Set-Valued Analysis **2:** 241–258, 1994.

Graphical Methods in First and Second-Order Differentiability Theory of Integral Functionals

DOMINIKUS NOLL*
University of Waterloo, Department of Combinatorics and Optimization, Waterloo, Ontario, N2L 3G1, Canada and Universität Stuttgart, Mathematisches Institut B, Pfaffenwaldring 57, 70550, Stuttgart, Germany

(Received: 1 December 1993)

Abstract. We discuss several notions of first and second-order differentiability for integral functionals on a Hilbert space.

Mathematics Subject Classification (1991). 49J52.

Key words: Second-order differentiability, Mosco convergence, Attouch–Wets convergence.

1. Introduction

Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space, and let $L^p_{\mathbb{R}^d}(\Omega, \mathcal{A}, \mu)$ (for $1 \leq p < \infty$) be the space of classes of measurable functions $u : \Omega \to \mathbb{R}^d$ having

$$\|u\|_p = \left(\int_\Omega |u(x)|^p \, \mu(dx) \right)^{1/p} < +\infty.$$

We consider an integral functional f of the form

$$f(u) = \int_\Omega \phi(x, u(x)) \, \mu(dx), \quad u \in L^p_{\mathbb{R}^d} \tag{1.1}$$

defined on the space $L^p_{\mathbb{R}^d}$, where $\phi : \Omega \times \mathbb{R}^d \to \mathbb{R}$ is $\mathcal{A} \otimes \mathcal{B}_{\mathbb{R}^d}$ measurable. While our results concerning first and second epiderivatives (Sections 3 and 4) apply to functionals (1.1) which may take on the value $+\infty$, a discussion of the classical first and second derivatives requires f to be finite everywhere. This is the case e.g. when ϕ satisfies a growth condition of the form

$$|\phi(x, u)| \leq C|u|^p + g(x) \cdot u + h(x) \tag{1.2}$$

for some $C \geqslant 0$, $g \in L^{p'}_{\mathbb{R}^d}$ $(1/p + 1/p' = 1)$, and $h \in L^1$. Further, when f is finite everywhere, we assume for convenience that f is a continuous function on the

* Supported in part by Deutsche Forschungsgemeinschaft (DFG) and Natural Sciences and Engineering Research Council of Canada (NSERC).

space $L^p_{\mathbb{R}^d}$, which is the case e.g. when $\phi(x,u)$ is continuous in u and satisfies the growth condition (1.2) (see [19, p. 93]). One might even assume that f is locally Lipschitz on $L^p_{\mathbb{R}^d}$, which is true for instance when the integrand satisfies a uniform Lipschitz condition of the form

$$|\phi(x,u) - \phi(x,v)| \leq C(x) \cdot |u - v| \tag{1.3}$$

for all $u, v \in \mathbb{R}^d$, a function $C(\cdot)$ in $L^{p'}_{\mathbb{R}+}$, and for almost all $x \in \Omega$. For $p = 1$, $C(x) = C$ is a constant, and in this case condition (1.3) is necessary *and* sufficient for f to be (locally) Lipschitz. The score of our present investigation lies in studying the differentiability properties of the functionals f considered as functions on $L^p_{\mathbb{R}^d}$.

Naturally, we first focus on the case where the integrand $\phi(x,u)$ satisfies some smoothness condition. There are several questions here, for instance, if $\phi(x,u)$ is of class C^1 in u, under what conditions will f be of class C^1 as a function on $L^p_{\mathbb{R}^d}$? Which is the right notion of differentiability to be employed, Fréchet or Gâteaux? Concerning second-order differentiability, if $\phi(x,u)$ is of class C^2 in u, what are the second order differentiability properties of f as a function on the space $L^p_{\mathbb{R}^d}$? And again, which is the right notion of second-order differentiability?

In the case of a convex functional (1.1) it is known that if the integrand $\phi(x,u)$ is of class C^1 in u, then f is itself of class C^1 in the sense that ∇f exists everywhere as a Fréchet derivative, and is norm to norm continuous as an operator $L^p_{\mathbb{R}^d} \to L^{p'}_{\mathbb{R}^d}$ (cf. [19]). Naturally, the same observation pertains to an even wider class of functionals (1.1) which are common in various applications such as variational problems, control, or Hamiltonian mechanics, namely when $f + g$ is convex for some function g of class C^1. Without the presence of convexity, stronger assumptions are needed to guarantee even the differentiability of f at u. For instance, a necessary requirement for f to be differentiable at u is of course that $x \to \nabla\phi(x, u(x))$ be an element of $L^{p'}_{\mathbb{R}^d}$, but this is not sufficient in general. The natural condition to guarantee Fréchet differentiability of the functional (1.1) at every u is that the integrand satisfies the uniform Lipschitz condition (1.3). See [19, 38, 33, 29] for a discussion. After all, the first-order differentiability theory of integral functionals (1.1) has been under discussion for a long time, and is basically well-understood.

Concerning second-order differentiability of the functional (1.1), the situation turns out to be considerably more complicated. To begin with, we observe that for $1 \leq p < 2$, the convex functional $f = \|\cdot\|_p$ on the space L^p is *nowhere* second-order differentiable when this notion is understood in the sense that the difference quotient

$$\frac{\nabla f(u + th) - \nabla f(u)}{t} \tag{1.4}$$

be convergent in norm (resp. weakly) for every fixed vector h (as $t \to 0$). The reason for this failure is that second-order differentiability in the quoted sense requires the function f to be *Lipschitz smooth* (as defined in [20], see also [14, 13, 16]), and this conditions fails for the norms $\|\cdot\|_p,\ 1 \leq p < 2$. As $\|\cdot\|_p^p$ is an integral functional on $L^p_{\mathbb{R}^d}$, we infer that even smoothness of the type C^∞ for the integrand need not imply second-order smoothness of the functional (1.1). Indeed, we may adjust the integrand near the origin by replacing $|\cdot|^p$ by a convex C^∞ function showing the mentioned behaviour. This shows in particular that the situation in the spaces L^p for $1 \leq p < 2$ is very different from the Hilbert space case $p = 2$, since in a space L^2, a convex integral functional with C^2 integrand is at least densely twice differentiable (see [29] and [13]).

