Dualization of the Euler and Hamiltonian inclusions\textsuperscript{1}

D.N. Bessis\textsuperscript{a,\*}, Yu.S. Ledyaev\textsuperscript{b}, R.B. Vinter\textsuperscript{a}

\textsuperscript{a}Department of Electrical Engineering, Imperial College of Science, Technology and Medicine, Exhibition Road, London, SW7 2AZ, UK
\textsuperscript{b}Department of Mathematics, Western Michigan University, Kalamazoo, MA, 49008-4510, USA

Received 1 December 1997; accepted 12 March 1999

Dedicated to Francis Clarke on the occasion of his 50th birthday

\textbf{Keywords:} Lagrangian; Hamiltonian; Dualization; Local epi-continuity; Local epi-Lipschitzness

1. Introduction

Optimal control emerged as a distinct field of research in the late 1950s with the derivation of the Pontryagin maximum principle, a fundamental set of necessary conditions on minimizing arcs which takes account of differential constraints. The Pontryagin maximum principle directly generates the Euler equation and other classical necessary conditions from the calculus of variations, when specialised to problems without dynamic constraints.

However, deeper connections between necessary conditions in optimal control and the calculus of variations were only subsequently revealed when Clarke developed new analytic machinery, now referred to as Nonsmooth Analysis (a term introduced by Clarke himself), for extending classical necessary conditions to apply to the generalized

\textsuperscript{1}We gratefully acknowledge the financial support of the EPSRC, the NSERC and the Russian Fund For Fundamental Research.

\* Corresponding author.

\textit{E-mail addresses:} dorothee.bessis@ima.unil.ch (D.N. Bessis), ledyaev@math-stat.wmich.edu (Yu.S. Ledyaev), r.vinter@ic.ac.uk (R.B. Vinter).
nonconvex problem of the calculus of variations:

\[
\text{Minimize } \int_0^1 L(x(t), \dot{x}(t)) \, dt + l(x(0), x(1))
\]

over some set of arcs \( x \)

with data the functions \( L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \) and \( l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \).

Since the Lagrangian is extended-valued, the above problem formulation can accommodate a wide range of constraints, through the introduction of penalty terms, including the differential constraints associated with optimal control. Thus necessary conditions for general nonconvex optimal control problems can be interpreted as necessary conditions for calculus of variations problems with nonsmooth data. These developments were pre-figured, in the context of fully convex problems, by Rockafellar [18].

Over the past two decades, there has been great interest in deriving necessary conditions for the above optimization problem, valid under different hypotheses and providing (in some cases) different information about minimizers. All these conditions assert the existence of an adjoint arc \( p \) satisfying certain conditions in relation to the minimizer \( x \) in question. The main distinction between them is the precise form of the ‘adjoint equation’ (i.e. the differential equation or related condition satisfied by \( p \)).

Classical conditions from the calculus of variations, namely the Euler condition

\[
(\dot{p}(t), p(t)) = \nabla L(\dot{x}(t), \dot{x}(t)), \tag{1}
\]

and Hamilton’s canonical equations

\[
(−\dot{p}(t), \dot{x}(t)) = \nabla H(\dot{x}(t), p(t)), \tag{2}
\]

in which the Hamiltonian \( H \) is defined to be, for all \((x, p)\),

\[
H(x, p) = \max_{v \in \mathbb{R}^n} \{ (p, v) − L(x, v) \},
\]

serve as prototypes for many of these conditions.

In Clarke’s early and seminal work in this area [1,3,6], there appear extensions to nonclassical settings of both the Euler and the Hamiltonian equations, namely the following. We find nonsmooth (and nonconvex) version of (1):

\[
(\dot{p}(t), p(t)) \in \text{co} \hat{c}_L L(\dot{x}(t), \dot{x}(t)) \quad \text{a.e. } t; \tag{3}
\]

and the following nonsmooth (and nonconvex) version of (2):

\[
(−\dot{p}(t), \dot{x}(t)) \in \text{co} \hat{c}_L H(\dot{x}(t), p(t)) \quad \text{a.e. } t. \tag{4}
\]

(Here \( \hat{c}_L \) and \( \hat{c}_H \) denote appropriately defined nonsmooth ‘subgradients’ of \( L \) and \( H \), respectively.)

These conditions are distinct (to the extent that particular problems can be constructed in which one can be used to exclude from the set of minimizers certain arcs when the other fails to do so, and vice versa), though Clarke did establish some relationship between them by showing that the Hamilton inclusion (4) can be derived in the limit by applying the extended Euler inclusion to a family of perturbed problems.
In subsequent work, aimed at sharpening necessary conditions of this nature and weakening the hypotheses under which they are valid, condition (3) has been replaced by the inclusion

$$\dot{p}(t) \in \text{co} \{ q : (q, p(t)) \in \partial L(\tilde{x}(t), \tilde{\xi}(t)) \} \quad \text{a.e. } t. \quad (5)$$

The difference is that ‘convexification’ is carried out only with respect to the velocity variable on the right; this partially convexified version of the Euler condition is some improvement on (3), in part because it provides, in theory, more information about minimizers and in part because (5) is preserved under certain kinds of limit taking for which (3) is not. The latter property can be useful when we attempt to derive new necessary conditions by applying known necessary conditions to a sequence of approximating problems and passing to the limit (‘bootstrapping’).

Likewise, condition (4) can be replaced by the following sharper, partially convexified, version [16,21]:

$$\dot{p}(t) \in \text{co} \{ q : (-q, \tilde{\xi}(t)) \in \partial L(\tilde{x}(t), p(t)) \} \quad \text{a.e. } t. \quad (6)$$

Some relevant references are [10,13–15]. A detailed discussion of the historical background of conditions akin to (3) and (4) is provided in a number of papers and books (e.g. [2,4–6,10–12,19,22,23]).

In a recent paper [19], Rockafellar has substantially unified the various necessary conditions that had previously been derived, related to the Euler condition and Hamilton’s system of equations: fix $t \in [0,1]$ such that $L(\tilde{x}(t), \tilde{\xi}(t)) < +\infty$ and suppose that

(i) $L$ is lower semi-continuous and $L(x, \cdot)$ is convex for each $x$,  
(ii) $L$ is locally epi-Lipschitz near $(\tilde{x}(t), \tilde{\xi}(t))$,  
(iii) $L$ is epi-continuous, then, given any $p(t)$ in $\mathbb{R}^n$, we have

$$\text{co} \{ q : (q, p(t)) \in \partial L(\tilde{x}(t), \tilde{\xi}(t)) \} = \text{co} \{ q : (-q, \tilde{\xi}(t)) \in \partial L(\tilde{x}(t), p(t)) \}. \quad (7)$$

(The terms ‘epi-continuous’ and ‘locally epi-Lipschitz near a point’ will be defined later.) Thus, conditions (3) and (4), which are distinct, can be viewed as different, coarser versions of same condition (5) (which is equivalent to (6)).

In related work [10], Ioffe has shown that when Rockafellar’s epi-continuity hypothesis (iii) is discarded, we still retain a one-sided version of the relationship, namely

$$\text{co} \{ q : (q, p(t)) \in \partial L(\tilde{x}(t), \tilde{\xi}(t)) \} \subseteq \text{co} \{ q : (-q, \tilde{\xi}(t)) \in \partial L(\tilde{x}(t), p(t)) \}. \quad (8)$$

The point here is that necessary conditions incorporating the partially convexified Euler condition (5) have been derived merely under hypotheses (i) and (ii) (see [10]), while the partially convexified Hamiltonian condition (6) has been derived directly, only under stronger hypotheses [13]. Relationship (8) therefore permits us to conclude, automatically, the validity of the partially dualized Hamiltonian condition under reduced hypotheses, via the partially dualized Euler condition.

We provide alternative proofs of the important relationships (7) and (8). Those of Rockafellar [19] and Ioffe [10] are based on mollification techniques and a sophisticated analysis of the relationship between subgradients of $L$ and smooth approximations.
of $L(x, \cdot)$. By contrast, the proofs in this paper employ a standard repertoire of techniques from nonsmooth analysis (exploitation of the Proximal Inequality, a Mean Value Inequality and a Mini–Max Theorem).

Furthermore, we establish that the two-sided relationship is valid under hypotheses weaker than those imposed in [19] (we relax Rockafellar’s epi-continuity hypothesis on $L$), and that the one-sided relationship is valid under hypotheses weaker than those imposed in [10] (we replace Ioffe’s local epi-Lipschitz hypothesis on $L$ by a weaker ‘dual’ hypothesis on a function related to $H$).

We can expect that, in the future, hypotheses will be further relaxed under which conditions akin to the Euler condition and Hamilton’s system of equations are known to be necessary conditions of optimality. If so, results relating the two kinds of conditions under unrestricted hypotheses, examples of which are reported here, will have a significant role in such developments.

Finally, we remark that certain aspects of the ‘finite dimensional’ results of this paper have analogues in Hilbert spaces. Such extensions will be explored in a future work.

Now some comments on our choice of notations. Throughout this article, we will restrict ourselves to $\mathbb{R}^n$, the Euclidean $n$-dimensional space, with inner product $\langle \cdot, \cdot \rangle$ and associated norm $| \cdot |$. $B_n := \{x \in \mathbb{R}^n : |x| < 1 \}$ denotes the open unit ball in $\mathbb{R}^n$ and $\overline{B}_n := \{x \in \mathbb{R}^n : |x| \leq 1 \}$ its closure.

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm \infty\}$ be a function. The sets $\text{dom} f := \{x \in \mathbb{R}^n : |f(x)| < +\infty\}$ and $\text{epi} f := \{(x,r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq r\}$ are called, respectively, the domain and the epigraph of $f$.

