GENERALIZED SECOND DERIVATIVES
OF CONVEX FUNCTIONS AND SADDLE FUNCTIONS

R. T. ROCKAFELLAR

ABSTRACT. The theory of second-order epi-derivatives of extended-real-valued functions is applied to convex functions on $\mathbb{R}^n$ and shown to be closely tied to proto-differentiation of the corresponding subgradient multifunctions, as well as to second-order epi-differentiation of conjugate functions. An extension is then made to saddle functions, which by definition are convex in one argument and concave in another. For this case a concept of epi-hypo-differentiability is introduced. The saddle function results provide a foundation for the sensitivity analysis of primal and dual optimal solutions to general finite-dimensional problems in convex optimization, since such solutions are characterized as saddle-points of a convex-concave Lagrangian function, or equivalently as subgradients of the saddle function conjugate to the Lagrangian.

1. INTRODUCTION

Generalized first derivatives of extended-real-valued functions have long been studied in nonsmooth analysis and from many points of view. They are essential in understanding the characterization of solutions to problems of optimization and also the possible rates of change that may be exhibited in the optimal values of such problems relative to perturbations. Generalized second derivatives are newer on the scene, and until recently little was known about them despite the potential for important applications. These applications are concerned with the sensitivity analysis of optimal solutions (minimizing or maximizing vectors) rather than merely optimal values.

Convexity theory has provided the springboard for much of nonsmooth analysis, and not surprisingly this has been true as well for attempts at defining generalized second derivatives. One line of development was started in 1980 by Lemarechal and Nurminski [1], who defined approximate second derivatives of convex functions in terms of the ε-subdifferential of convex analysis in [2, §23]. This line has been followed by Auslender [3] and explored in detail by Hiriart-Urruty [5, 6, 7, 8], Hiriart-Urruty and Seeger [9], Hiriart-Urruty, Strodiot, and

Hien Nguyen [10]. It has led to many interesting results, but unfortunately these seem limited in basic concept to the convex case and not readily extendable beyond.

Another approach was initiated by Aubin in defining the second derivatives of a convex function f in terms of first derivatives of the subgradient multifunction ∂f, the latter being expressed by the multifunction whose graph is, at a certain point, the tangent cone to the graph of ∂f in the sense of Clarke [11, 12]. This approach, which was applied by Aubin [13] to the sensitivity analysis of optimal solutions to dual pairs of problems in convex programming, was carried forward in the book of Aubin and Ekeland [14], which develops as well still other ideas for the generalized differentiation of multifunctions, i.e., set-valued mappings.

Aubin's theory of second derivatives of convex functions in [13] has turned out to be much less "general" than was first imagined. It was shown in Rockafellar [15] that the Clarke tangent cone to the graph of ∂f is always a *subspace* in the case of a convex function f. The nondegeneracy assumption employed by Aubin in his applications to sensitivity analysis actually requires that this subspace, in a context we shall not try to explain here, be the graph of a linear transformation. This means that f must be in fact twice differentiable in the classical sense of having a second-order Taylor expansion at the point in question.

The ideas connected with such expansions were explored in detail in [15]. This paper introduces "generalized" quadratic functions, which may be extended-real-valued, as giving second derivatives of convex functions at certain points of nonsmoothness. It uses the subdifferential convergence theorem of Attouch [16; 17, Theorem 3.66], to demonstrate the relationship between such derivatives and the graphical derivatives of Aubin [13]. It further provides an alternative modern proof of the theorem of Alexandroff [18], according to which a finite convex function on an open convex set has a second-order Taylor expansion at almost every point. (See Dudley [19] for a proof based instead on the theory of distributions.) An infinite-dimensional version of this approach to second derivatives of convex functions in terms of generalized quadratic functions and Clarke tangent cones to the graph of ∂f has been provided by Ndoutoume [20].

Second derivatives of nonconvex functions have been developed in quite a different way by Auslender [21] and Chaney [22–27], using special upper and lower limits of certain second-order difference quotients. This theory deals with functions that are locally lipschitzian, thus in particular finite valued. Nevertheless it is applicable by means of penalty representations of constraints to the derivation of necessary and sufficient conditions for optimality in nonlinear programming, as shown in the papers in question.

Another approach to generalized second derivatives has been followed in a series of papers of Ben-Tal and Zowe [28–30]. These authors, like Chaney, focus on finite-valued functions and rely on penalty representations to apply

their results to optimality conditions. Their derivatives are defined in terms of second-order limits taken along parabolic curves. Applications to marginal functions have been given by Seeger [31] and Shapiro [32]. Still other ideas of second derivatives that are of interest in optimization have been explored by Shapiro [33, 34], and by Cominetti and Correa [35].

Recently a theory of second derivatives which does not depend either on convexity or on finiteness of the functions involved has been set forth in our paper [36]. The definition in this case rests on epi-convergence of second-order difference quotients. The second derivatives so obtained are therefore called epi-derivatives. They are dual in a certain sense to parabolic derivatives that generalize the ones of Ben-Tal and Zowe mentioned above. The chief fact established in [36] is that for a very broad class of extended-real-valued functions, which includes most of the functions commonly encountered in nonlinear programming, such second-order epi-derivatives do exist and are given by usable formulas. We have worked from this fact in a follow-up paper [37] to show that optimality conditions can be developed in a "neo-classical" setting which provides new insights into matters of convergence. An extension of the class of functions covered in [36] has been achieved recently by Cominetti [48].

Our intention here is to return with these ideas of epi-derivatives to the context of convex analysis. One reason for giving special attention to *convex* functions is that, at present, there are no good tools that relate epi-convergence of *nonconvex* functions to graphical convergence of their subdifferential multifunctions as does Attouch's theorem in the convex case. We wish to use Attouch's theorem and also the continuity of the Legendre-Fenchel transform for convex functions to explore various connections of epi-differentiation with graphical differentiation and duality. At the same time we aim at providing additional examples of second-order epi-differentiation, beyond the ones in [36].

Our agenda does not end with convex functions, however, but passes from such a framework to the more difficult study of second derivatives of saddle functions (convex-concave functions). Such functions are fundamental in convex optimization because of their role as Lagrangians in the statement of saddle point conditions for optimality. The sensitivity analysis of saddle points, and therefore of solutions to primal and dual problems of optimization, is intimately tied with the understanding of generalized first derivatives of the subgradient multifunctions associated with Lagrangians. Such derivatives can in turn be expected to correspond to second derivatives of the Lagrangians themselves.

Although we are unable within the limits of this paper to elaborate on the details of such connections with sensitivity analysis, we do lay a full foundation for applications in that direction. Results of Attouch, Azé, and Wets [38] on epi-hypo-convergence of saddle functions are utilized in defining epi-hypo-derivatives of such functions. This phase of theory may at first appear somewhat formidable to the reader. This is because of the unavoidable need, already well recognized in convex analysis, of treating saddle functions as elements of equivalence classes rather than just as individuals.

The importance of the potential applications justifies the effort of studying saddle functions, however, and in any case the end results are relatively simple in outline. They completely mirror the ones for purely convex functions and provide a rich array of additional examples of generalized second-order differentiation.

We confine our treatment to the finite-dimensional case so as not to obscure the basic ideas with technical and notational complications over a multiplicity of topologies.

2. EPI-DERIVATIVES OF CONVEX FUNCTIONS

First- and second-order epi-derivatives of extended-real-valued functions, as defined in Rockafellar [36], require the concept of epi-convergence. Consider a family of functions $\varphi_t \colon \mathbb{R}^n \to \overline{\mathbb{R}}$, indexed by a real parameter $t > 0$. ($\overline{\mathbb{R}}$ denotes the extended real numbers.) Such a family is said to *epi-converge* to a function $\varphi \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ if their epigraphs, as subsets of $\mathbb{R}^{n+1}$, converge to the epigraph of φ in the Painlevé-Kuratowski sense.

Such set convergence can be characterized in many ways, which we will not attempt to review (see Salinetti and Wets [39], for instance.) Epi-convergence of φ_t to φ can itself be shown to be equivalent to the relation

$$(2.1) \qquad \liminf_{t \downarrow 0} \inf_{\xi' \to \xi} \varphi_t(\xi') = \varphi(\xi) = \limsup_{t \downarrow 0} \inf_{\xi \to \xi} \varphi_t(\xi'),$$

with mixed limits defined as in Rockafellar [40], namely

$$(2.2) \qquad \liminf_{t \downarrow 0} \inf_{\xi' \to \xi} = \lim_{\varepsilon \downarrow 0} \left[\liminf_{t \downarrow 0} \left[\inf_{|\xi' - \xi| \leq \varepsilon} \varphi_t(\xi') \right] \right],$$

$$(2.3) \qquad \limsup_{t \downarrow 0} \inf_{\xi' \to \xi} = \lim_{\varepsilon \downarrow 0} \left[\limsup_{t \downarrow 0} \left[\inf_{|\xi' - \xi| \leq \varepsilon} \varphi_t(\xi') \right] \right].$$

(We denote by $|\cdot|$ the euclidean norm.)

Epi-convergence was first introduced by Wijsman [41, 42], although not under that name. He looked at it only in the case of *convex* functions on $\mathbb{R}^n$ but showed that it was preserved under the Legendre-Fenchel transform. The notation was generalized appropriately to Banach spaces by Mosco [43] and it since has been developed extensively by many authors; here we defer to the recent text of Attouch [17].

Consider now a function $f \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and a point x, where $f(x)$ is finite. As in [36], we say that f is *epi-differentiable* at x if the first-order difference quotient functions

$$(2.4) \qquad \varphi_t(\xi) = [f(x + t\xi) - f(x)]/t \quad \text{for } t > 0$$

epi-converge as $t \downarrow 0$ to some function φ having $\varphi(0) \neq -\infty$. Then in place of $\varphi(\xi)$ we write $f'_x(\xi)$ and refer to f'_x as the (*first-order*) *epi-derivative function* for f at x.

For the corresponding second-order concept, we suppose f is epi-differentiable at x as above, and we fix also an auxiliary vector $v \in \mathbb{R}^n$. Then we look at the second-order difference quotient functions

$$(2.5) \qquad \varphi_t(\xi) = [f(x + t\xi) - f(x) - tv \cdot \xi]/\tfrac{1}{2}t^2 \quad \text{for } t > 0.$$

If these epi-converge as $t \downarrow 0$ to some function φ having $\varphi(0) \neq -\infty$, we say that f is *twice epi-differentiable* at x relative to v, and we speak of v as being an *epi-gradient* of f at x. Then we write $\varphi(\xi)$ as $f''_{x,v}(\xi)$ and call this the *second-order epi-derivative* function for f at x relative to v.

