Chapter 2

Measurable Multifunctions and Differential Inclusions

You cannot argue with someone who denies the first principles.

– Anonymous Medieval Proposition

2.1 Introduction

Differential inclusions

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

(2.1)

feature prominently in modern treatments of Optimal Control. This has come about for several reasons. One is that Condition (2.1), summarizing constraints on allowable velocities, provides a convenient framework for stating hypotheses under which optimal control problems have solutions and optimality conditions may be derived. Another is that, even when we choose not to formulate an optimal control problem in terms of a differential inclusion, in cases when the data are nonsmooth, often the very statement of optimality conditions makes reference to differential inclusions. It is convenient then at this stage to highlight important properties of multifunctions and differential inclusions of particular relevance in Optimal Control.

Section 2.3 deals primarily with measurability issues. The material is based on a by now standard definition of a measurable set-valued function, which is a natural analogue of the familiar concept of a measurable point-valued function. Criteria are given for a specified set-valued function to be measurable. These are used to establish that measurability is preserved under various operations on measurable multifunctions (composition, taking limits, etc.) frequently encountered in applications.

A key consideration is whether a set-valued function \( \Gamma(t) \) has a measurable selection \( \gamma(t) \). This means that \( \gamma \) is a Lebesgue measurable function satisfying

\[ \gamma(t) \in \Gamma(t) \quad \text{a.e. } t \in I. \]

The concept of a measurable selection has many uses, not least to define solutions to differential inclusions: in the case when \( I = [a, b] \), a solution \( x \) to (2.1) is an absolutely continuous function satisfying

\[ x(t) = x(a) + \int_a^t \gamma(s)ds, \]
where \( \gamma(t) \) is some measurable selection of \( F(t, x(t)) \).

Sections 2.4 and 2.5 concern solutions to the differential inclusion (2.1). The first important result is an existence theorem (the Generalized Filippov Existence Theorem), which gives conditions under which a solution (called an \( F \)-trajectory) may be found "near" an arc that approximately satisfies the differential inclusion. Of comparable significance is the Compactness of Trajectories Theorem which describes the closure properties of the set of \( F \)-trajectories. These results are used extensively in future chapters.

In Sections 2.6 through 2.8, we examine questions of when, and in what sense, Optimal Control problems have minimizers. The Compactness of Trajectories Theorem leads directly to simple criteria for existence of minimizers, for Optimal Control problems formulated in terms of the differential inclusion (2.1), when \( F \) takes as values convex sets. If \( F \) fails to be convex-valued, there may be no minimizers in a traditional sense. But existence of minimizers can still be guaranteed, if the domain of the optimal control problem is enlarged to include additional elements. This is the theme of relaxation, which is taken up in Section 2.7. Finally, in Section 2.8, broad criteria are established for existence of minimizers to the Generalized Bolza Problem (GBP). (GBP) subsumes a large class of Optimal Control problems, and is a natural framework for studying existence of minimizers under hypotheses of great generality.

### 2.2 Convergence of Sets

Take a sequence of sets \( \{A_i\} \) in \( \mathbb{R}^n \). There are number of ways of defining limit sets. For our purposes, "Kuratowski sense" limit operations are the most useful. The set

\[
\lim \inf_{i \to \infty} A_i
\]

(the Kuratowski \( \lim \inf \)) comprises all points \( x \in \mathbb{R}^n \) satisfying the condition: there exists a sequence \( x_i \to x \) such that \( x_i \in A_i \) for all \( i \).

The set

\[
\lim \sup_{i \to \infty} A_i
\]

(the Kuratowski \( \lim \sup \)) comprises all points \( x \in \mathbb{R}^n \) satisfying the condition: there exist a subsequence \( \{A_{i_j}\} \) of \( \{A_i\} \) and a sequence \( x_j \to x \) such that \( x_j \in A_{i_j} \) for all \( j \).

\( \lim \inf_{i \to \infty} A_i \) and \( \lim \sup_{i \to \infty} A_i \) are (possibly empty) closed sets, related according to

\[
\lim \inf_{i \to \infty} A_i \subset \lim \sup_{i \to \infty} A_i.
\]

In the event \( \lim \inf_{i \to \infty} A_i \) and \( \lim \sup_{i \to \infty} A_i \) coincide, we say that \( \{A_i\} \) has a limit (in the Kuratowski sense) and write

\[
\lim_{i \to \infty} A_i := \lim \inf_{i \to \infty} A_i (= \lim \sup_{i \to \infty} A_i).
\]
The sets $\liminf_{i \to \infty} A_i$ and $\limsup_{i \to \infty} A_i$ are succinctly expressed in terms of the distance function

$$d_A(x) = \inf_{y \in A} |x - y|,$$

thus

$$\liminf_{i \to \infty} A_i = \{x : \limsup_{i \to \infty} d_{A_i}(x) = 0\},$$

and

$$\limsup_{i \to \infty} A_i = \{x : \liminf_{i \to \infty} d_{A_i}(x) = 0\}.$$

Moving to a more general context, we take a set $D \subset R^k$ and a family of sets $\{S(y) \subset R^n : y \in D\}$, parameterized by points in $y \in D$. Fix a point $x \in R^k$. The set

$$\liminf_{y \to x} S(y)$$

(the Kuratowski lim inf) comprises all points $\xi$ satisfying the condition: corresponding to any sequence $y_i \to D x$, there exists a sequence $\xi_i \to \xi$ such that $\xi_i \in S(y_i)$ for all $i$. The set

$$\limsup_{y \to x} S(y)$$

(the Kuratowski lim sup) comprises all points $\xi$ satisfying the condition: there exist sequences $y_i \to D x$ and $\xi_i \to \xi$ such that $\xi_i \in S(y_i)$ for all $i$. (In the above, $y_i \to D x$ means $y_i \to x$ and $y_i \in D$ for all $i$.)

If $D$ is a neighborhood of $x$, we write $\liminf_{y \to x} S(y)$ in place of

$$\liminf_{y \to x} S(y),$$

etc.

Here, too, we have convenient characterizations of the limit sets in terms of the distance function on $R^n$: \[
\liminf_{y \to x} S(y) = \{\xi \in R^n : \limsup_{y \to x} d_{S(y)}(\xi) = 0\} \]

and

$$\limsup_{y \to x} S(y) = \{\xi \in R^n : \liminf_{y \to x} d_{S(y)}(\xi) = 0\}. $$

We observe that $\liminf_{y \to x} S(y)$ and $\limsup_{y \to x} S(y)$ are closed (possibly empty) sets, related according to

$$\liminf_{y \to x} S(y) \subset \limsup_{y \to x} S(y).$$
To reconcile these definitions, we note that, given a sequence of sets \( \{A_i\} \) in \( R^n \),

\[
\limsup_{y \to x} A_i = \limsup_{y \to x} S(y) \text{ etc.,}
\]

when we identify \( D \) with the subset \( \{1, 1/2, 1/3, \ldots\} \) of the real line, choose \( x = 0 \) and define \( S(y) = A_i \) when \( y = i^{-1}, i = 1, 2, \ldots \).

### 2.3 Measurable Multifunctions

Take a set \( \Omega \). A multifunction \( \Gamma : \Omega \leadsto R^n \) is a mapping from \( \Omega \) into the space of subsets of \( R^n \). For each \( \omega \in \Omega \) then, \( \Gamma(\omega) \) is a subset of \( R^n \). We refer to a multifunction as convex, closed, or nonempty depending on whether \( \Gamma(\omega) \) has the referred-to property for all \( \omega \in \Omega \).

Recall that a measurable space \( (\Omega, \mathcal{F}) \) comprises a set \( \Omega \) and a family \( \mathcal{F} \) of subsets of \( \Omega \) which is a \( \sigma \)-field, i.e.,

(i) \( \emptyset \in \mathcal{F} \),

(ii) \( F \in \mathcal{F} \) implies that \( \Omega \setminus F \in \mathcal{F} \), and

(iii) \( F_1, F_2, \ldots \in \mathcal{F} \) implies \( \bigcup_{i=1}^{\infty} F_i \in \mathcal{F} \).

**Definition 2.3.1** Let \( (\Omega, \mathcal{F}) \) be a measurable space. Take a multifunction \( \Gamma : \Omega \leadsto R^n \). \( \Gamma \) is measurable when the set

\[
\{ x \in \Omega : \Gamma(x) \cap C \neq \emptyset \}
\]

is \( \mathcal{F} \) measurable for every open set \( C \subset R^n \).

Fix a Lebesgue subset \( I \subset R \). Let \( \mathcal{L} \) denote the Lebesgue subsets of \( I \). If \( \Omega \) is the set \( I \) then "\( \Gamma : I \leadsto R^n \) is measurable" is taken to mean that the multifunction is \( \mathcal{L} \) measurable.

Denote by \( \mathcal{B}^k \) the Borel subsets of \( R^k \). The product \( \sigma \)-algebra \( \mathcal{L} \times \mathcal{B}^k \) (that is, the smallest \( \sigma \)-algebra of subsets of \( I \times R^k \) that contains all product sets \( A \times B \) with \( A \in \mathcal{L} \) and \( B \in \mathcal{B}^k \)) is often encountered in hypotheses invoked to guarantee measurability of multifunctions, validity of certain representations for multifunctions, etc.

A first taste of such results is provided by:

**Proposition 2.3.2** Take an \( \mathcal{L} \times \mathcal{B}^m \) measurable multifunction \( F : I \times R^n \leadsto R^k \) and a Lebesgue measurable function \( u : I \to R^m \). Then \( G : I \leadsto R^k \) defined by

\[
G(t) := F(t, u(t))
\]

is an \( \mathcal{L} \) measurable multifunction.
2.3 Measurable Multifunctions 65

**Proof.** For an arbitrary choice of set $A \in \mathcal{L}$ and $B \in \mathcal{B}^m$, the set

$$\{ t \in I : (t, u(t)) \in A \times B \}$$

is a Lebesgue subset because it is expressible as $A \cap u^{-1}(B)$ and $u$ is Lebesgue measurable. Denote by $\mathcal{D}$ the family of subsets $E \subset I \times \mathbb{R}^m$ for which the set

$$\{ t \in I : (t, u(t)) \in E \}$$

is Lebesgue measurable. $\mathcal{D}$ is a $\sigma$-field, as is easily checked. We have shown that it contains all product sets $A \times B$ with $A \in \mathcal{L}$, $B \in \mathcal{B}^m$. So $\mathcal{D}$ contains the $\sigma$-field $\mathcal{L} \times \mathcal{B}^m$.

Take any open set $W \subset \mathbb{R}^k$. Then, since $F$ is $\mathcal{L} \times \mathcal{B}^m$ measurable, $E := \{(t, u) : F(t, u) \cap W \neq \emptyset \}$ is $\mathcal{L} \times \mathcal{B}^m$ measurable. But then

$$\{ t \in I : F(t, u(t)) \cap W \neq \emptyset \} = \{ t \in I : (t, u(t)) \in E \}$$

is a Lebesgue measurable set since $E \in \mathcal{D}$. Bearing in mind that $W$ is an arbitrary open set, we conclude that $t \sim G(t) := F(t, u(t))$ is a Lebesgue measurable multifunction. □

Specializing to the point-valued case we obtain:

**Corollary 2.3.3** Consider a function $g : I \times \mathbb{R}^m \to \mathbb{R}^k$ and a Lebesgue measurable function $u : I \to \mathbb{R}^m$. Suppose that $g$ is $\mathcal{L} \times \mathcal{B}^m$ measurable. Then the mapping $t \to g(t, u(t))$ is Lebesgue measurable.

Functions $g(t, u)$ arising in Optimal Control to which Corollary 2.3.3 are often applied are composite functions of a nature covered by the following proposition.

**Proposition 2.3.4** Consider a function $\phi : I \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ satisfying the following hypotheses.

(a) $\phi(t, ., u)$ is continuous for each $(t, u) \in I \times \mathbb{R}^m$;

(b) $\phi(., x, .)$ is $\mathcal{L} \times \mathcal{B}^m$ measurable for each $x \in \mathbb{R}^n$.

Then for any Lebesgue measurable function $x : I \to \mathbb{R}^n$, the mapping $(t, u) \to \phi(t, x(t), u)$ is $\mathcal{L} \times \mathcal{B}^m$ measurable.

**Proof.** Let $\{r_j\}$ be an ordering of the set of $n$-vectors with rational coefficients. For each integer $k$ define

$$\phi_k(t, u) := \phi(t, r_j, u),$$

where $j$ is chosen ($j$ will depend on $k$ and $t$) such that

$$|x(t) - r_j| \leq 1/k \quad \text{and} \quad |x(t) - r_i| > 1/k \quad \text{for all } i \in \{1, 2, \ldots, j - 1\}.$$
(These conditions uniquely define $j$.)

Since $\phi(t, ., u)$ is continuous,

$$\phi_k(t, u) \to \phi(t, x(t), u) \text{ as } k \to \infty$$

for every $(t, u) \in I \times R^m$. It suffices then to show that $\phi_k$ is $\mathcal{L} \times \mathcal{B}^m$ measurable for an arbitrary choice of $k$.

For any open set $V \subset R^k$,

$$\phi_k^{-1}(V) = \{(t, u) \in I \times R^m : \phi_k(t, u) \in V\}$$

$$= \bigcup_{j=1}^{\infty} \{(t, u) \in I \times R^m : \phi(t, r_j, u) \in V\}$$

$$\cap \{(t, u) \in I \times R^m : |x(t) - r_j| \leq 1/k \text{ and } |x(t) - r_i| > 1/k \text{ for } i = 1, \ldots, j - 1\}.$$ 

Since the set on the right side is a countable union of $\mathcal{L} \times \mathcal{B}^m$ measurable sets, we have established that $\phi_k$ is an $\mathcal{L} \times \mathcal{B}^m$ measurable function. □

The $\mathcal{L} \times \mathcal{B}^m$ measurability hypothesis of Proposition 2.3.4 is unrestricted. It is satisfied, for example, by the Carathéodory functions.

**Definition 2.3.5** A function $g : I \times R^m \to R^k$ is said to be a Carathéodory function if

(a) $g(., u)$ is Lebesgue measurable for each $u \in R^m$,

(b) $g(t, .)$ is continuous for each $t \in I$.

**Proposition 2.3.6** Consider a function $g : I \times R^m \to R^k$. Assume that $g$ is a Carathéodory function. Then $g$ is $\mathcal{L} \times \mathcal{B}^m$ measurable.

**Proof.** Let $\{r_1, r_2, \ldots\}$ be an ordering of the set of $m$-vectors with rational components. For every positive integer $k$, $t \in I$, and $u \in R^m$, define

$$g_k(t, u) := g(t, r_j),$$

in which the integer $j$ is uniquely defined by the relationships:

$$|r_j - u| \leq 1/k \quad \text{and} \quad |r_i - u| > 1/k \quad \text{for } i = 1, \ldots, j - 1.$$ 

Since $g(t, .)$ is assumed to be continuous, we have

$$g_k(t, u) \to g(t, u)$$
as \( k \to \infty \), for each fixed \((t, u) \in I \times R^n\). It suffices then to show that \(g_k\) is \(\mathcal{L} \times \mathcal{B}^m\) measurable for each \(k\). However this follows from the fact that, for any open set \(V \subset R^k\), we have

\[
g_k^{-1}(V) = \{(t, u) \in I \times R^n : g_k(t, u) \in V\} = \bigcup_{j=1}^{\infty} \{(t, u) \in I \times R^n : g(t, r_j) \in V, |u - r_j| \leq 1/k, \text{ and } |u - r_i| > 1/k, i = 1, \ldots, j - 1\} = \bigcup_{j=1}^{\infty} \{(t \in I : g(t, r_j) \in V) \times \{u \in R^n : |u - r_j| \leq 1/k, \text{ and } |u - r_i| > 1/k \text{ for } i = 1, 2, \ldots, j - 1\}\)

and this last set is \(\mathcal{L} \times \mathcal{B}^m\) measurable, since it is expressible as a countable union of sets of the form \(A \times B\) with \(A \in \mathcal{L}\) and \(B \in \mathcal{B}^m\). □

The preceding propositions combine incidentally to provide an answer to the following question concerning an appropriate framework for the formulation of variational problems: under what hypotheses on the function \(L : I \times R^n \times R^n \to R\) is the integrand of the Lagrange functional

\[
\int L(t, x(t), \dot{x}(t))dt
\]

Lebesgue measurable for an arbitrary absolutely continuous arc \(x \in W^{1,1}\)? The two preceding propositions guarantee that the integrand is Lebesgue measurable if

(i) \(L(\cdot, x, \cdot)\) is \(\mathcal{L} \times \mathcal{B}^n\) measurable for each \(x \in R^n\) and

(ii) \(L(t_1, \cdot, u)\) is continuous for each \((t, u) \in I \times R^n\).

