Canad. J. Math. Vol. **51** (2), 1999 pp. 250–265

Convergence of Subdifferentials of Convexly Composite Functions

C. Combari, R. Poliquin and L. Thibault

Abstract. In this paper we establish conditions that guarantee, in the setting of a general Banach space, the Painlevé-Kuratowski convergence of the graphs of the subdifferentials of convexly composite functions. We also provide applications to the convergence of multipliers of families of constrained optimization problems and to the generalized second-order derivability of convexly composite functions.

Introduction

Attouch's Theorem states that a sequence of proper lower semicontinuous convex functions on a reflexive Banach space, Mosco epi-converges if and only if the graphs of the subdifferentials Painlevé-Kuratowski converge to the graph of the subdifferential of the limit function and a condition that fixes the constant of integration holds (*cf.* [1]). This theorem, with "slice" convergence in place of Mosco epi-convergence, was extended by Attouch and Beer [2] to general Banach spaces. While this result completely characterizes the Painlevé-Kuratowski convergence of the subdifferentials of convex functions, much less is known in the nonconvex case. Recently however Poliquin [21] in the context of a finite dimensional space gave an extension of Attouch's Theorem for the large class of possibly nonconvex *primal-lower-nice* functions. In addition, Levy, Poliquin and Thibault [17] gave partial extensions in infinite Hilbert spaces. Partial extensions of Attouch's Theorem to nonconvex functions are also provided in [13], [18] and [30]. Other types of convergence (*e.g.*, Attouch-Wets) have also been studied in [3] and [17].

A typical example of a primal lower nice function is the composition of an extended-real-valued convex function f defined on a Banach space Y with a twice continuously differentiable mapping F from a Banach space X into Y satisfying a qualification condition, *cf.* [10]. Although these convexly composite functions are primal-lower-nice on a general Banach space, it appears that the Hilbert structure is essential to derive an Attouch-like theorem for primal-lower-nice functions (it also appears to be essential in obtaining "integration" results and the "proto-differentiability" of subgradient mappings for such functions). In this paper we abandon the primal-lower-nice setting and instead we study directly in the setting of general Banach spaces the Painlevé-Kuratowski convergence of the graphs of the subdifferentials of convexly composite functions satisfying a qualification condition.

Received by the editors March 26, 1998; revised August 5, 1998.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant OGP41983 for the second author, and by NATO under grant CRG 960360 for the third author.

AMS subject classification: Primary: 49A52, 58C06, 58C20; secondary: 90C30.

Keywords: epi-convergence, Mosco convergence, Painlevé-Kuratowski convergence, primal-lower-nice functions, constraint qualification, slice convergence, graph convergence of subdifferentials, convexly composite functions.

A qualification condition is essential when working with convexly composite functions. This is easily understood because on Banach spaces that admit a locally uniformly rotund norm (which includes many well known spaces) every lower semicontinuous function can be written locally as the composition of a lower semicontinuous convex function and a smooth mapping (*cf.* [22]). We therefore cannot expect that in general the convergence of a sequence of arbitrary lower semicontinuous functions will be linked in any way to the convergence of the sequence of subgradients. We begin in Section 1 by reviewing the well-known Robinson qualification condition. We then give an alternate form of this condition which will be the basis of our qualification condition and of our uniform qualification condition.

This qualification condition serves many useful purposes. On the one hand, as in the convex case (*cf.* [6]), we may speak of the subdifferential of a convexly composite function since they are all the same under the qualification condition; see [10]. On the other hand the qualification condition guarantees (*cf.* [10]) that $\partial(f\circ F)(x) = \nabla^* F(x)\partial f\big(F(x)\big)$ where $\partial(f\circ F)(x)$ is the subdifferential of the convexly composite function $(f\circ F)$ at the point x, $\nabla^* F(x)$ is the adjoint of $\nabla F(x)$ and $\partial f\big(F(x)\big)$ is the subdifferential of f at $F(x)$. The qualification condition is also used in Proposition 1.1 to give a precise estimate for the norm of the vectors $y^* \in \partial f\big(F(x)\big)$ in terms of the norms of x^* where $x^* = \nabla^* F(x) y^*$; this is crucial in obtaining our main result. In the rest of this first section we review set-convergence, epi-convergence (we state a result of Combari and Thibault [9] on the epi-convergence of a sequence of convexly composite functions) and slice convergence.

We prove in the second section the main result of this paper. Let $\{(f_n\circ F_n) : n \in \mathrm{N}\} \cup \{(f\circ F)\}$ be a sequence of convexly composite functions that is *uniformly qualified* over the open set Ω (see Section 1). If f_n slice converges to f, F_n pointwise converges to F over Ω, and if for each $x\in\Omega$ we have

$$\lim_{\substack{n\to\infty\\ u\to x}} \nabla F_n(u) = \nabla F(x),$$

then $(f_n\circ F_n)$ strongly epi-converges to $(f\circ F)$ over Ω. Moreover the graph of the subdifferential of $(f_n\circ F_n)$ Painlevé-Kuratowski converges to the graph of the subdifferential of $(f\circ F)$ over $\Omega\cap \operatorname{dom}(f\circ F)$. Even in the context of a finite dimensional space this result is new because the continuity assumptions on ∇F_n and ∇F are not sufficient to ensure that the composite functions are primal lower nice. In addition, because we are dealing with convex functions that are not necessarily finite on the whole space, the results in [18], [30] are not applicable.

In Section 3 we give applications to the convergence of multipliers. Finally in Section 4 we give applications to the *proto-differentiation* of subgradient mappings of convexly composite functions and to *second-order epi-derivatives*.

1 Preliminaries

Throughout this paper X and Y will be two real *Banach spaces*. For any function $g\colon X\to \mathrm{R}\cup\{+\infty\}$ the Clarke subdifferential of g at x is denoted by $\partial g(x)$ (see [6] for a broad discussion of this subdifferential) and the effective domain of g by $\operatorname{dom} g := \{x\in X : g(x) < +\infty\}$.

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

We say that $f \circ F$ is *convexly composite* over an open subset Ω of X if $f\colon Y \to \mathrm{R} \cup \{+\infty\}$ is a proper lower semicontinuous convex function and $F\colon X \to Y$ is continuously differentiable over Ω. We will say that the convexly composite function $f \circ F$ is *qualified* at $x \in \Omega \cap \operatorname{dom}(f \circ F)$ if the Robinson qualification condition holds at x, *i.e.*,

$$\mathrm{R}_+\big(\operatorname{dom} f - F(x)\big) - \nabla F(x)(X) = Y. \tag{1.1}$$

Here $\nabla F(x)$ denotes the (Fréchet) derivative of F at x. In the finite dimensional setting (1.1) is equivalent to the following condition used in [25] and [26]

$$N\big(\operatorname{dom} f; F(x)\big) \cap \operatorname{Ker}\big(\nabla^* F(x)\big) = \{0\},$$

where $\nabla^* F(x)$ denotes the adjoint operator of $\nabla F(x)$ and $N(\operatorname{dom} f;\ )$ the normal cone (in the sense of convex analysis) to $\operatorname{dom} f$. It is important to realize that (1.1) is actually a local condition in the sense that when it is satisfied at x it also holds for all $x' \in \operatorname{dom}(f \circ F)$ near x, *cf.* [10].

