin{aligned} \operatorname{cl} \operatorname{dom} f^* &= R_{\mathbb{R}_-}^* f(1), \\ \operatorname{gph} R_{\mathbb{R}_-}^* f &= \operatorname{cl} \operatorname{pos} \{(1, x^*) \mid x^* \in \operatorname{dom} f^*\} . \end{aligned}$$

Note that, by [35, Theorem 13.3], the recession function of f is the support function of $\operatorname{dom} f^*$ (or $\operatorname{cl}\operatorname{dom} f^*$), so that the above relation can be used to translate facts about the recession mapping into facts about the recession function. Since $(R_K f)^{-1}(0) = (\operatorname{rge} R_K^* f)^*$ by Lemma 2.4, we obtain in particular

$$(\operatorname{rc}_{\mathbb{R}_-} f)^* = (R_{\mathbb{R}_-} f)^{-1}(0)^* = \operatorname{cl}\operatorname{rge} R_{\mathbb{R}_-}^* f = \operatorname{cl}\operatorname{pos}\operatorname{dom} f^*,$$

which agrees with [35, Theorem 14.2].

The special form of K-profile mappings has some interesting consequences also in graphical differentiation. Once the X-component in the argument of the graphical derivative of a K-profile mapping has been fixed, a natural choice for the U-component is the corresponding value of f. The *K-Jacobian* $D_K f(x) : X \rightrightarrows U$ of f at x is defined by $D_K f(x) = DS_{f,K}(x \mid f(x))$, or directly by

$$\operatorname{gph} D_K f(x) = T_{\operatorname{epi}_K f}(x, f(x)).$$

If f is extended-real-valued, then $D_{\mathbb{R}_-} f(x)$ is the epigraph of the closure of the directional derivative function of f at x. The upper adjoint of $D_K f(x)$ will be denoted by $D_K^* f(x)$.

Proposition 6.5. *Let f be a K-convex function from X to U, and let $x \in \operatorname{dom} f$. Then $D_K^* f(x)$ is given by*

$$D_K^* f(x)(u^*) = \begin{cases} \partial \langle u^*, f\rangle (x) & \text{if } u^* \in K^*, \\ \emptyset & \text{if } u^* \notin K^*. \end{cases}$$

Proof. Because $(\operatorname{gph} D_K f(x))^* = (T_{\operatorname{epi}_K f}(x, f(x)))^* = N_{\operatorname{epi}_K f}(x, f(x))$, we have

$$\begin{aligned} D_K^* f(x)(u^*) &= \left\{x^* \;\middle|\; (x^*, -u^*) \in N_{\operatorname{epi}_K f}(x, f(x))\right\} \\ &= \{x^* \mid \langle x' - x, x^*\rangle + \langle f(x) - u', u^*\rangle \le 0 \;\forall (x', u') \in \operatorname{epi}_K f\} \\ &= \{x^* \mid \langle x' - x, x^*\rangle + \langle f(x) - f(x') + v', u^*\rangle \le 0 \;\forall x' \in \operatorname{dom} f,\; v' \in K\}. \end{aligned}$$

Hence, for $u^* \notin K^*$, $D_K^* f(x)(u^*) = \emptyset$, and for $u^* \in K^*$,

$$D_K^* f(x)(u^*) = \{x^* \mid \langle u^*, f(x')\rangle \ge \langle u^*, f(x)\rangle + \langle x^*, x' - x\rangle \;\; \forall x' \in \operatorname{dom} f\} = \partial \langle u^*, f\rangle (x).$$

□

This is a convex version of the scalarization formula [27, Proposition 2.11] for locally Lipschitz functions. In the above proposition, Lipschitz continuity is replaced by K-convexity. If f is differentiable at x, so is $\langle u^*, f\rangle$, and $D_K^* f(x)$ reduces to the adjoint of the classical Jacobian on K^*:

$$D_K^* f(x)(u^*) = \begin{cases} \{\nabla f(x)^*(u^*)\} & \text{if } u^* \in K^*, \\ \emptyset & \text{if } u^* \notin K^*. \end{cases}$$

Proposition 6.5 gives the following one-to-one correspondence between $D_{\mathbb{R}_-}^* f(x)$ and the subdifferential of an extended-real-valued convex function:

$$\begin{aligned} \partial f(x) &= D_{\mathbb{R}_-}^* f(x)(1), \\ \operatorname{gph} D_{\mathbb{R}_-}^* f(x) &= \operatorname{cl}\operatorname{pos}\{(1, x^*) \mid x^* \in \partial f(x)\}. \end{aligned}$$

Through this, facts about K-Jacobians can be translated into facts about subdifferentials of extended-real-valued convex functions.

The following is obtained from Corollary 4.5.

Proposition 6.6. *Let $f : X \to U$ be K-convex, and $L(u) = S_{f,K}^{-1}(u) = \{x | f(x) - u \in K\}$. Then for any $x \in \operatorname{dom} f$*

(i) $D_K f(x)(0) \supset -K$, *and* $\operatorname{dom} D_K^* f(x) \subset K^*$, *with equality if* $x \in \operatorname{ri} \operatorname{dom} f$,
(ii) $D_K f(x)^{-1}(0) \supset T_{L(f(x))}(x)$, *and* $\operatorname{rge} D_K^* f(x) \subset N_{L(f(x))}(x)$, *with equality if* $f(x) \in \operatorname{ri} \operatorname{rge}_K f$,
(iii) $\operatorname{cl} \operatorname{rge} D_K f(x) = T_{\operatorname{rge}_K f}(f(x))$, *and* $D_K^* f(x)^{-1}(0) = -N_{\operatorname{rge}_K f}(f(x))$,
(iv) $\operatorname{cl} \operatorname{dom} D_K f(x) = T_{\operatorname{dom} f}(x)$, *and* $D_K^* f(x)(0) = N_{\operatorname{dom} f}(x)$.

By Proposition 6.5, (i) may be expressed as $\partial \langle u^*, f \rangle (x) \neq \emptyset$ for any $u^* \in K^*$ and $x \in \operatorname{ri} \operatorname{dom} f$. Part (ii) gives an expression for the normal cone of a "level set", and it generalizes the corresponding result [35, Theorem 23.7] for extended-real-valued functions. By Corollary 3.5(i), part (iv) implies that $\operatorname{rc} D_K^* f(x)(u^*) = N_{\operatorname{dom} f}(x)$ for any $u^* \in \operatorname{dom} D_K^* f(x)$, so that in particular, $D_K^* f(x)(u^*) = \partial \langle u^*, f \rangle (x)$ is nonempty and bounded for any $u^* \in K^*$, if and only if $x \in \operatorname{int} \operatorname{dom} f$ [35, Theorem 8.4]. These facts agree with [35, Theorem 23.4].

Combining (v) and (iv) of Corollary 5.3, we obtain the following.

Proposition 6.7. *Let f be a K-convex function f. Then*

$$(\operatorname{ri} S_{f,K})(x) = \begin{cases} f(x) - \operatorname{ri} K & \text{if } x \in \operatorname{ri} \operatorname{dom} f, \\ \emptyset & \text{otherwise}, \end{cases}$$

and for $x \in \operatorname{ri} \operatorname{dom} f$, $(\operatorname{cl} S_{f,K})(x) = f(x) - \operatorname{cl} K$. Hence,

$$\operatorname{ri} \operatorname{rge}_K f = \bigcup \{ f(x) - \operatorname{ri} K \mid x \in \operatorname{ri} \operatorname{dom} f \} .$$

7. Convex composite functions

In this section f will be a K-convex function from X to U, and g will be an L-convex function (L is a convex cone) from U to V. The composition of f and g is defined by

$$\operatorname{dom} g \circ f = \{ x \in \operatorname{dom} f \mid f(x) \in \operatorname{dom} g \} ,$$
$$(g \circ f)(x) = g(f(x)), \quad \forall x \in \operatorname{dom} g \circ f.$$

A classical criterion for L-convexity of the composition $g \circ f$ is that g is nonincreasing in the directions of K, in the sense that for any $u \in \operatorname{dom} g$ and $v \in K$, we have $u + v \in \operatorname{dom} g$ and

$$g(u + v) - g(u) \in L.$$

In our notation this can be expressed compactly as $K \subset \operatorname{rc}_L g$. Although often sufficient, this condition is far from being necessary, and it precludes some interesting applications we have in mind. As an example, let $X = U = \mathbb{R}$, $K = L = \mathbb{R}_-$, $f(x) = e^x$, and $g(u) = u^2$. Then $K \not\subset \operatorname{rc}_L g = \{0\}$, but still $(g \circ f)(x) = e^{2x}$ is convex.

