4-2003

Multiobjective Optimal Control Problems with Endpoint and State Constraints

Kirsty J. Eisenhart
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Mathematics Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/1224

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
MULTIOBJECTIVE OPTIMAL CONTROL PROBLEMS
WITH ENDPOINT AND STATE CONSTRAINTS

by

Kirsty J Eisenhart

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of
requirements for the
Degree of Doctor of Philosophy
Department of Mathematics

Western Michigan University
Kalamazoo, Michigan
April 2003
In this thesis we consider nonsmooth multiobjective optimal control problems in terms of a general preference on $\mathbb{R}^m$. The optimal control problems considered involve differential inclusion, endpoint constraints and state constraints. No convexity assumption is needed on the differential inclusion. Examples of common preferences are given, and the idea of approximating a preference is introduced. Euler-Lagrange necessary conditions and a form of the maximum principle are developed for closed preferences (and those that can be approximated by closed preferences) in terms of the limiting subdifferential. As a consequence, this is the first result in the literature for lexicographical order. Also this is the first time noncontinuous utility functions are considered. A transversality condition in terms of the limit supremum of normal cones to the level sets of the preference is also developed.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ACKNOWLEDGEMENTS

Let me begin by thanking the faculty and staff of the Mathematics and Statistics department of Western Michigan University for their devotion to their students. I especially want to thank the people involved in the Analysis seminar for their patience and comments during my first few presentations. Also, I would have been at a loss without my committee's helpful comments and suggestions throughout my dissertation process. I especially want to thank my advisor, Jim Zhu, for his wisdom and his ability to know when to let me struggle through a topic and when to guide me along.

I also wish to thank all of the people who believed in me. There are far too many to list, but the most influential were my parents, Jan and Jean Eisenhart. Without their belief and initiative I may have never gone to college. Without their support I might have listened to my guidance counselors and the world would have another nurse Hatchet to fear. Furthermore I wish to thank: my Mom for her philosophy that her children could do whatever they wanted and for showing us how to achieve our desires (for some unknown reason I listened to this lecture); my Dad for his philosophy that if it's worth doing then it's worth giving your best, and for showing me the many uses of duct tape (you can see irony runs deep in this family); my grandparents, aunts and uncles for instilling in me that no matter your age, experience, or education, your mother is always right, and the sooner you admit this the sooner you can eat; and my two siblings, Dara and Andy, who will never know the many things they have taught me through the years, basically because I’ll never tell.

The most important lessons of life are the ones which help you determine who you are. Beyond my family there are many people who have enlightened and/or challenged me in this way. Instead of listing them all I will list one person from each of several different stages in my life: Renee Westfall, Judy Manion, and Jason Schultz.
As usual, I saved the best for the end: a very special thanks goes to my spouse, Terry Grant, for always exceeding my expectations and insisting on several hugs throughout the day.

Kirsty J Eisenhart
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ......................................................... ii
LIST OF FIGURES ............................................................... v

1  INTRODUCTION ............................................................. 1

2  COMMON PREFERENCES OVER $\mathbb{R}^n$ ............................ 5

3  NONSMOOTH ANALYSIS BACKGROUND ....................... 9

4  MULTIFUNCTIONS AND TOOLS FOR LIMITING PROCESSES .... 24

5  APPROXIMATELY CLOSED PREFERENCES .................... 29

6  MAIN RESULTS ............................................................. 35

7  A NONCONTINUOUS UTILITY FUNCTION RESULT ........... 52

REFERENCES ................................................................. 55
LIST OF FIGURES

1 Lexicographical level sets in \( \mathbb{R}^2 \) .................................................. 7
2 Limiting \( \{ \prec \} \)-normal cone for lexicographical order in \( \mathbb{R}^2 \) .......... 33
3 \( l(x, y) \) for \( c := u(x, y) \geq 1 \) ............................................................ 52
4 \( l(x, y) \) for \( c := u(x, y) \in (0, 1) \) ....................................................... 53
1 INTRODUCTION

Let $\prec$ be a nonreflexive preference (binary relation) for elements of $\mathbb{R}^m$ where $\mathbb{R}^m$ represents Euclidean $m$-space. The main concern of this thesis is the multiobjective Mayer problem: minimize

$$\phi(x(a), x(b))$$

over absolutely continuous functions (arcs) $x : [a, b] \to \mathbb{R}^n$ with respect to the preference $\prec$ and subject to the differential inclusion

$$x(t) \in F(t, x(t)) \text{ a.e. } t \in [a, b].$$

the general endpoint constraints

$$(x(a), x(b)) \in \Omega \subset \mathbb{R}^{2n}.$$ (1.3)

and the state constraint

$$h(t, x(t)) \leq 0 \text{ for all } t \in [a, b].$$ (1.4)

In the notation above, the objective function $\phi$ is a vector valued function from $\mathbb{R}^{2n}$ to $\mathbb{R}^m$. $h$ is a function on $\mathbb{R}^{n+1}$. $T := [a, b]$ is a given time interval. and $F : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is a set-valued mapping (multifunction). We will label this Mayer problem by $(\mathcal{M}_\prec)$.

Given the fact that most practical decisions are made based on not one, but on a host of factors that are often competing, it is necessary to develop a theory that can deal with these multiple factors simultaneously. This is the essential difference between single objective and multiobjective optimal control problems. It should be noted that in [13], Debreu showed that some preferences can be represented by a continuous utility function. In other words, there exists a continuous function $u : \mathbb{R}^m \to \mathbb{R}$ such that
$y < r$ if and only if $u(y) < u(r)$. One could then use this utility function to reduce $(\mathcal{M}_\gamma)$ to a single objective problem by replacing the objective function $\phi$ with $u \circ \phi$. Unfortunately this provides little help at this time since the current single objective results require the objective function to be Lipschitz continuous. For these reasons this thesis will focus directly on multiobjective optimal control problems.

There has been extensive research done on single objective problems (see [9], [17], [21], [22], [25], [26], [28], [31], [42], [45], [47]). Similar to some of these studies we choose to model control problems with differential inclusion (1.2). This is a very natural model for a dynamic control system since it directly associates the set of possible velocities $\dot{x}$ with any given state $x$ and moment $t$. Recently Zhu [47] used nonsmooth penalty functions to formulate Hamiltonian inclusion conditions for multiobjective optimal control problems with convex-valued differential inclusion. Although not explicitly stated in the paper, this process also provides Euler-Lagrange inclusion conditions.

A classical result from calculus of variations is that given a strongly convex integrand (in terms of the velocity argument), the Euler equation and the Hamiltonian system are equivalent. A partial generalization of this fact is provided by Loewen and Rockafellar in [25], where they prove that a necessary condition for a minimum of an optimal control problem with convex-valued differential inclusion (1.2) is the existence of an adjoint arc which simultaneously fulfills the maximality principle. Clarke's form of the Hamiltonian inclusion, and Loewen and Rockafellar's form of Euler-Lagrange inclusion. Kaskosz and Lojasiewicz present examples in [22] which show that there is no such equivalence without the convexity assumption on $F$.

Many control systems do not lead to convex differential inclusions. Also, determining a Hamiltonian necessary condition for even a single valued problem with nonconvex differential inclusion is a difficult problem which, until recently, had been a longstanding open question: in [10] Clarke determined Hamiltonian necessary conditions for such problems. For these reasons this thesis will not assume any convexity
Recently Bellaassali and Jourani [3] derived Euler-Lagrange necessary conditions for multiobjective optimal control problems with nonconvex differential inclusion constraints. This can be viewed as a partial generalization of Zhu's results to nonconvex differential inclusions. The next natural extension is to add state constraints (1.4) since a variety of physical, economical, and engineering problems can be modeled in this fashion. Necessary conditions for problems with state constraints are developed in this thesis.

When first analyzing the problem of forming necessary conditions for multiobjective optimal control problems with nonconvex-valued differential inclusions, it seems natural to adapt Zhu's nonsmooth penalization method for multiobjective optimal control problems. The dilemma here is that a relaxation method is needed to convert the nonconvex constraint to a convex constraint, and such a relaxation will not guarantee the endpoint constraints. The next logical step is to analyze and try to adapt Euler-Lagrange necessary conditions for single objective optimal control problems with nonconvex-valued differential inclusion. One such method would be to extend Mordukhovich's discrete approximation method [31] to multiobjective problems: note that this was the first Euler-Lagrange result derived for problems with nonconvex differential inclusion. Another method would be similar to Ioffe's method found in [21] which separates the singular case out. In this thesis, similar to Bellaassali and Jourani [3], we chose to explore the latter. The advantage of this method is that a large class of preferences, mainly approximately closed preferences, can be handled.

In addition to handling state constraints, another feature of this paper is the introduction of the idea of approximating a preference by a family of preferences. As a result, necessary conditions are developed for an even larger family of preferences: approximately closed preferences. All previous literature on multiobjective problems required the preference to be closed. In the next section (Section 2) it is proven that lexicographical order is not a closed preference. Fortunately lexicographical order lies
in the family of approximately closed preferences, and as a result we obtain Euler-Lagrange necessary conditions for \( \mathcal{M}_\prec \) over lexicographical order. This is the first result for lexicographical order. This method of approximating preferences with a family of preferences also allows one to investigate other preferences not previously studied, such as preferences determined by discontinuous utility functions (see Section 7).

The two main tools used throughout this paper are nonsmooth analysis and variational methods [4] [16]. Specifically, properties of normal cones and subdifferentials of lower semicontinuous functions and locally Lipschitz continuous functions are utilized ([8], [11], [23], [24], [28], [29], [30], [32], [38], [39]). Other tools used include known necessary conditions for single valued optimal control problems ([7], [9], [17], [21], [22], [24], [25], [26], [28], [31], [39], [40], [42], [45]), recent results for multiobjective problems ([3], [47]), and compactness of trajectories ([9], [39], [40]). The contents of the remainder of this thesis are as follows. Section 2 contains definitions associated with preferences and examples of common preferences. Section 3 focuses on results from nonsmooth analysis which will be utilized in the thesis. Section 4 focuses on results from functional analysis and results dealing with multifunctions which will help in the limiting processes in the proof of the main theorem. The rest of the background material is found in Section 5 where the definition of an approximately closed preference and the definition of a solution to the Mayer problem \( \mathcal{M}_\prec \) are given. Once sufficient background is presented, the main result is given in Section 6 and Section 7 contains an example which illustrates an advantage of approximating preferences.
Throughout this paper a preference on a space \( Y \) will simply be a nonreflexive binary relation on \( Y \). In other words, a preference is a subset \( R \) of the cartesian product \( Y \times Y \) which does not contain \( (y, y) \) for some \( y \in Y \). We say \( y \) is preferred to \( r \) if and only if \( (y, r) \in R \), and since we are dealing with minimization problems we will write \( y < r \). The motivation for the nonreflexive assumption will become apparent in the result section after the definition of the solution to control problems \( \{ \mathcal{M}_x \} \).

In this thesis we will be optimizing over \( \mathbb{R}^m \). Here are a few examples of common preferences used when optimizing over \( \mathbb{R}^m \).

**Example 2.1.** Weak Pareto preference is defined by \( y < r \) if \( y_i \leq r_i \), \( i = 1, \ldots, m \), and \( y \neq r \).

**Example 2.2.** Strong Pareto preference is defined by \( y < r \) if \( y_i < r_i \), \( i = 1, \ldots, m \).

Note that the weak Pareto preferences can be defined in terms of a cone, that is to say \( y < r \) if \( y - r \in K \) and \( y \neq r \) where \( K = \mathbb{R}^m_+ := \{ x \in \mathbb{R}^m : x_i \leq 0, \quad i = 1, \ldots, m \} \). We generalize this below.

**Example 2.3.** To define a generalized weak Pareto preference let \( K \) be a closed cone in \( \mathbb{R}^m \). Then \( y < r \) if \( y - r \in K \) and \( y \neq r \).

Recall that the notion of a continuous utility function was discussed in the introduction. Below we formalize our definition.

**Example 2.4.** Let \( u : \mathbb{R}^m \to \mathbb{R} \) be a continuous function. Then the preference defined by this utility function is such that \( y < r \) if \( u(y) < u(r) \).

Our final example is commonly known as dictionary order.
Example 2.5. Lexicographical order is defined by \( y < r \) if there exists an integer \( q \in \{0, 1, \ldots, m - 1\} \) such that \( y_i = r_i \) for \( i = 1, \ldots, q \), and \( y_{q+1} < r_{q+1} \).

Throughout this paper we will use the following notation for the preference level sets:

\[
l(r) := \{s : s < r\}.
\]

Definition 2.6. A preference \( \prec \) is closed if the following two properties hold.

(A1) For all \( r \in \mathbb{R}^m \), we have \( r \in \overline{l(r)} \).

(A2) If \( t \in \overline{l(r)} \) and \( r < s \), then \( t < s \).

A pointed cone is one which satisfies \( K \cap -K = \{0\} \). The theorem below shows how assumption (A2) restricts a generalized weak Pareto preference, \( \prec_K \), which is defined by the closed cone \( K \).

Theorem 2.7. A generalized weak Pareto preference, \( \prec_K \), satisfies (A2) if and only if \( K \) is convex and pointed.

Proof. To show the forward direction, suppose \( \prec_K \) satisfies (A2). \( k_1, k_2 \in K \setminus \{0\} \), and \( \lambda \in (0, 1) \). Let \( r \) be an arbitrary element of \( \mathbb{R}^m \) and define \( t := r + \lambda k_1 \) and \( s := r - (1 - \lambda)k_2 \). Then since \( \lambda k_1 \neq 0 \), \( t \prec_K r \); similarly \( r \prec_K s \). Thus by (A2) \( t \prec_K s \) which means that \( \lambda k_1 + (1 - \lambda)k_2 = t - s \in K \). Thus \( K \) is convex.

Next let \( k \in K \cap -K \) and define \( t := r + k \) and \( s := r - (-k) \). If \( k \neq 0 \), then again by (A2) we have \( t \prec_K s \). But this means that \( 0 = t - s \in K \setminus \{0\} \), which is clearly a contradiction, and so \( k = 0 \).

To show the converse, suppose \( K \) is convex and pointed, and assume \( t \in \overline{l(r)} \) and \( r <_K s \). Then there are \( k_1, k_2 \in K \) such that \( t = r + k_1 \), \( s = r - k_2 \), and \( k_2 \neq 0 \). Since \( K \) is convex \( \frac{1}{2}(t-s) = \frac{1}{2}k_1 + \frac{1}{2}k_2 \in K \), so \( t \in \overline{l(s)} \). But if \( t = s \), then \( k_1 = -k_2 \neq 0 \) which contradicts the pointedness of \( K \), so \( t <_K s \).
Note that in the lexicographical ordering if \( t \in \overline{l(r)} \), then \( t - r \) lies in the cone \( K := \{ x \in \mathbb{R}^m : x_1 \leq 0 \} \). Given Theorem 2.7, since this cone is not pointed one might expect that lexicographical ordering does not satisfy (A2). The following illustration shows that indeed lexicographical ordering on \( \mathbb{R}^2 \) does not satisfy (A2).