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a function, and $\bar{x}$ a point in $\text{dom} f$. We now define various concepts of subgradients [6,8,9,17]. The proximal subdifferential of $f$ at $\bar{x}$, denoted $\partial_{p} f(\bar{x})$, is the (possibly empty) set of all vectors $\xi$ in $\mathbb{R}^n$ for which there exists $\sigma > 0$ such that, for all $x$ close to $\bar{x}$,

$$\langle \xi, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) + \sigma|x - \bar{x}|^2.$$ 

The limiting subdifferential of $f$ at $\bar{x}$, denoted $\partial_{L} f(\bar{x})$, is the (possibly empty) set of all vectors $\xi$ in $\mathbb{R}^n$ such that

$$\xi = \lim_{i \to \infty} \xi_i \quad \xi_i \in \partial_{p} f(x_i) \quad \text{for all } i \quad x_i \xrightarrow{f} \bar{x},$$

and the asymptotic limiting subdifferential of $f$ at $\bar{x}$, denoted $\partial_{L}^\infty f(\bar{x})$, is the set of all vectors $\xi$ in $\mathbb{R}^n$ such that

$$\xi = \lim_{i \to \infty} \lambda_i \xi_i \quad \xi_i \in \partial_{p} f(x_i) \quad \text{for all } i \quad x_i \xrightarrow{f} \bar{x}, \quad \lambda_i \downarrow 0.$$ 

If, moreover, $f$ is a convex function on $\mathbb{R}^n$, we have

$$\partial_{p} f(\bar{x}) = \partial_{L} f(\bar{x}) := \{\xi \in \mathbb{R}^n : \langle \xi, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{R}^n\}.$$ 

If $U$ is an open set in $\mathbb{R}^k$ and $L : U \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm \infty\}$ is a function, then the partial dualization of $L$ with respect to the second variable is the function $H : U \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm \infty\}$, defined, for all $(x,v)$ in $U \times \mathbb{R}^n$, by

$$H(x, p) := \sup_{v \in \mathbb{R}^n} \{\langle p, v \rangle - L(x, v)\}. \quad (9)$$
This article is organized as follows: Section 2 is devoted to the epi-continuity and the epi-Lipschitz properties; our main results are stated in Section 3 and proved in Section 4.

2. Epi-continuity and epi-Lipschitz properties

We now introduce various continuity concepts for a function \( L : U \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \), where \( U \) is an open subset of \( \mathbb{R}^k \), in which we focus attention on the properties of the multi-function \( x \mapsto epi L(x, \cdot) \).

**Definition 2.1.** We say that \( L \) is **epi-continuous** if
(a) \( L \) is lower semi-continuous on \( U \times \mathbb{R}^n \);
(b) for all \((x, v)\) in \( U \times \mathbb{R}^n \) and for each sequence \( \{x_i\} \) converging to \( x \), there exists a sequence \( \{v_i\} \) converging to \( v \) such that \( \limsup_{i \to \infty} L(x_i, v_i) \leq L(x, v) \).

In the sequel, a slightly less restrictive notion of epi-continuity will be encountered in which property (b) of Definition 2.1 is required to hold merely in a neighborhood of some specified point. Namely:

**Definition 2.2.** If \((\tilde{x}, \tilde{v})\) is a point in \( dom L \), we say that \( L \) is **locally epi-continuous around** \((\tilde{x}, \tilde{v})\), if there exist neighborhoods \( O, V \) and \( W \) of \( \tilde{x}, \tilde{v} \) and \( L(\tilde{x}, \tilde{v}) \), respectively, such that for all \( x \) in \( O \), we have:
(a) \( L \) is lower semi-continuous on \( O \times \mathbb{R}^n \);
(b) for all \((x, v, L(x, v))\) in \( O \times V \times W \) and for each sequence \( \{x_i\} \) converging to \( x \), we can find a sequence \( \{v_i\} \) converging to \( v \) such that \( \limsup_{i \to \infty} L(x_i, v_i) \leq L(x, v) \).

**Remark 2.1.** Continuous functions on \( U \times \mathbb{R}^n \) are epi-continuous in both senses. Clearly, Definition 2.2 is weaker than Definition 2.1. It is known that \( L \) is epi-continuous if and only if \( H \) (defined as in (9)) is epi-continuous (using Wisjman’s theorem [20]). There is no such convenient characterization of local epi-continuity in terms of \( H \), because \( H \) depends on the global properties of \( L \).

We shall make use too of a Lipschitz-type notion of continuity, namely the local epi-Lipschitz property, again relating to the properties of the multi-function \( x \mapsto epi L(x, \cdot) \). This property has also been referred to as ‘Aubin continuity’ in the literature [10,19].

**Definition 2.3.** If \((\tilde{x}, \tilde{v})\) is a point in \( dom L \), we say that \( L \) is **locally epi-Lipschitz around** \((\tilde{x}, \tilde{v})\) if there exist neighborhoods \( O, V \) and \( W \) of \( \tilde{x}, \tilde{v} \) and \( L(\tilde{x}, \tilde{v}) \), respectively, and a positive constant \( k \) such that, for all \( x', x'' \) in \( O \),
\[
epi L(x', \cdot) \cap (V \times W) \subseteq epi L(x'', \cdot) + k|x' - x''|B_{n+1}.
\]

If \( L \) satisfies a Lipschitz condition near \((\tilde{x}, \tilde{v})\), then certainly \( L \) is locally epi-Lipschitz around that point, but the converse is not true in general. Furthermore, local epi-continuity and epi-Lipschitz properties near a point are related as follows:
Proposition 2.1. If $L$ is lower semi-continuous, finite at $(\tilde{x}, \tilde{v})$ and locally epi-Lipschitz near that point, then $L$ is locally epi-continuous around $(\tilde{x}, \tilde{v})$.

What implications does the local epi-Lipschitzness of $L$ near a point $(\tilde{x}, \tilde{v})$ of $\text{dom} L$ have regarding the regularity properties of $H$? There is no satisfactory answer to this question in general, because the local epi-Lipschitz property of $L$ near $(\tilde{x}, \tilde{v})$ places restrictions on the values of $L$ only in a neighborhood of $\tilde{x}$, while the conjugate functional $H$ depends on the global properties of $L$. However, we can profitably explore the Lipschitz continuity properties of the related function, defined, for all $(x, p)$ in $U \times \mathbb{R}^n$ and for all $\varepsilon > 0$, by

$$H_{\varepsilon}(x, p) := \sup_{v \in \mathbb{R}^n} \{ \langle p, v \rangle - L(x, v) \}.$$  \hspace{1cm} (10)

Proposition 2.2. Let $U$ be an open subset of $\mathbb{R}^k$ containing $\tilde{x}$ and $L: U \times \mathbb{R}^n \to \mathbb{R} \cup \{ +\infty \}$ be a function finite at $(\tilde{x}, \tilde{v})$.

Suppose that

(i) $L$ is lower semi-continuous on $U \times \mathbb{R}^n$;
(ii) for each $x$ in $U$, the function $v \mapsto L(x, v)$ is convex on $\mathbb{R}^n$;
(iii) $L$ is locally epi-Lipschitz around $(\tilde{x}, \tilde{v})$.

Then for each $c > 0$, there exist strictly positive constants $\varepsilon$ and $\beta$ such that $H_{c, \varepsilon}(\cdot, p)$ (defined as in (10)) is Lipschitz continuous on $\tilde{x} + \beta B_k$, uniformly with respect to $p$'s in $cB_n$. To be more precise, for each $c > 0$, there exist strictly positive constants $\varepsilon$, $\beta$, and $k'$ (not depending on $p$) such that

$$|H_{c, \varepsilon}(x, p) - H_{c, \varepsilon}(x', p)| \leq k'(1 + |p|) |x - x'| \leq k'(1 + c)|x - x'|$$

for all $x, x'$ in $\tilde{x} + \beta B_k$ and for all $p$ in $cB_n$.

Before proving this proposition, consider the following example [19]: the function $L$ defined, for all $(x, v)$ in $\mathbb{R} \times \mathbb{R}$, by

$$L(x, v) := \begin{cases} x^2(4x^{4/3})^{-1} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \text{ and } v = 0, \\ +\infty & \text{if } x = 0 \text{ and } v \neq 0 \end{cases}$$

is not locally epi-Lipschitz near $(0, 0)$; but we can easily calculate, for any $\varepsilon > 0$ and any $(x, p)$ in $\mathbb{R} \times \mathbb{R}$,

$$H_{c, \varepsilon}(x, p) := \begin{cases} x^{4/3} p^2 & \text{if } |2px^{4/3}| < \varepsilon, \\ |p|e - (e^2/4x^{4/3}) & \text{if } |2px^{4/3}| \geq \varepsilon \end{cases}$$

and notice that, for any $c > 0$ and even for any $\varepsilon > 0$, there exists a neighborhood of 0 (which depends on $c$ and $\varepsilon$), on which $H_{c, \varepsilon}(\cdot, p)$ is Lipschitz continuous, uniformly with respect to $p$'s in $cB_1$. This shows that the Lipschitz continuity property of the localized conjugate functional $H_{c, \varepsilon}$ is weaker than the local epi-Lipschitz property for $L$ around $(\tilde{x}, \tilde{v})$.

The assertion of Proposition 2.2 follows directly from the next result for a general multifunction $F: (\mathbb{R}^k, \cdot \mid \cdot) \to (\mathbb{R}^l, \cdot \mid \cdot)$, where $(\cdot \mid \cdot \rangle, \mathcal{A} \rangle)$ and $(\cdot \mid \cdot \rangle, \mathcal{A} \rangle)$ are some (not necessarily Euclidean) norms and associated open unit balls on $\mathbb{R}^k$ and $\mathbb{R}^l$, respectively.
Recall that a multifunction $F$ is called pseudo-Lipschitz with constant $k > 0$ at a point $(\bar{x}, \bar{z})$ such that $\bar{z} \in F(\bar{x})$ if there exist strictly positive constants $\varepsilon_0$ and $\beta_0$ such that, for all $x', x$ in $\bar{x} + \beta_0 \mathcal{B}$,

$$
(\bar{z} + \varepsilon_0 \mathcal{B}) \cap F(x') \subseteq F(x) + k|x' - x|_\varnothing \mathcal{B}.
$$

(11)

This concept was introduced by Aubin and it is easy to see that the local epi-Lipschitz condition for $L$ coincides with the pseudo-Lipschitz condition for the multifunction $x \mapsto F(x) := \text{epi} L(x, \cdot)$.

