TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 347, Number 2, February 1995

EPI-DERIVATIVES OF INTEGRAL FUNCTIONALS WITH APPLICATIONS

PHILIP D. LOEWEN AND HARRY H. ZHENG

ABSTRACT. We study first- and second-order epi-differentiability for integral functionals defined on $L^2[0, T]$, and apply the results to obtain first- and second-order necessary conditions for optimality in free endpoint control problems.

1. INTRODUCTION

The notion of epi-differentiability for nonsmooth functions with extended real values was introduced and developed by Rockafellar [12, 13]. He shows [12, Theorem 4.5] that a sufficient condition for $f: \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ to be twice epi-differentiable at the point x is the existence of a local representation $f = g \circ F$ in which $F: \mathbb{R}^m \to \mathbb{R}^d$ is a C^2 mapping and $g: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a piecewise linear-quadratic convex function such that $g(F(x))$ is finite and a certain constraint qualification holds. Under these conditions f is called *fully amenable at* x. A large class of finite-dimensional optimization problems can be described in terms of fully amenable functions. Poliquin and Rockafellar [9, 10] develop a calculus of epi-derivatives and apply it to derive optimality conditions for certain mathematical programming problems.

Epi-differentiability has also been studied in reflexive Banach spaces. Do [5] treats convex integral functionals, while Cominetti [4] considers general amenable functions. A discussion of their results, of Noll's recent work [8], and of Levy's concurrent research on the same topic [6], appears at the end of Section 3.

In this paper, we study the epi-differentiability of the extended-valued functional $\mathcal{J}$ defined on $L^2([0, T]; \mathbb{R}^m)$ by

$$\mathcal{J}(u) = \int_0^T f(t, u(t))\, dt,$$

where the integrand $f(t, \cdot)$ is, for almost all t, a fully amenable function of a particular form, namely, the sum of a finite-valued fully amenable function and an indicator function. (See Section 3 for details.) We show that the the

Received by the editors January 5, 1993.

1991 *Mathematics Subject Classification*. Primary 49J52, 49K05; Secondary 58C20, 58C05.

Key words and phrases. Integral functionals, epiderivative, epiconvergence.

The first author's research was supported by the Natural Science and Engineering Research Council (NSERC) of Canada through grant 5-89441.

first and second epi-derivatives of $\mathcal{I}$ can be expressed in terms of the epi-derivatives of $f(t, \cdot)$—effectively justifying the operation of "differentiating under the integral sign". This reduces our infinite-dimensional differentiation problem to a finite-dimensional one for which results are already available, and lets us obtain optimality conditions for minimization problems involving $\mathcal{I}$. We also discuss the epi-differentiability of simple composite functionals involving $\mathcal{I}$, particularly Bolza functionals. We then apply our results to derive second-order necessary conditions for free endpoint control problems. Since epi-differentiability has a strong geometric foundation and can capture the local behavior of integral functionals near a given point, this method may also be useful in studying optimality conditions for problems with constraints on both endpoints.

We begin with some background material, then study the epi-differentiability of the nonconvex integral functional $\mathcal{I}$, and finally apply our results to a Bolza functional.

2. Background material

We write $\overline{\mathbb{R}}$ for the extended real line $\mathbb{R} \cup \{+\infty\}$. Throughout this section, $(X, \|\cdot\|)$ is a normed linear space; a function $f: X \to \overline{\mathbb{R}}$ is called *proper* when its domain, the set $\operatorname{dom} f = \{x \in X : f(x) < +\infty\}$, is nonempty. In cases where X is finite-dimensional, we think of its elements as column vectors and the elements of X^* (e.g., gradients and normal vectors) as row vectors.

Given an extended-real-valued function f on X and a point x in the domain of f, the first difference quotient for f at x is the function $f_{x,h}: X \to \overline{\mathbb{R}}$ defined by

$$f_{x,h}(y) = \frac{f(x+hy) - f(x)}{h}.$$

The epi-differentiability of f at x is characterized by the equation

$$\limsup_{h \downarrow 0} \operatorname{epi} f_{x,h} = \liminf_{h \downarrow 0} \operatorname{epi} f_{x,h} = \operatorname{epi} f'_x. \tag{1}$$

(Recall that for any $g: X \to \overline{\mathbb{R}}$, $\operatorname{epi} g = \{(x, r) : x \in \operatorname{dom} g,\ r \geq g(x)\}$.) We say that f is *epi-differentiable* at x if the first equation holds and the function f'_x defined by the second equation obeys $f'_x(0) > -\infty$. In this case the function f'_x is lower semicontinuous and satisfies $f'_x(0) = 0$: it is called the *epi-derivative of f at x*. This is the Kuratowski-Painlevé mode of epi-convergence, in which the strong topology is used throughout. Thus criteria (I) and (II) below are both equivalent to (1).

(I) For any $y \in X$ and any sequence $h_i \downarrow 0$ we have both
  - (i) Any sequence $y_i \to y$ obeys $\liminf_{i\to\infty} f_{x,h_i}(y_i) \geq f'_x(y)$; and
  - (ii) There exists a sequence $y_i \to y$ such that $\limsup_{i\to\infty} f_{x,h_i}(y_i) \leq f'_x(y)$.

(II) For all $y \in X$, we have

$$\liminf_{h\downarrow 0} \inf_{y' \to y} f_{x,h}(y') = \limsup_{h\downarrow 0} \inf_{y' \to y} f_{x,h}(y') = f'_x(y)$$

where, by definition,

$$\liminf_{h\downarrow 0}\inf_{y'\to y} f_{x,h}(y') = \lim_{\epsilon\downarrow 0}\liminf_{h\downarrow 0}\inf_{\|y'-y\|<\epsilon} f_{x,h}(y'),$$

$$\limsup_{h\downarrow 0}\inf_{y'\to y} f_{x,h}(y') = \lim_{\epsilon\downarrow 0}\limsup_{h\downarrow 0}\inf_{\|y'-y\|<\epsilon} f_{x,h}(y').$$

To arrive at a third characterization of epi-differentiability, we recall that any proper function $F: X \to \overline{\mathbb{R}}$ gives rise to a family $\{F^\lambda\}_{\lambda>0}$ of Moreau-Yosida approximates, where

$$F^\lambda(x) = \inf_{u\in X}\left\{F(u) + \lambda\|u-x\|^2\right\} \quad \forall x \in X.$$

Note that $F^\lambda(x)$ is nondecreasing as a function of λ. The following result of Attouch [1, Thm. 2.65] relates compound limits of the sort in criterion (II) to iterated limits involving the Moreau-Yosida approximates.