Some readers might not find the factor $\tfrac{1}{2}$ in (2.5) to their liking, because of the way it pops up in certain formulas that will be encountered later. We include it, however, in order to have second-order epi-derivatives agree with ordinary second derivatives in cases where both exist, as with classical functions of a single variable.

The basic properties of first- and second-order epi-derivatives of general functions have been laid out in [36] along with many examples. We refer the reader to that paper and proceed here directly to the special properties in the convex case.

Besides such epi-derivatives of functions, we shall work with proto-derivatives of multifunctions, as developed in Rockafellar [44]. For a multifunction (set-valued mapping) $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, this notion is defined in terms of a point x and an image $v \in F(x)$. Consider the difference quotient multifunctions

$$(2.6) \qquad \Phi_t(\xi) = [F(x + t\xi) - v]/t \quad \text{for } t > 0.$$

If the graph of Φ_t as a subset of $\mathbb{R}^n \times \mathbb{R}^m$ converges as $t \downarrow 0$ to another subset of $\mathbb{R}^n \times \mathbb{R}^m$ in the Painlevé-Kuratowski sense, we say F is *proto-differentiable* at x relative to v. The limit set is interpreted as the graph of another multifunction $F'_x: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and this is called the proto-derivative of F at x relative to v. For more background on this concept and its relationship to other forms of differentiation, the reader should turn to [44].

In the rest of this section, f will be a *closed proper convex* function on $\mathbb{R}^n$, and x will be a point of $\operatorname{dom} f$, the set of points where $f(x) < \infty$. As usual, the set of subgradients v of f at x will be denoted by $\partial f(x)$. It is to the multifunction $\partial f: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ that the notion of proto-differentiation will presently be applied.

Theorem 2.1. *For f to be (once) epi-differentiable at x, it is necessary and sufficient that $\partial f(x) \neq \varnothing$. Then f'_x is the support function of $\partial f(x)$,*

$$(2.7) \qquad f'_x(\xi) = \sup\{\xi \cdot v \mid v \in \partial f(x)\} \quad \text{for all } \xi,$$

and f'_x is therefore the closure of the ordinary directional derivative function $f'(x; \cdot)$, defined by

$$(2.8) \qquad f'(x; \xi) = \lim_{t \downarrow 0}[f(x + t\xi) - f(x)]/t.$$

In particular, f'_x is a closed proper convex function which is positively homogeneous of degree one and has $f'_x(0) = 0$.

Proof. The lower semicontinuous hull of the convex function $f'(x; \cdot)$ (which is the same as its closure in convex analysis, unless the value $-\infty$ is taken somewhere) is the function whose epigraph is the (closed) tangent cone to the convex set epi f at the point $(x, f(x))$, as defined in convex analysis; see [2, §23]. This cone is known to coincide with both the contingent cone to epi f at $(x, f(x))$ and the Clarke tangent cone to epi f at $(x, f(x))$ by convexity; see Clarke [11, 12]. In general, however, as explained in [44], the epi-differentiability of f at x is equivalent to the coincidence of the contingent cone in question and the so-called intermediate cone to epi f at $(x, f(x))$ (which lies between the contingent cone and the Clarke tangent cone), along with the property that the common cone does not include the "downward" pointing ray from $(0, 0)$ (i.e., the value $-\infty$ is not taken at 0). This cone is then the epigraph of f'_x .

Thus f is epi-differentiable at x if and only if the lower semicontinuous hull of $f'(x; \cdot)$ does not have the value $-\infty$ at 0 (which by convexity is true if and only if this function nowhere takes on $-\infty$). In this case the lower semicontinuous hull of $f'(x : \cdot)$ is the same function as f'_x , and it is also the same as the closure of the $f'(x; \cdot)$. From convex analysis it is known that the lower semicontinuous hull of $f'(x; \cdot)$ fails to have the value $-\infty$ at 0 if and only if $\partial f(x) \neq \varnothing$, in which case it is the support function of $\partial f(x)$. These facts justify all the assertions of the theorem. □

Theorem 2.2. *For f to be twice epi-differentiable at x relative to a vector v , it is necessary and sufficient that $v \in \partial f(x)$ and that the multifunction ∂f be proto-differentiable at x relative to v . Then the second-order epi-derivative function $f''_{x,v}$ is a closed proper convex function which is positively homogeneous of degree two and satisfies*

$$(2.9) \qquad f''_{x,v}(\xi) \geq 0 \quad \text{for all } \xi, \quad \text{with } f''_{x,v}(0) = 0.$$

Moreover the subdifferential of $\frac{1}{2}f''_{x,v}$ is the proto-derivative of ∂f at x relative to v :

$$(2.10) \qquad \partial(\tfrac{1}{2}f''_{x,v})(\xi) = (\partial f)'_{x,v}(\xi) \quad \text{for all } \xi.$$

Proof. The necessity of having $v \in \partial f(x)$ follows from [36, Proposition 2.8] in conjunction with Theorem 2.1. The fact that $f''_{x,v}$ is closed, proper, and positively homogeneous of degree two is covered by [36, Proposition 2.7]. The nonnegativity in (2.9) follows from the nonnegativity of the difference quotients φ_t in (2.5) when $v \in \partial f(x)$. These are, of course, closed proper convex functions of ξ having $\varphi_t(0) = 0$. Their limit $f''_{x,v}$, when it exists, is necessarily convex because convexity is preserved under epi-convergence. (A limit of convex sets is convex.)

The rest of our argument is based on the theorem of Attouch [16]. According to this theorem, a sequence of closed proper convex functions φ^k converges as

$k \to \infty$ to a closed proper convex function φ if and only if the graph of $\partial \varphi^k$ converges to the graph of $\partial \varphi$ in the Painlevé-Kuratowski sense and, besides this, there is a point ξ and vector $\eta \in \partial \varphi(\xi)$ such that for some sequences $\xi^k \to \xi$ and $\eta^k \to \eta$ with $\eta^k \in \partial \varphi(\xi^k)$ one has $\varphi^k(\xi^k) \to \varphi(\xi)$. This theorem can be applied equally well, with the obvious minor changes in statement, to a family φ_t converging to φ. We need only observe now that for the functions φ_t in (2.5) the fact that $\varphi_t(0) = 0$ takes care of the assumption about convergence of function values. Thus in our context the functions φ_t epi-converge if and only if the graphs of their subdifferentials converge as sets.

Actually we want to apply this not to φ_t but to $\frac{1}{2}\varphi_t$, because

$$(2.11) \qquad \partial(\tfrac{1}{2}\varphi_t)(\xi) = [\partial f(x + t\xi) - v]/t$$

from (2.5). Our conclusion is that the functions $\frac{1}{2}\varphi_t$ epi-converge to a function $\frac{1}{2}f''_{x,v}$ if and only if the multifunctions in (2.11) converge graphically to a multifunction $(\partial f)'_{x,v}$, in which case the latter is the subdifferential $\partial(\frac{1}{2}f''_{x,v})$. This is what we had to prove. $\square$

Although the epi-convergence that defines $f''_{x,v}$ must in general be expressed in terms of the limits in (2.1) for the second-order difference quotients (2.5), there is a situation worth recording where simple pointwise convergence suffices.

Proposition 2.3. *Suppose that the second-order difference quotients in (2.5) converge pointwise on some nonempty open set C to finite values, but pointwise to ∞ outside of the closure of C. Then $f''_{x,v}$ exists and one has $C = \operatorname{int}(\operatorname{dom} f''_{x,v})$. Conversely, if $f''_{x,v}$ exists and $\operatorname{int}(\operatorname{dom} f''_{x,v})$ is nonempty then this property of convergence holds with respect to $C = \operatorname{int}(\operatorname{dom} f''_{x,v})$.*

Proof. This merely invokes a general criterion of Wets [45] for the reduction of epi-convergence to pointwise convergence in the case of convex functions. $\square$

Duality is addressed next. We recall that for the closed proper function f^* which is conjugate to f in the sense of the Legendre-Fenchel transform, the subdifferential multifunction ∂f^* is the inverse of ∂f; cf. [2, §23].

Theorem 2.4. *The function f is twice epi-differentiable at x relative to the vector $v \in \partial f(x)$ if and only if its conjugate f^* is twice epi-differentiable at v relative to the vector $x \in \partial f^*(v)$. Then the functions $\frac{1}{2}f''_{x,v}$ and $\frac{1}{2}f^{*''}_{v,x}$ are conjugate to each other.*

Proof. We rely on Theorem 2.2. The existence of $(\partial f)'_{x,v}$ is equivalent to that of $(\partial f^*)'_{v,x}$, and indeed these two multifunctions are just inverse to each other. This is immediate from the fact that ∂f and ∂f^* are inverse to each other. Theorem 2.2 tells us then that $f''_{x,v}$ exists if and only if $f^{*''}_{v,x}$ exists. Further, it tells us that in this case the multifunctions $\partial(\frac{1}{2}f''_{x,v})$ and $\partial(\frac{1}{2}f^{*''}_{v,x})$ are inverse to each other.

A closed proper convex function is, however, uniquely determined by its subdifferential multifunction up to an additive constant [2, Theorem 24.3].

Therefore $\frac{1}{2} f_{v,x}^{*\prime\prime}$ can differ from the conjugate of $\frac{1}{2} f_{x,v}^{\prime\prime}$ by an additive constant at most. But both of the functions in question vanish at the origin. It follows that the additive constant must be zero, because the conjugate of a nonnegative function that vanishes at the origin is always another nonnegative function that vanishes at the origin. □

The foregoing result can be placed in a "global" perspective by means of the following terminology.

Definition 2.5. The closed proper convex function f will be called *essentially twice epi-differentiable* if at *every* point x where $\partial f(x) \neq \varnothing$ (which implies $f(x)$ finite), f is twice epi-differentiable relative to *every* $v \in \partial f(x)$. It will simply be called *twice epi-differentiable* if, in addition, $\partial f(x) \neq \varnothing$ at every x where $f(x)$ is finite.

This terminology is consistent with the facts in Theorem 2.2.

Definition 2.6. A multifunction $F: \mathbb{R}^n \Rightarrow \mathbb{R}^m$ will be called *proto-differentiable* if at *every* x where $F(x) \neq \varnothing$, F is proto-differentiable relative to *every* $v \in F(x)$.

Theorem 2.7. *The following properties are equivalent to each other:*
   (a) *f is essentially twice epi-differentiable;*
   (b) *f^* is essentially twice epi-differentiable;*
   (c) *∂f is proto-differentiable;*
   (d) *∂f^* is proto-differentiable.*

Proof. Combine Theorems 2.2 and 2.4. □

3. Examples of twice epi-differentiable convex functions

We turn now to the identification of some classes of convex functions that do exhibit twice epi-differentiability. Many examples can first of all be derived from the results of our earlier paper [36] by taking special advantage of convexity. The following concepts will be needed.

