A Uniqueness Theorem for Differential Inclusions

PETER R. WOLENSKI

Centre de Recherche Mathématiques, Université de Montreal,
C.D. 6128 Succursale A, Montreal, Quebec, Canada H3C 3J7

Communicated by Jack K. Hale

Received November 23, 1988

The reachable set multifunction of a Lipschitz differential inclusion is characterized in terms of a semigroup property and a type of set differentiation at \( t = 0 \). This generalizes the classical uniqueness theorem of ordinary differential equations. A notion of an escape time for differential inclusions is introduced and plays a major role in the main result. © 1990 Academic Press, Inc.

1. INTRODUCTION

Consider the differential inclusion

\[
\begin{align*}
\dot{x}(\cdot) & \in AC[0, T] \\
\dot{x}(t) & \in F(x(t)) \quad \text{a.e.} \quad t \in [0, T] \\
x(0) & = \xi,
\end{align*}
\]

where \( T > 0 \), \( F: \mathbb{R}^n \rightrightarrows \mathbb{R}^n \) is a given multifunction (or set-valued map), \( \xi \in \mathbb{R}^n \), and \( \dot{x}(t) \) is the derivative of \( x(\cdot) \). An expository account of the basic theory of differential inclusions can be found, for example, in Aubin and Cellina [1] or Clarke [2]. Of fundamental importance to a wide range of control and optimal control problems is the reachable set \( R^{(T)}(\xi) \), defined by

\[
R^{(T)}(\xi) = \{ x(T): x(\cdot) \text{ satisfies } (1.1) \}.
\]

The goal of this paper is a characterization of the collection of multifunctions \( \{ R^{(T)} \}_{T \geq 0} \).

The genesis of the main result theorem 3.2 below is the observation that, for \( s, t \geq 0 \), and \( \xi \in \mathbb{R}^n \), one has

\[
R^{(s+t)}(\xi) = \bigcup_{\eta \in R^{(s)}(\xi)} R^{(t)}(\eta) = (R^{(s)} \circ R^{(t)})(\xi).
\]
This simple consequence of the definition (1.2) illustrates that the collection of multifunctions \( \{ R^{(t)} \}_{t \geq 0} \) exhibits a semigroup property, where the binary operation is the composition of multifunctions. As in all semigroup theory, some natural questions arise: What is the infinitesimal generator? Can the generator be recovered from the semigroup by some type of "differentiation" at \( t = 0 \)? Under what conditions can semigroups be uniquely determined by these properties? This paper tackles such questions with a systematic treatment for a broad class of differential inclusions.

Although the semigroup property (1.3) is quite trivial, the natural questions mentioned above have not heretofore been adequately answered. A notable exception is contained in Roxin [6]. In Section 3, after the explicit statement of our main result, we will discuss Roxin's paper in detail. This will allow for a more direct comparison between our results and those of [6].

As a short preview, we can say that our main result Theorem 3.2 is a uniqueness theorem that characterizes the collection \( \{ R^{(T)} \}_{T \geq 0} \) when \( F \) is locally Lipschitz with compact convex values. In addition to resolving the questions from the abstract semigroup perspective we have outlined, one can view this result as a generalization of the classical uniqueness theorem for ordinary differential equations. We next explain how this is so.

The theory of ordinary differential equation is of course subsumed under the more general theory of differential inclusions. In (1.1), suppose \( F(\eta) = \{ f(\eta) \} \), where \( f: \mathbb{R}^n \to \mathbb{R}^n \) is, for simplicity, a globally Lipschitz function defined on \( \mathbb{R}^n \). Then (1.1) becomes simply the o.d.e.

\[
\begin{align*}
\dot{x}(t) &= f(x(t)), \quad t \in [0, T] \\
x(0) &= \xi.
\end{align*}
\]