![Figure 1. Lexicographical level sets in \( \mathbb{R}^2 \).](image)

The next example shows that lexicographical ordering on \( \mathbb{R}^m \) does not satisfy (A2) for any \( m \).

**Example 2.8.** Let \( \prec_L \) denote the lexicographical ordering on \( \mathbb{R}^m \) and consider the following points:

\[
\begin{align*}
t_n &= (-\frac{1}{n}, 1, 1, 0, \ldots, 0) & t &= (0, 1, 1, 0, \ldots, 0) \\
r &= (0, 0, 0, \ldots, 0) & s &= (0, 0, 1, 0, \ldots, 0)
\end{align*}
\]

Then \( r \prec_L s \) and for each \( n \), \( t_n \prec_L r \) but \( t \not\prec_L s \). Thus \( t \in \overline{l(r)} \) and \( r \prec_L s \) but \( t \not\prec_L s \).

The final example of this section is a family of preferences defined by pointed convex cones. This example will be utilized in Section 4 to show that lexicographical order is closable (which will be defined in Section 4).
Example 2.9. Consider the sequence of generalized weak Pareto preferences, \( \prec_j \), defined by the following cones:

\[
K_j := \{ y \in \mathbb{R}^m : \langle e_1, y \rangle \leq -\frac{1}{j} \| y \| \} = \{ y \in \mathbb{R}^m : y_1 \leq -\frac{1}{j} \| y \| \}.
\]

Note that \( e_1 = (1, 0, \ldots, 0) \). We claim that each of these preferences is closed since each \( K_j \) is a closed convex pointed cone.

To see this observe that each \( K_j \) is closed; and for each \( J \) if \( \lambda \geq 0 \) and \( k \in K_j \) then \( \lambda k \in K_j \), so \( K_j \) is a cone. To see that \( K_j \) is pointed, let \( -u, u \in K_j \). Then

\[
0 = \langle e_1, -u \rangle + \langle e_1, u \rangle \leq -\frac{1}{J} \| u \| \leq 0,
\]

which implies \( u = 0 \). Finally to see that each \( K_j \) is convex, let \( u, v \in K_j \) and \( \lambda \in [0, 1] \). Then

\[
\lambda u + (1 - \lambda) v \geq -\frac{1}{j} (\lambda \| u \| + (1 - \lambda) \| v \|) \geq \lambda u_1 + (1 - \lambda) v_1.
\]
In this section we will define the normal cones and subdifferentials used in this thesis and briefly list properties of these items. Additional details, methods and insight can be found in the following books and papers: Aubin and Frankowska [2], Borwein and Zhu [5] [6], Clarke [7] [8] [9], Clarke et al [11], Ioffe [18] [19], Mordukhovich [27] [28] [29] [30], Mordukhovich and Shao [32], Warga [43] [44], and Zhu [46]. Note that given a closed set $C$ the notation $x \rightharpoonup x$ signifies that $x_i \to x$ and $x_i \in C$ for all $i$. Given a function $f$, the notation $x \rightharpoonup x$ means $x_i \to x$ and $f(x_i) \to f(x)$. Also throughout this paper $B$ will represent the closed unit ball in Euclidean space.

**Definition 3.1.** Let $C$ be a closed subset of $\mathbb{R}^k$ and let $x \in C$.

(i) A proximal normal to $C$ at $x$ is an element $p \in \mathbb{R}^k$ for which there exists $\Delta > 0$ and $M_0 > 0$ such that the following proximal normal inequality holds:

$$
\langle p, y - x \rangle \leq M_0\|y - x\|^2 \text{ for all } y \in C \cap (x + \Delta B).
$$

The proximal normal cone to $C$ at $x$ is the set of proximal normals to $C$ at $x$ and is represented by $N^P(x; C)$.

(ii) The limiting normal cone to $C$ at $x$ is the following set:

$$
N(x; C) := \{ p : \exists \ x_i \overset{C}{\rightharpoonup} x, p_i \to p \text{ with } p_i \in N^P(x_i; C) \quad \forall \ i \}.
$$

Elements of $N(x; C)$ are referred to as limiting normals to $C$ at $x$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(iii) Given a preference, \( \prec \), the \( \prec \)-normal cone at \( x \) is defined in terms of the Kuratowski-Painlevé limit supremum of sets:

\[
N_\prec(x) := \limsup_{x_i \to x} N(x_i; \overline{l(x_2)})
\]

\[
= \left\{ p : \exists x_{1k} \to x. x_{2k} \to x. p_k \to p \text{ with } x_{1k} \in \overline{l(x_{2k})} \right\}.
\]

The proximal normal cone defined above is due to Clarke. The limiting normal cone was introduced by Mordukhovich in [27]. Bellaassali and Jourani motivated our definition of the preference normal in [3].

**Remark 3.2.** Note that the \( \prec \)-normal cone at \( x \) can be written in terms of the Kuratowski-Painlevé limit supremum of proximal normal cones:

\[
\left\{ p : \exists x_{1k} \to x. x_{2k} \to x. p_k \to p \text{ with } x_{1k} \in \overline{l(x_{2k})}, \text{ and } p_k \in N(x_{1k}; \overline{l(x_{2k})}) \right\}
\]

\[
= \left\{ p : \exists x_{1k} \to x. x_{2k} \to x. p_k \to p \text{ with } x_{1k} \in \overline{l(x_{2k})}, \text{ and } p_k \in N^P(x_{1k}; \overline{l(x_{2k})}) \right\}
\]

\[
= \limsup_{x_i \to x} N^P(x_i; \overline{l(x_2)}).
\]

**Remark 3.3.** Clearly by equation (3.1), if \( C_1 \subset C_2 \) then \( N^P(x; C_1) \supset N^P(x; C_2) \).

The next few propositions for normal cones and ideas for proofs of these propositions can be found in [8], [11], [28], [29], [24], and [39]. For a more thorough look at the proximal normal see Clarke et. al [11].
Proposition 3.4. There exists a $\Delta > 0$ and $M_0 > 0$ such that equation (3.1) holds if and only if there exists $M > 0$ such that

$$
(p, y - x) \leq M\|y - x\|^2 \text{ for all } y \in C.
$$

(3.2)

Proof. Clearly if equation (3.2) holds then equation (3.1) holds with $M_0 = M$. Assume there exists a $\Delta > 0$ and $M_0 > 0$ such that equation (3.1) holds. Let $M = \max\{M_0, \frac{\|p\|}{\Delta}\}$. Then equation (3.2) holds.

Proposition 3.5. The vector $p$ is an element of $N^P(x; C)$ if and only if there exists a point $z \in \mathbb{R}^k$ and a scalar $r > 0$ such that $x$ is the closest point in $C$ to $z$ and $p = r(z - x)$.

Proof. The proximal normal inequality (3.2) gives the following inequality:

$$
\|x + \frac{p}{2M} - y\|^2 = (x + \frac{p}{2M} - y, x + \frac{p}{2M} - y)
= \frac{1}{4M^2}\|p\|^2 + \frac{1}{M}(p, x - y) + \|x - y\|^2
= \frac{1}{4M^2}\|p\|^2 - \frac{1}{M}(p, y - x) + \|y - x\|^2
\geq \frac{1}{4M^2}\|p\|^2 - \|y - x\|^2 + \|y - x\|^2.
$$

Thus the proximal normal inequality (3.2) can be manipulated to the following equivalent statement:

There exists $M > 0$ such that $\|x + \frac{p}{2M} - y\| \geq \frac{1}{2M}\|p\|$ for all $y \in C$.

Setting $r = 2M$ and $z = x + \frac{p}{2M}$ yields the following:

$$
\min\{\|z - y\| : y \in C\} = \|z - x\| \text{ and } p = r(z - x).
$$

Thus $p \in N^P(x; C)$ implies there exists a point $z \in \mathbb{R}^k$ and a scalar $r > 0$ such that $x$ is the closest point in $C$ to $z$ and $p = r(z - x)$. Conversely if there exists a point $z \in \mathbb{R}^k$...
and a scalar \( r > 0 \) such that \( x \) is the closest point in \( C \) to \( z \) and \( p = r(z - x) \) then for all \( y \in C \)

\[
\langle p, y - x \rangle = r\langle z - x, y - x \rangle \\
\leq r\|z - x\|\|y - x\| \\
\leq r\|y - x\|\|y - x\|.
\]

so equation (3.2) holds with \( M = r \).

**Remark 3.6.** By the above proposition, \( N(x; C) \) is a cone. In other words, for all \( p \in N(x; C) \) and for all \( \lambda \geq 0 \), \( \lambda p \not\in N(x; C) \).

**Proposition 3.7.** If \( C \) is a closed set then \( N(x; C) \) contains a nonzero element if and only if \( x \) is in the boundary of \( C \).

**Proof.** To see this note that \( x \) in the boundary of \( C \) implies there exists a sequence \( z_n \not\in C \) such that \( z_n \to x \). Since \( C \) is closed, there exists \( x_n \in C \) such that \( x_n \) is the closest point to \( z_n \) in \( C \) and since \( z_n \to x \), \( x_n \to x \). Then by the above proposition 3.5, \( w_n := \frac{z_n - x_n}{\|z_n - x_n\|} \in N^F(x_n; C) \) and \( \|w_n\| = 1 \). Thus some subsequence of \( w_n \) converges to \( w \in N(x; C) \). To see the converse note that by the definition of \( N(x; C) \) and by Proposition 3.5, for each \( w \in N(x; C) \) there exists \( z_n \in \mathbb{R}^k \), \( r_n > 0 \), and \( x_n \to x \) such that \( p_n := r_n(z_n - x_n) \to w \) and \( x_n \) is the closest point in \( C \) to \( z_n \). Now if \( w \neq 0 \) then we may assume without loss of generality that \( \|p_n\| > 0 \). Thus \( \|z_n - x_n\| > 0 \) which means that each \( x_n \) is in the boundary of \( C \). Finally since the boundary of \( C \) is closed, \( x \) is in the boundary of \( C \).

**Proposition 3.8.** Let \( C \) be a convex set. Then the proximal normal cone and the limiting normal cone coincide with the convex normal cone:

\[
N(x; C) = N^F(x; C) = N_{\text{convex}}(x; C) := \{ p : \langle p, y - x \rangle \leq 0 \text{ for all } y \in C \}.
\]
Proof. Clearly if \( \langle p, y - x \rangle \leq 0 \) for all \( y \in C \) then \( p \in N^p(x; C) \) with \( M = 0 \).

Next let \( p \in N^p(x; C) \). Since \( C \) is convex \( y = (1 - \lambda)x + \lambda z \in C \) for each \( z \in C \) and \( \lambda \in [0, 1] \). Thus for all \( z \in C \) and each \( \lambda \in [0, 1] \) we have

\[
\lambda \langle p, z - x \rangle = \langle p, y - x \rangle \leq M\|y - x\|^2 = M\lambda^2\|z - x\|^2.
\]

Dividing each side by \( \lambda \) and then letting \( \lambda \to 0 \) yields the desired result.

Example 3.9. By Proposition 3.8 above, \( x \in \text{int } C \) implies that \( N(x; C) = \{0\} \).

Example 3.10. Let \( \prec \) represent lexicographical order. Since \( \overline{I(r)} = r + \{p : p_1 \leq 0\} \), by Proposition 3.8 for each \( x \not\in \text{int } \overline{I(r)} \),

\[
N^p(x; \overline{I(r)}) = \{ \lambda e_1 : \lambda \geq 0 \text{ and } e_1 = (1.0.0\ldots.0) \}.
\]

Thus

\[
N_\prec(x) = N(x; \overline{I(x)}) = \{ \lambda e_1 : \lambda \geq 0 \text{ and } e_1 = (1.0.0\ldots.0) \}.
\]

Proposition 3.7 shows the existence of a nontrivial element in the limiting normal cone. The following proposition guarantees the existence of a nontrivial element in the \( \prec \)-normal cone when \( \prec \) is a closed preference.

Proposition 3.11. If the preference \( \prec \) is closed, then for all \( r \in \mathbb{R}^m \) there exists a nonzero element in \( N_\prec(r) \).

Proof. By (A1) there exists \( r_k \to r \) with \( r_k \prec r \); and since \( \prec \) is nonreflexive, \( r \not\in \overline{I(r_k)} \). Thus, as in Proposition 3.7, there exist \( \theta_k \in \overline{I(r_k)} \) such that \( p_k := \frac{e - 2\theta_k}{\|r - \theta_k\|} \in N^p(\theta_k; \overline{I(r_k)}) \). Then by taking a subsequence if necessary, \( p_k \to p \) with \( \|p\| = 1 \); and since \( r_k \to r \) and \( \theta_k \to r \), \( p \in N_\prec(r) \).

In the following definition \( \text{dom}(f) \) represents the domain of the function \( f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\} \) and \( \epsilon f \) represent the epigraph of \( f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\} \). These are
defined by the following:

\[ \text{dom}(f) := \{ x : f(x) < \infty \} \], and

\[ \text{epi} f := \{ (x, \alpha) \in \mathbb{R}^k \times \mathbb{R} : \alpha \geq f(x) \} \].

**Definition 3.12.** Let \( f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\} \) be a lower semicontinuous (l.s.c.) function and let \( x \in \text{dom}(f) \).

(i) The proximal subdifferential of \( f \) at \( x \) is the set:

\[ \partial^P f(x) := \{ \xi : (\xi, -1) \in N^P((x, f(x)); \text{epi} f) \} \].

(ii) The limiting subdifferential of \( f \) at \( x \) is the set:

\[ \partial f(x) := \{ \xi : (\xi, -1) \in N((x, f(x)); \text{epi} f) \} \].

(iii) The asymptotic limiting subdifferential of \( f \) at \( x \) is the set:

\[ \partial^0 f(x) := \{ \xi : (\xi, 0) \in N((x, f(x)); \text{epi} f) \} \].

Elements of the set of subdifferentials are called subgradients.

This method of defining the subdifferential in terms of the normal cone at the epigraph of the function was first utilized by Rockafellar[33]. Since the above definition defines the proximal subdifferential in terms of the proximal normal cone, \( \xi \in \partial^P f(x) \) if \( \xi \) satisfies the proximal subgradient inequality. The following proposition illustrates this fact more precisely; a proof of this can be found in [11] or [39].

**Proposition 3.13.** Let \( f \) be a l.s.c. function and let \( x \in \text{dom}(f) \). Then \( \xi \in \partial^P f(x) \) if and only if there exist positive numbers \( \sigma \) and \( \Delta \) such that the following proximal
subgradient inequality holds:

\[ f(y) \geq f(x) + \langle \xi, y - x \rangle - \sigma \|y - x\|^2 \quad \text{for all } y \in x + \Delta B. \]  

(3.3)

**Remark 3.14.** Using the proximal subgradient inequality, one can show that if \( f \) is a Lipschitz continuous function of rank \( K \) near \( x \), then \( \partial f(x) \subset KB \). This is due to the fact that \( K \|y - x\| \geq |f(x) - f(y)| \geq \langle \xi, y - x \rangle - \sigma \|y - x\|^2 \) for all \( y \in x + \Delta B \), which implies that \( K \geq \|\xi\| - \sigma \|y - x\| \) for all \( y \in x + \Delta B \) with \( y \neq x \). Then allowing \( y \to x \) yields \( K \geq \|\xi\| \).

