Variational Analysis of a Composite Function:  
A Formula for the Lower Second Order Epi-derivative\(^1\)

ALEXANDER IOFFE

Department of Mathematics, The Technion, Haifa 32000, Israel

Submitted by Frank H. Clarke  
Received September 13, 1989

An exact formula is established for the lower second order epi-derivative of a function of the form \(g(F(x))\), where \(F\) is a smooth map from one Banach space into another and \(g\) is a convex function (generally, not everywhere finite). Unconstrained minimization of such functions typically arise as an equivalent (in one or another sense) reduction form for many important classes of constrained optimization problems. The formula is further applied to study epi-differentiability of the max-function \(f(x) = \max \{ f(q, x) : q \in Q \}\).

1. INTRODUCTION

Let \(X\) be a Banach space and \(f\) an extended real-valued function on \(X\) which is finite at \(\bar{x}\). In the article we consider functions

\[
    f(x) = g(F(x)),
\]

where \(F\) is a \(C^1\)-map of a neighborhood of \(\bar{x} \in X\) into another Banach space \(Y\) which is twice Fréchet differentiable at \(\bar{x}\) and \(g\) is a proper (i.e., everywhere \(> -\infty\) and not identically equal to \(+\infty\)) convex function on \(Y\) which is finite and l.s.c. at \(\bar{y} = F(\bar{x})\).

We say that \(f\) is convex—twice differentiable composite, or, simply, \(CC^2\)-composite at \(\bar{x}\), or that it belongs to the class \(CC^2(\bar{x})\).

Functions of this class have been recently recognized as being extremely important in optimization theory mainly because practically every "smooth" optimization problem with constraints can be equivalently (in one sense or another) reformulated in terms of unconstrained minimization of a \(CC^2\)-composite function.

Obviously this offers a tempting possibility of a different and very direct approach to optimization theory (and there has been a number of results already obtained in this way \([3, 4, 7–10, 15]\)), provided only that local analysis of \(CC^2\)-composite functions has been developed well enough.

\(^1\) This research was supported by the Fund for Promotion of Science at the Technion under Grant 100-820.
Epi-derivatives of various order seem to be extremely suitable for that purpose. The first order lower epi-derivative is the same as the lower contingent derivative. It found numerous and diverse applications in analysis and optimization (see [1], where the term epicontingent derivative is used, and references there). The second order lower epi-derivative \( f''_-(\tilde{x}, x^*; h) \) (or the upper \( f''_+(\tilde{x}, x^*; h) \) or, just the (two-sided) epi-derivative \( f''(\tilde{x}, x^*; h) \)) of a function \( f(\cdot) \) at \( \tilde{x} \) with respect to a given \( x^* \in X^* \) is defined as the function whose epigraph is the upper Kuratowski limit (or lower, or just the limit if it exists) of the epigraphs of the functions

\[
\phi_t(h) = t^{-2}[f(\tilde{x} + th) - f(x) - t\langle x^*, h \rangle]
\]
as \( t \to 0 \). Analytically,

\[
f''_-(\tilde{x}, x^*; h) = \liminf_{t \to 0} t^{-2}(f(\tilde{x} + th) - f(\tilde{x}) - t\langle x^*, h' \rangle).
\]

The basic property that distinguishes the lower second order epi-derivative from other second order subderivatives is the following.

**Proposition 1.** If \( f''_-(\tilde{x}, x^*; h) \geq 0 \) for all \( h \) then for any \( \varepsilon > 0 \) and any linear subspace \( L \subset X \) with \( \dim L < \infty \) the function

\[
h \mapsto f(\tilde{x} + h) - f(\tilde{x}) - t\langle x^*, h \rangle + \varepsilon \|h\|^2
\]
attains a local minimum on \( L \) at zero.

This means that, while other second order directional derivatives can also be used to obtain certain second order necessary condition for a minimum, only \( f''_-(\tilde{x}, x^*; h) \) or its equivalents are suitable for sufficient conditions.

Two-sided second order epi-derivatives were first considered by Rockafellar [15, 16]. The lower second order epi-derivative seems to have not yet appeared as an explicit object of study, although it must be observed that in case \( \dim X < \infty \) the second order subderivative introduced by Chaney [5], when it exists, coincides with \( f''_-(\tilde{x}, x^*; h) \). (This was probably obscured by complexity of Chaney’s definition which, in addition, a priori requires some connection between \( x^* \) and \( h \) at certain points around \( \tilde{x} \).) For that reason we have retained Chaney’s notation.

A subclass of \( CC^2 \)-composite functions on \( R^n \) corresponding to \( g \) being piecewise linear-quadratic was considered in [15]. It was shown that functions of this subclass are twice epi-differentiable and a formula for the second epi-derivative was found. Recall that a convex function is piecewise linear quadratic if its domain can be broken into a finite number of polyhedral pieces in such a way that the function be either affine or quadratic on each of them. This subclass is sufficient to cover standard problems of mathematical programming with finitely many constraints but
not much more. The example at the end of the next section shows that already a function on $\mathbb{R}^2$ which is the maximum of an affine and a linear function may not belong to this subclass. (Moreover, the function in the example is twice epi-differentiable but its second order epi-derivative cannot be computed by Rockafellar's formula.) An extension of Rockafellar's formula to a broader class of outer functions $g(\cdot)$ was given by Cominetti [6] who required that $g(\cdot)$ be Lipschitz on its domain and have the property that the second order epi-derivative exists and coincides with the usual second order directional derivative. This property is also strong enough and the above mentioned example at the end of Section 3 shows that very simple functions may not satisfy it.

There is no hope that the second order epi-derivative may exist in a general situation, and if we want to deal with general classes of optimization problems using reduction to composite optimization we must move our attention to subderivatives, as was done in the first order theory.

In the paper we establish several formulae for the lower second order epi-derivative of a $CC^2$-composite function both in the general situation described in the beginning and in some special cases. We also give some new criteria for second order epi-differentiability and corresponding formulae for epi-derivatives. Our intention is to apply these results later to study higher order conditions and parametric sensitivity both in composite and constrained optimization [11, 12].

The paper is organized in the following way. In Section 2 we give the statement of the main result and prove some corollaries specifying and simplifying the formula for certain particular situations. In particular, we somewhat generalize the mentioned Cominetti's extension of Rockafellar's result by proving it in a Banach space without assuming that $g(\cdot)$ is Lipschitz on its domain.

Section 3 contains the proof of the main formula. It must be observed that the proof is much simpler if $g$ is continuous. The main effort in the proof is applied to prove a "basic lemma" which shows that a certain higher order regularity estimate hold under simplified (but close to standard general) first order regularity assumptions.

Finally in Section 4 we consider the max-function

$$f(x) = \max_{q \in Q} f(q, x).$$

The main formula is specified for such functions (Theorem 2), and then we extend to such functions the theory developed by Rockafellar in [15] for composite functions on $\mathbb{R}^n$ with piecewise linear-quadratic $g(\cdot)$. Namely, in Theorem 3 we give sufficient conditions for the functions of this class to be twice (two-sided) epi-differentiable as well as a formula for the second epi-derivative under those conditions. The formula employs a simple
construction that seems to have first appeared in Kawasaki's formula [14] for the upper "parabolic" derivative similar to that introduced by Ben-Tal and Zowe [2]. We give a short proof of Kawasaki's result in Theorem 4 and conclude the paper by establishing in Theorem 5 the connection between the second epi-derivative and the parabolic derivative under the sufficient conditions of Theorem 3.

2. THE MAIN THEOREM: STATEMENT AND COROLLARIES

In what follows we fix a point \( x \in X \) and set for simplicity
\[
\bar{y} = F(x), \quad A = F'(x), \quad w(h) = \left( \frac{1}{2} \right) \cdot F''(x)(h, h).
\]
For \( x^* \in X^* \), we set
\[
\Omega(x^*) = \{ y^* \in \partial g(y) : Ay^* = x^* \}
\]
\[
K(x^*) = \{ h \in X : \langle y^*, Ah \rangle \leq \langle x^*, h \rangle, \forall y^* \in \partial g(\bar{y}) \},
\]
and denote by \( L \) the closed subspace of \( Y \) spanned by \( \text{dom} \ g - \bar{y} \). We adopt the following regularity hypothesis:

(H) \( \text{Im} A + L = Y \) and there is a \( \tilde{u} \in X \) such that \( g \) is finite and continuous with respect to \( \bar{y} + L \) at \( \bar{y} + A\tilde{u} \).