The above condition ignores the fact that the behavior of g outside the range

$$\operatorname{rge} f = \bigcup_{x \in \operatorname{dom} f} f(x)$$

of f has no effect on the composition. A more general condition is the following

$$u_1 \in \operatorname{dom} g,\ u_2 \in \operatorname{rge} f,\ u_2 - u_1 \in K \quad \Longrightarrow \quad u_2 \in \operatorname{dom} g,\ g(u_2) - g(u_1) \in L. \tag{$\mathcal{C}$}$$

Note that this is implied in particular by $K \subset \operatorname{rc}_L g$ which can be written as

$$u_1 \in \operatorname{dom} g,\ u_2 - u_1 \in K \quad \Longrightarrow \quad u_2 \in \operatorname{dom} g,\ g(u_2) - g(u_1) \in L.$$

For an extended-real-valued g, $(\mathcal{C})$ can be stated more simply as

$$u_2 \in \operatorname{rge} f,\ u_2 - u_1 \in K \quad \Longrightarrow \quad g(u_2) \le g(u_1),$$

or as

$$g_K(u) = g(u) \quad \forall u \in \operatorname{rge} f,$$

where $g_K(u) = \inf_{v \in K} g(u - v)$. This condition says that g can be replaced by g_K without effecting the composition. Since necessarily, $K \subset \operatorname{rc}_{\mathbb{R}_-} g_K$, we see that the composition $g_K \circ f$ is always convex. The following lemma shows that condition $(\mathcal{C})$ works for vector-valued functions as well.

Lemma 7.1. *Let f be a K-convex function from X to U, and g be an L-convex function from U to V. If $(\mathcal{C})$ holds, then*

$$S_{g \circ h, L} = S_{g,L} \circ S_{h,K},$$

and thus, $g \circ h$ is L-convex.

Proof. We have

$$\begin{aligned} \operatorname{dom} S_{g \circ f, L} &= \{x \in \operatorname{dom} f \mid f(x) \in \operatorname{dom} g\} \\ &\subset \{x \in \operatorname{dom} f \mid f(x) \in \operatorname{dom} g + K\} \\ &= \{x \in \operatorname{dom} f \mid (f(x) - K) \cap \operatorname{dom} g \neq \emptyset\} = \operatorname{dom}(S_{g,L} \circ S_{f,K}). \end{aligned}$$

Assume $x \in \operatorname{dom}(S_{g,L} \circ S_{f,K})$, *i.e.* there exists an $u \in K$, such that $f(x) - u \in \operatorname{dom} g$. Choosing $u_1 = f(x) - u$ and $u_2 = f(x)$ in $(\mathcal{C})$, we see that $f(x) \in \operatorname{dom} g$, so

$$\operatorname{dom} S_{g \circ f, L} = \operatorname{dom}(S_{g,L} \circ S_{f,K}).$$

For any $x \in \operatorname{dom} S_{g \circ f, L}$,

$$\begin{aligned} (S_{g,L} \circ S_{f,K})(x) &= S_{g,L}(S_{f,K}(x)) \\ &= \bigcup \{g(f(x) - u) - L \mid u \in K,\ f(x) - u \in \operatorname{dom} g\} \\ &= \bigcup \{g(f(x) - u) - g(f(x)) \mid u \in K,\ f(x) - u \in \operatorname{dom} g\} + g(f(x)) - L, \end{aligned}$$

where the first term contains zero, so that $(S_{g,L} \circ S_{f,K})(x) \supset g(f(x)) - L = S_{g \circ f, L}(x)$. On the other hand, $(\mathcal{C})$ implies that the first term is contained in $-L$, so that

$$(S_{g,L} \circ S_{f,K})(x) \subset -L + g(f(x)) - L = g(f(x)) - L = S_{g \circ f, L}(x).$$

$\square$

The following corresponds to Corollary 2.2.

Corollary 7.2.

(i) *If $A : X \to U$ is linear and g is an L-convex function from U to V, then $g{\circ}A$ is L-convex.*

(ii) *If f is a K-convex function from X to U, and $B : U \to V$ is linear, then $B{\circ}f$ is $B(K)$-convex.*

(iii) *If f_1 and f_2 are functions from X to U, K_1-convex and K_2-convex respectively, then $f_1 + f_2$ is $(K_1 + K_2)$-convex.*

(iv) *Let f be a K-convex function from X to U, and let D be a convex subset of U. If $(D + K) \cap (\operatorname{rge} f) \subset D$, then $f^{-1}(D)$ is convex. In particular, $f^{-1}(D)$ is convex whenever $K \subset \operatorname{rc} D$.*

Proof. Since A is affine, and since we always have $0 \in \operatorname{rc}_L g$, the condition $K \subset \operatorname{rc}_L g$ holds trivially in (i). In (ii), choose $g = B$ and $L = B(K)$. Then

$$\operatorname{rc}_L g = \{u \mid B(\bar{u} + \alpha u) - B(\bar{u}) \in B(K), \forall \alpha \geq 0, \forall \bar{u} \in U\} = \{u \mid B(u) \in B(K)\} \supset K,$$

so that $B{\circ}f$ is $B(K)$-convex. Part (iii) follows from (i) and (ii) by writing $f_1 + f_2 = B{\circ}f{\circ}A$, where $Ax = (x, x)$, $f(x_1, x_2) = (f_1(x_1), f_2(x_2))$, and $B(u_1, u_2) = u_1 + u_2$. The set $f^{-1}(D)$ in part (iv) may be written as $\operatorname{dom} g{\circ}f$, where $g = \delta_D$. Since $g = 0$ on D, the condition $(\mathcal{C})$ may be written as

$$u_1 \in D,\ u_2 \in (K + u_1) \cap \operatorname{rge} f \quad \Longrightarrow \quad u_2 \in D,$$

or $(K + D) \cap \operatorname{rge} f \subset D$. So $g{\circ}f$ is convex which implies the convexity of $\operatorname{dom} g{\circ}f = f^{-1}(D)$. □

The following is obtained by combining Lemma 7.1 and Theorem 3.3.

Theorem 7.3. *Let f be a K-convex function from X to U, and let g be an L-convex function from U to V, such that* $\operatorname{dom} g{\circ}f \neq \emptyset$ *and $(\mathcal{C})$ holds. If f is K-closed, and g is L-closed, then*

$$R_L(g{\circ}f) \supset R_L g{\circ}R_K f.$$

If in addition, $(-K) \cap \operatorname{rc}_L g$ is a subspace, then $g{\circ}f$ is L-closed, and the inclusion holds as an equality, with

$$R_L^*(g{\circ}f) = \operatorname{cl}(R_K^* h{\circ}R_L^* g).$$

Proof. Since by Lemma 7.1, $S_{g\circ f,L} = S_{g,L}{\circ}S_{f,K}$, we may apply Theorem 3.3. The special form of the subspace condition follows from the definition of $\operatorname{rc}_L g$ and Proposition 6.3(i). □

Corollary 7.4.