Definition 3.12 expresses the limiting subgradients and the asymptotic limiting subgradients in terms of the limiting normal cone to the epigraph. The next proposition takes the reverse view and can be found in Section 4.3 of Vinter [39] (see notes in Vinter for additional references).

**Proposition 3.15.** Let \( f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\} \) be a l.s.c. function and let \( x \in \text{dom } f \). Then \( N((x, f(x)) : \epsilon \text{pi } f) = \{ (\lambda \xi, -\lambda) : \lambda > 0, \xi \in \partial f(x) \} \cup \{ \partial^\infty f(x) \times \{0\} \} \), and so by Proposition 3.7 either \( \partial f(x) \) is nonempty or \( \partial^\infty f(x) \) contains a nonzero element.

The following proposition was explored by Kruger and Mordukhovich in [23]. The proposition shows that the existence of a nonzero asymptotic limiting subgradient of \( f \) at \( x \in \text{dom } f \) implies "non-Lipschitzian behavior" of \( f \) near \( x \).

**Proposition 3.16.** Let \( f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\} \) be a l.s.c. function and suppose \( x \in \text{dom } f \). Then \( f \) is Lipschitz continuous on a neighborhood of \( x \) if and only if \( \partial^\infty f(x) = \{0\} \).

**Remark 3.17.** Note that by the above two propositions (Proposition 3.15 and Proposition 3.16), if \( f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\} \) is a l.s.c. function which is Lipschitz continuous on a neighborhood of \( x \), then \( \partial f(x) \) is nonempty.
As one might expect, a limiting subgradient of \( f \) at \( x \) can be written as the limit of a sequence of subgradients of \( f \) at \( x \), where \( x_i \to x \). Before proving this fact we consider the following two lemmas.

**Lemma 3.18.** If \((p, -\lambda) \in N^P((x, r); \text{epi } f)\), then \( \lambda \geq 0 \).

**Proof.** By Proposition 3.4 there exists \( M > 0 \) such that for all \((y, s) \in \text{epi } f\),

\[
(p, y - x) - \lambda (s - r) = \langle p, -\lambda \rangle + \langle y, s \rangle - \langle x, r \rangle 
\leq M \|(y, s) - (x, r)\| \leq M \|y - x\|^2 + (s - r)^2.
\]

Now \((x, r + \frac{1}{n}) \in \text{epi } f\) for all \( n > 0 \) since \( f(x) \leq r < r + \frac{1}{n} \). Thus for all \( n > 0 \),

\[-\lambda \left( \frac{1}{n} \right) \leq \frac{M}{n}, \quad \text{which means that } \lambda \geq -\frac{M}{n}.\]

Taking the limit as \( n \to \infty \) yields \( \lambda \geq 0 \).

**Lemma 3.19.** If \((p, -\lambda) \in N^P((x, r); \text{epi } f)\) and \( \lambda > 0 \), then \( r = f(x) \).

**Proof.** Again, as in the above proof, by Proposition 3.4 there exists \( M > 0 \) such that for all \((y, s) \in \text{epi } f\), we have \( (p, y - x) - \lambda (s - r) \leq M \|y - x\|^2 + (s - r)^2 \). Suppose \( f(x) < r \). Then there is an \( N \) sufficiently large such that for all \( n \geq N \), \( f(x) < r - \frac{1}{n} \). Thus \((x, r - \frac{1}{n}) \in \text{epi } f\), and so for all \( n \geq N \) we have \( \lambda \leq \frac{M}{n} \). Taking the limit as \( n \to \infty \) and using Lemma 3.18 shows \( \lambda = 0 \).

**Proposition 3.20.** Given a l.s.c. function \( f \) with \( x \in \text{dom}(f) \) the limiting subdifferential of \( f \) at \( x \) can be written as the limit of sequences of proximal subgradients as follows:

\[
\partial f(x) = \{ \xi : \exists x_i \overset{\varepsilon}{\to} x \text{ and } \xi_i \to \xi \text{ with } \xi_i \in \partial^P f(x_i) \forall i \}.
\]

**Proof.** It is clear by the definition of the limiting subdifferential, Definition 3.12, that \( \{ \xi : \exists x_i \overset{\varepsilon}{\to} x \text{ and } \xi_i \to \xi \text{ with } \xi_i \in \partial^P f(x_i) \forall i \} \subset \partial f(x) \). To show the opposite inclusion, assume that \( \xi \in \partial f(x) \). By the definition of the limiting subdifferential.
\((\xi, -1) \in N(((x, f(x)) : \epsilon \triangleright f)\) which means there exists \((\xi_m, -\lambda_m) \rightarrow (\xi, -1)\) and there exists \((x_m, r_m) \rightarrow f((x, f(x)))\) such that \((\xi_m, -\lambda_m) \in N_P((x_m, r_m) : \epsilon \triangleright f)\). Then by Lemma 3.19, for \(m\) sufficiently large, \(r_m = f(x_m)\) and so \(\xi_m \in \partial P f(x_m)\) with \(\xi_m \rightarrow \xi\) and \(x_m \rightarrow x\).

**Remark 3.21.** Using a diagonalization argument one can show that

\[ \partial f(x) = \{ \xi : \exists x_i \rightarrow x \text{ and } \xi_i \rightarrow \xi \text{ with } \xi_i \in \partial f(x_i) \ \forall \ i \} . \]

Next we present a few calculus rules for the subdifferentials defined earlier. Versions of these propositions and their proofs can be found in [11], [28], and [39]. Note that limiting normal cones and limiting subgradients play a crucial role in the main theory of this paper. For a more complete approach to the calculus of limiting normal cones and limiting subgradients see [30] and [32].

**Proposition 3.22.** For all \(k \geq 0\), \(\partial (kf)(x) = k \partial f(x)\).

**Proposition 3.23.** Let \(f\) be a l.s.c. function and let \(x \in \text{dom } f\). If \(f\) attains a minimum at \(x\), then

\[ 0 \in \partial P f(x) \subseteq \partial f(x) . \]

**Proposition 3.24.** If \(f \in C^2(U)\), then for all \(x \in U\), \(\partial P f(x) = \partial f(x) = \{ \nabla f(x) \} \).

**Proposition 3.25.** Let \(f(x) = \|x\|\). Then for \(x \neq 0\), \(\partial P f(x) = \left\{ \frac{x}{\|x\|} \right\} \).

**Proposition 3.26 (Sum Rule).** Suppose \(f_1\) and \(f_2\) are l.s.c. functions which are bounded below and \(x \in \text{dom}(f_1) \cap \text{dom}(f_2)\). If one of the functions is Lipschitz continuous near \(x\), then \(\partial (f_1 + f_2)(x) \subseteq \partial f_1(x) + \partial f_2(x)\).
Proposition 3.27 (Chain Rule). Let \( \phi : \mathbb{R}^k \rightarrow \mathbb{R}^m \) be Lipschitz near \( x \), and let \( h : \mathbb{R}^m \rightarrow \mathbb{R} \) be Lipschitz near \( \phi(x) \). Then

\[
\partial (h \circ \phi)(x) \subset \{ \partial (\gamma \circ \phi)(x) : \gamma \in \partial h(\phi(x)) \}.
\]

In the next proposition the notation \( \text{dist}(x; C) \) represents the distance from \( x \) to \( C \):

\[
\text{dist}(x; C) := \inf_{y \in C} ||x - y||.
\]

Proposition 3.28. Suppose \( C \) is a closed set of \( \mathbb{R}^n \) and \( x \not\in C \). Then there exists \( c \in C \) such that

\[
\begin{cases}
x - c \\
\frac{||x - c||}{||x - c||}
\end{cases} \subseteq \partial^p \text{dist}(x; C) \subseteq N^p(c; C).
\]

Finally we give common forms of Ekeland's variational theorem (see [1] [6] [8] [9] [16] [24] [35] [39]) and an exact penalization theorem (see [6] [8] [9] [11] [35] [39]). This will be followed by two known finite Lagrangian necessary conditions for given single-objective optimal control problems which can be found in [39]. These will be utilized in the proof of the Main Lemma.

Theorem 3.29 (Ekeland's Variational Theorem). [16] Take a complete metric space \( \{X, m(\cdot, \cdot)\} \), a l.s.c. function \( f : X \rightarrow \mathbb{R} \cup \{+\infty\} \), a point \( x_0 \in \text{dom } f \), and numbers \( \varepsilon > 0 \) and \( \lambda > 0 \). Assume that

\[
f(x_0) \leq \inf_{x \in X} f(x) + \varepsilon.
\]

Then there exists \( \bar{x} \in X \) such that

(i) \( f(\bar{x}) \leq f(x_0) \).

(ii) \( m(x_0, \bar{x}) \leq \lambda \).

(iii) \( f(\bar{x}) \leq f(x) + \frac{\varepsilon}{\lambda} m(x, \bar{x}) \) for all \( x \in X \).
Theorem 3.30 (Exact Penalization Theorem). Let \((X, m(\cdot, \cdot))\) be a metric space, \(C \subset X\) and \(f : X \to \mathbb{R}\). Assume that \(f\) satisfies a Lipschitz condition on \(X\) with Lipschitz constant \(K\). Let \(x\) be a minimizer for the constrained minimization problem.

\[
\text{minimize } f(x) \text{ over points } x \in X \text{ satisfying } x \in C. \tag{3.4}
\]

Then for any \(K > K\), \(x\) is also a minimizer for the unconstrained minimization problem.

\[
\text{minimize } f(x) + K \, \text{dist}(x; C) \text{ over points } x \in X. \tag{3.5}
\]

If \(K > K\) and \(C\) is closed, then the converse assertion is also true: any minimizer \(x\) for the unconstrained problem (3.5) is also a minimizer for the constrained problem (3.4) and so, in particular, \(x \in C\).

Throughout the paper \(W^{1,1}\) represents the space of absolutely continuous functions \(x : [a, b] \to \mathbb{R}^n\) equipped with the following norm:

\[
\|x\|_* := \|x(a)\| + \int_a^b \|\dot{x}(t)\| \, dt.
\]

Also throughout this paper \(B\) represents the Euclidean closed ball and \(B_*\) represents the \(W^{1,1}\) closed ball:

\[
B := \{x \in \mathbb{R}^k : \|x\| \leq 1\} \quad \text{and} \quad B_* := \{x \in W^{1,1} : \|x\|_* \leq 1\}.
\]

One more comment on notation for this paper. \(\text{co}(A)\) will represent the convex hull of the set \(A\).

The next two theorems involve finite Lagrangian conditions for Bolza problems and utilize the following functions \(l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}\), \(L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}\), and \(h : [a, b] \times \mathbb{R}^n \to \mathbb{R}\). Additionally \(\dot{x}\) is a feasible trajectory to a given Bolza problem with objective function \(J\) if \(\dot{x}\) satisfies the constraints of the problem; and \(\dot{x}\) is a local minimizer to this Bolza problem if \(\dot{x}\) is a feasible trajectory and there exists \(\Delta_0\) such
that for all feasible trajectories \( x \in x + \Delta_0 B_* \), \( J(x) \leq J(x) \). Finally for Bolza problems with state constraints we will need to introduce the set \( C^+([a, b]) \). Here \( C^+([a, b]) := C^+([a, b]; \mathbb{R}^n) \) is a subset of the topological dual \( C^*([a, b]; \mathbb{R}^n) \) of the space of continuous functions \( C([a, b]; \mathbb{R}^n) \) with supremum norm. The norm on \( C^*([a, b]; \mathbb{R}^n) \) is the induced norm and is written as \( \| \cdot \|_{T.V.} \). Then \( C^+([a, b]) \) is the set of elements in \( C^*([a, b]; \mathbb{R}^n) \) which take nonnegative values on nonnegative-valued functions in \( C([a, b]; \mathbb{R}^n) \); in other words \( C^+([a, b]) \) represents the positive linear functionals on \( C([a, b]; \mathbb{R}^n) \). By the Riesz Representation Theorem, elements of \( C^*([a, b]; \mathbb{R}^n) \) can be identified with finite regular vector-valued signed measures on the Borel subsets of \([a, b]\), so we refer to elements \( \mu \in C^*([a, b]; \mathbb{R}^n) \) as "measures." Also by the Riesz Representation Theorem, for \( \mu \in C^+([a, b]; \mathbb{R}^n) \) the norm \( \| \cdot \|_{T.V.} \) coincides with the total variation of \( \mu \):

\[
\| \mu \|_{T.V.} = \int_a^b \mu(ds).
\]

**Theorem 3.31 (Finite Lagrangian Condition).** [39] Let \( x \) be a \( W^{1,1} \) local minimizer for

\[
\begin{aligned}
&\text{Minimize } J(x) := l(x(a), x(b)) + \int_a^b L(t, x(t), \dot{x}(t)) dt \\
&\text{over arcs } x \in W^{1,1}.
\end{aligned}
\]

Assume that \( J(x) < \infty \) and the following hypotheses hold.
(H1) \( l \) is l.s.c.

(H2) \( L(t, \cdot, \cdot) \) is l.s.c. for almost all \( t \in [a, b] \) and \( L(\cdot, x, \cdot) \) is \( \mathcal{L} \times B^n \) measurable for each \( x \in \mathbb{R}^n \). In other words, \( L \) is Lebesgue measurable in the time variable and Borel measurable in the velocity variable.

(H3) For every \( N > 0 \), there exists \( \Delta > 0 \) and \( k \in L^1 \) such that
\[
|L(t, x', v) - L(t, x, v)| \leq k(t)|x' - x|, \quad L(t, x(t), v) \geq -k(t)
\]
for all \( x', x \in x + \Delta B \) and \( v \in \dot{x} + NB \).

Then there exists an arc \( p \in W^{1,1} \) that satisfies

(i) \( \dot{p}(t) \in \text{co} \{ \eta : (\eta, p(t)) \in \partial L(t, x(t), \dot{x}(t)) \} \) a.e.;

(ii) \( (p(a), -p(b)) \in \partial l(x(a), x(b)) \); and

(iii) \( \langle p(t), \dot{x}(t) \rangle - L(t, x(t), \dot{x}(t)) \geq \langle p(t), v \rangle - L(t, x(t), v) \) \( \forall v \in \mathbb{R}^n \) a.e.

Theorem 3.32 (Finite Lagrange Condition with State Constraints). [39] Let \( x \) be a \( W^{1,1} \) local minimizer for

(FLS) \[
\begin{align*}
\text{Minimize } J(x) & : = l(x(a), x(b)) + \int_a^b L(t, x(t), \dot{x}(t)) \, dt \\
\text{over arcs } x & \in W^{1,1} \text{ satisfying} \\
h(t, x(t)) & \leq 0 \text{ for all } t \in [a, b].
\end{align*}
\]

Assume that \( J(\dot{x}) < \infty \) and the following hypotheses hold.
(G1) \( l \) is l.s.c.

(G2) \( L(\cdot, x, \cdot) \) is \( \mathcal{L} \times \mathcal{B}^n \) measurable for each \( x \in \mathbb{R}^n \) and \( L(t, \cdot, \cdot) \) is l.s.c. for a.e. \( t \in [a, b] \).