This is because Proposition 2.3.4 tells us that \((t, u) \to L(t, x(t), u)\) is \(\mathcal{L} \times \mathcal{B}^n\) measurable, in view of assumptions (i) and (ii), and since \(x : I \to R^n\) is Lebesgue measurable. Corollary 2.3.3 permits us to conclude, since \(t \to \dot{x}(t)\) is Lebesgue measurable, that \(t \to L(t, x(t), \dot{x}(t))\) is indeed Lebesgue measurable.

The following theorem, a proof of which is to be found in [27], lists important characterizations of closed multifunctions that are measurable. (Throughout, \(I\) is a Lebesgue subset of \(R\).)

**Theorem 2.3.7** Take a multifunction \(\Gamma : I \sim R^n\) and define \(D := \{t \in I : \Gamma(t) \neq \emptyset\}\). Assume that \(\Gamma\) is closed. Then the following statements are equivalent.

(a) \(\Gamma\) is an \(\mathcal{L}\) measurable multifunction.

(b) \(Gr \Gamma\) is an \(\mathcal{L} \times \mathcal{B}^n\) measurable set.
2. Measurable Multifunctions

(c) \( D \) is a Lebesgue subset of \( I \) and there exists a sequence \( \{\gamma_k : D \to \mathbb{R}^n\} \)

of Lebesgue measurable functions such that

\[
\Gamma(t) = \bigcup_{k=1}^{\infty} \{\gamma_k(t)\} \quad \text{for all } t \in D.
\] (2.2)

The representation of a multifunction in terms of a countable family of
Lebesgue measurable functions according to (2.2) is called the Castaing
Representation of \( \Gamma \).

Our aim now is to establish the measurability of a number of frequently
encountered multifunctions derived from other multifunctions.

**Proposition 2.3.8** Take a measurable space \((\Omega, \mathcal{F})\) and a measurable multifunction \( \Gamma : \Omega \rightharpoonup \mathbb{R}^n \). Then the multifunction \( \tilde{\Gamma} : \Omega \rightharpoonup \mathbb{R}^n \) is also meas-
urable in each of the following cases:

(i) \( \tilde{\Gamma}(y) := \overline{\Gamma(y)} \) for all \( y \in \Omega \).

(ii) \( \tilde{\Gamma}(y) := \text{co} \Gamma(y) \) for all \( y \in \Omega \).

**Proof.**

(i): \( \tilde{\Gamma} \) is measurable in this case since, for any open set \( W \in \mathbb{R}^n \),

\[
\{y \in \Omega : \overline{\Gamma(y)} \cap W \neq \emptyset\} = \{y \in \Omega : \Gamma(y) \cap W \neq \emptyset\}.
\]

(ii): Define the multifunction \( \Gamma^{(n+1)} : \Omega \rightharpoonup \mathbb{R}^{n \times (n+1)} \) to be

\[
\Gamma^{(n+1)}(y) := \Gamma(y) \times \ldots \times \Gamma(y) \quad \text{for all } y \in \Omega.
\]

Then \( \Gamma^{(n+1)} \) is measurable, since \( \{y \in \Omega : \Gamma^{(n+1)}(y) \cap W \neq \emptyset\} \) is obviously measurable for any set \( W \) in \( \mathbb{R}^{n \times (n+1)} \) that is a product of open sets of \( \mathbb{R}^n \), and therefore for any open set \( W \), since an arbitrary open set in the product space can be expressed as a countable union of such sets.

Define also

\[
\Lambda := \left\{ (\lambda_0, \ldots, \lambda_n) : \lambda_i \geq 0 \text{ for all } i, \sum_{i=0}^{n} \lambda_i = 1 \right\}.
\]

Take any open set \( W \) in \( \mathbb{R}^n \) and define

\[
W^{(n+1)} := \left\{ (w_0, \ldots, w_n) : \sum_i \lambda_i w_i \in W, (\lambda_0, \ldots, \lambda_n) \in \Lambda \right\}.
\]

Obviously, \( W^{(n+1)} \) is an open set.

We must show that the set

\[
\{y \in \Omega : \text{co} \Gamma(y) \cap W \neq \emptyset\}
\]
is measurable. But this follows immediately from the facts that \( \tilde{\Gamma}^{(n+1)} \) is measurable and \( W^{(n+1)} \) is open, since

\[
\{y : \text{co } \Gamma(y) \cap W \neq \emptyset\} = \{y : \Gamma^{(n+1)}(y) \cap W^{(n+1)} \neq \emptyset\}.
\]

\( \square \)

The next proposition concerns the measurability properties of limits of sequences of multifunctions. We make reference to Kuratowski sense limit operations, defined in Section 2.2.

**Theorem 2.3.9** Consider closed multifunctions \( \Gamma_j : I \rightrightarrows \mathbb{R}^n \), \( j = 1, 2, \ldots \). Assume that \( \Gamma_j \) is \( \mathcal{L} \) measurable for each \( j \). Then the closed multifunction \( \Gamma : I \rightrightarrows \mathbb{R}^n \) is also \( \mathcal{L} \) measurable when \( \Gamma \) is defined in each of the following ways.

(a) \( \Gamma(t) := \overline{\bigcup_{j \geq 1} \Gamma_j(t)} \);

(b) \( \Gamma(t) := \bigcap_{j \geq 1} \Gamma_j(t) \);

(c) \( \Gamma(t) := \limsup_{j \to \infty} \Gamma_j(t) \);

(d) \( \Gamma(t) := \liminf_{j \to \infty} \Gamma_j(t) \).

(c) and (d) imply in particular that if \( \{\Gamma_j\} \) has a limit as \( j \to \infty \), then \( \lim_{j \to \infty} \Gamma_j \) is measurable.

**Proof.** (a): \( (\Gamma(t) = \overline{\bigcup_{j \geq 1} \Gamma_j(t)} ) \)

Take any open set \( W \subset \mathbb{R}^n \). Then, since \( W \) is an open set, we have

\[
\{t \in I : \Gamma(t) \cap W \neq \emptyset\} = \{t \in I : \overline{\bigcup_{j \geq 1} \Gamma_j(t)} \cap W \neq \emptyset\} = \{t \in I : \bigcup_{j \geq 1} \Gamma_j(t) \cap W \neq \emptyset\} = \bigcup_{j \geq 1} \{t \in I : \Gamma_j(t) \cap W \neq \emptyset\}.
\]

This establishes the measurability of \( t \rightrightarrows \Gamma(t) \) since the set on the right side, a countable union of Lebesgue measurable sets, is Lebesgue measurable.

(b): \( (\Gamma(t) = \bigcap_{j \geq 1} \Gamma_j(t) ) \)

In this case,

\[
\text{Gr } \Gamma = \text{Gr } \{t \rightrightarrows \bigcap_{j \geq 1} \Gamma_j(t)\} = \bigcap_{j \geq 1} \text{Gr } \Gamma_j.
\]

\( \text{Gr } \Gamma \) then is \( \mathcal{L} \times \mathcal{B}^n \) measurable since each \( \text{Gr } \Gamma_j \) is \( \mathcal{L} \times \mathcal{B}^n \) measurable by Theorem 2.3.7. Now apply again Theorem 2.3.7.
(c): \((\Gamma(t) = \limsup_{j \to \infty} \Gamma_j(t))\)

The measurability of \(t \sim \Gamma(t)\) in this case follows from (a) and (b) and the following characterization of \(\limsup_{j \to \infty} \Gamma_j(t)\).

\[
\limsup_{j \to \infty} \Gamma_j(t) = \bigcap_{j \geq 1} \bigcup_{j \geq j} \Gamma_j(t).
\]

(d): \((\Gamma(t) = \liminf_{j \to \infty} \Gamma_j(t))\)

Define

\[
\Gamma_j^k(t) := \Gamma_j(t) + (1/k)B.
\]

Notice that \(\Gamma_j^k\) is measurable since, for any closed set \(W \subset R^n\),

\[
\{t : \Gamma_j^k(t) \cap W \neq \emptyset\} = \{t : \Gamma_j(t) \cap (W + (1/k)B) \neq \emptyset\}
\]

and the set \(W + (1/k)B\) is closed. The measurability of \(\Gamma\) in this case too follows from (a) and (b) in view of the identity

\[
\liminf_{j \to \infty} \Gamma_j(t) = \bigcap_{k \geq 1} \bigcup_{j \geq 1} \bigcap_{j \geq j} \Gamma_j^k(t).
\]

\(\square\)

**Proposition 2.3.10** Take a multifunction \(F : I \times R^n \sim R^k\) and a Lebesgue measurable function \(\bar{x} : I \to R^n\). Assume that

(a) for each \(x\), \(F(\cdot, x) : I \sim R^k\) is an \(L\) measurable, nonempty, closed multifunction;

(b) for each \(t\), \(F(t, \cdot)\) is continuous at \(x = \bar{x}(t)\), in the sense that

\[
y_i \to \bar{x}(t) \text{ implies } F(t, \bar{x}(t)) = \lim_{i \to \infty} F(t, y_i).
\]

Then \(G : I \sim R^k\) defined by

\[
G(t) := F(t, \bar{x}(t))
\]

is a closed \(L\) measurable multifunction.

**Proof.** Let \(\{r_i\}\) be an ordering of \(n\)-vectors with rational entries. For each integer \(l\) and for each \(t \in I\), define

\[
F_l(t) := F(t, r_j),
\]

in which \(j\) is chosen according to the rule

\[
|\bar{x}(t) - r_j| \leq 1/l \quad \text{and} \quad |\bar{x}(t) - r_i| > 1/l \quad \text{for} \quad i = 1, \ldots, j - 1.
\]
In view of the continuity properties of $F(t,.)$,

$$F(t, \bar{x}(t)) = \lim_{l \to \infty} F_l(t).$$

By Theorem 2.3.9 then it suffices to show that $F_l$ is measurable for arbitrary $l$.

We observe however that

$$\text{Gr } F_l = \bigcup_j (\text{Gr } F(.,r_j) \cap (A_j \times R^k)),$$

where

$$A_j := \{t : |\bar{x}(t) - r_j| \leq 1/l, |\bar{x}(t) - r_i| > 1/l \quad \text{for } i = 1, \ldots, j - 1\}.$$

Since $F(.,r_j)$ has an $\mathcal{L} \times \mathcal{B}^k$ measurable graph (see Theorem 2.3.7) and $\bar{x}$ is Lebesgue measurable, we see that $\text{Gr } F_l$ is $\mathcal{L} \times \mathcal{B}^k$ measurable. Applying Theorem 2.3.7 again, we see that the closed multifunction $F_l$ is measurable. $\Box$

Take a multifunction $\Gamma : I \leadsto R^k$. We say that a function $x : I \to R^k$ is a measurable selection for $\Gamma$ if

(i) $x$ is Lebesgue measurable, and

(ii) $x(t) \in \Gamma(t)$ a.e.

We obtain directly from Theorem 2.3.7 the following conditions for $\Gamma$ to have a measurable selection.

**Theorem 2.3.11** Let $\Gamma : I \leadsto R^k$ be a nonempty multifunction. Assume that $\Gamma$ is closed and measurable. Then $\Gamma$ has a measurable selection.

(In fact Theorem 2.3.7 tells us rather more than this: not only does there exist a measurable selection under the stated hypotheses, but the measurable selections are sufficiently numerous to “fill out” the values of the multifunction.)

The above measurable selection theorem is inadequate for certain applications in which the multifunction is not closed. An important extension (see [27], [146]) is

**Theorem 2.3.12** (Aumann’s Measurable Selection Theorem) Let $\Gamma : I \leadsto R^k$ be a nonempty multifunction. Assume that

$$\text{Gr } \Gamma \text{ is } \mathcal{L} \times \mathcal{B}^k \text{ measurable.}$$

Then $\Gamma$ has a measurable selection.
This can be regarded as a generalization of Theorem 2.3.11 since if $\Gamma$ is closed and measurable then, by Theorem 2.3.7, $\text{Gr } \Gamma$ is automatically $\mathcal{L} \times \mathcal{B}^k$ measurable.

Of particular significance in applications to Optimal Control is the following measurable selection theorem involving the composition of a function and a multifunction.

**Theorem 2.3.13 (The Generalized Filippov Selection Theorem)**

Consider a nonempty multifunction $U : I \leadsto R^m$ and a function $g : I \times R^m \to R^n$ satisfying

(a) the set $\text{Gr } U$ is $\mathcal{L} \times \mathcal{B}^m$ measurable;

(b) the function $g$ is $\mathcal{L} \times \mathcal{B}^m$ measurable.

Then for any measurable function $v : I \to R^n$, the multifunction $U' : I \leadsto R^m$ defined by

$$U'(t) := \{u \in U(t) : g(t, u) = v(t)\}$$

has an $\mathcal{L} \times \mathcal{B}^m$ measurable graph. Furthermore, if

$$v(t) \in \{g(t, u) : u \in U(t)\} \quad \text{a.e.} \quad (2.3)$$

then there exists a measurable function $u : I \to R^m$ satisfying

$$u(t) \in U(t) \quad \text{a.e.} \quad (2.4)$$

$$g(t, u(t)) = v(t) \quad \text{a.e.} \quad (2.5)$$

Notice that Condition (2.3) is just a rephrasing of the hypothesis

$U'(t)$ is nonempty for a.e. $t \in I$

and the final assertion can be expressed in measurable selection terms as

the multifunction $U'$ has a measurable selection,

a fact that follows directly from Aumann's Selection Theorem.