In Hilbert space $p = 2$, for historical reasons, let us quote the following scenario from the classical work by Palais and Smale [32] on infinite-dimensional Morse Theory. The authors consider integral functionals (1.1) (or more generally integral functionals depending on a differential operator as for instance discussed in [29]) having smooth integrand. In order to build a theory in analogy to the finite-dimensional case, they wish to deal with smooth functionals (1.1), and they claim that the following growth condition (stated here in terms of the functional (1.1)) should guarantee the latter to be of class C^2:

$$\left| \frac{\partial^2 \phi(x,u)}{\partial u_i \partial u_j} \right| \leq C < \infty \tag{1.5}$$

for $i, j = 1, \ldots, d$, all $u \in \mathbb{R}^d$, and almost all x. Although this is a natural idea, our analysis of the functionals (1.1), (1.2) satisfying the condition (1.5) obtained in [29] shows that the statement quoted from [32] is only correct when class C^2 is understood in the following weak sense: the difference quotient (1.4) converges pointwise in norm, in other terms, ∇f is Gâteaux differentiable, and the Hessian operator $\nabla^2 f$ is norm to weakly continuous. Examples presented in [29] show that the difference quotient (1.4) may fail to converge uniformly over $\|h\|_2 \leq 1$ *for all* u, i.e., ∇f is not Fréchet differentiable at any u, and that $\nabla^2 f$ may fail to be norm to norm continuous throughout.

Suppose now in the case $p = 2$ we have an integral functional with smooth integrand, but not necessarily satisfying (1.5). Is it true that f is second-order differentiable in the Gâteaux sense at those $u \in L^p_{\mathbb{R}^d}$ where the boundedness condition (1.5) *is* satisfied? Surprisingly, even for convex f, this need not be the case, i.e., the difference quotient (1.4) need not even converge pointwise weakly. Nevertheless, in this situation, we wish to have a kind of substrate for the classical second derivative, which we would then call a *generalized second derivative.* Indeed, the obvious candidate for a generalized Hessian operator H_u of f at u is

$$\tfrac{1}{2}\langle H_u h, h\rangle = \tfrac{1}{2}\int_\Omega \langle \nabla^2\phi(x,u(x))h(x), h(x)\rangle\,\mu(dx), \tag{1.6}$$

which is bounded symmetric and linear as a consequence of (1.5). This idea is made precise by employing the theory of graphical convergence. Namely, in the above situation, we can prove that the second difference quotient of f at u epi-converges to the quadratic form (1.6), (see Section 2), or equivalently, that the difference quotient (1.4) proto-converges to the limit H_u, a concept of convergence which is usually weaker than pointwise type convergence of (1.4). This concept of a generalized differentiability has been proposed by R.T. Rockafellar [34, 35, 36], see also [30], in finite dimensions, and we will show here that a similar approach is successful in the second-order theory of the integral functionals of type (1.1). In contrast with the situation in finite dimensions, however, there are several different notions of graphical convergence, such as epi, Mosco or Attouch–Wets convergence, each running for the mandate of replacing pointwise type convergence notions, and one of our issues here is to clarify which of them has to be elected to allow for a reasonable theory.

We end by observing that formula (1.6) for the generalized Hessian provides useful information even for $p \neq 2$. In this case, for H_u to be fully defined, it is necessary that the operator $\nabla^2\phi(\cdot, u(\cdot))$ maps $L^p_{\mathbb{R}^d}$ into $L^{p'}_{\mathbb{R}^d}$. As we can easily see, in the case $p = 1$, and when $(\Omega, \mathcal{A}, \mu)$ has no atoms, this is certainly impossible unless the Hessian operators are zero almost everywhere, which tells us roughly that there is no reasonable second order differentiability theory for integral functionals on L^1 spaces.

2. Notions of Differentiability

In this section we recall various notions of first and second-order differentiability and discuss their interrelation.

Let f be a real-valued (or more generally extended real-valued) continuous function defined on a Banach space E. For fixed x and $y^* \in E^*$ we consider the first and second order difference quotients of f at x:

$$.noindent \quad \delta_{f,x,t}(h) = \frac{f(x+th)-f(x)}{t},$$

$$\Delta_{f,x,y^*,t}(h) = \frac{f(x+th)-f(x)-\langle y^*, th\rangle}{t^2}, \tag{2.1}$$

considered for every fixed $t \neq 0$ as functions of $h \in E$. Here y^* will be the gradient $\nabla f(x)$ of f at x, or more generally a subgradient $y^* \in \partial f(x)$ in the sense of convex analysis when f is convex.

While first-order differentiability of f is described by the pointwise convergence behaviour of the first difference quotient $\delta_{f,x,t}$ as $t \to 0$, one would naturally wish to describe the second-order differentiability properties of f by using the convergence of the difference quotient (1.4) of ∇f at x to a bounded and symmetric linear operator H_x, called the Hessian of f at x, and noted $\nabla^2 f(x)$. There are at least four different notions of convergence of the pointwise type, which might be used for (1.4).

DEFINITION 2.1. If (1.4) converges pointwise in h with respect to the norm, resp. pointwise in h and in the weak star topology, then f is said to be second-order Gâteaux differentiable, resp. second-order weak star Gâteaux differentiable, at x. On the other hand, if (1.4) converges uniformly on the ball $\|h\| \leq 1$ and in norm resp. uniformly on the ball $\|h\| \leq 1$ and in the weak star topology, then f is said to be second-order Fréchet differentiable at x, resp. second-order weak star Fréchet differentiable, at x. □

There is a different approach to second-order differentiability which is motivated by the situation of convex analysis. Here one wishes to discuss second-order notions without having a first-order derivative at all points in a neighbourhood x. Namely, consider convergence of the second-order difference quotient (2.1), for instance, pointwise, or uniform convergence on compact, resp. bounded, sets.

DEFINITION 2.2. We write $x \in D_f^2$ if the second-order difference quotient (2.1) of f at x converges uniformly on compact sets to some purely quadratic limit function $q_x : E \to \mathbb{R}$ having domain $\mathrm{dom}(q_x) = E$. That is, $\Delta_{f,x,y^*,t} \to q_x$ uniformly on compact sets as $t \to 0$. □

It is clear that, at least for f locally Lipschitz, $x \in D_f^2$ implies that f is differentiable at x, so $\partial f(x)$ is singleton. In [13] it was proved that for a convex f, $x \in D_f^2$ implies the even stronger fact that f is Fréchet differentiable at x.

Recall here that a function q: $E \to \mathbb{R} \cup \{\infty\}$ is called *purely quadratic* if its domain is a linear subspace of E, and q admits a representation of the form

$$q(h) = \tfrac{1}{2}\langle Th, h\rangle, \quad h \in \mathrm{dom}(q)$$

with a closed and symmetric linear operator T: $\mathrm{dom}(q) \to E^*$. (A characterization of the convex purely quadratic functions in terms of the graph of ∂q has been obtained in [13] in the case of a Hilbert space. For an extension to the setting of

Banach spaces see [28].) In contrast, a function q (fully defined or not) is called quadratic if $q(\lambda h) = \lambda^2 q(h)$, and convergence $\Delta_{f,x,y^*,t} \to q$ for quadratic limit functions might as well be used for a definition of second-order differentiability (see, for instance, [34, 35]). Notice, for instance, that for a function $f\colon \mathbb{R}^2 \to \mathbb{R}$, convergence $\Delta_{f,x,y^*,t} \to q$ with quadratic but not purely quadratic q having $\mathrm{dom}(q) = \mathbb{R}^2$ implies that the second partial derivatives $f_{x_i x_j}$ at $x = (x_1, x_2)$ exist, but that $f_{x_1 x_2}(x) \neq f_{x_2 x_1}(x)$.