Theorem 1. *Let each $F_i: X \to \overline{\mathbb{R}}$ be proper. If there exist $r \geq 0$, $x_0 \in X$, such that*

$$\inf_i F_i(x) \geq -r(\|x-x_0\|^2 + 1) \qquad \forall x \in X,$$

then the following equations hold for every x in X:

$$\liminf_{i\to\infty}\inf_{x'\to x} F_i(x') = \lim_{\lambda\to\infty}\liminf_{i\to\infty} F_i^\lambda(x),$$

$$\limsup_{i\to\infty}\inf_{x'\to x} F_i(x') = \lim_{\lambda\to\infty}\limsup_{i\to\infty} F_i^\lambda(x).$$

It follows from Attouch's theorem that if $x \in \operatorname{dom} f$ is a point where each difference quotient $f_{x,h}$ $(h > 0)$ is proper and there exist $r \geq 0$ and $y_0 \in X$ such that

$$\inf_{h>0} f_{x,h}(y) \geq -r(\|y-y_0\|^2+1) \qquad \forall y \in X, \tag{2}$$

then the epi-derivative f'_x can be characterized as follows:

(III) $\displaystyle\lim_{\lambda\to\infty}\liminf_{h\downarrow 0} f^\lambda_{x,h}(y) = \lim_{\lambda\to\infty}\limsup_{h\downarrow 0} f^\lambda_{x,h}(y) = f'_x(y) \qquad \forall y \in X.$

Now suppose that f is epi-differentiable at x. A vector $w \in X^*$ is called an *epi-gradient of f at x* if

$$\langle w\ ,\ y\rangle \leq f'_x(y) \qquad \forall y \in X.$$

The set of vectors w satisfying this condition is written $\partial f(x)$. For any w in $\partial f(x)$, the second difference quotient of f at x relative to w is the function $f_{x,w,h}: X \to \overline{\mathbb{R}}$ defined by

$$f_{x,w,h}(y) = \frac{f(x+hy) - f(x) - h\langle w\ ,\ y\rangle}{h^2/2}.$$

The function f is called *twice epi-differentiable* at x relative to w with *second epi-derivative* $f''_{x,w}$ if the following equations hold with $f''_{x,w}(0) > -\infty$:

$$\limsup_{h\downarrow 0}\operatorname{epi} f_{x,w,h} = \liminf_{h\downarrow 0}\operatorname{epi} f_{x,w,h} = \operatorname{epi} f''_{x,w}. \tag{3}$$

Again, the first equation is the epi-differentiability criterion, while the second defines the function $f''_{x,w}$. There are alternative characterizations of $f''_{x,w}$, just as there are of f'_x.

The simplest interesting class of epi-differentiable functions is the set of convex functions. Rockafellar [12] proves the following results. Suppose $X = \mathbb{R}^l$ and $\sigma: X \to \mathbb{R}$ is convex and finite-valued. Then σ is epi-differentiable at any point $\widehat{\alpha}$, where its epi-derivative coincides with the directional derivative

$$\sigma'_{\widehat{\alpha}}(\alpha) = \lim_{h\downarrow 0} \frac{\sigma(\widehat{\alpha}+h\alpha)-\sigma(\widehat{\alpha})}{h}. \tag{4}$$

This is the support function of the set of epigradients $\partial\sigma(\widehat{\alpha})$, which agrees with the usual subdifferential of convex analysis. A sufficient condition for the convex function σ to be twice epi-differentiable at $\widehat{\alpha}$ is that it be *piecewise linear-quadratic* near $\widehat{\alpha}$, i.e., that some open cube centred at $\widehat{\alpha}$ admit a decomposition into finitely many polyhedral cells, in each of which σ is either quadratic or affine. In this case any γ in $\partial\sigma(\widehat{\alpha})$ obeys

$$\sigma''_{\widehat{\alpha},\gamma}(\alpha) = \sigma''_{\widehat{\alpha}}(\alpha) + \Psi_{\Sigma(\gamma)}(\alpha),$$

where

$$\begin{aligned} \sigma''_{\widehat{\alpha}}(\alpha) &= \lim_{h\downarrow 0} \frac{\sigma(\widehat{\alpha}+h\alpha)-\sigma(\widehat{\alpha})-h\sigma'_{\widehat{\alpha}}(\alpha)}{h^2/2}, \\ \Sigma(\gamma) &= \left\{\alpha : \langle\alpha\,,\,\gamma\rangle = \sigma'_{\widehat{\alpha}}(\alpha)\right\}. \end{aligned} \tag{5}$$

At the other extreme, consider a closed convex subset C of $\mathbb{R}^k$, and the indicator function $\Psi_C: \mathbb{R}^k \to \overline{\mathbb{R}}$ defined by setting $\Psi_C(\alpha) = 0$ if $\alpha \in C$, and $\Psi_C(\alpha) = +\infty$ otherwise. This function is also epi-differentiable at any point of its domain: given $\widehat{\alpha} \in C$, we have

$$(\Psi_C)'_{\widehat{\alpha}}(\alpha) = \Psi_{T_C(\widehat{\alpha})}(\alpha) \qquad \forall\alpha\in\mathbb{R}^k,$$

where $T_C(\widehat{\alpha})$ is the usual closed convex cone of tangents to the set C at the point $\widehat{\alpha}$. The epi-derivative is the support function of the normal cone (whose elements are row vectors)

$$N_C(\widehat{\alpha}) = \left\{\zeta\in(\mathbb{R}^k)^* : \langle\zeta\,,\,\alpha-\widehat{\alpha}\rangle \le 0\ \forall\alpha\in C\right\}. \tag{6}$$

The second epiderivative of Ψ_C is sure to exist if C is polyhedral, i.e., the intersection of finitely many closed half-spaces.

After proving the preceding statements about convex functions, Rockafellar [12] generalizes them to sums and smooth compositions. As before, let $\sigma: \mathbb{R}^l \to \mathbb{R}$ be a finite-valued convex function, and let $C \subseteq \mathbb{R}^k$ be a closed convex set. Let $F: \mathbb{R}^m \to \mathbb{R}^l$ and $G: \mathbb{R}^m \to \mathbb{R}^k$ be continuously differentiable, and consider the function

$$f(u) = \sigma(F(u)) + \Psi_C(G(u)).$$

The domain of f is the set $U = \{u\in\mathbb{R}^m : G(u)\in C\}$: we call points in U *feasible.* Associated with each feasible point u is the following constraint qualification:

$$\text{If } \eta\in N_C(G(u)) \text{ and } \eta G'(u) = 0, \text{ then } \eta = 0. \tag{CQ}$$

This is well known as the dual form of the Mangasarian-Fromovitz constraint qualification. It implies that the tangent cone to U at u is

$$T_U(u) = \{v\in\mathbb{R}^m : G'(u)v \in T_C(G(u))\}. \tag{7}$$

We now quote Rockafellar [12, Theorem 4.5], calling on the notation of (4)–(7).

Theorem 2. *If u is feasible and satisfies (CQ), then f is both epi-differentiable and (Clarke-) subdifferentially regular at u, with*

$$f'_u(v) = \sigma'_{F(u)}(F'(u)v) + \Psi_{T_U(u)}(v). \tag{8}$$

The function f'_u is the support function of the epigradient set

$$\partial f(u) = \partial\sigma(F(u))F'(u) + N_C(G(u))G'(u). \tag{9}$$

If in addition F and G are C^2, σ is piecewise linear quadratic, and C is polyhedral, then f is twice epi-differentiable at u relative to any $w \in \partial f(u)$, with

$$f''_{u,w}(v) = \sigma''_{F(u)}(F'(u)v) + \max_{(\gamma,\eta)\in\Gamma(u,w)} \{v^T[\gamma F + \eta G]''(u)v\} + \Psi_{\Sigma(u,w)}(v), \tag{10}$$

where

$$\Gamma(u, w) = \{(\gamma, \eta) \in \partial\sigma(F(u)) \times N_C(G(u)) : \gamma F'(u) + \eta G'(u) = w\}$$

is a nonempty, bounded, polyhedral convex set, and

$$\Sigma(u, w) = \left\{v \in \mathbb{R}^m : v \in T_U(u), \quad \sigma'_{F(u)}(F'(u)v) = \langle w, v\rangle\right\}$$

is a polyhedral convex cone.