We show below that under the additional assumption that the values of $F$ are convex sets, the multifunction

$$
x \mapsto F_\varepsilon(x) := (\bar{z} + \varepsilon \mathcal{B}) \cap F(x)
$$

(12)

is Lipschitz continuous on the ball $\bar{x} + \beta \mathcal{B}$ for sufficiently small strictly positive numbers $\varepsilon$ and $\beta$. More precisely, we have the following lemma (which is valid in the more general setting of Banach spaces):

**Lemma 2.3.** Let $F : (\mathbb{R}^k, |\cdot|_\varnothing) \to (\mathbb{R}^l, |\cdot|_\varnothing)$ be a convex-valued multifunction satisfying (11). Then for any $\varepsilon$ in $(0, \varepsilon_0]$ and for any $\beta$ in $(0, \varepsilon/4k)$, the multifunction $F_\varepsilon$ (defined as in (12)) is Lipschitz continuous with constant $5k$ on the ball $\bar{x} + \beta \mathcal{B}$.

**Proof.** Taking $x' = \bar{x}$ in (11), we obtain that, for all $x$ in $\bar{x} + \beta \mathcal{B}$,

$$
\bar{z} \in F(x) + k|x - \bar{x}|_\varnothing \mathcal{B},
$$

this implies that there exists a point $z_\varepsilon$ in $F(x)$ such that

$$
|z_\varepsilon - \bar{z}|_\varnothing \leq k|x - \bar{x}|_\varnothing < \varepsilon/2,
$$

(13)

which proves that $F_\varepsilon$ is nonempty valued on $\bar{x} + \beta \mathcal{B}$.

To prove that $F_\varepsilon$ is Lipschitz on the same ball, we consider arbitrary points $x', x$ in it and $z'$ in $F_\varepsilon(x')$. It follows from (11) that there exists $z$ in $F(x)$ such that

$$
|z - z'|_\varnothing \leq k|x - x'|_\varnothing.
$$

Let us assume that $z \notin \bar{z} + \varepsilon \mathcal{B}$ (otherwise, the proposition is easily deduced). This means that

$$
r := |z - \bar{z}|_\varnothing > \varepsilon.
$$

Recall that there exists a point $z_\varepsilon$ in $F(x)$ satisfying (13). Due to the convexity of $F(x)$, the point $z''$ of the segment $[z_\varepsilon; z]$ taken such that $|z'' - \bar{z}|_\varnothing = \varepsilon$, belongs to $F(x)$ and, consequently, to $F_\varepsilon(x)$. We show that

$$
|z'' - z'|_\varnothing \leq 5k|x - x'|_\varnothing,
$$

(14)

which implies that

$$
F_\varepsilon(x') \subseteq F_\varepsilon(x) + 5k|x - x'|_\mathcal{B}
$$

and $F_\varepsilon$ is Lipschitz continuous, as required.
To prove (14), we estimate the quantity \( \rho := |z - z''| \) in terms of the distance \( d:= r - \varepsilon \) from \( z \) to the ball \( \hat{z} + \varepsilon \mathcal{B}_\star \). By the definition of \( z'' \), we have that
\[
z'' = (1 - t)z + tz_x,
\]
for some \( t \in (0, 1) \) and
\[
\varepsilon = |z'' - \hat{z}| \leq |(1 - t)(z - \hat{z}) + t(z_x - \hat{z})| \leq (1 - t)r + t|z_x - \hat{z}|.
\]
We obtain from the last inequality that
\[
t(r - |z_x - \hat{z}|) \leq r - \varepsilon.
\]
This implies, by (13), that
\[
t \leq \frac{r - \varepsilon}{r - \varepsilon/2}.
\]
Note that because of this relation,
\[
\rho := |z'' - z| \leq (r - \varepsilon) \frac{r + \varepsilon/2}{r - \varepsilon/2}.
\]
(Here, we use the fact that \( |z_x - z| \leq |z - \hat{z}| + |z_x - \hat{z}| \leq r + \varepsilon/2 \).) Note also that
\[
r := |z - \hat{z}| \leq |z - z'| + |z' - \hat{z}| \leq k|x - x'| + \varepsilon < \frac{\varepsilon}{2}
\]
by choice of the constant \( \beta \). Hence, we have from the inequalities above (and the inequality \( r > \varepsilon \)) that
\[
|z'' - z| = \rho \leq 4(r - \varepsilon) = 4d.
\]
But \( d \leq |z - z'| \) since \( |z' - \hat{z}| \leq \varepsilon \). Thus, we obtain that
\[
|z'' - z'| \leq |z'' - z| + |z - z'| \leq 5k|x - x'|,
\]
which means that (14) holds. □

**Proof of Proposition 2.2.** Let us fix some \( c > 0 \) and consider an arbitrary vector \( p \) such that \( |p| < c \). From the local epi-Lipschitzness of \( L \), without loss of generality, there exist strictly positive constants \( \beta, \varepsilon, k \) and \( x \) such that for all \( x, x' \) in \( \bar{x} + \beta \mathcal{B}_n \),
\[
\text{epi } L(x, \cdot) \cap ((\bar{v} + \varepsilon \mathcal{B}_n) \times (L(\bar{x}, \bar{v}) + z \mathcal{B}_n)) \subseteq \text{epi } L(x', \cdot) + k|x - x'| \mathcal{B}_{n+1}.
\]
We can assume \( \beta \) and \( \varepsilon \) small enough such that \( (2c+1)\varepsilon < \alpha, \varepsilon < \alpha, k\beta < \varepsilon, \beta < \varepsilon/4k \) and such that, for all \( (x, v) \) in \( (\bar{x} + \beta \mathcal{B}_n) \times (\bar{v} + \varepsilon \mathcal{B}_n) \), we have
\[
L(x, v) > L(\bar{x}, \bar{v}) - \alpha,
\]
using the lower semi-continuity of \( L \) at the point \( (\bar{x}, \bar{v}) \).

Recall our notation \( |\cdot| \) for the Euclidean norm on any Euclidean space whatever be the dimension. Next, consider for an element \( z = (v, a) \in \mathbb{R}^n \times \mathbb{R}^1 \) the following norm:
\[
|z| = \max \left\{ |v|; |a| \right\}.
\]
Let us denote the open unit ball for this norm as \( \mathcal{B}_\star \). Since all norms on \( \mathbb{R}^n \times \mathbb{R}^1 \) are equivalent, there exists \( M > 0 \) such that for all \( z \) in \( \mathbb{R}^n \times \mathbb{R}^1 \approx \mathbb{R}^{n+1} \),
\[
|z| \leq M|z| \star.
\]
Now, because of (15), the multifunction $F : (\mathbb{R}^k, |\cdot|) \to (\mathbb{R}^n \times \mathbb{R}^1, |\cdot|)$, defined by $x \mapsto F(x) := \text{epi} L(x, \cdot)$, is pseudo-Lipschitz with constant $k$ at the point $(\bar{x}, \bar{z})$ where $\bar{z} = (\bar{L}, \bar{ar{L}})$; indeed, recalling that $\varepsilon < \alpha$, we have, for all $x, x'$ in $\bar{x} + \beta B_k$,

$$(\bar{z} + \varepsilon \mathcal{B}_*) \cap F(x') \subseteq F(x) + k|x - x'| \mathcal{B}_*$$

that is exactly (11) with $\varepsilon_0 := \varepsilon$ and $\beta_0 := \beta$.

Since $\beta < \varepsilon/4k$, it follows from Lemma 2.3 that the multifunction from $(\mathbb{R}^{n+1}, |\cdot|)$ into the subsets of $(\mathbb{R}^{n+1}, |\cdot|)$:

$$x \mapsto F_x(x) := (\bar{z} + \varepsilon \mathcal{B}_*) \cap \text{epi} L(x, \cdot)$$

is Lipschitz continuous with constant $5k$ on $\bar{x} + \beta B_k$.

Now, given $x$ in $\bar{x} + \beta B_k$, we consider a maximizing vector $\mathbf{v}$ in $\bar{v} + \epsilon \mathcal{B}_a$ such that

$$\langle p, \mathbf{v} \rangle - L(x, \mathbf{v}) = H_{\varepsilon, \bar{v}}(x, p)$$

and we show that such a vector $\mathbf{v}$ satisfies

$$L(x, \mathbf{v}) \in L(\bar{x}, \bar{v}) + x \mathcal{B}_1.$$  \hfill (17)

From (15), since $k\beta < \varepsilon$, there is a vector $\mathbf{v}'$ in $\bar{v} + \epsilon \mathcal{B}_a$ such that

$$L(\bar{x}, \bar{v}) \geq L(x, \mathbf{v}') - k|x - \bar{x}| \geq L(x, \mathbf{v}') - k\beta \geq L(x, \mathbf{v}') - \varepsilon.$$  This implies that

$$H_{\varepsilon, \bar{v}}(x, p) \geq \langle p, \mathbf{v}' \rangle - L(x, \mathbf{v}') \geq \langle p, \bar{v} \rangle - L(\bar{x}, \bar{v}) - (c + 1)\varepsilon.$$

If $\mathbf{v}$ belonging to $\bar{v} + \epsilon \mathcal{B}_a$ satisfies (16), then we obtain from the previous inequality that

$$L(x, \mathbf{v}) \leq L(\bar{x}, \bar{v}) + (2c + 1)\varepsilon \leq L(\bar{x}, \bar{v}) + \alpha,$$

which implies, together with the lower bound for $L(x, \mathbf{v})$, that (17) is valid.