(i) *Let $A : X \to U$ be linear, and let g be an L-convex function from U to V, such that* $\operatorname{dom} g{\circ}A \neq \emptyset$*. If g is L-closed, then $g{\circ}A$ is L-closed, and*

$$R_L(g{\circ}A) = R_L g{\circ}A,$$
$$R_L^*(g{\circ}A) = \operatorname{cl}(A^*{\circ}R_L^* g).$$

(ii) *Let f be a K-convex and $B : U \to V$ be linear. If f is K-closed, then*

$$R_{B(K)}(B \circ f) \supset B \circ R_K f.$$

If in addition, $K \cap B^{-1}(0)$ is a subspace, then $B \circ f$ is $B(K)$-closed, and the inclusion holds as an equality, with

$$R^*_{B(K)}(B \circ f) = R^*_K f \circ B^*.$$

In particular, $\langle u^*, f \rangle$ is closed for any $u^* \in \operatorname{ri} K^*$.

(iii) *Let f_i be a K_i-convex and K_i-closed function from X to U, such that $\operatorname{dom} f_1 \cap \operatorname{dom} f_2 \neq \emptyset$. Then*

$$R_{K_1+K_2}(f_1 + f_2) \supset R_{K_1} f_1 + R_{K_2} f_2.$$

If in addition, $(-K_1) \cap K_2$ is a subspace, then $f_1 + f_2$ is $(K_1 + K_2)$-closed, and the inclusion holds as an equality, with

$$R^*_{K_1+K_2}(f_1 + f_2) = \operatorname{cl}(R^*_{K_1} f_1 + R^*_{K_2} f_2).$$

(iv) *Let f and g be as in the theorem. Then*

$$\operatorname{rc}_L g \circ f \supset (R_K f)^{-1}(\operatorname{rc}_L g),$$

with equality if $(-K) \cap \operatorname{rc}_L g$ is a subspace.

(v) *Let f be a K-convex function from X to U, and let D be a convex subset of U, such that $f^{-1}(D) \neq \emptyset$ and $(D + K) \cap \operatorname{rge} f \subset D$. If f is K-closed, and D is closed, then*

$$\operatorname{rc} f^{-1}(D) \supset (R_K f)^{-1}(\operatorname{rc} D).$$

If $(-K) \cap \operatorname{rc} D$ is a subspace, then $f^{-1}(D)$ is closed, and equality holds.

Lemma 7.1, Proposition 6.7 and Corollary 5.3 yield the following.

Lemma 7.5. *Let f be a K-convex function from X to U, and let D be a convex subset of U such that $(D + K) \cap (\operatorname{rge} f) \subset D$. Then $\operatorname{ri} \operatorname{rge}_K f \cap \operatorname{ri} D \neq \emptyset$ if and only if there is an $x \in \operatorname{ri} \operatorname{dom} f$ such that $h(x) \in \operatorname{ri} D$, in which case*

$$\operatorname{ri} f^{-1}(D) = \{x \in \operatorname{ri} \operatorname{dom} f \mid f(x) \in \operatorname{ri} D\}.$$

The following gives a chain rule for K-Jacobians.

Theorem 7.6. *Let f be a K-convex function from X to U, and let g be an L-convex function from U to V, such that $(\mathcal{C})$ holds. Then*

$$D_L(g \circ f)(x) \subset \operatorname{cl}(D_L g(f(x)) \circ D_K f(x)),$$
$$D^*_L(g \circ f)(x) \supset D^*_K f(x) \circ D^*_L g(f(x)),$$

with equalities if $\operatorname{ri} \operatorname{rge}_K f \cap \operatorname{ri} \operatorname{dom} g \neq \emptyset$ which is equivalent to

$$\exists x \in \operatorname{ri} \operatorname{dom} f : \quad f(x) \in \operatorname{ri}(\operatorname{dom} g + K).$$

Proof. The chain rules are obtained by combining Lemma 7.1 and Theorem 4.3. The alternative form of the constraint qualification is obtained by applying the previous lemma to f and $\operatorname{dom} g$. □

In the case $K \subset \operatorname{rc}_L g$, we have $\operatorname{dom} g + K = \operatorname{dom} g$ so that the constraint qualification simplifies slightly.

Corollary 7.7.

(i) *If $A : X \to U$ is linear, and g is an L-convex function from U to V, then*

$$D_L(g \circ A)(x) \subset D_L g(Ax) \circ A,$$
$$D_L^*(g \circ A)(x) \supset A^* \circ D_L^* g(Ax),$$

with equalities if $\operatorname{rge} A \cap \operatorname{ri} \operatorname{dom} g \neq \emptyset$.

(ii) *If f is a K-convex function from X to U, and $B : U \to V$ is a linear, then*

$$D_{B(K)}(B \circ f)(x) = \operatorname{cl}(B \circ D_K f(x)),$$
$$D_{B(K)}^*(B \circ f)(x) = D_K^* f(x) \circ B^*.$$

(iii) *If f_i is a K_i-convex function from X to U, then*

$$D_{K_1+K_2}(f_1 + f_2)(x) \subset \operatorname{cl}(D_{K_1} f_1(x) + D_{K_2} f_2(x)),$$
$$D_{K_1+K_2}^*(f_1 + f_2)(x) \supset D_{K_1}^* f_1(x) + D_{K_2}^* f_2(x),$$

with equalities if $\operatorname{ri} \operatorname{dom} f_1 \cap \operatorname{ri} \operatorname{dom} f_2 \neq \emptyset$.

(iv) *Let f be a K-convex function from X to U, and let g be a convex function from U to $\mathbb{R}$, such that $(\mathcal{C})$ holds. Then*

$$\partial(g \circ f)(x) \supset D_K^* f(x)(\partial g(f(x))),$$

with equality if $\operatorname{ri} \operatorname{rge}_K f \cap \operatorname{ri} \operatorname{dom} g \neq \emptyset$.

(v) *Let f be a K-convex function from X to U, and let D be convex subset of U, such that $f^{-1}(D) \neq \emptyset$ and $(D + K) \cap \operatorname{rge} f \subset D$. Then*

$$N_{f^{-1}(D)}(x) \supset D_K^* f(x)(N_D(f(x))),$$

with equality if $\operatorname{ri} \operatorname{rge}_K f \cap \operatorname{ri} D \neq \emptyset$.

Part (iv) gives a subdifferential chain rule for extended-real-valued compositions. By Proposition 6.5, it can be restated completely in terms of subdifferentials as

$$\partial(g \circ f)(x) \supset \bigcup \{\partial \langle u^*, f \rangle (x) \mid u^* \in \partial g(f(x)) \cap K^*\}.$$

This formula is closely related to [15, Corollary 8.1] which gives the subdifferential in terms of ϵ-subdifferentials of g and f, for the case $K \subset \operatorname{rc}_L g$. For locally Lipschitz continuous f, there is a nonconvex inclusion version of this for generalized subgradients [37]. In the special case where f is differentiable at x, the above formula simplifies to

$$\partial(g \circ f)(x) \supset \nabla f(x)^*(\partial g(f(x))).$$

If $f = A$ for a linear $A : X \to U$, we may choose $K = \{0\}$, so that condition $(\mathcal{C})$ holds trivially, and the constraint qualification reduces to $\operatorname{rge} A \cap \operatorname{ri} \operatorname{dom} g \neq \emptyset$. Thus, we obtain [35, Theorem 23.9] with

$$\partial(g \circ A)(x) \supset A^*(\partial g(A(x))).$$

8. Composite model of convex programming

Let f_0 and g be extended-real-valued convex functions on X and U, respectively, and let f be a K-convex function from X to U, such that $(\mathcal{C})$ holds. Consider the problem

$$\text{minimize } f_0 + g\circ f \text{ over } \operatorname{dom}(f_0 + g\circ f), \tag{$\mathcal{P}$}$$

where $\operatorname{dom}(f_0 + g\circ f) = \{x \in \operatorname{dom} f_0 \cap \operatorname{dom} f \mid f(x) \in \operatorname{dom} g\}$.