Note that the theorem's assertion of Clarke regularity implies that the epigradient set $\partial f(u)$ coincides with Clarke's generalized gradient of f at u (see [3]), which is given in the regular case as the set of (row vectors) ζ for which

$$\liminf_{v\to u} \frac{f(v) - f(u) - \langle\zeta, v-u\rangle}{\|v-u\|} \geq 0.$$

3. Epi-derivatives of Integral Functionals

Consider the following integral functional defined on $L^2([0, T]; \mathbb{R}^m)$:

$$\mathscr{I}(u) = \int_0^T f(t, u(t))\,dt.$$

Here the integrand has the form $f(t, u) = \sigma(F(t, u)) + \Psi_C(G(t, u))$, where the functions $F: [0, T] \times \mathbb{R}^m \to \mathbb{R}^l$ and $G: [0, T] \times \mathbb{R}^m \to \mathbb{R}^k$ are measurable in t and continuously differentiable in u, the function σ is convex and finite-valued on $\mathbb{R}^l$, and the set C in $\mathbb{R}^k$ is nonempty, closed and convex. We write $U(t) = \{u \in \mathbb{R}^m : G(t, u) \in C\}$, and note that every u in L^2 for which $\mathscr{I}(u)$ is finite must obey $u(t) \in U(t)$ almost everywhere. Throughout this section we deal with a fixed function u in L^2 with $\mathscr{I}(u)$ finite. Our main results are Theorem 3, which deals with first-order derivatives, and Theorem 5, which treats the second-order case.

In the remainder of the paper, we reserve the notation $\langle\cdot, \cdot\rangle$ for the inner product in L^2; inner products in other spaces will be labelled explicitly, and finite-dimensional inner products will be indicated by simple juxtaposition.

Theorem 3. *Suppose there exist a constant $c \geq 0$ and an integrable function ϕ such that for almost all t in $[0, T]$, $u(t) \in U(t)$ satisfies both the constraint qualification (CQ) and the growth condition*

$$\inf_{h>0} \frac{\sigma(F(u(t)+hx)) - \sigma(F(u(t)))}{h} \geq \phi(t) - c|x|^2 \quad \forall x \in \mathbb{R}^m.$$

Then $\mathcal{I}$ is epi-differentiable at u, with epiderivative given by (see (8))

$$\mathcal{I}_u'(v) = \int_0^T f'_{t,u(t)}(v(t))\,dt \quad \forall v \in L^2. \tag{11}$$

The proof of Theorem 3 relies on some standard material from the theory of integral functionals. (See Rockafellar [11, Theorems 2A and 3A].) Consider a function $f : [0, T] \times \mathbb{R}^m \to \overline{\mathbb{R}}$ for which $u \mapsto f(t, u)$ is lower semicontinuous for each fixed t. Such an f is called a *normal integrand* if the epigraph multifunction $t \rightrightarrows \operatorname{epi} f(t, \cdot)$ is measurable. This is true if and only if f itself is $\mathcal{L} \times \mathcal{B}$ measurable. If f is a normal integrand for which the left side of equation (12) below is finite, then the equation itself is valid:

$$\inf_{u \in L^2([0,T];\mathbb{R}^m)} \int_0^T f(t, u(t))\,dt = \int_0^T \inf_{u \in \mathbb{R}^m} f(t, u)\,dt. \tag{12}$$

Proof of Theorem 3. By hypothesis, the first difference quotient obeys the estimate below:

$$\begin{aligned}\mathcal{I}_{u,h}(v) &\geq \int_0^T \frac{\sigma(F(t, u(t)+hv(t))) - \sigma(F(t, u(t)))}{h}\,dt \\ &\geq \int_0^T \phi(t)\,dt - c\|v\|_2^2.\end{aligned}$$

Thus $\mathcal{I}_{u,h}$ is proper and condition (2) holds. By criterion (III) in Section 2, the epi-differentiability of $\mathcal{I}$ at u is equivalent to

$$\lim_{\lambda\to\infty} \liminf_{h\downarrow 0} \mathcal{I}^\lambda_{u,h}(v) = \lim_{\lambda\to\infty} \limsup_{h\downarrow 0} \mathcal{I}^\lambda_{u,h}(v). \tag{‡}$$

For each $\lambda > 0$, we have

$$\begin{aligned}\mathcal{I}^\lambda_{u,h}(v) &= \inf_{x\in L^2} \left\{\mathcal{I}_{u,h}(x) + \lambda\|x - v\|_2^2\right\} \\ &= \inf_{x\in L^2} \int_0^T \left[f_{t,u(t),h}(x(t)) + \lambda|x(t) - v(t)|^2\right] dt.\end{aligned}$$

Now σ and Ψ_C are normal integrands, and $(t, x) \mapsto (F, G)(t, u(t)+hx)$ is measurable in t and continuous in x, so the function $(t, x) \mapsto f_{t,u(t),h}(x)$ is a normal integrand by [11, Proposition 2N]. Note that taking $x = 0$ gives

$$\inf_{x\in L^2} \int_0^T \left[f_{t,u(t),h}(x(t)) + \lambda|x(t) - v(t)|^2\right] dt \leq \lambda \int_0^T |v(t)|^2\,dt < \infty.$$

We may therefore apply (12) to take the infimum under the integral sign above and deduce that $\mathcal{I}^\lambda_{u,h}(v) = \int_0^T f^\lambda_{t,u(t),h}(v(t))\,dt$. But

$$f_{t,u(t),h}(x) \geq \phi(t) - c|x|^2 \geq k_1(t) - 2c|x - v(t)|^2,$$

where $k_1(t) := \phi(t) - 2c|v(t)|^2$ is integrable. Thus whenever $\lambda \geq 2c$ we have

$$k_1(t) \leq f^\lambda_{t,u(t),h}(v(t)) = \inf_{x\in\mathbb{R}^m} \left\{ f_{t,u(t),h}(x) + \lambda|x - v(t)|^2 \right\} \leq \lambda|v(t)|^2.$$

By Fatou's lemma (assuming $\lambda \geq 2c$) we have

$$\liminf_{h\downarrow 0} \mathscr{I}^\lambda_{u,h}(v) \geq \int_0^T \liminf_{h\downarrow 0} f^\lambda_{t,u(t),h}(v(t))\, dt.$$

For any fixed t, the integrand on the right is nondecreasing as a function of λ, and is bounded below by $k_1(t)$. Thus the monotone convergence theorem applies, to give

$$\lim_{\lambda\to\infty} \liminf_{h\downarrow 0} \mathscr{I}^\lambda_{u,h}(v) \geq \int_0^T \lim_{\lambda\to\infty} \liminf_{h\downarrow 0} f^\lambda_{t,u(t),h}(v(t))\, dt = \int_0^T f'_{u(t)}(v(t))\, dt.$$

Similarly, we have

$$\lim_{\lambda\to\infty} \limsup_{h\downarrow 0} \mathscr{I}^\lambda_{u,h}(v) \leq \int_0^T \lim_{\lambda\to\infty} \limsup_{h\downarrow 0} f^\lambda_{t,u(t),h}(v(t))\, dt = \int_0^T f'_{u(t)}(v(t))\, dt.$$

Combining these two inequalities confirms (‡). □

Here is an immediate consequence of Theorem 3 in terms of epigradients.