We mention that the name Filippov's Selection Theorem usually attaches to the final assertion of the theorem concerning existence of a measurable function $u : I \to R^m$ satisfying (2.4) and (2.5) under (2.3) and strengthened forms of Hypotheses (a) and (b), namely,

(a)' $U$ is a closed measurable multifunction, and

(b)' $g$ is a Carathéodory function.
Proof. By redefining \( U'(t) \) on a nullset if required, we can arrange that \( U'(t) \) is nonempty for all \( t \in I \). In view of the preceding discussion it is required to show merely that \( \text{Gr } U' \) is \( \mathcal{L} \times \mathcal{B}^m \) measurable. But this follows directly from the relationship

\[
\text{Gr } U' = \phi^{-1}\{0\} \cap \text{Gr } U
\]

in which \( \phi(t,u) := g(t,u) - v(t) \), since under the hypotheses both \( \phi^{-1}\{0\} \) and \( \text{Gr } U \) are \( \mathcal{L} \times \mathcal{B}^m \) measurable sets. \( \square \)

The relevance of Filippov’s Theorem in a Control Systems context is illustrated by the following application. Take a function \( f : [S,T] \times R^n \times R^m \to R^n \) and a multifunction \( U : [S,T] \rightsquigarrow R^m \). The class of state trajectories for the control system

\[
\dot{x}(t) = f(t,x(t),u(t)) \quad \text{a.e. } t \in [S,T] \\
u(t) \in U(t) \quad \text{a.e. } t \in [S,T]
\]

comprises every absolutely continuous function \( x : [S,T] \to R^n \) for some measurable \( u : [S,T] \to R^m \). It is often desirable to interpret the state trajectories as solutions of the differential inclusion \( \dot{x}(t) \in F(t,x(t)) \) with

\[
F(t,x) := \{ f(t,x,u) : u \in U(t) \}.
\]

The question then arises whether the state trajectories for the control system are precisely the absolutely continuous functions \( x \) satisfying

\[
\dot{x}(t) \in F(t,x(t)) \quad \text{a.e.} \tag{2.6}
\]

Clearly a necessary condition for an absolutely continuous function to be a state trajectory for the control system is that (2.6) is satisfied. Filippov’s Theorem tells us that (2.6) is also a sufficient condition (i.e., the differential inclusion provides an equivalent description of state trajectories) under the hypotheses:

(i) \( f(\cdot , x, \cdot) \) is \( \mathcal{L} \times \mathcal{B}^m \) measurable and \( f(t, \cdot , u) \) is continuous;

(ii) \( \text{Gr } U \) is \( \mathcal{L} \times \mathcal{B}^m \) measurable.

To see this, apply the Generalized Filippov Selection Theorem with \( g(t,u) = f(t,x(t),u) \) and \( v = \dot{x} \). (The relevant hypotheses, (b) and (2.3), are satisfied in view of Proposition 2.3.4 and since \( \dot{x}(t) \in F(t,x(t)) \) a.e.) This yields a measurable function \( u : [S,T] \to R^m \) satisfying \( \dot{x}(t) = f(t,x(t),u(t)) \) and \( u(t) \in U(t) \) a.e. thereby confirming that \( x \) is a state trajectory for the control system.

Having once again an eye for future Optimal Control applications, we now establish measurability of various derived functions and the existence
of measurable selections for related multifunctions. The source of a number of useful results is the following theorem concerning the measurability of a "marginal" function.

**Theorem 2.3.14** Consider a function $g : I \times R^k \to R$ and a closed nonempty multifunction $\Gamma : I \rightrightarrows R^k$. Assume that

(a) $g$ is a Carathéodory function;

(b) $\Gamma'$ is a measurable multifunction.

Define the extended-valued function $\eta : I \to R \cup \{-\infty\}$

$$\eta(t) = \inf_{\gamma \in \Gamma'(t)} g(t, \gamma) \quad \text{for } t \in I.$$  

Then $\eta$ is a Lebesgue measurable function. Furthermore, if we define

$$I' := \{t \in I : \inf_{\gamma' \in \Gamma'(t)} g(t, \gamma') = g(t, \gamma) \text{ for some } \gamma \in \Gamma(t)\}$$

(i.e., $I'$ is the set of points $t$ for which the infimum of $g(t, \cdot)$ is achieved over $\Gamma(t)$) then $I'$ is a Lebesgue measurable set and there exists a measurable function $\gamma : I' \to R^k$ such that

$$\eta(t) = g(t, \gamma(t)) \quad \text{a.e. } t \in I'. \quad (2.7)$$

**Proof.** Since $\Gamma$ is closed, nonempty, and measurable, it has a Castaing representation in terms of some countable family of measurable functions $\{\gamma_i : I \to R^k\}$. Since $g(t, \cdot)$ is continuous and $\{\gamma_i(t)\}$ is dense in $\Gamma(t)$,

$$\eta(t) = \inf\{g(t, \gamma_i(t)) : i \text{ an integer}\}$$

for all $t \in I$.

Now according to Proposition 2.3.6 and Corollary 2.3.3, $t \to g(t, \gamma_i(t))$ is a measurable function. It follows from a well-known property of measurable functions that $\eta$, which we have expressed as the pointwise infimum of a countable family of measurable functions, is Lebesgue measurable and $\text{dom}\{\eta\} := \{t \in I : \eta(t) > -\infty\}$ is a Lebesgue measurable set.

Now apply the Generalized Filippov Selection Theorem (identifying $\eta$ with $v$ and $\Gamma$ with $U$, and replacing $I$ by $\text{dom}\{\eta\}$). If $I' = \emptyset$, there is nothing to prove. Otherwise, since

$$I' = \{t \in \text{dom}\{\eta\} : \{\gamma \in \Gamma(t) : g(t, \gamma) = \eta(t)\} \neq \emptyset\},$$

$I'$ is a nonempty, Lebesgue measurable set and there exists a measurable function $\gamma : I' \to R^k$ such that

$$\eta(t) = g(t, \gamma(t)) \quad \text{a.e. } t \in I'.$$
2.4 Existence and Estimation of $F$-Trajectories

Fix an interval $[S, T]$ and a relatively open set $\Omega \subset [S, T] \times \mathbb{R}^n$. For $t \in [S, T]$, define

$$\Omega_t := \{x : (t, x) \in \Omega\}.$$  

Take a continuous function $y : [S, T] \to \mathbb{R}^n$ and $\varepsilon > 0$. Then the $\varepsilon$ tube about $y$ is the set

$$T(y, \varepsilon) := \{(t, x) \in [S, T] \times \mathbb{R}^n : t \in [S, T], |x - y(t)| \leq \varepsilon\}.$$  

Consider a multifunction $F : \Omega \rightrightarrows \mathbb{R}^n$. An arc $x \in W^{1,1}([S, T]; \mathbb{R}^n)$ is an $F$-trajectory if $\text{Gr} x \in \Omega$ and

$$\dot{x}(t) \in F(t, x(t)) \quad \text{a.e. } t \in [S, T].$$  

Naturally we would like to know when $F$-trajectories exist. We make extensive use of a local existence theorem that gives conditions under which an $F$-trajectory exists near a nominal arc $y \in W^{1,1}([S, T]; \mathbb{R}^n)$. This theorem provides important supplementary information about how "close" to $y$ the $F$-trajectory $x$ may be chosen. Just how close will depend on the extent to which the nominal arc $y$ fails to satisfy the differential inclusion, as measured by the function $\Lambda_F$,

$$\Lambda_F(y) := \int_S^T \rho_F(t, y(t), \dot{y}(t)) \, dt.$$  

Here

$$\rho_F(t, x, v) := \inf\{|\eta - v| : \eta \in F(t, x)\}.$$  

(We make use of function $\Lambda_F(y)$ only when the integrand above is Lebesgue measurable and $F$ is a closed multifunction; then $\Lambda_F(y)$ is a nonnegative number that is zero if and only if $y$ is an $F$-trajectory.)

We pause for a moment however to list some relevant properties of $\rho_F$ that, among other things, give conditions under which the integral $\Lambda_F(y)$ is well-defined.

**Proposition 2.4.1** Take a multifunction $F : [S, T] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

(a) Fix $(t, x) \in \text{dom} F$. Then

$$|\rho_F(t, x, v) - \rho_F(t, x, v')| \leq |v - v'| \quad \text{for all } v, v' \in \mathbb{R}^n.$$  

(b) Fix $t \in [S, T]$. Suppose that, for some $\varepsilon > 0$ and $k > 0$, $\bar{x} + \varepsilon B \subset \text{dom} F$ and

$$F(t, x) \subset F(t, x') + k|x - x'|B \quad \text{for all } x, x' \in \bar{x} + \varepsilon B.$$  

Then

$$|\rho_F(t, x, v) - \rho_F(t, x', v')| \leq |v - v'| + k|x - x'|$$  

for all $v, v' \in \mathbb{R}^n$ and $x, x' \in \bar{x} + \varepsilon B$.  

(c) Assume that $F$ is $\mathcal{L} \times B^n$ measurable. Then for any Lebesgue measurable functions $y : [S, T] \to R^n$ and $v : [S, T] \to R^n$ such that $Gry \subset \text{dom} \ F$, we have that $t \to \rho_F(t, y(t), v(t))$ is a Lebesgue measurable function on $[S, T]$.

Proof.

(a) Choose any $\epsilon > 0$ and $v, v' \in R^n$. Since $F(t, x) \neq \emptyset$, there exists $\eta \in F(t, x)$ such that

$$
\rho_F(t, x, v) \geq |v - \eta| - \epsilon \\
\geq |v' - \eta| - |v - v'| - \epsilon \\
\geq \rho_F(t, x, v') - |v - v'| - \epsilon.
$$

(The second line follows from the triangle inequality.) Since $\epsilon > 0$ is arbitrary, and $v$ and $v'$ are interchangeable, it follows that

$$
|\rho_F(t, x, v') - \rho_F(t, x, v')| \leq |v - v'|.
$$

(b) Choose any $x, x' \in \bar{B}$ and $v, v' \in R^n$. Take any $\delta$. Since $F(t, x') \neq \emptyset$, there exists $\eta' \in F(t, x')$ such that

$$
\rho_F(t, x', v') > |v' - \eta'| - \delta.
$$

Under the hypotheses, there exists $\eta \in F(t, x)$ such that $|\eta - \eta'| \leq k|x - x'|$. Of course, $\rho_F(t, x, v) \leq |v - \eta|$. It follows from these relationships and the triangle inequality that

$$
\rho_F(t, x, v) - \rho_F(t, x', v') \\
\leq |v - \eta| - |v' - \eta'| + \delta \leq |v - v'| + |v' - \eta| - |v' - \eta'| + \delta \\
\leq |v - v'| + |\eta - \eta'| + \delta \leq |v - v'| + k|x - x'| + \delta.
$$

Since the roles of $(x, v)$ and $(x', v')$ are interchangeable and $\delta > 0$ is arbitrary, we conclude that

$$
|\rho_F(t, x, v) - \rho_F(t, x', v')| \leq |v - v'| + k|x - x'|.
$$

(c) For each $v$, the function $t \to \rho_F(t, y(t), v)$ is measurable. This follows from the identity

$$
\{ t \in [S, T] : \rho_F(t, y(t), v) < \alpha \} = \{ t : F(t, y(t)) \cap (v + \alpha \text{int} \ B) \neq \emptyset \},
$$

valid for any $\alpha \in R$, and the fact that $t \to F(t, y(t))$ is measurable (see Proposition 2.3.2).

By Part (a) of the lemma, the function $\rho_F(t, y(t), \cdot)$ is continuous, for each $t \in [S, T]$. It follows that $(t, v) \to \rho_F(t, y(t), v)$ is a Carathéodory
function. But then \( t \rightarrow \rho_F(t,y(t),v(t)) \) is (Lebesgue) measurable on \([S,T]\), by Proposition 2.3.6 and Corollary 2.3.3. \( \Box \)

The following proposition, concerning regularity properties of projections onto continuous convex multifunctions, is also required.

**Proposition 2.4.2** Take a continuous multifunction \( \Gamma : [S,T] \rightrightarrows R^n \) and a function \( u : [S,T] \rightarrow R^n \). Assume that

(i) \( u \) is continuous,

(ii) \( \Gamma(t) \) is nonempty, compact, and convex for each \( t \in [S,T] \), and \( \Gamma \) is continuous; i.e., there exists \( o : R^+ \rightarrow R^+ \) with \( \lim_{\alpha \downarrow 0} o(\alpha) = 0 \) such that

\[
\Gamma(s) \subset \Gamma(t) + o(|t-s|)B \quad \text{for all } t,s \in [S,T].
\]

Let \( \hat{u} : [S,T] \rightarrow R^n \) be the function defined according to

\[
|u(t) - \hat{u}(t)| = \min_{u' \in \Gamma(t)} |u(t) - u'| \quad \text{for all } t.
\]

(There is a unique minimizer for each \( t \) since \( \Gamma(t) \) is nonempty, closed, and convex.)

Then \( \hat{u} \) is a continuous function.

**Proof.** Suppose that \( \hat{u} \) is not continuous. Then there exist \( \epsilon > 0 \) and sequences \( \{s_i\} \) and \( \{t_i\} \) in \([S,T]\) such that \( |t_i - s_i| \rightarrow 0 \) and

\[
|\hat{u}(t_i) - \hat{u}(s_i)| > \epsilon, \quad \text{for all } i. \quad (2.8)
\]

Since the multifunction \( \Gamma \) is compact-valued and continuous and has bounded domain, \( \text{Gr } \Gamma \) is bounded. It follows that, for some \( K > 0, \)

\[
u(s_i), u(t_i), \hat{u}(s_i), \hat{u}(t_i) \in KB.
\]

Since \( \Gamma \) is continuous, there exist \( \{y_i\} \) and \( \{z_i\} \) such that

\[
y_i \in \Gamma(t_i), \quad z_i \in \Gamma(s_i) \quad \text{for all } i \quad (2.9)
\]

and

\[
|y_i - \hat{u}(s_i)| \rightarrow 0, \quad |z_i - \hat{u}(t_i)| \rightarrow 0 \quad \text{as } i \rightarrow \infty. \quad (2.10)
\]

Since \( \Gamma \) is a convex multifunction, \( u(s_i) - \hat{u}(s_i) \) and \( u(t_i) - \hat{u}(t_i) \) are normal vectors to \( \Gamma(s_i) \) and \( \Gamma(t_i) \), respectively. Noting (2.9), we deduce from the normal inequality for convex sets that

\[
(u(s_i) - \hat{u}(s_i)) \cdot (z_i - \hat{u}(s_i)) \leq 0
\]

\[
(u(t_i) - \hat{u}(t_i)) \cdot (y_i - \hat{u}(t_i)) \leq 0.
\]
It follows that
\[
(u(s_i) - \hat{u}(s_i)) \cdot (\hat{u}(t_i) - \hat{u}(s_i)) \leq (u(s_i) - \hat{u}(s_i)) \cdot (\hat{u}(t_i) - z_i)
\]
\[
(u(t_i) - \hat{u}(t_i)) \cdot (\hat{u}(s_i) - \hat{u}(t_i)) \leq (u(t_i) - \hat{u}(t_i)) \cdot (\hat{u}(s_i) - y_i).
\]
Adding these inequalities gives
\[
(u(s_i) - u(t_i)) \cdot (\hat{u}(t_i) - \hat{u}(s_i)) + |\hat{u}(t_i) - \hat{u}(s_i)|^2 
\leq (u(s_i) - \hat{u}(s_i)) \cdot (\hat{u}(t_i) - z_i) + (u(t_i) - \hat{u}(t_i)) \cdot (\hat{u}(s_i) - y_i).
\]
We deduce that
\[
|\hat{u}(t_i) - \hat{u}(s_i)|^2 \leq |u(s_i) - u(t_i)| \cdot |\hat{u}(t_i) - \hat{u}(s_i)| 
+ |u(s_i) - \hat{u}(s_i)| \cdot |\hat{u}(t_i) - z_i| + |u(t_i) - \hat{u}(t_i)| \cdot |\hat{u}(s_i) - y_i|.
\]
It follows that
\[
|\hat{u}(t_i) - \hat{u}(s_i)|^2 \leq 2K(|u(s_i) - u(t_i)| + |\hat{u}(t_i) - z_i| + |\hat{u}(s_i) - y_i|).
\]
But the right side has limit 0 as \(i \to \infty\), since \(u\) is continuous, and by (2.10). We conclude that \(|\hat{u}(t_i) - \hat{u}(s_i)| \to 0\) as \(i \to \infty\). We have arrived at a contradiction of (2.8). \(\hat{u}\) must therefore be continuous. □

We refer to the "truncation" function \(tr_\epsilon : \mathbb{R}^n \to \mathbb{R}^n\), defined to be
\[
tr_\epsilon(\xi) := \begin{cases} 
\xi & \text{if } |\xi| \leq \epsilon \\
\epsilon|\xi|^{-1}\xi & \text{if } |\xi| > \epsilon.
\end{cases} \tag{2.11}
\]
Recall that \(T(y, \epsilon)\) denotes the \(\epsilon\) tube about the arc \(y\):
\[
T(y, \epsilon) := \{(t, x) \in [S, T] \times \mathbb{R}^n : t \in [S, T], |x - y(t)| \leq \epsilon\}.
\]

**Theorem 2.4.3 (Generalized Filippov Existence Theorem)** Let \(\Omega\) be a relatively open set in \([S, T] \times \mathbb{R}^n\). Take a multifunction \(F : \Omega \rightharpoonup \mathbb{R}^n\), an arc \(y \in W^{1,1}([S, T]; \mathbb{R}^n)\), a point \(\xi \in \mathbb{R}^n\), and \(\epsilon \in (0, +\infty) \cup \{+\infty\}\) such that \(T(y, \epsilon) \subset \Omega\). Assume that

(i) \(F(t, x')\) is a closed nonempty set for all \((t, x') \in T(y, \epsilon)\), and \(F\) is \(L \times B^n\) measurable;

(ii) there exists \(k \in L^1\) such that
\[
F(t, x') \subset F(t, x'') + k(t)|x' - x''|B \tag{2.12}
\]
for all \(x', x'' \in y(t) + \epsilon B\), a.e. \(t \in [S, T]\).
Assume further that
\[
K \left( |\xi - y(S)| + \int_S^T \rho_F(t, y(t), \dot{y}(t)) dt \right) \leq \epsilon,
\]  
where \( K := \exp \left( \int_S^T k(t) dt \right) \).