A powerful consequence of (1.1), which was established in [10], is that under (1.1) all types of subdifferentials (see [11], [28], [29]) for the convexly composite function $f \circ F$ coincide with the Clarke subdifferential. Moreover,

$$\partial(f \circ F)(x) = \nabla^* F(x) \partial f\big(F(x)\big). \tag{1.2}$$

Here the set on the right side of the equality is given by $\{\nabla^* F(x) y^* : y^* \in \partial f\big(F(x)\big)\}$. Recall that for the convex function f, the Clarke subdifferential coincides with the subdifferential in the sense of convex analysis, *i.e.*, $\partial f(y) = \{y^* \in Y^* : \langle y^*, y' - y\rangle + f(y) \le f(y'), \forall y' \in Y\}$. When (1.2) is fulfilled we will say that $f \circ F$ is *subdifferentially qualified* at x. Obviously the convexly composite function $f \circ F$ is subdifferentially qualified at x whenever it is qualified at x. On the global side, we will say that $f \circ F$ is qualified (resp. subdifferentially qualified) over Ω if (1.1) (resp. (1.2)) holds for any $x \in \Omega \cap \operatorname{dom}(f \circ F)$.

In this paper we need to develop a uniform qualification condition for the sequence of convexly composite functions $\{(f_n \circ F_n) : n \in \mathrm{N}\} \cup \{f \circ F\}$. It is impossible to develop such a condition based on (1.1). Fortunately in [10] a condition equivalent to (1.1) is given which can easily be made uniform. Indeed it is proven in [10] that condition (1.1) holds if and only if there exist $r > 0$, $s > 0$ such that

$$s\mathrm{B}_Y \subset \Big(\{f \le r + f\big(F(x)\big)\} - F(x)\Big) - \nabla F(x)(r\mathrm{B}_X), \tag{1.3}$$

where B_X denotes the closed unit ball of X centered at the origin and $\{f \le p\} := \{y \in X : f(y) \le p\}$. We will say that the sequence of convexly composite functions $\{(f_n \circ F_n) : n \in \mathrm{N}\} \cup \{f \circ F\}$ is *uniformly qualified* over Ω if for each $x \in \Omega \cap \operatorname{dom}(f \circ F)$ there exist $s > 0$ and $r > 0$ such that (1.3) holds and such that there exists $N \in \mathrm{N}$ for which one has

$$s\mathrm{B}_Y \subset \Big(\{f_n \le r + f\big(F(x)\big)\} - F(x)\Big) - \nabla F(x)(r\mathrm{B}_X) \tag{1.4}$$

for all $n \ge N$. For some other uniform qualification conditions, obtained via the Attouch-Wets convergence, we refer the reader to Guillaume [14], [15].

The following result will be needed in the sequel.

Proposition 1.1 *For $r > 0$, and $s > 0$, assume that (1.3) is satisfied at $x \in \operatorname{dom}(f \circ F)$ for the convexly composite function $f \circ F$. Then for every $x^* = \nabla^* F(x) y^*$ with $y^* \in \partial f\big(F(x)\big)$ one has*

$$s\|y^*\| \le r(1 + \|x^*\|).$$

Proof Consider any $y \in \mathrm{B}_Y$. By (1.3) there exist $b \in \mathrm{B}_X$ and $y' \in Y$ with $f(y') \le r + f\big(F(x)\big)$ such that $sy = y' - F(x) + \nabla F(x)(rb)$. Then

$$\begin{aligned}\langle y^*, sy\rangle &= \langle y^*, y' - F(x)\rangle + \langle y^*, \nabla F(x)(rb)\rangle\\ &\le f(y') - f\big(F(x)\big) + \langle y^*, \nabla F(x)(rb)\rangle\\ &\le r + \langle \nabla^* F(x) y^*, rb\rangle\\ &\le r + r\|x^*\|\end{aligned}$$

and hence $s\|y^*\| \le r(1 + \|x^*\|)$. ■

In the remainder of this section we review the fundamental concepts of epi-convergence, Mosco epi-convergence, Painlevé-Kuratowski convergence, and slice convergence. Among other things, this will enable us to introduce the notation we will need in the next sections. In addition we review in Theorem 1.2, the relationship between slice convergence of convex functions and Painlevé-Kuratowski convergence of the graphs of the subdifferentials. Finally in Theorem 1.3 we give a condition that is sufficient for the epi-convergence of convexly composite functions.

First recall that Attouch's Theorem [1] gives on a reflexive Banach space the equivalence between the Painlevé-Kuratowski convergence of the subdifferentials of convex functions and the Mosco epi-convergence of these functions. A sequence of functions f_n from Y into $\mathrm{R} \cup \{+\infty\}$ Mosco epi-converges to f if for any $y \in Y$, $f(y) \le \liminf_{n\to\infty} f_n(y_n)$ for any y_n weakly convergent to y and there exists y_n strongly converging to y with $\lim_{n\to\infty} f_n(y_n) = f(y)$. This can also be stated in terms of epi-limits (or Γ-limits) as follows. With respect to a topology τ on Y, the (sequential) epi-limits inferior and superior are defined by

$$(Li_e f_n)(y) := \inf\{\liminf_{n\to\infty} f_n(y_n) : y_n \xrightarrow{\tau} y\}$$

and

$$(Ls_e f_n)(y) := \inf\{\limsup_{n\to\infty} f_n(y_n) : y_n \xrightarrow{\tau} y\}.$$

One says that the sequence $(f_n)_n$ epi-converges (or Γ-converges) to f (see [1], [12]) with respect to τ if $Ls_e f_n = f = Li_e f_n$. Therefore, f_n Mosco converges to f if and only if it epi-converges to f both in the strong and the weak topology of Y.

Next we recall the definition of the Painlevé-Kuratowski convergence of sets. For a sequence C_n of subsets of Y recall that $\liminf_{n\to\infty} C_n = \{y \in Y : y = \lim_{n\to\infty} y_n$ with $y_n \in C_n\}$ and $\limsup_{n\to\infty} C_n = \{y \in Y : y = \lim_{n\to\infty} y_{s(n)} y_n \in C_n\}$. One says that (C_n)

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

Painlevé-Kuratowski converges to C and one writes $C_n \to^{pk} C$ if $\limsup_{n\to\infty} C_n = C = \liminf_{n\to\infty} C_n$. Note that in terms of distance function

$$\liminf_{n\to\infty} C_n = \{y \in Y : \lim_{n\to\infty} d(y, C_n) = 0\}. \tag{1.5}$$

Let (M_n) be a sequence of set-valued mappings from X into X^* and S be a nonempty subset of X. Recall that the graph of a set-valued mapping M is defined by $\operatorname{gph} M := \{(x, x^*) \in X \times X^* : x^* \in M(x)\}$. The sequential limits inferior and superior with respect to the strong topology on X and the weak-star topology on X^* will be denoted by ${}^*\liminf_{n\to\infty} \operatorname{gph} M_n$ and ${}^*\limsup_{n\to\infty} \operatorname{gph} M_n$. We will also write

$$\operatorname{gph} M \subset \liminf_{n\to\infty} \operatorname{gph} M_n \text{ over } S \quad (\text{resp.} \limsup_{n\to\infty} \operatorname{gph} M_n \subset \operatorname{gph} M \text{ over } S)$$

when any $(x, x^*) \in \operatorname{gph} M$ with $x \in S$ belongs to $\liminf_{n\to\infty} \operatorname{gph} M_n$ (resp. when any $(x, x^*) \in \limsup_{n\to\infty} \operatorname{gph} M_n$ with $x \in S$ belongs to $\operatorname{gph} M$). When both inclusions hold, we will say that the graphs of M_n Painlevé-Kuratowski converge to the graph of M over S with respect to the norm $\times$ norm topology of $X \times X^*$ and we will write $\operatorname{gph} M_n \to^{pk} \operatorname{gph} M$ over S. In the same way one defines the Painlevé-Kuratowski convergence with respect to the norm $\times$ weak star topology of $X \times X^*$ of the graphs of M_n to the graph of M over S and we write $\operatorname{gph} M_n \to^{*pk} \operatorname{gph} M$ over S.