The interrelation between the convergence of (1.4), resp. (2.1), has been discussed by many authors, mostly in the finite-dimensional case (see [15, 7, 22, 13, 11, 34]). We just mention the result obtained in [13, §2] for convex f in any Banach space, which states that pointwise convergence of the second difference quotient (2.1) corresponds to second-order weak star Gâteaux differentiability, while uniform convergence of (2.1) on bounded sets corresponds to second-order Fréchet differentiability of f. It is clear that this is no longer true in the nonconvex case, as for instance shown by the example $f(x) = x^3 \cos(1/x)$.

Let us now focus on first and second-order differentiability notions which are based on graphical type convergence of the difference quotients (2.1). We shall discuss here the notions of norm epi convergence, Mosco convergence, and Attouch–Wets convergence.

Let f_n, f be real-valued (or more generally extended real-valued) functions defined on the Banach space E. The sequence (f_n) is said to *epi converge* to the limit f if the following conditions (α), (β) are satisfied:

(α) Given any $x \in E$, there exist $x_n \to x$ (norm) such that $f_n(x_n) \to f(x)$.

(β) Given any $x \in E$, a sequence $n_k \nearrow \infty$ of indices, and a sequence $x_k \to x$ (norm), we have $f(x) \leq \liminf_{k\to\infty} f_{n_k}(x_k)$.

The sequence (f_n) is said to *Mosco converge* to the limit f if conditions (α) and $(\tilde\beta)$ are satisfied, where:

$(\tilde\beta)$ Given any $x \in E$, a sequence $n_k \nearrow \infty$ of indices, and a sequence $x_k \rightharpoonup x$ (weakly), we have $f(x) \leq \liminf_{k\to\infty} f_{n_k}(x_k)$.

We use the notations $f_n \xrightarrow{e} f$ and $f_n \xrightarrow{m} f$. Clearly Mosco convergence implies epi convergence. Notice that the constant sequence $f_n = f$ fails to converge to the limit f unless f is weakly lower semi-continuous. This explains why Mosco convergence is usually restricted to the context of convex functions. Here, however, we shall give credit to Mosco convergence in a more general setting. Similarly, epi convergence requires the functions f_n, f to be lower semi-continuous in the norm topology, but this is a reasonable requirement even in the nonconvex case. We refer to [2, 4, 17, 37, 13] for a discussion of these notions.

Concerning the concept of Attouch–Wets convergence or equivalently, convergence with respect to the epi distance, we have to recall the following notions. Let C, D be subsets of E, then the *excess* of C, D is defined as $\mathrm{ex}(C, D) := \sup_{x \in C} d(x, D)$. For $\rho > 0$ let

$$\mathrm{haus}_\rho(C, D) = \max\{\mathrm{ex}(C_\rho, D), \mathrm{ex}(D_\rho, C)\},$$

where $C_\rho = C \cap B(0, \rho)$. Notice that for bounded sets C, D and ρ large enough, this is just the usual Hausdorff distance of the sets C and D. Now the sequence (f_n) is said to be Attouch–Wets convergent to the limit f, henceforth noted $f_n \xrightarrow{\mathrm{aw}} f$, if for all $\rho > 0$ sufficiently large, we have

$$\mathrm{haus}_\rho(\mathrm{epi}\, f_n, \mathrm{epi}\, f) \to 0. \tag{2.2}$$

See [3] and [5] for details on this notion of graphical convergence. Let us mention here that $f_n \xrightarrow{\mathrm{aw}} f$ implies $f_n \xrightarrow{m} f$ when f is weakly (sequentially) lower semi-continuous, and implies $f_n \xrightarrow{e} f$ when f is lower semi-continuous in norm.

It seems natural to apply these notions of convergence to the first and second-order difference quotients (2.1) of a function f. As it turns out, the result of this investment is quite different on the first and second-order level. While graphical convergence of $\delta_{f,x,t}$ (as $t \to 0$) does not really provide much new insight (see Section 3), we will see that graphical convergence of $\Delta_{f,x,y^*,t}$ in fact does (see Section 4). Let us mention that, in infinite dimensions, a systematic account on the use of graphical second-order notions in the context of differentiability has been developed quite recently by J.M. Borwein and D. Noll [13], D. Noll [29], and also [23], [28], [17], [27], [24].

3. First-Order Theory

In this paragraph we discuss graphical convergence notions for the first-order difference quotient. As it turns out, these coincide with pointwise type convergence under fairly reasonable side conditions, in particular when the function f under consideration is locally Lipschitz, and therefore do not provide much new insight. Namely, we have the following (more or less standard) result (compare with [18 Thm. 2.18]).

PROPOSITION 3.1. *Let f be a locally Lipschitz function on a Banach space E. Let $x \in E$, and suppose $\delta_{f,x,t} \xrightarrow{e} \delta$ as $t \to 0$. Then $\delta_{f,x,t} \to \delta$ pointwise and hence uniformly on compact sets. Conversely, pointwise convergence $\delta_{f,x,t} \to \delta$ implies epi convergence $\delta_{f,x,t} \xrightarrow{e} \delta$. Moreover, in these cases, the limit function δ is fully defined.*

Proof. Due to the local Lipschitz assumption, pointwise convergence $\delta_t \to \delta$ implies uniform convergence on compact sets, and hence epi convergence. Notice here that, due to the local Lipschitz assumption, δ_t is uniformly bounded, and hence δ is fully defined. This proves the first part. Now suppose $\delta_t \xrightarrow{e} \delta$ for a lower semi-continuous extended real-valued limit function δ. Let h and a sequence $t_n \to 0$ be fixed. Using condition (α) find $h_n \to h$ (norm) such that $\delta_{t_n}(h_n) \to \delta(h)$. As f is Lipschitz in a neighbourhood of x, with constant C say, we find

$$\left|\delta_{t_n}(h_n) - \delta_{t_n}(h)\right| \leq C\|h_n - h\|$$

for n large enough. This proves $\delta_{t_n}(h) \to \delta(h)$. Moreover, the local Lipschitz condition again guarantees that δ_t is uniformly bounded and so δ is fully defined. □

This generalizes a result obtained in [35] for the class of convex functions. Concerning Mosco convergence of the first difference quotient, we have the following result.