The classical result is that (1.4) admits a unique solution \( x(t) = r^{(t)}(\xi) \) on \([0, T]\) for all \( \xi \in \mathbb{R}^n \). For our purposes, it is convenient to state this in the following equivalent manner. There is only one collection of functions \( \{ r^{(t)} \}_{t \geq 0} \) defined on \( \mathbb{R}^n \) for which the following hold for all \( \xi \in \mathbb{R}^n \):

\[
 r^{(s+t)}(\xi) = (r^{(s)} \circ r^{(t)})(\xi)
\]

and

\[
\frac{d}{dt} r^{(t)}(\xi) \big|_{t = 0} = f(\xi).
\]

It is this latter form of the o.d.e. uniqueness theorem that can be generalized to differential inclusions. To realise this, however, one should be slightly cautious in interpreting (1.5) and (1.6). For differential inclusions, the trajectories of (1.1) are quite distinct from the reachable set mul-
tifunction \( R^{(T)}(\cdot): \mathbb{R}^n \to \mathbb{R}^n \) that is defined via these trajectories, whereas in o.d.e. theory, the distinction between trajectories of (1.4) (which are functions on \([0, T]\)) and the endpoints of these trajectories (which are elements in \(\mathbb{R}^n\)) can be easily blurred. Our uniqueness theorem for differential inclusions is not in terms of the trajectories of (1.1), but rather in terms of the collection \( \{R^{(T)}\}_{T \geq 0} \) of reachable set multifunctions.

The differential inclusion semigroup property (1.3) is directly analogous to the property (1.5). What is the differential inclusion analogue of (1.6)? Will it insure a uniqueness property? A viable candidate to generalizing (1.6) is the fact that

\[
\text{dist}_H \left( \frac{R^{(T)}(\xi) - \xi}{t}, F(\xi) \right) \to 0 \quad \text{as} \quad t \downarrow 0
\]

holds for all \( \xi \) (where \( \text{dist}_H \) denotes the Hausdorff distance). Recently Frankowska has proved and used (1.7) in [4] to show that \( F \) is the "infinitesimal generator" of the semigroup. However, as Example 4.3 below illustrates, the pointwise convergence property (1.7) is not sufficient to uniquely characterize the semigroup \( \{R^{(T)}\}_{T \geq 0} \). Rather, to obtain the uniqueness result, one must assume that the limit in (1.7) is taken uniformly over \( \xi \) in a compact set.

Let us briefly return again to the o.d.e. (1.4). The property (1.7) (where \( R^{(T)}(\xi) = \{r^{(T)}(\xi)\} \)), with the limit taken either pointwise or uniformly on a compact set, is equivalent to (1.6). The classical uniqueness theorem says that the collection \( \{r^{(T)}\}_{T \geq 0} \) is the only set of functions with this property that also satisfies (1.5). With differential inclusions, to say that \( \{R^{(T)}\}_{T \geq 0} \) are the only multifunctions satisfying (1.3) and (1.7), one must take the limit uniformly in (1.7), even if the values of \( F \) are singletons. Again see Example 4.3.

The o.d.e. uniqueness theorem requires merely a local Lipschitz assumption on an open set. For a more complete result, the concept of an escape time is introduced to capture the entire time interval for which a trajectory exists. These added features are also incorporated into our differential inclusion uniqueness result. We introduce the appropriate modification of an escape time in Section 2.

For simplicity we only prove our result in autonomous form. That is, the values of \( F \) in (1.1) do not explicitly depend on \( t \). This greatly simplifies the notation. We state the nonautonomous analogues in Section 10. The proofs of these requires only straightforward modifications of the ones we give, though they become notationally cumbersome, and therefore are omitted.

A listing of the section titles reveals the plan for the rest of the paper.

1. Introduction
2. Notation and Definitions
2. NOTATION AND DEFINITIONS

The notation is fairly standard. The letter $B$ always denotes the closed unit ball of $\mathbb{R}^n$. If $A \subseteq \mathbb{R}^n$, we denote the closure of $A$ by $\text{cl } A$. For $T > 0$, the spaces of absolutely continuous and continuously differentiable function on $[0, T]$ are denoted by $AC[0, T]$ and $C^1[0, T]$, respectively. The norm $\| \cdot \|$ will always denote the sup norm on $[0, T]$, unless otherwise noted. It will be clear from the context which $T > 0$ is being used in the norm.