This regularity assumption is not the most general but simple and convenient to work with and general enough to cover the most important applications. In particular, it is automatically satisfied if \( g \) is continuous or if \( g \) is as in Example 1 and the original problem with constraints satisfies the standard regularity condition. If both \( X \) and \( Y \) are finite dimensional, (H) coincides with the "basic constraint qualification" of Rockafellar [15].

THEOREM 1. If the regularity hypothesis (H) is satisfied, then for a given \( x^* \) we have the following alternative:

(a) \( \text{either } f''_-(\bar{x}, x^*; h) = -\infty \text{ for some } h \),

(b) \( \text{or } \Omega(x^*) \neq \emptyset \text{ and for any } h \)

\[
f''_-(\bar{x}, x^*; h) = \lim_{h' \to h} \lim_{t \to +0} \lim_{w \to w(h)} \frac{g(\bar{y} + th' + t^2w) - g(\bar{y}) - t\langle x^*, h' \rangle}{t^2}
\]
\[
= \lim_{h' \to h} \lim_{t \to +0} \frac{g(\bar{y} + th' + t^2w(h)) - g(\bar{y}) - t\langle x^*, h' \rangle}{t^2}
\]

If, in addition, \( f''_-(\bar{x}, x^*; h) < \infty \text{ for some } h \), then \( h \in K(x^*) \).
At first glance the formula looks like a chain rule and
\[ g(h, w) = \liminf_{h \to h' \atop t \to +0} t^{-2}(g(\tilde{y} + tAh' + t^2w) - g(\tilde{y}) - t\langle x^*, h' \rangle) \] (2.1)

resembles the second order directional derivative of Ben-Tal and Zowe [2] (if we ignore the fact that the limit inferior is used instead of the limit). A substantial difference, however, is caused by the fact that \( h \) is involved in the limit process. If, say, \( A \) is onto, then (2.1) does not depend on \( w \) at all and therefore cannot be represented as a composition of \( Ah, w(h) \) and a certain derivative of \( g \).

**Corollary 1.** If \( \Omega(x^*) \neq \emptyset \), then
\[ f''(x, x^*; h) \leq \max_{y^* \in \Omega(x^*)} \langle y^*, w(h) \rangle + \liminf_{h \to h' \atop t \to +0} t^{-2}(g(\tilde{y} + tAh') - g(\tilde{y}) - t\langle x^*, h' \rangle). \]

**Proof.** We observe first that for any \( z \in X \)
\[ g(h, w + Az) = g(h, w) + \langle x^*, z \rangle \] (2.2)
and that
\[ f''(x, x^*; h) = g(h, w(h)). \]

We have further
\[ g(\tilde{y} + tAh' + t^2w) - g(\tilde{y}) - t\langle x^*, h' \rangle \leq (1 - t) \left[ g(\tilde{y} + tAh' \frac{h'}{1-t}) - g(\tilde{y}) - t\langle x^*, \frac{h'}{1-t} \rangle \right] + t\left[ g(\tilde{y} + tw) - g(\tilde{y}) \right]. \]

Therefore
\[ g(h, w) \leq \liminf_{h \to h' \atop t \to +0} \frac{g(\tilde{y} + tAh'/(1-t)) - g(\tilde{y}) - t\langle x^*, (h'/(1-t)) \rangle}{t^2} \]
\[ + \lim_{t \to 0} t \left[ g(\tilde{y} + tw) - g(\tilde{y}) \right] \]
\[ = \liminf_{h \to h' \atop t \to +0} \frac{g(\tilde{y} + tAh') - g(\tilde{y}) - t\langle x^*, h' \rangle}{t^2} + g'(\tilde{y}, w). \]
We note further that owing to (H), the function
\[ \eta(w) = \inf_z \left[ g'(\tilde{y}; w + Az) - \langle x^*, z \rangle \right] \]

is continuous, hence coinciding with its second conjugate. It is an easy exercise from convex analysis to show that
\[ \eta^{**}(w) = \max \left\{ \langle y^*, w \rangle : y^* \in \partial g(\tilde{y}), A^*y^* = x^* \right\} = \max \left\{ \langle y^*, w \rangle : y^* \in \Omega(x^*) \right\}. \]

Therefore for any \( w \) we have by (2.2)
\[ \mathcal{G}(h, w) = \inf_z \left[ \mathcal{G}(h, w + Az) - \langle x^*, z \rangle \right] \leq \lim_{h' \to h, t \to 0} t^{-2}(g(\tilde{y} + tAh') - g(\tilde{y}) - t\langle x^*, h' \rangle) \]
\[ + \inf_z \left\{ g'(\tilde{y}; w + Az) - \langle x^*, z \rangle \right\} \]
\[ = \lim_{h' \to h, t \to 0} t^{-2}(g(\tilde{y} + tAh') - g(\tilde{y}) - t\langle x^*, h' \rangle) \]
\[ + \max \left\{ \langle y^*, w \rangle : y^* \in \Omega(x^*) \right\}. \]

I am indebted to R. Poliquin for the observation that, likewise, we can prove that
\[ f''(\tilde{x}, x^*; h) \geq \min_{y^* \in \Omega(x^*)} \langle y^*, w \rangle \]
\[ + \lim_{h' \to h, t \to 0} t^{-2}(g(\tilde{y} + tAh') - g(\tilde{y}) - t\langle x^*, h' \rangle) \]
which follows at the same way from the inequality
\[ g(\tilde{y} + tAh' + t^2w) - g(\tilde{y}) - t\langle x^*, h' \rangle \]
\[ \geq (1 + t) \left[ g\left( \tilde{y} + tA \frac{h'}{1+t} \right) - g(\tilde{y}) - t\langle x^*, \frac{h'}{1+t} \rangle \right] \]
\[ + t\left[ g(\tilde{y} + tw) - g(\tilde{y}) \right]. \]

We therefore have the following.
**Corollary 2.** Suppose that \( \Omega(x^*) \) is a singleton. Then

\[
f''(x, x^*; h) = \langle y^*, w(h) \rangle + \liminf_{h \to 0} \left( g(y + tAh) - g(y) - t \langle x^*, h \rangle \right),
\]

where \( y^* \) is the unique element of \( \Omega(x^*) \).

**Corollary 3.** If \( \Omega(x^*) \neq \emptyset \), then

\[
f''(x, x^*; h) \geq \max_{y^* \in \Omega(x^*)} \left[ \langle y^*, w(h) \rangle + g''(\bar{y}, y^*; Ah) \right].
\]

**Proof.** Indeed, setting \( z(t, h') = Ah' + tw(h) \), and taking an arbitrary \( y^* \in \Omega(x^*) \), we get from the theorem

\[
f''(x, x^*; h) \geq \liminf_{t \to 0} t^{-2}(g(y + tz(t, h')) - g(y) - t \langle y^*, Ah \rangle)
\]

\[
- g(\bar{y}) - t \langle y^*, z(t, h') \rangle + t^2 \langle y^*, w(h) \rangle
\]

\[
\geq \langle y^*, w(h) \rangle + \liminf_{z \to 0} t^{-2}(g(y + z - t^2(\bar{y} + z - t \langle y^*, z \rangle))
\]

\[
= \langle y^*, w(h) \rangle + g''(\bar{y}, y^*; Ah).
\]

Combining Corollaries 1 and 3, we can obtain further useful results. Here is one.