PROPOSITION 3.2. *Let f be a locally Lipschitz function on a Banach space E. Let $x \in E$, then the following statements are equivalent:*

(1) *f is weakly Hadamard differentiable at x;*
(2) *$\delta_{f,x,t} \xrightarrow{m} \langle \nabla f(x), \cdot \rangle$ as $t \to 0$.*

Proof. Recall that weak Hadamard differentiability of f at x means that $\delta_{f,x,t}$ converges to $\langle \nabla f(x), \cdot \rangle$ uniformly on weakly compact sets as $t \to 0$. Now it is clear that condition $(\tilde{\beta})$ of Mosco convergence of $\delta_{f,x,t}$ gives the estimate

$$\liminf_{n \to \infty} \left(\delta_{f,x,t_n}(h_n) - \langle \nabla f(x), h_n \rangle\right) \geq 0 \tag{3.1}$$

for every fixed $t_n \to 0$ and h_n converging weakly. But replacing t_n by $-t_n$ and h_n by $-h_n$ gives rise to a similar estimate, which by $\delta_{f,x,-t_n}(-h_n) = -\delta_{f,x,t_n}(h_n)$ is the reverse estimate to (3.1). This shows (1) and (2) being in fact equivalent. □

Let us now consider Attouch–Wets convergence of the first difference quotient. The following no longer comes as a surprise.

PROPOSITION 3.3. *Let f be a locally Lipschitz function on a Banach space E. Let $x \in E$, then the following statements are equivalent:*

(1) *$\delta_{f,x,t} \xrightarrow{\text{aw}} \langle y^*, \cdot \rangle$ as $t \to 0$;*
(2) *f is Fréchet differentiable at x with gradient y^*.*

Proof. Observe that $\delta_t \xrightarrow{\text{aw}} \delta := \langle y^*, \cdot \rangle$ implies $\delta_t \xrightarrow{e} \delta$ and hence implies pointwise convergence by Proposition 3.1. Suppose now that convergence fails to be uniform on the unit ball. Find $\|h_n\| \leq 1$ and $t_n \to 0$ such that

$$|\delta_{t_n}(h_n) - \delta(h_n)| \geq \epsilon \tag{3.2}$$

for some $\epsilon > 0$ and all n. By assumption we have $\text{ex}((\text{epi}\,\delta_{t_n})_\rho, \text{epi}\,\delta) \to 0$ for an appropriate ρ, so using the fact that f is locally Lipschitz, we find h_n' such that

$$\|h_n - h_n'\| + |\delta_{t_n}(h_n) - \delta(h_n') - \sigma_n| \to 0 \qquad (n \to \infty) \tag{3.3}$$

for certain $\sigma_n \geq 0$. On the other hand, $\text{ex}((epi\ \delta)_\rho, epi\ \delta_{t_n}) \to 0$ provides a sequence h_n'' such that

$$\|h_n - h_n''\| + |\delta(h_n) - \delta_{t_n}(h_n'') - \tau_n| \to 0 \qquad (n \to \infty) \tag{3.4}$$

for certain $\tau_n \geq 0$. Now (3.2), (3.3), (3.4), and the fact that $\delta(h_n) - \delta(h_n') \to 0$ (because of $\|h_n - h_n'\| \to 0$) combine to give the estimate $\liminf |\delta_{t_n}(h_n) - \delta_{t_n}(h_n'')| \geq \epsilon$, and this in tandem with $\|h_n - h_n''\| \to 0$ contradicts the local Lipschitz behaviour of f at x. Hence δ_t converges uniformly on the unit ball. As for the converse, the fact that uniform convergence on bounded sets implies Attouch–Wets convergence is straightforward but a little tedious. □

The reader might observe here that the proofs of Proposition 3.1 and 3.3 did not really use the linearity of the limit function δ so they still apply when we consider only one-sided limits ($t \to 0^+$), as for instance in [34, 35, 36]. On the contrary, using two-sided limits turned out to be essential for the implication (2) ⇒ (1) in the proof of Proposition 3.2. Another observation is of course that we did not in all places need the fact that f is locally Lipschitz. It would have been sufficient to assume that the difference quotient is bounded in a neighbourhood of 0. Let us further observe that the question on when Mosco convergence of $\delta_{f,x,t}$ coincides with epi convergence resp. with Attouch–Wets convergence has been decided by J. Borwein, M. Fabian and J. Vanderwerff [10, 12]. Namely, Mosco convergence and epi convergence coincide iff every weak star convergent sequence in E^* is Mackey convergent, while on the other hand, Mosco convergence of $\delta_{f,x,t}$ coincides with Attouch–Wets convergence iff E is sequentially reflexive, which is to say that every Mackey convergent sequence in E^* is convergent in the dual norm. Equivalently, E must not contain a copy of ℓ_1. In particular, this is the case for Asplund spaces. (See [9]).

For the remainder of this paragraph we consider first order epi derivatives of integral functionals (1.1). We no longer assume the functional f to be locally Lipschitz, trying to express the convergence $\delta_{f,u,t} \xrightarrow{e} \delta$ in terms of the integrand.

It becomes clear by looking at easy examples that the boundedness condition (1.5) at a point u does not by itself guarantee epi differentiability of f at u, although the Hessian type operator (1.6) is defined. What is needed in addition is a uniform boundedness below condition on the integrand.

PROPOSITION 3.4. *Let f be an integral functional (1.1), (1.2) on the space $L^p_{\mathbb{R}^d}$, $(1 \leq p < \infty)$. Let $u \in L^p_{\mathbb{R}^d}$, and suppose $\phi(x, \cdot)$ is Fréchet differentiable at $u(x)$ for almost all x. Suppose the difference quotient $\delta_t(x, \cdot) := \delta_{\phi(x,\cdot),u(x),t}$ of $\phi(x, \cdot)$ at $u(x)$ satisfies the following condition:*

$$\delta_t(x, \xi) + C|\xi|^p + g_1(x) \cdot \xi + g_0(x) \geqslant 0 \tag{3.5}$$

for some $C > 0$, $g_1 \in L^{p'}_{\mathbb{R}^d}$ and $g_0 \in L^1$, all $0 < |t| \leq 1$, $\xi \in \mathbb{R}^d$ and almost all x. Then, (a) f is epi differentiable at u, i.e., $\delta_{f,u,t} \xrightarrow{e} \langle \nabla f(u), \cdot \rangle$. Here $\nabla f(u)$ is defined by $\nabla f(u)(x) = \nabla \phi(x, u(x))$.

Moreover, (b) for $p = 1$, and if $(\Omega, \mathcal{A}, \mu)$ has no atoms, condition (3.5) is also necessary for f to be epi differentiable at u.