Corollary 4. *Under the conditions of* 3, *the set of epigradients of* $\mathscr{I}$ *at* u *is*

$$\partial\mathscr{I}(u) = \left\{ w \in L^2 \,:\, w(t) \in \partial f(t, u(t)) \ \text{a.e.}\ t \in [0, T] \right\}.$$

Here $\partial f(t, u(t))$ *refers to the set of epigradients in* u *for fixed* t; *see* (9).

Proof. If $w \in L^2$ satisfies $w(t) \in \partial f(t, u(t))$ a.e., then we have

$$\mathscr{I}'_u(v) = \int_0^T f'_{t,u(t)}(v(t))\, dt \geq \int_0^T w(t)v(t)\, dt = \langle w\,,\, v\rangle$$

for any v in L^2, so w lies in $\partial\mathscr{I}(u)$.

On the other hand, if w belongs to $\partial\mathscr{I}(u)$, then for any $v \in L^2$ we have

$$\langle w\,,\, v\rangle \leq \mathscr{I}'_u(v) = \int_0^T f'_{t,u(t)}(v(t))\, dt,$$

that is,

$$(*) \qquad \inf_{v\in L^2} \int_0^T \left[f'_{t,u(t)}(v(t)) - w(t)v(t) \right] dt \geq 0.$$

Now $(f)'_{t,u(t)}(x) = g'_{(F,G)(t,u(t))}((F, G)'(t, u(t))x) = \lim_{h\downarrow 0} g_h(t, x)$, where

$$g_h(t, x) = \frac{\sigma(F(t, u(t) + hx)) - \sigma(F(t, u(t))) + \Psi_C(G(t, u(t) + hx))}{h}.$$

Since g_h is a normal integrand, it must be $\mathscr{L} \times \mathscr{B}$ measurable. Therefore $(f)'_{t,u(t)}(x)$, being the pointwise limit of $g_h(t, x)$, must also be $\mathscr{L} \times \mathscr{B}$ measurable. It is also lower semicontinuous with respect to x by the definition of the epi-derivative. Thus $(t, x) \mapsto (f)'_{t,u(t)}(x)$ is a normal integrand. By applying (12) to $(*)$ we get

$$\int_0^T \inf_{v\in\mathbb{R}^m} \left[f'_{t,u(t)}(v) - w(t)v \right] dt \geq 0.$$

Since the integrand here is obviously nonpositive (take $v = 0$), it must actually vanish for almost all t. That is, $w(t) \in \partial f(t, u(t))$ almost everywhere. □

Now we turn to the second order epi-differentiability of $\mathcal{J}$.

Theorem 5. *Suppose both functions F and G are C^2 in x, the convex function σ is piecewise linear quadratic, and the set C is polyhedral. Suppose further that there exist $c \geq 0$ and $\phi \in L^1[0, T]$ such that for almost all t in $[0, T]$, the set $U(t)$ is convex, the vector $u(t) \in U(t)$ satisfies the constraint qualification (CQ), and for any γ in $\partial\sigma(F(t, u(t)))$ one has the growth condition*

$$\inf_{h>0} \frac{\sigma(F(t, u(t) + hx)) - \sigma(F(t, u(t))) - h\gamma F'(t, u(t))x}{h^2/2} \geq \phi(t) - c|x|^2 \qquad \forall x \in \mathbb{R}^m.$$

Then $\mathcal{J}$ is twice epi-differentiable at u. Its second epi-derivative relative to w in $\partial\mathcal{J}(u)$ is (*see* (10))

$$\mathcal{J}''_{u,w}(v) = \int_0^T f''_{t,u(t),w(t)}(v(t))\,dt \quad \forall\, v \in L^2. \tag{13}$$

Proof. Since $w(t)$ is an epi-gradient of f at $u(t)$, Theorem 2 provides $\gamma(t) \in \partial\sigma(F(t, u(t)))$ and $\eta(t) \in N_C(G(t, u(t)))$ such that

$$w(t) = \gamma(t)F'(t, u(t)) + \eta(t)G'(t, u(t)).$$

Fix x in $\mathbb{R}^m$. If $u(t) + hx \notin U(t)$, then $f_{t,u(t),w(t),h}(x) = \infty$. Otherwise, the convexity of $U(t)$ implies $x \in T_{U(t)}(u(t))$, that is, $G'(t, u(t))x \in T_C(G(t, u(t)))$, so $\eta(t)G'(t, u(t))x \leq 0$. Thus our assumed growth condition leads to the estimate

$$\begin{aligned} f(t, u(t) + hx) &- f(t, u(t)) - hw(t)x \\ &= \sigma(F(t, u(t) + hx)) - \sigma(F(t, u(t))) - hw(t)x \\ &= \sigma(F(t, u(t) + hx)) - \sigma(F(t, u(t))) \\ &\qquad - h\gamma(t)F'(t, u(t))x - h\eta(t)G'(t, u(t))x \\ &\geq \left(\phi(t) - c|x|^2\right) h^2/2. \end{aligned}$$

In either case, we have $f_{t,u(t),w(t),h}(x) \geq \phi(t) - c|x|^2$ for all x. Thus for any $0 < h < 1$ and $v \in L^2$, the second difference quotient $\mathcal{J}_{u,w,h}(v) = \int_0^T f_{t,u(t),w(t),h}(v(t))\,dt$ is at least $\int_0^T \phi(t)\,dt - c\|v\|_2^2$, which confirms the growth condition (2) of Attouch's theorem. By the second-order cognate of criterion (III) in Section 2, the second-order epi-differentiability of $\mathcal{J}$ at u relative to w is equivalent to

$$\lim_{\lambda\to\infty} \liminf_{h\downarrow 0} \mathcal{J}^\lambda_{u,w,h}(v) = \lim_{\lambda\to\infty} \limsup_{h\downarrow 0} \mathcal{J}^\lambda_{u,w,h}(v).$$

For each $\lambda > 0$, we have

$$\begin{aligned} \mathcal{J}^\lambda_{u,w,h}(v) &= \inf_{x\in L^2} \left\{\mathcal{J}_{u,w,h}(x) + \lambda\|x - v\|_2^2\right\} \\ &= \inf_{x\in L^2} \int_0^T \left(f_{t,u(t),w(t),h}(x(t)) + \lambda|x(t) - v(t)|^2\right) dt. \end{aligned}$$

Since $f_{t,u(t),w(t),h}(x)$ is a normal integrand and $f_{t,u(t),w(t),h}(x) \geq \phi(t) - 2c|v(t)|^2 - 2c|x-v(t)|^2$, we also have

$$k_1(t) := \phi(t) - 2c|v(t)|^2 \leq f^{\lambda}_{t,u(t),w(t),h}(v(t)) \leq \lambda|v(t)|^2$$

for $\lambda \geq 2c$. Thanks to Fatou's lemma and the monotone convergence theorem (see the proof of Theorem 3), the result follows. □

Let us apply our differentiation theorem to the integrand

$$(14) \qquad f(t, u) = \max_{1\leq i\leq l_1} \{f_i(t, u)\} + \sum_{i=l_1+1}^{l} \rho_i(d_i(f_i(t, u))) + \Psi_C(G(t, u)).$$

Here each $\rho_i\colon [0, \infty) \to \mathbb{R}$ has the form $\rho_i(x) = \frac{1}{2}r_i x^2 + a_i x + b_i$ with $r_i \geq 0$ and $a_i \geq 0$. So ρ_i is a nondecreasing affine ($r_i = 0$) or convex quadratic ($r_i > 0$) function. Each d_i is a distance function associated with either an infinite interval or a single point:

$$d_i(z) = \begin{cases} \max\{0, z - c_i\}, & \text{if } i = l_1+1, \dots, l_2, \\ |z - c_i|, & \text{if } i = l_2+1, \dots, l. \end{cases}$$

Thus f can be intepreted as a maximum of finitely many smooth functions, plus a standard augmented penalty function (optional—set $\rho_i \equiv 0$), plus an infinite penalty function (optional—choose $C = \mathbb{R}^m$).