For general Banach spaces, the Painlevé-Kuratowski convergence of the graphs of subdifferentials of convex functions is equivalent to the slice convergence of these functions. Before stating this result, recall that a sequence C_n of closed convex subsets of Y slice converges to a closed convex subset C if for each closed and bounded convex subset B of Y one has (see Beer [5])

$$D(B, C) = \lim_{n\to\infty} D(B, C_n),$$

where $D(B, C) = \inf\{\|b - c\| : b \in B \text{ and } c \in C\}$. The slice convergence of a sequence of convex functions f_n to f is defined as the slice convergence of $\operatorname{epi} f_n$ to $\operatorname{epi} f$. Here $\operatorname{epi} f = \{(y, \alpha) \in Y \times \mathrm{R} : \alpha \geq f(y)\}$. Beer [5] showed that the slice convergence of a sequence of convex functions f_n to f is equivalent to the Mosco epi-convergence of f_n and f_n^* to f and f^* respectively. Here f^* denotes the Fenchel conjugate of f, that is $f^*(y^*) = \sup\{\langle y^*, y\rangle - f(y) : y \in Y\}$. As in Beer [4], [5] let

$$\Delta f := \{\big(y, y^*, f(y)\big) \in Y \times Y^* \times \mathrm{R} : y^* \in \partial f(y)\}.$$

***Theorem 1.2* (Attouch-Beer [2] and Beer[4])** *Let $\{f_n : n \in \mathrm{N}\} \cup \{f\}$ be a sequence of proper lower semicontinuous convex functions from Y into $\mathrm{R} \cup \{+\infty\}$. The following assertions are equivalent:*

(i) *f_n slice converges to f;*
(ii) $\operatorname{gph} \partial f_n \to^{pk} \operatorname{gph} \partial f$ *and there exist* $(a, a^*) \in \operatorname{gph} \partial f$ *and a sequence* $(a_n, a_n^*) \in \operatorname{gph} \partial f_n$ *such that* $\big(a_n, a_n^*, f_n(a_n)\big) \to \big(a, a^*, f(a)\big)$*;*
(iii) $\Delta f_n \to^{pk} \Delta f$.

In general Banach spaces the equivalence between (i) and (iii) was first established by Beer [4]. Then Attouch and Beer [2] proved the equivalence between (i) and (ii). Note that a unified and simpler proof of the equivalence between the three assertions has been given recently by Combari and Thibault [8]. All these results are extensions of Attouch's Theorem in reflexive Banach spaces.

We will also need the following theorem which gives a sufficient condition for the epi-convergence of a sequence of convexly composite functions.

Theorem 1.3 **(Combari and Thibault [9])** *Let $\{F_n : n \in \mathrm{N}\} \cup \{F\}$ be a sequence of mappings from X to Y that are C^1 near x and let f_n be a sequence of proper lower semicontinuous convex functions from Y into $\mathrm{R} \cup \{+\infty\}$ that epi-converges to f with respect to the strong topology. Assume that $f\big(F(x)\big)$ is finite, F_n pointwise converges to F around x, and that*

$$\lim_{\substack{n\to\infty\\ u\to x}} \nabla F_n(u) = \nabla F(x).$$

Finally, assume that there exist some $r > 0$ and $s > 0$ such that for all n

$$s\mathrm{B}_Y \subset \Big(\{f_n \le r + f\big(F(x)\big)\} - F(x)\Big) - \nabla F(x)(r\mathrm{B}_X).$$

Then $(f_n \circ F_n)$ strongly epi-converges to $(f \circ F)$ on an open neighborhood of x.

2 Main Result

Our main theorem will be a consequence of the two following propositions. Throughout this section, Ω denotes an open *convex* subset of X.

Proposition 2.1 *Let $\{(f_n \circ F_n) : n \in \mathrm{N}\} \cup \{(f \circ F)\}$ be a sequence of convexly composite functions with $(f \circ F)$ subdifferentially qualified over Ω. Assume that f_n slice converges to f, F_n pointwise converges to F over Ω and for each $x \in \Omega$*

$$\lim_{\substack{n\to\infty\\ u\to x}} \nabla F_n(u) = \nabla F(x).$$

Assume further that for each $x \in \Omega$ there exists a sequence x_n converging to x such that $\big(f_n \circ F_n\big)(x_n)$ converges to $(f \circ F)(x)$. Then

$$\operatorname{gph} \partial(f \circ F) \subset \liminf_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n) \quad \textit{over } \Omega.$$

Proof Let $(x, x^*) \in \operatorname{gph} \partial(f \circ F)$ with $x \in \Omega$ and let $\varepsilon > 0$. Note that $x \in \operatorname{dom}(f \circ F)$. Since $(f \circ F)$ is subdifferentially qualified over Ω, there exists $y^* \in \partial f\big(F(x)\big)$ such that $x^* = \nabla^* F(x) y^*$. By (iii) in Theorem 1.2 there exist $(y_n, y_n^*) \in \operatorname{gph} \partial f_n$ such that $(y_n, y_n^*) \to \big(F(x), y^*\big)$ and $f_n(y_n) \to f\big(F(x)\big)$. Fix some real positive number $r < \varepsilon$ and some integer N_1 such that

$$2\|y_n^*\|\,\|\nabla F_n(x') - \nabla F(x)\| \le \varepsilon/6 \tag{2.1}$$

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

for all $n \geq N_1$ and $x' \in x + 2r\mathrm{B}$ (this is possible under our assumptions). According to our assumptions there exists a sequence x_n converging to x with $\big(f_n \circ F_n\big)(x_n)$ converging to $(f \circ F)(x)$. It is easily seen by the Inequality Mean Value Theorem that $F_n(x_n) \to F(x)$. So we can find an integer $N \geq N_1$ such that for every $n \geq N$ one has $\|x_n - x\| < r$ and

$$|f_n(y_n) - f_n\big(F_n(x_n)\big)| + \|y_n^*\| \, \|F_n(x_n) - y_n\| \leq \varepsilon r/2. \tag{2.2}$$

For each $u \in x + 2r\mathrm{B}$ we have (by definition of ∂f_n)

$$\langle y_n^*, F_n(u) - y_n\rangle \leq f_n\big(F_n(u)\big) - f_n(y_n).$$

Hence

$$\langle y_n^*, F_n(u) - F_n(x_n)\rangle \leq (f_n \circ F_n)(u) - f_n(y_n) + \|y_n^*\| \, \|F_n(x_n) - y_n\|.$$