**Corollary 4.** Assume that \( \Omega(x^*) \neq \emptyset \), \( g(\cdot) \) is twice epi-differentiable at \( \bar{y} \) along \( h \) and

\[
g''(\bar{y}, y^*; Ah) = \lim_{t \to +0} t^{-2}(g(y + tAh) - g(y) - t \langle y^*, Ah \rangle) \quad (2.3)
\]

for any \( y^* \in \Omega(x^*) \). Then \( f \) is twice epi-differentiable at \( \bar{x} \) along \( h \) with respect to \( x^* \) and

\[
f''(x, x^*; h) = \max_{y^* \in \Omega(x^*)} \langle y^*, w(h) \rangle
\]

\[
+ \lim_{t \to +0} t^{-2}(g(y + tAh) - g(y) - t \langle x^*, h \rangle)
\]

**Proof.** Corollaries 1 and 3 actually imply that under the assumptions

\[
f''(x, x^*; h) = \max_{y^* \in \Omega(x^*)} \langle y^*, w(h) \rangle
\]

\[
+ \lim_{t \to +0} t^{-2}(g(y + tAh) - g(y) - t \langle x^*, h \rangle).
\]
However, the same arguments as in the proof of Corollary 3 easily show that by virtue of (2.3)

\[ f''_+(x, x^*; h) \leq \max_{y^* \in \Omega(x^*)} \langle y^*, w(h) \rangle + \lim_{t \to 0} t^{-2}(g(y + tAh) - g(y) - t\langle x^*, h \rangle). \]

The last corollary extends Cominetti's generalization \[6\] of Rockafellar's formula \[15\] for second order epi-derivative of a composite function on \( \mathbb{R}^n \) with piecewise linear-quadratic outer function \( g \). (In \[6\] it is assumed in addition to the assumptions of Corollary 4 that \( g \) is Lipschitz on its domain.)

In particular, an obvious and useful consequence of the corollary is

**Corollary 5.** Suppose in addition to (H) that \( g \) is a polyhedral function. If \( \Omega(x^*) \neq \emptyset \), then for any \( h \in K(x^*) \)

\[ f''_-(\bar{x}, x^*; h) = \max_{y^* \in \Omega(x^*)} \langle y^*, w(h) \rangle. \]

The main assumption of Corollary 4 is, of course, (2.3) which means that the second order epi-derivative coincides with the usual second order directional derivative. The following simple example shows that this condition is fairly restrictive.

**Example.** Let \( X = \mathbb{R}, Y = \mathbb{R}^2, F(x) = (-x^2, x), g(y) = g(\xi, \eta) = \max\{0, \xi + \eta^2\} \). Then \( f(x) \equiv 0 \) and for \( \bar{x} = 0 \) we have \( \bar{y} = 0, A = (0, 1), w(h) = (-h^2, 0), K(0) = \mathbb{R}, \Omega(0) = \partial g(0) = \{(\lambda, 0) : 0 \leq \lambda \leq 1\} \).

Therefore \( \max \{ \langle y^*, w(h) \rangle : y^* \in \Omega(0) \} = 0 \) and \( \alpha = \lim_{t \to 0} t^{-2} \max\{0, r^2h^2\} = h^2 \),

so

\[ \alpha + \max_{y^* \in \Omega(0)} \langle y^*, w(h) \rangle = h^2 \neq f''_-(\bar{x}, x^*; h) = 0. \]

The function \( g \) in this example is the maximum of a linear and a quadratic function, hence it is twice epi-differentiable, as follows from \[15\] but \( g \) is not piecewise linear-quadratic.

Another circumstance worth mentioning in connection with this example is that, whereas the formula in Theorem 1 is valid independently of specific \( g \) and \( F \) chosen to represent \( f \), the formula in Corollary 4 is not.
3. Proof of the Theorem

We begin the proof with the analysis of the regularity condition (H). Let \( \phi(u) \) be a mapping from a neighborhood of the origin in \( X \) into \( Y \) such that

\[
\phi(0) = 0 \quad \text{and} \quad \phi(h) - \phi(h') - A(h - h') = r(h, h') \|h - h'\|, \tag{3.1}
\]

where \( r(h, h') \to 0 \) as \( h, h' \to 0 \). The latter means that \( \phi(\cdot) \) is strictly differentiable at the origin and \( \phi'(0) = A \). The condition \( \text{Im } A + L = Y \) then means that the operator \( (u, v) \to Au + v \) sends \( X \times Y \) onto \( Y \). The theorem of Ljusternik (see [13]) now implies that there are \( \varepsilon > 0, C > 0 \) such that

\[
\|u\| < \varepsilon, \quad \|w\| < \varepsilon \Rightarrow \exists x \in X \quad \text{s.t.} \quad \phi(x) + w \in L
\]

and

\[
\|x - u\| \leq C \cdot \text{dist}(L, \phi(u) + w). \tag{3.2}
\]

Suppose also that \( a(x) \) is a real-valued function defined and Lipschitz near the origin. The following lemma plays the crucial role in the proof.

**Basic Lemma.** We assume that (H) holds and \( \partial g(\bar{y}) \neq \emptyset \). Suppose we are given bounded sequences \( \{h_m\} \subset X, \{w_m\} \subset Y \), a sequence \( \{v_m\} \subset Y \) converging to zero and sequences \( \{t_m\} \) and \( \{\alpha_m\} \) of positive numbers converging to zero and such that

\[
g(\bar{y} + \phi(t_m h_m) + \alpha_m(w_m + v_m)) < \infty, \quad m = 1, 2, \ldots. \tag{3.3}
\]

Then there is a sequence \( \{h_m'\} \subset X \) such that \( \|h_m' - h_m\| \to 0 \) and

\[
g(\bar{y} + \phi(t_m h_m') + \alpha_m w_m) + a(t_m h_m') \\
\leq g(\bar{y} + \phi(t_m h_m) + \alpha_m(w_m + v_m)) + a(t_m h_m) + \alpha_m \gamma_m,
\]

where \( \gamma_m \to 0 \).

**Proof.** By (3.3)

\[
\phi(t_m h_m) + \alpha_m(w_m + v_m)) \in \text{dom } g - \bar{y} \subset L. \tag{3.4}
\]

Choose \( \varepsilon_m \to 0 \) such that

\[
\|v_m\|/\varepsilon_m \to 0, \tag{3.5}
\]

set

\[
\delta_m = (\varepsilon_m \alpha_m)/t_m,
\]

and take \( \tilde{u} \in X \) such that \( g \) is continuous at \( \bar{y} + A\tilde{u} \) with respect to \( \bar{y} + L \).
For
\[ e_m = (1 - \delta_m) h_m + \delta_m \tilde{u} \]
we have by (3.4)
\[
\begin{align*}
\text{dist}(L, \phi(t_m e_m) + \alpha_m w_m) & \leq \text{dist}(L, \phi(t_m h_m) + \alpha_m w_m) + \|\phi(t_m e_m) - \phi(t_m h_m)\| \\
& \leq \alpha_m \|v_m\| + c_1 \delta_m t_m \|h_m - \tilde{u}\| \\
& = \alpha_m (\|v_m\| + c_1 \varepsilon_m \|\tilde{u} - h_m\|),
\end{align*}
\]
(3.6)
where \( c_1 \) is a certain constant.