(G3) For every \( N > 0 \), there exists \( \delta > 0 \) and \( k \in L^1 \) such that, for a.e. \( t \in [a, b] \):

\[
|L(t, x', v) - L(t, x, v)| \leq k(t)|x' - x|, \quad L(t, x(t), v) \geq -k(t)
\]

for all \( x', x \in x + \Delta B \) and \( v \in \dot{x} + NB \).

(G4) There exists \( \varepsilon > 0 \) and a constant \( k_h \) such that \( h \) is upper semicontinuous on \( \{(t, x) : t \in [a, b], x \in x(t) + \varepsilon B\} \) and

\[
|h(t, x) - h(t, x')| \leq k_h|x - x'|
\]

for all \( x, x' \in x + \varepsilon B \) and \( t \in [a, b] \).

Then there exists an arc \( p \in W^{1,1}([a, b], \mathbb{R}^n) \), \( \lambda \geq 0 \), a measure \( \mu \in \mathcal{C}^{+,+}(a, b) \), and a \( \mu \)-integrable function \( \zeta : [a, b] \to \mathbb{R}^n \) such that

(i) \( \lambda + \|p\|_{L^\infty} + \|\mu\|_{TV} = 1 \);

(ii) \( \tilde{p}(t) \in \text{co} \left\{ \eta : (\eta, p(t) + \int_{a}^{t} \zeta(s) \mu(ds), -\lambda) \right\} \in N \left( (\dot{x}(t), \dot{x}(t), L(t, x(t), \dot{x}(t)), \text{epi} L(t, \cdot, \cdot)) \right) \) a.e.:

(iii) \( (p(a), -[p(b) + \int_{a}^{b} \zeta(s) \mu(ds)], -\lambda) \in N \left( (\dot{x}(a), \dot{x}(b), L(x(a), x(b)), \text{epi} l) \right) \)
(iv) $\left\langle p(t) + \int_{[a, t]} \zeta(s) \mu(ds) \right\rangle \dot{x}(t) - \lambda L(t, x(t), \dot{x}(t)) \geq$

$\left\langle \left( p(t) + \int_{[a, t]} \zeta(s) \mu(ds) \right) \cdot v \right\rangle - \lambda L(t, x(t), v) \forall \ v \in \mathbb{R}^n \ a.e.$ and

(v) $\zeta(t) \in \partial^2 h(t, x(t)) \mu \ a.e.$ where $\partial^2 h(t, x)$ is defined as follows:

$\partial^2 h(t, x) := \text{co} \left\{ \xi : \exists \ (t_i, x_i) \overset{h}{\rightarrow} (t, x) \ \text{with} \ h(t_i, x_i) > 0 \right\}$

for all $i$ and $\nabla_x h(t_i, x_i) \rightarrow \xi$.

Note, as indicated, the last two propositions (Proposition 3.31 and Proposition 3.32) along with their proofs can be found in [39]. For further references see Vinter’s notes.
4 MULTIFUNCTIONS AND TOOLS FOR LIMITING PROCESSES

In this section we will provide some technical results which will be utilized in the proof of the main theorem. Many of these results involve multifunctions. A multifunction \( G : X \to Y \) is simply a set-valued map. This means that for each \( x \in X \), \( G(x) \) is a subset of \( Y \). Given \( S \subset X \), we say that \( G \) is closed on \( S \) provided \( G(x) \) is a closed subset of \( Y \) for each \( x \in S \). Similarly \( G \) is compact, convex, or nonempty on \( S \) provided the subset \( G(x) \) possesses the appropriate property for each \( x \in S \). The graph of \( G \) is defined by \( \text{Gr} G := \{(x, y) : x \in X \text{ and } y \in G(x)\} \), and \( G \) is called closed if \( \text{Gr} G \) is a closed subset of \( X \times Y \). Clearly if \( G \) is closed, then \( G \) is closed on \( S \) for all \( S \subset X \).

If \( X \) and \( Y \) are Banach spaces, then \( G \) is upper semicontinuous (u.s.c.) at \( x \in X \) if for each \( \epsilon > 0 \) there exists \( \delta > 0 \) such that

\[
G(z) \subset G(x) + \epsilon B_Y \text{ for all } z \in x + \delta B_X.
\]

Here \( B_X \) and \( B_Y \) represent the closed unit balls in \( X \) and \( Y \), respectively. As usual, \( G \) is u.s.c. on the set \( S \) if \( G \) is u.s.c. at \( x \) for each \( x \in S \).

The multifunction \( G : S \to Y \) is said to be measurable if for each open set \( O \subset Y \), the set

\[
\{x \in S : G(x) \cap O \neq \emptyset\}
\]

is measurable. Note that an equivalent definition exists in terms of closed subsets of \( Y \).

**Proposition 4.1.** Let \( S \) be a nonempty subset of \( \mathbb{R}^q \) and let \( G : S \to \mathbb{R}^m \) be a closed multifunction. Suppose there exists a compact subset \( K \subset \mathbb{R}^m \) such that for all \( x \in S \), \( G(x) \subset K \). Then
(i) $G$ is u.s.c. on $S$, and

(ii) if $x \in S$ with $G(x) = \emptyset$ then there exists $\Delta > 0$ such that $G(x) = \emptyset$ for all $x \in S \cap (x + \Delta B)$.

Proof. To show (i) suppose the hypotheses are met, but $G$ is not u.s.c. on $S$. Then there exists $\varepsilon > 0$, $\{x_n\} \subset S$, $\{y_n\} \subset \mathbb{R}^n$, and $x \in S$ such that $x_n \to x$, $y_n \in G(x_n)$ for each $n$, and $y_n \notin G(x) + \varepsilon B$ for each $n$. Since each $y_n \in K$ and $K$ is a compact subset of $\mathbb{R}^n$, we may assume (by re-enumerating if need be) that $y_n$ converges to some $y \in K$. Then since $Gr G$ is closed, $y_n \in G(x_n)$ for each $n$, and $(x_n, y_n) \to (x, y)$, we must have $y \in G(x)$ which contradicts the fact that $y_n \notin G(x) + \varepsilon B$ for each $n$.

To show (ii) assume to the contrary. Then there exists $x_n \to x$ such that for each $n$ there exists $y_n \notin G(x_n) \subset K$. As in the first part of this proof, we may assume $y_n \to y$. Then since $G$ is closed we have the following contradiction: $y \notin G(x)$.

The next theorem can be found in a convex analysis text (such as [33]).

**Theorem 4.2 (Carathéodory Theorem).** Let $A \subset \mathbb{R}^n$. If $x \in co(A)$, then $x = \sum_{j=1}^{n+1} \lambda_j x_j$ where $x_j \in A$ for each $j$, $\lambda_j \geq 0$ for each $j$, and $\sum_{j=1}^{n+1} \lambda_j = 1$.

**Remark 4.3.** Note that by Carathéodory, if $A$ is a compact subset of $\mathbb{R}^n$ then $co(A)$ is also a compact subset of $\mathbb{R}^n$.

**Proposition 4.4.** Let $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and suppose $f$ is Lipschitz continuous of rank $K_f$ on $(x + \varepsilon B) \times \mathbb{R}^n$. Define $G(w, x, v, s) := co\{ (\eta, w) : (\eta, w) \in \partial f(x, v) + \{0\} \times s B \}$. Then $G$ is a closed multifunction and for all $\sigma \in (0, 1)$, $G$ is u.s.c. on $S_\sigma := \mathbb{R}^n \times (x + \sigma \varepsilon B) \times \mathbb{R}^n \times [0, \sigma \varepsilon]$.

Proof. Fix $\sigma \in (0, 1)$. Since $f$ is Lipschitz continuous of rank $K_f$ on $(x + \sigma \varepsilon B) \times \mathbb{R}^n \subseteq (x + \varepsilon B) \times \mathbb{R}^n$, by Remark 3.14, $\partial f(x, v)$ is bounded by $K_f$. Thus $G(w, x, v, s) \subseteq K_f B$ for each $(w, x, v, s) \in S_\sigma$. So if $Gr G$ is closed in $S_\sigma \times \mathbb{R}^n$, then we obtain our result using Proposition 4.1.
To show that $GRG$ is a closed subset of $V := S_x \times \mathbb{R}^n$, assume $(w^m, x^m, v^m, s^m, \eta^m) \rightarrow (w, x, v, s, \eta)$ in $V$ and $\eta^m \in G(w^m, x^m, v^m, s^m)$ for each $m$. Then by the Carathéodory Theorem 4.2, $\eta^m = \sum_{j=1}^{n+1} \lambda_j^m \eta_j^m$ with $\lambda_j^m \geq 0$, $\sum_{j=1}^{n+1} \lambda_j^m = 1$, and $(\eta_j^m, w^m) \in \partial f(x^m, v^m) + \{0\} \times s^m B$. Since $\sigma < 1$, eventually $(x^m, v^m) \in (x + \varepsilon B) \times \mathbb{R}^n$, and so by Remark 3.14, $(\eta_j^m, w^m)$ is bounded by $K_f + s^m \leq K_f + \varepsilon$. Thus for each $j$ a subsequence of $\eta_j^m$ converges to some $\eta_j$ as $m \rightarrow \infty$. Without loss of generality we may assume $\eta_j^m \rightarrow \eta_j$ for each $j$. Similarly since $\lambda_j^m \in [0, 1]$, we may assume for each $j$ that $\lambda_j^m \rightarrow \lambda_j$, $\lambda_j \geq 0$, and $\sum_{j=1}^{n+1} \lambda_j = 1$. Then $\eta^m \rightarrow \eta := \sum_{j=1}^{n+1} \lambda_j \eta_j$ and since $f$ is Lipschitz continuous by Remark 3.21, $(\eta, w) \in \partial f(x, v) + \{0\} \times sB$ for each $j$. Hence $\eta \in G(w, x, v, s)$ which means $GRG$ is closed on $V$ which in turn implies that $G$ is u.s.c. on $S_x$.

The next result is due to Vinter and Pappas [40].

**Proposition 4.5.** Let $G$ be a compact convex multifunction from $[a, b] \times \mathbb{R}^n$ to $\mathbb{R}^n$ having closed graph, and whose images are contained in a given bounded set. Let $\nu_i$ be a sequence of positive measures converging weak$^*$ to $\nu_0$, and let $x_i$ be a sequence of arcs converging uniformly to $x$, such that, for each $i$ there is a measurable function $\gamma_i \in G(t, x(t))$ $\nu_i$-a.e. Then there exists a measurable function $\gamma_0$, which is $\nu_0$ integrable, such that $\gamma_0(t) \in G(t, x(t))$ $\nu_0$-a.e. and such that some subsequence of the measures $\gamma_i \nu_i$ converges weak$^*$ to $\gamma_0 \nu_0$.

The next two theorems rely on linear operator theory, the first of which is a specific case of Theorem iv.8.9 of Dunford [15].

**Theorem 4.6 (Dunford-Pettis Theorem).** Let $K \subset L^1([a, b], \mathbb{R}^n)$. Then $K$ is weakly sequentially precompact if and only if $K$ is bounded and for each $\varepsilon_0 > 0$ there is a $\delta_0 > 0$ such that for each measurable set $E$ of $[a, b]$ and for each $f \in K$, if $\mu(E) < \delta_0$ then $|\int_E f \, d\mu| < \varepsilon_0$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Theorem 4.7 (Mazur's Lemma). Let $X$ be a Banach space and suppose that $\{x_n\}$ is a sequence of $X$ which converges weakly to $x$. Then a sequence of convex combinations of $\{x_n\}$ converges strongly to $x$.

This last result of this section is a simpler form of the compactness of trajectories theorem which can be found in [11] or [39].

Proposition 4.8. Let $\varepsilon > 0$. Let $\{x_k\}$ and $\{p_k\}$ be sequences in $W^{1,1}$ such that $x_k \to x$ in $W^{1,1}$. Let $m \in L^1[a,b]$ with $m(t) \geq 0$ a.e., let $\{y_k\}$ be a sequence of measurable functions such that $y_k(t) \to 0$ a.e., and let $\{s_k\}$ be a sequence of positive numbers such that $s_k \to 0$. Suppose $G : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is a convex multifunction such that the following are satisfied:

1. the sequence $\{p_k(t)\}$ is bounded;
2. given functions $p'(t)$, $x'(t)$, $v'(t)$, and $s'(t)$ if $\gamma(t) \in G(p'(t), x'(t), v'(t), s'(t))$ and $x'(t) \in x(t) + \varepsilon B$ for all $t$, then $||\gamma(t)|| \leq m(t)$ a.e.;
3. for all $k$, $p_k(t) \in G(t, p_k(t) + y_k(t), x_k(t), s_k(t))$ a.e. $t \in [a,b]$;
4. for almost all $t \in [a,b]$, $(p, x, v, s) \to G(t, p, x, v, s)$ is a closed multifunction which is u.s.c. on the interior of $\mathbb{R}^n \times (x + \varepsilon B) \times \mathbb{R}^n \times [0, \varepsilon]$.

Then there is a subsequence of $\{p_k\}$ which converges uniformly to an arc $p$ which satisfies the following equation:

$$\dot{p}(t) \in G(t, p(t), x(t), \dot{x}(t), 0) \text{ a.e. } t \in [a,b].$$
Proof. Define $\mathcal{K} := \left\{ f \in L^1([a,b], \mathbb{R}^n) : \|f\| \leq m(t) \ a.e. \ t \in [a,b] \right\}$. Clearly $\mathcal{K}$ is a bounded subset of $L^1([a,b], \mathbb{R}^n)$, and given $\varepsilon_0 > 0$ there exists $\delta_0 > 0$ such that $\mu(E) < \delta_0$ implies $\int_E m(t) \ dt < \frac{\delta_0}{\sqrt{n}}$. Then for all measurable sets with $\mu(E) < \delta_0$, $|\int_E f \ d\mu| < \varepsilon_0$ for each $f \in \mathcal{K}$. Note that $\hat{p}_k \in \mathcal{K}$ for each $k$ by hypothesis (ii). Thus by Dunford-Pettis (Theorem 4.6), taking a subsequence if needed, $\hat{p}_k \rightharpoonup \hat{p}$ weakly in $\mathcal{K} \subset L^1$. Since $\{p_k(a)\}$ is bounded, again by taking a subsequence if needed, $p_k(a) \rightarrow A$ in $\mathbb{R}^n$. Let $p(t) := A + \int^t_0 \dot{p}(s) \ ds$, then $p_k(t) \rightarrow p(t)$ uniformly on $[a,b]$ and $\hat{p}_k \rightharpoonup \hat{p}$ weakly. Then by Mazur's lemma (Theorem 4.7), there exists a convex combination $\{q^1_j \}_{j=1}^\infty$ of $\{p_k\}_{k=1}^\infty$ which converges to $p$ in $W^{1,1}$. Similarly for each $k$ there exists a convex combination $\{q^k_j \}_{j=1}^\infty$ of $\{p_j\}_{j=1}^\infty$ which converges to $p$ in $W^{1,1}$. Let $q_k = q^k_k$. Then $q_k$ converges to $p$ in $W^{1,1}$, and $\dot{q}_k = \sum_{j=1}^\infty \lambda^k_j \dot{p}_j$ with $\lambda^k_j \geq 0$ and $\sum_{j=1}^\infty \lambda^k_j = 1$.