We express the integrand in the simple form

$$f(t, u) = \sigma(F(t, u)) + \Psi_C(G(t, u))$$

by introducing the notation

$$\sigma(\alpha) = \max_{1\leq i\leq l_1} \{\alpha_i\} + \sum_{i=l_1+1}^{l} \rho_i(d_i(\alpha_i)),$$
$$F(t, u) = (f_1(t, u), \dots, f_l(t, u)).$$

We also denote

$$I(t, u) = \left\{ 1 \leq j \leq l_1 : f_j(t, u) = \max_{1\leq i\leq l_1} f_i(t, u) \right\},$$
$$S(t, u) = \left\{ \gamma \in \mathbb{R}^{l_1}_+ : \sum_{i=1}^{l_1} \gamma_i = 1;\ \gamma_i = 0,\ i \notin I(t, u) \right\},$$
$$p_i(t, u) = \rho_i'(d_i(f_i(t, u))) = r_i d_i(f_i(t, u)) + a_i,$$
$$S_i(t, u) = p_i(t, u)\partial d_i(f_i(t, u)) \quad \text{for} \quad i = l_1+1, \dots, l.$$

In this notation, the epiderivative and subgradient of the convex function σ are given by

$$\sigma'_{F(t,u)}(\alpha) = \max_{j\in I(t,u)} \{\alpha_j\} + \sum_{i=l_1+1}^{l} p_i(t, u)(d_i)'_{f_i(t,u)}(\alpha_i)$$
$$\partial\sigma(F(t, u)) = S(t, u) \times S_{l_1+1}(t, u) \times \cdots \times S_l(t, u).$$

According to Theorem 2, the function $f(t, \cdot)$ is twice epi-differentiable at any point $u \in U(t)$ where the constraint qualification (CQ) holds. Its first epi-derivative with respect to u is given by

$$f'_{t,u}(v) = \sigma'_{F(t,u)}(F'(t, u)v) + \Psi_{T_U(u)}(v).$$

Its second epi-derivative in u, relative to some $w \in \partial f(t, u)$, is given by

$$f''_{t,u,w}(v) = \sum_{i=l_1+1}^{l} r_i \left[(d_i)'_{f_i(t,u)}(f'_i(t, u)^T v)\right]^2 \\ + \max_{(\gamma,\eta)\in\Gamma(t,u,w)} \{v^T[\gamma F + \eta G]''(t, u)v\} + \Psi_{\Sigma(t,u,w)}(v),$$

where

$$\Gamma(t, u, w) = \{(\gamma, \eta) \in \partial\sigma(F(t, u)) \times N_C(G(t, u)) : \\ \gamma F'(t, u) + \eta G'(t, u) = w\},$$
$$\Sigma(t, u, w) = \left\{v \in T_{U(t)}(u) : \sigma'_{F(t,u)}(F'(t, u)v) + \Psi_{T_U(u)}(v) = wv\right\}.$$

These finite-dimensional calculations underlie the following infinite-dimensional result.

Theorem 6. *With f defined in* (14), *let u be a function such that $u(t) \in U(t)$ satisfies (CQ) for almost all t. Suppose there exists a constant $c \geq 0$ such that for almost all t, any x and y in $\mathbb{R}^m$ satisfy*

(i) $y^T f''_i(t, x)y \geq -c|y|^2$ *for* $i = 1, \ldots, l_2$,
(ii) $|y^T f''_i(t, x)y| \leq c|y|^2$ *for* $i = l_2 + 1, \ldots, l$,
(iii) $p_i(t, u(t)) \leq c$ *for* $i = l_1 + 1, \ldots, l$.

Suppose further that $|F'(t, u(t))|^2$ is integrable on $[0, T]$. Then $\mathcal{J}$ is epi-differentiable at u, and its epi-derivative is given by $\mathcal{J}'_u(v) = \int_0^T f'_{t,u(t)}(v(t))\,dt$. Furthermore, if $U(t)$ is convex, then $\mathcal{J}$ is twice epi-differentiable at u and its second epi-derivative relative to any $w \in \partial\mathcal{J}(u)$ is given by $\mathcal{J}''_{u,w}(v) = \int_0^T f''_{t,u(t),w(t)}(v(t))\,dt$.

Proof. The conclusions are precisely those of Theorems 3 and 5: we need only establish the growth conditions of those two results. The arguments in both cases are similar, so we discuss only the more demanding second-order condition of Theorem 5. For any $\gamma \in \partial\sigma(F(t, u(t)))$, the definition of subgradient and the mean value theorem imply

$$\frac{\sigma(F(u(t) + hx)) - \sigma(F(u(t)))}{h^2/2} \geq \gamma\left(\frac{F(t, u(t) + hx) - F(t, u(t))}{h^2/2}\right) \\ = \gamma\left(\frac{hF'(t, u(t))x}{h^2/2} + F''(t, \tilde{u})(x, x)\right).$$

Now $\gamma \in \partial\sigma(F(t, u(t)))$, so $0 \leq \gamma_i \leq 1$ for $i = 1, \ldots, l_1$, and $0 \leq \gamma_i \leq p_i(t, u(t)) \leq c$ for $i = l_1 + 1, \ldots, l_2$, and $|\gamma_i| \leq p_i(t, u(t)) \leq c$ for $i =$

$l_2 + 1, \ldots, l$. Thus our previous estimate implies

$$\begin{aligned}\frac{\sigma(F(t, u(t) + hx)) - \sigma(F(t, u(t))) - h\gamma F'(t, u(t))x}{h^2/2} &\geq -\sum_{i=1}^{l_1} c\gamma_i |x|^2 - \sum_{i=l_1+1}^{l} cp_i(t, u(t))|x|^2 \\ &\geq -\tilde{c}|x|^2\end{aligned}$$

for some constant $\tilde{c}$. This establishes the growth condition of Theorem 5, and hence the result. □

Comparison to other work. The generalization of epi-convergence from finite to infinite-dimensional spaces calls for some discussion of the topologies involved. Mosco epi-convergence, defined by changing " $y_i \to y$ " to " y_i converges weakly to y " in criterion I(i) of Section 2, is appropriate in many problems, including the analysis of convex integral functionals—see Do [5]. However, the Mosco epi-limit of any functional is, by definition, weakly lower semicontinuous, and it is well known that every weakly lower semicontinuous integral functional is convex. Thus our problem of treating *nonconvex* integrands, and obtaining correspondingly nonconvex results, can be solved only in some stronger topology.

Cominetti's work on general amenable functions [4] on reflexive Banach spaces deals exclusively with epidifferentiability in the Mosco sense. Thus, although [4, Theorem 4.4, page 860] and our Theorem 5 look similar, the former does not give useful second-order information for the integral functionals considered here. (The problem can be traced to the weak-continuity assumption on the second-derivative mapping in [4, Theorem 4.4].) We note, however, that Cominetti's Theorem 4.4 has other uses, and that in particular its finite-dimensional instance generalizes Theorem 2 of this paper.