Definition 3.1. A function $g: \mathbb{R}^d \to \overline{\mathbb{R}}$ is *piecewise linear-quadratic* if its effective domain $D = \{u \in \mathbb{R}^d \mid g(u) < \infty\}$ can be expressed as the union of finitely many polyhedral convex sets, relative to each of which g is at most quadratic.

Theorem 3.2. *If $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ has the form*

$$(3.1) \qquad\qquad f(x) = g(Ax + a)$$

for some $A \in \mathbb{R}^{m \times n}$, $a \in \mathbb{R}^n$, and function $g: \mathbb{R}^d \to \overline{\mathbb{R}}$ such that g is convex and piecewise linear-quadratic, then f is not only convex but twice epi-differentiable.

Proof. It is obvious that f is convex. In fact f is itself piecewise linear-quadratic. Hence by [36, Theorem 3.1], f is twice epi-differentiable at every point of its effective domain. □

Corollary 3.3. *If* $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ *is a polyhedral convex function, or in particular if* f *is the indicator of a polyhedral convex set, then* f *is twice epi-differentiable.*

Proof. For convex functions, polyhedrality as defined in convex analysis [2, §19] coincides with piecewise linearity in the sense of Definition 3.1. We therefore have the case of Theorem 3.2 where $A = I$ and $a = 0$. □

Theorem 3.4. *Suppose that* $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ *has the form*

$$(3.2) \qquad f(x) = g(G(x)) \quad with\ G(x) = (G_1(x), \dots, G_d(x)),$$

where $g: \mathbb{R}^d \to \overline{\mathbb{R}}$ *is convex and piecewise linear-quadratic and the notation is chosen so that the component functions* $G_k: \mathbb{R}^n \to \mathbb{R}$ *are convex of class* $\mathscr{C}^2$ *for* $k = 1, \dots, p$, *but affine for* $k = p + 1, \dots, d$. *Assume that* $g(u) = g(u_1, \dots, u_d)$ *is nondecreasing with respect to the variables* $u_1, \dots, u_p$, *and that there exist* $\bar{x} \in \mathbb{R}^n$ *and* $\bar{u} \in \mathbb{R}^d$ *with* $g(\bar{u}) < \infty$ *such that*

$$(3.3) \qquad \begin{aligned} G_k(\bar{x}) &< \bar{u}_k \quad for\ k = 1, \dots, p, \\ &= \bar{u}_k \quad for\ k = p + 1, \dots, d. \end{aligned}$$

Then f *is twice epi-differentiable.*

Proof. The convexity of f is easy to verify from the definition of convexity. Our line of argument for the rest of the theorem will be to reduce the situation to one where the "basic constraint qualification" of [36, Definition 4.1] is satisfied and then draw the conclusion of twice epi-differentiability from [36, Theorem 4.5], which treats functions of the form (3.2) under hypotheses not involving convexity.

Introduce for the affine functions G_k, $k = p + 1, \dots, d$, the notation $G_k(x) = c_k x + \gamma_k$ with the indices chosen so that $c_{p+1}, \dots, c_q$ is a maximal linearly independent subset of the vectors $c_{p+1}, \dots, c_d$ (for some $q \leq d$). Then there exist coefficients α_{kl} and β_k such that for all $x \in \mathbb{R}^n$ one has

$$G_k(x) = \sum_{l=p+1}^{q} \alpha_{kl} G_l(x) + \beta_k \quad for\ k = q + 1, \dots, q.$$

Let

$$\tilde{g}(u_1, \dots, u_q) = g\left(u_1, \dots, u_q, \sum_{l=1}^{q} \alpha_{q+1,l} u_l + \beta_{q+1}, \dots, \sum_{l=1}^{q} \alpha_{dl} u_l + \beta_d\right),$$

$$\widetilde{G}(x) = (G_1(x), \dots, G_q(x)).$$

Then $\tilde{g}$ is another piecewise linear-quadratic convex function and $f(x) = \tilde{g}(\widetilde{G}(x))$. Moreover the hypotheses of the theorem are satisfied for this representation of f as well as they were for the original one, but with the added provision that the gradients of the constraint functions designated as affine, namely the functions $G_{p+1}, \dots, G_q$ are now linearly independent. In view of the possibility of this reformulation, we can proceed henceforth under the

assumption that the functions $G_{p+1}, \ldots, G_d$ themselves have linearly independent gradients.

Consider now an arbitrary point x where $f(x)$ is finite, or in other words such that the vector $G(x)$ belongs to the set $D = \{u \in \mathbb{R}^d \mid g(u) < \infty\}$, which is a polyhedron by Definition 3.1. Let $y = (y_1, \ldots, y_d)$ be any normal vector to D at $G(x)$ in the sense that

$$(3.4) \qquad y \cdot (u - G(x)) \leq 0 \quad \text{for every } u \in D.$$

If we can show that

$$(3.5) \qquad \sum_{k=1}^{d} y_k \nabla G_k(x) = 0 \quad \text{implies} \quad y = (0, \ldots, 0),$$

this will mean that the basic constraint qualification of [36, Definition 4.1] is satisfied. Then by [36, Theorem 4.5], f will have $\partial f(x) \neq \emptyset$ and be twice epi-differentiable at x relative to every $v \in \partial f(x)$, which is our desired conclusion.

The assumption that g is nondecreasing in the variables $u_1, \ldots, u_p$ entails the property that whenever $u \in D$ and u' is a vector in $\mathbb{R}^d$ satisfying

$$u'_k \leq u_k \quad \text{for } k = 1, \ldots, p, \quad \text{and} \quad u'_k = u_k \quad \text{for } k = p+1, \ldots, d,$$

then $u' \in D$. From this and (3.4) it follows that

$$y_k \geq 0 \quad \text{for } k = 1, \ldots, p.$$

The function $h(x) = y \cdot G(x) = \sum_{k=1}^{d} y_k G_k(x)$ is therefore convex. Unless actually

$$(3.6) \qquad y_k = 0 \quad \text{for } k = 1, \ldots, p,$$

condition (3.4) as applied to the vector $\bar{u}$ in (3.3) would yield $y \cdot (G(\bar{x}) - G(x)) < 0$, or in other words that $h(\bar{x}) < h(x)$. This would be incompatible with the equation in (3.5), which says that $\nabla h(x) = 0$ and consequently by the convexity of h that x furnishes the global minimum of h on $\mathbb{R}^n$. The equation in (3.6) thus necessitates (3.7) and reduces to $\sum_{k=p+1}^{d} y_k c_k = 0$, where c_k is the (constant) gradient of the affine function G_k for $k = p+1, \ldots, d$. But these gradient vectors are supposed to be linearly independent. Therefore $y_k = 0$ for $k = p+1, \ldots, d$ in addition to (3.6). We conclude that the implication (3.5) is valid, which is all we needed to establish. $\square$

Corollary 3.5. *If f is of any one of the following types, then f is a closed proper convex function that is twice epi-differentiable.*

(a) $f = \max\{g_1, \ldots, g_r\}$ *on $\mathbb{R}^n$, where the functions $g_1, \ldots, g_r$ are convex and of class $\mathscr{C}^2$ on $\mathbb{R}^n$.*

(b) $f = \delta_C$, *the indicator of a nonempty set $C \subset \mathbb{R}^n$, where C consists of the points satisfying a constraint system of the form*

$$(3.7) \qquad \begin{aligned} x \in X \quad \text{and} \quad f_i(x) &\leq 0 \quad \text{for } i = 1, \ldots, s, \\ &= 0 \quad \text{for } i = s+1, \ldots, m, \end{aligned}$$

with X *a polyhedron,* f_i *a convex function of class* $\mathscr{C}^2$ *for* $i = 1, \ldots,$
s, *and* f_i *affine for* $i = s + 1, \ldots, m$, *under the assumption that there*
exists an $\bar{x} \in C$ *satisfying* $f_i(\bar{x}) < 0$ *for every* $i \in [1, s]$ *for which* f_i
is not affine.

(c) $f = f_0 + \delta_C$, *where* C *is a set of the type in* (b) *and* f_0 *is a convex*
function of class $\mathscr{C}^2$, *or more generally a max of finitely many such*
functions as in (a).

Proof. The general case of (c) subsumes (a) and (b), so we can concentrate on
it exclusively. We define

$$(3.8) \qquad G(x) = (g_1(x), \ldots, g_r(x), f_1(x), \ldots, f_m(x), l_1(x), \ldots, l_n(x)),$$

where

$$l_j(x) = x_j \quad (\text{the } j\text{th coordinate of } x),$$

and

$$g(u) = g(u_1, \ldots, u_{r+m+n})$$
$$= \begin{cases} \max\{u_1, \ldots, u_r\} & \text{if } (u_{r+1}, \ldots, u_{r+s}) \leq 0, \\ & \text{if } (u_{r+s+1}, \ldots, u_{r+m}) = 0, \\ & \text{if } (u_{r+m+1}, \ldots, u_{r+m+n}) \in X, \\ \infty & \text{otherwise}. \end{cases}$$

Then $f(x) = g(G(x))$. Obviously g is a polyhedral convex function, hence
piecewise linear, and g is nondecreasing in the variables $u_1, \ldots, u_{r+s}$. The
assumption about a point $\bar{x}$ gives condition (3.3) in terms of the component
functions in (3.8). The hypothesis of Theorem 3.4 is therefore fulfilled. $\square$

Remark 3.6. Actual formulas for the second-order epi-derivatives in Theorems
3.3 and 3.4 carry over by our proofs from the corresponding ones given in the
nonconvex setting in [36, Theorems 3.1 and 4.5]. In the case of Corollary 3.5,
the details as worked out in [36] give the following. Suppose that

$$(3.9) \qquad f(x) = \max\{g_1(x), \ldots, g_r(x)\} + \delta_C(x)$$

as in (c), and let $x \in C$. Let $K(x)$ be the set of indices $k \in [1, r]$ for which
the max in (3.9) is attained, and let $I(x)$ be the set of indices $i \in [1, s]$ for
which $f_i(x) = 0$ (the active inequality indices). Let $M(x)$ denote the set of all
multiplier vectors $(z, y) = (z_1, \ldots, z_r, y_1, \ldots, y_m)$ in $\mathbb{R}^r \times \mathbb{R}^m$ such that

$$z_k \geq 0 \quad \text{for } k \in K(x), \qquad z_k = 0 \quad \text{otherwise}; \qquad \sum_{k=1}^{r} z_k = 1,$$

$$y_i \geq 0 \quad \text{for } i \in I(x), \qquad y_i = 0 \quad \text{for other } i \in [1, s].$$