We have now all the ingredients to easily verify the statements of Proposition 2.2, with $k' := 5M$ and $\varepsilon, \beta$ carefully chosen as before. Let us take arbitrary points $x', x$ in the ball $\bar{x} + \beta B_k$ and a vector $\mathbf{v} \in \bar{v} + \epsilon \mathcal{B}_a$ such that (16) holds. Then (17) is valid and $(\mathbf{v}, L(x, \mathbf{v})) \in F_x(x)$ due to the definition of $\mathcal{B}_*$ and the choice of $\mathbf{v}$. But $F_x$ is Lipschitz continuous, so we have that, for $(\mathbf{v}, L(x, \mathbf{v}))$, there exists $\mathbf{v}' \in \bar{v} + \epsilon \mathcal{B}_a$ and $\mathbf{a} \geq L(x', \mathbf{v}')$ such that

$$|\langle \mathbf{v}, L(x, \mathbf{v}) \rangle - \langle \mathbf{v}', \mathbf{a} \rangle| \leq M|\langle \mathbf{v}, L(x, \mathbf{v}) \rangle - \langle \mathbf{v}', \mathbf{a} \rangle| \leq 5kM|x - x'|.$$

This implies that

$$H_{\varepsilon, \bar{v}}(x, p) = \langle p, \mathbf{v} \rangle - L(x, \mathbf{v}) \leq \langle p, \mathbf{v}' \rangle - L(x', \mathbf{v}')$$

$$|p||\mathbf{v} - \mathbf{v}'| + \langle \mathbf{a}, \mathbf{v}' \rangle - L(x, \mathbf{v}) \leq H_{\varepsilon, \bar{v}}(x', p) + 5kM(\|p\| + 1)|x - x'|.$$

Since the roles of $x$ and $x'$ are interchangeable, and this inequality is valid for all $p$ such that $|p| < c$, the proposition is proved. □
3. Main theorems

Recall that if $U$ is an open set in $\mathbb{R}^n$, and $L: U \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a function, the conjugate functional $H: U \times \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ is defined, for all $(x, p)$ in $U \times \mathbb{R}^n$, by

$$H(x, p) := \sup_{v \in \mathbb{R}^n} \{ (p, v) - L(x, v) \} \quad (18)$$

and the localized conjugate functional (around $\bar{v}$ in $\mathbb{R}^n$) $H_{\bar{v}}: U \times \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ is defined, for all $(x, p)$ in $U \times \mathbb{R}^n$ and all $\varepsilon > 0$, by

$$H_{\bar{v}}(x, p) := \sup_{v \in \mathbb{R}^n : \| v - \bar{v} \| \leq \varepsilon} \{ (p, v) - L(x, v) \}. \quad (19)$$

The first dualization theorem covers situations where the conjugate functional $H(\cdot, \cdot)$ is Lipschitz continuous with respect to the first variable, uniformly with respect to the second variable.

**Theorem 3.1.** Let $U$ and $P$ be open subsets of $\mathbb{R}^n$ containing, respectively, $\bar{x}$ and $\bar{p}$. Suppose that

- (H1) for each $x$ in $U$, the function $v \mapsto L(x, v)$ is convex on $\mathbb{R}^n$;
- (H2) $L$ is locally epi-continuous around $(\bar{x}, \bar{v})$;
- (H3) $H(\cdot, \cdot)$ is Lipschitz continuous on $U$, uniformly with respect to $p$’s in $P$.

Then if $(\bar{q}, \bar{p})$ belongs to $\partial L(\bar{x}, \bar{v})$, we have

$$\bar{q} \in \text{co}\{ q \in \mathbb{R}^n : (q, \bar{v}) \in \partial L(\bar{x}, \bar{v}) \}.$$

In particular, we deduce

$$\text{co}\{ q \in \mathbb{R}^n : (q, \bar{p}) \in \partial L(\bar{x}, \bar{v}) \} \subseteq \text{co}\{ q \in \mathbb{R}^n : (q, \bar{v}) \in \partial L(\bar{x}, \bar{p}) \}.$$

From the point of view of “dualizing” the extended Euler inclusion, hypothesis (H3) of the preceding theorem is somewhat restrictive. A variant, better suited to this purpose, imposes instead a Lipschitz continuity hypothesis on $H_{\bar{v}}$, for some $\varepsilon > 0$.

**Theorem 3.2.** Let $U$ and $P$ be open subsets of $\mathbb{R}^n$ containing, respectively, $\bar{x}$ and $\bar{p}$. Suppose that

- (H1) for each $x$ in $U$, the function $v \mapsto L(x, v)$ is convex on $\mathbb{R}^n$;
- (H2) $L$ is locally epi-continuous around $(\bar{x}, \bar{v})$;
- (H3) there exists $\varepsilon > 0$ such that $H_{\bar{v}}(\cdot, \cdot)$ is Lipschitz continuous on $U$, uniformly with respect to $p$’s in $P$.

Then if $(\bar{q}, \bar{p})$ belongs to $\partial L(\bar{x}, \bar{v})$, we have

$$\bar{q} \in \text{co}\{ q \in \mathbb{R}^n : (q, \bar{v}) \in \partial L(\bar{x}, \bar{v}) \}.$$

In particular, we deduce

$$\text{co}\{ q \in \mathbb{R}^n : (q, \bar{p}) \in \partial L(\bar{x}, \bar{v}) \} \subseteq \text{co}\{ q \in \mathbb{R}^n : (q, \bar{v}) \in \partial L(\bar{x}, \bar{p}) \}.$$
Remark 3.1. The hypotheses of Theorem 3.2 imply that $H$ is lower semi-continuous and finite at $(\tilde{x}, \tilde{p})$, and that it nowhere takes the value $-\infty$ on $U \times P$, but they do not imply that $H$ is either lower semi-continuous on a neighborhood of $(\tilde{x}, \tilde{p})$ or that it is proper on $U \times \mathbb{R}^n$.

Corollary 3.3. (Ioffe [10]). The assertions of Theorem 3.2 remain true when (H2) and (H3) are replaced by

(i) $L$ is lower semi-continuous on $U \times \mathbb{R}^n$;
(ii) $L$ is locally epi-Lipschitz around $(\tilde{x}, \tilde{v})$.

For, it suffices to invoke Propositions 2.1 and 2.2. Let us note however that the hypotheses of Theorem 3.2 are weaker than those of Corollary 3.3. Indeed, just consider $\tilde{x} = \tilde{v} = \tilde{p} = 0$ and the function $L$ defined by $L(x, v) := (1 + |x|^{1/2})|v|$ for all $(x, v)$ in $\mathbb{R} \times \mathbb{R}$; then we have, for all $(x, p)$ in $\mathbb{R} \times \mathbb{R}$ and for all $\varepsilon > 0$,

$$H(x, p) = \begin{cases} 0 & \text{if } |p| \leq 1 + |x|^{1/2}, \\ |p|e - (1 + |x|^{1/2})e & \text{if } |p| > 1 + |x|^{1/2}. \end{cases}$$

Obviously, taking $P := (1/2)B_1$, we find that $H(x, p) = 0$ for all $(x, p)$ in $\mathbb{R} \times P$, hence the continuous function $L$ satisfies all the hypotheses of Theorem 3.2, but is certainly not locally epi-Lipschitz around $(0, 0)$.

Finally, we shall state a complete dualization theorem, under which we have the equivalence of the extended Euler inclusion and the Hamiltonian inclusion, as discussed in the introduction:

Theorem 3.4. Let $U$ be an open subset of $\mathbb{R}^n$ containing $\tilde{x}$. Let $L : U \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a function finite at $(\tilde{x}, \tilde{v})$.

Suppose that

(H1) for each $x$ in $U$, the function $v \mapsto L(x, v)$ is convex on $\mathbb{R}^n$;

(H2) $L$ is lower semi-continuous on $U \times \mathbb{R}^n$;

(H3) $L$ is locally epi-Lipschitz around $(\tilde{x}, \tilde{v})$;

(H4) $H$ is locally epi-continuous around $(\tilde{x}, \tilde{p})$.

Then if $(\tilde{q}, \tilde{v})$ belongs to $\partial L H(\tilde{x}, \tilde{p})$, we have

$$\tilde{q} \in \co\{q \in \mathbb{R}^n : (-q, \tilde{p}) \in \partial L H(\tilde{x}, \tilde{v})\}.$$ 

In particular, we have the following equality:

$$\co\{q \in \mathbb{R}^n : (q, \tilde{p}) \in \partial L H(\tilde{x}, \tilde{v})\} = \co\{q \in \mathbb{R}^n : (-q, \tilde{v}) \in \partial L H(\tilde{x}, \tilde{p})\}. \quad (20)$$

Corollary 3.5. (Rockafellar [19]). Assertion (20) of Theorem 3.4 remains true when (H2) and (H4) are replaced by the epi-continuity of $L$ on $U \times \mathbb{R}^n$.

This is indeed a simple consequence of Theorem 3.4 and Remark 2.1.

Note also that we can substitute, in Theorem 3.4, the following (not necessarily equivalent) condition to (H3):

(H3)$'$ $H$ is locally epi-Lipschitz near $(\tilde{x}, \tilde{p})$. 


Indeed, we can interchange the role of \( L \) and \( H \), since \( L \) is convex in the second variable.