Noll [8] discusses the (strong) epi-differentiability of integral functionals whose integrands are finite and have second-order Taylor expansions. Thus his integrands are at least Fréchet differentiable. He shows that in this case, the second epi-derivative is a quadratic functional [8, Theorem 3.1]. In contrast, we deal with extended real-valued, nonsmooth integrands. Our results include the case where $f(t, u) = F(t, u)$ is smooth and scalar-valued as a function of u: we simply choose $l = 1$, $\sigma(\alpha) = \alpha$, $G \equiv 0$, and $C = \mathbb{R}^k$. In this case $\partial f(t, u) = \{F'(t, u)\}$, and it is easy to calculate $\Gamma(t, u, w) = \{(1, 0)\}$ and $\Sigma(t, u, w) = \mathbb{R}$, so $f''_{t,u,w}(v) = v^T F''(t, u)v$. If we add the assumption that $\inf_{x \in \mathbb{R}^m} F''(t, x) > -\infty$, then our Theorem 3 shows that $\mathcal{I}$ is twice epi-differentiable at u with second epi-derivative

$$\mathcal{I}''_{u,w}(v) = \int_0^T v^T(t)F''(t, u(t))v(t)\, dt,$$

as expected. Of course, membership in the class C^2 is a stronger condition on F than the existence of a second-order Taylor expansion. But even in this case, we have something Noll [8] does not cover. He considers only integrands bounded by quadratic functions [8, (3.2)], or, in the C^2 case, integrands whose Hessian matrices are bounded [8, Theorem 4.2]. We only impose a lower bound on the Hessian matrix. Thus our theory pertains to arbitrary smooth convex integrands (such as $f(t, u) = e^u$ in the case $m = 1$), whereas Noll's requires that the integrand grow at most quadratically. As for functionals in Sobolev spaces, or in our terms Bolza functionals, a similar remark applies.

Levy [6] discusses the strong epi-differentiability of integral functionals in L^p spaces and applies the results to the sensitivity analysis of set-valued functions. His paper contains results analogous to Theorems 3 and 5, which we proved independently at about the same time. (The proof techniques, however, are different: Levy chooses sequences satisfing criterion (I) directly, whereas we pass to Moreau-Yosida approximates and prove (III).) Indeed, a draft of [6] arrived in time for us to state Theorem 3 and Theorem 5 in a form that facilitates direct comparison. Our Theorem 6, which identifies a significant class of functions to which the general theory applies, has no counterpart in [6]. Finally, our applications are disjoint from those in [6]: whereas Levy concentrates on sensitivity analysis, we discuss the epi-differentiability of Bolza functionals and the resulting necessary conditions in optimal control.

4. Bolza functionals

Consider the functional $\mathscr{J}: L^2([0, T], \mathbb{R}^m) \to \overline{\mathbb{R}}$ defined by

$$\mathscr{J}(u) = \int_0^T [L(t, x(t), u(t)) + \Psi_{U(t)}(u(t))]\, dt,$$

where $U(t) = \{u \in \mathbb{R}^m : G(t, u) \in C\}$ as before, and the function x is given in terms of u by $x(t) = a(t)+E(u)(t)$ for some operator $E \in \mathscr{L}(L^2([0, T]; \mathbb{R}^m), L^2([0, T]; \mathbb{R}^n))$ and function $a \in L^2$. Note that the choices $m = n$ and $E(u)(t) = x_0 + \int_0^t u(r)\, dr$ imply $\dot{x} = u$ a.e., so that $\mathscr{J}$ takes the form of a classical Bolza functional with an infinite penalty for derivative values $\dot{x}(t)$ outside the prescribed velocity set $U(t)$. In this section we study the epi-differentiability of $\mathscr{J}$.

As in Section 3, we assume that C is a closed convex subset of $\mathbb{R}^k$, while $G(t, u): [0, T] \times \mathbb{R}^m \to \mathbb{R}^k$ is measurable in t and C^2 in u. The integrand L must satisfy the following conditions:

(1) The function $L(t, x, u)$ is measurable in t and C^2 in (x, u);
(2) Both $\int_0^T |L(t, 0, 0)|^2\, dt < \infty$ and $\int_0^T |L'(t, 0, 0)|^2\, dt < \infty$;
(3) There exists a constant $c \geq 0$ such that $|L''(t, x, u)| \leq c$ for all (t, x, u) in $[0, T] \times \mathbb{R}^{n+m}$.

(In (2) and (3), the gradient L' and Hessian L'' refer to the vector argument (x, u) in $\mathbb{R}^{n+m}$.)

To differentiate $\mathscr{J}$, we first decompose $\mathscr{J} = \mathscr{I}(\Phi(u))$, where $\mathscr{I}(u, v) = \mathscr{I}_1(x, u) + \mathscr{I}_2(u)$ is the functional on $L^2 \times L^2$ defined in terms of

$$\mathscr{I}_1(x, u) = \int_0^T L(t, x(t), u(t))\, dt,$$

$$\mathscr{I}_2(u) = \int_0^T \Psi_{U(t)}(u(t))\, dt,$$

$$\Phi(u) = (a + E(u), u).$$

Note that the functional $\mathscr{I}_1$ is continuously (Fréchet) differentiable and twice weakly Gâteaux differentiable, but in general it is not twice continuously differentiable unless L is precisely a quadratic function. (See Noll [8] for a detailed example and further references.)

Since $\mathscr{J}_1 \circ \Phi \in C^1$, we have $\partial\mathscr{J}(u) = \mathscr{J}_1'(\Phi(u))\Phi'(u) + \partial\mathscr{J}_2(u)$. Thus every $w \in \partial\mathscr{J}(u)$ gives rise to a functional $w_1 := w - \mathscr{J}_1'(\Phi(u))\Phi'(u)$ in $\partial\mathscr{J}_2(u)$, so that $w_1(t) \in N_{U(t)}(u(t))$ a.e. by Corollary 4. For any $v \in L^2$, $\Phi'(u)v = (E(v), v)$: writing $y = E(v)$ gives

$$\begin{aligned}\langle w , v\rangle &= \langle \mathscr{J}_1'(\Phi(u))\Phi'(u) , v\rangle + \langle w_1 , v\rangle \\ &= \langle \mathscr{J}_1'(\Phi(u)) , \Phi'(u)v\rangle_{L^2\times L^2} + \langle w_1 , v\rangle \\ &= \int_0^T [L_x(t)y(t) + L_u(t)v(t) + w_1(t)v(t)]\, dt \\ &= \langle (L_x , L_u + w_1) , \Phi'(u)v\rangle_{L^2\times L^2}.\end{aligned}$$

So we have $w = \Phi'(u)^*\overline{w}$, where $\overline{w}(t) = (L_x(t), L_u(t) + w_1(t))$.

We now describe the epi-differentiability of $\mathscr{J}$.