Then there exists an \( F \)-trajectory \( x \) satisfying \( x(S) = \xi \) such that
\[
||x - y||_{L^\infty} \leq |x(S) - y(S)| + \int_S^T |\dot{x}(t) - \dot{y}(t)| dt
\leq K \left( |\xi - y(S)| + \int_S^T \rho_F(t, y(t), \dot{y}(t)) dt \right).
\]  
\[\tag{2.14}\]

Now suppose that (i) and (ii) are replaced by the stronger hypotheses

(i)' \( F(t, x') \) is a nonempty, compact, convex set for all \( (t, x') \in T(y, \epsilon) \);

(ii)' there exists a function \( o(.) : R^+ \to R^+ \) and \( k_\infty > 0 \) such that
\[
\lim_{\alpha \to 0} o(\alpha) = 0 \quad \text{and} \quad F(s', x') \subset F(s'', x'') + k_\infty |x' - x''| B + o(|s' - s''|) B
\]  
for all \( (s', x'), (s'', x'') \in T(y, \epsilon) \).

Then, if \( y \) is continuously differentiable, \( x \) can be chosen also to be continuously differentiable.

(If \( \epsilon = +\infty \) then in the above hypotheses \( T(y, \epsilon) \) and \( \epsilon B \) are interpreted as \( [S, T] \times R^n \) and \( R^n \), respectively; the left side of Condition (2.13) is required to be finite.)

Remarks

(i) The hypotheses invoked in the first part of Theorem 2.4.3 do not require \( F \) to be convex-valued. For many developments in Optimal Control the requirement that \( F \) is convex is crucial; fortunately, proving this basic existence theorem is not one of them.

(ii) The proof of Theorem 2.4.3 is by construction. The iterative procedure used is a generalization to differential inclusions of the well-known Picard iteration scheme for obtaining a solution to the differential equation
\[
\dot{x}(t) = f(t, x(t)), \quad x(S) = \xi.
\]

An initial guess \( y \) at a solution is made. It is then improved by "successive approximations" \( x_0, x_1, x_2, \ldots \). These arcs are generated by the recursive equations
\[
x_{i+1}(t) = \xi + \int_S^t f(s, x_i(s)) ds
\]  

\[\tag{2.15}\]
with starting condition $x_0(t) = y(t) + (\xi - y(S))$.

**Proof.** We may assume without loss of generality that $\epsilon = \infty$. Indeed if $\epsilon = \bar{\epsilon}$ for some finite $\bar{\epsilon}$ then we consider $\tilde{F}$ in place of $F$, where

$$\tilde{F}(t,x) := F(t,y(t) + tr\bar{\epsilon}(x-y(t))).$$

(See (2.11) for the definition of $tr\bar{\epsilon}$.)

$\tilde{F}$ satisfies the hypotheses (in relation to $y$) with $\epsilon = \infty$ and with the same $k \in L^1$ as before. Of course

$$\tilde{F}(t,x) = F(t,x) \quad \text{for } x \in y(t) + \bar{\epsilon}B.$$  

Now apply the $\epsilon = +\infty$ case of the theorem to $\tilde{F}$. This gives an $\tilde{F}$-trajectory $x$ such that $x(S) = \xi$ and (2.14) is satisfied (when $\tilde{F}$ replaces $F$). If, however,

$$K \left( |\xi - y(S)| + \int_S^T \rho_F(t,y(t),\dot{y}(t))dt \right) \leq \bar{\epsilon},$$

then the theorem tell us that

$$||x - y||_{L^\infty} \leq \epsilon,$$

and therefore $x$ is an $F$-trajectory because $F(t,.)$ and $\tilde{F}(t,.)$ coincide on $y(t) + \bar{\epsilon}B$. This justifies setting $\epsilon = +\infty$.

It suffices to consider only the case $\xi = y(S)$. To show this, suppose that $\xi \neq y(S)$. Replace the underlying time interval $[S,T]$ by $[S-1,T]$ and $\xi$ by $\hat{\xi} = y(S)$. Replace also $F$ by $\tilde{F} : [S-1,T] \times R^n \sim N^n$ and $y$ by $\hat{y} : [S-1,T] \rightarrow R^n$, defined as follows.

$$\tilde{F}(t,x) := \begin{cases} F(t,x) & \text{for } (t,x) \in [S,T] \times R^n \\ \{\xi - y(S)\} & \text{for } (t,x) \in [S-1,S) \times R^n \end{cases}$$

and

$$\hat{y}(t) := \begin{cases} y(t) & \text{for } t \geq S \\ y(S) & \text{for } t < S. \end{cases}$$

Now apply the special case of the theorem to find $\hat{x} \in W^{1,1}([S-1,T]; R^n)$ such that $\hat{x}(S-1) = \hat{y}(S-1) = \hat{\xi} \equiv y(S)$. Take $x$ to be the restriction of $\hat{x}$ to $[S,T]$. We readily deduce that

$$||\hat{x} - \hat{y}||_{L^1([S-1,T]; R^n)} \leq K \int_{S-1}^T \rho_{\tilde{F}}(t,\hat{y}(t),\dot{\hat{y}}(t))dt.$$  

Now, to derive the desired estimate, we have merely to note that

$$||\hat{x} - \hat{y}||_{L^1([S-1,T]; R^n)} = |\xi - y(S)| + ||\hat{x} - \hat{y}||_{L^1([S,T]; R^n)}$$
and
\[ \int_{S-1}^{T} \rho_F(t, \tilde{y}(t), \dot{y}(t)) dt = |\xi - y(s)| + \int_{S}^{T} \rho_F(t, y(t), \dot{y}(t)) dt. \]

Henceforth, then, we assume that \( \epsilon = +\infty \) and \( y(S) = \xi \); we must find an \( F \)-trajectory \( x \) such that \( x(S) = \xi \) and
\[ ||\dot{x} - \dot{y}||_{L^1([S, T]; \mathbb{R}^n)} \leq K \int_{S}^{T} \rho_F(t, y(t), \dot{y}(t)) dt. \]

Write \( x_0(t) = y(t) \). According to Theorem 2.3.11, we may choose a measurable function \( v_1 \) satisfying
\[ v_1(t) \in F(t, x_0(t)) \quad \text{a.e. } t \in [S, T] \]
and
\[ \rho_F(t, x_0(t), \dot{x}_0(t)) = |v_1(t) - \dot{x}_0(t)| \quad \text{a.e. } t \in [S, T]. \]
This is because
\[ t \sim G(t) := \{ v \in F(t, x_0(t)) : \rho_F(t, x_0(t), \dot{x}_0(t)) = |v - \dot{x}_0(t)| \} \]
is a closed, nonempty, measurable multifunction. (We use Proposition 2.3.2 and the fact that
\[ (t, v) \to g(t, v) := \rho_F(t, x_0(t), \dot{x}_0(t)) - |v - \dot{x}_0(t)| \]
is a Carathéodory function.) Under the hypotheses, \( t \to \rho_F(t, x_0(t), \dot{x}_0(t)) \) is integrable. Since \( \dot{x}_0 \) is integrable, \( v_1 \) is integrable too. We may therefore define \( x_1 \) according to
\[ x_1(t) := y(S) + \int_{S}^{t} v_1(s) ds. \]
Note that
\[ \rho_F(t, x_0(t), \dot{x}(t)) = 0 \quad \text{a.e.} \]

Again appealing to Theorem 2.3.11, we choose a measurable function \( v_2 \) satisfying
\[ v_2(t) \in F(t, x_1(t)) \quad \text{a.e.} \]
and
\[ |v_2(t) - \dot{x}_1(t)| = \rho_F(t, x_1(t), \dot{x}_1(t)) \quad \text{a.e.} \tag{2.16} \]
In view of the Lipschitz continuity properties of \( \rho_F(t, , , v) \) and since \( \dot{x}_1 \) is integrable, we readily deduce from the integrability of \( t \to \rho_F(t, x_0(t), \dot{x}_0(t)) \)
that \( t \to \rho_F(t, x_1(t), \dot{x}_1(t)) \) is also integrable. It then follows from (2.16) that \( v_2 \) is integrable and we may define
\[
x_2(t) = y(S) + \int_S^t v_2(s)ds.
\]
We proceed in this way to construct a sequence of absolutely continuous arcs \( \{x_m\} \) satisfying
\[
\rho_F(t, x_m(t), \dot{x}_{m+1}(t)) = 0 \quad \text{a.e.}
\]
\[
|\dot{x}_{m+1}(t) - \dot{x}_m(t)| = \rho_F(t, x_m(t), \dot{x}_m(t)) \quad \text{a.e.}
\]
for \( m = 0, 1, 2, \ldots \) and
\[
\rho_F(t, x_0(t), \dot{x}_0(t)) = \rho_F(t, y(t), \dot{y}(t)) \quad \text{a.e.}
\]
Notice that
\[
||x_1 - x_0||_{L^\infty} \leq \int_S^T |\dot{x}_1(t) - \dot{x}_0(t)|dt = \int_S^T \rho_F(t, x_0(t), \dot{x}_0(t))dt = \Lambda_F(y). \tag{2.17}
\]
Applying Proposition 2.4.1, we deduce that, for \( m \geq 1 \) and a.e. \( t \),
\[
|\dot{x}_{m+1}(t) - \dot{x}_m(t)| \leq \rho_F(t, x_{m-1}(t), \dot{x}_m(t)) + k(t)|x_m(t) - x_{m-1}(t)| = k(t)|x_m(t) - x_{m-1}(t)|. \tag{2.18}
\]
Since \( x_{m+1}(S) = x_m(S) \), it follows that, for all \( m \geq 1 \) and a.e. \( t \in [S, T] \),
\[
|x_{m+1}(t) - x_m(t)| \leq \int_S^t k(t_1)|x_m(t_1) - x_{m-1}(t_1)|dt_1 \leq \int_S^t k(t_1)\int_S^{t_1} k(t_2) \ldots \int_S^{t_{m-1}} k(t_m)|x_1(t_m) - x_0(t_m)|dt_m \ldots dt_2 dt_1 \leq S_m(t)\Lambda_F(y), \tag{2.19}
\]
in view of (2.17). Here
\[
S_m(t) := \int_S^t k(t_1)\int_S^{t_1} k(t_2) \ldots \int_S^{t_{m-1}} k(t_m)dt_m \ldots dt_2 dt_1.
\]
The right side can be reduced, one indefinite integral at a time, with the help of the integration by parts formula. There results:
\[
S_m(t) = \frac{\left(\int_S^t k(t)dt\right)^m}{m!}.
\]
It follows from (2.17) through (2.19) that, for any integers $M > N \geq 0$, we have

$$
||\dot{x}_M - \dot{x}_N||_{L^1} \\
\leq ||\dot{x}_M - \dot{x}_{M-1}||_{L^1} + \ldots + ||\dot{x}_{N+1} - \dot{x}_N||_{L^1} \\
\leq \left[ \frac{(\int_{S}^{T} k(t)dt)^{M-1}}{(M-1)!} + \ldots + \frac{(\int_{S}^{T} k(t)dt)^{N}}{N!} \right] \Lambda_F(y). \tag{2.20}
$$

(Here $\left(\int_{S}^{T} k(t)dt\right)^m / m! := 1$ when $m = 0$.) It is clear from this inequality that $\{\dot{x}_m\}$ is a Cauchy sequence in $L^1$. It follows that

$$\dot{x}_m \to v \text{ in } L^1,$$

for some $v \in L^1$. Define $x \in W^{1,1}$ according to

$$x(t) := \xi + \int_{S}^{t} v(s)ds.$$

Since

$$||x - x_m||_{L^\infty} \leq \int_{S}^{T} |v(s) - \dot{x}_m(s)|ds$$

and $\dot{x}_m \to v$ in $L^1$, we know that

$$x_m \to x \text{ uniformly.}$$

By extracting a subsequence (we do not relabel), we can arrange that

$$\dot{x}_m \to \dot{x} \text{ a.e.}$$

Define $O$ to be the subset of points $t \in [S, T]$ such that $\dot{x}_m(t) \in F(t, x_{m-1}(t))$ for all index values $m = 1, 2, \ldots$ and such that $\dot{x}_m(t) \to \dot{x}(t)$. Take any $t \in O$. Then

$$\dot{x}_m(t) \in F(t, x_{m-1}(t)).$$

Since $F(t, \cdot)$ has a closed graph, we obtain in the limit

$$\dot{x}(t) \in F(t, x(t)).$$

But $O$ has full measure. It follows that $x$ is an $F$-trajectory.

Next observe that by setting $N = 0$ in inequality (2.20) we arrive at

$$||\dot{x}_M - \dot{y}||_{L^1} \leq \exp \left( \int_{S}^{T} k(t)dt \right) \Lambda_F(y)$$
for $M = 1, 2, \ldots$. Since $\dot{x}_M \rightarrow \dot{x}$ in $L^1$ as $M \rightarrow \infty$ we deduce that

$$||\dot{x} - \dot{y}||_{L^1} \leq \exp \left( \int_S^T k(t) dt \right) \Lambda_F(y).$$

This is the required estimate.

It remains to prove that $x$ can be chosen continuously differentiable when the "comparison function" $y$ is continuously differentiable, under the additional hypotheses.