If $A \subseteq \mathbb{R}^n$ and $a \in \mathbb{R}^n$, the distance from $a$ to $A$ is defined by $\text{dist}(a, A) = \inf \{ \| a - a' \| : a' \in A \}$. For two nonempty compact subsets $A_0$ and $A_1$ of $\mathbb{R}^n$, the Hausdorff distance between $A_0$ and $A_1$ is defined by $\text{dist}_H(A_0, A_1) = \inf \{ \delta : A_0 \subseteq A_1 + \delta B, A_1 \subseteq A_0 + \delta B \}$.

For a multifunction $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we write $\text{dom } G$ for the set $\{ \xi : G(\xi) \neq \emptyset \}$. If $X \subseteq \text{dom } G$ and $G$ has compact values on $X$, then $G$ is said to be Lipschitz (of order $\lambda > 0$) on $X$ if $\text{dist}_H(G(\xi), G(\xi')) \leq \lambda |\xi - \xi'|$ for all $\xi, \xi' \in X$. $G$ is locally Lipschitz on $X$ if $G$ is Lipschitz on each compact subset of $X$.

Suppose $\{ G_t \}_{t > 0}$ and $G$ are all compact-valued multifunctions. Then we say that $G_t$ converges to $G$ as $t \downarrow 0$ uniformly on compact subsets of $X$ if $X \subseteq \text{dom } G$ and for all compact $K \subseteq X$ and $\epsilon > 0$, there exists $t_0 > 0$ such that for all $\xi \in K$ and $0 < t \leq t_0$, we have $K \subseteq \text{dom } G_t$ and $\text{dist}_H(G_t(\xi), G(\xi)) < \epsilon$. This directly generalizes to multifunctions the usual definition of uniform convergence of functions.

Again consider the differential inclusion (1.1). We have already defined the reachable set $R^{(T)}(\xi)$ in (1.3). We use the notation $R^{(\leq T)}(\xi)$ to denote the set $\bigcup_{0 \leq t \leq T} R^{(t)}(\xi)$. It is also convenient to label the trajectories of (1.1) by defining $S^{(T)}(\xi) = \{ x(\cdot) : x(\cdot) \text{ satisfies (1.1)} \}$. The only new concept introduced in this section is the following: Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $X \subseteq \text{dom } F$ is open, and $\xi \in X$. We define the escape time $T_x(\xi)$ from $X$ (with base point $\xi$) by

$$T_x(\xi) := \sup \{ T : \text{cl } R^{(\leq T)}(\xi) \text{ is compact in } X \}. \quad (2.1)$$
One can easily check that the definition (2.1) extends the classical o.d.e. definition of an escape time if (1.1) is representable as an o.d.e. in the form (1.4). The basic assumptions on \( F \) introduced in the next section will render "cl" in (2.1) superfluous.

3. Statement of the Main Result

Let \( X \) be an open subset of \( \mathbb{R}^n \). Throughout the rest of the paper, with the exceptions of Section 10 and the examples in Section 4, we make the following basic assumptions on a fixed multifunction \( F: \mathbb{R}^n \to \mathbb{R}^n \):

\[
\text{dom } F = X, \\
F(\xi) \text{ is compact and convex for all } \xi \in X, \text{ and} \\
F \text{ is locally Lipschitz on } X.
\]

The first theorem is a collection of information about reachable sets, most of which is known or trivial. The second theorem states that this information uniquely determines reachable sets, hence we label it the uniqueness theorem.

**Theorem 3.1.** The following hold:

(a) for each \( \xi \in X \) and \( 0 \leq T < T_X(\xi) \), \( R^{(T)}(\xi) \) is nonempty and compact,

(b) for each compact \( K \subseteq X \), we have \( \inf_{\xi \in K} T_X(\xi) > 0 \),

(c) for all \( s, t \geq 0 \) and \( \xi \in X \), \( R^{(s+t)}(\xi) = R^{(s)}(R^{(t)}(\xi)) \), and

(d) the multifunctions \( \{(1/t)(R^{(t)} - I)\}_{t > 0} \) converge to \( F \) as \( t \downarrow 0 \) uniformly on compact subsets of \( X \).