It follows from (3.2) that for every \( m \) there is an \( h'_m \) such that
\[
\phi(t_m h'_m) + \alpha_m w_m \in L
\]
(3.7)
and (by (3.6))
\[
\|h'_m - e_m\| \leq C \alpha_m (\|v_m\| + c_1 \varepsilon_m \|h_m - \tilde{u}\|).
\]
(3.8)

Since \( g \) is continuous at \( \tilde{v} + A\tilde{u} \) with respect to \( \tilde{v} + L \), there are \( \eta > 0 \) and \( k > 0 \) such that
\[
g(\tilde{v} + A\tilde{u} + v) - g(\tilde{v}) \leq k \quad \text{if} \quad v \in L, \quad \|v\| \leq \eta.
\]
(3.9)

Define \( z_m \) by
\[
\phi(t_m h'_m) + \alpha_m w_m = (1 - \delta_m)(\phi(t_m h_m) + \alpha_m (w_m + v_m)) + \delta_m t_m (A\tilde{u} + z_m).
\]
(3.10)

Then \( z_m \in L \) by (3.4) and (3.7) and
\[
z_m = \frac{1}{\varepsilon_m \cdot \alpha_m} \left[ \phi(t_m h'_m) - (1 - \delta_m) \phi(t_m h_m) \right.
\]
\[
- \delta_m t_m A\tilde{u} + \delta_m \alpha_m w_m + (1 - \delta_m) \alpha_m v_m \right].
\]
We have by (3.5)
\[
\frac{1}{\varepsilon_m \cdot \alpha_m} \left[ \delta_m \alpha_m w_m + (1 - \delta_m) \alpha_m v_m \right] \to 0.
\]
(3.11)

On the other hand, since
\[
e_m - h_m = \delta_m (\tilde{u} - h_m),
\]
we have
\[
\phi(t_m h_m') - (1 - \delta_m) \phi(t_m h_m) - \delta_m t_m A \bar{u}
\]
\[
= \left[ \phi(t_m h_m') - \phi(t_m h_m) - t_m A(h_m' - h_m) \right] \\
+ \delta_m \left[ \phi(t_m h_m) - t_m A h_m \right] + t_m A(h_m' - e_m).
\]

By (3.8), (3.5)
\[
(\varepsilon_m \alpha_m)^{-1} A(h_m' - e_m) \to 0 \quad (3.12)
\]
and by (3.1)
\[
t_m^{-1} [\phi(t_m h_m) - t_m A h_m] \to 0. \quad (3.13)
\]

It also follows from (3.1), (3.8), and the definition of \( e_m \) that
\[
\| \phi(t_m h_m') - \phi(t_m h_m) - t_m A(h_m' - h_m) \| \\
\leq r_m \| h_m' - h_m \| \leq r_m (\| h_m' - e_m \| + t_m \| e_m - h_m \| ) \\
\leq r_m \varepsilon_m \alpha_m (C \| v_m \| \varepsilon_m + c_1 \| h_m - \bar{u} \| + \| h_m - \bar{u} \| ), \quad (3.14)
\]
where \( r_m \to 0. \)

Therefore
\[
(\varepsilon_m \alpha_m)^{-1} [\phi(t_m h_m') - \phi(t_m h_m) - t_m A(h_m' - h_m)] \to 0.
\]

Together with (3.1), (3.12), this shows that \( z_m \to 0 \), and for sufficiently large \( m \) we have by (3.9)
\[
\| z_m \| \leq \eta, \quad g(\bar{y} + A \bar{u} + z_m) - g(\bar{y}) \leq k. \quad (3.15)
\]

By (3.10)
\[
g(\bar{y} + \phi(t_m h_m') + \alpha_m w_m) \\
\leq (1 - \delta_m) g(\bar{y} + \phi(t_m h_m) + \alpha_m(w_m + v_m)) \\
+ \delta_m g(\bar{y} + t_m (A \bar{u} + z_m))
\]
or
\[
g(\bar{y} + \phi(t_m h_m') + \alpha_m w_m) + a(t_m h_m') \\
\leq g(\bar{y} + \phi(t_m h_m) + \alpha_m(w_m + v_m)) + a(t_m h_m) \\
- \delta_m [g(\bar{y} + \phi(t_m h_m) + \alpha_m(w_m + v_m)) - g(\bar{y})] \\
+ \delta_m [g(\bar{y} + t_m (A \bar{u} + z_m)) - g(\bar{y})] \\
+ a(t_m h_m') - a(t_m h_m). \quad (3.16)
\]
Since \( a(\cdot) \) is Lipschitz continuous, we conclude using the same estimate as in (3.14) that
\[
|a(t_m h_m') - a(t_m h_m)| \leq \lambda_m \alpha_m, \quad \text{where } \lambda_m \to 0. \tag{3.17}
\]
We have also (by convexity of \( g \))
\[
g(\tilde{y} + t_m(A\tilde{u} + z_m)) - g(\tilde{y}) \leq t_m(g(\tilde{y} + A\tilde{u} + z_m) - g(\tilde{y}))
\]
so by (3.15)
\[
\delta_m \left[ g(\tilde{y} + t_m(A\tilde{u} + z_m)) - g(\tilde{y}) \right] \leq k \cdot \varepsilon_m \alpha_m. \tag{3.18}
\]
And, finally, if \( y^* \in \partial g(\tilde{y}) \), then
\[
g(\tilde{y} + \phi(t_m h_m') + \alpha_m(w_m + v_m)) - g(\tilde{y}) \geq \langle y^*, \phi(t_m h_m') + \alpha_m(w_m + v_m) \rangle \geq -\eta t_m,
\]
where \( \eta \) is a positive constant. Therefore
\[
-\delta_m \left[ g(\tilde{y} + \phi(t_m h_m') + \alpha_m(w_m + v_m)) - g(\tilde{y}) \right] \leq \eta \cdot \varepsilon_m \alpha_m. \tag{3.19}
\]
By comparing (3.16)-(3.19), we complete the proof.

**Proof of the Theorem.** If \( f''_-(\bar{x}, x^*; h) > -\infty \), then, obviously
\[
\liminf_{h', h \to 0} t^{-1} \left[ f(x + tAh') - f(x) \right] \geq \langle x^*, h \rangle.
\]
If this is true for all \( h \), we have
\[
g'(\tilde{y}; Ah) \geq \langle x^*, h \rangle,
\]
which, in view of (H) means that \( x^* \in A^* \partial g(\tilde{y}) \). Hence \( \Omega(x^*) \neq \emptyset \).

On the other hand, if \( y^* \in \Omega(x^*) \), then
\[
f(\bar{x} + th) - f(\bar{x}) \geq \langle y^*, F(\bar{x} + th) - F(\bar{x}) \rangle
\]
\[
= \langle y^*, tAh + t^2w(h) \rangle + o(t^2)
\]
\[
= t^2 \langle y^*, w(h) \rangle + o(t^2),
\]
which implies that \( f''_-(\bar{x}, x^*; h) > -\infty \).

Suppose now that the sequences of \( t_m \to 0 \) and \( h_m \to h \) is such that
\[
t_m^{-2} \left[ f(\bar{x} + t_m h_m) - f(\bar{x}) - t_m \langle x^*, h_m \rangle \right] \to f''_-(\bar{x}, x^*; h) = a. \tag{3.20}
\]
We have

\[ f(\bar{x} + t_m h_m) = g(\bar{y} + t_m A h_m + t_m^2 w_m), \]

where \( w_m \to w(h) \). Set \( v_m = w_m - w(h) \) and apply the basic lemma with

\[ \phi(h) = A h, \quad a(h) = \langle x^*, h \rangle, \quad a_m = t_m^2, \]

to find a sequence \( \{h'_m\} \) converging to \( h \) such that

\[ \limsup_{m \to \infty} \frac{g(\bar{y} + t_m A h'_m + t_m^2 w(h)) - \bar{g}(y) - t_m \langle x^*, h'_m \rangle}{t_m^2} \leq a. \quad (3.21) \]

On the other hand, consider arbitrary sequences of \( t_m \to +0, \ h_m \to h, \)
and \( w_m \to w(h) \), and set in the basic lemma \( \phi(u) = F(x + u) - F(x) \) and \( a(u) \) and \( a_m \) as above. Then

\[ \phi(t_m h_m) = t_m A h_m + t_m^2 w(h) + t_m^2 v'_m, \]

where \( \|v'_m\| \to 0 \). Therefore

\[ t_m A h_m + t_m^2 w_m = \phi(t_m h_m) + t_m^2 v_m \]

with \( \|v_m\| \to 0 \). Applying Basic Lemma, we find a sequence \( \{h'_m\} \to h \) such that

\[
\begin{align*}
&\leq \liminf_{m \to \infty} t_m^{-2} \left[ f(\bar{x} + t_m h'_m) - f(\bar{x}) - t_m \langle x^*, h'_m \rangle \right] \\
&= \liminf_{m \to \infty} t_m^{-2} \left[ g(\bar{y} + \phi(t_m h'_m)) - g(\bar{y}) - t_m \langle x^*, h'_m \rangle \right] \\
&\leq \liminf_{m \to \infty} t_m^{-2} \left[ g(\bar{y} + \phi(t_m h_m) + t_m^2 v_m) - g(\bar{y}) - t_m \langle x^*, h_m \rangle \right] \\
&= \liminf_{m \to \infty} t_m^{-2} \left[ g(\bar{y} + t_m A h_m + t_m^2 w_m) - g(\bar{y}) - t_m \langle x^*, h_m \rangle \right]. \quad (3.22)
\]

Together with (3.21), this proves the formula for \( f''(\bar{x}, x^*; h) \). (We have used the basic lemma in which we assumed that the quantities are finite. But, of course, (3.21) is valid if \( a = \infty \) and (3.22) is valid if the right \( \lim \inf \) is equal to \( \infty \).)