Proof. (a) We have to check conditions (α) and (β) for the epi convergence. Concerning condition (α), notice that it suffices to find a dense subset Φ of $L^p_{\mathbb{R}^d}$ such that $\delta_t := \delta_{f,u,t}$ converges pointwise on Φ. Now, by assumption, for almost every x we find $t(x) > 0$ such that

$$|\delta_t(x, \xi) - \langle \nabla \phi(x, u(x)), \xi \rangle| \leq 1 \tag{3.6}$$

for all $|\xi| \leq 1$ and all $|t| \leq t(x)$. We may assume here that the function $x \to t(x)$ is measurable. Let $E_k = \{x \in \Omega : t(x) \geq 1/k\}$, then $\Omega \setminus \bigcup E_k$ is a null set. Let the Ω_k be an increasing sequence with union Ω such that $\mu(\Omega_k) < \infty$. Now let $\Phi_{k,m,n}$ be the set of all $h \in L^p_{\mathbb{R}^d}$ such that $h(x) = 0$ for $x \notin E_k \cup \Omega_m$ and having $\|h\|_\infty \leq n$. Let Φ be the union of all $\Phi_{k,m,n}$, then Φ is certainly dense. We show that $\delta_t(h) \to \delta(h) = \langle \nabla f(u), h \rangle$ for $h \in \Phi_{k,m,n}$. Indeed, the estimate (3.6) shows that the $\delta_t(x, h(x))$ have common integrable majorant $n \cdot \chi_{E_k \cap \Omega_m} + \langle \nabla \phi(x, u(x)), h(x) \rangle$, whence dominated convergence gives

$$\begin{aligned}
&\lim \int_\Omega \delta_t(x, h(x))\, \mu(dx) \\
&\quad = \int_\Omega \lim \delta_t(x, h(x))\, \mu(dx) \\
&\quad = \langle \nabla f(u), h \rangle =: \delta(h),
\end{aligned}$$

thus completing the proof of condition (α).

Concerning condition (β), let $t_n \to 0$ and $h_n \to h$ (norm) be fixed. Condition (3.5) allows for applying Fatou's Lemma, which gives

$$\liminf \int_\Omega (\delta_{t_n}(x, h_n(x)) + C|h_n(x)|^p + g_1(x) \cdot h_n(x) + g_0(x))\, \mu(dx) \geqslant$$
$$\int_\Omega \liminf (\delta_{t_n}(x, h_n(x)) + C|h_n(x)|^p +$$
$$+ g_1(x) \cdot h_n(x) + g_0(x))\, \mu(dx). \tag{3.7}$$

Now we fix a subsequence for which the liminf is attained, and then pass to another subsequence $h_{n'}$ which converges almost everywhere. As a consequence, the liminf of the integrand on the right hand side of (3.7) then equals $\nabla\phi(x, u(x)) \cdot h(x) + C|h(x)|^p + g_1(x) \cdot h(x) + g_0(x)$. This ends the proof of condition (β).

(b) We prove a little more on route. Assume that in an arbitrary Banach space E, we have $\delta_{f,u,t} \xrightarrow{e} \delta$ with a lower semi-continuous limit function δ, and where f is assumed continuous at u. Then we find $\alpha > 0$ and $t_0 > 0$ such that

$$\delta_{f,u,t}(h) + \alpha\|h\| \geq 0 \tag{3.8}$$

is satisfied for all $\|h\| \leq 1$, $|t| \leq t_0$. Indeed, assuming the contrary, we find $|t_n| \leq t_0$, $\|h_n\| \leq 1$ such that

$$\delta_{t_n}(h_n) + n\|h_n\| < 0$$

for all n. First assume $\|h_n\| \to 0$, at least for a subsequence. Find $\rho_n \to \infty$ such that $\rho_n h_n \to 0$, but $n\rho_n\|h_n\| \to \infty$. Then we have

$$\rho_n \delta_{t_n}(h_n) = \delta_{t_n/\rho_n}(\rho_n h_n) \to -\infty, \tag{3.9}$$

a contradiction, since $t_n/\rho_n \to 0$ and $\rho_n h_n \to 0$, whence the limit inferior of (3.9) ought to be minorized by $\delta(0) > -\infty$ as a consequence of condition (β) of epi convergence. Next consider the case where $\|h_n\| \geq \epsilon > 0$ for all n. Then $\delta_{t_n}(h_n) \to -\infty$, and since f is assumed continuous at u, this implies $t_n \to 0$. Now find $\sigma_n \to 0$ such that $n\sigma_n \to \infty$ and $t_n/\sigma_n \to 0$. Then we have

$$\sigma_n \delta_{t_n}(h_n) = \delta_{t_n/\sigma_n}(\sigma_n h_n) \to -\infty,$$

a contradiction, since by condition (β) of epi convergence the limit inferior of this expression ought to be minorized by $\delta(0) > -\infty$, as above. This proves (3.8).

Now let us consider the case $p = 1$. We claim that here (3.8) implies the estimate (3.5), (with $p = 1$). Indeed, assume the set of $x \in \Omega$ where (3.5) (with $g_1 = 0$, $g_0 = 0$) is violated has positive measure. Then we find $\eta > 0$ and some $\xi \in \mathbb{R}^d$ such that the set $\{x \in \Omega : \delta_t(x, \xi) + C|\xi| < -\eta\}$ has positive measure. Choose a subset A of this set having $0 < \mu(A) < t_0/|\xi|$, (μ has no atoms!), and define $h = \xi \cdot \chi_A$. Then (3.8) is violated for h, a contradiction. This proves the claim, and hence the necessity of (3.5) for the epi convergence of δ_t in the case $p = 1$. □

Notice that the above argument, carried out in the case $p = 1$, relies on the fact that the norm $\|\cdot\|_1$ is an integral functional. This is not the case for the $\|\cdot\|_p$ for $p > 1$, whence we do not know whether condition (3.5) is necessary in these cases.

4. Second-Order Theory – Convex Case

In this section we discuss Mosco convergence and Attouch–Wets convergence of the second-order difference quotient (2.1) of a convex function f. We show that both notions may be analyzed with the help of the Young–Fenchel conjugate. Here our main interest lies in studying integral functionals, so we focus on the Hilbert space case. Let us first recall the following result, obtained in [13], which deals with Mosco convergence of the second-order difference quotient (2.1).

PROPOSITION 4.1. *Let f be a continuous convex function on a separable Hilbert space H. Let $x \in H$, $y \in \partial f(x)$, $g = f + \frac{1}{2}\|\cdot\|^2$. Then the following statements are equivalent:*

(1) $\Delta_{f,x,y,t} \xrightarrow{m} q$ *for a purely quadratic convex limit function q;*
(2) $\Delta_{g,x,x+y,t} \xrightarrow{m} q + \frac{1}{2}\|\cdot\|^2$ *for a purely quadratic convex function q;*
(3) g^* *is second-order differentiable at $x + y$;*
(4) ∇g^* *is norm Gâteaux differentiable at $x + y$;*
(5) *The resolvent operator $J_{f^*} = (\mathrm{id} + \partial f^*)^{-1}$ is norm Gâteaux differentiable at $x + y$.*

Moreover, in these cases, the following are equivalent:

(a) *Statement 1. is true with* $\mathrm{dom}(q) = H$*;*
(b) g^* *is second-order differentiable at $x + y$, and $\|\nabla^2 g^*(x + y)\| < 1$;*
(c) J_{f^*} *is norm Gâteaux differentiable at $x + y$, and $\|\nabla J_{f^*}(x + y)\| < 1$.*

Remarks. (1) Here second-order differentiability of g^* at $x+y$ is to be understood in the sense that $x + y \in D^2_{g^*}$, which by convexity is equivalent to pointwise norm convergence of the difference quotient $(1/t)(\nabla g^*(x + y + th) - \nabla g^*(x + y))$ as $t \to 0$, (cf. [13]).