In part (d), "\( I \)" refers to the multifunction that takes \( \xi \) into the set \( \{\xi\} \). Hence for \( \xi \in X \), \( (1/t)(R^{(t)} - I)(\xi) \) is defined as the set \( \{(1/t)(\eta - \xi) : \eta \in R^{(t)}(\xi)\} \).

**Theorem 3.2 (Uniqueness Theorem).** Suppose real numbers \( \{T_G(\xi)\}_{\xi \in X} \) and multifunctions \( \{G^{(t)}\}_{t \geq 0} \) are given so that (a)-(d) of Theorem 3.1 hold where \( T_G \) and \( G^{(t)} \) replace \( T_X \) and \( R^{(t)} \), respectively. Then for all \( \xi \in X \), we have

(i) for all \( T \geq 0 \), \( R^{(T)}(\xi) \subseteq \text{cl } G^{(T)}(\xi) \), and

(ii) for all \( 0 \leq T < T_X(\xi) \), \( R^{(T)}(\xi) = G^{(T)}(\xi) \).

As mentioned in the Introduction, Theorem 3.2 needs to be compared
with the results in Roxin [6]. In [5, 6], Roxin developed an axiomatic approach to the study of control systems. A collection of sets \( \{ G^t(\xi) \} \) is parametrized by time \( t \) and state vector \( \xi \) and is assumed to satisfy a list of axioms. All of the axioms are properties of reachable sets for a globally Lipschitz differential inclusion (in this situation, all escape times are \( +\infty \)). Properties (a) and (c) of Theorem 3.1 are among these axioms, but also included are "uniform" continuity assumptions of the multifunctions \( \xi \Rightarrow G^t(\xi) \) and \( \xi \Rightarrow G^t(\xi) \). The conclusion of [6, Theorem 7.1] is that if the limsup of \( (1/t)(G^t(\xi) - \xi) \) as \( t \downarrow 0 \) equals \( F(\xi) \), and \( F \) is globally Lipschitz on \( \mathbb{R}^n \), then \( G^t(\xi) = R^t(\xi) \) for all \( \xi \in \mathbb{R}^n \) and \( t \geq 0 \). Hence the conclusion of Roxin's result and our Theorem 3.2 is the same.

The major contribution of our results here are twofold. First, we localize the problem to arbitrary open sets and replace Lipschitz by local Lipschitz; this requires the introduction of escape times. Second, we make no a priori continuity assumptions on \( G \). The latter makes our proof considerably more difficult. Roxin uses uniform continuity to prove the uniform convergence in Theorem 3.1(d) (in his more restricted setting), and then uses (d) and the uniform continuity again to show \( G^t(\xi) = R^t(\xi) \). In summary, continuity plays a prominent role in his proof.

Why are we adamant in avoiding continuity assumptions? In addition to generalizing Roxin's results, the continuity absence will be greatly significant in a future paper [9], where some of the techniques used here are applied to generalized problems of Lagrange. In the Lagrange problems, one must necessarily handle certain unbounded sets, the properties of which are not sufficiently captured by the employment of a metric. This makes the corresponding continuity assumptions awkward and difficult to handle. Since globally Lipschitz differential inclusions are subsumed in the theory of generalized Lagrange problems, it is significant that continuity assumptions can be avoided. We refer the reader to [7] or [9] for further details.

4. Examples

Examples 4.1. This first example elucidates the uniform convergence property of (d) in the special case when \( F \) is merely a linear mapping.