When $g = \delta_D$, such that $(D + K) \cap \operatorname{rge}_K f \subset D$, $(\mathcal{P})$ may be written with more explicit constraints as

$$\text{minimize } f_0(x) \text{ subject to } f(x) \in D.$$

With $f(x) = Ax$ and $K = \{0\}$, condition $(\mathcal{C})$ is satisfied for any g, and we obtain the Fenchel-Rockafellar model [35, Section 31]

$$\text{minimize } f_0(x) + g(Ax).$$

Proposition 8.1.

(i) *Assume that f_0 and g are closed, f is K-closed, and $-K \cap \operatorname{rc}_{\mathbb{R}_-} g$ is a subspace. Then the condition*

$$0 \in \operatorname{ri}(\operatorname{dom} f_0^* + R_K^* f(\operatorname{dom} g^*)), \tag{$\mathcal{C}_1$}$$

guarantees that $(\mathcal{P})$ has a solution. Moreover, $(\mathcal{C}_1)$ holds if and only if there is a $u^* \in \operatorname{ri}\operatorname{dom} g^*$ such that $\operatorname{ri}\operatorname{dom} f_0^* \cap -\operatorname{ri}\operatorname{dom}\langle u^*, f\rangle^* \neq \emptyset$.

(ii) *The condition*

$$0 \in \partial f_0(\bar{x}) + D_K^* f(\bar{x})(\partial g(f(\bar{x}))), \tag{KKT}$$

is always sufficient for optimality of $\bar{x}$ in $(\mathcal{P})$, and if

$$\exists x \in \operatorname{ri}\operatorname{dom} f_0 \cap \operatorname{ri}\operatorname{dom} f: \quad f(x) \in \operatorname{ri}(\operatorname{dom} g + K), \tag{$\mathcal{C}_2$}$$

it is also necessary.

Proof. The function $f_0 + g\circ f$ has a minimizer if and only if its $\mathbb{R}_-$-range is closed and bounded from below. By Proposition 6.3(iii) a sufficient condition for the closedness is that $f_0 + g\circ f$ is closed and its recession cone is a subspace. This also implies the boundedness. By Corollary 7.4(iii) (or [35, Theorem 9.3]) and Theorem 7.3, the conditions in (i) guarantee that $f_0 + g\circ f$ is closed. Also, by Lemma 2.4

$$\begin{aligned} \operatorname{rc}_{\mathbb{R}_-}(f_0 + g\circ f) &= (R_{\mathbb{R}_-}(f_0 + g\circ f))^{-1}(0) \\ &= \operatorname{cl}\operatorname{rge} R_{\mathbb{R}_-}^*(f_0 + g\circ f) = \operatorname{cl}\operatorname{pos} R_{\mathbb{R}_-}^*(f_0 + g\circ f)(1), \end{aligned}$$

where the last equality holds since the range space of $f_0 + g\circ f$ is the real-line. Thus $\operatorname{rc}_{\mathbb{R}_-}(f_0 + g\circ f)$ is a subspace if and only if $0 \in \operatorname{ri} R_{\mathbb{R}_-}^*(f_0 + g\circ f)(1)$, where by Corollary 7.4, Theorem 7.3 and Proposition 6.4

$$\begin{aligned} R_{\mathbb{R}_-}^*(f_0 + g\circ f)(1) &= (R_{\mathbb{R}_-}^* f_0 + R_K^* f\circ R_{\mathbb{R}_-}^* g)(1) \\ &= R_{\mathbb{R}_-}^* f_0(1) + R_K^* f(R_{\mathbb{R}_-}^* g(1)) = \operatorname{dom} f_0^* + R_K^* f(\operatorname{dom} g^*). \end{aligned}$$

This proves the first part of (i). By the remarks after Proposition 6.4, $(\mathrm{rc}_{\mathbb{R}_-}\, g)^* = \mathrm{cl}\,\mathrm{pos}\,\mathrm{dom}\, g^*$, so that by Lemma 3.2, $-K \cap \mathrm{rc}_{\mathbb{R}_-}\, g$ is a subspace if and only if $\mathrm{ri}\, K^* \cap \mathrm{ri}\,\mathrm{cl}\,\mathrm{pos}\,\mathrm{dom}\, g^* \neq \emptyset$ or equivalently $\mathrm{ri}\, K^* \cap \mathrm{ri}\,\mathrm{dom}\, g^* \neq \emptyset$. Since $\mathrm{dom}\, R_K^* f = K^*$, Corollary 5.3(iv) implies that

$$\begin{aligned}\mathrm{ri}(\mathrm{dom}\, f_0^* + R_K^* f(\mathrm{dom}\, g^*)) &= \mathrm{ri}\,\mathrm{dom}\, f_0^* + \mathrm{ri}\, R_K^* f(\mathrm{dom}\, g^*)\\ &= \mathrm{ri}\,\mathrm{dom}\, f_0^* + (\mathrm{ri}\, R_K^* f)(\mathrm{ri}\,\mathrm{dom}\, g^*).\end{aligned}$$

The second claim of (i) now follows from Proposition 6.4 and Corollary 5.3(v).

By (iii) (or [35, Theorem 23.8]) and (iv) of Corollary 7.7,

$$\partial(f_0 + g\circ f)(x) \supset \partial f_0(x) + D_K^* f(\partial g(x)),$$

which proves the sufficiency. Applying Corollary 5.3(vi) to $S = S_{f,K}$ and $T = S_{g,\mathbb{R}_-}$, we obtain

$$\begin{aligned}\mathrm{ri}\,\mathrm{dom}(g\circ f) &= \{x \in \mathrm{ri}\,\mathrm{dom}\, f \mid \mathrm{ri}(f(x) - K) \cap \mathrm{ri}\,\mathrm{dom}\, g \neq \emptyset\}\\ &= \{x \in \mathrm{ri}\,\mathrm{dom}\, f \mid f(x) \in \mathrm{ri}(\mathrm{dom}\, g + K)\},\end{aligned}$$

so that $(\mathcal{C}_2)$ implies that the conditions for equality in both (iii) and (iv) of Corollary 7.7 are satisfied, so that (KKT) becomes necessary for $0 \in \partial(f_0 + g\circ f)(\bar{x})$. □

Condition (KKT) means that there exists a $\bar{u}^*$ such that

$$\begin{aligned}0 &\in \partial f_0(\bar{x}) + \partial\langle \bar{u}^*, f\rangle(x),\\ \bar{u}^* &\in \partial g(f(\bar{x})).\end{aligned}$$

The vector $\bar{u}^*$ may be interpreted as a Kuhn-Tucker vector for $(\mathcal{P})$. If g is closed, these conditions can be derived from the conjugate duality framework where $\bar{u}^*$ is a solution to a maximization problem dual to $(\mathcal{P})$ [36]. The above derivation is based solely on subdifferential calculus, and the closedness of g is irrelevant.

When $g = \delta_K$, the condition $\bar{u}^* \in \partial g(f(\bar{x}))$ reduces to the familiar complementarity condition

$$f(\bar{x}) \in K, \quad \bar{u}^* \in K^*, \quad \langle \bar{u}^*, f(\bar{x})\rangle = 0,$$

and conditions $(\mathcal{C}_1)$ and $(\mathcal{C}_2)$ may be written as

$$\begin{gathered}0 \in \mathrm{ri}(\mathrm{dom}\, f_0^* + \mathrm{rge}\, R_K^* f),\\ \exists x \in \mathrm{ri}\,\mathrm{dom}\, f_0 \cap \mathrm{ri}\,\mathrm{dom}\, f: \quad f(x) \in \mathrm{ri}\, K,\end{gathered}$$

the latter being a generalization of the classical Slater condition. For the Fenchel-Rockafellar model $f(x) = f_0(x) + g(Ax)$ with $K = \{0\}$, the subspace condition in Proposition 8.1(i) is trivially satisfied, and conditions $(\mathcal{C}_1)$ and $(\mathcal{C}_2)$ become

$$\begin{gathered}\exists x \in \mathrm{ri}\,\mathrm{dom}\, f_0: \quad A(x) \cap \mathrm{ri}\,\mathrm{dom}\, g \neq \emptyset,\\ \exists u^* \in \mathrm{ri}\,\mathrm{dom}\, g^*: \quad A^*(u^*) \cap \mathrm{ri}\,\mathrm{dom}\, f_0^* \neq \emptyset,\end{gathered}$$

which agree with those in [35, Corollary 31.2.1].