Then

$$v \in \partial f(x) \Leftrightarrow \exists (z, y) \in M(x)$$

$$(3.10) \qquad \qquad \text{with } v - \sum_{k=1}^{r} z_k \nabla g_k(x) - \sum_{i=1}^{m} y_i \nabla f_i(x) \in N_X(x),$$

where $N_X(x)$ is the normal cone to X at x. Furthermore, in terms of the polyhedral cone

$$\Xi(x) = \{\xi \in T_X(x) \mid \nabla g_k(x) \cdot \xi \leq 0 \text{ for all } k \in K(x),$$
$$\nabla f_i(x) \cdot \xi \leq 0 \text{ for all } i \in I(x),$$
$$\nabla f_i(x) \cdot \xi = 0 \text{ for } i = s + 1, \ldots, m\},$$

one has for any v as in (3.10) and the corresponding set $M_v(x)$ of multiplier vectors (z, y) appearing in (3.10) that

$$(3.11) \quad f''_{x,v}(\xi) = \begin{cases} \displaystyle\max_{(z,y)\in M_v(x)} \xi \cdot \left[\sum_{k=1}^{r} z_k \nabla^2 g_k(x) + \sum_{i=1}^{s} y_i \nabla^2 f_i(x)\right] \xi < \infty \\ \hspace{6cm} \text{when } \xi \in \Xi(x), \\ \infty \quad \text{when } \xi \notin \Xi(x). \end{cases}$$

Theorem 3.7. *Let f be a closed proper convex function on $\mathbb{R}^n$, and let φ be a function constructed from f in either of the following ways.*

    (a) *$\varphi = f + g$, where g is a convex function of class $\mathscr{C}^2$ on $\mathbb{R}^n$;*
    (b) *$\varphi = f \Box g$ (infimal convolution), where g is the conjugate of a convex function of class $\mathscr{C}^2$ on $\mathbb{R}^n$.*

    Then φ is a closed proper convex function, and φ is essentially twice epi-differentiable if and only if f is essentially twice epi-differentiable.

Proof. Case (a) merely extends [36, Proposition 2.10] to the terminology of Definition 2.5. The sum of a closed proper convex function and a finite convex function is well known to be a closed proper convex function. Case (b) follows from Theorem 2.4 by duality with (a): one has $f \Box h^* = (f^* + h)^*$ by [2, Theorem 16.4] when h is a finite convex function. $\quad\Box$

Corollary 3.8. *Let f be a closed proper convex function on $\mathbb{R}^n$ and let f^λ (for $\lambda > 0$) be the Moreau-Yosida approximate given by*

$$f^\lambda(x) = \min_{w \in \mathbb{R}^n} \left\{ f(w) + \frac{1}{2\lambda} |w - x|^2 \right\}.$$

Then f is essentially twice epi-differentiable if and only if f^λ is twice epi-differentiable.

Proof. Apply case (b) of Theorem 3.7 with $g(x) = (1/2\lambda)|x|^2$. $\quad\Box$

Corollary 3.9. *Let C be a nonempty closed convex set in $\mathbb{R}^n$. Let $d_C^2(x) = \text{dist}^2(x, C)$ and let $P_C(x)$ be the projection of x on C (nearest point mapping). Then the following properties are equivalent to each other:*

    (a) *the indicator δ_C is twice epi-differentiable;*
    (b) *the convex function d_C^2 is twice epi-differentiable;*
    (c) *the projection mapping P_C is proto-differentiable.*

Proof. For $f = \delta_C$ we have $d_C^2 = f^\lambda$ for $\lambda = \frac{1}{2}$. The equivalence between (a) and (b) is therefore a special case of the equivalence in Corollary 3.8. The equivalence between (b) and (c) follows from Theorem 2.4 because $P_C = \frac{1}{2}\nabla[d_C^2]$. □

Corollary 3.9 can be applied in particular, of course, to the vast class of convex sets C covered by Corollary 3.5(b).

4. Saddle functions and epi/hypo-convergence

A function $K \colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ is a *saddle function* if $K(x, y)$ is convex in x and concave in y. The convention that convexity occurs in the first argument and concavity in the second is of course arbitrary, but we shall rely on this particular asymmetry to keep straight a number of definitions. Thus for any function $K \colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$, whether a saddle function or not, we define the *effective domain* of K by

$$(4.1) \quad \operatorname{dom} K = \operatorname{dom}_1 K \times \operatorname{dom}_2 K = \{x \mid K(x, \cdot) > -\infty\} \times \{y \mid K(\cdot, y) < \infty\}$$

and say that K is *proper* when this is nonempty. Furthermore, we denote by $\operatorname{cl}_1^* K$ the function obtained by taking the (convex) closure of K in the first argument: for each y, the function $(\operatorname{cl}_1^* K)(\cdot, y)$ is the greatest closed convex function majorized by $K(\cdot, y)$, i.e., the biconjugate of $K(\cdot, y)$. (Recall that in convex analysis, a convex function is *closed* either if it is proper and lower semicontinuous, or if it is the constant function ∞ or the constant function $-\infty$.) In parallel fashion, we denote by $\operatorname{cl}_2^* K$ the function obtained by taking the (concave) closure of K in the second argument.

Two saddle functions K and L are called *equivalent* if $\operatorname{cl}_1^* K = \operatorname{cl}_1^* L$ and $\operatorname{cl}_2^* K = \operatorname{cl}_2^* L$. A saddle function K is said to be *closed* if K is equivalent to $\operatorname{cl}_1^* K$ and $\operatorname{cl}_2^* K$. This definition of closedness stems from the fact that the class of saddle functions K satisfying $K = \operatorname{cl}_1^* K = \operatorname{cl}_2^* K$ is too narrow. It can be shown, for instance, that starting from an arbitrary function $K \colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$, the functions $\operatorname{cl}_1^* \operatorname{cl}_2^* K$ and $\operatorname{cl}_2^* \operatorname{cl}_1^* K$ will always be closed saddle functions, and if K is itself a saddle function and proper, then they will be equivalent and only depend on the equivalence class containing K. In this sense, therefore, a closure operation exists for saddle functions K even though $\operatorname{cl}_1^* \operatorname{cl}_2^* K$ and $\operatorname{cl}_2^* \operatorname{cl}_1^* K$ generally do not coincide.

The theory of equivalence classes of closed saddle functions on $\mathbb{R}^n \times \mathbb{R}^m$ was first developed in [46]; see also [2, §§33–37]. It is deeply connected with minimax theorems and the study of Lagrange multipliers and duality in convex optimization. Each equivalence class is actually an interval consisting of all the saddle functions K satisfying $\underline{K} \leq K \leq \overline{K}$, where $\underline{K}$ and $\overline{K}$ are saddle functions satisfying

$$(4.2) \qquad\qquad \operatorname{cl}_2^* \underline{K} = \overline{K} \quad \text{and} \quad \operatorname{cl}_1^* \overline{K} = \underline{K}.$$

We sometimes emphasize this fact by writing $[\underline{K}, \overline{K}]$ for the equivalence class.

All the saddle functions K in a closed, proper equivalence class $[\underline{K}, \overline{K}]$ have the same effective domain $\operatorname{dom} K$, where they are all finite. Their values

agree when $x \in \mathrm{ri}(\mathrm{dom}_1 K)$ or when $y \in \mathrm{ri}(\mathrm{dom}_2 K)$. They also have the same subgradients, which are defined by

$$\partial K(x, y) = \partial_1 K(x, y) \times \partial_2 K(x, y)$$

(4.3) $\qquad = \{(v, u) \mid \text{the function } K_{v, u}(\xi, \eta) = K(\xi, \eta) - v \cdot \xi - u \cdot \eta$

$$\text{has a saddle point at } (\xi, \eta) = (x, y)\},$$

where $\partial_1 K(x, y)$ is the set of subgradients of $K(\cdot, y)$ at x, and $\partial_2 K(x, y)$ is the set of subgradients of $K(x, \cdot)$ at y.

Later, when we come to the extension to saddle functions of the results in §§2 and 3, we shall be involved very strongly with several fundamental relations of conjugacy. First there is the concept of conjugacy for saddle functions themselves. Given any closed proper saddle function K, one defines

(4.4) $\qquad \overline{K}^*(v, u) = \inf_y \sup_x \{v \cdot x + u \cdot y - K(x, y)\},$

(4.5) $\qquad \underline{K}^*(v, u) = \sup_x \inf_y \{v \cdot x + u \cdot y - K(x, y)\}.$

These are closed proper saddle functions which are equivalent to each other and satisfy

(4.6) $\qquad \mathrm{cl}_2^* \underline{K}^* = \overline{K}^* \quad \text{and} \quad \mathrm{cl}_1^* \overline{K}^* = \underline{K}^*.$

Moreover they depend only on the equivalence class $[\underline{K}, \overline{K}]$ that contains K.