This ensures that there exists $x_n' \in \{tu + (1-t)x_n : t \in [0,1]\} \subset x + 2r\mathrm{B}$ such that

$$\begin{aligned}
\langle \nabla^* F_n(x_n) y_n^*, u - x_n\rangle &\leq (f_n \circ F_n)(u) - f_n(y_n) + \|y_n^*\| \, \|F_n(x_n) - y_n\| \\
&\quad + \|y_n^*\| \, \|\nabla F_n(x_n') - \nabla F_n(x_n)\| \, \|u - x_n\| \\
&\leq (f_n \circ F_n)(u) - (f_n \circ F_n)(x_n) + |f_n(y_n) - f_n\big(F_n(x_n)\big)| \\
&\quad + \|y_n^*\| \, \|F_n(x_n) - y_n\| \\
&\quad + \|y_n^*\| \, \|\nabla F_n(x_n') - \nabla F_n(x_n)\| \, \|u - x_n\|.
\end{aligned} \tag{2.3}$$

For each $n \geq N$ and each $u \in x + 2r\mathrm{B}$ we have by (2.1)

$$\begin{aligned}
\|y_n^*\| \, \|\nabla F_n(x_n') - \nabla F_n(x_n)\| &\leq \|y_n^*\| \big(\|\nabla F_n(x_n') - \nabla F(x)\| + \|\nabla F(x) - \nabla F_n(x_n)\|\big) \\
&\leq \varepsilon/6.
\end{aligned}$$

Let $x_n^* := \nabla^* F_n(x_n) y_n^*$. Therefore, according to (2.2) and (2.3) we obtain for each $n \geq N$

$$\langle x_n^*, u - x_n\rangle \leq (f_n \circ F_n)(u) - (f_n \circ F_n)(x_n) + \varepsilon r \quad \text{for all } u \in x + 2r\mathrm{B}.$$

Applying (for each $n \geq N$) the Ekeland variational principle to the function $(f_n \circ F_n) - \langle x_n^*, \cdot\rangle$ we find u_n with $\|u_n - x_n\| \leq r$ and such that

$$(f_n \circ F_n)(u_n) - \langle x_n^*, u_n\rangle \leq (f_n \circ F_n)(u) - \langle x_n^*, u\rangle + \varepsilon\|u - u_n\|$$

for all $u \in x + 2r\mathrm{B}$. As $\|u_n - x\| < 2r$, the point u_n is an unconstrained local minimum of the function $u \mapsto (f_n \circ F_n)(u) - \langle x_n^*, u\rangle + \varepsilon\|u - u_n\|$. Therefore 0 is in the Clarke subdifferential of this function at the point u_n. Hence for each $n \geq N$ (by the subdifferential calculus rules for the Clarke subdifferential; see [6]) $x_n^* \in \partial(f_n \circ F_n)(u_n) + \varepsilon\mathrm{B}$. Hence there exists u_n^* such that $(u_n, u_n^*) \in \operatorname{gph} \partial(f_n \circ F_n)$ and $\|u_n^* - x_n^*\| \leq \varepsilon$. Since $x_n^* = \nabla^* F_n(x_n) y_n^* \to x^*$, we may suppose that $\|x_n^* - x^*\| \leq \varepsilon$ for all $n \geq N$. Thus for each $n \geq N$ we have

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

$(u_n, u_n^*) \in \operatorname{gph} \partial(f_n \circ F_n)$ and $\|(x, x^*) - (u_n, u_n^*)\| \le 2\varepsilon$. According to (1.5) we conclude that $(x, x^*) \in \liminf_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n)$. ■

Remark It is easily shown that the proof of the above proposition is also valid when a presubdifferential in the sense of [11], [28] (contained in the Clarke subdifferential) is used in place of the Clarke subdifferential. ■

In the proof of the next proposition, we employ Fréchet subgradients. For a function $g\colon X \to \mathrm{R} \cup \{+\infty\}$ and $x \in \operatorname{dom} g$ the Fréchet subdifferential of g at x is defined by

$$\partial_{\mathrm{Fre}} g(x) := \{x^* \in X : \liminf_{u\to x} \|u - x\|^{-1}\big(g(u) - g(x) - \langle x^*, u - x\rangle\big) \ge 0\}.$$

Recall from [10] that under (1.1) the Fréchet subdifferential $\partial_{\mathrm{Fre}}(f \circ F)(x)$ coincides with the Clarke subdifferential $\partial(f \circ F)(x)$.

Proposition 2.2 *Let $\{(f_n \circ F_n) : n \in \mathrm{N}\} \cup \{(f \circ F)\}$ be a sequence of convexly composite functions that is uniformly qualified over Ω. Assume that $(f \circ F) = Li_e(f_n \circ F_n)$ or that $f = Li_e f_n$ with respect to the strong topology over Ω. In addition assume that F_n pointwise converges to F over Ω and that for each $x \in \Omega$*

$$\lim_{\substack{n\to\infty \\ u\to x}} \nabla F_n(u) = \nabla F(x).$$

Then over $\Omega \cap \operatorname{dom}(f \circ F)$

$$^*\liminf_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n) \subset \operatorname{gph} \partial(f \circ F).$$

If instead of assuming that $(f \circ F) = Li_e(f_n \circ F_n)$, one assumes that $f_n \circ F_n$ epi-converges to $f \circ F$ with respect to the strong topology over Ω, then one has

$$^*\limsup_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n) \subset \operatorname{gph} \partial(f \circ F).$$

Proof Let $(x, x^*) \in {}^*\liminf_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n)$ with $x \in \Omega \cap \operatorname{dom}(f \circ F)$. By the comment preceding the statement of this proposition, it will be enough to show that $x^* \in \partial_{\mathrm{Fre}}(f \circ F)(x)$.

Fix positive real numbers r and s as given by (1.4) and take a sequence (x_n, x_n^*) with $x_n^* \in \partial(f_n \circ F_n)(x_n)$, x_n converging to x and x_n^* weak-star converging to x^*. It is easily seen (using our assumptions) that $F_n(x_n)$ converges to $F(x)$. Since $(f \circ F) = Li_e(f_n \circ F_n)$ or that $f = Li_e f_n$ with respect to the strong topology over Ω one has

$$f\big(F(x)\big) \le \liminf_{n\to\infty} f_n\big(F_n(x_n)\big).$$

As $(f \circ F)(x)$ is finite, there exists an integer N_1 such that $f\big(F(x)\big) \le f_n\big(F_n(x_n)\big) + r$ and $f_n\big(F_n(x_n)\big) < +\infty$ for all $n \ge N_1$. Thus by (1.4) and Lemma 2 in [24] there exists an integer $N \ge N_1$ such that for all $n \ge N$

$$\frac{1}{2}s\mathrm{B} \subset \Big(\{f_n \le 2r + f_n\big(F_n(x_n)\big)\} - F_n(x_n)\Big) - \nabla F_n(x_n)(2r\mathrm{B}) \tag{2.4}$$

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

and hence $\partial(f_n \circ F_n)(x_n) = \nabla^* F_n(x_n)\partial f_n\big(F_n(x_n)\big)$. Without loss of generality we may suppose that this holds for all $n \in \mathrm{N}$. Choose $y_n^* \in \partial f_n\big(F_n(x_n)\big)$ such that $x_n^* = \nabla^* F_n(x_n) y_n^*$. By (2.4) and Proposition 1.1 one has $s\|y_n^*\| \le 4r(1 + \|x_n^*\|)$ and hence the sequence y_n^* is bounded.