(2) It is important to observe here that the statement $\mathrm{dom}(q) = H$ does *not* imply that $x \in D^2_f$, i.e., the limit in (a) is not necessarily pointwise. See [13, §3] for counterexamples. This means that second-order Mosco differentiability does in fact represent a concept of generalized second-order differentiability which

deserves being studied at equal rights with classical second-order notions. In [13], we proposed the notation $x \in GD_f^2$ for statement (a) above.

(3) As we have seen in the previous section, on the first-order level, graphical convergence of $\delta_{f,x,t}$ plus an extra condition, saying that f be locally Lipschitz at x, already implies pointwise type convergence of $\delta_{f,x,t}$. In second-order theory, we may ask for a similar extra condition which allows for improving graphical convergence of $\Delta_{f,x,y,t}$ to get pointwise type convergence. This condition has been singled out in [13], and in terms of the function f it says that f has to be *Lipschitz smooth* at x. More formally, there must exist $C > 0$ and $\eta > 0$ such that

$$|f(x+h) - f(x) - \langle \nabla f(x), h \rangle| \leq C\|h\|^2$$

is satisfied for all $\|h\| \leq \eta$. However, while the local Lipschitz behaviour of f is a fairly general condition, this is not the case for Lipschitz smoothness, which explains to some extent why graphical convergence plays a more important role on the second order level.

Let us now establish a result similar to Proposition 4.1 providing a dual version for Attouch–Wets convergence of the second-order difference quotient. As we will see, in contrast with Mosco convergence, this will show that the use of Attouch–Wets convergence on the second-order level is much more limited. This is in contrast with the situation on the first-order level, where Attouch–Wets convergence is quite useful. Before getting started, we recall the following notion from [13]. A convex function f is said to be *strongly second-order differentiable* at x, if the second order difference quotient (2.1) converges uniformly on bounded sets to a purely quadratic and fully defined limit.

THEOREM 4.2. *Let f be a continuous convex function on a separable Hilbert space H. Let $x \in H$, $y \in \partial f(x)$, $g = f + \frac{1}{2}\|\cdot\|^2$. Then the following statements are equivalent:*

(1) $\Delta_{f,x,y,t} \xrightarrow{\text{aw}} q$ *for a purely quadratic convex q;*
(2) $\Delta_{g,x,x+y,t} \xrightarrow{\text{aw}} q + \frac{1}{2}\|\cdot\|^2$ *for a purely quadratic convex q;*
(3) g^* *is strongly second-order differentiable at $x + y$;*
(4) ∇g^* *is Fréchet differentiable at $x + y$;*
(5) J_{f^*} *is Fréchet differentiable at $x + y$.*

Moreover, in these cases, the following are equivalent

(a) *Statement (1) is true with* $\mathrm{dom}(q) = H$;
(b) f *is strongly second-order differentiable at x;*
(c) g^* *is strongly second-order differentiable at $x+y$ and $\|\nabla^2 g^*(x+y)\| < 1$;*
(d) J_{f^*} *is Fréchet differentiable at $x + y$, and we have $\|\nabla J_{f^*}(x + y)\| < 1$.*

Proof. Let us first prove (1) $\Leftrightarrow$ (2). As above, $g = f + \frac{1}{2}\|\cdot\|^2$. Observe that the Attouch- Wets topology is induced by the family of pseudo-metrics

$$d_\lambda(\phi, \psi) = |\phi_1(0) - \psi_1(0)| + \sup_{\|x\| \leq \lambda} \|J_\phi(x) - J_\psi(x)\|,$$

$\lambda > 0$, (see [5]). Now we are dealing with second-order difference quotients, or rather, with functions ϕ, ψ satisfying $\phi_1(0) = \psi_1(0) = 0$, and for these, we obtain the following useful relation:

$$d_\lambda(\phi + \tfrac{1}{2}\|\cdot\|^2, \psi + \tfrac{1}{2}\|\cdot\|^2) = d_{\frac{\lambda}{2}}(\tfrac{1}{2}\phi, \tfrac{1}{2}\psi),$$

which uses the equality $J_{\phi+\frac{1}{2}\|\cdot\|^2}(h) = J_{\frac{1}{2}\phi}(\frac{1}{2}h)$. From this it becomes clear that $\Delta_{f,x,y,t} \overset{\text{aw}}{\to} q$ is equivalent to

$$\Delta_{g,x,x+y,t} = \Delta_{f,x,y,t} + \Delta_{\frac{1}{2}\|\cdot\|^2,x,x,t} = \Delta_{f,x,y,t} + \tfrac{1}{2}\|\cdot\|^2 \overset{\text{aw}}{\to} q + \tfrac{1}{2}\|\cdot\|^2,$$

proving the equivalence of (1) and (2).

Next observe that (2) $\Leftrightarrow$ (3). Indeed, according to [3], Attouch–Wets convergence $\Delta_{f,x,y,t} \overset{\text{aw}}{\to} q$ is equivalent to uniform convergence $\Delta_{f,x,y,t} \,\square\, \frac{1}{2}\|\cdot\|^2 \to q \,\square\, \frac{1}{2}\|\cdot\|^2$ on bounded sets. Now observe that $\Delta_{f \,\square\, \frac{1}{2}\|\cdot\|^2, x+y,y,t} = \Delta_{f,x,y,t} \,\square\, \frac{1}{2}\|\cdot\|^2$, and recall Moreau's identity

$$\Delta_{f \,\square\, \frac{1}{2}\|\cdot\|^2, x+y,y,t} + \Delta_{g^*,x+y,x,t} = \tfrac{1}{2}\|\cdot\|^2,$$

(cf. [26]). Then $\Delta_{g^*,x+y,y,t}$ converges uniformly on bounded sets to the fully defined purely quadratic convex limit $q^* \,\square\, \frac{1}{2}\|\cdot\|^2$, which is just statement (3). An alternative reasoning to obtain the equivalence of (2) and (3) would be the following. Notice that (2) is equivalent to the statement $\Delta_{g^*,x+y,x,t} \overset{aw}{\to} (q + \frac{1}{2}\|\cdot\|^2)^* = q^* \,\square\, \frac{1}{2}\|\cdot\|^2$, since Attouch–Wets convergence is invariant under Young–Fenchel conjugation (see [5, 3]). But now observe that g^* is Lipschitz smooth at $x + y$, and hence $\Delta_{g^*,x+y,x,t}$ is uniformly bounded (see [13]), and hence equi-Lipschitzian, and this implies uniform convergence of $\Delta_{g^*,x+y,x,t}$ on bounded sets.