The following gives a Karush-Kuhn-Tucker theorem for a fairly general class of variational inequalities.

Proposition 8.2. *Let f be a K-convex function from X to U, and let g be an extended-real-valued convex function on U, such that $(\mathcal{C})$ holds. For any single-valued mapping $S : X \to X^*$, consider the variational inequality*

$$\langle S(x), y - x\rangle \geq (g\circ f)(x) - (g\circ f)(y) \quad \forall y \in X, \tag{$\mathcal{VI}$}$$

and the inclusion

$$0 \in S(x) + D_K^* f(x)(\partial g(x)). \tag{$\mathcal{P}$}$$

Then every solution $\bar{x}$ of $(\mathcal{P})$ solves $(\mathcal{VI})$, and if $\operatorname{ri}\operatorname{rge}_K f \cap \operatorname{ri}\operatorname{dom} g \neq \emptyset$, *the converse holds.*

In particular, a sufficient condition for an $\bar{x}$ to solve the variational inequality

$$\langle S(x), y - x\rangle \geq 0 \quad \forall y \in C,$$

where $C = \{x \mid f(x) \in K\}$, is the existence of a $\bar{u}^*$ such that

$$\begin{gathered} 0 \in S(\bar{x}) + \partial\langle \bar{u}^*, f\rangle(\bar{x}), \\ f(\bar{x}) \in K, \quad \bar{u}^* \in K^*, \quad \langle \bar{u}^*, f(\bar{x})\rangle = 0. \end{gathered}$$

If there exists an $x \in \operatorname{ri}\operatorname{dom} f$ such that $f(x) \in \operatorname{ri} K$, this condition is necessary as well.

Proof. By the definition of a subgradient, $(\mathcal{VI})$ is equivalent to $-S(x) \in \partial(g\circ f)(x)$, where by Corollary 7.7(iv), $\partial(g\circ f)(x) \supset D_K^* f(x)(\partial g(x))$, with equality if $\operatorname{ri}\operatorname{rge}_K f \cap \operatorname{ri}\operatorname{dom} g \neq \emptyset$. The second variational inequality is $(\mathcal{VI})$ in the case $g = \delta_K$. The Karush-Kuhn-Tucker condition is obtained from $(\mathcal{P})$ by using Proposition 6.5, and the fact that $u^* \in \partial\delta_K(f(x)) = N_K(f(x))$ is equivalent to the given complementarity condition. We have used the alternative form of the constraint qualification in Theorem 7.6. □

The similarity of the above conditions with the KKT-conditions in optimization suggests that there is some kind of duality involved with variational inequalities. This is indeed the case; associated with $(\mathcal{P})$ there is a "dual inclusion" whose solutions are the vectors $\bar{u}^*$ satisfying the above conditions with some $\bar{x}$ [31].

K-convex functions give a natural framework for studying convex "vector minimization" problems too. In vector minimization it is not so clear what a "solution" should mean. The most classical concept is the so called *Pareto-efficiency.* The following is closely related to *proper efficiency* as defined by Borwein [5]. We will say that a vector $\bar{u} \in D$ is *K-minimal* in $D \subset U$, if $T_D(\bar{u}) \cap K$ is a subspace. If K is closed, this is equivalent to $\operatorname{ri} N_D(\bar{u}) \cap \operatorname{ri} -K^* \neq \emptyset$, by Lemma 3.2. For other solution concepts and further references on vector minimization, see Borwein [8] and the references there in.

Lemma 8.3. *Let K be closed and let $D \subset U$ be convex, such that $-K \subset \operatorname{rc} D$. Then $\bar{u}$ is K-minimal in D, if and only if $N_D(\bar{u}) \cap \operatorname{ri} -K^* \neq \emptyset$. If D is closed, then a K-minimizer exists if and only if $\operatorname{rc} D \cap K$ is a subspace.*

Proof. The recession condition implies $(\operatorname{rc} D)^* \subset -K^*$, so that since $\operatorname{rc} D \subset T_D(\bar{u})$, we have $N_D(\bar{u}) \subset -K^*$. By [35, Corollary 6.5.2], this means that K-minimality of $\bar{u}$ can be expressed as $N_D(\bar{u}) \cap \operatorname{ri} -K^* \neq \emptyset$. When D is closed, so is $\operatorname{rc} D$, so that by Lemma 3.2,

$\operatorname{rc} D \cap K$ is a subspace if and only if $\operatorname{ri}(\operatorname{rc} D)^* \cap \operatorname{ri} -K^* \neq \emptyset$, where by Corollary 14.2.1, Theorem 23.4 and Corollary 23.5.1 of [35]

$$\operatorname{ri}(\operatorname{rc} D)^* = \operatorname{ri}\operatorname{dom}\delta_D^* = \operatorname{ri}\operatorname{dom}\partial\delta_D^* = \operatorname{ri}\operatorname{rge}\partial\delta_D = \operatorname{ri}\operatorname{rge} N_D.$$

Thus, $\operatorname{rc} D \cap K$ is a subspace if and only if $\operatorname{ri}\operatorname{rge} N_D \cap \operatorname{ri} -K^* \neq \emptyset$. Since $\operatorname{rge} N_D \subset -K^*$, this is equivalent to $\operatorname{rge} N_D \cap \operatorname{ri} -K^* \neq \emptyset$. □

We will say that an $\bar{x} \in \operatorname{dom} f$ is a *K-minimizer* of a function f if $f(\bar{x})$ is K-minimal in $\operatorname{rge}_K f$. The second part of the following proposition generalizes the existence criterion [35, Theorem 27.1(b)] for minimization of extended-real-valued functions.

Proposition 8.4. *Let f be K-convex for a closed K. Then $\bar{x}$ is K-minimal for f if and only if $\bar{x}$ minimizes $\langle u^*, f\rangle$ for some $u^* \in \operatorname{ri} K^*$. If f is K-closed, and $\operatorname{rc}_K f$ is a subspace, then a K-minimizer exists.*

Proof. Since $\operatorname{rge}_K f - K = \operatorname{rge}_K f$, we have $-K \subset \operatorname{rc}\operatorname{rge}_K f$, so by the above lemma, $\bar{x}$ is K-minimal if and only if $N_{\operatorname{rge}_K f}(f(\bar{x})) \cap \operatorname{ri} -K^* \neq \emptyset$, where by Propositions 6.6 and 6.5

$$N_{\operatorname{rge}_K f}(f(\bar{x})) = -D_K^* f(\bar{x})^{-1}(0) = -\left\{u^* \mid 0 \in \partial\langle u^*, f\rangle(\bar{x})\right\}.$$

This proves the first claim. If $\operatorname{rc}_K f$ is a subspace, Proposition 6.3(iii) implies that $\operatorname{rge}_K f$ is closed, and $\operatorname{rc}\operatorname{rge}_K f = \operatorname{rge} R_K f$. So by Lemma 8.3, it suffices to show that $\operatorname{rge} R_K f \cap K$ is a subspace.