Thus $[\underline{K}^*, \overline{K}^*]$ is an equivalence class of closed proper saddle functions; it is said to be *conjugate* to $[\underline{K}, \overline{K}]$. For any saddle function K^* in the class $[\underline{K}^*, \overline{K}^*]$ one has

(4.7) $\qquad \overline{K}(x, y) = \inf_u \sup_v \{v \cdot x + u \cdot y - K^*(v, u)\},$

(4.8) $\qquad \underline{K}(x, y) = \sup_v \inf_x \{v \cdot x + u \cdot y - K^*(v, u)\},$

so that the class $[\underline{K}, \overline{K}]$ is in turn conjugate to $[\underline{K}^*, \overline{K}^*]$. Furthermore, one has

(4.9) $\qquad (v, u) \in \partial K(x, y) \Leftrightarrow (x, y) \in \partial K^*(v, u).$

For any closed proper saddle function K, the partial conjugate functions F and G defined by

(4.10) $\qquad F(x, u) = \sup_y \{K(x, y) - u \cdot y\},$

(4.11) $\qquad G(v, y) = \inf_x \{K(x, y) - v \cdot x\}$

are called the *parents* of K. They depend only on the equivalence class $[\underline{K}, \overline{K}]$ containing K, and indeed this class can be recovered from its parents by the formulas

(4.12) $\qquad \overline{K}(x, y) = \inf_u \{F(x, y) + u \cdot y\},$

(4.13) $\qquad \underline{K}(x, y) = \sup_v \{G(v, y) + v \cdot x\}.$

From this it is clear that

$$(4.14) \qquad G(v, y) = \inf_{x, u} \{F(x, u) - v \cdot x + u \cdot y\} = -F^*(v, -y),$$

$$(4.15) \qquad F(x, u) = \sup_{v, y} \{G(v, y) - u \cdot y + v \cdot x\} = -G^*(-x, u).$$

Indeed, F is a closed proper convex function on $\mathbb{R}^n \times \mathbb{R}^m$ and G is a closed proper concave function on $\mathbb{R}^n \times \mathbb{R}^m$. Moreover, for any such F and G related by (4.14)–(4.15), the formulas (4.12) and (4.13) determine an equivalence class $[\underline{K}, \overline{K}]$. Thus there is a three-way one-to-one correspondence between the closed proper convex functions F on $\mathbb{R}^n \times \mathbb{R}^m$, the closed proper concave functions G on $\mathbb{R}^n \times \mathbb{R}^m$, and the equivalence classes of closed proper saddle functions K on $\mathbb{R}^n \times \mathbb{R}^m$. Under this correspondence one has

$$(4.16) \quad (v, u) \in \partial K(x, y) \Leftrightarrow (v, -y) \in \partial F(x, u) \Leftrightarrow (-x, u) \in \partial G(v, y).$$

Let us turn now to the notions of convergence that fit with this setting. For a family of functions F_t for $t > 0$ we can consider epi-convergence to a function F as $t \downarrow 0$, as in §2. Quite similarly there is a notion of *hypo-convergence* of a family of functions G_t to a function G. The latter merely involves interchanging the roles of "inf" and "sup" in the definition given for epi-convergence. When the functions F_t and F are closed proper convex, and G_t and G are the corresponding closed proper concave functions under (4.9), (4.10), these notions of convergence are equivalent to each other by the well-known theorem of Wijsman [42] on the continuity of the Legendre-Fenchel transform.

In view of the three-way correspondence mentioned above, there must be an equivalent notion of convergence for the corresponding saddle function classes $[\underline{K}_t, \overline{K}_t]$ and $[\underline{K}, \overline{K}]$, but what is it? An answer has been provided by Attouch, Azé, and Wets [38] in a general setting of Banach spaces. Here we specialize and modify their result so as to present the answer in a somewhat simpler and sharper form which is more convenient for the case at hand.

Definition. A family of closed proper saddle functions K_t for $t > 0$ will be called *modulated* as $t \downarrow 0$ if for some $\rho \geq 0$ sufficiently large and $\tau > 0$ sufficiently small, one has for all $t \in (0, \tau)$ that

$$(4.17) \qquad \inf_{|x| \leq \rho} \overline{K}_t(x, y) \leq \rho(1 + |y|) \quad \text{for all } y,$$

$$(4.18) \qquad \sup_{|y| \leq \rho} \underline{K}_t(x, y) \geq -\rho(1 + |x|) \quad \text{for all } x.$$

Such a family will be said to *epi/hypo-converge* (*in the modulated sense*) to a closed proper saddle function K if it is modulated as $t \downarrow 0$ and satisfies

$$(4.19) \qquad \limsup_{\substack{y' \to y \\ t \downarrow 0}} \inf_{x' \to x} \overline{K}_t(x', y') \leq \overline{K}(x, y),$$

$$(4.20) \qquad \liminf_{\substack{x' \to x \\ t \downarrow 0}} \sup_{y' \to y} \underline{K}_t(x', y') \geq \underline{K}(x, y).$$

Note. For simplicity in this paper, when we use the term epi/hypo-convergence we shall always mean epi/hypo-convergence in the modulated sense, as just defined. This contrasts other senses of epi/hypo-convergence described in [38] *and* [47].

Observe that this property of convergence is really one of *equivalence classes* rather than of individual functions. The same is true of the modulation property by itself.

Observe further that the modulation property is satisfied trivially in the special case where the functions K_t all have a saddle point at $(0, 0)$ with saddle value $K_t(0, 0) = 0$. Then one can simply take $\rho = 0$. This is in fact the case that will occupy us in our subsequent work with second derivatives.

The following theorem ensures in particular that for modulated families K_t there is at most one epi/hypo-limit K up to equivalence.

Theorem 4.1. *Let K_t for $t > 0$ and K be closed proper saddle functions with convex parents F_t and F, and concave parents G_t and G. Then the following are equivalent:*

(a) K_t *epi/hypo-converges to K as $t \downarrow 0$,*

(b) G_t *hypo-converges to G as $t \downarrow 0$,*

(c) F_t *epi-converges to F as $t \downarrow 0$.*

Proof. Only a modest addition to the arguments of Attouch, Azé, and Wets [38, Theorem 3.1] will be needed. As a preliminary, however, we must show that property (4.17) in the definition of "modulated" is equivalent (for ρ sufficiently large) to

$$(4.21) \qquad \min_{|x| \leq \rho, \, |u| \leq \rho} F_t(x, u) \leq \rho,$$

whereas (4.18) is equivalent to

$$(4.22) \qquad \max_{|v| \leq \rho, \, |y| \leq \rho} G_t(v, y) \leq -\rho.$$

Clearly (4.17) can be written in the form

$$(4.23) \qquad \sup_y \inf_x \overline{K}_t^\rho(x, y) \leq \rho,$$

with

$$\overline{K}_t^\rho(x, y) = \begin{cases} \overline{K}_t(x, y) - \rho|y| & \text{if } |x| \leq \rho, \\ \infty & \text{if } |x| > \rho. \end{cases}$$

For ρ sufficiently large, $\overline{K}_t^\rho$ is a closed proper saddle function with $\text{dom}_1 \overline{K}_t^\rho$ bounded. For such a function we have

$$\sup_y \inf_x \overline{K}_t^\rho(x, y) = \min_x \sup_y \overline{K}_t^\rho(x, y)$$

by [2, Corollary 37.3.1]. Thus (4.23) is equivalent (for ρ sufficiently large) to

$$(4.24) \qquad \rho \geq \min_{|x| \leq \rho} \sup_y \{\overline{K}_t(x, y) - \rho|y|\}.$$

The inner supremum in (4.24) can be written by Fenchel's duality theorem [2, Theorem 31.1] and the conjugacy relations (4.10), (4.12) between $\overline{K}_t(x;\cdot)$ and $F_t(x,\cdot)$ as

$$\inf_{|u|\le\rho} F_t(x,u).$$

Substituting this into (4.24) we get the desired inequality (4.21). The proof that (4.18) is equivalent to (4.22) runs along similar lines—or one can just invoke duality.

Now let $\widetilde{K}$ and $\underline{K}$ denote the functions on the left in (4.19) and (4.20), so that these two inequalities, which characterize epi/hypo-convergence, can be written as

$$(4.25) \qquad\qquad \widetilde{K} \le \overline{K} \quad\text{and}\quad \underline{K} \ge \underline{K}.$$

What Attouch, Azé, and Wets state in effect in [38, Theorem 3.1] is that if either (4.21) or (4.22) hold for $0 < t < \tau$ (where τ is sufficiently small and ρ sufficiently large), then (b) and (c) are equivalent to

$$(4.26) \qquad\qquad \mathrm{cl}_1^*\,\widetilde{K} \le \underline{K} \quad\text{and}\quad \mathrm{cl}_2^*\,\underline{K} \ge \overline{K}.$$

The equivalence of (b) with (c) was already known to be true even without such assumptions, and indeed it is obvious that (b) implies (4.21) and (c) implies (4.22). The only real issue then in deriving the present version of the result is whether, under both of these assumptions, (4.25) is equivalent to (4.26).

Certainly (4.25) does imply (4.26) by (4.2). The pattern of argument used in [38], on which we shall rely here as well, consists in showing first that the inequalities

$$\limsup_{t\downarrow 0}\ \inf_{(x',u')\to(x,u)} F_t(x',u') \le F(x,u),$$

$$\liminf_{t\downarrow 0}\ \sup_{(v',y')\to(v,y)} G_t(v',y') \ge G(v,y),$$

which follow from (b) and (c), imply (4.26). Actually the proof that is furnished for [38, Theorem 3.1] shows that they imply the stronger condition (4.25). The remaining part of the argument in [38] is to verify that (4.26) yields the inequalities

$$\liminf_{t\downarrow 0}\ \inf_{(x',u')\to(x,u)} F_t(x',u') \ge F(x,u),$$

$$\limsup_{t\downarrow 0}\ \sup_{(v',y')\to(v,y)} G_t(v',y') \le G(v,y),$$

which are known to imply (b) and (c) under either of the assumptions (4.21) or (4.22). This pattern thus in fact yields a circle of implications proving that (4.25) and (4.26) are equivalent, as required. □

From Theorem 4.1, we can derive the continuity of the conjugacy correspondence under epi/hypo-convergence.

Theorem 4.2. *For closed proper saddle functions K_t and K having conjugates K_t^* and K^*, the following are equivalent:*

(a) K_t *epi/hypo-converges to K as $t \downarrow 0$;*

(b) K_t^* *epi/hypo-converges to K^* as $t \downarrow 0$.*

Proof. By writing (4.10) and (4.11) in the form

$$(4.27) \qquad -F(x, u) = \inf_y \{u \cdot y - K(x, y)\},$$

$$(4.28) \qquad -G(v, y) = \sup_x \{v \cdot x - K(x, y)\},$$

we obtain from (4.4) and (4.5) that

$$(4.29) \qquad \overline{K}^*(v, u) = \inf_y \{x \cdot v - G(v, y)\},$$

$$(4.30) \qquad \underline{K}^*(v, u) = \sup_x \{y \cdot u - F(x, u)\}.$$

The functions $-G$ and $-F$ thus serve as the convex and concave parents of K^*. Obviously F_t epi-converges to F if and only if $-F_t$ hypo-converges to $-F_t$. By Theorem 4.1, the first property is equivalent to the present property (a), while the second is equivalent to the present (b). $\square$

The next theorem offers a small improvement over the result of Attouch, Azé, and Wets [38, Theorem 3.6] on the graphical convergence of subdifferentials of saddle functions. The improvement consists in translating the "normalizing condition" on function values into the saddle function context.