Theorem 7. *If the constraint qualification (CQ) holds at $u(t)$ for almost all t, then $\mathscr{J}$ is epi-differentiable at u. Its epi-derivative is given by*

$$\mathscr{J}_u'(v) = \mathscr{J}_{(x,u)}'(y, v) = \int_0^T \left[L_x(t)y(t) + L_u(t)v(t) + \Psi_{T_{U(t)}}(u(t))\right] dt,$$

where $y = E(v)$. If, in addition, the set $U(t)$ is convex, then $\mathscr{J}$ is twice epi-differentiable at u relative to $w \in \partial\mathscr{J}(u)$, and its second epi-derivative is

$$\mathscr{J}_{u,w}''(v) = \int_0^T \Big[L_2(t, v(t), y(t)) + \max_{\eta\in\Gamma(t)} \{v(t)^T(\eta^T G)''(t, u(t))v(t)\} + \Psi_{\Sigma(t)}(v(t))\Big]\, dt.$$

Here $w_1 := w - \mathscr{J}_1'(\Phi(u))\Phi'(u)$ and

$$\begin{aligned}L_2(t, v, y) &= v^T L_{uu}(t)v + 2y^T L_{ux}(t)v + y^T L_{xx}(t)y, \\ \Gamma(t) &= \{\eta \in N_C(G(t, u(t))) : \eta G'(t, u(t)) = w_1(t)\}, \\ \Sigma(t) &= \{v \in T_{U(t)}(u(t)) : w_1(t)v = 0\}.\end{aligned}$$

Proof. Note that $\mathscr{J} = \mathscr{J}_1 + \mathscr{J}_2$ is epi-differentiable, since $\mathscr{J}_1$ is continuously differentiable and $\mathscr{J}_2$ is epi-differentiable by Theorem 3. This is relevant because $\mathscr{J}_{u,h}(v) = \mathscr{J}_{(x,u),h}(y, v)$ where $y = E(v)$. By criterion (I) in Section 2, our first-order assertions regarding $\mathscr{J}$ will follow if for any point $v \in L^2$ and any sequence $h_i \downarrow 0$, both

(i) for any sequence v_i converging to v, we have $\liminf_{i\to\infty} \mathscr{J}_{u,h_i}(v_i) \geq \mathscr{J}_{(u,x)}'(v, y)$; and

(ii) there exists a sequence v_i converging to v, such that $\limsup_{i\to\infty} \mathscr{J}_{u,h_i}(v_i) \leq \mathscr{J}_{(u,x)}'(v, y)$.

Given any sequence $v_i \to v$ in L^2, let $y_i = E(v_i)$. Then $(v_i, y_i) \to (v, y)$ in $L^2 \times L^2$. Since $\mathscr{J}$ is epi-differentiable at (u, x), we have

$$\liminf_{i\to\infty} \mathscr{J}_{(x,u),h_i}(y_i, v_i) \geq \mathscr{J}_{(x,u)}'(y, v),$$

so condition (i) holds. It remains to prove (ii). This is trivial if $\mathscr{J}_{(x,u)}'(y, v) = \infty$, so we assume that $\mathscr{J}_{(u,x)}'(v, y)$ is finite. The epi-differentiability of $\mathscr{J}$ implies that there exists a sequence $(w_i, v_i) \to (y, v)$ such that

$$\limsup_{i\to\infty} \mathscr{J}_{(x,u),h_i}(w_i, v_i) \leq \mathscr{J}_{(x,u)}'(y, v).$$

Since the right side is finite, we have $u(t) + h_i v_i(t) \in U(t)$ almost everywhere for i sufficiently large. Now by definition,

$$\mathscr{I}_{(x,u),h_i}(w_i, v_i) = \mathscr{I}_{(x,u),h_i}(y_i, v_i) + \epsilon_i$$

where

$$\epsilon_i = \int_0^T \frac{L(t, x + h_i w_i, u + h_i v_i) - L(t, x + h_i y_i, u + h_i v_i)}{h_i}\, dt.$$

Using the mean value theorem, we estimate

$$\begin{aligned}
&|L(t, x + y, u + v) - L(t, x + z, u + v)| \\
&\quad = |L_x(t, x + \theta y + (1 - \theta)z, u + v)||y - z| \\
&\quad \le (|L_x(t, x, u)| + c|y| + c|z| + c|v|)|y - z|.
\end{aligned}$$

It follows that

$$|\epsilon_i| \le \int_0^T (|L_x(t, x, u)| + c|w_i| + c|y_i| + c|v_i|)\, |w_i - y_i|\, dt.$$

The sequence of integrals $\int_0^T (|L_x(t, x(t), u(t))| + c|v_i(t)| + c|w_i(t)| + c|y_i(t)|)^2\, dt$ is bounded and $\int_0^T |w_i(t) - y_i(t)|^2\, dt \to 0$. Therefore $\epsilon_i \to 0$ as $i \to \infty$. So

$$\limsup_{i\to\infty} \mathscr{J}_{u,h_i}(y_i) = \limsup_{i\to\infty} \mathscr{I}_{(x,u),h_i}(w_i, v_i) \le \mathscr{I}'_{(x,u)}(y, v).$$

By criterion (I), the functional $\mathscr{J}$ is epi-differentiable at u with epi-derivative $\mathscr{I}'_{(u,x)}(v, y)$.

Next, suppose $w \in \partial\mathscr{J}(u)$. The second difference quotient of $\mathscr{J}$ at u relative to w is

$$\mathscr{J}_{u,w,h}(v) = \mathscr{I}_{(x,u),(L_x, L_u + w_1),h}(y, v).$$

Note that $\mathscr{I}$ is twice epi-differentiable at (x, u) relative to $(L_x, L_u + w_1)$, so a discussion similar to that for the first epi-derivative shows that $\mathscr{J}$ is twice epi-differentiable at u. The only difference is the estimate of ϵ_i. This time ϵ_i equals

$$\int_0^T \frac{L(t, x + h_i w_i, u + h_i v_i) - L(t, x + h_i y_i, u + h_i v_i) - h_i L_x(t, x, u)(w_i - y_i)}{h_i^2/2}\, dt.$$

The mean value theorem provides the estimate

$$\begin{aligned}
&|L(t, x + y, u + v) - L(t, x + z, u + v) - L_x(t, x, u)(y - z)| \\
&\quad = |(L_x(t, x + \theta y + (1 - \theta)z, u + v) - L_x(t, x, u))\,(y - z)| \\
&\quad \le |y - z|(c|y| + c|z| + c|v|),
\end{aligned}$$

which implies that

$$|\epsilon_i| \le \int_0^T 2c(|w_i(t)| + |y_i(t)| + |v_i(t)|)|w_i(t) - v_i(t)|\, dt.$$

Consequently $\epsilon_i \to 0$ as $i \to \infty$, as required. $\square$

5. An optimal control problem

In this section, we apply our results to obtain necessary conditions for optimality in the following free endpoint control problem:

$$(P)\qquad \begin{aligned} &\text{minimize} && \int_0^T L(t, x(t), u(t))\, dt, \\ &\text{subject to} && \dot{x}(t) = A(t)x(t) + B(t)u(t) \text{ a.e.}, \ x(0) = x_0, \\ & && u(t) \in U(t) \text{ a.e.} \end{aligned}$$

Here the state x evolves in $\mathbb{R}^n$ and the control u takes values in $\mathbb{R}^m$, so the measurable matrix-valued functions A and B have dimensions $n\times n$ and $n\times m$ respectively. As before, the control set is $U(t) = \{u \in \mathbb{R}^m : G(t, u) \in C\}$.