Consider any real number $\varepsilon > 0$, and choose (according to the assumptions of the proposition) $\rho > 0$ such that $\|y_n^*\| \, \|\nabla F_n(x') - \nabla F_n(x'')\| \le \epsilon$ for all $x', x'' \in x + 2\rho \mathrm{B}$ and $n \in \mathrm{N}$. For any $u \in x + \rho\mathrm{B}$ and $y_n \to F(u)$ we have (because $y_n^* \in \partial f_n\big(F_n(x_n)\big)$)

$$\langle y_n^*, y_n - F_n(u)\rangle + \langle y_n^*, F_n(u) - F_n(x_n)\rangle \le f_n\big(y_n\big) - f_n\big(F_n(x_n)\big)$$

and hence for some $x_n' := t_n u + (1 - t_n)x_n$ with $t_n \in [0, 1]$

$$\begin{aligned} &\langle y_n^*, y_n - F_n(u)\rangle + \langle \nabla^* F_n(x_n) y_n^*, u - x_n\rangle \\ &\quad \le f_n(y_n) - f_n\big(F_n(x_n)\big) \\ &\qquad + \|y_n^*\| \, \|\nabla F_n(x_n') - \nabla F_n(x_n)\| \, \|u - x_n\|. \end{aligned} \tag{2.5}$$

Suppose first that $(f \circ F) = Li_e(f_n \circ F_n)$, and consider any sequence u_n in $x + 2r\mathrm{B}$ converging to u. By (2.5), with $y_n = F_n(u_n)$, we have for all n large enough

$$\langle x_n^*, u_n - x_n\rangle \le (f_n \circ F_n)(u_n) - (f_n \circ F_n)(x_n) + \varepsilon\|u - x_n\|$$

and hence

$$\langle x^*, u - x\rangle \le \liminf_{n\to\infty} (f_n \circ F_n)(u_n) - \liminf_{n\to\infty} (f_n \circ F_n)(x_n) + \varepsilon\|u - x\|.$$

This ensures that

$$\begin{aligned} \langle x^*, u - x\rangle &\le \big(Li_e(f_n \circ F_n)\big)(u) - \liminf_{n\to\infty} (f_n \circ F_n)(x_n) + \varepsilon\|u - x\| \\ &\le \big(Li_e(f_n \circ F_n)\big)(u) - \big(Li_e(f_n \circ F_n)\big)(x) + \varepsilon\|u - x\| \\ &= (f \circ F)(u) - (f \circ F)(x) + \varepsilon\|u - x\|. \end{aligned} \tag{2.6}$$

Suppose now that $f = Li_e f_n$. Taking the limit inferior in (2.5) we obtain (since $F_n(x_n) \to F(x)$)

$$\begin{aligned} \langle x^*, u - x\rangle &\le \liminf_{n\to\infty} f_n(y_n) - \liminf_{n\to\infty} f_n\big(F_n(x_n)\big) + \epsilon\|u - x\| \\ &\le \liminf_{n\to\infty} f_n(y_n) - f\big(F(x)\big) + \epsilon\|u - x\| \end{aligned}$$

and hence taking the infimum over all sequences y_n converging strongly to $F(u)$ we have that

$$\langle x^*, u - x\rangle \le f\big(F(u)\big) - f\big(F(x)\big) + \epsilon\|u - x\|. \tag{2.7}$$

It is easily seen that the above inequalities (2.6) and (2.7) imply that $x^* \in \partial_{\mathrm{Fre}}(f \circ F)(x)$. This completes the proof of the first part of the proposition. The second part of the proposition is a direct consequence of the preceding one because $(x, x^*) \in^* \limsup \operatorname{gph} \partial(f_n \circ F_n)$ means that $(x, x^*) \in^* \liminf \operatorname{gph} \partial(f_k \circ F_k)$ for a subsequence $k \in K \subset \mathrm{N}$ and because epi-convergence is preserved by subsequences. ■

Remark In fact if one just considers the case where f_n epi-converges to f the proof is much simpler. Indeed let $(x, x^*) \in^* \limsup_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n)$ with $x \in \Omega \cap \operatorname{dom}(f \circ F)$. Taking a subsequence, we may suppose that there exists $(x_n, x_n^*) \in \operatorname{gph} \partial(f_n \circ F_n)$ with x_n norm converging to x and x_n^* weak-star converging to x^*. As in the beginning of the proof of the proposition one obtains that the sequence y_n^* is bounded (here $y_n^* \in \partial f_n\big(F_n(x_n)\big)$ and $x_n^* = \nabla^* F_n(x_n) y_n^*$). Fix a weak-star cluster point y^* of this sequence. Then for any $y \in Y$ and any y_n converging to y one has

$$\langle y_n^*, y_n - F_n(x_n)\rangle + f_n\big(F_n(x_n)\big) \le f_n(y_n).$$

Hence by taking the limit superior on both sides of the inequality we have

$$\langle y^*, y - F(x)\rangle + f\big(F(x)\big) \le \langle y^*, y - F(x)\rangle + \liminf_{n\to\infty} f_n\big(F_n(x_n)\big) \le \limsup_{n\to\infty} f_n(y_n).$$

Taking the infimum over all sequences y_n converging to y we have

$$\langle y^*, y - F(x)\rangle + f\big(F(x)\big) \le (Ls_e f_n)(y) = f(y).$$

This means that $y^* \in \partial f\big(F(x)\big)$ and hence

$$x^* = \nabla^* F(x) y^* \in \nabla^* F(x) \partial f\big(F(x)\big) = \partial(f \circ F)(x).$$

Our main theorem is thus a consequence of Propositions 2.1 and 2.2 and of Theorem 1.3.

Theorem 2.3 *Let $\{(f_n \circ F_n) : n \in \mathrm{N}\} \cup \{(f \circ F)\}$ be a sequence of convexly composite functions that is uniformly qualified over Ω. Assume that f_n slice converges to f, F_n pointwise converges to F over Ω and that for each $x \in \Omega$*

$$\lim_{\substack{n\to\infty \\ u\to x}} \nabla F_n(u) = \nabla F(x).$$

Then $(f_n \circ F_n)$ strongly epi-converges to $(f \circ F)$ over Ω and over $\Omega \cap \operatorname{dom}(f \circ F)$ one has

$$\operatorname{gph} \partial(f_n \circ F_n) \to^{pk} \operatorname{gph} \partial(f \circ F) \quad \textit{and} \quad \operatorname{gph} \partial(f_n \circ F_n) \to^{*pk} \operatorname{gph} \partial(f \circ F).$$

Proof First note that slice convergence of f_n to f implies in particular that f_n epi-converges to f in the strong topology. This shows that the assumptions of Proposition 2.2 are verified and we conclude that

$$\limsup_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n) \subset {}^*\limsup_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n) \subset \operatorname{gph} \partial(f \circ F)$$

over $\Omega \cap \operatorname{dom}(f \circ F)$ (the first inclusion always holds).