**Remark 3.2.** Let us also notice that, under (H2), hypothesis (H3) of Theorem 3.4 is equivalent to the following nondegeneracy condition on the asymptotic limiting subgradient (see the section on Mordukhovich’s criterion in [20], for example):

if \( (q, 0) \) belongs to \( \hat{\partial}^\infty \! L(\bar{x}, \bar{v}) \), then \( q = 0 \);

which is in turn equivalent to:

there exists \( K > 0 \) such that for all \( (x, v, L(x, v)) \) close enough to \( (\bar{x}, \bar{v}, L(\bar{x}, \bar{v})) \), the proximal normal cone to \( epi L \) at the point \( (x, v, L(x, v)) \) is included in \( \{ (\alpha, \beta, \gamma): |\alpha| \leq K(\beta, \gamma) \} \).

**Remark 3.3.** The first assertion of Theorem 3.1, related to \( \bar{q} \), remains valid with (H2) replaced by the weaker following conditions:

(a) \( L \) is lower semi-continuous on \( U \times \mathbb{R}^n \);

(b) there exist neighborhoods \( V \) of \( \bar{x} \), \( W \) of \( L(\bar{x}, \bar{v}) \) and \( Y \) of \( (\bar{q}, \bar{p}) \) such that for all \( (x, v, L(x, v)) \) in \( U \times V \times W \) for which \( \partial_v L(x, v) \cap Y \neq \emptyset \) and for all sequences \( \{x_i\} \) converging to \( x \), we can find a sequence \( \{v_i\} \) converging to \( v \) such that \( \limsup_{i \to \infty} L(x_i, v_i) \leq L(x, v) \).

It is a kind of local epi-continuity near \( (\bar{x}, \bar{v}) \), but with respect to \( (\bar{q}, \bar{p}) \). This remark is also relevant for the other theorems of this section.

4. Proofs of the main results

A key role in the proofs is played by the following version of the Mini–Max Theorem, in which compactness of the level sets replaces the usual compactness hypothesis.

**Proposition 4.1.** Let \( h: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \) be a convex, lower semi-continuous and proper function. Let \( \sigma \) be a positive constant and \( (\bar{x}, \bar{y}) \) be a vector in \( \mathbb{R}^n \times \mathbb{R}^n \). Let \( F: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \) be the following function:

\[
F(x, y) := \langle x, y - \bar{y} \rangle + \sigma |x - \bar{x}|^2 - h(y) \quad \text{for all } (x, y) \in \mathbb{R}^n \times \mathbb{R}^n.
\]

Then there exists a point \((x^*, y^*) \) in \( \mathbb{R}^n \times dom h \) such that, for all \((x, y) \) in \( \mathbb{R}^n \times \mathbb{R}^n \),

\[
F(x^*, y) \leq F(x, y^*) \leq F(x, y^*).
\]

Furthermore, we have

\[
x^* = \bar{x} - \frac{1}{2\sigma}(y^* - \bar{y}).
\]

**Proof of Theorem 3.1.** In view of (H3), we can assume, by reducing the size of \( U \) and \( P \) if required, that \( H \) is finite-valued and continuous on \( U \times P \), and that the multifunction \( (x, p, v) \to \{ q \in \mathbb{R}^n: (q, v) \in \partial_v H(x, p) \} \), defined on \( U \times P \times \mathbb{R}^n \), is uniformly bounded (i.e. there exists a ball in \( \mathbb{R}^n \) containing all the values of the multifunction) and
has closed graph (recall the closure properties of the subgradient multifunction \( \partial_l H \)).

By Carathéodory’s Theorem, we easily deduce that the multifunction \( (x, p, v) \mapsto \text{co}\{q \in \mathbb{R}^n : (q, v) \in \partial_l H(x, p)\} \), defined on \( U \times P \times \mathbb{R}^n \), is also uniformly bounded and has closed graph. By (H2), \( L \) is lower semi-continuous on \( U \times \mathbb{R}^n \), and there exist a bounded neighborhood \( V \) of \( \bar{v} \) and a neighborhood \( W \) of \( L(\bar{x}, \bar{v}) \), such that, following a further reduction in the size of \( U \), given any \( (x, v, L(x, v)) \) in \( U \times V \times W \) and any sequence \( \{x_i\} \) converging to \( x \), there exists a sequence \( \{v_i\} \) converging to \( v \) such that \( \lim_{i \to \infty} L(x_i, v_i) \leq L(x, v) \).

These facts will be used in the sequel.

Take a point \( \bar{q} \) in \( \mathbb{R}^n \) such that \( (\bar{q}, \bar{p}) \in \partial_l L(\bar{x}, \bar{v}) \). We must show that

\[-\bar{q} \in \text{co}\{q \in \mathbb{R}^n : (q, v) \in \partial_l L(\bar{x}, \bar{p})\}.\]

We note at the outset that it suffices to treat only the case when

\[(\bar{q}, \bar{p}) \in \partial_p L(\bar{x}, \bar{v}). \tag{21}\]

This is because if merely \((\bar{q}, \bar{p}) \in \partial_l L(\bar{x}, \bar{v})\), there exist sequences \( \{(x_i, v_i)\} \) converging with \( L \), to \((\bar{x}, \bar{v})\) and \( \{(q_i, p_i)\} \) converging to \((\bar{q}, \bar{p})\) such that \((q_i, p_i) \in \partial_p L(x_i, v_i)\) for all \( i \). Applying the special case of the theorem gives, for \( i \) sufficiently large,

\[-q_i \in \text{co}\{q \in \mathbb{R}^n : (q, v_i) \in \partial_l L(x_i, p_i)\}.\]

Taking the limit as \( i \to \infty \), we then deduce, using the properties of the multifunction \((x, p, v) \mapsto \text{co}\{q \in \mathbb{R}^n : (q, v) \in \partial_l H(x, p)\} \), that \(-\bar{q} \in \text{co}\{q \in \mathbb{R}^n : (q, v) \in \partial_l H(x, p)\} \).

This confirms that, without loss of generality, we can assume (21).

By modifying the neighborhoods \( \bar{U} \) and \( \bar{V} \) if required, we can arrange that for some constant \( \sigma > 0 \), \( M(x, v) \geq 0 \) for all \((x, v)\) in \( \bar{U} \times \bar{V} \), where

\[M(x, v)=L(x, v) - L(\bar{x}, \bar{v}) - \langle \bar{q}, x - \bar{x} \rangle - \langle \bar{p}, v - \bar{v} \rangle - \sigma |x - \bar{x}|^2 + \sigma |v - \bar{v}|^2. \tag{22}\]

Since for each \( x \) in \( \bar{U} \), the function \( M(x, \cdot) \) is lower semi-continuous and strictly convex, there exists a unique minimizer \( v_x \) over the compact set \( \bar{P} \). We can deduce from (H2) that

\[\lim_{x \to \bar{x}} M(x, v_x) \leq 0.\]

By the lower semi-continuity of \( M(\cdot, \cdot) \) on \( U \times \mathbb{R}^n \) and since \( \bar{v} \) is the unique minimizer of \( M(\bar{x}, \cdot) \) over \( \bar{P} \), we have

\[\lim_{x \to \bar{x}} v_x = \bar{v}. \tag{23}\]

For all \( x \) sufficiently close to \( \bar{x} \), then \( v_x \) is interior to \( \bar{V} \) and so, since \( M(x, \cdot) \) is convex, \( v_x \) is the unique global minimizer. For all such \( x \), we have

\[0 \leq M(x, v_x) = \min_{v \in \mathbb{R}^n} M(x, v). \tag{24}\]

From (22), \( \bar{v} \) minimizes \( v \mapsto L(\bar{x}, v) + \sigma |v - \bar{v}|^2 - \langle \bar{p}, v - \bar{v} \rangle \) over \( \mathbb{R}^n \); hence \( \bar{p} \) belongs to \( \partial_p L(\bar{x}, \bar{v})(\bar{v}) \), from which we conclude that

\[L(\bar{x}, \bar{v}) = \langle \bar{p}, \bar{v} \rangle - H(\bar{x}, \bar{p}). \tag{25}\]
Representing $L(x, \cdot)$ as the conjugate function of $H(x, \cdot)$, we obtain

$$
\min_{v \in \mathbb{R}^n} M(x, v) = \min_{v \in \mathbb{R}^n} \left\{ \sup_{p \in \mathbb{R}^n} \left[ \langle p, v \rangle - H(x, p) \right] - L(\bar{x}, \bar{v}) - \langle \bar{q}, x - \bar{x} \rangle - \langle \bar{p}, v - \bar{v} \rangle + \sigma |x - \bar{x}|^2 + \sigma |v - \bar{v}|^2 \right\}
$$

$$= \min_{v \in \mathbb{R}^n} \sup_{p \in \mathbb{R}^n} K_s(v, p) + H(\bar{x}, \bar{p}) - \langle \bar{q}, x - \bar{x} \rangle + \sigma |x - \bar{x}|^2
$$

by (25). Here,

$$K_s(v, p) := \langle p - \bar{p}, v \rangle + \sigma |v - \bar{v}|^2 - H(x, p).
$$

It follows from Proposition 4.1, that

$$\min_{v \in \mathbb{R}^n} \sup_{p \in \mathbb{R}^n} K_s(v, p) = \min_{v \in \mathbb{R}^n} \sup_{p \in \mathbb{R}^n} K_s(v, p) = \langle p - \bar{p}, \bar{v} \rangle - (4\sigma)^{-1} |p - \bar{p}|^2 - H(x, p).
$$

in which

$$p_s := \bar{p} - 2\sigma (v_s - \bar{v}). \quad (26)
$$

We also have that, for any $p$ in $\mathbb{R}^n$,

$$\min_{v \in \mathbb{R}^n} \sup_{p \in \mathbb{R}^n} K_s(v, p) \geq \min_{v \in \mathbb{R}^n} \sup_{p \in \mathbb{R}^n} K_s(v, p) = \langle p - \bar{p}, \bar{v} \rangle - (4\sigma)^{-1} |p - \bar{p}|^2 - H(x, p).
$$