Next define the set

$$E := \bigcup_{n=1}^\infty \left\{ t \in [a,b] : \hat{p}_k(t) \notin G(t, p_k(t) + y_k(t), x_k(t), \dot{x}_k(t), s_k) \right\}$$

$$\cup \left\{ t \in [a,b] : \dot{q}_k(t) \neq \dot{p}(t) \right\} \cup \left\{ t \in [a,b] : \dot{x}_k(t) \neq \dot{x}(t) \right\}$$

$$\cup \left\{ t \in [a,b] : y_k \neq 0 \right\} \cup \left\{ t \in [a,b] : (p(t, x, c, s) \rightarrow (t, p(t, x, c, s) is not closed \right\}$$

$$\cup \left\{ t \in [a,b] : (p(t, x, c, s) \rightarrow (t, p(t, x, c, s) is not u.s.c. \right\}$$

Then $E$ is a set of measure zero such that $\hat{p}_k(t) \in G(t, p_k(t) + y_k(t), x_k(t), \dot{x}_k(t), s_k)$ for all $t \notin E$. Fix $t \notin E$. By (iv), for each $m > 0$ there exists $M$ such that $\|\dot{q}_k(t) - \dot{p}(t)\| \leq \frac{1}{m}$ and $\hat{p}_k \in G(t, p_k(t) + y_k(t), x_k(t), \dot{x}_k(t), s_k) \subset G(t, p(t), \dot{x}(t), \dot{x}(t), 0) + \frac{1}{m} B$ for each $k \geq M$. Since $G$ is a convex multifunction and $\dot{q}_k$ is a convex combination of $\{p_k\}_{k=1}^\infty$, $\dot{q}_k \in G(t, p(t), \dot{x}(t), \dot{x}(t), 0) + \frac{1}{m} B$. Thus $\dot{p}(t) \in G(t, p(t), \dot{x}(t), \dot{x}(t), 0) + \frac{1}{m}$ for each $m > 0$. Finally since $(p(t, x, c, s) \rightarrow G(t, p(t, x, c, s)$ is a closed multifunction, $G(t, p(t), \dot{x}(t), \dot{x}(t), 0)$ is a closed subset of $\mathbb{R}^n$. Thus allowing $m \rightarrow \infty$ yields

$$\dot{p}(t) \in G(t, p(t), \dot{x}(t), \dot{x}(t), 0) \text{ for all } t \notin E.$$
Let \( \prec \) be a nonreflexive preference for elements of \( \mathbb{R}^m \), and recall the multiobjective Mayer problem:

\[
\begin{align*}
\text{Minimize} & \quad \phi(x(a), x(b)) \\
\text{subject to} & \quad \dot{x}(t) \in F(t, x(t)) \text{ a.e. } t \in [a, b] \\
& \quad (x(a), x(b)) \in \Omega \subset \mathbb{R}^2 \\
& \quad h(t, x(t)) \leq 0 \text{ for all } t \in [a, b]
\end{align*}
\]

(5.1) - (5.3)

This section develops the definition of a solution to the above Mayer problem and creates the framework for the main results by defining an approximately closed preference and developing a normal cone in terms of this approximation. The new normal cone is then rewritten in terms of proximal normal cones.

**Definition 5.1.** A feasible trajectory of \( (\mathcal{M}_{\prec}) \) is an absolutely continuous function over \([a, b]\) which satisfies the differential inclusion (5.1), the endpoint constraint (5.2), and the state constraint (5.3).

**Definition 5.2.** A local solution of \( (\mathcal{M}_{\prec}) \) is a feasible trajectory \( x \) for which there exists \( \Delta_0 > 0 \) such that no feasible trajectory \( \tilde{x} \) of \( (\mathcal{M}_{\prec}) \) satisfies both

\[
\phi(x(a), x(b)) < \phi(x(a), \tilde{x}(b)) \quad \text{and} \quad \|x - \tilde{x}\| \leq \Delta_0.
\]

Note that if \( x \) is a solution of \( (\mathcal{M}_{\prec}) \), then \( \phi(x(a), x(b)) \neq \phi(x(a), x(b)) \). This is another instance in which the nonreflexive assumption is utilized. This assumption will also simplify the subderivative calculations in the proof of the Main Lemma 6.3 and, as seen in Proposition 3.11, this assumption will guarantee the existence of a nonzero
element in the $\prec$-normal cone of a closed preference $\prec$.

The following definition of approximately closed preferences relies on the definition of a solution to $(\mathcal{M}_\prec)$.

**Definition 5.3.** A preference $\prec$, is approximately closed at $x$, a solution of $(\mathcal{M}_\prec)$, if there exists a sequence of closed preferences $\{\prec_J\}$ and a sequence $\{x_J\} \subset W^{1,1}$ such that for each $J$, $x_J$ is a solution of $(\mathcal{M}_{\prec_J})$ and $x_J \to x$ in $W^{1,1}$. The sequence of pairs $\{(\prec_J, x_J)\}$ is called a closed approximator of the preference $\prec$ at $x$.

The example below illustrates that for any solution $x$ of $(\mathcal{M}_\prec)$ over the lexicographical ordering, there is a closed approximator to the lexicographical order at $x$.

**Example 5.4.** Let $\prec$ represent lexicographical order on $\mathbb{R}^m$ and consider the sequence of generalized weak Pareto preferences, $\prec_J$, defined by the closed convex pointed cones of Example 2.9:

$$K_J := \left\{ y \in \mathbb{R}^m : (e_1, y) \leq -\frac{1}{J} \|y\| \right\} = \left\{ y \in \mathbb{R}^m : y_1 \leq -\frac{1}{J} \|y\| \right\}.$$

Then for each solution $x$ of $(\mathcal{M}_\prec)$, $\{(\prec_J, x)\}$ is a closed approximator of lexicographical order on $\mathbb{R}^m$ at $x$.

To see this note that for all $u, v \in \mathbb{R}^m$, if $u \prec_J v$ then $u - v \in K_J/\{0\}$. Thus $u_1 - v_1 \leq -\frac{1}{J}\|u - v\| < 0$, which in turn means that $u \prec v$. In other words for each $v \in \mathbb{R}^m$, $l_J(v) \subset l(v)$, where $l_J(v)$ represents the level set at $v$ corresponding to $\prec_J$ and $l(v)$ represents the level set at $v$ corresponding to lexicographical order. Thus, if $\hat{x}$ is a solution to $(\mathcal{M}_\prec)$, then $\hat{x}$ is a solution to $(\mathcal{M}_{\prec_J})$ for each $J \geq 0$. for if $x$ is a feasible solution of $(\mathcal{M}_{\prec_J})$ with $\circ(x(a), x(b)) \prec_J \circ(\hat{x}(a), \hat{x}(b))$, then $x$ is a feasible solution of $(\mathcal{M}_\prec)$ with $\circ(x(a), x(b)) \prec \circ(\hat{x}(a), \hat{x}(b))$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The above example can be generalized as in the following proposition. Again we will use the notation \( l(c) \) to represent the level set at \( c \) corresponding to \(<\) and \( l_j(c) \) to represent the level set at \( c \) corresponding to \(<_j\).

**Proposition 5.5.** Let \( x \) be a solution of \((\mathcal{M}_\prec)\) and let \( \gamma = \circ(x(a), x(b)) \). Assume there exists a sequence of closed preferences \( \{<_j\} \) such that for each \( J \) there exists a \( \Delta > 0 \) for which

\[
l_j(z) \cap (z + \Delta B) \subset l(z) \text{ for all } z \in \gamma + \Delta B.
\]

Then the sequence of pairs \( \{(-<_j, x)\} \) is a closed approximator of \( \prec \) at \( x \).

A preference which satisfies equation (5.4) in the above proposition at \( \gamma \) will be called **closable at** \( \gamma \). As usual, a preference will be called closable if the preference is closable at each \( \gamma \). As will be seen in Section 7, some discontinuous utility functions are contained in the family of closable preferences, and as such necessary conditions can be explored for Mayer problems with these discontinuous utility functions. It should also be noted that closable preferences can be dealt with simply in terms of level sets instead of sequences of Mayer problems. In the more general setting of closed approximators one may be able to examine necessary conditions for Mayer problems over more complex preferences.

Next we develop a definition for a preference normal cone when the preference is approximately closed. Note that, as with the \( \prec \)-normal cone, this normal cone can also be rewritten in terms of a Kuratowski-Painlevé limit supremum of proximal normal cones.

**Definition 5.6.** Let \( \{(-<_j, x_j)\} \) be a closed approximator of the preference \( \prec \) at \( \hat{x} \) and let \( \gamma := \circ(\hat{x}(a), \hat{x}(b)) \). Then the limiting \( \{<_j\} \)-normal cone at \( \gamma \) corresponding to \( \{(-<_j, x_j)\} \) is defined as follows:
\[ N_{\{\prec J\}}(\gamma) := \limsup_{\gamma \to \gamma} N_{\prec J}(\gamma) \]

\[ = \left\{ p : \exists J_k \to \infty, \gamma_k \to \gamma, p_k \to p \text{ with } p_k \in N_{\prec J_k}(\gamma_k) \right\} \]

\[ = \left\{ p : \exists J_k \to \infty, y_k \to \gamma, z_k \to \gamma, p_k \to p \text{ with } \right. \]
\[ y_k \in \overline{I_{J_k}(z_k)} \text{ and } p_k \in N^P(y_k; \overline{I_{J_k}(z_k)}) \}

\[ = \limsup_{\gamma \to \gamma} N^P(y; \overline{I_J(z)}). \]

Clearly \( N^P(y; \overline{I_J(z)}) \subseteq N_{\{\prec J\}}(\gamma) \). More interesting is the case in which the closed approximator can be chosen such that these two normal cones coincide, in other words, the case in which the limiting \( \{\prec J\} \)-normal cone at \( \gamma \) does not become larger than the original normal cone.

**Example 5.7.** Clearly if \( \prec \) is closed and \( x \) is a solution of \((M_x)\), then \( \{\prec, x\} \) is a closed approximator and \( N_{\{\prec, x\}}(\gamma) = N_\prec(\gamma) \).

Next consider lexicographical order and the closed approximator of Example 5.4: generalized weak Pareto preferences, \( \prec J \), defined by the closed convex pointed cones

\[ K_J := \left\{ y \in \mathbb{R}^m : \langle e_1, y \rangle \leq -\frac{1}{J} ||y|| \right\} = \left\{ y \in \mathbb{R}^m : y_1 \leq -\frac{1}{J} ||y|| \right\}. \]

The following picture illustrates that \( N^P(y; \overline{I_J(z)}) = N_{\{\prec J\}}(\gamma) \) for this closed approximator of lexicographical order on \( \mathbb{R}^2 \).
The following example verifies that $N^P(\gamma;\overline{l_j(\gamma)}) = N_{|x_j|}(\gamma)$ for lexicographical order in general.

**Example 5.8.** Let $\prec$ represent lexicographical order over $\mathbb{R}^m$, let $x$ be a solution of $(M_\prec)$, let $\gamma = \sigma(x(a),x(b))$, and let $<_j$ be the closed approximation as in Example 5.4.

Then $N_{|x_j|}(\gamma) = N_\prec(\gamma) = \{\lambda e_1 : \lambda \geq 0 \text{ and } e_1 = (1,0,0,\ldots,0)\}$.

To see this let $p \in N_{|x_j|}(\gamma)$. By re-enumerating, if necessary, one may assume there exists $p_j \in N^P(y_j;l_j(z_j)) = N^P(y_j;l_j(z_j))$ such that $p_j \rightarrow p$, $y_j \rightarrow \gamma$, $z_j \rightarrow \gamma$, and $y_j \in l_j(z_j) = z_j + K_j$. Let $y_{j+} = z_j + k_j$. Note that for each $J$, $N^P(y_J;l_J(z_J)) = N^P(y_J - z_J;K_J) = N^P(k_J;K_J)$ and since each $K_J$ is convex by Proposition 3.8

$$\langle p_j, k_j \rangle \leq 0 \quad \text{for all } k \in K_J = \bigcup_{m=1}^{J} K_m.$$ 

Thus

$$\langle p, k - k_J \rangle \leq 0 \quad \text{for all } k \in \bigcup_{m=1}^{J} K_m.$$ 

which implies that

$$\langle p, k \rangle - \|p_J - p\| \|k - k_J\| - \|p\| \|k_J\| \leq 0 \quad \text{for all } k \in \bigcup_{m=1}^{J} K_m.$$
Since $y, z \to \cdot$, $k_J \to 0$. Thus letting $J \to \infty$ yields

$$\langle p, k \rangle \leq 0 \quad \text{for all } k \in \bigcup_{m=1}^{\infty} K_m = \{k \in \mathbb{R}^m : k_1 \leq 0\}.$$ 

Thus $p = (\lambda, 0, \ldots, 0)$ for some $\lambda \geq 0$ and, as seen in Example 3.10,

$$N_{\omega I}(\cdot) = N_{x}^{\nu}(\cdot) = N^{P}(x; \overline{I(r)}) = \{\lambda \epsilon_1 : \lambda \geq 0\} \quad \forall x, r \in \mathbb{R}^m.$$
6 MAIN RESULTS

Throughout this section assume the objective function $\phi = (\phi_1, \phi_2, \ldots, \phi_m)$ is a Lipschitz vector function on $\mathbb{R}^2b$ of rank $K_\phi$. $\Omega$ is a closed subset of $\mathbb{R}^n$. and $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ is a closed multifunction with nonempty values.

**Theorem 6.1 (Main Theorem).** Let $x$ be a local solution of $(M^\ast)$. Let $\{ (i, j, x_j) \}$ be a closed approximator of a preference $\prec$ at $x$, and let $\gamma := \phi(x(i), x(j))$. Further assume that there exists $\varepsilon > 0$ such that the following two properties hold:

(A3) $F$ is $C_{\mathbb{R}^n}$ measurable and there exists $k \in L^1$ and $\beta \geq 0$ such that, for a.e. $t \in [a, b]$.

$$F(t, x') - (\dot{x}(t) + \lambda^2 B) \subseteq F(t, x) + (k(t) + \beta \lambda^2)\|x' - x\|B$$

for all $\lambda \geq 0$ and $x', x \in x(t) + \varepsilon B$.

(A4) $h$ is upper semicontinuous on $\{(t, x) : t \in [a, b], x \in x(t) + \varepsilon B\}$ and there exists a constant $k_h$ such that

$$|h(t, x) - h(t, x')| \leq k_h|x - x'|$$

for all $x, x' \in x(t) + \varepsilon B$ and $t \in [a, b]$.