The previous observation also implies that the assumptions of Theorem 1.3 are verified at any $x \in \Omega \cap \operatorname{dom}(f \circ F)$. As a consequence of Theorem 1.3, we conclude that for any $x \in \Omega \cap \operatorname{dom}(f \circ F)$, the sequence $(f_n \circ F_n)$ epi-converges to $(f \circ F)$ on an open neighborhood of x. For a point $x \in \Omega$ but $x \notin \operatorname{dom}(f \circ F)$ we easily have that $(f \circ F)(x) = \lim_{n\to\infty}(f_n \circ F_n)(x_n)$ for any sequence x_n converging to x (this is because f_n epi-converges to f and $F_n(x_n)$ converges to $F(x)$). From this we conclude that $(f_n \circ F_n)$ strongly epi-converges to $(f \circ F)$ over Ω. It follows that for any $x \in \Omega$ there exists a sequence x_n that converges to x with $(f_n \circ F_n)(x_n)$ converging to $(f \circ F)(x)$. From this and the fact that a convexly composite function is subdifferentially qualified when it is qualified, we deduce that the assumptions of Proposition 2.1 are verified on Ω. From Proposition 2.1, we obtain that over Ω one has

$$\operatorname{gph} \partial(f \circ F) \subset \liminf_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n) \subset {}^{*}\liminf_{n\to\infty} \operatorname{gph} \partial(f_n \circ F_n)$$

(the second inclusion always holds). This completes the proof of the theorem. ∎

Remarks (1) We are grateful to a referee for pointing out that the result in Theorem 2.3 also holds in the norm $\times$ weak-star topology of $X \times X^*$.

(2) We point out that contrary to Theorem 1.3, where the functions are assumed to epi-converge, we suppose here in Theorem 2.3 that the functions f_n slice converge. This stronger assumption is natural. Indeed if F_n is the identity mapping over X (with $X = Y$) one has $f_n \circ F_n = f_n$ and the natural condition (see Theorem 1.2) for the Painlevé-Kuratowski convergence of $\operatorname{gph} \partial f_n$ is the slice convergence of f_n (but not the epi-convergence).

3 Applications to the Convergence of Multipliers

It is often useful in optimization theory to approximate a given problem with a sequence of hopefully simpler problems. In this context it is then important to know whether the sequence of multipliers will convergence to a multiplier for the original problem. In this section we show how the results of the previous sections can be applied to the convergence of multipliers. For the sake of simplicity and to avoid messy technical details we have limited our discussion to problems with inequality constraints.

Let $g_i \colon X \to \mathrm{R}$ for $i = 0, 1, \ldots, p$ be a family of $p + 1$ convexly composite functions with $g_i = f_i \circ F_i$. Consider also for each integer $n \in \mathrm{N}$ a family of $p + 1$ convexly composite functions $g_{i,n} = f_{i,n} \circ F_{i,n}$. Denote by (P) the mathematical programming problem

$$(P) \qquad \text{Minimize } g_0(x) \text{ subject to } g_i(x) \le 0 \text{ for all } i = 1, \ldots, p,$$

and recall that for a local solution x of (P), the vector $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_p) \in \mathrm{R}^{p+1}$ is a multiplier if $\lambda_i \ge 0$ for $i = 0, \ldots, p$, $\lambda_i g_i(x) = 0$ for $i = 1, \ldots, p$, $\|\lambda\| = 1$, and

$$0 \in \lambda_0 \partial g_0(x) + \lambda_1 \partial g_1(x) + \cdots + \lambda_p \partial g_p(x). \tag{3.1}$$

We refer to Clarke [6] for the existence of multipliers for problems with locally Lipschitz data. Denote by (P_n) the corresponding problem associated with $g_{i,n}$ for $i = 0, \ldots, p$. The

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

following theorem concerns the convergence of multipliers of optimization problems given by convexly composite functions. For other results with equi-lower semidifferentiable functions over some separable Banach spaces, we refer the reader to Zolezzi [30]. For bounds on Kuhn-Tucker points of perturbed optimization problems with nondifferentiable convex data, see Schultz [27].

Theorem 3.1 *Assume that for each $i = 0, \ldots, p$ the sequence $\{f_{i,n} \circ F_{i,n}\} \cup \{f_i \circ F_i\}$ satisfies for a local solution point x of (P) the assumptions of Theorem 2.3. Assume also that the sequence $\{g_{i,n}\}$ is equi-Lipschitz around x and that the closed unit ball of X^* is weak-star sequentially compact. Let x_n be a sequence of local solutions of (P_n) that converges to x. Then any limit of multipliers of (P_n) at x_n is a multiplier of (P) at x.*

Proof Let $\lambda^n = (\lambda_0^n, \lambda_1^n, \ldots, \lambda_p^n)$ be a multiplier of (P_n) at x_n that converges to some point $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_p)$. According to (3.1) take $\zeta_i^n \in \partial g_{i,n}(x_n)$ for $i = 0, \ldots, p$ such that

$$0 = \lambda_0^n \zeta_0^n + \lambda_1^n \zeta_1^n + \cdots + \lambda_p^n \zeta_p^n.$$

We may suppose that ζ_i^n weak-star converges to some $\zeta_i \in X^*$ for $i = 0, 1, \ldots, p$; this follows from the equi-Lipschitz assumption and the hypothesis on the Banach space X. Then $0 = \lambda_0\zeta_0 + \lambda_1\zeta_1 + \cdots + \lambda_p\zeta_p$ and by Proposition 2.2 one has that $\zeta_i \in \partial g_i(x)$ for each $i = 0, 1, \ldots, p$. As it is obvious that $\lambda_i \geq 0$ it remains only to show that $\lambda_i g_i(x) = 0$ for $i = 1, \ldots, p$ in order to conclude that $\lambda = (\lambda_0, \ldots, \lambda_p)$ is a multiplier for (P) at x. Fix i in $\{1, \ldots, p\}$. By (1.2) write $\zeta_i^n = \xi_i^n \circ \nabla F_{i,n}(x_n)$ with $\xi_i^n \in \partial f_{i,n}\big(F_{i,n}(x_n)\big)$. As $f_{i,n}$ epi-converges to f_i, there exist $y_{i,n} \to F_i(x)$ with $f_{i,n}(y_{i,n}) \to f_i\big(F_i(x)\big)$. Then, because $\lambda_i^n \geq 0$ and $f_{i,n}$ is convex we have

$$0 = \lambda_i^n f_{i,n}\big(F_{i,n}(x_n)\big) \leq \lambda_i^n f_{i,n}(y_{i,n}) - \lambda_i^n \langle \xi_i^n, y_{i,n} - F_{i,n}(x_n)\rangle. \tag{3.2}$$

By Proposition 1.1 the sequence ξ_i^n is bounded and as in the proof of Proposition 2.1 it is not difficult to see that $F_{i,n}(x_n) \to F_i(x)$. Therefore passing to the limit in (3.2) we obtain that $0 \leq \lambda_i f_i\big(F_i(x)\big) = \lambda_i g_i(x)$. From this we conclude that $0 = \lambda_i g_i(x)$ since the reverse inequality is verified because x is an admissible point of (P). This completes the proof. ∎

4 Applications to Second-Order Epi-Derivatives

Let $g\colon X \to \mathrm{R} \cup \{+\infty\}$. For $t > 0$ and $x^* \in \partial g(x)$ define the second-order difference quotients $g_{x,x^*,t}\colon X \to \mathrm{R} \cup \{+\infty\}$ by

$$g_{x,x^*,t}(\xi) := \frac{g(x + t\xi) - g(x) - t\langle x^*, \xi\rangle}{(1/2)t^2}. \tag{4.1}$$

We say, following Rockafellar [25], that g is *twice strongly epi-differentiable* at x relative to x^* if the second-order difference quotients $g_{x,x^*,t}$ strongly epi-converge as $t \downarrow 0$ to a proper function. The limit function is denoted by g''^{e}_{x,x^*}. With the obvious modifications we can also define the second-order slice epi-derivative g''^{s}_{x,x^*}.