Assume $u \in K \cap \operatorname{rge} R_K f$, *i.e.* there is an x such that $u \in K \cap R_K f(x)$. Since $\operatorname{rc}_K f = (R_K f)^{-1}(0) = (R_K f)^{-1}(K)$, this implies that $x \in \operatorname{rc}_K f$. Since $\operatorname{rc}_K f$ is a subspace, we have $-x \in \operatorname{rc}_K f$, or $0 \in R_K f(-x)$. By the sublinearity of $R_K f$, $u + 0 \in R_K f(x - x)$, which by Proposition 6.3(i) means that $u \in -K$. Thus, $K \cap \operatorname{rge} R_K f \subset K \cap -K$, which by Proposition 6.3(i) must hold as an equality, so that $K \cap \operatorname{rge} R_K f$ is a subspace. □

Using Proposition 8.4 and the calculus rules of Section 6, one could derive conditions and existence criteria for K-minimality for more structured models of vector minimization.

9. The eigenvalue function on Hermitian matrices

Let $\mathcal{H}$ be the real vector space of $n \times n$ Hermitian matrices, and define an inner product on $\mathcal{H}$ by

$$\langle A, B\rangle = \operatorname{tr}(AB), \quad \forall A, B \in \mathcal{H}.$$

Since $\operatorname{tr}(AB) = \sum_{ij} A_{ij} B_{ij}$, it is clear that $\langle A, B\rangle = \langle B, A\rangle$, and

$$\langle A, CBC^*\rangle = \operatorname{tr}[(ACB)C^*] = \operatorname{tr}[C^*(ACB)] = \langle C^*AC, B\rangle,$$

for any $C \in \mathbb{C}^{n\times n}$ and its adjoint $C^* = \bar{C}^T$. Recall that any $A \in \mathcal{H}$ can be expressed as $A = U \operatorname{diag}(\lambda(A))U^*$, where $\lambda(A) \in \mathbb{R}^n$ is the vector of eigenvalues of A in nonincreasing order, and U is unitary: $U^*U = UU^* = I$ [17]. The set of unitary matrices will be denoted by $\mathcal{U}$.

Our aim is to analyze the function $\lambda : \mathcal{H} \to \mathbb{R}^n$ in the framework of Sections 6 and 7. This is based on the following; see [21] and the references there in.

Lemma 9.1 (Fan-Theobald). *For any $A, H \in \mathcal{H}$,*

$$\langle A, H \rangle \leq \langle \lambda(A), \lambda(H) \rangle \,,$$

with equality if and only if $U^*AU = \operatorname{diag}(\lambda(A))$, and $U^*HU = \operatorname{diag}(\lambda(H))$ for some $U \in \mathcal{U}$.

Define the closed convex cone

$$K = \left\{ u \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{k} u_i \leq 0, \; k = 1, \ldots, n-1, \; \sum_{i=1}^{n} u_i = 0 \right\},$$

whose polar cone is

$$K^* = \{ u^* \in \mathbb{R}^n \mid u_1^* \geq \cdots \geq u_n^* \} \,.$$

The following scalarization formula for λ has been observed at least in [11, 23]. The *convex hull* of a set C will be denoted by $\operatorname{co} C$.

Proposition 9.2. *For any $u^* \in K^*$ and $A \in \mathcal{H}$,*

$$\langle u^*, \lambda(A) \rangle = \sup_{H \in \mathcal{H}} \{ \langle A, H \rangle \mid \lambda(H) = u^* \} \,,$$

where

$$\operatorname*{argmax}_{H \in \mathcal{H}} \{ \langle A, H \rangle \mid \lambda(H) = u^* \} = \{ U \operatorname{diag}(u^*) U^* \mid U \in \mathcal{U} : U^*AU = \operatorname{diag}(\lambda(A)) \} \,.$$

Thus, λ is K-convex and K-closed, and for any $u^* \in K^*$

$$D_K^* \lambda(A)(u^*) = \operatorname{co} \{ U \operatorname{diag}(u^*) U^* \mid U \in \mathcal{U} : A = U \operatorname{diag}(\lambda(A)) U^* \} \,.$$

Proof. Let $H \in \mathcal{H}$ be such that $\lambda(H) = u^*$. Then by Lemma 9.1 equality in $\langle A, H \rangle \leq \langle \lambda(A), u^* \rangle$ holds if and only if $U^*AU = \operatorname{diag}(\lambda(A))$ and $U^*HU = \operatorname{diag}(\lambda(H))$ for some $U \in \mathcal{U}$. Solving for H, we obtain the attainment criterion. For any $u^* \in K^*$, $\langle u^*, \lambda \rangle$ is convex as a pointwise supremum of linear functions. Since K is closed, Lemmas 6.1 and 6.2 imply the K-convexity and K-closedness, respectively, of λ. The formula for $D_K^* \lambda(A)$ follows from Proposition 6.5 and [35, Corollary 23.5.3]. $\square$

Choosing $u_i^* = 1$ for $i = 1, \ldots, m$ and zero otherwise, we see that the sum of m largest eigenvalues A is a convex function of A [16, 30].

Corollary 9.3. *Let $u^* \in K^*$, and define $L(u^*) = \{ H \in \mathcal{H} \mid \lambda(H) - u^* \in K \}$. Then*

$$\operatorname{co} \{ H \in \mathcal{H} \mid \lambda(H) = u^* \} = L(u^*).$$

Indeed, we have the expression $\langle u^*, \lambda \rangle = \sigma_{L(u^*)}$, and the function λ is sublinear in the sense that

$$\lambda(A + B) - \lambda(A) - \lambda(B) \in K \quad \forall A, B \in \mathcal{H}.$$

Thus, the K-epigraph of λ is a convex cone, so that $R_K \lambda = S_{\lambda, K}$.

Proof. Because $\lambda(A) \in K^*$ for any $A \in \mathcal{H}$, we have $\langle v^* - u^*, \lambda(A)\rangle \leq 0$ whenever $v^* - u^* \in K$. Thus, for any $u^* \in K^*$

$$\begin{aligned}\langle u^*, \lambda\rangle (A) &= \sup\{\langle v^*, \lambda\rangle (A) \mid v^* - u^* \in K\}\\ &= \sup\{\langle A, H\rangle \mid \lambda(H) = v^*,\ v^* - u^* \in K\}\\ &= \sup\{\langle A, H\rangle \mid \lambda(H) - u^* \in K\} = \sigma_{L(u^*)}(A).\end{aligned}$$

This implies that $\operatorname{cl}\operatorname{co}\{H \in \mathcal{H} \mid \lambda(H) = u^*\} = \operatorname{cl}\operatorname{co} L(u^*)$, where the closures are superfluous by continuity of λ and the boundedness of the sets. Since λ is K-convex, the set $L(u^*) = S^{-1}_{\lambda,K}(u)$ is convex. By the sublinearity of the functions $\sigma_{L(u^*)}$

$$\langle u^*, \lambda\rangle (A + B) - \langle u^*, \lambda\rangle (A) - \langle u^*, \lambda\rangle (B) \leq 0, \quad \forall A, B \in \mathcal{H},\ \forall u^* \in K^*,$$

which by closedness of K is equivalent to the given inclusion. □

The above sublinearity property was recognized in [12, Lemma 2.1]. Choosing $u_i^* = 1$ for $i = 1, \dots, m$ and zero otherwise, and using the definition of K, we obtain

$$\operatorname{co}\left\{UU^* \;\middle|\; U \in \mathbb{C}^{n\times m} : U^*U = I\right\} = \{H \in \mathcal{H} \mid \lambda(H) \in [0,1]^n,\ \operatorname{tr} H = m\},$$

which is a complex version of [16, Proposition 2.1].

We will next study the condition $(\mathcal{C})$ in the case $f = \lambda$. Since $\operatorname{diag}(u) \in \mathcal{H}$ for any $u \in \mathbb{R}^n$, we see that $\operatorname{rge}\lambda = K^*$. The sets of *permutation matrices* and *doubly stochastic matrices* will be denoted by $\mathcal{P}$ and $\mathcal{DS}$, respectively. That is, $P \in \mathcal{P}$ means that P has exactly one 1 in each row and each column, and all other entries are 0's, whereas $P \in \mathcal{DS}$ means that P has nonnegative entries and all the row and column sums equal 1.