It remains to show that \( h \in K(x^*) \) if \( |f''(\bar{x}, x^*; h)| < \infty \). Take a \( y^* \in \partial g(\bar{y}) \). Then for any \( v \)

\[ g(\bar{y} + v) - g(\bar{y}) \geq \langle y^*, v \rangle. \]
so that for any \( t_m \to 0, h_m \to h \) satisfying (3.20), we have

\[
0 = \lim_{m \to \infty} t_m^{-1} \left[ f(\bar{x} + t_m h_m) - f(\bar{x}) - t_m \langle x^*, h_m \rangle \right]
\]

\[
= \lim_{m \to \infty} t_m^{-1} \left[ g(\bar{y} + t_m Ah_m + r_m) - g(\bar{y}) - t_m \langle x^*, h_m \rangle \right]
\]

(\text{where } \|r_m\| = o(t_m))

\[
\geq \lim_{m \to \infty} \left[ \langle y^*, Ah_m \rangle - \langle x^*, h_m \rangle \right] = \langle y^*, Ah \rangle - \langle x^*, h \rangle.
\]

This completes the proof of the theorem.

4. AN APPLICATION: THE MAX-FUNCTION

This is the function of the form

\[
f(x) = \max_{q \in Q} f(q, x). \quad (4.1)
\]

Here \( Q \) is a compact metrizable space and \( f(q, x) \) is continuous near \( \bar{x} \) (as a function of \( (q, x) \)) together with its first and second derivatives with respect to \( x: f_x(q, x) \) and \( f_{xx}(q, x) \). It is obvious that \( f(x) \) belongs to class \( CC^2(\bar{x}) \). Here is the “dictionary” of the notation:

\[
Y = C(Q)
\]

\[
F: x \to y(q) = f(q, x)
\]

\[
g(y) = \max_{q \in Q} y(q)
\]

\[
\bar{y}(q) = f(q, \bar{x}); \quad Ah = f_x(q, \bar{x});
\]

\[
\Omega(x^*) = \left\{ \mu \in P(Q_0) : \int f_x(q, \bar{x}) \, d\mu = x^* \right\},
\]

where \( Q_0 = \{ q \in Q : f(q, \bar{x}) = \max_{p \in Q} f_x(p, \bar{x}) \} \) and \( P(Q_0) \) is the collection of probability measures supported on \( Q_0 \). Then

\[
\hat{\partial}g(\bar{y}) = P(Q_0);
\]

\[
K(x^*) = \{ h : f_x(q, \bar{x})h \leq \langle x^*, h \rangle, \forall q \in Q_0 \}.
\]

We also use the following notation in the rest of the section:

\[
u(q) = f(q, \bar{x}) - f(\bar{x});
\]

\[
w(q) = f_{xx}(q, \bar{x})(h, h).
\]
ANALYSIS OF A COMPOSITE FUNCTION

PROPOSITION 2. If \( Q_0 = Q \), then for any \( h \in K(x^*) \)

\[
f''(\bar{x}, x^*; h) = \max_{\mu \in \Omega(x^*) \cdot Q_0} \int w(q, h) \, d\mu.
\]

(4.2)

Proof. In this case \( \bar{y}(q) = \text{const} - \bar{y} \) and

\[
g(y) = \max_q (y(q) - \bar{y}) + \bar{y},
\]

so we can apply Corollary 4.

Let us turn to the general case

\( Q_0 \neq Q \).

For \( t > 0 \) we set

\[
c(t, q, h') = t^{-2}[u(q) + tv(q)h + t^2w(q, h)].
\]

The notation

\[
\mu \to \Omega(x^*)
\]

will mean that

\[
\mu \in P(Q), \quad \left\| \int v(q) \, d\mu \right\| \to 0, \quad \left\| \int u(q) \, d\mu \right\| \to 0.
\]

(4.3)

As \( u(q) < 0 \) outside of \( Q_0 \), this implies that the distance from \( \mu \) to \( \Omega(x^*) \)
(w.r.t. any metric that induces the weak* topology on \( P(Q) \)) tends to zero
(and is equivalent to this property if \( \dim X < \infty \)).

We use the notation \( \lim \inf \sup \) in the same sense as in [15].

THEOREM 2. Suppose that \( \Omega(x^*) \neq \emptyset \) and \( h \in K(x^*) \). Then

\[
f''(\bar{x}, x^*; h) = \lim_{t \to +0} \inf_{\mu \to \Omega(x^*)} \sup_{\mu \in \Omega(x^*)} \int c(t, q, h) \, d\mu.
\]

Proof. 1. Using the Ky Fan minimax theorem and taking into account that \( c(t, q, h) \) is linear continuous w.r.t. \( h' \) and \( P(Q) \) is convex weak* compact, we have for \( t > 0 \)

\[
\min_{\|h - h'\| < \varepsilon} \max_{q \in Q} c(t, q, h')
\]

= \[
\min_{\|h - h'\| < \varepsilon} \max_{\mu \in P(Q)} \int c(t, q, h') \, d\mu
\]

= \[
\max_{\mu \in P(Q)} \left[ \int c(t, q, h) \, d\mu - st^{-1} \left\| \int v(q) \, d\mu \right\| \right].
\]

(4.4)
Denote by $b_{\varepsilon}(t, \mu)$ the quantity in the square brackets, and let

$$a_{\varepsilon} = \lim \inf_{t \to +0} \max_{\mu \in \mathcal{P}(Q)} b_{\varepsilon}(t, \mu).$$

Obviously, $a_{\varepsilon}$ does not decrease as $\varepsilon \to 0$. It is also clear that $a_{\varepsilon} > -\infty$ because for $\mu \in \Omega(x^*)$, $b_{\varepsilon}(t, \mu) = \int w(q, h) \, d\mu$.

According to Theorem 1

$$a = f^\prime(\bar{x}, x^*; h) = \lim \inf_{t \to +0} \max_{\mu \in \mathcal{P}(Q)} c(t, q, h')$$

and therefore by (4.4)

$$a = \sup_{\varepsilon > 0} a_{\varepsilon} = \sup_{\varepsilon > 0} \lim \inf_{t \to +0} \max_{\mu \in \mathcal{P}(Q)} b_{\varepsilon}(t, \mu). \quad (4.5)$$

2. To prove the theorem we must show that

(a) for any sequence of $t_m \to +0$ there is a sequence of measures $\mu_m \to \Omega(x^*)$ such that

$$\lim \sup_{m \to \infty} \int c(t_m, q, h) \, d\mu_m \geq a;$$

(b) there is a sequence of $t_m \to +0$ such that

$$\lim \sup_{m \to \infty} \int c(t_m, q, h) \, d\mu_m \leq a.$$

The proof of the second is elementary. Fix an $\varepsilon > 0$ and let $t_m \to +0$ be such that

$$a_{\varepsilon} = \lim_{m \to \infty} \max_{\mu \in \mathcal{P}(Q)} b_{\varepsilon}(t_m, \mu).$$

Then for any $\mu_m \to \Omega(x^*)$ we have by (4.4)

$$\lim \sup_{m \to \infty} \int c(t_m, q, h) \, d\mu_m$$

$$= \lim \sup_{m \to \infty} b_{\varepsilon}(t_m, \mu_m)$$

$$\leq \lim \sup_{m \to \infty} \max_{\mu \in \mathcal{P}(Q)} b_{\varepsilon}(t_m, \mu) = a_{\varepsilon} \leq a.$$
3. It remains to prove (a). Fix a sequence of $t_m \to +0$. Then by (4.4) for any $\varepsilon > 0$

$$a_\varepsilon \leqslant \limsup_{m \to \infty} \max_{\mu \in P(Q)} b_\varepsilon(t_m, \mu) = b_\varepsilon.$$ Clearly, $b_\varepsilon$ do not increase as $\varepsilon \to +0$ and $b = \lim b_\varepsilon \geqslant a_\varepsilon$.