Then there exists an arc $p \in W^{1,1}([a, b], \mathbb{R}^n)$. $\lambda \geq 0$. $w \in N_{(x, \gamma)}$ with $\|w\| = 1$. a measure $\mu \in C^\infty(a, b)$. and a $\mu$-integrable function $\zeta : [a, b] \rightarrow \mathbb{R}^n$ such that
(i) \( \lambda + \|p\|_L^\infty + \|\mu\|_{TV} = 1. \)

(ii) \( \dot{p}(t) \in \text{co}\left\{ \eta : \langle \eta, p(t) + \int_{[a,t]} \zeta(s) \mu(ds) \rangle \in N \left( (x(t), \dot{x}(t)); \text{graph } F(t, \cdot) \right) \right\} \text{ a.e.:} \)

(iii) \( \left( p(a), -[p(b) + \int_{[a,t]} \zeta(s) \mu(ds)] \right) \in \lambda \partial \langle w, \alpha(\cdot, \cdot) \rangle(x(a), x(b)) + N((x(a), x(b)); \Omega); \)

(iv) \( \left( p(t) + \int_{[a,t]} \zeta(s) \mu(ds), \dot{x}(t) \right) \geq \left( p(t) + \int_{[a,t]} \zeta(s) \mu(ds), v \right) \)

for all \( v \in F(t, x(t)) \text{ a.e.: and} \)

(v) \( \zeta(t) \in \partial_{\geq} h(t, x(t)) \text{ a.e.} \)

Note that the Lipschitz condition on \( F(t, \cdot) \) is called integrable sub-Lipschitz and was introduced by Loewen and Rockafellar in [25]. Also note that the majority of the proof of this theorem relies on the proof of the Main Lemma 6.3 which is stated in terms of closed preferences. Once we establish a necessary condition for closed preferences, we complete the proof of the Main Theorem 6.1 by taking the limit of the results for the closed approximator. Before stating our Main Lemma we give one more definition. The following distance function is used in this definition:

\[
\rho_t(x, v) := \inf \{ \|v - y\| : y \in F(t, x) \}.
\]

\textbf{Definition 6.2.} Let \( G \) be the solution set of the system

\[ \dot{x} \in F(t, x(t)) \text{ a.e., } (x(a), x(b)) \in \Omega. \]

Let \( \bar{x} \in G. \) If there exist \( \alpha > 0 \) and \( q > 0 \) such that for all \( x \in \bar{x} + qB \), we have
\[ \text{dist}(\mathbb{x}; G) \leq \alpha \left[ \text{dist}(x(a), x(b)); \Omega \right] + \int_a^b \rho_i(x(t), \dot{x}(t))dt \]

then the system is called semi-normal at \( x \).

**Lemma 6.3 (Main Lemma).** Let \( x \) be a local solution of \((M_\varepsilon)\) and let \( \varepsilon := \varphi(x(a), x(b)) \). Suppose \( \varepsilon \) is closed and there exists \( \varepsilon > 0 \) such that \((A3)\) and \((A4)\) hold. Then there exists an arc \( p \in W^{1,1}([a, b], \mathbb{R}^n) \), a \( \lambda \geq 0 \), an \( w \in N_\varepsilon(\gamma) \) with \( \|w\| = 1 \), a measure \( \mu \in C^{\varepsilon}(a, b) \), and a \( \mu \)-integrable function \( \zeta : [a, b] \rightarrow \mathbb{R}^n \) such that

\begin{enumerate}
  \item \( \lambda + \|p\|_{L^\infty} + \|\mu\|_{TV} = 1 \);
  \item \( \hat{p}(t) \in \text{co}\{\eta : (\eta, p(t)) + \int_{[a, t]} \zeta(s)\mu(ds) \} \in N(x(t), \dot{x}(t)); \text{graph } F(t, \cdot)\} \) a.e.:
  \item \( \left( \hat{p}(a), \hat{p}(b) + \int_{[a, b]} \zeta(s)\mu(ds) \right) \geq \lambda \partial(w, \varphi(\dot{x}(a), x(b)) + N(x(a), x(b)); \Omega) \)
  \item \( \left( p(t) + \int_{[a, t]} \zeta(s)\mu(ds), \dot{x}(t) \right) \geq \left( p(t) + \int_{[a, t]} \zeta(s)\mu(ds), v \right) \) for all \( v \in F(t, x(t)) \) a.e.; and
  \item \( \zeta(t) \in \partial^2 h(t, x(t)) \mu \) a.e.
\end{enumerate}

**Proof of Main Lemma.** Without loss of generality assume \( \varepsilon \leq 1 \) and let \( \Delta < \min\{\Delta_0, \varepsilon\} \) where \( \Delta_0 \) is as in the definition of a solution of \((M_\varepsilon)\), and define the two sets:

\[ G := \left\{ x \in W^{1,1} : \dot{x} \in F(t, x(t)) \text{ a.e.. } (x|a), x(b) \in \Omega, \|x - \dot{x}\| \leq \Delta \right\} \text{, and} \]

\[ S := G \cap \left\{ x \in W^{1,1} : h(t, x(t)) \leq 0 \text{ for all } t \in [a, b] \right\}. \]

From here we break the proof into two cases.}

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Case 1: Assume the system composed of differential inclusion constraint (5.1) and endpoint constraint (5.2) is seminormal at \( x \) (without loss of generality assume \( q = \Delta \), where \( q \) is as in Definition 6.2). Note that by (A1) we can pick \( \gamma_k < \gamma \) such that \( \|\gamma_k - \gamma\| < \frac{1}{k} \). Let \( \Theta_k := \bar{I}(\gamma_k) \), and define \( g : W^{1,1} \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{\infty\} \) by

\[
g(x, \gamma) := \begin{cases} 
\|c(x(a), x(b)) - \gamma\| & \text{if } x \in x + \Delta B_x \\
+\infty & \text{otherwise}.
\end{cases}
\]

Then \( g \) is l.s.c., \( g(x, \gamma) \geq 0 \) everywhere, and \( g(x, \gamma) < \frac{1}{k} \). Thus if \( S \) is closed in \( W^{1,1} \) then we can apply Ekeland's Theorem 3.29 to the closed subspace \( S \times \Theta_k \subseteq W^{1,1} \times \mathbb{R}^m \).

Lemma 6.4. \( S \) is a closed subset of \( W^{1,1} \).

Proof. We will prove \( S \) is the intersection of two closed sets: \( G \) and \( H := \{ x \in W^{1,1} : h(t, x(t)) \leq 0 \text{ for all } t \in [a, b] \} \). Let \( x_n \rightarrow x \) in \( W^{1,1} \) with \( x_n \in S \Rightarrow G \cap H \) for each \( n \).

Note that by (A4) for each \( t \in [a, b] \).

\[
h(t, x(t)) - h(t, x_n(t)) \leq k_h \|x(t) - x_n(t)\| \leq k_h \|x(t) - x_n(t)\|_*.
\]

Thus for all \( t \in [a, b] \).

\[
h(t, x(t)) \leq k_h \|x(t) - x_n(t)\|_* + h(t, x_n(t)) \leq k_h \|x(t) - x_n(t)\|_*.
\]

Taking the limit yields \( x \in H \).

To show that \( x \in G \) note that by construction, \( \|x - x\|_* \leq \Delta < \varepsilon \). Also since \( \Omega \) is closed, \( (x(a), x(b)) \in \Omega \). Let

\[
Z := \{ t \in [a, b] : \dot{x}_n(t) \neq \dot{x}(t) \} \cup \left( \bigcup_{n=1}^{\infty} \{ t \in [a, b] : \dot{x}_n(t) \notin F(t, x_n(t)) \} \right).
\]

Then \( Z \) has zero measure. Use (A3) to pick \( N \) sufficiently large such that \( \dot{x}(t) \in
\[ \dot{x}(t) + \lambda B. \text{ Then by (A3), for all } n \text{ and for all } t \notin Z. \]

\[
\text{dist}(\dot{x}(t); F(t, x(t))) \leq \text{dist}(\dot{x}(t); F(t, x_n(t))) + (k(t) + \lambda \lambda) \|x(t) - x_n(t)\|
\]

\[
\leq \|\dot{x}(t) - \dot{x}_n(t)\| + (k(t) + \lambda \lambda) \|x(t) - x_n(t)\|.
\]

Hence \(\text{dist}(\dot{x}(t); F(t, x(t))) = 0\) for all \(t \notin Z\). Since \(F\) is a closed-valued multifunction, \(\dot{x}(t) \in F(t, x(t))\) for all \(t \notin Z\); so \(G\) is closed.

Continuing with the proof of the Main Lemma 6.3, by Ekeland's Theorem 3.29 there exists \((x_k, \theta_k) \in S \times \Theta_k\) such that

\[
g(x_k, \theta_k) \leq g(x, \gamma_k) < \frac{1}{k^2},
\]

\[
\|x_k - x\| + \|\theta_k - \gamma_k\| < \frac{1}{k}, \text{ and } \quad \|x_k - x\| + \|\theta - \theta_k\| < \frac{1}{k} \quad (6.1)
\]

\[
g(x_k, \theta_k) \leq g(x, \theta) + \frac{1}{k} \|x - x_k\| + \|\theta - \theta_k\| \quad (6.2)
\]

for all \((x, \theta) \in S \times \Theta_k\).

Fixing the appropriate element in (6.2) yields the following.

For all \(\theta \in \Theta_k\), \(g(x_k, \theta_k) \leq g(x, \theta) + \frac{1}{k} \|\theta - \theta_k\| \quad (6.3)\)

For all \(x \in S\), \(g(x_k, \theta_k) \leq g(x, \theta_k) + \frac{1}{k} \|x - x_k\| \quad (6.4)\)

By (6.3) and the sum rule, Proposition 3.26, we have

\[
0 \leq \partial g(x_k, \cdot)(\theta_k) + \frac{1}{k} B + N(\theta_k; \Theta_k)
\]

\[
= \partial \left( \|\phi(x_k(a), x_k(b)) - \cdot\| \right)(\theta_k) + \frac{1}{k} B + N(\theta_k; \Theta_k).
\]

By (A2) and the definition of a solution, \(\dot{x}, \phi(x_k(a), x_k(b)) \neq \theta_k\), so by Proposition 3.25

\[
w_k := -\partial \left( \|\phi(x_k(a), x_k(b)) - \cdot\| \right)(\theta_k) = \frac{\phi(x_k(a), x_k(b)) - \theta_k}{\|\phi(x_k(a), x_k(b)) - \theta_k\|} \in \frac{1}{k} B + N(\theta_k; \Theta_k).
\]
Since \( g(\cdot, \theta_k) \) is Lipschitz continuous on \( x + \Delta B \) and \( S \subseteq x + \Delta B \), by (6.4) and Theorem 3.30, for all \( x \in x + \Delta B \) such that \( h(t, x(t)) \leq 0 \), we have

\[
g(x_k, \theta_k) \leq g(x, \theta_k) + \frac{1}{k} \| x - x_k \|_* + (K_\phi + 1) \text{dist}(x; G).
\]

Using the seminormal assumption, \( x_k \) is a local \( W^{1,1} \) minimizer of \((FLS)\) with

\[
l(c, d) := ||e(c, d) - \theta_k|| + \frac{1}{k} ||r - x_k || + \alpha(K_\phi + 1) \text{dist}((c, d); \Omega) \text{ and}
\]

\[
L(t, x, v) := \frac{1}{k} ||v - x_k|| + \alpha(K_\phi + 1) \rho_t(x, v).
\]

Now by Theorem 3.32, Remark 3.14, and Proposition 3.15 there exist \( p_k \in W^{1,1} \), \( \lambda_k > 0 \), a measure \( \mu_k \in C_\text{w}^\infty(a, b) \), and a \( \mu \)-integrable function \( \zeta_k : [a, b] \to \mathbb{R}^n \) such that

\[
\lambda_k + \| p_k \|_{L^\infty} + \| \mu_k \|_{TL^1} = 1; \quad \lambda_k \geq \| \mu_k \|_{L^\infty} . \quad (6.5)
\]

\[
\mathcal{P}k(t) \in co\left\{ \eta \in \eta, p_k(t) + \int_{[a, t]} \zeta_k(s) \mu_k(ds) \right\} \lambda_k \partial L(t, x_k(t), \dot{x}_k(t)) a.e.: \quad (6.6)
\]

\[
\left\{ p_k(a), -|p_k(b) + \int_{[a, b]} \zeta_k(s) \mu_k(ds) | \right\} \in \lambda_k \partial l(x_k(a), x_k(b)): \quad (6.7)
\]

\[
\langle p_k(t) + \int_{[a, t]} \zeta_k(s) \mu_k(ds), \dot{x}_k(t) \rangle \\
\geq \langle p_k(t) + \int_{[a, t]} \zeta_k(s) \mu_k(ds), v \rangle - \lambda_k L(t, x_k(t), v) \quad \forall v \in \mathbb{R}^n \ a.e. . \quad (6.8)
\]

\[
= \langle p_k(t) + \int_{[a, t]} \zeta_k(s) \mu_k(ds), v \rangle - \frac{\lambda_k}{k} ||v - \dot{x}_k|| \text { for all } v \in F(t, x_k(t)) \ a.e.: \quad (6.9)
\]

\[
\zeta_k(t) \in \partial^2 h(t, x_k(t)) \mu_k \ a.e. .
\]

The following two lemmas, found in Section 7.2 of [39], will prove helpful in simplifying some of the above expressions.
Lemma 6.5. Take a multifunction $\Gamma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and a point $(x, v) \in \text{graph} \ \Gamma$. Define 

$$
\rho_\Gamma(x, v) := \inf \{ \| v - y \| : y \in \Gamma(x) \},
$$

and assume that $\Gamma$ has as values closed nonempty sets. Also assume that there exist $\varepsilon > 0$, $r > 0$, $k > 0$, and $K > 0$ such that for all $x', x \in x + \varepsilon B$ we have

$$
\Gamma(x') \cap (v + rB) \subseteq \Gamma(x) \cap k\|x' - x\|B \quad \text{and}
$$

$$
\rho_\Gamma(x, v) \leq K\|x - x\|.
$$

Then:

(a) for each $x \in \mathbb{R}^n$, $\rho_\Gamma(x, \cdot)$ is Lipschitz continuous with Lipschitz constant $1$. For each $v \in v + (r/3)B$, $\rho_\Gamma(\cdot, v)$ is Lipschitz continuous on $x + \min\{\varepsilon, r/(3K)\}B$ with Lipschitz constant $k$;

(b) fix elements $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$ and $(w, p) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$
\|x - x\| < \min\{\varepsilon, r/(3K)\}, \quad \|v - v\| < r/3, \quad \text{and} \quad (w, p) \in \partial \rho_\Gamma(x, v).
$$

Then

$$
\|w\| \leq k \quad \text{and} \quad \|p\| \leq 1.
$$

Furthermore

$$
v \in \Gamma(x) \quad \text{implies} \quad (w, p) \in \mathcal{N}((x, v); \text{graph} \ \Gamma)
$$

and

$$
v \notin \Gamma(x) \quad \text{implies} \quad p = \frac{v - u}{\|v - u\|}
$$

for some $u \in \Gamma(x)$ such that $\|v - u\| = \min\{\|v - y\| : y \in \Gamma(x)\}$.
Lemma 6.6. Take a multifunction $\Gamma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and a point $(x, v) \in \text{graph } \Gamma$. Define $p_{\Gamma}(x, v) := \inf \{ \| v - y \| : y \in \Gamma(x) \}$, and assume that $\Gamma$ has as values closed nonempty sets. Also assume that there exist $\varepsilon > 0$, $\beta > 0$, and $k > 0$ such that for all $x', x \in x + \varepsilon B$ and for all $N \geq 0$ we have

$$\Gamma(x') \cap (v + N B) \subseteq \Gamma(x) + (k + \beta N)\| x' - x \| B.$$  

Then, for any $v \in \mathbb{R}^n$, $p_{\Gamma}(\cdot, v)$ is Lipschitz continuous on $x + \varepsilon B$ with Lipschitz constant $k(1 + 3\varepsilon) + 2\beta(\| v - v' \|)$. 