The simplest example of a nonsmooth K-convex function is given by the following vector-valued generalization of the max-function. Define the function $\operatorname{dec} : \mathbb{R}^n \to \mathbb{R}^n$, by

$$\operatorname{dec}(u)_i = \text{the } i\text{th largest component of } u.$$

It is easy to check that for any $u \in \mathbb{R}^n$ and $u^* \in K^*$, $\langle u^*, u\rangle \leq \langle u^*, \operatorname{dec}(u)\rangle$, so that

$$\langle u^*, \operatorname{dec}\rangle (u) = \sup_{P\in\mathcal{P}} \langle u^*, Pu\rangle \tag{9.1}$$

which is a pointwise supremum of linear functions, and hence convex. Thus, dec is K-convex by Lemma 6.1. Note that dec may also be viewed as a restriction of λ to the diagonal matrices [24]. Proofs of the following can be found for example in [25].

Lemma 9.4.
(i) (Birkhoff) $\mathcal{DS} = \operatorname{co}\mathcal{P}$
(ii) (Hardy-Littlewood-Polya) $\operatorname{dec}(u_2) - \operatorname{dec}(u_1) \in K$ *if and only if* $u_2 = Pu_1$ *for some* $P \in \mathcal{DS}$.

A function g will be called *permutation invariant* if $g(Pu) = g(u)$ for all $u \in \operatorname{dom} g$ and $P \in \mathcal{P}$, i.e. if the value of g is not changed if we reorder its argument vector. The following is an obvious vector-valued generalization of [25, Proposition C.2].

Lemma 9.5. *If g is L-convex and permutation invariant, then it satisfies $(\mathcal{C})$ with $f = \lambda$.*

Proof. Assume $u_1 \in \operatorname{dom} g$, $u_2 \in \operatorname{rge} f = K^*$, $u_2 - u_1 \in K$. It follows from the expression (9.1) that $\langle u^*, \operatorname{dec}(u_1)\rangle \geq \langle u^*, u_1\rangle$ for any $u^* \in K^*$, so that $u_1 - \operatorname{dec}(u_1) \in K$. Together with $u_2 - u_1 \in K$ this implies $u_2 - \operatorname{dec}(u_1) \in K$. Since $u_2 \in K^*$ implies $u_2 = \operatorname{dec}(u_2)$, Lemma 9.4 implies that u_2 can be written as a convex combination $u_2 = \sum_i \alpha_i P_i u_1$ for some $P_i \in \mathcal{P}$. Thus, by permutation invariance and L-convexity of g

$$g(u_2) - g(u_1) = g\left(\sum_i \alpha_i P_i u_1\right) - \sum_i \alpha_i g(P_i u_1) \in L.$$

□

The following facts about spectrally defined matrix functions now follow easily. The *horizon* (or recession) *function* f^∞ of a closed convex function f is defined by

$$\operatorname{epi} f^\infty = \operatorname{rc} \operatorname{epi} f.$$

Proposition 9.6. *Let g be extended-real-valued, convex and permutation invariant. Then*

(i) *$g\circ\lambda$ is convex,*

(ii) *The subdifferential of $g\circ\lambda$ satisfies*

$$\partial(g\circ\lambda)(A) \supset \bigcup \operatorname{co}\{U \operatorname{diag}(u^*)U^* \mid U \in \mathcal{U} : U^*AU = \operatorname{diag}(\lambda(A)),\ u^* \in \partial g(\lambda(A))\},$$

with equality if $g(u) < \infty$ for some $u \in \operatorname{int} K^*$.

(iii) *If g is closed and $(-K)\cap\operatorname{rc}_K g$ is a subspace, then $(g\circ\lambda)$ is closed and $(g\circ\lambda)^\infty = g^\infty\circ\lambda$.*

Proof. Part (i) is obtained by combining Lemmas 7.1, 9.5 and Proposition 9.2, and (ii) follows from Corollary 7.7(iv) and Proposition 9.2 by noting that the interior of K^* is nonempty.

Since λ is K-closed by Proposition 9.2, the conditions in part (iii) imply by Theorem 7.3 that $g\circ\lambda$ is closed, and

$$R_{\mathbb{R}_-}(g\circ\lambda) = R_{\mathbb{R}_-} g\circ R_K\lambda.$$

Using the definition of the horizon function, and the formula $R_K\lambda = S_{\lambda,K}$ in Corollary 9.3, this can be written as

$$S_{(g\circ\lambda)^\infty,\mathbb{R}_-} = S_{g^\infty,\mathbb{R}_-}\circ S_{\lambda,K}.$$

Permutation invariance of g implies that of g^∞, so that by Lemmas 9.5 and 7.1,

$$S_{g^\infty,\mathbb{R}_-}\circ S_{\lambda,K} = S_{g^\infty\circ\lambda,\mathbb{R}_-}.$$

Thus, $S_{(g\circ\lambda)^\infty,\mathbb{R}_-} = S_{g^\infty\circ\lambda,\mathbb{R}_-}$, which means that $(g\circ\lambda)^\infty = g^\infty\circ\lambda$. □

Part (i) was obtained in [10] for real-valued functions and in [23] for extended-real-valued functions. It can be shown that the conclusions of (ii) and (iii) remain valid even without the qualification conditions; see [23, Theorem 3.1] and [38, Theorem 8.1], respectively. This is based on a different approach that makes direct use of the special structure of the problem. Somewhat similarly, using polyhedral refinements of the basic results in [35], one can show that the qualification conditions in our main results can be relaxed if the mappings are polyhedral.

In minimization algorithms, it is often sufficient to find a single subgradient of a function at a given point [14]. Proposition 9.6(ii) gives the following simple procedure: form a spectral decomposition $A = U\operatorname{diag}(\lambda(A))U^*$ of A (There are many efficient algorithms for doing this [13]). Find a subgradient u^* of g at $\lambda(A)$. Then $U\operatorname{diag}(u^*)U^* \in \partial(g\circ\lambda)(A)$.

In [24], Lewis introduced the concept of a *normal decomposition system.* This provides a unified framework which includes the above framework of Hermitian matrices as a special case. Another example is the setting of $m\times n$ complex matrices where λ is replaced by the singular-value function [22], [24, Section 7]. In general, a normal decomposition system involves a function γ which generalizes λ. By [24, Theorem 2.4] γ is K-convex, where K is the polar cone of the range of γ. The above analysis generalizes immediately to this setting, in which the horizon function formula in Corollary 9.6(iii) would seem to be new.

In problems like semidefinite programming and eigenvalue optimization, one is interested in matrix-valued functions and the behavior of the associated eigenvalues. Except for [39], semidefinite programming has mainly concentrated on affine matrix functions. We let $P \subset \mathcal{H}$ denote the convex cone of positive semidefinite Hermitian matrices.

Lemma 9.7. *Let g be an extended-real-valued convex and permutation invariant function on $\mathbb{R}^n$, and let A be a P-convex function from X to $\mathcal{H}$. If g is coordinate-wise nonincreasing, then the function $g\circ\lambda\circ A$ is convex, and*

$$\partial(g\circ\lambda\circ A)(x) \supset \bigcup \{\partial \langle H, A\rangle (x) \mid H \in \partial(g\circ\lambda)(A(x))\},$$

with equality if there exists an $x \in \operatorname{ri}\operatorname{dom} A$ such that $A(x) \in \operatorname{ri}\operatorname{dom}(g\circ\lambda)$.