Construct a sequence $\{x_i\}$ as above. Under the additional hypotheses, $t \rightarrow F(t, x_0(t))$ is a continuous multifunction. In view of Proposition 2.4.2, $\dot{x}_1$ is a continuous function. ($\dot{x}_1(t)$ is the projection of $\dot{y}(t) = \dot{x}_0(t)$ onto $F(t, x_0(t))$.) Arguing inductively, we conclude that

$$t \rightarrow \dot{x}_i(t) \text{ is continuous for all } i \geq 0.$$ 

Fix any $i \geq 0$. For arbitrary $t$, $\dot{x}_{i+1}(t) \in F(t, x_i(t))$ and $\dot{x}_{i+2}(t)$ minimizes $v \rightarrow |\dot{x}_{i+1}(t) - v|$ over $v \in F(t, x_{i+1}(t))$, by construction. Under the hypotheses, we can find $w \in F(t, x_{i+1}(t))$ such that

$$|w - \dot{x}_{i+1}(t)| \leq k_\infty |x_{i+1}(t) - x_i(t)| \text{ for all } t.$$ 

Since $x_{i+2}(S) = x_{i+1}(S)$ and $i$ was chosen arbitrarily, we conclude that

$$|\dot{x}_{i+2}(t) - \dot{x}_{i+1}(t)| \leq k_\infty |x_{i+1}(t) - x_i(t)| \text{ for all } t.$$ 

Now Hypothesis (ii)' implies Hypothesis (ii) with $k(t) = k_\infty$ for all $t$. By (2.19) then, for any integers $M > N \geq 2$, we have

$$||\dot{x}_M - \dot{x}_N||_C \leq k_\infty \left[ \left( \int_S^T k_\infty dt \right)^{M-2} \left( M - 2 \right)! + \cdots + \frac{\left( \int_S^T k_\infty dt \right)^{N-2}}{(N-1)!} \right] \Lambda_F(y).$$

It follows that $\dot{x}_i$ is a Cauchy sequence in $C$. But $C$ is complete, so the sequence has a strong $C$ limit, some continuous function $v$. But $v$ must coincide with $\dot{x}$, the strong $L^1$ limit of $\{\dot{x}_i\}$, following adjustment on a nullset. It follows that $x$ is a continuously differentiable function. □

Naturally, if we specialize to the point-valued case, we recover an existence theorem for differential equations. In this important special case the solution is unique. We require

**Lemma 2.4.4 (Gronwall's Inequality)** Take an absolutely continuous function $z : [S, T] \rightarrow R^n$. Assume that there exist nonnegative integrable functions $k$ and $v$ such that

$$\left| \frac{d}{dt} z(t) \right| \leq k(t)|z(t)| + v(t) \text{ a.e. } t \in [S, T].$$


Then
\[ |z(t)| \leq \exp \left( \int_S^t k(\sigma) d\sigma \right) \left[ |z(S)| + \int_S^t \exp \left( -\int_S^\tau k(\sigma) d\sigma \right) v(\tau) d\tau \right] \]
for all \( t \in [S, T] \).

**Proof.** Since \( z \) is absolutely continuous so too is \( t \to |z(t)| \). Let \( O \subset [S, T] \) be the subset of points \( t \) such that \( z(\cdot) \) and \( |z(\cdot)| \) are both differentiable at \( t \). \( O \) has full measure and, it is straightforward to show,
\[ \frac{d}{dt} |z(t)| \leq |\dot{z}(t)| \quad \text{for all } t \in O. \]

Now define the absolutely continuous function
\[ \eta(t) := \exp \left( -\int_S^t k(\sigma) d\sigma \right) |z(t)| \quad \text{for all } t \in [S, T]. \]

Then for every \( t \in O \) we have
\[ \dot{\eta}(t) = \exp \left( -\int_S^t k(\sigma) d\sigma \right) \left[ \frac{d}{dt} |z(t)| - k(t)|z(t)| \right] \]
\[ \leq \exp \left( -\int_S^t k(\sigma) d\sigma \right) \left[ |\dot{z}(t)| - k(t)|z(t)| \right] \]
\[ \leq \exp \left( -\int_S^t k(\sigma) d\sigma \right) v(t). \]

It follows that for each \( t \in [S, T] \),
\[ \eta(t) \leq \eta(S) + \int_S^t \exp \left( -\int_S^\tau k(\sigma) d\sigma \right) v(\tau) d\tau. \]

Since \( \eta(S) = |z(S)| \) and
\[ |z(t)| = \exp \left( \int_S^t k(\sigma) d\sigma \right) \eta(t), \]
we deduce
\[ |z(t)| \leq \exp \left( \int_S^t k(\sigma) d\sigma \right) \left[ |z(S)| + \int_S^t \exp \left( -\int_S^\tau k(\sigma) d\sigma \right) v(\tau) d\tau \right]. \]

\( \Box \)

**Corollary 2.4.5 (ODEs: Existence and Uniqueness of Solutions)**
Take a function \( f : [S, T] \times \mathbb{R}^n \to \mathbb{R}^n \), an arc \( y \in W^{1,1}([S, T]; \mathbb{R}^n), \epsilon \in (0, \infty) \cup \{ +\infty \} \), and a point \( \xi \in \mathbb{R}^n \). (The case \( \epsilon = +\infty \) is interpreted as in the statement of Theorem 2.4.3.) Assume that
(i) \( f(\cdot, x) \) is measurable for each \( x \in \mathbb{R}^n \);

(ii) there exists \( k \in L^1 \) such that

\[
|f(t, x') - f(t, x'')| \leq k(t)|x' - x''|
\]

for all \( x', x'' \in y(t) + \epsilon B \), a.e. \( t \in [S, T] \).

Assume further that

\[
K \left( |\xi - y(S)| + \int_S^T |\dot{y}(t) - f(t, y(t))|dt \right) \leq \epsilon,
\]

where \( K := \exp \left( \int_S^T k(t)dt \right) \).

Then there exists a unique solution to the differential equation

\[
\begin{align*}
\dot{x}(t) &= f(t, x(t)) \quad \text{a.e. } t \in [S, T] \\
x(S) &= \xi
\end{align*}
\]

that satisfies

\[
\|x - y\|_{L^\infty} \leq \|x(S) - y(S)\| + \int_S^T |\dot{x}(t) - \dot{y}(t)|dt
\]

\[
\leq K \left( |\xi - y(S)| + \int_S^T |\dot{y}(t) - f(t, y(t))dt \right).
\]

**Proof.** All the assertions of the corollary follow immediately from the Generalized Filippov Existence Theorem (Theorem 2.4.3), with the exception of “uniqueness.” Suppose however that there are two solutions \( x' \) and \( x'' \) to the differential equation which satisfy \( ||x' - y||_{L^\infty} \leq \epsilon \) and \( ||x'' - y||_{L^\infty} \leq \epsilon \). Define \( z(t) := x'(t) - x''(t) \). Then under the hypotheses, for almost every \( t \in [S, T] \),

\[
|\dot{z}(t)| = |f(t, x'(t)) - f(t, x''(t))|
\]

\[
\leq k(t)|x'(t) - x''(t)| = k(t)|z(t)|.
\]

Since \( z(S) = 0 \), it follows from Gronwall’s Lemma that \( z \equiv 0 \). We conclude that \( x' = x'' \). Uniqueness is proved. \( \square \)

### 2.5 Perturbed Differential Inclusions

Consider a sequence of arcs whose elements satisfy perturbed versions of a “nominal” differential inclusion such that the perturbation terms in some
sense tend to zero as we proceed along the sequence. When can we extract a subsequence with limit a solution to the nominal differential inclusion? As we show, this is possible under unrestrictive hypotheses on the differential inclusion and on the nature of the perturbations. Necessary conditions in Nonsmooth Optimal Control are usually obtained by deriving necessary conditions for simpler perturbed versions of the Optimal Control problem of interest and passing to the limit. The significance of the results of this section is that they justify the limit-taking procedures.

Use is made of a characterization of subsets of $L^1$ that are relatively sequentially compact.

**Theorem 2.5.1 (Dunford–Pettis Theorem)** Let $S$ be a bounded subset of $L^1([S,T], \mathbb{R}^n)$. Then the following conditions are equivalent.

(i) Every sequence in $S$ has a subsequence converging to some $L^1$ function, with respect to the weak $L^1$ topology;

(ii) For every $\epsilon > 0$ there exists $\delta > 0$ such that for every measurable set $D \subset [S,T]$ and $x \in S$ satisfying $\text{meas} \{D\} < \delta$, we have $\int_D x(t)dt < \epsilon$.

**Proof.** See [61]. $\Box$

When Condition (ii) above is satisfied, we say “the family of functions $S$ is equi-integrable.” A simple criterion for equi-integrability (a sufficient condition to be precise) is that the family of functions is “uniformly integrably bounded” in the sense that there exists an integrable function $\alpha \in L^1$ such that

$$|x(t)| \leq \alpha(t) \quad \text{a.e. } t \in [S,T]$$

for all $x \in S$.

We require also certain properties of the Hamiltonian $H(t,x,p)$ associated with a given multifunction $F : [S,T] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, defined at points $(t,x,p) \in \text{dom} F \times \mathbb{R}^n$.

$$H(t,x,p) := \sup_{v \in F(t,x)} p \cdot v. \tag{2.21}$$

**Proposition 2.5.2** Consider a multifunction $F : [S,T] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, which has as values closed sets.

(a) Fix $(t,x) \in \text{dom} F$. Assume that there exists $c \geq 0$ such that $F(t,x) \subset cB$. Then

$$|H(t,x,p)| \leq c|p| \quad \text{for every } p \in \mathbb{R}^n,$n

and $H(t,x,.)$ is Lipschitz continuous with Lipschitz constant $c$. 
(b) Fix $t \in [S, T]$. Take convergent sequences $x_i \to x$ and $p_i \to p$ in $\mathbb{R}^n$. Assume that there exists $c \geq 0$ such that

$$(t_i, x_i) \in \text{dom } F \, \text{ and } \, F(t, x_i) \subset cB \, \text{ for all } \, i,$$

and that $\text{Gr } F(t, \cdot)$ is closed. Then

$$\limsup_{i \to \infty} H(t, x_i, p_i) \leq H(t, x, p).$$

(c) Take measurable functions $x : [S, T] \to \mathbb{R}^n$, $p : [S, T] \to \mathbb{R}^n$ such that $\text{Gr } x \subset \text{dom } F$. Assume that

$F$ is $\mathcal{L} \times \mathcal{B}^n$ measurable.

Then $t \to H(t, x(t), p(t))$ is a measurable function.

Proof.

(a): Take $(t, x) \in \text{dom } F$. Choose any $p, p' \in \mathbb{R}^n$. Since $F(t, x)$ is a compact, nonempty set, there exists $v \in F(t, x)$ such that

$$H(t, x, p) = p \cdot v.$$

Of course, $H(t, x, p') \geq p' \cdot v$. But then

$$H(t, x, p) - H(t, x, p') \leq (p - p') \cdot v \leq |v||p - p'|.$$

Since $|v| \leq c$, and the roles of $p$ and $p'$ can be interchanged, we deduce that

$$H(t, x, p) - H(t, x, p') \leq c|p - p'|.$$

Notice, in particular, that $|H(t, x, p)| \leq c|p|$, since $H(t, x, 0) = 0$.

(b): Fix $t \in [S, T]$ and take any sequences $x_i \to x$ and $p_i \to p$ in $\mathbb{R}^n$. $\limsup_i H(t, x_i, p_i)$ can be replaced by $\lim_i H(t, x_i, p_i)$, following extraction of a suitable subsequence (we do not relabel). For each $i$, $F(t, x_i)$ is a nonempty compact set (since $(t, x_i) \in \text{dom } F$); consequently, there exists $v_i \in F(t, x_i)$ such that $H(t, x_i, p_i) = p_i \cdot v_i$. But $|v_i| \leq c$ for $i = 1, 2, \ldots$. We can therefore arrange, by extracting another subsequence, that $v_i \to v$ for some $v \in \mathbb{R}^n$. Since $\text{Gr } F(t, \cdot)$ is closed, $v \in F(t, x)$ and so

$$\lim_i H(t, x_i, p_i) = \lim_i p_i \cdot v_i = p \cdot v \leq H(t, x, p).$$

(c): According to Proposition 2.3.2 and Theorem 2.3.11, we can find a measurable selection $v(.)$ of $t \sim F(t, x(t))$. Fix $k > 0$ and define for all $(t, x) \in [S, T] \times \mathbb{R}^n$,

$$F_k(t, x) := F(t, x) \cap (v(t) + kB)$$
and

\[ F_k(t, x) = \begin{cases} 
F_k(t, x) & \text{if } (t, x) \in \text{dom } F'_k \\
v(t) + kB & \text{otherwise}.
\end{cases} \]

It is a straightforward exercise to show that \( F_k \) is \( \mathcal{L} \times \mathcal{B}^n \) measurable. Since \( \text{dom } F_k = [S, T] \times R^n \), we can define

\[ H_k(t, x, p) := \max \{ p \cdot v : v \in F_k(t, x) \} \]

for all \((t, x, p) \in [S, T] \times R^n \times R^n\). Clearly

\[ H_k(t, x(t), p(t)) \to H(t, x(t), p(t)) \] as \( k \to \infty \),

for all \( t \). It therefore suffices to show that \( t \to H_k(t, x(t), p(t)) \) measurable.

However, by Corollary 2.3.3, this will be true if we can show that \( (t, (x, p)) \to H_k(t, (x, p)) \) is \( \mathcal{L} \times \mathcal{B}^{2n} \) measurable.

Fix \( r \in R \). We complete the proof by showing that

\[ D := \{ (t, x, p) \in [S, T] \times R^n \times R^n : H_k(t, x, p) \geq r \} \]

is \( \mathcal{L} \times \mathcal{B}^{2n} \) measurable. Let \( \{ v_i \} \) be a dense subset of \( R^n \). Take a sequence \( \epsilon \downarrow 0 \). It is easy to verify that \( D = D \), where

\[ D := \cap_{j=1}^{\infty} \cup_{i=1}^{\infty} (A_{i, j} \times B_{i, j}), \]

in which

\[ A_{i, j} := \{ (t, x) \in [S, T] \times R^n : (v_i + \epsilon_j B) \cap F_k(t, x) \neq \emptyset \} \]

and

\[ B_{i, j} := \{ p \in R^n : p \cdot v_i > r - \epsilon_j \}. \]

(That \( D \subset D \) is obvious; to show that \( D \subset D \), we exploit the fact that \( \text{Gr } F_k(t, .) \) is a compact set.)

But \( D \) is obtainable from \( \mathcal{L} \times \mathcal{B}^{2n} \) measurable sets, by means of a countable number of union and intersection operations. It follows that \( D \) is \( \mathcal{L} \times \mathcal{B}^{2n} \) measurable, as claimed. \( \square \)

With these preliminaries behind us we are ready to answer the question posed at the beginning of the section.

**Theorem 2.5.3 (Compactness of Trajectories)** Take a relatively open subset \( \Omega \subset [S, T] \times \mathbb{R}^n \) and a multifunction \( F : \Omega \rightrightarrows \mathbb{R}^n \).

Assume that, for some closed multifunction \( X : [S, T] \rightrightarrows \mathbb{R}^n \) such that \( \text{Gr } X \subset \Omega \), the following hypotheses are satisfied.

(i) \( F \) is a closed, convex, nonempty multifunction.

(ii) \( F \) is \( \mathcal{L} \times \mathcal{B}^n \) measurable.
(iii) For each $t \in [S,T]$, the graph of $F(t,.)$ restricted to $X(t)$ is closed.

Consider a sequence $\{x_i\}$ of $W^{1,1}([S,T];\mathbb{R}^n)$ functions, a sequence $\{r_i\}$ in $L^1([S,T];\mathbb{R})$ such that $\|r_i\|_{L^1} \to 0$ as $i \to \infty$, and a sequence $\{A_i\}$ of measurable subsets of $[S,T]$ such that $\text{meas} A_i \to |T - S|$ as $i \to \infty$.

Suppose that:

(iv) $\text{Gr} \ x_i \subset \text{Gr} \ X$ for all $i$;

(v) $\{\dot{x}_i\}$ is a sequence of uniformly integrably bounded functions on $[S,T]$ and $\{x_i(S)\}$ is a bounded sequence;

(vi) there exists $c \in L^1$ such that

$$F(t, x_i(t)) \subset c(t)B$$

for a.e. $t \in A_i$ and for $i = 1, 2, \ldots$.