From (24), we conclude that

$$0 \leq M(x, v_s) = -\langle \bar{q}, x - \bar{x} \rangle + \sigma |x - \bar{x}|^2 + \langle p_s - \bar{p}, \bar{v} \rangle - (4\sigma)^{-1} |p_s - \bar{p}|^2 - H(x, p_s) + H(\bar{x}, \bar{p}).
$$

and that, for all $p$ in $\mathbb{R}^n$,

$$0 = M(\bar{x}, \bar{v}) \geq \langle p - \bar{p}, \bar{v} \rangle - (4\sigma)^{-1} |p - \bar{p}|^2 - H(\bar{x}, p) + H(\bar{x}, \bar{p}).
$$

It follows from these last two inequalities that for all $x$ near $\bar{x}$ and for all $p$ in $\mathbb{R}^n$,

$$H(\bar{x}, p) - H(x, p_s) + (4\sigma)^{-1} \left\{ |p - \bar{p}|^2 - |p_s - \bar{p}|^2 \right\} + \langle p_s - p, \bar{v} \rangle - \langle \bar{q}, x - \bar{x} \rangle + \sigma |x - \bar{x}|^2 \geq 0. \quad (27)
$$

For fixed $x$ close to $\bar{x}$, define the function $\phi_s(\cdot)$ whose argument $z$ is partitioned as $z = (y, p)$, to be

$$\phi_s(z) := H(y, p) - H(x, p_s) + (4\sigma)^{-1} \left\{ |p - \bar{p}|^2 - |p_s - \bar{p}|^2 \right\} + \langle p_s - p, \bar{v} \rangle - \langle \bar{q}, x - y \rangle + \sigma |x - y|^2.
$$

In consequence of the definition of $\phi_s$ and also by (27), we have that

$$\phi_s(\bar{x}, \bar{p}) \geq 0 \quad \text{for all } p \text{ in } \mathbb{R}^n,
$$

$$\phi_s(x, p_s) = 0. \quad (28)
$$

By (23) and (26), we know that $\lim_{x \to \bar{x}} p_s = \bar{p}$. Therefore, we can choose sequences $\{e_i\}$ and $\{\delta_i\}$ decreasing to 0 such that $\{|e_i|/\delta_i^2\}$ also decreases to 0, and for all $i$,

$$|x - \bar{x}| \leq e_i \quad \text{implies } |p_s - \bar{p}| \leq \delta_i/2.$$
Fix $i$ in $\mathbb{N}$. For each $x$ in $\tilde{x} + e_iB_n$ such that $x \neq \tilde{x}$, apply the generalized Mean Value Theorem [7] to $\phi_z(\cdot, \cdot)$, around the point $(x, p_x)$ and with respect to the nonempty convex and compact subset $Z := \{ (\tilde{x}, p) \in \tilde{p} + \delta_iB_n \}$ of $\mathbb{R}^{2n}$. Since, by relations (28),
\[
\tilde{r} := \min_{z \in Z} \phi_z(z) - \phi_z(x, p_x) = \min_{p \in \tilde{p} + \delta_iB_n} \phi_z(x, p) \geq 0
\]
then for the real number $\tilde{r} := -e_i|\tilde{x} - \tilde{x}| < 0$ which is strictly less than $\tilde{r}$, and for $e_i > 0$, we can find an element $(\tilde{\zeta}'(x), \tilde{\zeta}'(x))$ in $\partial p H(y'(x), p'(x))$ for some $(y'(x), p'(x))$ in $\mathbb{R}^n \times \mathbb{R}^n$ such that
\[
|y'(x) - x| \leq 2e_i, \quad |p'(x) - \tilde{p}| \leq e_i + \delta_i
\]
and such that for all $p$ in $\tilde{p} + \delta_iB_n$,
\[
-e_i|\tilde{x} - \tilde{x}| < \langle \zeta'(x) + \tilde{q} + 2\sigma(y'(x) - x), \tilde{x} - x \rangle
\]
\[
+ \langle \tilde{\zeta}'(x) - \tilde{v} + (2\sigma)^{-1}(p'(x) - \tilde{p}), p - p_x \rangle.
\]
Since $|p_x - \tilde{p}| \leq \delta_i/2$, taking the minimum over $p$ in the final term on the right, we arrive at
\[
-e_i^2 \leq -e_i|\tilde{x} - \tilde{x}| \leq \langle \zeta'(x) + \tilde{q} + 2\sigma(y'(x) - x), \tilde{x} - x \rangle
\]
\[
- \delta_i |\tilde{\zeta}'(x) - \tilde{v} + (2\sigma)^{-1}(p'(x) - \tilde{p})|.
\]
This last inequality is valid for all $x$ in $\tilde{x} + e_iB_n$ such that $x \neq \tilde{x}$.

Now $|\zeta'(x)|$ is bounded on $\tilde{x} + e_iB_n$ by a constant independent of $i$ and our choice of $y'(x), p'(x)$ (see the remarks at the beginning of the proof). It follows that there exists a positive constant $N$ (independent of $i$) such that, for all $x$ in $\tilde{x} + e_iB_n$ such that $x \neq \tilde{x}$,
\[
-e_i^2 \leq (N + |\tilde{q}| + 4\sigma e_i) e_i - (\delta_i/2)|\tilde{\zeta}'(x) - \tilde{v}| + (\delta_i/2)(2\sigma)^{-1}(e_i + \delta_i).
\]
Knowing that $\{e_i\}$ and $\{\delta_i\}$ are two sequences decreasing to 0 such that $\{e_i, \delta_i^{-1}\}$ also converges to 0, we can find a sequence $\{\gamma_i\}$ decreasing to 0 such that for all $i$,
\[
\sup_{x \in \tilde{x} + e_iB_n} |\zeta'(x) - \tilde{v}| < \gamma_i.
\]
Inequality (29) also tells us that
\[
-e_i|\tilde{x} - \tilde{x}| \leq \langle \zeta'(x) + \tilde{q} + 2\sigma(y'(x) - x), \tilde{x} - x \rangle
\]
\[
\leq \langle \zeta'(x) + \tilde{q}, \tilde{x} - x \rangle + 4\sigma e_i|x - \tilde{x}|
\]
for all $x$ in $\tilde{x} + e_iB_n$. This means that
\[
0 \leq \langle \zeta'(x) + \tilde{q}, \tilde{x} - x \rangle + \max_{e \in (\gamma_i/2)(1 + 4\sigma \delta_iB_n)} \langle e, \tilde{x} - x \rangle
\]
for all $x$ in $\tilde{x} + e_iB_n$. Inequalities (30) and (31) yield, for all $i$ and for all $x$ in $\tilde{x} + e_iB_n$ (including $x = \tilde{x}$),
\[
\sup_{\zeta \in \hat{Q} + \delta_iB_n} \langle \zeta, \tilde{x} - x \rangle \geq 0,
\]
where we define
\[ S_i := \{ \xi \in \mathbb{R}^n : (\xi, \bar{v}) \in \partial_p H(\bar{x}', p') + ((e_i/2)(1 + 4\sigma)B_n) \times (\gamma_2B_n), \]
\[ |\bar{x}' - \bar{x}| \leq 3e_i, |p' - \bar{p}| \leq e_i + \delta_i \}. \]

We conclude from (32) that, for all \( i \),
\[ -q \in \overline{\partial S_i}. \]

But in view of the remarks at the beginning of the proof, the \( S_i \)'s are uniformly bounded.

Since \( H \) is continuous on a neighborhood of \((\bar{x}, \bar{p})\), we obtain the result from Carathéodory’s theorem and the closure properties of \( \partial L \).

**Proof of Theorem 3.2.** Take a point \( \bar{q} \) such that \((\bar{q}, \bar{p}) \in \partial L(\bar{x}, \bar{v})\). We must show that
\[ -\bar{q} \in \text{co}\{q \in \mathbb{R}^n : (q, \bar{v}) \in \partial L(\bar{x}, \bar{p})\}. \]

For \( \epsilon > 0 \) given by (H3), define, for all \((x, v)\) in \( \mathbb{R}^n \times \mathbb{R}^n \),
\[ \mathcal{L}_{\epsilon, \bar{v}}(x, v) := \begin{cases} L(x, v) & \text{if } v \in \bar{v} + \epsilon B_n, \\ +\infty & \text{otherwise} \end{cases}. \]

Since \( L \) and \( \mathcal{L}_{\epsilon, \bar{v}} \) coincide on a neighborhood of \((\bar{x}, \bar{v})\), then \((\bar{q}, \bar{p}) \in \partial L, \mathcal{L}_{\epsilon, \bar{v}}(\bar{x}, \bar{v})\). But the hypotheses of Theorem 3.1 are satisfied when \( \mathcal{L}_{\epsilon, \bar{v}} \) replaces \( L \). We deduce that
\[ -\bar{q} \in \text{co}\{q \in \mathbb{R}^n : (q, \bar{v}) \in \partial L, \mathcal{L}_{\epsilon, \bar{v}}(\bar{x}, \bar{p})\}. \] (33)

It follows from Carathéodory’s theorem and the closure properties of the proximal subdifferential that \( \bar{q} \) is a convex combination of \((n + 1)\) points of the form
\[ q = \lim_i q_i \]
such that, for all \( i \),
\[ (-q_i, v_i) \in \partial_p H_{\epsilon, \bar{v}}(x_i, p_i) \] (34)
for some sequences \( \{v_i\} \) converging to \( \bar{v} \) and \( \{(x_i, p_i)\} \) converging to \((\bar{x}, \bar{p})\). Take such a point \( q \). Let \( \{v_i\}, \{(x_i, p_i)\} \) be sequences associated with \( q \) as above. We claim that
\[ H_{\epsilon, \bar{v}}(\bar{x}, \bar{p}) = H(\bar{x}, \bar{p}) \] (35)
and
\[ H_{\epsilon, \bar{v}}(x_i, p_i) = H(x_i, p_i). \] (36)