The equivalence of (3) and (4) is just [13, Theorem 3.1], which relates convergence of the second-order difference quotient of f to convergence of the first-order difference quotient of ∂f. Finally, the equivalence of (4) and (5) is immediate from Moreau's identity $\nabla g^* + J_{f^*} = \text{id}$. (Notice that $\nabla g^* = J_f$.)

Let us now prove the additional statement. The equivalence of (a), (c) and (d) follows from the corresponding part of Proposition 4.1, since Attouch–Wets convergence implies Mosco convergence. So we are left to prove that statement (d) implies (b). But this is a consequence of Lemma 7.4(3) in [13], which tells that in the case of Fréchet differentiability of J_{f^*} at $x+y$, $\|\nabla J_{f^*}(x+y)\| < 1$ is equivalent to saying that the *one-sided Lipschitz constant*

$$b = b(J_{f^*}, x+y) = \lim_{\delta\to 0} \sup_{0<\|h\|\leq\delta} \frac{\langle J_{f^*}(x+y+h) - J_{f^*}(x+y), h\rangle}{\|h\|^2}$$

of J_{f^*} at $x+y$ satisfies $b < 1$ (see [13, §7]). The latter, when combined with statement (3), implies pointwise convergence $\Delta_{f,x,y,t} \to q$ [13, Theorem 7.5], which in tandem with Attouch–Wets convergence of $\Delta_{f,x,y,t}$ gives uniform convergence on bounded sets. This ends the proof of the addendum. □

Remarks. (1) In two places we used the following fact: Suppose $f_n \overset{\text{aw}}{\to} f$, and that the sequence (f_n) is locally equi-Lipschitzian. Then convergence is uniform on bounded sets, and f is fully defined. The reasoning is as in the proof of Proposition 3.3.

(2) Theorem 4.2 shows that the use of Attouch-Wets convergence for second-order difference quotients is limited. Namely, even in a separable Hilbert space, a locally Lipschitz operator $T : H \to H$ cannot in general be expected to have points of Fréchet differentiability, while by the work of N. Aronszajn [1], F. Mignot [25] and others, it is known that such T has sufficiently many points of Gâteaux differentiability.

(3) Theorem 4.2 has an application to the differentiability of the metric projection P_C onto a closed convex set in Hilbert space. It gives the main step towards proving the fact that P_C is Fréchet differentiable at a point $x \notin C$ if and only if the boundary of C is second-order Attouch–Wets smooth at $P_C x$. See [31] for details.

EXAMPLE. We produce a convex integral functional f on the Hilbert space $L^2[0,1]$ where J_{f^*} is nowhere Fréchet differentiable, and hence $\Delta_{f,x,y,t}$ for *no* x and $y \in \partial f(x)$ is Attouch–Wets convergent.

Let $C = \{x \in L^2 : \|x\|_\infty \leq 1\}$, and let P_C be the orthogonal projection onto C. Then, according to [21, §5], P_C is nowhere Fréchet differentiable. Let f be the support function of the set C, then it follows that $J_{f^*} = P_C$, whence f is as desired. Notice that f is in fact an integral functional, namely, $f = \|\cdot\|_1$, considered as a function on L^2.

An even better example which is of class $C^{1,1}$ is obtained by taking the function $f_1 = f \,\square\, \frac{1}{2}\|\cdot\|^2$, which by the Moreau identities has Fréchet derivative $\nabla f_1 = J_{f^*} = P_C$, and which is given by the formula (cf. [21]):

$$f_1(x) = \tfrac{1}{2}\|x\|^2 - \tfrac{1}{2}\|x - P_C x\|^2.$$

Indeed, f_1 is everywhere Lipschitz smooth, hence Attouch–Wets convergence of $\Delta_{f_1,z,t}$ at any z would imply uniform convergence on bounded sets, hence Fréchet differentiability of ∇f_1 at z, which was seen to be impossible.

5. Second-Order Theory – General Case

Dealing with nonconvex functions, it is clear that Mosco convergence and Attouch-Wets convergence have to be replaced by a more flexible notion of graphical convergence which does not essentially rely on the weak lower semi-continuity of the function f under consideration. It turns out that the right choice is epi convergence with respect to the norm topology. The following characterization of second-order epi differentiability for integral functionals (1.1), (1.2) on Hilbert space $L^2_{\mathbb{R}^d}$ was obtained in [29].

THEOREM 5.1. *Let f be an integral functional on $L^2_{\mathbb{R}^d}$ satisfying (1.1), (1.2). Let $u \in L^2_{\mathbb{R}^d}$ be fixed satisfying $u(x) \in D^2_{\phi(x,\cdot)}$ for almost all x. Then the following statements are equivalent:*

(1) *f is second-order epi differentiable at u, i.e., $\Delta_{f,u,v,t} \xrightarrow{e} q$ for a purely quadratic and fully defined limit function q, $(v(x) = \nabla\phi(x, u(x)))$;*

(2) *The eigenvalues of the Hessian matrices $\nabla^2\phi(x, u(x))$ are essentially uniformly bounded, and moreover, the second-order difference quotients $\Delta_t(x,\cdot) = \Delta_{\phi(x,\cdot),u(x),v(x),t}$ of the $\phi(x,\cdot)$ at $u(x)$ satisfy the following uniform boundedness below condition:*

$$\Delta_t(x,\xi) + \alpha|\xi|^2 \geq 0 \tag{5.1}$$

for some $\alpha > 0$, all $|t| \leq 1$, $\xi \in \mathbb{R}^d$ and almost all x.

Moreover, in these cases, q is given by

$$q(h) = \tfrac{1}{2}\int_\Omega \langle \nabla^2\phi(x, u(x))h(x), h(x)\rangle\, \mu\,(dx). \tag{5.2}$$

The proof, which may be found in [29], proceeds in a way similar to the proof of Proposition 3.4. Notice that in [29] we used this result as a basic tool to discuss the second-order differentiability properties of the integral functionals (1.1), (1.2) having smooth integrand.

Notice that Theorem 5.1 adjusts the result by R.S. Palais and S. Smale (cf. [32]) mentioned in the introduction. A corrected version of their result was also given in I.V. Skrypnik (cf. [6, p. 25]), stating that the only integral functionals of class C^2 are those having integrand a polynomial of degree ≤ 2. See [8] for a proof of this fact and the related Theorem of Vainberg.

As a consequence of Theorem 5.1, we obtain the important fact that Mosco convergence and epi convergence of the second-order difference quotients coincide when f is a convex integral functional on an L^2-space. This is strongly in contrast with the first order theory, where Mosco convergence certainly coincides with Attouch-Wets convergence when the space is reflexive.