Suppose further that

$$\dot{x}_i(t) \in F(t, x_i(t)) + r_i(t)B \quad \text{a.e. } t \in A_i.$$

Then along some subsequence (we do not relabel)

$$x_i \to x \text{ uniformly and } \dot{x}_i \to \dot{x} \text{ weakly in } L^1$$

for some $x \in W^{1,1}([S,T];\mathbb{R}^n)$ satisfying

$$\dot{x}(t) \in F(t, x(t)) \quad \text{a.e. } t \in [S,T].$$

**Proof.** The $\dot{x}_i$s are uniformly integrably bounded on $[S,T]$. According to the Dunford–Pettis Theorem, we can arrange, by extracting a subsequence (we do not relabel), that $\dot{x}_i \to v$ weakly in $L^1$ for some $L^1$ function $v$. Since $\{x_i(S)\}$ is a bounded sequence we may arrange by further subsequence extraction that $x_i(S) \to \xi$ for some $\xi \in \mathbb{R}^n$. Now define

$$x(t) := \xi + \int_S^t v(s)ds.$$

By weak convergence, $x_i(t) \to x(t)$ for every $t \in [S,T]$. Clearly too $\dot{x}_i \to \dot{x}$ weakly in $L^1$.

Now consider the Hamiltonian $H(t, \xi, p)$ defined by (2.21). Choose any $p$ and any Lebesgue measurable subset $V \subset [S,T]$. For almost every $t \in V \cap A_i$, we have

$$H(t, x_i(t), p) \geq p \cdot \dot{x}_i(t) - r_i(t)|p|.$$ 

Since all terms in this inequality are integrable, we deduce

$$\int_{V \cap A_i} p \cdot \dot{x}_i(t)dt - \int_{V \cap A_i} r_i(t)|p|dt \leq \int_{V \cap A_i} H(t, x_i(t), p)dt.$$
Because the $\dot{x}$'s are uniformly integrably bounded, $\text{meas}[A_i] \to |T - S|$, $\dot{x} \to \dot{x}$ weakly in $L^1$, and $\|\tau_i\|_{L^1} \to 0$, we see that the left side of this relationship has limit $\int_V p \cdot \dot{x}(t)dt$. It follows that

$$\int_V p \cdot \dot{x}(t)dt \leq \limsup_{i \to \infty} \int_V \chi_i(t)H(t, x_i(t), p)dt.$$ 

Here $\chi_i$ denotes the characteristic function of the set $A_i$.

Since $\chi_i(t)H(t, x_i(t), p)$ is bounded above by $c(t)|p|$, we deduce from Fatou's Lemma that

$$\int_V p \cdot \dot{x}(t) \leq \int_V \limsup_i \chi_i(t)H(t, x_i(t), p)dt.$$ 

From the upper semicontinuity properties of $H$ then (see Proposition 2.5.2)

$$\int_V \left( H(t, x(t), p) - p \cdot \dot{x}(t) \right) dt \geq 0.$$ 

Let $\{p_i\}$ be an ordering of the set of $n$-vectors having rational coefficients. Define $D \subset [S, T]$ to be the subset of points $t \in [S, T]$ such that $t$ is a Lebesgue point of $t \to H(t, x(t), p_i) - p_i \cdot \dot{x}(t)$ for all $i$. $D$ is a set of full measure. For any $t \in D \cap [S, T]$ and any $i$

$$H(t, x(t), p_i) - p_i \cdot \dot{x}(t) = \lim_{\delta \downarrow 0} \frac{1}{\delta} \int_t^{t+\delta} [H(\sigma, x(\sigma), p_i) - p_i \cdot \dot{x}(\sigma)] d\sigma \geq 0.$$ 

Since $H(t, x(t), \cdot)$ is continuous for each $t$, it follows that

$$\sup \{ p \cdot e - p \cdot \dot{x}(t) : e \in F(t, x(t)) \} \geq 0 \quad \text{for all } p \in R^n \quad \text{a.e. } t \in [S, T].$$

But $F(t, x(t))$ is closed and convex. We deduce with the help of the separation theorem that

$$\dot{x}(t) \in F(t, x(t)) \quad \text{a.e. } t \in [S, T].$$

We have confirmed that $x$ is an $F$-trajectory.

2.6 Existence of Minimizing $F$-Trajectories

Take a relatively open subset $\Omega \subset [S, T] \times R^n$, a multifunction $F : \Omega \rightrightarrows R^n$, a closed multifunction $X : [S, T] \rightrightarrows R^n$ with the property that $\text{Gr} X \subset \Omega$, and a closed set $C \subset R^n \times R^n$. We define the set of feasible $F$-trajectories (associated with the constraint sets $X(t)$, $S \leq t \leq T$, and $C$) to be

$$\mathcal{R}_F(X, C) := \{ x \in C([S, T]; R^n) : x \text{ is an } F\text{-trajectory} \quad x(t) \in X(t) \text{ for all } t \in [S, T] \text{ and } (x(S), x(T)) \in C \}.$$
We deduce from the results of the previous section the following criteria for compactness of the set of feasible $F$-trajectories.

**Proposition 2.6.1** Take $\Omega$, $F$, $X$, and $C$ as above. Assume that

(i) $F$ is a closed, $\mathcal{L} \times \mathcal{B}^n$ measurable multifunction;

(ii) for each $t \in [S,T]$, the graph of $F(t,.)$ restricted to $X(t)$ is closed;

(iii) there exist $\alpha \in L^1$ and $\beta \in L^1$ such that

$$F(t,x) \subset (\alpha(t)|x| + \beta(t))B \text{ for all } (t,x) \in \text{Gr } X;$$

(iv) either $X(s)$ is bounded for some $s \in [S,T]$ or one of the following two sets

$$C_0 := \{x_0 \in R^n : (x_0,x_1) \in C \text{ for some } x_1 \in R^n \}$$

$$C_1 := \{x_1 \in R^n : (x_0,x_1) \in C \text{ for some } x_0 \in R^n \}$$

is bounded. Assume further that

(v) $F(t,x)$ is convex for all $(t,x) \in \text{Gr } X$.

Then $\mathcal{R}_F(X,C)$ is compact with respect to the supremum norm topology.

**Proof.** Since the supremum norm topology is a metric topology, it suffices to prove sequential compactness. Accordingly, take any sequence of feasible $F$-trajectories $\{x_i\}$. We must show that there exists an $F$-trajectory $x$ satisfying the constraints $x(t) \in X(t)$ for all $t \in [S,T]$ and $(x(S),x(T)) \in C$ such that

$$x_i \to x \text{ uniformly}$$

along some subsequence. But these conclusions can be drawn from Theorem 2.5.3 provided we can show that the set $\mathcal{R}_F(X,C)$ is bounded with respect to the supremum norm. By Hypothesis (iv) however there exists $k > 0$ and $\bar{s} \in [S,T]$ such that for any feasible $F$-trajectory $y$ we have

$$|y(\bar{s})| \leq k.$$

By Hypothesis (iii),

$$|\dot{y}(t)| \leq \alpha(t)|y(t)| + \beta(t) \text{ a.e.}$$

It follows from the Gronwall Lemma (applied "backwards" in time on the interval $[S,\bar{s}]$ and "forwards" on $[\bar{s},T]$) that

$$|y(t)| \leq K \text{ for all } t \in [S,T],$$

where

$$K = e^{\|\alpha\|_{L^1}} (k + \|\beta\|_{L^1}).$$
We have confirmed that \( \mathcal{R}(X, C) \) is bounded with respect to the supremum norm. \( \square \)

It is a simple step now to supply conditions for existence of solutions to the optimal control problem

\[
(P) \begin{cases}
\text{Minimize } g(x(S), x(T)) \quad \text{over } x \in W^{1,1}([S,T]; \mathbb{R}^n) \\
\text{which satisfy} \\
\quad \dot{x}(t) \in F(t, x(t)) \quad \text{a.e. } t \in [S,T], \\
\quad x(t) \in X(t) \quad \text{for all } t \in [S,T], \\
\quad (x(S), x(T)) \in C,
\end{cases}
\]

in which \( g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) is a given lower semicontinuous function.

Indeed \((P)\) can be equivalently formulated as a problem of seeking a minimizer of a lower semicontinuous function over a compact subset of \( C([S,T]; \mathbb{R}^n) \) equipped with the supremum norm, namely,

\[
\text{minimize } \psi(y) \quad \text{over } y \in \mathcal{R}_F(X, C),
\]

where

\[
\psi(y) := g(y(S), y(T)).
\]

\((P)\) therefore has a minimizer provided \( \mathcal{R}_F(X, C) \) is nonempty. We have proved:

**Proposition 2.6.2** Take \( \Omega, F, X, C, \) and \( g \) as above. Assume that

(i) \( F \) is a closed, \( \mathcal{L} \times \mathcal{B}^n \) measurable multifunction;

(ii) for each \( t \in [S,T] \), the graph of \( F(t,.) \) restricted to \( X(t) \) is closed;

(iii) there exist \( \alpha \in L^1 \) and \( \beta \in L^1 \) such that

\[
F(t,x) \subseteq (\alpha(t)|x| + \beta(t))B \quad \text{for all } (t,x) \in \text{Gr } X;
\]

(iv) either \( X(s) \) is bounded for some \( s \in [S,T] \) or one of the following two sets

\[
\begin{align*}
C_0 & := \{ x_0 \in \mathbb{R}^n : (x_0,x_1) \in C \text{ for some } x_1 \in \mathbb{R}^n \} \\
C_1 & := \{ x_1 \in \mathbb{R}^n : (x_0,x_1) \in C \text{ for some } x_0 \in \mathbb{R}^n \}
\end{align*}
\]

is bounded.

Assume further that

(a) the set of feasible \( F \)-trajectories \( \mathcal{R}_F(X, C) \) is nonempty;

(b) \( F(t,x) \) is convex for each \( (t,x) \in \text{Gr } X \).

Then \((P)\) has a minimizer.
2.7 Relaxation

Suppose that an optimization problem of interest fails to have a minimizer. “Relaxation” is the procedure of adding extra elements to the domain of the optimization problem to ensure existence of minimizers.

For a relaxation scheme to be of interest it usually needs to be accompanied by the information that an element \( \bar{x} \) in the extended domain can be approximated by an element \( y \) in the original domain of the optimization problem (to the extent that we can arrange that the cost of \( y \) is arbitrarily close to that of \( \bar{x} \)). In these circumstances, we can find a suboptimal element for the original problem (i.e., one whose cost is arbitrarily close to the infimum cost) by finding a minimizer in the extended domain and approximating it.

Relaxation is now examined in connection with the optimization problem (\( P \)) of the preceding section, which for convenience we reproduce.

\[
\begin{align*}
(P) \quad \text{Minimize } g(x(S), x(T)) & \text{ over } x \in W^{1,1}([S, T]; \mathbb{R}^n) \text{ which satisfy } \\
\dot{x}(t) & \in F(t, x(t)) \quad \text{a.e. } t \in [S, T], \\
x(t) & \in X(t) \quad \text{for all } t \in [S, T], \\
(x(S), x(T)) & \in C.
\end{align*}
\]

We impose the hypotheses of Proposition 2.6.2 with the exception of the convexity hypothesis

\[ F(t, x) \text{ is convex for all } (t, x) \in \text{Gr } X. \]

In these circumstances (\( P \)) may fail to have a minimizer. The point is illustrated by the following example.

**Example 2.1** Consider problem (\( P \)) when the state vector \( x \) is 2-dimensional. We write the components of the state vector \( x = (y, z) \). Set

\[
[S, T] = [0, 1], \quad \Omega = [0, 1] \times \mathbb{R}^2, \quad X(t) = \mathbb{R}^2, \\
g((y_0, z_0), (y_1, z_1)) = z_1, \\
C = \{((y_0, z_0), (y_1, z_1)) : z_0 = 0\}, \\
F(y, z) := (-1) \cup (+1) \times \{|y|\}.
\]

Notice that, if an arc \((y, z)\) satisfies the constraints, then

\[ z(t) = \int_0^t |y(s)|ds. \quad (2.22) \]

Evidently then this special case of (\( P \)) is a disguised version, arrived at
through state augmentation, of the optimization problem

\[
(P)' \quad \begin{cases} 
\text{Minimize } J(y) = \int_0^1 |y(s)| ds \\
\text{over } y \in W^{1,1}([0, 1]; \mathbb{R}) \text{ satisfying} \\
\dot{y}(t) \in \{-1\} \cup \{+1\} \text{ a.e.}
\end{cases}
\]

(If \((y, z)\) is feasible for \((P)\) then \(y\) and \(z\) are related by (2.22), \(y\) is feasible for \((P)'\), and the costs are the same; also, if \(y\) is feasible for \((P)'\) then \((y, z)\), with \(z\) given by (2.22), is feasible for \((P)\) and again the costs are the same.)

However, as we show, \((P)'\) has no minimizers. It follows that, in this case, \((P)\) fails to have a solution. The fact that all the hypotheses of Proposition 2.6.2 are satisfied with the exception of (ii) confirms that we cannot dispense with the convexity hypothesis.

To show that \((P)'\) does not have a solution, notice first of all that

\[ J(y) \geq 0 \quad (2.23) \]

for all arcs \(y\) satisfying the constraint of \((P)'\). Consider next the sequence of feasible arcs \(\{y_i\}\),

\[ y_i(t) = \int_0^t v_i(s) ds, \]

where

\[ v_i(s) = \begin{cases} 
+1 & \text{for } s \in A_i \cap [0, 1] \\
-1 & \text{for } s \notin A_i \cap [0, 1]
\end{cases} \]

and

\[ A_i = \bigcup_{j=0}^\infty [(2i)^{-1}2j, (2i)^{-1}(2j+1)]. \]

An easy calculation yields:

\[ J(y_i) = 2^{-(i+1)} \text{ for } i = 1, 2, \ldots. \]

Since \(J(y_i) \to 0\) as \(i \to \infty\) we conclude from (2.23) that the infimum cost is 0. If there exists a minimizer \(\tilde{y}\) then

\[ J(\tilde{y}) = \int_0^1 |\tilde{y}(s)| ds = 0. \]

This implies that \(\tilde{y} \equiv 0\). It follows that

\[ \dot{\tilde{y}}(t) = 0 \text{ a.e.} \]

But then \(\tilde{y}\) fails to satisfy the differential inclusion

\[ \dot{\tilde{y}}(t) \in \{-1\} \cup \{+1\} \text{ a.e.} \]

It follows that no minimizer exists.
The pathological feature of the above problem is that the limit point \( y \) of any minimizing sequence satisfies only the convexified differential inclusion

\[
\dot{y}(t) \in \text{co} \ (\{-1\} \cup \{+1\}) \quad \text{a.e.}
\]

In light of this example, a natural relaxation procedure for us to adopt is to allow arcs that satisfy the convexified differential inclusion

\[
\dot{x}(t) \in \text{co} F(t, x(t)) \quad \text{a.e.} \quad t \in [S, T].
\]  

(2.24)

Accordingly, an arc \( x \in W^{1,1} \) satisfying this dynamic constraint is called a "relaxed" \( F \)-trajectory. When it is necessary to emphasize the distinction with relaxed \( F \)-trajectories, we sometimes call \( F \)-trajectories "ordinary" \( F \)-trajectories.

As earlier discussed, for this concept of relaxed trajectory to be useful we need to know that arcs satisfying the "relaxed" dynamic constraint (2.24) can be adequately approximated by arcs satisfying the original, unconvexified, constraint. That this can be done is a consequence of the Generalized Filippov Existence Theorem and the following theorem of R. J. Aumann on integrals of multifunctions. (See [27].)