Remark 6.7. Note that by (A3) we obtain (6.10) for almost all $t \in [a, b]$ at $(x(t), \dot{x}(t))$ with $\Gamma_t = F(t, \cdot)$. $r = N$ and $k = k(t) + 3N$. Also as in the proof of Lemma 6.4, for $p_t := p_{\Gamma}$, we have

$$p_t(x, \dot{x}(t)) = \text{dist}(\dot{x}(t); F(t, x)) \leq \text{dist}(\dot{x}(t); F(t, x(t))) + (k(t) + 3N)\| x(t) - x \| = (k(t) + 3N)\| \dot{x}(t) - x \|.$$ 

Thus (A3) also yields (6.11) with $K = k(t) + 3N$.

Remark 6.8. By the above remark, Lemma 6.6, and part a) of Lemma 6.5, for almost all $t \in [a, b]$ and for all $N \geq 0$, $p_t(x, v)$ is Lipschitz continuous on $(x(t) + \min \{ \varepsilon, N/(3(k(t) + 3N)) \} B) \times \mathbb{R}^n$ with rank less than or equal to $\sqrt{2}(k(t)(1 + 3\varepsilon) + 2\beta + 1)$. 

Taking the limit as $N \rightarrow \infty$ yields, for almost all $t \in [a, b]$, $p_t(x, v)$ is Lipschitz continuous on $(x(t) + \varepsilon B) \times \mathbb{R}^n$ with rank less than or equal to $\sqrt{2}(k(t)(1 + 3\varepsilon) + 2\beta + 1) \leq \sqrt{2}(k(t)(1 + \beta) + 2\beta + 1)$.

Remark 6.9. Similar to the above remark, we can consider the results of part b) of Lemma 6.5 as $N \rightarrow \infty$. For almost all $t \in [a, b]$, if $(x, v) \in (x(t) + \varepsilon B) \times \mathbb{R}^n$ and
where $w, p \in \partial \rho_t(x, v)$, then

$$v \in F(t, x) \implies (w, p) \in \mathcal{N}((x, v); \text{graph } F(t, \cdot))$$

and

$$v \notin F(t, x) \implies p = \frac{v - u}{\|v - u\|}$$

for some $u \in F(t, x)$ such that $\|v - u\| = \min \{\|v - y\| : y \in F(t, x)\}$.

Continuing with the main proof, by the sum rule, Proposition 3.26.

$$\partial L(t, x_k(t), \dot{x}_k(t)) \subset \{0\} \times \partial \left( \frac{1}{k} \|\cdot - \dot{x}_k(t)\| \right)(\dot{x}_k(t)) + \partial \left( \alpha(K_\alpha + 1) \rho_t(x_k(t), \dot{x}_k(t)) \right)$$

and by Proposition 3.22 and Remark 3.14, respectively.

$$\partial L(t, x_k(t), \dot{x}_k(t)) \subset \{0\} \times \frac{1}{k} \partial \|\cdot - \dot{x}_k(t)\| (\dot{x}_k(t)) + \alpha(K_\alpha + 1) \rho_t(x_k(t), \dot{x}_k(t))$$

Further by Remark 6.9, for almost all $t \in [a, b]$.

$$\partial L(t, x_k(t), \dot{x}_k(t)) \subset \{0\} \times \frac{1}{k} \partial \|\cdot - \dot{x}_k(t)\| (\dot{x}_k(t)) + \alpha(K_\alpha + 1) \rho_t(x_k(t), \dot{x}_k(t))$$

Now (6.5) implies that $0 < \lambda_k \leq 1$, so

$$\lambda_k \partial L(t, x_k(t), \dot{x}_k(t)) \subset \{0\} \times \frac{1}{k} B + \mathcal{N}((x_k(t), \dot{x}_k(t)); \text{graph } F(t, \cdot))$$

and (6.6) becomes

$$\dot{p}_k(t) \in \text{co} \left\{ \eta : (\eta, p_k(t)) + \int_{[a, t]} \zeta_k(s) \mu_k(ds) \right\}$$

$$\quad \subset \{0\} \times \frac{1}{k} B + \alpha(K_\alpha + 1) \rho_t(x_k(t), \dot{x}_k(t))$$

and

$$\dot{p}_k(t) \in \text{co} \left\{ \eta : (\eta, p_k(t)) + \int_{[a, t]} \zeta_k(s) \mu_k(ds) \right\}$$

$$\quad \subset \{0\} \times \frac{1}{k} B + \mathcal{N}((x_k(t), \dot{x}_k(t)); \text{graph } F(t, \cdot))$$
Similarly by the sum rule, Proposition 3.26, and Proposition 3.28
\[ \lambda_k \partial l(x_k(a), x_k(b)) \]
\[ \subseteq \partial l(x_k(a), x_k(b)) \]
\[ \subseteq \partial \| \phi \cdot \cdot \cdot - \theta_k \| (x_k(a), x_k(b)) + \frac{1}{k} B \times \{0\} + N((x_k(a), x_k(b)); \Omega), \]
and then by the chain rule, Proposition 3.27, (6.7) becomes
\[ \left( p_k(a), -\left[ p_k(b) + \int_{[t_1, t]} \xi_k(s) \mu_k(ds) \right] \right) \]
\[ \equiv \lambda_k \partial (w_k, \phi \cdot \cdot \cdot ) (x_k(a), x_k(b)) + \frac{1}{k} B \times \{0\} + N((x_k(a), x_k(b)); \Omega). \] (6.15)

What remains is to show convergence of the above sequences, which will be accomplished with the aid of Proposition 4.5 and Proposition 4.8, respectively. First consider the multifunction
\[ G_1(t, x) := \begin{cases} \partial^2 h(t, x): & \text{if } x \in x(t) + \varepsilon B \\ \{0\} & \text{otherwise}. \end{cases} \]
Clearly \( G_1 \) is convex for all \((t, x)\), and since the function \( x \to h(t, x) \) is Lipschitz continuous on \( x(t) + \varepsilon B \) for all \( t \in [a, b] \), if \( \gamma \in G_1(t, x) \) then \( \| \gamma \| \leq k_h \). In view of Remark 4.3, to show \( G_1(t, x) \) is compact one need only show \( \{ \xi: \exists (t_i, x_i) \xrightarrow{h} (t, x) \text{ with } h(t_i, x_i) > 0 \text{ for all } i \text{ and } \sup_{t_i} h(t_i, x_i) \to \xi \} \) is closed for all \( x \in x(t) + \varepsilon B \) and this can be accomplished with a diagonalization argument. Thus \( G_1(t, x) \) is a compact convex multifunction whose images are contained in \( k_h B \), and by construction, for all \( k \), \( \xi_k \in G_1(t, x_k(t)) \) \( \mu - a.e. \) and \( \{x_k\} \) is a sequence of arcs which converge uniformly to \( x \). Now by equation (6.5), \( \| \mu_k \| \leq 1 \), so (by re-enumerating if necessary) we may assume that \( \mu_k \) converges weak* to \( \mu \). Thus by Proposition 4.5, there exists a measurable function \( \xi \), which is \( \mu \)-integrable, such that \( \xi \in G_1(t, x(t)) \) \( \mu - a.e. \) and some subsequence of \( \xi_kd\mu_k \) converges weak* to \( \xi d\mu \). Without loss of generality we will assume
\[ \xi_k d\mu_k \text{ converges weak* to } \xi d\mu. \] (6.16)
Then \( y_k(t) := \int_{[t,t]} \zeta_k(s) \mu_k(ds) - \int_{[a,b]} \zeta(s) \mu(ds) \) is a sequence of continuous functions which converge to zero. Next let \( c = \alpha(K_o + 1) \) and define
\[
G_2(t, p, x, v, s) := \begin{cases} 
\text{co} \left\{ \eta : (\eta, p(t) + \int_{[a,b]} \zeta(s) \mu(ds)) \in \partial \mu(x, v) + \{0\} \times sB \right\} & \text{if } x \in x(t) + sB \\
\{0\} & \text{otherwise.} 
\end{cases}
\]

By Remark 6.8, \( \rho_i(x, v) \) is Lipschitz continuous on \((x(t) + sB) \times \mathbb{R}^n\) of rank \( m(t) := \sqrt{2}(k(t)(1 + J) + 2J + 1) \in L^1[a,b] \), so for all \( \gamma \in G_2(t, p, x, v, s) \), \( \|\gamma\| \leq c m(t) \). Also since \( \rho_i(x, v) \) is Lipschitz continuous, by Proposition 4.4, \((p, x, v, s) \rightarrow G_2(t, p, x, v, s) \) is a closed multifunction which is u.s.c. on the interior of \( \mathbb{R}^n \times (x(t) + sB) \times \mathbb{R}^n \times [0, s] \).

Equation (6.5) shows that the sequence \( \{p_k(a)\} \) is bounded, and equation (6.13) shows that for all \( k \), \( p_k(t) \in (G_2(t, p_k(t) + y_k(t), x_k(t), \dot{x}_k(t), s_k(t)) \). Thus by Proposition 4.8 and calculations similar to that of equation (6.14), there exists a subsequence of \( \{p_k\} \) which converges uniformly to an arc \( p \) which satisfies
\[
\hat{p}(t) \in \text{co} \left\{ \eta : (\eta, p(t) + \int_{[a,b]} \zeta(s) \mu(ds)) \in N \left( (x(t), \dot{x}(t)); \text{graph } F(t, \cdot) \right) \right\} \quad \text{a.e.} \quad (6.17)
\]

Again we will re-enumerate if need be.

Our limiting process has thus far has given us property (i), (ii), and (v). To obtain (iii) first note that by construction \( w_k \in N(\theta_k; \Theta_k) + \frac{1}{k}B \) and \( \|w_k\| = 1 \) for each \( k \). Since \( \theta_k \rightarrow \gamma \), \( \gamma_k \rightarrow \gamma \) and \( \Theta_k := \overline{(\{\gamma_k\})} \), we may assume \( w_k \rightarrow w \in N_\omega(\gamma) \) with \( \|w\| = 1 \). Next note that by equation (6.16), \( (p_k(a), -[p_k(b) + \int_{[a,b]} \zeta_k(s) \mu_k(ds)]) \rightarrow (p(a), -[p(b) + \int_{[a,b]} \zeta(s) \mu(ds)]) \). By the sum rule (Proposition 3.26) and since \( \langle w_k - w, o(\cdot, \cdot) \rangle \) is Lipschitz continuous of rank \( \|w_k - w\| K_o \).
Then by equation 6.15, taking the limit of the above with the aid of Remark 4.21 yields (iii).

Finally, to show (iv), define the following set:

\[ Z_0 := \{ t \in [a, b] : \dot{x}_k(t) \neq \dot{x}(t) \} \cup \left( \bigcup_{n=1}^{\infty} \{ t \in [a, b] : \dot{x}_k(t) \text{ does not satisfy (6.8)} \} \right). \]

and choose \( N \) sufficiently large such that \( \dot{x}_k(t) \in \dot{x}(t) + NB \) for \( k \) sufficiently large. Note that by construction \( Z_0 \) is a set of measure zero. Recall (6.8) shows that \( x_k \) is a minimizer of \( \langle A_k, v \rangle + \lambda_k \| v - \dot{x}_k(t) \| \) over \( v \in F(t, x_k(t)) \) with \( A_k = -\left( p_k(t) + \int_{[t,t]} \zeta_k(s) \mu_k(ds) \right) \).

Then by Theorem 3.30 for all \( v \in \mathbb{R}^n \), \( t \in Z_0 \) and for \( k \) sufficiently large,

\[
\langle A_k, \dot{x}_k(t) \rangle \leq \langle A_k, v \rangle + \lambda_k \| v - \dot{x}_k(t) \| + (\| A_k \| + 1) \text{dist}(v;F(t, x_k(t)))
\leq \langle A_k, v \rangle + \lambda_k \| v - \dot{x}(t) \| + \lambda_k \| \dot{x}(t) - \dot{x}_k(t) \|
+ (\| A_k \| + 1) \left[ \text{dist}(v;F(t, \dot{x}(t))) + (k(t) + 3N) \| \dot{x}(t) - x_k(t) \| \right].
\]

(6.19)
Let \( A = -\left(p(t) + \int_{(s,t)} \gamma(s) \mu(ds)\right) \). Then \( A_k \to A \), so for sufficiently large \( k \), for all \( t \in Z_0 \), and for a given \( v \in \mathbb{R}^n \).

\[
\langle A_k, \dot{x}_k(t) \rangle \leq \|A_k - A\| \|v\| + \langle A, v \rangle + \frac{\lambda_k}{k} \|v - \dot{x}(t)\| + \frac{\lambda_k}{k} \|\dot{x}(t) - \dot{x}_k(t)\| + (\|A_k\| + 2) \left[ dist(v; F(t, x(t))) + (k(t) + \lambda_N) \|x(t) - x_k(t)\| \right].
\]

Taking the limit as \( k \to \infty \) yields for each given \( v \in \mathbb{R}^n \) and for all \( t \in Z_0 \).

\[
\langle A, \dot{x}(t) \rangle \leq \langle A, v \rangle + (\|A\| + 1) dist(v; F(t, x(t))).
\]

Then \((iv)\) is achieved with Theorem 3.30.