Let X be the fundamental matrix function associated with A, i.e., the unique solution of the initial-value problem $X'(t) = A(t)X(t)$, $X(0) = I$. Using X, we define the operator $E: L^2([0, T]; \mathbb{R}^m) \to L^2([0, T]; \mathbb{R}^n)$ and the function $a \in L^2([0, T]; \mathbb{R}^n)$ by

$$(15)\qquad E(u)(t) = \int_0^t X(t)X(s)^{-1}B(s)u(s)ds, \qquad a(t) = X(t)x_0.$$

Then the controlled dynamics above reduce to the equation $x = a + E(u)$, so problem (P) takes the form

$$\text{minimize } \mathscr{J}(u) = \int_0^T \left[L(t, a(t) + E(u)(t), u(t)) + \Psi_{U(t)}(u(t))\right] dt.$$

We will apply the analysis of $\mathscr{J}$ in Section 4 to derive optimality conditions for (P).

The pre-Hamiltonian for problem (P) is

$$H(t, x, p, u) = p^T(A(t)x + B(t)u) - L(t, x, u).$$

We also consider the extended pre-Hamiltonian function below, which incorporates the control constraints through a multiplier vector $\eta \in (\mathbb{R}^k)^*$:

$$\mathscr{H}(t, x, p, u, \eta) = p^T(A(t)x + B(t)u) - L(t, u, x) - \sum_{i=1}^k \eta_i g_i(t, u).$$

Our final result concerns first- and second-order necessary conditions for optimality in problem (P), formulated in terms of the adjoint arc p defined by

$$(16)\qquad -\dot{p}(t)^T = H_x(t) = p(t)^T A(t) - L_x(t, x(t), u(t)), \quad p(T) = 0.$$

Theorem 8. *Let (x, u) give the minimum in (P), such that the constraint qualification (CQ) holds at $u(t)$ for almost all t. Then for all (y, v) satisfying the linearized system*

$$(17)\qquad \dot{y}(t) = A(t)y(t) + B(t)v(t), \quad y(0) = 0, \quad v(t) \in T_{U(t)}(u(t)),$$

one has

$$(18)\qquad 0 \le \mathscr{J}_u'(v) = \int_0^T (L_x(t)y(t) + L_u(t)v(t))\, dt.$$

If, in addition, the control set $U(t)$ is convex, then every such pair (y, v) satisfies (19)

$$0 \le \mathscr{J}''_{u,0}(v) = \int_0^T \left[L_2(t, y(t), v(t)) + \max_{\eta \in \Gamma(t)} \left\{ v(t)^T (\eta G)''(t, u(t)) v(t) \right\} + \Psi_{\Sigma(t)}(v(t)) \right] dt,$$

where

$$\begin{aligned} L_2(t, y, v) &= v^T L_{uu}(t) v + 2y^T L_{ux}(t) v + y^T L_{xx}(t) y, \\ \Gamma(t) &= \{\eta \in N_C(G(t, u(t))) : \mathscr{H}_u(t) = 0\}, \\ \Sigma(t) &= \left\{ v \in T_{U(t)}(u(t)) : (p(t)^T B(t) - L_u(t)) v = 0 \right\}. \end{aligned}$$

Proof. The first-order condition $\mathscr{J}_u'(v) \ge 0$ for all v is a direct consequence of the definition of the epi-derivative. The calculation of $\mathscr{J}_u'$ implicit in (17) and (18) follows from Theorem 7 and the definition of E. Likewise, the second-order condition $\mathscr{J}''_{u,0}(v) \ge 0$ for all v is obvious: the point of (19) is its formula for the second epi-derivative. This follows from Theorem 7, once we evaluate $w_1 = -\mathscr{J}_1'(\Phi(u))\Phi'(u)$. For any $v \in L^2$, we write $y = E(v)$ to obtain

$$\begin{aligned} \langle \mathscr{J}_1'(\Phi(u))\Phi'(u), v \rangle &= \langle \mathscr{J}_1'(\Phi(u)), \Phi'(u) v \rangle_{L^2 \times L^2} \\ &= \langle \mathscr{J}_1'(x, u), (y, v) \rangle_{L^2 \times L^2} \\ &= \int_0^T (L_x(t) y(t) + L_u(t) v(t))\, dt \\ &= \int_0^T [(\dot{p}(t)^T + p(t)^T A(t)) y(t) + L_u(t) v(t)]\, dt \\ &= \int_0^T [(\dot{p}(t)^T y(t) + p(t)^T \dot{y}(t)) - p(t)^T B(t) v(t) + L_u(t) v(t)]\, dt \\ &= \int_0^T (L_u(t) - p(t)^T B(t)) v(t)\, dt \\ &= \left\langle L_u - p^T B, v \right\rangle. \end{aligned}$$

Thus $w_1(t) = -[\mathscr{J}_1'(\Phi(u))\Phi'(u)](t) = p(t)^T B(t) - L_u(t)$, and the sets $\Gamma(t)$ and $\Sigma(t)$ of Theorem 7 reduce to those of the current statement. □

Remark. If the gradients $g_i'(t, u(t))$ corresponding to the equality and active inequality indices defining the set $U(t)$ are linearly independent for almost all t, then the multipliers η_i are unique, and Theorem 8 reduces to Loewen and Zheng [7, Theorem 3.4].

References

1. H. Attouch, *Variational convergence for functions and operators*, Pitman, Boston, 1984.
2. J. P. Aubin and H. Frankowska, *Set-valued analysis*, Birkháuser, Boston, 1990.
3. F. H. Clarke, *Optimization and nonsmooth analysis*, Wiley-Interscience, New York, 1983.
4. R. Cominetti, *On pseudo-differentiability*, Trans. Amer. Math. Soc. **324** (1988), 843–865.
5. C. Do, *Generalized second derivatives of convex functions in reflexive Banach spaces*, Trans. Amer. Math. Soc. (to appear).

6. A. Levy, *Second-order epi-derivatives of integral functionals with fully amenable integrands*, Set-Valued Analysis (to appear).

7. P. D. Loewen and H. Zheng, *Generalized conjugate points for optimal control problems*, Nonlinear Analysis, Theory Methods & Applications (to appear).

8. D. Noll, *Second order differentiability of integral functionals on Sobolev spaces and L^2 spaces*, Z. Reine Angew. Math. (to appear).

9. R. A. Poliquin and R. T. Rockafellar, *A calculus of epi-derivatives with applications to optimization*, Canad. J. Math. **45** (1993), 879-896.

10. ———, *Amenable functions in optimization*, preprint.

11. R. T. Rockafellar, *Integral functionals, normal integrands and measurable selections*; in Nonlinear Operators and the Calculus of Variations, Lecture Notes in Math., vol. 543, Springer-Verlag, 1976, pp. 157–207.

12. ———, *First- and second-order epi-differentiability in nonlinear programming*, Trans. Amer. Math. Soc. **307** (1988), 75–107.

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

14. ———, *Proto-differentiability of set-valued mappings and its applications in optimization*, Analyse Non Linéaire (H. Attouch, J.P. Aubin, F. Clarke, I. Ekeland, eds.), 1989, pp. 449–482.

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

16. ———, *Nonsmooth analysis and parametric optimization*, Methods of Nonconvex Analysis (A. Cellina, ed.), Lecture Notes in Math., vol. 1446, Springer-Verlag, 1990, pp. 137–151.

Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z2

E-mail address: `loew@math.ubc.ca`

Department of Electrical Engineering, Imperial College of Science, Technology, and Medicine, London SW7 2BT, United Kingdom