Theorem 4.3. *Let K_t and K be closed proper saddle functions with convex parents F_t and F and concave parents G_t and G, as well as conjugates K_t^* and K^*. Then the following conditions are equivalent to the ones in Theorem 4.1 and Corollary 4.2.*

(a) *∂K_t converges graphically to ∂K as $t \downarrow 0$, and one can find $x_t \to x$, $y_t \to y$, $u_t \to u$, $v_t \to v$ such that $(v_t, u_t) \in \partial K_t(x_t, y_t)$ and $K_t(x_t, y_t) \to K(x, y)$;*

(b) *∂F_t converges graphically to ∂F as $t \downarrow 0$ and one can find $x_t \to x$, $y_t \to y$, $u_t \to u$, $v_t \to v$, such that $(v_t, -y_t) \in \partial F_t(x_t, u_t)$ and $F_t(x_t, u_t) \to F(x, u)$;*

(c) *∂G_t converges graphically to ∂G as $t \downarrow 0$, and one can find $x_t \to x$, $y_t \to y$, $u_t \to u$, $v_t \to v$, such that $(-x_t, u_t) \in \partial G_t(v_t, y_t)$ and $G_t(v_t, y_t) \to G(v, y)$;*

(d) *∂K_t^* converges graphically to ∂K^* as $t \downarrow 0$, and one can find $x_t \to x$, $y_t \to y$, $u_t \to u$, $v_t \to v$ such that $(x_t, y_t) \in \partial K_t^*(v_t, u_t)$ and $K_t^*(v_t, u_t) \to K^*(v, u)$.*

Proof. From (4.16) we have

$$(v_t, u_t) \in \partial K_t(x_t, y_t) \Leftrightarrow (v_t, -y_t) \in \partial F_t(x_t, u_t) \Leftrightarrow (-x_t, u_t) \in \partial G_t(v_t, y_t).$$

From (4.9) we also have

$$(x_t, y_t) \in \partial K_t^*(v_t, u_t) \Leftrightarrow (v_t, u_t) \in \partial K_t(x_t, y_t).$$

These equivalent conditions imply by the conjugacy relations that

$$F_t(x_t, u_t) + u_t \cdot y_t = K_t(x_t, y_t) = G_t(v_t, y_t) + v_t \cdot x_t,$$

$$K_t^*(v_t, u_t) = v_t \cdot x_t + u_t \cdot y_t - K_t(x_t, y_t).$$

Such equations establish that (a), (b), (c), and (d) of the present theorem are equivalent to each other. By the theorem of Attouch [16], (b) of the present theorem is equivalent to (b) of Theorem 4.1, and this gives us our result. □

5. Epi/hypo-derivatives of saddle functions

Let K be a closed proper saddle function on $\mathbb{R}^n \times \mathbb{R}^m$ with convex and concave parents F and G as above, and let (x, y) be a point of dom K. We define K to be *epi/hypo-differentiable* at (x, y) if the difference quotients

(5.1) $k_t(\xi, \eta) = [K(x + t\xi, y + t\eta) - K(x, y)]/t$ for $t > 0$,

which are closed proper saddle functions, epi/hypo-converge (as defined in §4) to a closed proper saddle function k. Then in place of $k(\xi, \eta)$ we write $K'_{x,y}(\xi, \eta)$, bearing in mind that the symbol $K'_{x,y}$ only represents an element of an equivalence class of saddle functions, i.e., is only defined up to equivalence. The upper and lower elements of the class are denoted by $\overline{K}'_{x,y}$ and $\underline{K}'_{x,y}$.

If K is epi/hypo-differentiable at (x, y) in this sense, we consider further, for any vector pair (v, u), the second-order difference quotients

(5.2) $k_t(\xi, \eta) = [K(x + t\xi, y + t\eta) - K(x, y) - tv \cdot \xi - tu \cdot \eta]/\frac{1}{2}t^2$ for $t > 0$.

Again, these are closed proper saddle functions. If they epi/hypo-converge to a closed proper saddle function k as $t \downarrow 0$, we say that K is *twice epi/hypo-differentiable* at (x, y) relative to (v, u). The function k is denoted then by $K''_{x,y;u,v}$, which thus merely denotes an arbitrary member of a certain equivalence class of closed proper saddle functions. The upper and lower elements of this class are written as $\overline{K}''_{x,y;u,v}$ and $\underline{K}''_{x,y;u,v}$.

It may not be obvious that these concepts really depend only on the equivalence class containing K, but this is one of the consequences of Theorems 5.1 and 5.2 below.

It is worth noting that in the second-derivative case (5.2) the difference quotient satisfies $k_t(0, 0) = 0$ for all t. Thus in this case k_t is certainly modulated as $t \downarrow 0$, and the question of epi/hypo-convergence of k_t to k revolves simply around the inequalities

$$\limsup_{\substack{\eta' \to \eta \\ t \downarrow 0}} \inf_{\xi' \to \xi} [\overline{K}(x + t\xi', y + t\eta') - K(x, y)]/\tfrac{1}{2}t^2 \le \overline{k}(\xi, \eta),$$

$$\liminf_{\substack{\xi' \to \xi \\ t \downarrow 0}} \inf_{\eta' \to \eta} [\underline{K}(x + t\xi', y + t\eta') - K(x, y)]/\tfrac{1}{2}t^2 \ge \underline{k}(\xi, \eta).$$

Theorem 5.1. *For K to be epi/hypo-differentiable at (x, y), it is necessary and sufficient that $\partial K(x, y) \neq \varnothing$. Then*

$$(5.3) \qquad K'_{x,y}(\xi, \eta) = \sup_{v \in \partial_1 K(x,y)} \xi v + \inf_{u \in \partial_2 K(x,y)} \eta u$$

whenever (ξ, η) is such that the "sup" is not ∞, or the inf is not $-\infty$. (If $\infty - \infty = \infty$ is used in this formula for (ξ, η) such that the "sup" is ∞ and the "inf" is $-\infty$, one gets $\overline{K}'_{x,y}$. If $\infty - \infty = -\infty$ is used instead, one gets $\underline{K}'_{x,y}$. In this sense, the formula covers the entire equivalence class.)

Proof. Taking k_t to be the function in (5.1), we demonstrate first that k_t is modulated as $t \downarrow 0$ if and only if $\partial K(x, y) \neq \varnothing$. Suppose, to begin with, that $\partial K(x, y) \neq \varnothing$; let $(v, u) \in \partial K(x, y)$. By (4.3) we have

$$K(x + t\xi, y) \geq K(x, y) - t\xi \cdot v \quad \text{for all } \xi,$$

$$K(x, y + t\eta) \leq K(x, y) - t\eta \cdot u \quad \text{for all } \eta,$$

so that for $\rho = \max\{|u|, |v|\}$ we have

$$k_t(\xi, 0) \geq -\xi \cdot v \geq -\rho|\xi| \quad \text{for all } \xi,$$

$$k_t(0, \eta) \leq -\eta \cdot u \leq -\rho|\eta| \quad \text{for all } \eta.$$

Then in particular the convex function $k_t(\cdot, 0)$ is proper, so that the function $\underline{k}_t(\cdot, 0) = \text{cl}_1^* k(\cdot, 0)$ is just the lower semicontinuous hull of $k_t(\cdot, 0)$ and satisfies $\underline{k}_t(\xi, 0) \geq -\rho|\xi|$. Similarly $\overline{k}_t(0, \eta) \leq \rho|\eta|$. Trivially the definition of k_t being modulated as $t \downarrow 0$ is then satisfied:

$$\inf_{|\xi| \leq \rho} \overline{k}_t(\xi, \eta) \leq \rho(1 + |\eta|) \quad \text{for all } \eta,$$

$$\sup_{|\eta| \leq \rho} \underline{k}_t(\xi, \eta) \geq \rho(1 + |\xi|) \quad \text{for all } \xi.$$

Conversely, suppose that k_t is modulated as $t \downarrow 0$ and consider the parents of k_t, which are the functions

$$(5.4) \quad f_t(\xi, u) = \sup_\eta \{k_t(\xi, \eta) - \eta \cdot u\} = [F(x + t\xi, u) + y \cdot u - K(x, y)]/t,$$

$$(5.5) \quad g_t(v, \eta) = \inf_\xi \{k_t(\xi, \eta) - \xi \cdot v\} = [G(v, y + t\eta) + x \cdot v - K(x, y)]/t.$$

By the argument given at the beginning of the proof of Theorem 4.1, the modulation property implies for $\rho > 0$ sufficiently large and $\tau > 0$ sufficiently small that

$$(5.6) \qquad \min_{|\xi| \leq \rho, |u| \leq \rho} f_t(\xi, u) \leq \rho \quad \text{when } 0 < t < \tau,$$

$$(5.7) \qquad \max_{|v| \leq \rho, |\eta| \leq \rho} g_t(v, \eta) \geq -\rho \quad \text{when } 0 < t < \tau.$$

For any sequence $t_i \downarrow 0$, there exist by (5.6) sequences $\{\xi_i\}$ and $\{u_i\}$ with $|\xi_i| \leq \rho$ and $|u_i| \leq \rho$ such that

$$F(x + t_i\xi_i, u_i) \leq t_i\rho + K(x, y) - y \cdot u_i.$$

Passing to subsequences if necessary, we can suppose that u_i converges to some u. Then

$$F(x, u) \leq K(x, y) - y \cdot u,$$

and inasmuch as

$$F(x, u) = \sup_{y'} \{K(x, y') - y' \cdot u\}$$

we are able to conclude that

$$K(x, y') \leq K(x, y) + u \cdot (y' - y) \quad \text{for all } y',$$

which means that $u \in \partial_2 K(x, y)$. In a similar way, (5.7) yields the existence of some v such that $v \in \partial_1 K(x, y)$. Thus $\partial K(x, y) \neq \varnothing$, as claimed.

Concentrating now on the case where $\partial K(x, y) \neq \varnothing$, we work toward the derivative formula (5.3). Consider the functions

$$\theta_1(\xi) = \lim_{t \downarrow 0} [K(x+t\xi, y) - K(x, y)]/t, \quad \theta_2(\eta) = \lim_{t \downarrow 0} [K(x, y+t\eta) - K(x, y)]/t.$$

The first limit is monotone decreasing, and the second is monotone increasing. We know from convex analysis that θ_1 is a positively homogeneous, proper convex function whose (convex) closure is

$$(5.8) \qquad \underline{\theta}_1(\xi) = \sup_{v \in \partial_1 K(x, y)} \xi \cdot v.$$

whereas θ_2 is a positively homogeneous, proper concave function whose (concave) closure is

$$(5.9) \qquad \overline{\theta}_2(\eta) = \inf_{u \in \partial_2 K(x, y)} \eta \cdot u.$$

An equivalence class of closed, proper saddle functions is therefore given by the interval $[\underline{k}, \overline{k}]$, where

$$\overline{k}(\xi, \eta) = \underline{\theta}_1(\xi) + \overline{\theta}_2(\eta) \quad \text{with } \infty - \infty = \infty,$$

$$\underline{k}(\xi, \eta) = \underline{\theta}_1(\xi) + \overline{\theta}_2(\eta) \quad \text{with } \infty - \infty = -\infty.$$

We want to show for k in this class that k_t epi/hypo-converges to k.