**Theorem 2.7.1 (Aumann's Theorem)** Take a Lebesgue measurable multifunction \( \Gamma : [S, T] \rightarrow R^n \) that is closed and nonempty. Assume that there exists \( c \in L^1 \) such that

\[
\Gamma(t) \in c(t)B \quad \text{for all} \quad t \in [S, T].
\]

Then

\[
\int_S^T \Gamma(s)ds = \int_S^T \text{co} \Gamma(s)ds,
\]

where

\[
\int_S^T A(s)ds := \left\{ \int_S^T \gamma(s)ds : \gamma \text{ is a Lebesgue measurable selection of } A \right\}
\]

for \( A = \Gamma \) and \( A = \text{co} \Gamma \).

The ground has now been prepared for:

**Theorem 2.7.2 (Relaxation Theorem)** Take a relatively open set \( \Omega \subset [S, T] \times R^n \) and an \( \mathcal{L} \times \mathcal{B}^n \) measurable multifunction \( F : \Omega \rightarrow R^n \) that is closed and nonempty. Assume that there exist \( k \in L^1 \) and \( c \in L^1 \) such that

\[
F(t, x') \subset F(t, x'') + k(t)B \quad \text{for all} \quad (t, x'), (t, x'') \in \Omega
\]

and

\[
F(t, x) \subset c(t)B \quad \text{for all} \quad (t, x) \in \Omega.
\]
Take any relaxed $F$-trajectory $x$ with $\text{Gr} x \subset \Omega$ and any $\delta > 0$. Then there exists an ordinary $F$-trajectory $y$ that satisfies $y(S) = x(S)$ and
\[
\max_{t \in [S,T]} |y(t) - x(t)| < \delta.
\]

**Proof.** Choose $\epsilon > 0$ such that $T(x, 2\epsilon) \subset \Omega$ and let $\alpha$ be such that
\[
0 < \alpha < \min \left\{ \frac{\epsilon}{K \ln K}, \epsilon, \frac{\delta}{1 + K \ln K} \right\},
\]
where $K := \exp(||k||_{L^1})$. Let $h > 0$ be such that
\[
\int_I c(t)dt < \alpha/2
\]
for any subinterval $I \subset [S, T]$ of length no greater than $h$.

Let \{\(S = t_0, t_1, \ldots, t_k = T\)\} be a partition of $[S, T]$ such that $\text{meas} I_i < h$ for $i = 1, 2, \ldots, k$, where $I_i := [t_{i-1}, t_i)$ for $i = 1, \ldots, k - 1$ and $I_k = [t_{k-1}, t_k]$. The multifunction $t \mapsto F(t, x(t))$ is Lebesgue measurable (see Proposition 2.3.2) and satisfies
\[
F(t, x(t)) \subset c(t)B \text{ for all } t \in [S, T].
\]

Recalling that $x$ is a co-$F$-trajectory, we deduce from Aumann’s Theorem (Theorem 2.7.1) that there exist measurable functions $f_i : [S, T] \to \mathbb{R}^n$ such that $f_i(t) \in F(t, x(t))$ a.e. $t \in I_i$ and
\[
\int_{I_i} f_i(t)dt = \int_{I_i} \dot{x}(t)dt
\]
for $i = 1, \ldots, k$. Define
\[
f(t) := \sum_{i=1}^{k} f_i(t) \chi_{I_i}(t),
\]
where $\chi_{I_i}$ denotes the indicator function of $I_i$ and set
\[
z(t) = x(S) + \int_{S}^{t} f(s)ds \text{ for } t \in [S, T].
\]

Fix $t \in [S, T]$. Then for some $j \in \{1, \ldots, k\}$,
\[
|z(t) - x(t)| = \left| \sum_{i=1}^{j-1} \int_{I_i} (f_i(\sigma) - \dot{x}(\sigma))d\sigma + \int_{I_j \cap [S,t]} (f_j(\sigma) - \dot{x}(\sigma))d\sigma \right|
\leq 0 + 2 \int_{I_j} c(t)dt < \alpha.
\]
It follows that
\[ ||z - x||_{L^\infty} < \alpha. \]  
(2.25)

Since \( \alpha < \epsilon \), we have
\[ T(z, \epsilon) \subset \Omega. \]

Notice that, since \( \dot{x}(t) \in F(t, x(t)) \) a.e. and in view of Proposition 2.4.1,
\[ \rho_F(t, z(t), \dot{z}(t)) \leq \rho_F(t, x(t), \dot{x}(t)) + k(t)|x(t) - z(t)| \]
\[ < k(t)\alpha \quad \text{a.e.} \]

We have
\[ \Lambda_F(z) := \int_S^T \rho_F(t, z(t), \dot{z}(t))dt < \alpha \ln K. \]

Since \( \alpha K \ln K < \epsilon \), we deduce from the Generalized Filippov Existence Theorem that there exists an \( F \)-trajectory \( y \) such that \( y(S) = x(S) \) and
\[ ||y - z||_{L^\infty} \leq K \Lambda_F(z) < \alpha K \ln K. \]

By (2.25) then
\[ ||y - x||_{L^\infty} \leq ||y - z||_{L^\infty} + ||z - x||_{L^\infty} \leq \alpha(K \ln K + 1). \]

Since however
\[ \alpha(K \ln K + 1) < \delta, \]

we conclude that the \( F \)-trajectory \( y \) satisfies
\[ ||y - x||_{L^\infty} < \delta. \]

\( \square \)

The optimization problem obtained when we replace the dynamic constraint \( \dot{x} \in F \) in \( (P) \) by \( \dot{x} \in \text{co} F \) is denoted \( (P)_{\text{relaxed}} \). Minimizers for \( (P)_{\text{relaxed}} \) are called relaxed minimizers for \( (P) \).

The following proposition provides information clarifying the relationship between \( (P) \) and \( (P)_{\text{relaxed}} \).

**Proposition 2.7.3** Take \( \Omega, F, X, C, \) and \( g \) as above. Assume that

(i) \( F \) is a compact, \( \mathcal{L} \times \mathcal{B}^n \) measurable multifunction;

(ii) for each \( t \in [S, T] \), the graph of \( F(t, .) \) restricted to \( X(t) \) is closed;

(iii) there exist \( \alpha \in L^1 \) and \( \beta \in L^1 \) such that
\[ F(t, x) \subset (\alpha(t)|x| + \beta(t))B \quad \text{for all} \quad (t, x) \in \text{Gr} X; \]
(iv) either $X(s)$ is bounded for some $s \in [S, T]$ or one of the following two sets

\[
C_0 := \{x_0 \in R^n : (x_0, x_1) \in C \text{ for some } x_1 \in R^n \}
\]

\[
C_1 := \{x_1 \in R^n : (x_0, x_1) \in C \text{ for some } x_0 \in R^n \}
\]

is bounded.

Assume further that the set of feasible $F$-trajectories $\mathcal{R}_F(X, C)$ is nonempty. Then $(P)_{\text{relaxed}}$ has a minimizer.

If, in addition, we assume that

(a) there exists $k \in L^1$ such that

\[
F(t, x) \subset F(t, x') + k(t)|x - x'|B \text{ for all } (t, x), (t, x') \in \Omega,
\]

(b) $g$ is continuous,

and, for some relaxed minimizer $\tilde{x}$ and $\epsilon > 0$,

(c)\[\tilde{x}(t) + \epsilon B \subset X(t) \text{ for all } t \in [S, T],\] \tag{2.26}

(d)\[\text{either } (\tilde{x}(S) + \epsilon B) \times \{\tilde{x}(T)\} \subset C \text{ or } \{\tilde{x}(S)\} \times (\tilde{x}(T) + \epsilon B) \subset C,\] \tag{2.27}

then

\[\inf (P)_{\text{relaxed}} = \inf (P).\]

The right and left sides of the last relationship denote the infimum cost for $(P)$ and $(P)_{\text{relaxed}}$, respectively.

**Proof.** Existence of a minimizer for $(P)_{\text{relaxed}}$ follows immediately from Proposition 2.6.2 applied to the modified version of $(P)$ in which $\text{co } F$ replaces $F$. (Notice that $\text{co } F$ inherits the measurability properties of $F$ according to Proposition 2.3.8, as well as the linear growth properties.)

Suppose that there exists a relaxed minimizer $\tilde{x}$ and $\epsilon > 0$ such that

\[
\tilde{x}(t) + \epsilon B \subset X(t) \text{ for all } t \in [S, T]
\]

and

\[
\{\tilde{x}(S)\} \times (\tilde{x}(T) + \epsilon B) \subset C
\]

(The case $(\tilde{x}(S) + \epsilon B) \times \{\tilde{x}(T)\} \subset C$ is treated by "reversing time".) Take any $\alpha > 0$. Then noting the continuity of the function $g$ and applying the
Relaxation Theorem, Theorem 2.7.2, we can find an ordinary $F$-trajectory $\bar{y}$ such that
\[ \bar{y}(S) = \bar{x}(S), \|\bar{y} - \bar{x}\|_{L^\infty} < \epsilon \]
and
\[ g(\bar{y}(S), \bar{y}(T)) < g(\bar{x}(S), \bar{x}(T)) + \alpha. \]
Clearly $\bar{y}$ satisfies the constraints
\[ (\bar{y}(S), \bar{y}(T)) \in C \text{ and } \bar{y}(t) \in X(t) \text{ for all } t \in [S, T]. \]
In other words, $\bar{y}$ is a feasible (ordinary) $F$-trajectory. It follows that
\[ \inf(P)_{\text{relaxed}} = g(\bar{x}(S), \bar{x}(T)) > g(\bar{y}(S), \bar{y}(T)) - \alpha \geq \inf(P) - \alpha. \]
Since $\alpha > 0$ is arbitrary and
\[ \inf (P)_{\text{relaxed}} \leq \inf (P), \]
we conclude that
\[ \inf (P)_{\text{relaxed}} = \inf (P). \]
\[ \square \]

Notice the crucial role of the “interiority” hypotheses (2.26) and (2.27). If all relaxed minimizers violate these hypotheses then the Relaxation Theorem does not automatically imply that the infimum costs of $(P)$ and $(P)_{\text{relaxed}}$ coincide; while it is true that any relaxed minimizer $\bar{x}$ can be uniformly approximated by an $F$-trajectory $\bar{y}$, we cannot in general guarantee that $y$ will satisfy the constraints for it to qualify as a feasible $F$-trajectory.

2.8 The Generalized Bolza Problem

In Section 2.6 we gave conditions for the existence of minimizers in the context of minimization problems over some class of arcs satisfying a given differential inclusion $\dot{x}(t) \in F(t, x(t))$. These conditions restricted attention to problems for which the velocity sets $\dot{F}(t, x)$ are bounded. For traditional variational problems and also for many Optimal Control problems of interest there are no constraints on permitted velocities. To deal in a unified manner with problems with bounded and unbounded velocity sets, it is convenient to adopt a new framework for the optimization problems involved, namely, to regard them as special cases of the Generalized Bolza Problem:

\[ (GBP) \left\{ \begin{array}{l}
\text{Minimize } \Lambda(x) := l(x(S), x(T)) + \int_S^T L(t, x(t), \dot{x}(t)) dt \\
\text{over arcs } x \in W^{1,1}([S, T]; \mathbb{R}^n),
\end{array} \right. \]
in which \([S, T]\) is a given interval, and \(l : R^n \times R^n \to R \cup \{+\infty\}\) and \(L : [S, T] \times R^n \times R^n \to R \cup \{+\infty\}\) are given extended-valued functions. Provided we arrange that \(t \to L(t, x(t), \dot{x}(t))\) is measurable and minorized by an integrable function for every \(x \in W^{1,1}\) (our hypotheses take care of this), \(\Lambda\) will be a well-defined \(R \cup \{+\infty\}\) valued functional on \(W^{1,1}\).

Notice that the functions \(l\) and \(L\) are permitted to takes values \(+\infty\). So they can be used implicitly to take account of constraints. For example, the “differential inclusion” problem

\[
\begin{aligned}
&\text{Minimize } g(x(S), x(T)) \\
&\text{over arcs } x \in W^{1,1}([S, T]; R^n) \text{ satisfying} \\
&\dot{x}(t) \in F(t, x(t)), \\
&(x(S), x(T)) \in C
\end{aligned}
\]

is a special case of the Generalized Bolza Problem in which

\[
L(t, x, v) = \begin{cases} 
0 & \text{if } v \in F(t, x) \\
+\infty & \text{otherwise}
\end{cases}
\]

and

\[
l(x_0, x_1) = \begin{cases} 
g(x_0, x_1) & \text{if } (x_0, x_1) \in C \\
+\infty & \text{otherwise}
\end{cases}
\]

In existence theorems covering problems with unbounded velocity sets, superlinear growth hypotheses on the cost integrand are typically invoked to compensate for unbounded velocity sets. The key advantage of the Generalized Bolza Problem as a vehicle for such theorems is that unrestricted hypotheses ensuring existence of minimizers, which require coercivity of the cost integrand in precisely those “directions” in which velocities are unconstrained, can be economically expressed as conditions on the extended-valued function \(L\).

**Theorem 2.8.1 (Generalized Bolza Problem: Existence of Minimizers)** Assume that the data for (GBP) satisfy the following hypotheses:

(H1): \(l\) is lower semicontinuous and there exist a lower semicontinuous function \(l^0 : R^+ \to R\) satisfying

\[
\lim_{r \to +\infty} l^0(r) = +\infty
\]

and either

\[
l^0(|x_0|) \leq l(x_0, x_1) \quad \text{for all } x_0, x_1
\]

or

\[
l^0(|x_1|) \leq l(x_0, x_1) \quad \text{for all } x_0, x_1.
\]

(H2): \(L\) is \(\mathcal{L} \times B^{n \times n}\) measurable.

(H3): \(L(t, \ldots)\) is lower semicontinuous for each \(t \in [S, T]\).
(H4): For each \((t, x) \in [S, T] \times R^n\), \(L(t, x, .)\) is convex and \(\text{dom } L(t, x, .) \neq \emptyset\).

(H5): For all \(t \in [S, T]\), \(x \in R^n\), and \(v \in R^n\),

\[L(t, x, v) \geq \theta(|v|) - \alpha|x|,\]

for some \(\alpha \geq 0\) and some lower semicontinuous convex function \(\theta : R^+ \to R^+\) satisfying

\[\lim_{r \uparrow +\infty} \theta(r)/r = +\infty.\]

Then (GBP) has a minimizer. (We allow the possibility that \(\Lambda(x) = +\infty\) for all \(x \in W^{1,1}\). In this case all arcs \(x\) are regarded as minimizers.)

Comment

The proof of this theorem, which follows shortly, exploits properties of the Hamiltonian

\[H(t, x, p) := \sup_{v \in R^n} \{p \cdot v - L(t, x, v)\}.\]

To a large extent, the role of the growth condition (H5) is to ensure that the Hamiltonian has the required properties to furnish existence theorems. Loosely speaking, growth conditions on the Lagrangian translate into (one-sided) boundedness conditions on the Hamiltonian. One direction for generalizing this theorem is to replace (H5) by less restrictive conditions imposed directly on the Hamiltonian. Rockafellar has shown that the conclusions of Theorem 2.8.1 remain valid when (H5) is replaced by:

(H5') For all \(t \in [S, T]\) and \(x, p \in R^n\),

\[H(t, x, p) \leq \mu(t, p) + |x|(\sigma(t) + \rho(t)|p|),\]

for some integrable functions \(\sigma(.), \rho(.)\), and some function \(\mu(., .)\) such that \(\mu(., p)\) is integrable for each \(p \in R^n\).