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

Similarly, for a set-valued mapping $\Gamma\colon X \rightrightarrows X^*$ with $x^* \in \Gamma(x)$ define the first-order difference quotients $\Gamma_t\colon X \rightrightarrows X^*$ by

$$\Gamma_t(\xi) := \frac{\Gamma(x+t\xi) - x^*}{t}. \tag{4.2}$$

We say that Γ is Painlevé-Kuratowski (PK) proto-differentiable at x relative to x^* with proto-derivative Γ'^{pk}_{x,x^*} if $\operatorname{gph}\Gamma_t$ Painlevé-Kuratowski (PK) converges to $\operatorname{gph}\Gamma'^{pk}_{x,x^*}$.

Let $x^* \in \partial g(x)$ where $g := (f \circ F)$ is a convexly composite function on X. Assume that there exists $y^* \in \partial f\big(F(x)\big)$ with $x^* = \nabla^* F(x) y^*$ (this is the case for example when $(f \circ F)$ is subdifferentially qualified at x). We have

$$\begin{aligned} g_{x,x^*,t}(\xi) &= \frac{f\big(F(x+t\xi)\big) - f\big(F(x)\big) - t\langle y^*, \nabla F(x)\xi\rangle}{(1/2)t^2} \\ &= \frac{f\big(F(x) + t\big[\langle \nabla F(x)\xi + r(t,\xi)\big]\big) - f\big(F(x)\big) - t\langle y^*, \nabla F(x)\xi + r(t,\xi)\rangle}{(1/2)t^2} \\ &\qquad + (2/t)\langle y^*, r(t,\xi)\rangle. \end{aligned}$$

Here $r(t,\xi) = (1/t)[F(x+t\xi) - F(x) - t\nabla F(x)\xi]$. With

$$G_t(\xi) = \nabla F(x)\xi + r(t,\xi), \tag{4.3}$$

we have

$$g_{x,x^*,t}(\xi) = f_{F(x),y^*,t}\big(G_t(\xi)\big) + (2/t)\langle y^*, r(t,\xi)\rangle. \tag{4.4}$$

Note that for each $\xi \in X$, we have $r(t,\xi) \to 0$ as $t \downarrow 0$, which means that $G_t(\xi)$ converges to $\nabla F(x)\xi$ as $t \downarrow 0$.

We say that the derivative of F has a *first order expansion* at the point x for the mapping D if D is continuous and

$$\nabla F(x+t\xi)\xi = \nabla F(x)\xi + tD(\xi) + o(|t\xi|).$$

The mapping D giving the approximating term must not only be continuous but positively homogeneous: $D(\lambda\xi) = \lambda D(\xi)$ for $\lambda > 0$, and in particular $D(0) = 0$. Differentiability of ∇F at x in the classical sense is the case where D happens to be a linear mapping.

With the obvious modifications, the notion of uniform qualification can be adapted to a family indexed by $t > 0$.

Theorem 4.1 *Let $g = (f \circ F)$ be a convexly composite function with $F\colon X \to Y$ and $f\colon Y \to$ R $\cup \{+\infty\}$. Assume that ∇F has a first-order expansion at x given by the mapping D. Let $x^* \in \partial g(x)$. Assume further that $x^* = \nabla^* F(x) y^*$ with $y^* \in \partial f\big(F(x)\big)$ and that f is twice slice epi-differentiable at $F(x)$ relative to y^*. In addition assume that the family of convexly composite functions $\{(f_{F(x),y^*,t} \circ G_t) : t > 0\} \cup \{\big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)\}$ is uniformly qualified over some open neighborhood Ω of the origin. Then g is twice strongly epi-differentiable at x*

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

relative to x^*, and ∂g is Painlevé-Kuratowski proto-differentiable at x relative to x^*. Moreover, for all $\xi \in X$

$$g''^{e}_{x,x^*}(\xi) = \big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)(\xi) + \langle y^*, D(\xi)\rangle$$

and

$$\begin{aligned}(\partial g)'^{pk}_{x,x^*}(\xi) &= \partial(1/2)g''^{e}_{x,x^*}(\xi)\\ &= \partial(1/2)\big(f''^{s}_{F(x),y^*} \circ \nabla F(x) + y^* \circ D\big)(\xi).\end{aligned}$$

Proof Our assumptions imply that $(f_{F(x),y^*,t} \circ G_t)(\xi)$ strongly epi-converges as $t \downarrow 0$ to the function $\big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)$ over Ω. Because the second-order difference quotients are positively homogeneous of degree 2, the convergence is actually over X. The first-order expansion of ∇F at x combined with (4.4) and Theorem 1.3 shows that g is twice strongly epi-differentiable at x for x^*. The rest of the Theorem follows easily from Theorem 2.3 and the fact that $\partial(1/2)g_{x,x^*,t}(\xi) = (1/t)\big(\partial g(x + t\xi) - x^*\big)$. ■

Next we give a condition that implies that the difference quotients are uniformly qualified over some neighborhood of the origin.

Proposition 4.2 *In Theorem 4.1, the family $\{(f_{F(x),y^*,t} \circ G_t) : t > 0\} \cup \{\big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)\}$ is uniformly qualified over X if the operator $\nabla F(x)$ is surjective.*

Proof As $\nabla F(x)$ is surjective, one knows according to the Banach open mapping Theorem that there exists $\rho > 0$ such that

$$\mathrm{B}_Y \subset \nabla F(x)(\rho \mathrm{B}_X). \tag{4.5}$$

Note that for each ξ, $G_t(\xi)$ converges to $\nabla F(x)\xi$ as $t \downarrow 0$. Consider the family of convexly composite functions $\{(f_{F(x),y^*,t} \circ G_t) : t > 0\} \cup \{\big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)\}$ and fix any $\xi \in \operatorname{dom}\big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)$. As $f_{F(x),y^*,t}$ epi-converges to $f''^{s}_{F(x),y^*}$ as t goes to 0 there exist $y_t \to \nabla F(x)\xi$, $t_0 > 0$ such that for all positive $t < t_0$ we have $f_{F(x),y^*,t}(y_t) < (1/2) + f''^{s}_{F(x),y^*} \circ \nabla F(x)(\xi)$ and $\|y_t - \nabla F(x)(\xi)\| \le (1/2)$. Then for all positive $t < t_0$ (because of (4.5)) we have for $r := \max\{\rho, 1/2\}$

$$\mathrm{B}_Y \subset \Big(\{f_{F(x),y^*,t} \le r + \big(f''^{s}_{F(x),y^*} \circ \nabla F(x)(\xi)\big)\} - \nabla F(x)(\xi)\Big) + (1/2)\mathrm{B}_Y - \nabla F(x)(r\mathrm{B}_X).$$

Using Lemma 2 in Robinson [24] we obtain that for all positive $t < t_0$

$$(1/3)\mathrm{B}_Y \subset \Big(\{f_{F(x),y^*,t} \le r + \big(f''^{s}_{F(x),y^*} \circ \nabla F(x)(\xi)\big)\} - \nabla F(x)(\xi)\Big) - \nabla F(x)(r\mathrm{B}_X).$$

This shows that the above family is uniformly qualified over X and hence the proof is complete. ■

As a final application, we turn our attention to the maximum of finitely many C^2 functions. The proof of the following corollary follows immediately from Theorem 4.1, Proposition 4.2 and the fact the function $f(y) := \max\{y_1, \ldots, y_m\}$ is twice (slice) epi-

https://doi.org/10.4153/CJM-1999-013-6 Published online by Cambridge University Press

differentiable on R^m (*cf.* [25]). In this case, $f''^{s}_{y,y^*}(\xi)$ is simply the indicator of the normal cone to the subdifferential of f at the point y^*; see [25] for details.