If $b = \infty$, then there is a sequence of $\mu_m \in P(Q)$ such that the limit superior of

$$\xi_m = \int c(t_m, q, h) \, d\mu_m$$

equals $\infty$. Let $m(k)$ be a sequence of indices such that $\xi_{m(k)}$ tends to infinity, let $\mu' \in \Omega(x^*)$ and

$$\lambda_m = \begin{cases} \frac{1}{\sqrt{\xi_{m(k)}}}, & \text{if } m = m(k), k = 1, 2, \ldots; \\ 0, & \text{otherwise} \end{cases}$$

$$\mu'_m = \lambda_m \mu_m + (1 - \lambda_m) \mu'.$$

Then $\lambda_m \to 0$ and, consequently, $\mu_m \to \Omega(x^*)$. On the other hand,

$$\lambda_m \xi_{m(k)} = \int c(t_{m(k)}, q, h) \, d\mu'_{m(k)} = \lambda_m \xi_{m(k)} + \int w(q, h) \, d\mu' \to \infty.$$ This proves (a) in the case when the sequence $\{t_m\}$ is such that $b = \infty$.

Suppose now that $b < \infty$. Let $\varepsilon$ be such that

$$\int c(t_m, q, h) \, d\mu'_m = \int c(t_m, q, h) \, d\mu'_m \to \infty.$$ We have

$$a_\varepsilon \leqslant \liminf_{m \to \infty} b_\varepsilon(t_m, \mu_m)$$

$$\leqslant \limsup_{m \to \infty} b_\varepsilon(t_m, \mu^{\varepsilon}_m)$$

$$\leqslant \limsup_{m \to \infty} b_{\varepsilon/2}(t_m, \mu^{\varepsilon}_m)$$

$$\leqslant b_{\varepsilon/2} - \varepsilon \cdot \limsup_{m \to \infty} \left( (2t_m)^{-1} \left\| \int v(q) \, d\mu^\varepsilon_m \right\| \right).$$
and, consequently,
\[
\limsup_{m \to \infty} \left\| \int v(q) \, d\mu_m \right\| \leq \limsup_{m \to \infty} (2t_m/\varepsilon)(b - a) = 0.
\]

It follows that
\[
a < \liminf_{m \to \infty} b_{m}(t_m, \mu_m) = \liminf_{m \to \infty} \int c(t_m, q, h) \, d\mu_m
\]
\[
\leq \liminf_{m \to \infty} \left[ t_m^{-2} \int u(q) \, d\mu_m + t_m^{-1} \|h\| \cdot \left\| \int v(q) \, d\mu_m \right\| + \int w(q, h) \, d\mu_m \right]
\]
\[
\leq \max_q w(q, h) + \liminf_{m \to \infty} \left( t_m^{-2} \int u(q) \, d\mu_m \right)
\] (4.6)

In other words,
\[
\liminf_{m \to \infty} \left( t_m^{-2} \int u(q) \, d\mu_m \right) > -\infty
\]

which may happen only if \( \int u(q) \, d\mu \to 0 \) since \( u(q) < 0 \) for all \( q \).

Thus \( \mu_m \to \Omega(x^*) \) and the first two relations in (4.6) give
\[
\liminf_{m \to \infty} \int c(t_m, q, h) \, d\mu_m \geq a.
\]

Since \( a > a < \infty \), (a) easily follows.

Remark 1. Slightly changing the argumentation, we can prove the following: if \( \Omega(x^*) \neq \emptyset \), \( h \in K(x^*) \), then ([15])
\[
f''_+(\bar{x}, x^*; h) \triangleq \limsup_{t \to +0} \inf_{h \to h} \frac{(f(\bar{x}+th')-f(\bar{x})-t\langle x^*, h' \rangle)}{t^2}
\]
\[
= \limsup_{t \to +0} \int c(t, q, h) \, d\mu.
\]

Recall that \( f \) is twice epi-differentiable with respect to \( x^* \) if \( f''_+(\bar{x}, x^*; h) \) and \( f''_-(\bar{x}, x^*; h) \) coincide for all \( h \). The following theorem gives a sufficient condition for the max-function to be twice epi-differentiable. Moreover, it follows from the proof of the theorem that the condition is actually almost necessary.
Consider the function \( e(q, h) \) on \( Q \) defined as
\[
e(q, h) = \begin{cases} 
0, & \text{if } q \in \text{int } Q_0; \\
l(q, h), & \text{if } q \notin Q_0; \\
\limsup_{p \to q, p \notin Q_0} l(p, h), & \text{if } q \in Q_0 \setminus \text{int } Q_0,
\end{cases}
\]
where for \( q \notin Q_0 \)
\[
l(q, h) = -\left((v(q)h)^+\right)^2/4u(q)
\]
(and, as usual, \( z^+ = \max\{z, 0\} \)).

Since \( u(q) < 0 \) if \( q \notin Q_0 \), the function is well defined, nonnegative, and upper semicontinuous on \( Q \).

In what follows we adopt the convention
\[
\int e(q, h) \, d\mu = \infty \quad \forall \mu \in \Omega(x^*) \quad \text{if } e(q, h) = \infty \text{ for some } q
\]
(such a \( q \) necessarily belongs to \( Q_0 \setminus \text{int } Q \)).

**THEOREM 3.** If \( \Omega(x^*) \neq \emptyset \) and \( h \in K(x^*) \), then
\[
\limsup_{\mu \to 0} \int c(t, q, h) \, d\mu \leq \max_{\mu \in \Omega(x^*)} \int \left[ e(q, h) + w(q, h) \right] \, d\mu. \quad (4.7)
\]

If in addition
\[
(H1) \quad Q \text{ is locally connected and for any } q \in Q_0 \setminus \text{int } Q_0 \text{ with } e(q, h) > 0 \quad \text{the limit } \lim_{p \to q, p \in Q_0} l(p, h) \exists \text{ then actually}
\]
\[
\liminf_{t \to 0} \sup_{\mu \in \Omega(x^*)} \int c(t, q, h) \, d\mu = \max_{\mu \in \Omega(x^*)} \int \left[ e(q, h) + w(q, h) \right] \, d\mu.
\]

Thus, in this case \( f \) is twice epi-differentiable with respect to \( x^* \) and
\[
f''(\bar{x}, x^*; h) = \max_{\mu \in \Omega(x^*)} \int \left[ e(q, h) + w(q, h) \right] \, d\mu.
\]

**Proof.** 1. We first show that (4.7) is valid. If \( e(q, h) = \infty \) for some \( q \in Q_0 \), this follows from the convention. So suppose that \( e(q, h) < \infty \) everywhere. We observe first that
\[
c(t, q, h) \leq l(q, h) + w(q, h) \quad \text{if } q \notin Q_0. \quad (4.8)
\]
Indeed, in this case \( u(q) < 0 \); if also \( v(q) h \leq 0 \) then \( l(q, h) = 0 \) and \( c(t, q, h) \leq w(q, h) \) by definition. If \( v(q) h > 0 \), then
\[
\max_t (t^{-2}u(q) + t^{-1}v(q) h) = l(q, h) \tag{4.9}
\]
and the maximum is attained at
\[
t = -2u(q)/(v(q) h) > 0. \tag{4.10}
\]
On the other hand,
\[
c(t, q, h) = w(q, h) \quad \text{if} \quad q \in Q_0. \tag{4.11}
\]
It follows from (4.8), (4.11) that
\[
c(t, q, h) \leq e(q, h) + w(q, h)
\]
for all \( t > 0, q \in Q \) and, as \( e(q, h) \) is u.s.c., (4.7) follows.