COROLLARY 5.2. *Let f be a convex integral functional on $L^2_{\mathbb{R}^d}$. Let $u \in L^2_{\mathbb{R}^d}$ be fixed. Then the following statements are equivalent:*

(1) *$\nabla^2\phi(x, u(x))$ exists for almost all x, with eigenvalues being essentially bounded;*
(2) *$\Delta_{f,u,v,t} \xrightarrow{e} q$ for a purely quadratic and fully defined q;*
(3) *$\Delta_{f,u,v,t} \xrightarrow{m} q$ for a purely quadratic and fully defined q.*

Moreover, in these cases, q has the form (5.2).

Proof. Notice that (1) and (2) are equivalent as a consequence of Theorem 5.1, since the uniform boundedness below condition (5.1) is automatically satisfied with any $\alpha > 0$, as ϕ is convex, and hence $\Delta_t(x, \cdot) \geq 0$ almost everywhere. Since clearly (3) implies (2), it remains to observe that (1) and (3) are equivalent. This was proved in [13] using duality techniques based on Proposition 4.1. □

We do not know whether Corollary 5.2 remains valid for more general classes of continuous convex functions defined on a separable Hilbert space.

References

1. Aronszajn, N.: Differentiability of Lipschitz mappings between Banach spaces, *Studia Math.* **LVII** (1976), 147–190.
2. Attouch, H.: Familles d'opérateurs maximaux monotones et mesurabilité, *Ann. Mat. Pura Appl.* **120** (1979), 35–111.
3. Attouch, H., Ndoutoume, J. L., and Théra, M.: On the equivalence between epi-convergence of sequences of functions and graph convergence of their derivatives, to appear.
4. Attouch, H. and Wets, R. J.-B.: A convergence theory for saddle functions, *Trans. Amer. Math. Soc.* **280** (1983), 1–41.
5. Attouch, H. and Wets, R. J.-B.: Isometries for the Legendre-Fenchel transform, *Trans. Amer. Math. Soc.* **296** (1986), 33–60.
6. Aubin, J.-P. and Ekeland, I.: *Applied Nonlinear Analysis*, Wiley, New York, 1984.
7. Bangert, V.: Analytische Eigenschaften konvexer Funktionen auf Riemannschen Mannigfaltigkeiten, *J. reine angew. Math.* **307** (1979), 309–324.
8. Blot, J.: Démonstration des théorèmes de Vainberg et Skrypnik, Publications du Département de Mathématiques de l'Université de Limoges.
9. Borwein, J. M.: Asplund spaces are sequentially reflexive, to appear.
10. Borwein, J. M. and Fabian, M.: On convex functions having points of Gâteaux differentiability which are not points of Fréchet differentiability' to appear.
11. Borwein, J. M. and Fabian, M.: Generic second order Gâteaux differentiability of convex functions, to appear.
12. Borwein, J. M., Fabian, M., and Vanderwerff, J.: Locally Lipschitz functions and bornological derivatives, to appear.

13. Borwein, J. M. and Noll, D.: Second order differentiability of convex functions in Banach spaces, *Trans. Amer. Math. Soc.* to appear.
14. Borwein, J. M. and Preiss, D.: A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, *Trans. Amer. Math. Soc.* **303** (1987), 517–527.
15. Busemann, H.: *Convex Surfaces*, Interscience Publishers, New York 1955.
16. Deville, R., Godefroy, G., and Zizler, V.: A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions, to appear.
17. Do, C. N.: Generalized second-order derivatives of convex functions in reflexive Banach spaces, *Trans. Amer. Math. Soc.* to appear.
18. Dolecki, S., Salinetti, G., and Wets, R.: Convergence of functions: equi-semicontinuity, *Trans. Amer. Math. Soc.* **276** (1983), 409–429.
19. Ekeland, I.: *Convexity Methods in Hamiltonian Mechanics*, Springer-Verlag, New York, 1989.
20. Fabian, M.: Lipschitz smooth points of convex functions and isomorphic characterization of Hilbert spaces, *Proc. London Math. Soc.* **751** (1985), 113–126.
21. Fitzpatrick, S. and Phelps, R. R.: Differentiability of the metric projection in Hilbert space, *Trans. Amer. Math. Soc.* **270** (1982), 483–501.
22. Hiriart–Urruty, J. B. and Seeger, A.: 'Calculus rules on a new set-valued second derivative for convex functions, *Nonlinear Anal. Theory Meth. Appl.* **13** (1989), 721–738.
23. Kato, N.: On the second derivatives of convex functions in Hilbert space, *Proc. Amer. Math. Soc.* **106** (1989).
24. Loewen, P. D. and Zeng, H.: Epi-derivatives of integral functionals with applications, to appear.
25. Mignot, F.: Contrôle dans les inéquations variationelles elliptiques, *J. Funct. Anal.* **22** (1976), 130–185.
26. Moreau, J. J.: Proximité et dualité dans un espace Hilbertien, *Bull. Soc. Math. France* **93** (1965), 273–299.
27. Ndoutoume, J. L.: Calcul différentiel généralisé du second ordre, *Publ. AVAMAC, Univ. de Perpignan*, 1987.
28. Ndoutoume, J. L. and Théra, M.: Generalized second-order derivatives of convex functions in locally convex topological vector spaces, to appear.
29. Noll, D.: Second order differentiability of integral functionals on Sobolev spaces and L^2-spaces, *J. reine angew. Math.* **436** (1993), 1–17.
30. Noll, D.: Generalized second fundamental form for Lipschitzian hypersurfaces by way of second epi deivatives, *Canadian Math. Bull.* **35**(4) (1992), 523–536.
31. Noll, D.: Directional differentiability of the metric projection in Hilbert space, *Pacific J. Math.*, to appear.
32. Palais, R. S. and Smale, S.: A generalized Morse Theory, *Bull. Amer. Math. Soc.* **70** (1964), 165–172.
33. Rockafellar, R. T.: *Conjugate Duality and Optimization*, SIAM Publ., Philadelphia, 1974.
34. Rockafellar, R. T.: Maximal monotone relations and the second derivatives of nonsmooth functions, *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **2** (1985), 167–184.
35. Rockafellar, R. T.: Generalized second derivatives of convex functions and saddle functions, *Trans. Amer. Math. Soc.* **322** (1990), 51–77.
36. Rockafellar, R. T.: Second order optimality conditions in non-linear programming obtained by way of epi derivatives, *Math. Oper. Res.* **14** (1989), 462–484.
37. Salinetti, G. and Wets, R. J.-B.: On the relation between two types of convergence for convex functions, *J. Math. Anal. Appl.* **60** (1977), 211–226.
38. Skrypnik, I. V.: On the application of Morse's method to nonlinear elliptic equations, *Dokl. Akad. Nauk SSSR* **202** (1972), 202ff.