The parent functions for k are

$$(5.10) \qquad f(\xi, u) = \underline{\theta}_1(\xi) + \delta(u \mid \partial_2 K(x, y)),$$

$$(5.11) \qquad g(v, \eta) = \overline{\theta}_2(\eta) - \delta(v \mid \partial_1 K(x, y)).$$

It will be proved that

$$(5.12) \qquad \limsup_{t \downarrow 0} \inf_{(\xi', u') \to (\xi, u)} f_t(\xi', u') \leq f(\xi, u),$$

$$(5.13) \qquad \liminf_{t \downarrow 0} \sup_{(v', \eta') \to (v, \eta)} g_t(v', \eta') \geq g(v, \eta).$$

Because of the usual relations between convex and concave parents, in (4.14) and (4.15), the inequality (5.13) is equally well written as

$$(5.14) \qquad \limsup_{t \downarrow 0} \inf_{(v', \eta') \to (v, \eta)} f_t^*(v', \eta') \leq f^*(v, \eta).$$

From (5.12) and (5.14) it will follow that f_t epi-converges to f as $t \downarrow 0$ (cf. [38]), and consequently by Theorem 4.1 that k_t epi/hypo-converges to k as $t \downarrow 0$. This will finish the proof of the theorem.

By duality we need only prove (5.12), and for this it will suffice to establish that

$$(5.15) \qquad \limsup_{t \downarrow 0} \; \inf_{(x', u') \to (\xi, u)} f_t(\xi', u') \leq \underline{\theta}_1(\xi) \quad \text{when } u \in \partial_2 K(x, y).$$

The condition $u \in \partial_2 K(x, y)$ gives $F(x, u) = K(x, y) - u \cdot y$ (in particular $F(x, u)$ must thus be finite) and allows us to express

$$(5.16) \qquad f_t(\xi', u') = [F(x + t\xi', u') - F(x, u) + y \cdot (u' - u)]/t.$$

With such a u fixed from now on, we denote by $\varphi(\xi)$ the left side of (5.15), so that

$$(5.17) \qquad \varphi(\xi) = \lim_{\varepsilon \downarrow 0} \left[\limsup_{t \downarrow 0} \left[\min_{\substack{|\xi' - \xi| \leq \varepsilon \\ |u' - u| \leq \varepsilon}} f_t(\xi', u') \right] \right].$$

The function φ is lower semicontinuous, in particular. Our aim is to demonstrate that $\varphi \leq \underline{\theta}_1$.

From (5.17) we see that for every vector $\omega \in \mathbb{R}^m$ it is true that

$$\begin{aligned} \varphi(\xi) &\leq \lim_{\varepsilon \downarrow 0} \left[\limsup_{t \downarrow 0} [f_t(\xi, u + t\omega)] \right] \\ &= \limsup_{t \downarrow 0} [F(x + t\xi, u + t\omega) - F(x, y) + ty \cdot \omega]/t \\ &= F'(x, y; \xi, \omega) + y \cdot \omega. \end{aligned}$$

Therefore

$$(5.18) \qquad \varphi(\xi) \leq \inf_{\omega} \{ F'(x, u; \xi, \omega) + y \cdot \omega \}.$$

The infimum here is a function $\varphi_0(\xi)$ which is convex and positively homogeneous in ξ, with $\varphi_0(0) < \infty$. This is because $F'(x, u; \xi, \omega)$ is convex and positively homogeneous in (ξ, ω), with $F'(x, u; 0, 0) = 0$. If we can prove that

$$(5.19) \qquad \underline{\theta}_1(\xi) = \liminf_{\xi' \to \xi} \varphi_0(\xi'),$$

we will know from the inequality $\varphi \leq \varphi_0$ in (5.18) and the lower semicontinuity of φ that $\varphi \leq \underline{\theta}_1$, as desired.

Because φ_0 is convex and positively homogeneous with $\varphi_0(0) < \infty$, the right side of (5.19) is the support function of the set

$$C = \{ v \mid v \cdot \xi \leq \varphi_0(\xi) \text{ for all } \xi \}$$

[2, Corollary 13.2.1], if $C \neq \varnothing$. Thus with (5.8) in view, we have brought our task down to showing that $C = \partial_1 K(x, y)$.

Since $\varphi_0(\xi)$ denotes the right side of (5.18), we have $v \in C$ if and only if

$$v \cdot \xi - y \cdot \omega \leq F'(x, u; \xi, \omega) \quad \text{for all } \xi, \omega.$$

This inequality means that $(v, -y) \in \partial F(x, u)$. But the latter is equivalent by (4.16) to having $v \in \partial_1 K(x, y)$ and $u \in \partial_2 K(x, y)$. We chose u as an element of $\partial_2 K(x, y)$, so it is indeed true that $v \in C$ if and only if $v \in \partial_1 K(x, y)$. The proof of Theorem 5.1 is now complete. $\square$

Theorem 5.2. *The following conditions are equivalent and necessitate having $(v, u) \in \partial K(x, y)$:*

    (a) *K is twice epi/hypo-differentiable at (x, y) relative to (v, u);*
    (b) *F is twice epi-differentiable at (x, u) relative to $(v, -y)$;*
    (c) *G is twice hypo-differentiable at (v, y) relative to $(-x, u)$.*

When these conditions hold, the epi-hypo-derivative $K''_{x,y;v,u}$ is a closed proper saddle function which is positively homogeneous of degree two and has a saddle point at $(0, 0)$:

$$(5.20) \quad K''_{x,y;v,u}(\lambda\xi, \lambda\eta) = \lambda^2 K''_{x,y;v,u}(\xi, \eta) \quad \text{for } \lambda > 0,$$

$$K''_{x,y;v,n}(0, 0) = 0,$$

$$(5.21) \quad K''_{x,y;v,u}(\xi, 0) \geq K''_{x,y;v,u}(0, 0) \geq K''_{x,y;v,n}(0, \eta) \quad \text{for all } \xi, \eta.$$

Moreover the convex parent of $\frac{1}{2}K''_{x,y;v,u}$ is $\frac{1}{2}F''_{x,u;v,-y}$, and the concave parent is $\frac{1}{2}G''_{v,y;-x,v}$.

Proof. For each $t > 0$ the function k_t in (5.2) gives

$$(5.22) \quad \tfrac{1}{2}k_t(\xi, \eta) = [K(x + t\xi, y + t\eta) - K(x, y) - t\xi \cdot v - t\eta \cdot u]/t^2.$$

The convex parent of this saddle function $\frac{1}{2}k_t$ is readily calculated to be $\frac{1}{2}f_t$, where

$$(5.23) \quad f_t(\xi, \omega) = [F(x + t\xi, u_t\omega) - (K(x, y) - y \cdot u) - t\xi \cdot v + t\omega \cdot y]/\tfrac{1}{2}t^2.$$

At the same time, the concave parent is $\frac{1}{2}g_t$, where

$$(5.24) \quad g_t(\zeta, \eta) = [G(v + t\zeta, y + t\eta) - (K(x, y) - x \cdot v) + t\zeta \cdot x - t\eta \cdot u]/\tfrac{1}{2}t^2.$$

By Theorem 4.1, $\frac{1}{2}k_t$ epi/hypo-converges to a closed proper saddle function $\frac{1}{2}k$ if and only if f_t epi-converges to f, or equivalently g_t hypo-converges to g, where f and g denote the functions giving $\frac{1}{2}f$ and $\frac{1}{2}g$ as the convex and concave parents of $\frac{1}{2}k$. To obtain our result, except for (5.20) and (5.21), we simply have to verify that under these circumstances one must have

$$(5.25) \quad K(x, y) - y \cdot u = F(x, u) \quad \text{and} \quad K(x, y) - x \cdot v = G(v, y),$$

so that

$$(5.26) \quad f_t(\xi, \omega) = [F(x + t\xi, u + t\omega) - F(x, u) - t(\xi, \omega) \cdot (v, -y)]/\tfrac{1}{2}t^2,$$

(5.27) $g_t(\zeta, \eta) = [G(v + t\zeta, \, y + t\eta) - G(v, \, y) - t(\zeta, \eta) \cdot (-x, \, v)]/\tfrac{1}{2}t^2.$

In particular it will follow that $(v, \, u) \in \partial K(x, \, y)$, since the equations in (5.23) are equivalent respectively by the parenting formulas (4.10) and (4.11) to $u \in \partial_2 K(x, \, y)$ and $v \in \partial_1 K(x, \, y)$. The epi-limit of f_t and hypo-limit of g_t, if any, will have to be $F''_{x, u; v, -y}$ and $G''_{v, y; -x, u}$.

Suppose f_t epi-converges to a function $f \not\equiv \infty$. Then in particular there will exist sequences $t_i \downarrow 0$, ξ_i converging to some ξ, and ω_i converging to some ω, and a value $\rho > 0$, such that

$$f_{t_i}(\xi_i, \, \omega_i) \leq \rho \quad \text{for all } i.$$

From the definition (5.23) of f_t we get

$$F(x + t_i\xi_i, \, u + t_i\omega_i) \leq (K(x, \, y) - y \cdot u) + t_i(\xi_i v - \omega_i y) + \tfrac{1}{2}t_i^2 \rho,$$

and by taking the "lim inf" on both sides as $i \to \infty$ we obtain

$$F(x, \, u) \leq K(x, \, y) - y \cdot u.$$

The opposite inequality always holds by (4.10), so the first equation in (5.25) has been established as valid. The second equation may be proved in a dual manner, starting from the assumption that g_t hypo-converges to a function $g \not\equiv -\infty$.

We are left with verifying (5.20) and (5.21). These follow from the partial conjugacy between $\tfrac{1}{2}K''_{x, y; v, u}$ and its convex parent $F''_{x, u; v, -y}$. The latter is known from Theorem 2.2 to be positively homogeneous of degree two and to have minimum value 0 at $(0, 0)$.   □

Theorem 5.3. *The following conditions are equivalent:*

    (a) K *is twice epi/hypo-differentiable at* $(x, \, y)$ *relative to* $(v, \, u)$;

    (b) K^* *is twice epi/hypo-differentiable at* $(v, \, u)$ *relative to* $(x, \, y)$.