2. Assume now that (H1) is valid. If \( e(q, h) = 0 \) on \( Q_0 \), then by (4.11)
\[
\liminf_{t \to +0} \sup_{\mu \in \Omega(x^*)} \int c(t, q, h) \, d\mu = \max_{\mu \in \Omega(x^*)} \int w(q, h) \, d\mu = \max_{\mu \in \Omega(x^*)} \int [e(q, h) + w(q, h)] \, d\mu.
\]
Thus the theorem is true if \( e(q, h) = 0 \) on \( Q_0 \), so we assume in what follows that
\[
Q_0^+ = \{ q \in Q_0 : e(q, h) > 0 \} \neq \emptyset.
\]
Take a \( q \in Q_0^+ \). Then \( v(p) h > 0 \) for \( p \notin Q_0 \) sufficiently close to \( q \) and, as
\[
\limsup_{\substack{p \to q \\
p \notin Q_0}} [v(q) h]^2/[ -u(q)]
\]
exists and is positive,
\[
\lim_{\substack{p \to q \\
p \notin Q_0}} [-u(p)]/v(p) h = 0 \tag{4.12}
\]
since \( v(p) h \to 0 \) when \( p \to q \in Q_0 \) (recall that \( h \in K(x^*) \), so \( v(q) h \leq 0 \)).
We must show that for any sequence of $t_m \to +0$

$$\liminf_{m \to \infty} \sup_{\mu \in \Omega(x^*)} \int c(t_m, q, h) \, d\mu$$

$$\geq \max_{\mu \in \Omega(x^*)} \left[ \int [e(q, h) + w(q, h)] \, d\mu \right]. \quad (4.13)$$

Suppose there is a $q \in Q_0^+$ with $e(q, h) = \infty$. By the assumption, $q$ has a connected neighborhood. In view of (4.12) this means that

$$-u(p)/[v(p)h]$$

assumes all values between certain $\tau_0 > 0$ and $0$ as $p \to q$ in the complement of $Q_0$. Therefore since for a given sequence $\{t_m\}$ there is an $m_0$ such that $0 < t_m < 2\tau_0$ if $m \geq m_0$, we can find $p_m \notin Q_0$ such that $p_m \to q$ and

$$t_m = -2u(p_m)/[v(p_m)h]$$

as follows from (4.9), (4.10)

$$\xi_m = c(t_m, p_m, h) = l(p_m, h) + w(p_m, h) \to \infty.$$ (Take $\lambda_m = 1/\sqrt{\xi_m}$, take any $\mu \in \Omega(x^*)$ and set

$$p_m = \lambda_m e(p_m) + (1 - \lambda_m)\mu,$$ where $e(p)$ is the unit mass at $p$.)

This proves the theorem in the case when $e(q, h) = \infty$ somewhere in $Q_0$.

3. It remains to prove that (4.13) is valid also if

$$0 < \max_{q \in Q} e(q, h) < \infty. \quad (4.15)$$

Take a $\mu \in \Omega(x^*)$ at which the maximum in the right-hand part of (4.13) is attained. Since $Q_0$ is compact, $v(q)$ and $u(q)$ are both continuous and $e(q, h)$ is u.s.c., there exists a sequence $\{v_k\}$ of probability measures on $Q_0$ with finite supports such that
\[ v_k \to \mu \text{ weakly}^* \]
\[ \left\| \int v(q) \, dv_k \right\| \to 0 \]  
(4.16)
\[ \int u(q) \, dv_k \to 0 \]
\[ \lim \sup_{k \to \infty} \int e(q, h) \, dv_k \geq \int e(q, h) \, d\mu. \]

Such measures can be constructed, for example, as follows. Let \( d(\cdot, \cdot) \) be any metric on \( Q \) compatible with the weak-star topology. For any \( k \), let \( \{ U_{ik} \}, i = 1, 2, \ldots, m(k) \), be a collection of open subsets of \( Q \) with diameters smaller than \( 1/k \), which covers \( Q_0 \) and has the property that
\[ \| v(p) - v(q) \| \leq 1/k, \quad |u(p) - u(q)| \leq 1/k, \]
if \( p, q \in U_{ik} \), \( i = 1, \ldots, m(k) \).

Let \( q_{ik} \) be defined by
\[ q_{ik} \in \text{cl} \, U_{ik}, \]
\[ e(q_{ik}, h) = \max \{ e(q, h) : q \in \text{cl} \, U_{ik} \}, \quad i = 1, \ldots, m(k). \]

Then the measures
\[ v_k = \sum_{i=1}^{m(k)} \alpha_{ik} \cdot e(q_{ik}), \]
where
\[ \alpha_{1k} = \mu(U_{1k}), \quad \alpha_{2k} = \mu(U_{2k} \setminus U_{1k}), \ldots, \quad \alpha_{mk} = \mu \left( U_{mk} \bigcup_{i=1}^{m-1} U_{ik} \right), \ldots \]
satisfy (4.16).

Let now
\[ \alpha(p) = -\frac{[v(p)h]}{2u(p)}. \]

This function is continuous on \( Q_1 = Q \setminus Q_0 \) and, as follows from (4.12), \( \alpha(p) \to \infty \) when \( p \to q \in Q_0^+ \). Since \( Q \) is locally connected, it follows that for any \( q \in Q_0^+ \) and any \( \varepsilon > 0 \) the function \( \alpha(p) \) assumes on \( P(\varepsilon, q) = \{ p \in Q_1 : d(p, q) \leq \varepsilon \} \) all values between certain \( \alpha_0 \) (depending on \( q \) and \( \varepsilon \)) and \( \infty \). In other words,
\[ t(q, \varepsilon) = \inf \{ t > 0 : t^{-1} \notin \alpha(P(q, \varepsilon)) \} > 0 \]
for any $q \in Q^+_0$ and $\varepsilon > 0$. This means that for any $k, r = 1, 2, \ldots$ there are $m = m(k, r)$ such that $m(k, r) \to \infty$ as $r \to \infty$ and

$$t_m < t(q_{ik}, r^{-1}) \quad \text{if} \quad m \geq m(k, r),$$

$$i = 1, 2, \ldots, m(k) \text{ and } q_{ik} \in Q^+_0.$$ 

Therefore for $m \geq m(k, r)$ and $i = 1, 2, \ldots, m(k)$ such that $q_{ik} \in Q^+_0$ we can find $p_{imk} \in Q_1$ such that

$$\alpha(p_{imk}) = t_m^{-1} \quad \text{and} \quad d(q_{ik}, p_{imk}) \leq r^{-1}.$$  \hspace{1cm} (4.17)

If $q_{ik} \notin Q^+_0$ (that is if $e(q_{ik}, h) = 0$), we set $p_{imk} = q_{ik}$.

We now define measures $\mu_{nk}$ as follows:

$$\mu_{nk} = \sum_{i = 1}^{m(k)} \alpha_{ik} \cdot e(p_{imk}), \quad \text{if} \quad m(k, r + 1) > m \geq m(k, r).$$

It is clear that (by (4.17))

$$\left\| \int v(q) \, d\mu_{nk} - \int v(q) \, dv_k \right\| \to 0$$

$$\int u(q) \, d\mu_{nk} \to \int u(q) \, dv_k$$ \hspace{1cm} (4.18)

$$\mu_{nk} \to v_k \text{ weakly*}$$

when $m \to \infty$.