**Case 2:** Assume the system composed of (5.1) and (5.2) is not seminormal at \( x \). From Definition 6.2 of seminormal let \( \alpha = k \), and let \( \gamma = \frac{1}{k} \), and define

\[
q(x) := \text{dist}(\{(x(a), x(b)): \Omega\}) + \int_a^b \rho_i(x(t), \dot{x}(t)) dt.
\]

Then there exists \( x_k \in W^{1,1} \) such that \( \|x_k - x\|_* \leq \frac{1}{k} \) and

\[
dist(x_k; G) > k(q(x_k)). \tag{6.20}
\]

Note that by Remark 6.8, \( q \) is Lipschitz continuous on \( \bar{x} + \epsilon B_* \). Let \( \epsilon_k = \sqrt{q(x_k)} \), \( \lambda_k = \min\{\epsilon_k, k\epsilon_k^2\} \), and \( s_k = \frac{\epsilon_k^2}{\lambda_k} \). Since \( x_k \to \bar{x} \) in \( W^{1,1} \) and \( x \in G \), \( \epsilon_k, s_k \to 0 \). Now by construction \( q(x) \geq 0 \) for all \( x \in W^{1,1} \), so

\[
q(x_k) \leq \inf_{x \in W^{1,1}} q(x) + \epsilon_k^2.
\]

Therefore by Ekeland's Theorem 3.29, there exists \( z_k \in \bar{x} + \epsilon B_* \) such that

\[
q(z_k) \leq q(x_k) < \epsilon_k^2.
\]

\[
\|z_k - x_k\|_* < \lambda_k, \text{ and } \tag{6.21}
\]

\[
q(z_k) \leq q(z_k) + s_k \|z - z_k\|_* \quad \text{for all } z \in \bar{x} + \epsilon B_* \tag{6.22}
\]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Note that by (6.21) $z_k \to x$ in $W^{1,1}$ and $z_k \notin G$, since

$$\kappa \zeta_k^2 < \text{dist}(x_k; G)$$

$$\leq \|z_k - x_k\|_* + \text{dist}(z_k; G)$$

$$\leq \lambda_k + \text{dist}(z_k; G)$$

$$\leq \kappa \zeta_k^2 + \text{dist}(z_k; G).$$

Consider the following.

$$l_k(c, d) := \text{dist}((c, d); \Omega) + s_k\|c - z_k(a)\|$$

$$L_k(t, x, v) := \rho_t(x, v) + s_k\|v - \dot{z}_k(t)\|.$$

Since $x_k \to x$ in $W^{1,1}$, for $k$ sufficiently large $z_k$ is an interior point of $x + \varepsilon B$, so $z_k$ is a $W^{1,1}$ local minimizer for $(FL)$ with $l = l_k$ and $L = L_k$. Thus by Theorem 3.31 and the sum rule (Proposition 3.26), there exists $p_k \in W^{1,1}$ such that

$$\hat{p}_k(t) \in \text{co}\{\eta : (\eta, p_k(t)) \in \partial \rho_t(z_k(t), \dot{z}_k(t)) + \{0\} \times s_k B\} \quad \text{a.e.}$$

$$\begin{align*}
(p_k(a), -p_k(b)) &\in \partial \text{dist}((z_k(a), z_k(b)); \Omega) + s_k B \times \{0\}; \quad \text{and} \\
(p_k(t), \dot{z}_k(t)) - \rho_t(z_k(t), \dot{z}_k(t)) &\geq (p(t), v) - \rho_t(z_k(t), v) - s_k\|v - \dot{z}_k(t)\| \quad \forall v \in \mathbb{R}^n \ \text{a.e.} \\
&\geq (p(t), v) - \rho_t(z_k(t), v) - s_k\|v - \dot{z}(t)\| - s_k\|\dot{z}(t) - \dot{z}_k(t)\| \quad \forall v \in \mathbb{R}^n \ \text{a.e.}
\end{align*}$$

Again by Remark 6.8, for almost all $t \in [a, b]$, $\rho_t(\cdot, \cdot)$ is Lipschitz continuous on $(\dot{z}(t) + \varepsilon B) \times \mathbb{R}^n$, so similar to the proof of case 1, define

$$G_3(t, p, x, v, s) := \begin{cases} 
\text{co}\{\eta : (\eta, p) \in \partial \rho_t(x, v) + \{0\} \times s B\} : & \text{if } x \in \dot{x}(t) + \varepsilon B \\
\{0\} & \text{otherwise.}
\end{cases}$$
Then using the same argument we used on \( G_2 \), we see that \( G_3 \) is a multifunction such that for all \( \gamma \in G_3(t, p, x, v, s) \), \( \|\gamma\| \leq m(t) \). Also \((p, x, v, s) \mapsto G_3(t, p, x, v, s)\) is a closed multifunction which is u.s.c. on the interior of \( R^n \times (x(t) + \varepsilon B) \times R^n \times [0, \varepsilon] \). \( \hat{p}_k(t) \in G_3(t, p_k(t), x_k(t), \hat{x}_k(t), s_k) \) a.e. for all \( k \), and \( G_3 \) satisfies the rest of the hypothesis of Proposition 4.8. Thus there exists a subsequence of \( \{p_k\} \) which converges uniformly to an arc \( p \) which satisfies

\[
\dot{p}(t) \in \operatorname{co} \left\{ \eta : (\eta, p(t)) \in \partial p(t) \left( (x(t), \dot{x}(t)) \right) \right\} \quad \text{a.e.}
\]

\[
\subset \operatorname{co} \left\{ \eta : (\eta, p(t)) \in N \left( (x(t), \dot{x}(t)) ; \text{graph } F(t, \cdot) \right) \right\} \quad \text{a.e.} \tag{6.26}
\]

Note that the last inclusion is due to Remark 6.9, and again we will re-enumerate if need be.

By Remark 3.11 there exists \( w \in N_+ (\gamma) \) with \( \|w\| = 1 \). Let \( \lambda = 0, \mu = 0 \), and \( \zeta(t) \equiv 0 \). Then \((v)\) of the Main Lemma is trivially true. Using an argument similar to that of equation 6.18 in case 1 transforms (6.24) into \((iii)\), and using an argument similar to that of equation 6.19 transforms (6.25) into \((iv)\).

To show \((i)\) recall that \( z_k \notin G \) for each \( k \). Thus for each \( k \) either

\[
|z_k(a), z_k(b)| \notin \Omega \quad \text{or} \tag{6.27}
\]

\[
|z_k(t)| \notin F(t, z_k(t)) \text{ on some set of positive measure.} \tag{6.28}
\]

Therefore if (6.28) is true for infinitely many \( k \), then by Remark 6.9, \( \|p_k(t)\| = 1 \) on a set of positive measure. Thus \( \|p_k\|_{L^\infty} = 1 \) and since \( p_k(t) \) converges uniformly to \( p(t) \) on \([a, b], \|p\|_{L^\infty} = 1 \). Hence \((i)\) is satisfied.

Finally if (6.27) is satisfied for infinitely many \( k \), then for each of these \( k \) there exists \( a_k \in s_k B \) such that \( \|(p_k(a) + a_k, p_k(b))\| = 1 \). Thus \( \|p(a), -p(b)\| = 1 \) which means that either \( \|p(a)\| > 0 \) or \( \|p(b)\| > 0 \). Either way since \( p \) is continuous on \([a, b]\) there is an open interval in \([a, b]\) such that \( \|p(t)\| > 0 \). Thus \( \|p\|_{L^\infty} > 0 \). Let \( \hat{p} := \frac{p(t)}{\|p(t)\|} \).
Then \( \hat{p} \) along with previously defined \( \lambda, \upsilon, \mu, \) and \( \zeta \) satisfy (i) through (v) of the Main Lemma.

**Remark 6.10.** In both cases for \( c_2 := \max\{c, 1\} \), we have

\[
\hat{p}(t) \in \text{co}\left\{ \eta : (\eta, p(t) + \int_{[x, t]} \zeta(s) \mu(ds)) \in c_2 \partial \rho_t(x_J(t), \dot{x}_J(t)) \right\} \text{ a.e. .}
\]

so in addition to our listed results in the Main Lemma we also have

\[
\|\hat{p}(t)\| \leq \int_{[a, b]} c_2 \eta(t) \text{ a.e. .}
\]

The next section is devoted to the limiting process needed to prove the Main Theorem 6.1

**Proof of Main Theorem.** Assume each \( J \) is sufficiently large that \( \|x - x_J\|_* < \frac{\varepsilon}{2} \).

Then for each \( J \) (A3) and (A4) hold at \( x_J \) with \( \varepsilon = \frac{\varepsilon}{2} \), so by the Main Lemma 6.3 for each \( J \) there exists arc \( p_J \in W^{1,1}([a, b], \mathbb{R}^n) \) and \( w_J \in N_{x_J}(\gamma_J) \) with \( \|w_J\| = 1 \) and \( l_J(\gamma_J) \) corresponding to the level set created by \( \phi \) at \( \gamma_J := \phi(x_J(a), x_J(b)) \).

As before, since \( \|w_J\| = 1 \) we may assume \( w_J \to w \in \mathbb{R}^n \), and since \( \phi \) is continuous \( w \in N_{x_J}(\gamma) \).

Also by the Main Lemma 6.3 and by Remark 6.10 for each \( J \), there exists \( \lambda_J \geq 0 \), a measure \( \mu_J \in C^\infty([a, b]) \), and a \( \mu_J \)-integrable function \( \zeta_J : [a, b] \to \mathbb{R}^n \) such that

(i) \( \lambda_J + \|p_J\|_{L^\infty} + \|\mu_J\|_{T.A.} = 1 \);

(ii) \( \hat{p}(t) \in \text{co}\left\{ \eta : (\eta, p_J(t) + \int_{[x, t]} \zeta_J(s) \mu_J(ds)) \in c_2 \partial \rho_t(x_J(t), \dot{x}_J(t)) \right\} \text{ a.e. :}

(iii) \[
\left( p_J(a), -\left[p_J(b) + \int_{[a, b]} \zeta_J(s) \mu_J(ds)\right] \right) \\
\in \lambda_J \partial \omega((\omega, \dot{\omega}) (x_J(a), x_J(b)) + N((x_J(a), x_J(b)); \Omega);
\]
(iv) \[ \langle p_J(t) + \int_{[a,t]} \zeta_J(s) \mu_J(ds), \dot{x}_J(t) \rangle \geq \langle p_J(t) + \int_{[a,t]} \zeta_J(s) \mu_J(ds), v \rangle \]
for all \( v \in F(t, x_J(t)) \) a.e.; and

(v) \[ \zeta_J(t) \in \partial^2 h(t, x_J(t)) \text{ } \mu_J \text{ a.e.} \]

As in the proof of the Main Lemma, we utilize the multifunction \( G_1 \) and Proposition 4.5 to obtain (v) and then use multifunction \( G_2 \) with \( c \) replaced by \( c_2 \). We also utilize Proposition 4.8 and arguments found in equation (6.18) (with \( w \in X_{\zeta_J} \) from above) and equation (6.19) to obtain (i), (ii), (iii) and (iv).
Consider the following cost function, $\phi$, and utility function, $u$, defined on $\mathbb{R}^2$:

$$
\phi(x, y) = (x, y)
$$

$$
u(x, y) = \begin{cases} 
x + y & : y > -x \\
x^2 + y^2 & : y = -x \\
1 & : \text{otherwise}
\end{cases}
$$

and consider the preference, $\prec$, defined by the utility function $u$. In this example we will limit our solutions of ($M_\prec$) to the constant functions by letting $F(t, x) = 0$ for all $(t, x) \in \mathbb{R} \times \mathbb{R}^2$. These will be the only constraints used in this example, so let $\Omega = \mathbb{R}^4$ and $h(t, x) = -1$ for all $(t, x) \in \mathbb{R} \times \mathbb{R}^2$. Clearly $F$ satisfies ($A3$) and $h$ satisfies ($A4$). Notice that the level sets for points away from the origin are as illustrated below.

Figure 3. $l(x, y)$ for $c := u(x, y) \geq 1$. 

52
Next consider the following two families of cones:

\[ K^B_j := \{ (x, y) \in \mathbb{R}^2 : y \geq \frac{J - 1}{J} x \} \cap \{ (x, y) \in \mathbb{R}^2 : y \geq \frac{J + 1}{J} x \} \text{ and} \]

\[ K^L_j := \{ (x, y) \in \mathbb{R}^2 : y \leq \frac{J - 1}{J} x \} \cap \{ (x, y) \in \mathbb{R}^2 : y \leq \frac{J + 1}{J} x \}. \]

Let \( \prec_j^B \) be the generalized weak Pareto preference defined by the cone \( K^B_j \) and \( \prec_j^L \) be the generalized weak Pareto preference defined by the cone \( K^L_j \). Clearly \( K^B_j \) and \( K^L_j \) are closed pointed convex cones for each \( J \), so \( \prec_j^B \) and \( \prec_j^L \) are closed preferences. It is also evident from the two illustrations that if \( \gamma \neq (0, 0) \) and if \( \gamma \notin \overline{I(\gamma)} \), then Equation (5.4) is satisfied at \( \gamma \) in terms of either \( \prec_j^B \) or \( \prec_j^L \). Thus for \( \gamma \neq (0, 0) \), \( \prec \) is closable at \( \gamma \).

Then either the pair \( \{ (\prec_j^B, \gamma) \} \) or \( \{ (\prec_j^L, \gamma) \} \) is a closed approximator of \( \prec \) at \( \gamma \neq (0, 0) \), and this example satisfies the hypotheses of the Main Theorem 6.1.

Suppose \( x(t) \) is a solution of this Mayer problem. By part (v) of the results of the Main Theorem 6.1, \( \zeta(t) \in \partial^2_x h(t, x(t)) \) \( \mu \) a.e. where \( \partial^2_x h(t, x) \) is defined as follows:

\[
\partial^2_x h(t, x) := \text{co}\{ \xi : \exists (t_i, x_i) \xrightarrow{\Delta} (t, x) \text{ with } h(t_i, x_i) > 0 \}
\]

for all \( i \) and \( \nabla_x h(t_i, x_i) \rightarrow \xi \).
Since $h(t, x) = -1$, $\mu$ must be the zero measure. Thus the Main Theorem 6.1 guarantees the existence of an arc $p \in W^{1,1}([a, b], \mathbb{R}^n)$. $\lambda \geq 0$, and $w \in \mathcal{N}_{\llangle \gamma \rrangle}(\gamma)$ with $\|w\| = 1$ and $\gamma := \mathcal{O}(x(a), x(b)) = (x(a), x(b))$ such that

(i) $\lambda + \|p\|_{L^\infty} = 1$:

(ii) \( \dot{p}(t) \in \text{co}\left\{ \eta : (\eta, p(t)) \in \mathcal{N}\left((x(t), \dot{x}(t)); \text{graph } F(t, \cdot)\right) \right\} \text{ a.e.} \)

\[ = \text{co}\left\{ \eta : (\eta, p(t)) \in \mathcal{N}\left((x(t), \dot{x}(t)); \{(x, y) : y = 0\}\right) \right\} \text{ a.e.} \]

\[ = \{0\} \text{ a.e.; and} \]

(iii) \((p(a), -p(b)) \in \lambda \mathcal{I}(w, \mathcal{O}(\cdot, \cdot))(x(a), x(b)) + \mathcal{N}(x(a), x(b); \Omega)\)

\[ = \lambda w. \]

By (ii), $\dot{p}(t) = 0$ a.e., so $p(t) = k$ a.e. for some constant $k$. By (i), $\lambda + |k| = 1$, so then by (iii) $\lambda \neq 0$ and $w = \left(\frac{k}{\lambda}, \frac{1}{\lambda}\right)$. Next note that, by construction, if $\llangle \chi_r \rrangle$ was used to approximate $<$ at $\gamma$, then $\mathcal{N}_{\llangle \gamma \rrangle}(\gamma) = \{(a, a) : a \geq 0\}$; and if $\llangle \chi_f \rrangle$ was used to approximate $<$ at $\gamma$, then $\mathcal{N}_{\llangle \gamma \rrangle}(\gamma) = \{(-a, -a) : a \geq 0\}$. Either way $w = \left(\frac{k}{\lambda}, \frac{1}{\lambda}\right) \not\in \mathcal{N}_{\llangle \gamma \rrangle}(\gamma)$ for $\gamma \neq (0, 0)$. Thus the only possible minimizer for this problem is $x(t) \equiv 0$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
REFERENCES


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.