Let us verify these assertions. Eq. (34) implies that, for each \( i \), a constant \( \sigma_i > 0 \) can be found such that, for all \((x, p)\) in some neighborhood of \((x_i, p_i)\),
\[ H_{\epsilon, \bar{v}}(x, p) - H_{\epsilon, \bar{v}}(x_i, p_i) \geq -(q_i, x - x_i) + \langle v_i, p - p_i \rangle - \sigma_i |(x, p) - (x_i, p_i)|^2. \] (37)

It follows that, for each \( i \) and for all \( p \) in some neighborhood of \( p_i \),
\[ H_{\epsilon, \bar{v}}(x_i, p) - H_{\epsilon, \bar{v}}(x_i, p_i) \geq \langle v_i, p - p_i \rangle - \sigma_i |p - p_i|^2. \]
Since $H_{i,\delta}(x, \cdot)$ is convex, we deduce that

$$v_i \in \partial p H_{i,\delta}(x_i, \cdot)(p_i).$$

But this implies that for $i$ sufficiently large,

$$H_{i,\delta}(x_i, p_i) = \langle p_i, v_i \rangle - \mathcal{L}_{i,\delta}(x_i, v_i) = \langle p_i, v_i \rangle - L(x_i, v_i).$$

(38)

For $i$ sufficiently large, the maximum of the concave function $v \mapsto \langle p_i, v \rangle - \mathcal{L}_{i,\delta}(x_i, v)$ over $\bar{v} + \delta B_n$ is achieved at the interior point $v_i$. We conclude that (36) is true. Since $H_{i,\delta}$ is continuous on a neighborhood of $(\bar{x}, \bar{p})$ and since the sequence $\{(x_i, v_i)\}$ converges to $(\bar{x}, \bar{v})$, it follows from (38) and the lower semi-continuity of $L$ that

$$H_{i,\delta}(\bar{x}, \bar{p}) = \langle \bar{p}, \bar{v} \rangle - \lim_{i} L(x_i, v_i) \leq \langle \bar{p}, \bar{v} \rangle - L(\bar{x}, \bar{v}) \leq H_{i,\delta}(\bar{x}, \bar{p}).$$

We conclude that $v \mapsto \langle \bar{p}, v \rangle - \mathcal{L}_{i,\delta}(\bar{x}, v)$ has a local maximum at $v = \bar{v}$. Since $-L(\bar{x}, \cdot)$ is concave, this maximum is in fact a global maximum. We deduce (35). The claim is verified.

Since $H_{i,\delta}(x, p) \leq H(x, p)$ for all $(x, p)$, we deduce from (38) and (36) that, for $i$ sufficiently large, there exists $\sigma_i > 0$ such that, for all $(x, p)$ in some neighborhood of $(x_i, p_i)$,

$$H(x, p) - H(x_i, p_i) \geq \langle v_i, x - x_i \rangle + \langle v_i, p - p_i \rangle - \sigma_i \| (x, p) - (x_i, p_i) \|^2.$$

This implies, for all $i$ sufficiently large,

$$\langle -q_i, v_i \rangle \in \partial p H(x_i, p_i).$$

Since $H_{i,\delta}$ is continuous on a neighborhood of $(\bar{x}, \bar{p})$, (35) and (36) imply

$$\lim_{i} H(x_i, p_i) = \lim_{i} H_{i,\delta}(x_i, p_i) = H_{i,\delta}(\bar{x}, \bar{p}) = H(\bar{x}, \bar{p}).$$

We conclude from these relationships that $q$ satisfies

$$\langle -q, \bar{v} \rangle \in \partial L H(\bar{x}, \bar{p}).$$

But $\bar{q}$ is expressible as a convex combination of such $q$'s. It follows that

$$\bar{q} \in co\{q \in \mathbb{R}^n: \langle -q, \bar{v} \rangle \in \partial L H(\bar{x}, \bar{p})\}.$$

This is what we set out to prove. ⊡

**Proof of Theorem 3.4.** Reducing the size of $U$ if necessary, we deduce from (H1) and (H2) that $H$ never takes the value $-\infty$ on $U \times \mathbb{R}^n$, and that the conjugate functional of $H$ is $L$.

First, take a point $(\bar{q}, \bar{v}) \in \partial p H(\bar{x}, \bar{p})$. We will show that

$$-\bar{q} \in co\{q \in \mathbb{R}^n: \langle q, \bar{p} \rangle \in \partial L H(\bar{x}, \bar{v})\}.$$

We will follow the proof of Theorem 3.1 as much as we can, provided we replace $L$ by $H$ and interchange the roles of $v$'s and $p$'s. In particular, we define, for fixed $x$ in $U$, the following function $\phi_x(\cdot)$ whose argument $z$ is partitioned as $z = (y, v)$:

$$\phi_x(z) := L(y, v) - L(x, v_x) + \frac{1}{4\sigma} \left\{ \| v - \bar{v} \|^2 - \| v_x - \bar{v} \|^2 \right\} + \langle v_x - v, \bar{p} \rangle - \langle \bar{q}, x - y \rangle + \sigma \| x - y \|^2.$$
We can assume, without loss of generality, that \( \sigma > 1 \). The first delicate point is to apply the generalized Mean Value Theorem to \( \phi_x \).

Reducing the size of \( U \) if necessary, there exists \( \varepsilon > 0 \) such that \( H_{,\varepsilon}(\cdot, p) \) is Lipschitz continuous on \( U \), uniformly with respect to \( p \)'s in the bounded open ball \( P':=(\|\bar{p}\|+1)B_n \) and there exist open subsets \( V \) of \( \mathbb{R}^n \) and \( W \) of \( \mathbb{R}^n \) containing, respectively, \( \bar{v} \) and \( L(\bar{x}, \bar{v}) \), such that for all \( (x, v, L(x, v)) \) belonging to \( U \times V \times W \), we have

(i) \( V \subseteq (\bar{v} + \varepsilon B_n) \) and \( (L(\bar{x}, \bar{v}) + x\bar{B}_n) \subseteq W \) for some \( \varepsilon > 0 \);
(ii) if \( (x, v) \in U \times V \) and if \( L(x, v) \leq L(\bar{x}, \bar{v}) + \varepsilon \), then \( L(x, v) \in W \);
(iii) if \( \{x_i\} \) is a sequence converging to \( x \), then there exists a sequence \( \{v_i\} \) converging to \( v \) such that \( \limsup_{i \to \infty} L(x_i, v_i) \leq L(x, v) \);
(iv) there exists a positive constant \( K \) (we can assume \( K > 1 \)) such that the proximal normal cone to \( \text{epi} L \) at the point \( (x, v, L(x, v)) \) is included in the set \( \{(\alpha, \beta, \gamma) : |\alpha| \leq K|\beta, \gamma|\} \).

We have invoked the lower semi-continuity of \( L \), Propositions 2.1 and 2.2, as well as Remark 3.2, to ensure these relations. Now let us choose sequences \( \{\varepsilon_i\} \) and \( \{\delta_i\} \) decreasing to \( 0 \) satisfying

(a) \( \{\varepsilon_i \cdot \delta_i^{-1}\} \) is strictly bounded above by \( (4K)^{-1} \) and decreases to \( 0 \);
(b) if \( |x - \bar{x}| \leq \varepsilon_i \), then \( |v - \bar{v}| \leq \delta_i/2 \) for all \( i \);
(c) \( (\bar{x} + 3\varepsilon_i\bar{B}_n) \times (\bar{v} + (\varepsilon_i + \delta_i)\bar{B}_n) \times (\bar{p} + (\varepsilon_i + \delta_i)\bar{B}_n) \) is a subset of \( U \times V \times P \) for all \( \bar{v} \);
(d) \( \bar{p}((3+5\sigma+2|\bar{p}|+2|\bar{q}|)) < \varepsilon \).

Recall that (b) is possible since \( \lim_{v \to \bar{v}} v = \bar{v} \). Furthermore, if \( |x - \bar{x}| \leq 3\varepsilon_i \), \( |v - \bar{v}| \leq \varepsilon_i + \delta_i \) and \( L(x, v) \leq L(\bar{x}, \bar{v}) + \delta_i (3+5\sigma+2|\bar{p}|+2|\bar{q}|) \), then \( L(x, v) \in W \).

Fix \( i \in \mathbb{N} \) and \( x \) in \( \bar{x} + \varepsilon_i\bar{B}_n \) such that \( x \neq \bar{x} \). Let us apply the generalized Mean Value Theorem [7] to the lower semi-continuous function \( \phi_x(\cdot, \cdot) \) defined on the open set \( U \times V \), around the point \( (x, v_x) \) and with respect to the nonempty convex and compact subset \( Z:=\{(\bar{x}, v) : v \in \bar{v} + \delta_i\bar{B}_n\} \) of \( \mathbb{R}^{2n} \). Since, as before,

\[
\tilde{r} := \min_{Z} \phi_x(z) - \phi_x(x, v_x) = \min_{v \in \bar{v} + \delta_i\bar{B}_n} \phi_x(\bar{x}, v) \geq 0,
\]

then for the real number \( \tilde{r} := -\varepsilon_i |x - \bar{x}| < 0 \) which is strictly less than \( \tilde{r} \), and for \( \varepsilon_i > 0 \), we can find an element \( (\tilde{\xi}(x), \tilde{\xi}(x)) \) in \( \partial_{L}(y'(x), v'(x)) \) for some point \( (y'(x), v'(x)) \) in \( \mathbb{R}^n \times \mathbb{R}^n \), such that

\[
|y'(x) - x| \leq 2\varepsilon_i, \quad |v'(x) - \bar{v}| \leq \varepsilon_i + \delta_i, \quad \phi_x(y'(x), v'(x)) < \inf_{z \in Z} \phi_x(z) + |\bar{r}| + \varepsilon_i
\]

and such that, for all \( v \) in \( \bar{v} + \delta_i\bar{B}_n \),

\[
-e_i |x - \bar{x}| < \langle \tilde{\xi}(x) + \tilde{q} + 2\sigma(y'(x) - x), \bar{x} - x \rangle + \langle \tilde{\xi}(x) - \bar{p} + (2\sigma)^{-1}(v'(x) - \bar{v}) - v, v_x \rangle.
\]