Corollary 4.3 *Let $g(u) = \max_{1\le i\le m} f_i(u)$ where $f_i \colon X \to \mathrm{R}$ are C^2 functions. Let $F(x) := \big(f_1(x), \dots, f_m(x)\big)$. Assume that $\nabla f_i(x)$ are linearly independent for $i \in I(x) := \{i \mid f_i(x) = g(x)\}$. Then g is twice strongly epi-differentiable at x relative to any $x^* \in \partial g(x)$, and ∂g is Painlevé-Kuratowski proto-differentiable at x relative to x^*. Moreover, for all $\xi \in X$ and $y^* \in \partial f\big(F(x)\big)$ with $x^* = y^* \circ \nabla F(x)$ we have*

$$g''^{e}_{x,x^*}(\xi) = \big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)(\xi) + \langle y^*, \nabla^2 F(x)(\xi)\xi\rangle$$

and

$$\begin{aligned}(\partial g)'^{pk}_{x,x^*}(\xi) &= \partial(1/2) g''^{e}_{x,x^*}(\xi)\\ &= \partial(1/2)\big(f''^{s}_{F(x),y^*} \circ \nabla F(x)\big)(\xi) + \nabla^2(y^* \circ F)(x)(\xi).\end{aligned}$$

Remark In finite dimensional spaces, Poliquin [19] first showed that the maximum of finitely many C^2 functions is proto-differentiable and gave a formula relating the proto-derivative of the subgradient mapping to the subdifferential of the second-order epi-derivative (all this without the linear independence condition above). Precise formulas were then given in Poliquin-Rockafellar [23]. Penot [18] then showed, in spaces that satisfy "Condition (A)", that the subgradient mapping of the maximum of finitely many C^2 functions is proto-differentiable. Here we do not assume "Condition (A)" however we assume linear independence.

References

[1] H. Attouch, *Variational Convergence for Functions and Operators.* Pitman, Boston-London-Melbourne, 1984.

[2] H. Attouch and G. Beer, *On the convergence of subdifferentials of convex functions.* Arch. Math. (Basel) **60**(1993), 389–400.

[3] H. Attouch, J. L. Ndoutoume and M. Théra, *Epigraphical convergence of functions and convergence of their derivatives in Banach spaces.* Exp. No. 9, Sem. Anal. Convexe **20**(1990), 9.1–9.45.

[4] G. Beer, *The slice topology: A viable alternative to Mosco convergence in nonreflexive spaces.* Nonlinear Anal. **19**(1992), 271–290.

[5] ———, *Topologies on Closed and Closed Convex Sets.* Kluwer Academic Publishers, Dordrecht-Boston-London, 1993.

[6] F. H. Clarke, *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983.

[7] R. Cominetti, *On pseudo-differentiability.* Trans. Amer. Math. Soc. **324**(1991), 843–865.

[8] C. Combari and L. Thibault, *On the graph convergence of subdifferentials of convex functions.* Proc. Amer. Math. Soc. **126**(1998), 2231–2240.

[9] ———, *Epi-convergence of convexly composite functions in Banach spaces.* Submitted.

[10] C. Combari, A. Elhilali Alaoui, A. Levy, R. Poliquin and L. Thibault, *Convex composite functions in Banach spaces and the primal-lower-nice property.* Proc. Amer. Math. Soc. **126**(1998), 3701–3708.

[11] R. Correa, A. Jofré and L. Thibault, *Subdifferential characterization of convexity.* In: Recent Advances in Nonsmooth Optimization (Eds. D-Z. Du, L. Qi and R. S. Womersley), World Scientific Publishing, 1995, 18–23.

[12] G. Dal Maso, *An introduction to Gamma-convergence.* Birkhauser, 1993.

[13] R. Deville, *Stability of subdifferentials of nonconvex functions in Banach spaces.* Set-Valued Anal. **2**(1994), 141–157.

[14] S. Guillaume, *Evolution equations governed by the subdifferential of a convex composite function in finite dimensional equations.* Discrete Contin. Dynam. Systems **2**(1996), 23–52.

[15] ———, *Problèmes d'optimisation et d'évolution en analyse nonconvexe de type convexe composite.* Thesis, Université Montpellier II, 1996.

[16] A. B. Levy, *Second-order variational analysis with application to sensitivity in optimization.* Ph.D. Thesis, University of Washington, 1994.

[17] A. B. Levy, R. A. Poliquin and L. Thibault, *Partial extensions of Attouch theorem and applications to proto-derivatives of subgradient mappings.* Trans. Amer. Math. Soc. **347**(1995), 1269–1294.

[18] J. P. Penot, *On the interchange of subdifferentiation and epi-convergence.* J. Math. Anal. Appl. **196**(1995), 676–698.

[19] R. A. Poliquin, *Proto-differentiation of subgradient set-valued mappings.* Canad. J. Math. **42**(1990), 520–532.

[20] ———, *Integration of subdifferentials of nonconvex functions.* Nonlinear Anal. **17** (1991), 385–398.

[21] ———, *An extension of Attouch's theorem and its application to second order epi-differentiation of convexly composite functions.* Trans. Amer. Math. Soc. **332**(1992), 861–874.

[22] R. A. Poliquin, J. Vanderwerff and V. Zizler, *Renormings and convex composite representation of functions.* Bull. Polish Acad. Sci. Math. **42**(1994), 9–19.

[23] R. A. Poliquin and R. T. Rockafellar, *Proto-derivative formulas for basic subgradient mapping in mathematical programming.* Set-Valued Anal. **2**(1994), 275–290.

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

[25] R. T. Rockafellar, *First and second-order epi-differentiability in nonlinear programming.* Trans. Amer. Math. Soc. **307**(1988), 75–107.

[26] ———, *Generalized second derivatives of convex functions and saddle functions.* Trans. Amer. Math. Soc. **320**(1990), 810–822.

[27] R. Schultz, *Estimates for Kuhn-Tucker points of perturbed convex programs.* Optimization **19**(1988), 29–43.

[28] L. Thibault, *A note on the Zagrodny mean value theorem*. Optimization **35**(1995), 127–130.

[29] L. Thibault and D. Zagrodny, *Integration of subdifferentials of lower semicontinuous functions on Banach spaces.* J. Math. Anal. Appl. **189**(1995), 33–58.

[30] T. Zolezzi, *Convergence of generalized gradients.* Set-Valued Anal. **2**(1994), 381–393.

Université Montpellier II
Laboratoire Analyse Convexe
Place Eugène Bataillon
34095 Montpellier Cedex 5
France

University of Alberta
Department of Mathematical Sciences
Edmonton, Alberta
T6G 2G1