Proof. If g is coordinate-wise nonincreasing, we have $P \subset \operatorname{rc}_{\mathbb{R}_-}(g\circ\lambda)$, which implies the convexity of $g\circ\lambda\circ G$. The rest is just an application of the chain rule. $\square$

The above condition for convexity of $g\circ\lambda\circ A$ generalizes [39, Proposition 3] which addressed the case where A is differentiable and g is the standard barrier function $g(u) = -\sum_{i=1}^n \ln u_i$ for $\mathbb{R}^n_+$. Note that in computing the matrices $H \in \partial(g\circ\lambda)(A(x))$ in the above formula, one may use Proposition 9.6(ii).

The P-convex matrix functions have been studied by many authors; for reference see [25, 18]. In general, A could be a function of a matrix, or it might have been constructed from such functions. For example, the functions $A_1(H) = H^{-1}$ and $A_2(H) = -H^{1/2}$ are P-convex functions from $\mathcal{H}$ to $\mathcal{H}$, with $\operatorname{dom} A_1 = \operatorname{dom} A_2 = \operatorname{int} P$, and $\operatorname{rc}_P A_1 = \operatorname{rc}_P A_2 = P$. Also, if $L \in \mathcal{H}_m$ ($m\times m$ Hermitian matrices) is positive semidefinite, then the function $A(C) = C^*LC$ from $\mathbb{C}^{m\times n}$ to $\mathcal{H}$ is P-convex. Any of these functions could be combined according to Lemma 7.1 or Corollary 7.2 to obtain other P-convex functions. For example, if $C : X \to \mathbb{C}^{m\times n}$ is affine, then

$$A_1(x) = C(x)^*LC(x), \quad A_2(x) = [C(x)^*LC(x)]^{-1}, \quad A_3(x) = -[C(x)^*LC(x)]^{1/2}$$

are a P-convex functions from X to $\mathcal{H}$. Also, Corollary 7.2(ii) implies that for any P-convex function A from X to $\mathcal{H}$ and a $B \in \mathbb{C}^{n\times m}$, the function

$$C(x) = B^*A(x)B$$

is (B^*PB)-convex from X to $\mathcal{H}_m$. Note that B^*PB is contained in the cone of positive semidefinite matrices in $\mathcal{H}_m$, and it is equal to it if the rank of B is at least m. Nonsmoothness enters this framework naturally when $A(x) = \operatorname{diag}[a_1(x), \ldots, a_n(x)]$, with a_i convex and nonsmooth.

References

[1] J.-P. Aubin: Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions, Advances in Mathematics, Supplementary studies, Ed. L. Nachbin (1981) 160–232.

[2] J.-P. Aubin, I. Ekeland: Applied Nonlinear Analysis, John Wiley & Sons, 1984.

[3] J.-P. Aubin, H. Frankowska: Set-Valued Analysis, Birkhäuser, 1990.

[4] J. M. Borwein: Multivalued convexity and optimization: A unified approach to inequality and equality constraints, Math. Programming 13 (1977) 183–199.

[5] J. M. Borwein: Proper efficient points for maximizations with respect to cones, SIAM J. Control Optim. 15 (1977) 57–63.

[6] J. M. Borwein: Continuity and differentiability properties of convex operators, Proc. London Math. Soc. 44(3) (1982) 420–444.

[7] J. M. Borwein: Automatic continuity and openness of convex relations, Proc. Amer. Math. Soc. 99(1) (1987) 49–55.

[8] J. M. Borwein: Super efficiency in vector optimization, Trans. Amer. Math. Soc. 338(1) (1993) 105–122.

[9] B. D. Craven: Control and Optimization, Chapman and Hall, 1995.

[10] C. Davis: All convex invariant functions of Hermitian matrices, Archiv der Mathematik 8 (1957) 276–278.

[11] S. Friedland: Inverse eigenvalue problems, Linear Algebra Appl. 17 (1977) 15–51.

[12] S. Friedland: Convex spectral functions, Linear and Multilinear Algebra 9 (1981) 299–316.

[13] G. H. Golub, C. F. Van Loan: Matrix Computations, John Hopkins Univ. Press, Baltimore, 1983.

[14] J.-B. Hiriart-Urruty, C. Lemarechal: Convex Analysis and Minimization Algorithms, Springer-Verlag, 1993.

[15] J.-B. Hiriart-Urruty, M. Moussaoui, A. Seeger, M. Volle: Subdifferential calculus without qualification conditions, using approximate subdifferentials: a survey, Nonlinear Analysis 24(12) (1995) 1727–1754.

[16] J.-B. Hiriart-Urruty, D. Ye: Sensitivity analysis of all eigenvalues of a symmetric matrix, Numer. Math. 70 (1995) 45–72.

[17] R. A. Horn, C. R. Johnson: Matrix Analysis, Cambridge University Press, Cambridge, 1985.

[18] R. A. Horn, C. R. Johnson: Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[19] S. S. Kutateladze: Convex operators, Russian Math. Surveys 34(1) (1979) 181–214.

[20] V. L. Levin: Subdifferentials of convex mappings and composite functions, Siberian Math. J. 13 (1972) 903–909.

[21] A. S. Lewis: Von Neumann's lemma and a Chevalley-type theorem for convex functions on Cartan subspaces, Trans. Amer. Math. Soc., submitted.

[22] A. S. Lewis: The convex analysis of unitarily invariant matrix functions, J. Convex Analysis 2(1) (1995) 173–183.

[23] A. S. Lewis: Convex analysis on the Hermitian matrices, SIAM J. Optim. 6 (1996) 164–177.

[24] A. S. Lewis: Group invariance and convex matrix analysis, SIAM J. Matrix Anal. Appl. 17(4) (1996) 927–949.

[25] A. W. Marshall, I. Olkin: Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.

[26] B. S. Mordukhovich: Metric approximations and necessary conditions for general classes of nonsmooth extremal problems, Soviet Math. Dokl. 22(2) (1980) 526–530.

[27] B. S. Mordukhovich: Generalized differential calculus for nonsmooth and set-valued mappings, J. Math. Anal. Appl. 183 (1994) 250–288.

[28] B. S. Mordukhovich, Y. Shao: Nonconvex differential calculus for infinite-dimensional multifunctions, Set-Valued Anal. 4(3) (1996) 205–236.

[29] M. L. Overton: Large-scale optimization of eigenvalues, SIAM J. Optim. 2(1) (1992) 88–120.

[30] K. L. Overton, R. S. Womersley: Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrix, Math. Programming 62 (1993) 321–357.

[31] T. Pennanen: Dualization of generalized equations of maximal monotone type, to appear in SIAM J. Optim.

[32] T. Pennanen, J. Eckstein: Generalized Jacobians of vector-valued convex functions, Rutcor Research Reports (1997) 6–97.

[33] S. M. Robinson: Regularity and stability for convex multivalued functions, Math. Oper. Res. 1(2) (1976) 130–143.

[34] R. T. Rockafellar: A Monotone convex analogue of linear algebra, In: W. Fenchel (ed.): Proceedings of the Colloquium on Convexity, Copenhagen (1965) 261–276.

[35] R. T. Rockafellar: Convex Analysis, Princeton University Press, 1970.

[36] R. T. Rockafellar: Conjugate Duality and Optimization, SIAM, 1974.

[37] R. T. Rockafellar, R. J-B Wets: Variational Analysis, Springer-Verlag, 1998.

[38] A. Seeger: Convex analysis of spectrally defined matrix functions, SIAM J. Optim. 7(3) (1997) 679–696.

[39] A. Shapiro: First and second order analysis of nonlinear semidefinite programs, Math. Programming 77 (1997) 301–320.

[40] A. Shapiro: On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints, SIAM J. Optim. 7(2) (1997) 508–518.

[41] M. Thera: Subdifferential calculus of convex operators, J. Math. Anal. Appl. 80(1) (1981) 78–91.

[42] L. I. Trudzik: Continuity properties of vector-valued convex functions, J. Austral. Math. Soc. (Series A) 36(3) (1984) 404–415.

[43] L. Vandenberghe, S. Boyd: Semidefinite programming, SIAM Review 38(1) (1996) 49–95.