The ensuing analysis calls upon some properties of convex functions, which it is now convenient to summarize. We say that a convex function \(f : R^n \to R \cup \{+\infty\}\) is proper if it is lower semicontinuous and \(\text{dom } f \neq \emptyset\). The conjugate of \(f\) is the function \(f^* : R^n \to R \cup \{+\infty\}\), defined by the Legendre-Fenchel transformation:

\[f^*(y) = \sup_{x \in R^n} \{y \cdot x - f(x)\}.\]

Important facts are that, for any proper, convex function \(f : R^n \to R \cup \{+\infty\}\), its conjugate \(f^*\), too, is proper, convex. Furthermore, \(f\) can be recovered from \(f^*\) by means of a second application of the Legendre-Fenchel transformation:

\[f(x) = \sup_{y \in R^n} (y \cdot x - f^*(y)).\]

One consequence of these relationships is
Proposition 2.8.2 (Jensen’s Inequality) Take any proper, convex function \( f : R^n \to R \cup \{+\infty\} \). Then, for any set \( I \subset R \) of positive measure and any \( v \in L^1(I; R^n) \), \( t \to f(v(t)) \) is a measurable function, minorized by an integrable function, and

\[
\int_I f(v(t))dt \geq |I| f \left( |I|^{-1} \int_I |v(t)|dt \right),
\]

where \( |I| \) denotes the Lebesgue measure of \( I \).

**Proof.** Take any \( y \in \text{dom} \ f^* \). Then, for each \( t \in I \),

\[
f(v(t)) \geq v(t) \cdot y - f^*(y).
\]

It follows that the (measurable) function \( t \to f(v(t)) \) is minorized by an integrable function. Also,

\[
\int_I f(v(t))dt \geq \left( \int_I v(t)dt \right) \cdot y - f^*(y)|I|
\]

\[
= |I| \left[ \left( |I|^{-1} \int_I v(t)dt \right) \cdot y - f^*(y) \right].
\]

This inequality is valid for all \( y \in \text{dom} \ f^* \). Maximizing over \( \text{dom} \ f^* \) and noting that \( f \) is obtained from \( f^* \) by applying the Legendre–Fenchel transformation, we obtain

\[
\int_I f(v(t))dt \geq |I| f \left( |I|^{-1} \int_I |v(t)|dt \right),
\]

as claimed. \( \square \)

**Proof of Theorem 2.8.1** We assume that \( l^0(x_0) \leq l(x_0, x_1) \) for all \((x_0, x_1)\). (The case \( l^0(x_1) \leq l(x_0, x_1) \) for all \((x_0, x_1)\) is treated similarly.) Under the hypotheses, \( l^0 \) and \( \theta \) are bounded below. By scaling and adding a constant to \( \int Ldt + l \) (this does not effect the minimizers) we can arrange that \( l^0 \geq 0 \) and \( \theta \geq 0 \). We can also arrange that the constant \( \alpha \) is arbitrarily small.

Choose \( \alpha \) such that \( e^{\alpha |T - S| |T - S|} \alpha < 1 \).

Since \( \theta \) has superlinear growth, we can define \( k : R^+ \to R^+ \):

\[
k(\beta) := \sup \{ r \geq 0 : r = 0 \text{ or } \theta(r) \leq \beta r \}.
\]

**Step 1:** Fix \( M \geq 0 \). We show that the level set

\[
S_M := \{ x \in W^{1,1} : \Lambda(x) \leq M \}
\]
is weakly sequentially precompact; i.e., any sequence \( \{x_t\} \) in \( S_M \) has a subsequence that converges, with respect to the weak \( W^{1,1} \) topology, to some point in \( W^{1,1} \).

Take any \( x \in S_M \) and define the \( L^1 \) function

\[
q(t) := L(t, x(t), \dot{x}(t)).
\]

Then

\[
|\dot{x}(t)| \leq k(1) + \theta(|\dot{x}(t)|) \leq k(1) + q(t) + \alpha |x(t)| \quad \text{a.e.} \quad (2.28)
\]

But, for each \( t \in [S, T] \),

\[
\int_S^t q(s)ds = \Lambda(x) - \int_t^T q(s)ds \leq M - \int_t^T \theta(|\dot{x}(s)|)ds + \alpha \int_t^T |x(s)|ds
\]

\[
\leq M + \alpha |T - S| |x(t)| + \int_t^T (\alpha |T - S| |\dot{x}(s)| - \theta(|\dot{x}(s)|)) ds
\]

\[
\leq M + \alpha |T - S| |x(t)| + \alpha |T - S|^2 k(\alpha |T - S|).
\]

It follows from (2.28) and Gronwall's Inequality that

\[
|x(t)| \leq e^{\alpha (t-S)} \left[ |x(S)| + \int_S^T (k(1) + q(s))ds \right]
\]

\[
\leq e^{\alpha (T-S)} \left[ |x(S)| + k(1)|T - S| + M + \alpha |T - S|^2 k(\alpha |T - S|) + \alpha |T - S| |x(t)| \right].
\]

Therefore

\[
|x(t)| \leq A |x(S)| + B, \quad (2.29)
\]

where the constants \( A \) and \( B \) (they do not depend on \( x \)) are

\[
A := \left( 1 - \alpha |T - S| e^{\alpha |T - S|} \right)^{-1} e^{\alpha |T - S|}
\]

and

\[
B := \left( 1 - \alpha |T - S| e^{\alpha |T - S|} \right)^{-1} (k(1)|T - S| + M + \alpha |T - S|^2 k(\alpha |T - S|)).
\]

We deduce from (2.29) and the fact that \( l(x_0, x_1) \geq l^0(|x_0|) \) that

\[
|x(S)| \leq K, \quad (2.30)
\]

where \( K > 0 \) is any constant (it can be chosen independent of \( x \)) such that

\[
l^0(r) - \alpha [Ar + B] > M \quad \text{for all } r \geq K.
\]
Now, for any set \( I \subset [S, T] \) of positive measure, Jensen’s inequality yields
\[
\theta(|I|^{-1} \int_I |\dot{x}(t)|dt) \leq |I|^{-1} \int_I \theta(|\dot{x}(t)|)dt \\
\leq |I|^{-1} \int_S^T \theta(|\dot{x}(t)|)dt \\
\leq |I|^{-1} \left( \int_S^T L(t, x(t), \dot{x}(t))dt + \alpha \int_S^T |x(t)|dt \right) \\
\leq |I|^{-1}(M + \alpha AK + \alpha B).
\]
We conclude that, if \( \int_I |\dot{x}(t)|dt > 0 \),
\[
\frac{\theta(\int_I |\dot{x}(t)|dt/|I|)}{\int_I |\dot{x}(t)|dt/|I|} \leq \frac{(M + \alpha AK + \alpha B)}{\int_I |\dot{x}(t)|dt}.
\]
Since \( \theta \) has superlinear growth, it follows from this inequality that there exists a function \( \omega : R^+ \to R^+ \) (which does not depend on \( x \)) such that \( \lim_{\sigma \to 0} \omega(\sigma) = 0 \) and
\[
\int_I |\dot{x}(t)|dt \leq \omega(|I|) \quad \text{for all measurable } I \subset [S, T]. \tag{2.31}
\]
Take any sequence \( \{x_i\} \) in \( S_M \). Then, by (2.30), \( \{x_i(S)\} \) is a bounded sequence. On the other hand, \( \{\dot{x}_i\} \) is an equicontinuous sequence, by (2.31). Invoking the Dunford–Pettis criterion for weak sequential compactness in \( L^1 \), we deduce that, along a subsequence,
\[
x_i(S) \to x(S) \quad \text{and} \quad \dot{x}_i \to \dot{x}_i \text{ weakly in } L^1.
\]
Otherwise expressed,
\[
x_i \to x \quad \text{weakly in } W^{1,1},
\]
for some \( x \in W^{1,1} \). This is what we set out to prove.

**Step 2:** Take an \( L \times B^n \) function \( \phi : [S, T] \times R^n \to R \cup \{+\infty\} \) that satisfies the conditions:

(a): for each \( t \in [S, T] \), \( \phi(t, \cdot) \) is lower semicontinuous and finite at some point, and

(b): for some \( \tilde{p} \in L^\infty \), the function \( t \to \phi(t, \tilde{p}(t)) \) is minorized by an integrable function.

We show that
\[
\int_S^T \phi(t)dt = \sup_{p(\cdot) \in L^\infty} \int_S^T \phi(t, p(t))dt, \tag{2.32}
\]
where
\[ \hat{\phi}(t) := \sup_{p \in R^n} \phi(x, p). \]
(Note that, under the hypotheses, \( \hat{\phi} \) is measurable and minorized by an integrable function. So the left side of (2.32) is well defined. The right side is interpreted as the supremum of the specified integral over \( p(.) \)s such that the integrand is minorized by some integrable function.)

For any \( p(.) \in L^\infty, \hat{\phi}(t) \geq \phi(t, p(t)) \) for all \( t \). It immediately follows that (2.32) holds, when \( \geq \) replaces \( = \).

It suffices then to validate (2.32) when \( \leq \) replaces \( = \). To this end, choose any \( r \in R^n \) such that
\[ \int_S^T \hat{\phi}(t) dt > r. \]

We can also choose \( K > 0 \) and \( \epsilon > 0 \) such that, writing
\[ \check{\phi}(t) := \min \{ \hat{\phi}(t), K \}, \]
we have
\[ \int_S^T (\check{\phi}(t) - \epsilon) dt > r. \]
Define the multifunction
\[ \Gamma(t) := \{ p \in R^n : \phi(t, p) > \check{\phi}(t) - \epsilon \}. \]
Under the hypotheses, \( \Gamma \) takes values nonempty (open) sets and \( \text{Gr } \Gamma \) is \( \mathcal{L} \times \mathcal{B}^n \) measurable. According to Aumann’s Measurable Selection Theorem then, \( \Gamma \) has a measurable selection, which we write \( \bar{p}(.) \).

However, since \( t \to \phi(t, \bar{p}(t)) \) and \( t \to \phi(t, \bar{p}(t)) \) are minorized by integrable functions, we can find a measurable set \( E \) such that \( \bar{p} \) restricted to \( [S, T] \setminus E \) is essentially bounded and
\[ \int_{[S, T] \setminus E} \phi(t, \bar{p}(t)) dt + \int_E \phi(t, \bar{p}(t)) dt > r. \]
It follows that
\[ \int_S^T \phi(t, p(t)) dt > r, \]
in which \( p(.) \) is the essentially bounded function
\[ p(t) := \begin{cases} \bar{p}(t) & \text{if } t \in I \\ \bar{p} & \text{otherwise.} \end{cases} \]
Since \( r \) is an arbitrary strict lower bound on \( \int_S^T \hat{\phi}(t) dt \), the desired inequality is confirmed.
Step 3: We show that \( \Lambda(.) \) is weakly sequentially lower semicontinuous (w.r.t. the \( W^{1,1} \) topology).

Since weak \( L^1 \) convergence implies uniform convergence, we deduce from the lower semicontinuity of \( \lambda \) that \( x \to \lambda(x(S), x(T)) \) is weakly sequentially lower semicontinuous. It remains therefore to show that

\[
\tilde{\Lambda}(x) := \int_s^T L(t, x(t), \dot{x}(t))dt
\]

is also weakly sequentially lower semicontinuous. Take any \( x \in W^{1,1} \). Then, since \( L(t, x(t),) \) is a proper, convex function for each \( t \),

\[
\tilde{\Lambda}(x) = \int_s^T \sup_{p \in R^n} [p \cdot \dot{x}(t) - H(t, x(t), p)] dt
\]

\[
= \sup_{p(.) \in L^\infty} \int_s^T [p(t) \cdot \dot{x}(t) - H(t, x(t), p(t))] dt.
\]

(We have used the results of Step 2 to justify the last equality. Note that the function \( \phi(t, p) = p \cdot \dot{x}(t) - H(t, x(t), p) \) satisfies the relevant hypotheses. In particular, \( \phi(t, \bar{p}(t)) \) is minorized by an integrable function for the choice \( \bar{p} \equiv 0 \).)

For fixed \( p \in L^\infty \), consider now the integral functional

\[
\Lambda_p(x) := \int_s^T [p(t) \cdot \dot{x}(t) - H(t, x(t), p(t))] dt.
\]

We claim that \( \Lambda_p \) is weakly sequentially lower semicontinuous.

To verify this assertion, take any weakly convergent sequence \( y_i \to y \) in \( W^{1,1} \). Then \( \dot{y}_i \to \dot{y} \) weakly in \( L^1 \) and \( y_i \to y \) uniformly. Since \( H(t, \ldots, p(t)) \) is upper semicontinuous for each \( t \) and the \( H(t, y_i(t), p(t)) \)'s are minorized by a common integrable function (this last property follows from Hypothesis (H5)), we deduce from Fubini's Theorem that

\[
\liminf_{i \to \infty} \Lambda_p(y_i) = \liminf_{i \to \infty} \int_s^T [p(t) \cdot \dot{y}_i(t) - H(t, y_i(t), p(t))] dt
\]

\[
\geq \int_s^T [p(t) \cdot \dot{y}(t) - H(t, y(t), p(t))] dt
\]

\[
= \Lambda_p(y).
\]

Weak sequential lower semicontinuity of \( \Lambda_p \) is confirmed.

We have shown that, for each \( x \in W^{1,1} \),

\[
\tilde{\Lambda}(x) = \sup_{p(.) \in L^\infty} \Lambda_p(x).
\]

But the upper envelope of a family of weakly sequentially lower semicontinuous functionals on \( W^{1,1} \) is also weakly sequentially lower semicontinuous. It follows that \( \tilde{\Lambda} \) is weakly sequentially lower semicontinuous.
Conclusion: We have shown in Steps 1 and 3 that $\Lambda$ is sequentially lower semicontinuous and that the level sets of $\Lambda$ are sequentially compact with respect to the weak $W^{1,1}$ topology. These properties guarantee existence of a minimizer. (We allow the possibility that $\Lambda(x) = +\infty$ for all $x$. In this case all $x$s are minimizers.) □

2.9 Notes for Chapter 2

A standard reference for properties of measurable multifunctions is [27]. More recent texts covering the subject matter of Section 2.3 are [9], [12], and [125].

There is a substantial literature on differential inclusions. Material in Sections 2.4 through 2.6, which restricts attention largely to differential inclusions involving multifunctions that are Lipschitz continuous with respect to the state, only scratches the surface of this extensive field. (See [9] and [57].) The important existence theorem, Theorem 2.3.3, and accompanying estimates are essentially due to Filippov. The proof given here is an adaptation of that in [9], to allow for measurable time dependence. The Compactness of Trajectories Theorem (Theorem 2.4.3) is from [38]. Implicit in this theorem are early ideas for establishing existence of optimal controls under a convexity hypothesis on the “velocity set” associated with Tonelli, L. C. Young, Filippov, Gramkrelidze, Roxin, and others.

Relaxation is a much broader topic than is conveyed by Section 2.7, which focuses on relaxation schemes for Optimal Control problems formulated in terms of a differential inclusion. Relaxation schemes for many other classes of optimization problems have been proposed including, for example, variational problems in several independent variables [5] and Optimal Control problems with time delay [126]. Of particular importance are those introduced by Warga for Optimal Control problems with dynamic constraints in the form of a differential equation parameterized by control functions [147] and those introduced by Gamkrelidze [76].

Tonelli’s techniques for proving existence of minimizers to variational problems with unbounded derivatives [135], based on establishing weak lower semicontinuity of integral functionals and weak compactness of level sets, underlie much research in this area to this day. Far-reaching extensions of Tonelli’s existence theorems were achieved by Rockafellar. (See, for example, [122].) An accessible treatment of existence issues in the Calculus of Variations and Optimal Control is to be found in [88].