When these conditions hold, the saddle functions $\tfrac{1}{2}K''_{x, y; v, u}$ *and* $\tfrac{1}{2}K^{*''}_{v, u; x, y}$ *are conjugate to each other.*

Proof. This follows at once from Theorem 5.2 and the fact that the convex and concave parents of K^* are given by $-G$ and $-F$, respectively. That was shown in the proof of Theorem 4.2.   □

Theorem 5.4. *The following conditions are equivalent to each other:*

    (a) ∂K *is proto-differentiable at* $(x, \, y)$ *relative to* $(v, \, u)$, *and* $(\zeta, \, \omega) \in (\partial K)'_{x, y; v, u}(\xi, \, \eta)$;

    (b) ∂F *is proto-differentiable at* $(x, \, u)$ *relative to* $(v, \, -y)$, *and* $(\zeta, \, -\eta) \in (\partial F)'_{x, u; v, -y}(\xi, \, \omega)$;

    (c) ∂G *is proto-differentiable at* $(v, \, y)$ *relative to* $(-x, \, u)$, *and* $(-\xi, \, \omega) \in (\partial G)'_{v, y; -x, u}(\zeta, \, \eta)$;

    (d) ∂K^* *is proto-differentiable at* $(v, \, u)$ *relative to* $(x, \, y)$, *and* $(\xi, \, \eta) \in (\partial K^*)'_{v, u; x, y}(\zeta, \, \omega)$.

Proof. The equivalences (4.9) and (4.16) tell all. They give us

$$(\zeta, \omega) \in [\partial K(x + t\xi, \ y + t\eta) - (v, \ u)]/t$$
$$\Leftrightarrow (\zeta, -\eta) \in [\partial F(x + t\xi, \ u + t\omega) - (v, \ -y)]/t$$
$$\Leftrightarrow (-\xi, \ \omega) \in [\partial G(v + t\zeta, \ y + t\eta) - (-x, \ u)]/t$$
$$\Leftrightarrow (\xi, \ \eta) \in [\partial K^*(v + t\zeta, \ u + t\omega) - (x, \ y)]/t,$$

and everything then follows. $\square$

Theorem 5.5. *The following conditions are equivalent and result in the formula* $\partial(\frac{1}{2}K''_{x,y;v,u}) = (\partial K)'_{x,y;v,n}$:

(a) K *is twice epi/hypo-differentiable at* (x, y) *relative to* (v, u) ;

(b) ∂K *is proto-differentiable at* (x, y) *relative to* (v, u) .

Proof. Aplying Theorem 2.2 to F, we get the equivalence of the proto-differentiability in Theorem 5.4(b) with the twice epi-differentiability in Theorem 5.2(b). We also get

$$(5.28) \qquad\qquad \partial(\tfrac{1}{2}F''_{x,u;v,-y}) = (\partial F)'_{x,u;v,-y}.$$

Because $\frac{1}{2}F''_{x,u;v,-y}$ is the convex parent of $\frac{1}{2}K''_{x,y;v,u}$ by Theorem 5.2, we have also by the general relations (4.16), specialized to these second-derivative functions, that

$$(5.29) \quad (\zeta, \omega) \in \partial(\tfrac{1}{2}K''_{x,y;v,u})(\xi, \eta) \leftrightarrow (\zeta, -\eta) \in \partial(\tfrac{1}{2}F''_{x,u;v,-y})(\xi, \omega).$$

The combination of (5.28) and (5.29) with the equivalence between the derivative relations in conditions (a) and (b) of Theorem 5.4 yields our conclusion. $\square$

REFERENCES

1. C. Lemarechal and E. Nurminski, *Sur la différentiabilité de la fonction d'appui du sousdifférential approaché*, C.R. Acad. Sci Paris **290** (1980), 855–858.

2. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

3. A. Auslender, *Differential properties of the support function of the ε-subdifferential of a convex function*, Math. Programming **24** (1982), 257–268 (also C.R. Acad. Sci. Paris **292** (1981), 221–224).

4. J.-B. Hiriart-Urruty, *Approximating a second-order directional derivative for nonsmooth convex functions*, SIAM J. Control. Optim. **20** (1982), 381–404.

5. ____, *Limiting behavior of the approximate first and second-order directional derivatives for a convex function*, Nonlinear Anal. Theory Methods Appl. **6** (1982), 1309–1326.

6. ____, *The approximate first-order and second-order directional derivatives of a marginal function in convex optimization*, J. Optim. Theory Appl. **48** (1986), 127–140.

7. ____, *The approximate first-order and second-order directional derivatives for a convex function*, Mathematical Theories of Optimization, Lecture Notes in Math., vol. 979, Springer-Verlag, 1983, pp. 144–177.

8. ____, *Calculus rules on the approximate second-order directional derivative of a convex function*, SIAM J. Control Optim. **22** (1984), 381–404.

9. J.-B. Hiriart-Urruty and A. Seeger, *Calculus rules on a new set-valued second-order derivative for convex functions*, Nonlinear Anal. Theory Methods Appl. **13** (1989), 721–738.

10. J.-B. Hiriart-Urruty, J.-J. Strodiot and V. Hien Nguyen, *Generalized Hessian matrix and second-order optimality conditions for problems with* $\mathscr{C}^{1,1}$ *data,* Appl. Math. Optim. **11** (1984), 43–56.

11. F. H. Clarke, *Generalized gradients and applications,* Trans. Amer. Math. Soc. **205** (1975), 247–262.

12. ____, *Nonsmooth analysis and optimization,* Wiley, 1983.

13. J. P. Aubin, *Lipschitz behavior of solutions to convex minimization problems,* Math. Oper. Res. **9** (1984), 87–111.

14. J. P. Aubin and I. Ekeland, *Applied nonlinear analysis,* Wiley, 1984.

15. R. T. Rockafellar, *Maximal monotone relations and the second derivatives of nonsmooth functions,* Ann. Inst. H. Poincaré. Anal. Non Linéaire **2** (1985), 167–184.

16. H. Attouch, *Convergence de fonctions convexes, des sous-différentiels, et semi-groupes associés,* C.R. Acad. Sci. Paris **284** (1977), 539–542.

17. ____, *Variational convergence for functions and operators,* Pitman, London, 1984.

18. A. D. Alexandroff, *The existence almost everywhere of the second differential of a convex function and some associated properties of convex surfaces,* Uchen. Zap. Gos. Univ. Ser. Mat. **37** (1939), 3–35. (Russian)

19. R. M. Dudley, *On second derivatives of convex functions,* Math. Scand. **41** (1977), 159–174, and **46** (1980), 61.

20. J. L. Ndoutoume, *Épi-convergence en calcul différentiel généralisé: Résultats de convergence et approximation,* Thèse, Université de Perpignan, 1987.

21. A. Auslender, *Stability in mathematical programming with nondifferentiable data,* SIAM J. Control. Optim. **22** (1984), 239–254.

22. R. W. Chaney, *On sufficient conditions in nonsmooth optimization,* Math. Oper. Res. **7** (1982), 463–475.

23. ____, *Second-order sufficient conditions for nondifferentiable progrmming problems,* SIAM J. Control Optim. **20** (1982), 20–33.

24. ____, *A general sufficiency theorem for nonsmooth nonlinear programming,* Trans. Amer. Math. Soc. **276** (1983), 235–245.

25. ____, *On second-order directional derivatives for nonsmooth functions,* J. Nonlinear Anal. Theory Math. Appl. (to appear).

26. ____, *Second-order necessary conditions in semismooth optimization,* SIAM J. Control. Optim. **25** (1987), 1072–1081.

27. ____, *Second-order sufficient conditions in nonsmooth optimization,* Math. Oper. Res. **13** (1988), 660–673.

28. A. Ben-Tal, *Second order theory for extremum problems,* System Analysis and External Methods (A.V. Fiacco and K. Kostaneta, eds.), Lecture Notes in Economics and Mathematical Sciences, Springer-Verlag, 1980, pp. 336-356.

29. A. Ben-Tal and J. Zowe, *A unified theory of first and second-order conditions for extremum problems in topological vector spaces,* Math. Programming Studies **19** (1982), 39–76.

30. ____, *Directional derivatives in nonsmooth optimization,* J. Optim. Theory Appl. **47** (1985), 483–490.

31. A. Seeger, *Second order directional derivatives in parametric optimization problems,* Math. Oper. Res. **13** (1988), 124–139.

32. A. Shapiro, *Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton,* Appl. Math. Optim. **18** (1988), 215–229.

33. ____, *Second-order derivatives of extremal-value functions and optimality conditions for semi-infinite programs,* Math. Oper. Res. **10** (1985), 207–219.

34. ____, *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections,* SIAM J. Control Optim. **26** (1988), 628–645.

35. R. Cominetti and R. Correa, *A generalized second order derivative in nonsmooth optimization*, SIAM J. Control Optim. (to appear).

36. R. T. Rockafellar, *First and second-order epi-differentiability in nonlinear programming*, Trans. Amer. Math. Soc. **307** (1988), 75–108.

37. _____, *Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives*, Math. Oper. Res. **14** (1989), 462–484.

38. H. Attouch, D. Azé, and R. J-B Wets, *Convergence of convex-concave saddle functions: continuity properties of the Legendre-Fenchel transform and applications to convex programming and mechanics*, Ann. Inst. H. Poincaré Anal. Non Linéaire **5** (1988), 537–572.

39. G. Salinetti and R. J.-B. Wets, *On the convergence of sequences of convex sets in finite dimensions*, SIAM Rev. **21** (1979), 18–33.

40. R. T. Rockafellar, *Generalized directional derivatives and subgradients of nonconvex functions*, Canad. J. Math. **32** (1980), 157–180.

41. R. A. Wijsman, *Convergence of sequencs of convex sets, cones and functions*, I, Bull. Amer. Math. Soc. **70** (1964), 186–188.

42. _____, *Convergence of sequences of convex sets, cones and functions*, II, Trans. Amer. Math. Soc. **123** (1966), 32–45.

43. U. Mosco, *Convergence of convex sets and solutions to variational inequalities*, Adv. in Math. **3** (1969), 510–585.

44. R. T. Rockafellar, *Proto-differentiability of set-valued mappings and its applications in optimization*, Ann. Inst. H. Poincaré Anal. Non Linéaire **6** (1989), 449–482.

45. R. J-B Wets, *Convergence of convex functions, variational inequalities and convex optimization problems*, Variational Inequalities and Complementarity Problems (R. Cottle, F. Gianessi, and J.-L. Lions, eds.), Wiley, 1980, pp. 375–403.

46. R. T. Rockafellar, *Minimax theorems and conjugate saddle functions*, Math. Scand. **14** (1964), 151–173.

47. H. Attouch and R. J.-B. Wets, *A convergence theory for saddle functions*, Trans. Amer. Math. Soc. **280** (1983), 1–41.

48. R. Cominetti, *Analyse du second ordre de problèmes d'optimisation non différentiable*, thèse, Université de Clermont-Ferrand, France, 1989.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON 98195