On the other hand, since (by (4.9) and (4.10))

$$t_m^{-2} u(p) + t_m^{-1} v(p) h = l(p, h)$$

if $\alpha(p) = t_m^{-1}$, we have (by (4.17))

$$c(t_m, p_{imk}, h) = e(p_{imk}, h) + w(p_{imk}, h)$$

and it follows from the assumptions that

$$c(t_m, p_{imk}, h) \to e(q_{ik}, h) + w(q_{ik}, h), \quad i = 1, \ldots, m(k)$$

when $m \to \infty$. Therefore

$$\int c(t_m, q, h) \, d\mu_{nk} \to \int [e(q, h) + w(q, h)] \, dv_k \quad \text{as} \quad m \to \infty.$$  

Together with (4.16) and (4.18), this completes the proof of the theorem.
Remark 2. The following result was proved by Kawasaki [14] (in the case when $x^* = 0$).

**Theorem 4.** Assume that $\Omega(x^*) \neq \emptyset$ and that $h \in K(x^*)$. Then

$$d^2_+ f(\bar{x}, x^*; (h, z)) = \max_{q \in Q_0(h)} [e(q, h) + w(q, h) + v(q)z],$$

where

$$d^2_+ f(\bar{x}, x^*; (h, z)) = \lim_{t \to 0} \sup \lim_{n \to \infty} t^{-2} [f(\bar{x} + th + (t^2/2)z) - f(\bar{x}) - t\langle x^*, h \rangle].$$

is the upper "parabolic" second derivative in the spirit of Ben-Tal and Zowe [2] and

$$Q_0(h) = \{ q \in Q_0 : v(q)h = 0 \}.$$

Remark 3. We observe that this definition of parabolic second derivative differs slightly from that given in [15].

**Proof.** We first observe that $Q_0(h) \neq \emptyset$ (because $v(q)$ is continuous, $\Omega(x^*) \neq \emptyset$ and, consequently, $\max\{v(q)h : h \in Q_0\} \geq 0$ for all $h$).

It is clear that

$$d^2_+ f(\bar{x}, x^*; (h, z)) = \lim_{n \to \infty} \max_{q \in Q_0(h)} \sup \lim_{t \to 0} t^{-2} c(t, q, h) + v(q)z, \quad (4.19)$$

where $c(\cdot, \cdot, \cdot)$ is as above. Therefore

$$d^2_+ f(\bar{x}, x^*; (h, z)) \geq \max_{q \in Q_0(h)} [w(q, h) + v(q)z] \quad (4.20)$$

because $c(t, q, h) = w(q, h)$ for $q \in Q_0$.

Suppose $q \in Q_0(h)$ and $q_n \to q$ are such that $l(q_n, h) \to e(q, h) > 0$. Then by definition $q_n \notin Q_0$, $u(q_n) < 0$, $v(q_n)h > 0$ and $v(q_n)h \to 0$. Consequently, as follows from the definition of $l(q, h)$,

$$t_n = -2u(q_n)/v(q_n)h \to 0. \quad (4.21)$$

Therefore (see (4.9), (4.10))

$$d^2_+ f(\bar{x}, x^*; (h, z)) = \lim_{n \to \infty} t_n^{-2} c(t_n, q_n, h) + v(q_n)z] = e(q, h) + w(q, h) + v(q)z. \quad (4.22)$$
Together with (4.20) this shows that
\[ d^2_+ f(\bar{x}, x^*; (h, z)) \geq \max_{q \in Q_0(h)} [e(q, h) + w(q, h) + v(q)z]. \] (4.23)

Assume now that \( \tau_n \to 0, q_n \to 0 \) are such that
\[ \tau_n^{-2}(c(\tau_n, q_n, h) + v(q_n)z) \to d^2_+ f(\bar{x}, x^*; (h, z)). \]

We may assume, of course that \( q_n \) converges to a certain \( q \). Then \( q \in Q_0 \) for otherwise \( u(q) < 0 \) and \( \tau_n^{-2}c(\tau, q, h) \) has the order of \( \tau_n^{-2}u(q) \to -\infty \).
Likewise, we conclude that \( q \in \Omega_0(h) \) for otherwise we would have \( v(q)h < 0 \) which again implies that \( d^2f(\bar{x}, x^*; (h, z)) = -\infty \). The latter is impossible in view of Theorem 1 because \( \Omega(x^*) \neq \emptyset \).

If \( \lim \sup \tau_n^{-2}[u(q_n) + \tau_n v(q_n)h] \leq 0 \), then
\[ d^2_+ f(\bar{x}, x^*; (h, z)) \leq w(q, h) + v(q)z \]
\[ \leq e(q, h) + w(q, h) + v(q)z. \] (4.24)

Otherwise (see again (4.9), (4.10))
\[ 0 < \lim \sup_{n \to \infty} \tau_n^{-2}[u(q_n) + \tau_n v(q_n)h] \]
\[ \leq \lim \sup_{n \to \infty} l(q_n, h) = e(q, h), \]

and we again arrive at (4.24) in view of (4.19).

How are this and the above results connected? Of course, the upper directional derivative has a distant relation to minimization. But under the conditions of the second part of Theorem 3 the epi-derivative can be calculated with the help of the upper parabolic derivative.

**Theorem 5.** Assume that \( \Omega(x^*) \neq \emptyset \) and (H1) holds for all \( h \in K(x^*) \). Then

(a) \( f \) is twice epi-differentiable at \( \bar{x} \) with respect to \( x^* \);
(b) the parabolic derivative (finite or infinite) with respect to \( x^* \) exists for any \( h \) and \( z \);
(c) \( f''(\bar{x}, x^*; h) = \inf_z d^2f(\bar{x}, x^*; (h, z)) \).

**Proof.** Part (a) is a part of Theorem 3. If \( h \in K(x^*) \), then (b) also follows from the proof of Theorem 3 for, if (H1) holds, then, as shown there, for any sequence of \( t \to 0 \) and any \( q \) such that \( e(q, h) \) is positive we can find a sequence of \( q_n \to q \) such that
\[ \tau_n^{-2}[u(q_n) + \tau_n v(q_n)h] \to e(q, h). \]
Finally, if \( h \notin K(x^*) \), then there is a \( q \in Q_0 \) such that \( v(q)h > 0 \) and for such a \( q \) \( t_n^{-2}c(t_n, q, h) + v(q)z \to \infty \) for any sequence \( t_n \to 0 \). Thus, (b) is valid.

It remains to prove (c) in the case when \( h \in K(x^*) \) and \( e(q, h) < \infty \) for all \( q \in Q_0 \). (If \( h \notin K(x^*) \) then \( f'_-(\bar{x}, x^*; h) = \infty \) by Theorem 1; if \( e(q, h) = \infty \) for some \( q \in Q \) then \( f'_-(\bar{x}, x^*; h) = \infty \) by Theorem 3 and it is obvious that \( f'_-(\bar{x}, x^*; h) \leq d_{\infty}^2 f(\bar{x}, x^*; (h, z)) \) for any \( z \).)

Thus, we must prove that

\[
\inf_{z} \max_{q \in Q_0(h)} [e(q, h) + w(q, h) + v(q)z]
\]

This easily follows from the Ky Fan theorem according to which

\[
\inf_{z} \max_{q \in Q_0(h)} [e(q, h) + w(q, h) + v(q)z] = \max_{\mu \in \Omega(x^*)} \left[ \int [e(q, h) + w(q, h)] d\mu + \inf_{z} \int v(q)z d\mu \right].
\]

The infimum inside the square brackets is distinct from \(-\infty\) if and only if \( \int v(q) d\mu = 0 \), that is, if \( \mu \in \Omega(x^*) \).

**Acknowledgments**

The research was initiated and, to a large extent, completed during my visits to the Department of Mathematics of the Ohio State University and to The Centre de Recherche Mathématiques of the University of Montreal in Summer 1989. It is my pleasure to thank both institutions and, specifically, Boris Mityagin and Frank Clarke for the excellent research opportunities I was given.

*Note added in proof.* J.-P. Penot brought to my attention the fact that it is possible to prove the main theorem under the regularity condition of Robinson: \( Q \in \text{int}[\text{dom } g + \lim A] \) which is weaker than \((H)\).

**References**