Of course, we have

\[
\phi_x(y'(x), v'(x)) < \phi_x(\bar{x}, \bar{v}) + \varepsilon_i |x - \bar{x}| + \varepsilon_i,
\]
implying (by definition of $\phi$, since $|x - \bar{x}| \leq \delta/2$ and since $\sigma > 1$) that

$$L(y'(x), v'(x)) \leq L(\bar{x}, \bar{v}) + \frac{1}{4\sigma}(e_i + 3\delta)^2 + \sigma(2\delta)^2$$

$$+ (e_i + 3\delta)|\bar{p}| + 2e_i|\bar{q}| + (\sigma + 1)\epsilon^2 + e_i$$

$$\leq L(\bar{x}, \bar{v}) + \delta_i(3 + 5\sigma + 2|\bar{p}| + 2|\bar{q}|).$$

Hence, $L(y'(x), v'(x))$ belongs to $W$. Therefore, the point $(y'(x), L(y'(x), v'(x)))$ belongs to $U \times V \times W$ and since $(\zeta'(x), \zeta'(x), -1)$ is a proximal normal to $epi L$ at the point $(y'(x), v'(x), L(y'(x), v'(x)))$, we deduce that

$$|\zeta'(x)| \leq K|\zeta'(x) - (-1)|.$$

Suppose first that $|\zeta'(x)| \geq 1$.

Then certainly, since $K \geq 1$, we have

$$|\zeta'(x)| \leq 2K|\zeta'(x)|.$$

Since $|e_i - \bar{v}| \leq \delta_i/2$ and since $|x - \bar{x}| \leq \epsilon$, taking the minimum over $v$ in the final term on the right, we arrive at, for all $x$ in $\bar{x} + e_i \bar{B}_n$, with $x \neq \bar{x}$:

$$-\epsilon_i^2 \leq -e_i|x - \bar{x}| \leq (\zeta'(x) + \bar{q} + 2\sigma(y'(x) - x), \bar{x} - x)$$

$$- \frac{\delta_i}{2} |\zeta'(x) - \bar{p} + (2\sigma)^{-1}(v'(x) - \bar{v})|$$

thus, we can assume $i$ big enough so that

$$\sup_{x \in \bar{x} + e_i \bar{B}_n, x \neq \bar{x}} |\zeta'(x) - \bar{p}| < 1,$$

implying that for all $x$ in $\bar{x} + e_i \bar{B}_n$ (with $x \neq \bar{x}$)

$$|\zeta'(x)| < |\bar{p}| + 1.$$

Hence, even if $|\zeta'(x)| < 1$, we have for all $x$ in $\bar{x} + e_i \bar{B}_n$ (with $x \neq \bar{x}$),

$$\zeta'(x) \in P.$$
Denote, for each $i$,
\[ S'_i := \{ q \in \mathbb{R}^n : (q, \bar{p}) \in \partial_i L(x', v') + ((\varepsilon_i/2)(1 + 4\sigma)\overline{B}_n) \times (\gamma_i\overline{B}_n), \]
\[ \left| x' - \bar{x} \right| \leq 3\varepsilon_i, \left| v' - \bar{v} \right| \leq \varepsilon_i + \delta_i, \]
\[ L(x', v') \leq L(\bar{x}, \bar{v}) + \delta_i(3 + 5\sigma + 2|\bar{p}| + 2|\bar{q}|). \]

We have for all $i$,
\[ -\bar{q} \in \overline{S'}_i. \]

Let us show that the $S'_i$’s are closed and uniformly bounded, so that, using Carathéodory’s theorem, we can assert that
\[ -\bar{q} \in \text{co}\{q \in \mathbb{R}^n : (q, \bar{p}) \in \partial_i L(\bar{x}, \bar{v})\}. \]

First, fix $q$ in $S'_i$, i.e. for some $(x', v', L(x', v'))$ such that
\[ \left| x' - \bar{x} \right| \leq 3\varepsilon_i, \left| v' - \bar{v} \right| \leq \varepsilon_i + \delta_i, L(x', v') \leq L(\bar{x}, \bar{v}) + \delta_i(3 + 5\sigma + 2|\bar{p}| + 2|\bar{q}|), \]
we have
\[ (q, \bar{p}) = (q_1, p_1) + (q_2, p_2) \]
with $(q_1, p_1) \in \partial_i L(x', v') = \partial_i \mathcal{L}_{\varepsilon, \delta}(x', v')$ and $(q_2, p_2) \in ((\varepsilon_i/2)(1 + 4\sigma)\overline{B}_n) \times (\gamma_i\overline{B}_n)$. Apply Theorem 3.1 to the function $\mathcal{L}_{\varepsilon, \delta}$ with $P := (1 + |\bar{p}|)\overline{B}_n$, around the points $x', v', p_1$ and $q_1$ to deduce that
\[ -q_1 \in \text{co}\{q' \in \mathbb{R}^n : (q', v') \in \partial_i H_{\varepsilon, \delta}(x', p_1)\}. \]

By the Lipschitz continuity property of $H_{\varepsilon, \delta}$ on $P$, we easily obtain the uniform boundedness of the sequence $\{S'_i\}$.

Second, fix $i$ and consider a sequence $\{q_n\}$ in $S'_i$ converging to a point $q_0$, i.e. we can find a sequence $\{(x_n, v_n, L(x_n, v_n))\}$ such that, for all $n$,
\[ |x_n - \bar{x}| \leq 3\varepsilon_i, |v_n - \bar{v}| \leq \varepsilon_i + \delta_i, L(x_n, v_n) \leq L(\bar{x}, \bar{v}) + \delta_i(3 + 5\sigma + 2|\bar{p}| + 2|\bar{q}|) \]
and we have
\[ (q_n, \bar{p}) = (q'_n, p'_n) + (q''_n, p''_n), \]
with $(q'_n, p'_n) \in \partial_i L(x_n, v_n) = \partial_i \mathcal{L}_{\varepsilon, \delta}(x_n, v_n)$ and $(q''_n, p''_n) \in ((\varepsilon_i/2)(1 + 4\sigma)\overline{B}_n) \times (\gamma_i\overline{B}_n)$ for all $n$. Up to a subsequence, we can assume that $\{(x_n, v_n)\}$ converges to some point $(x_0, v_0)$ such that $|x_0 - \bar{x}| \leq 3\varepsilon_i, |v_0 - \bar{v}| \leq \varepsilon_i + \delta_i$, and that $\{(q''_n, p''_n)\}$ converges to some point $(q''_0, p''_0) \in ((\varepsilon_i/2)(1 + 4\sigma)\overline{B}_n) \times (\gamma_i\overline{B}_n)$, implying that $\{(q'_n, p'_n)\}$ converges to $(q_0, \bar{p}) - (q''_0, p''_0)$. Apply Theorem 3.1 to $\mathcal{L}_{\varepsilon, \delta}$ around the points $x_n, v_n, p'_n$ and $q''_n$, with $P := (1 + |\bar{p}|)\overline{B}_n$, and deduce that
\[ -q'_{n} \in \text{co}\{q \in \mathbb{R}^n : (q, v_n) \in \partial_i H_{\varepsilon, \delta}(x_n, p'_n)\}. \]

Since $(q_n, v_n)$ is in $\partial_i H_{\varepsilon, \delta}(x_n, p'_n)$, and since $H_{\varepsilon, \delta} (\cdot, \cdot)$ is continuous in a neighborhood of $(x_n, p'_n)$ and since $H(x, \cdot)$ is convex on $\mathbb{R}^n$ (for $x$ close to $\bar{x}$, hence to $x_n$), then
certainly $v_n$ belongs to $\partial P; (x_n, v_n)$, i.e. $L;v_n(x_n, v_n) + H;v_n(x_n, p') = \langle p_n', v_n \rangle$ for all $n$. Hence, by continuity of $H;v_n$, we have

\[
\limsup_{n \to \infty} L;v_n(x_n, v_n) = \lim_{n \to \infty} \{ \langle p_n', v_n \rangle - H;v_n(x_n, p') \}
\]

\[
= \langle \tilde{p} - \tilde{p}_0', v_0 \rangle - H;v_n(x_0, \tilde{p} - \tilde{p}_0')
\]

\[
\leq \sup_{p \in \mathbb{R}^n} \{ \langle p, v_0 \rangle - H;v_n(x_0, p) \}
\]

\[
= L;v_n(x_0, v_0)
\]

because $L;v_n$ is lower semi-continuous, convex and proper (by its local epi-Lipschitz property). Now using the lower semi-continuity of $L;v_n$ in a neighborhood of $(\tilde{x}, \tilde{v})$, we obtain that

\[
\lim_{n \to \infty} L;v_n(x_n, v_n) = L;v_n(x_0, v_0);
\]

from there, it is easy to see that $S'$ is closed. So we have deduced that if $(\tilde{q}, \tilde{v}) \in \partial P; (\tilde{x}, \tilde{p})$, then $-\tilde{q} \in \mathcal{C}o\{q \in \mathbb{R}^n : (q, \tilde{p}) \in \partial L;v_n(\tilde{x}, \tilde{v})\}$.

It remains to show that this assertion remains true when $(\tilde{q}, \tilde{v}) \in \partial P; (\tilde{x}, \tilde{p})$, but this is straightforward using similar arguments as the ones previously seen.

References