Suppose \( A \) is an \( n \times n \) matrix, \( X = \mathbb{R}^n \), and \( F(\xi) = \{ A\xi \} \) for all \( \xi \in \mathbb{R}^n \). Then \( R^T(\xi) = \{ e^{TA}\xi \} \), where \( e^{TA} \) is the exponential of the matrix \( TA \). Note that

\[
\left| \frac{R^t(\xi) - \xi}{t} - A\xi \right| \leq \left\| \frac{e^{tA} - I}{t} - A \right\| \cdot |\xi|, \tag{4.1}
\]

\[
\left| \frac{R^t(\xi) - \xi}{t} - A\xi \right| \leq \left\| \frac{e^{tA} - I}{t} - A \right\| \cdot |\xi|, \tag{4.1}
\]

\[
\left| \frac{R^t(\xi) - \xi}{t} - A\xi \right| \leq \left\| \frac{e^{tA} - I}{t} - A \right\| \cdot |\xi|, \tag{4.1}
\]
UNIQUENESS THEOREM FOR DIFFERENTIAL INCLUSIONS

where \( \| \cdot \| \) is matrix norm. As \( t \downarrow 0 \), the matrix norms approach zero, and thus the values on the left side of (4.1) approach zero uniformly over \( \xi \) in a compact set.

**Example 4.2.** This next example reveals the importance of the local Lipshitz assumption on \( F \). Without some structure on \( F \) stronger than continuity, this example shows some of the difficulties in characterizing the reachable set with only a semigroup property and a "differentiation" at \( t = 0 \).

Let \( X = \mathbb{R}^1 \) and \( F(x) = \{ (\text{sgn } x) \sqrt{|x|} \} \). Then

\[
R^{(t)}(\xi) = \left\{ \frac{1}{t} (\text{sgn } \xi) (t + 2 \sqrt{|\xi|}) \right\}^2 \quad \text{if } \xi \neq 0
\]

\[
\left[ \frac{- \frac{1}{2} t^2, \frac{1}{2} t^2} \right] \quad \text{if } \xi = 0
\]

(the closed brackets refer to a closed interval). If we define

\[
G^{(t)}(\xi) = \begin{cases} R^{(t)}(\xi) & \text{if } \xi \neq 0 \\ \left[ - \frac{1}{2} t^2, \frac{1}{2} t^2 \right] & \text{if } \xi = 0, \end{cases}
\]

then \( \{ G^{(t)} \} \) satisfies (a)-(d) of Theorem 3.1 (with \( T_\xi(\xi) = + \infty \) for all \( \xi \)). Obviously \( F \) is not locally Lipschitz around \( \xi = 0 \).

**Example 4.3.** This third example shows that a pointwise convergence in (d) may hold for a semigroup that is not the reachable set.

Let \( X = \mathbb{R}^1 \) and \( F(x) = \{ -x \} \). Then \( R^{(t)}(\xi) = \{ e^{-t} \xi \} \). Define

\[
G^{(t)}(\xi) = \begin{cases} R^{(t)}(\xi) & \text{if } \xi \leq 1 \text{ or } t < \ln \xi \\ [0, e^{-t} \xi] & \text{if } \xi > 1 \text{ and } t \geq \ln \xi. \end{cases}
\]

Then \( \{ G^{(t)} \} \) is a semigroup where \( (G^{(t)}(\xi) - \xi) / t \to F(\xi) \) pointwise; however, this convergence is not uniform around \( \xi = 1 \).

**Example 4.4.** We next show how the reachable set may fail to be closed in \( X \) beyond the escape time. Open brackets refer to coordinate in \( \mathbb{R}^2 \), and closed brackets denote a closed interval in \( \mathbb{R}^1 \).

Let \( X = \mathbb{R}^2 \setminus \{(1, 1)\} \), and for \( (x, y) \in X \), define

\[
F(x, y) = [ |x|, \max\{|x|, 1\} ] \times [ |y|, \max\{|y|, 1\} ].
\]

Then if \( 0 \leq t < 1 \), we have \( R^{(t)}(0, 0) = [0, t] \times [0, t] \). But whenever \( t \leq 1 \) ( = \( T_x(0, 0) \)), \( R^{(t)}(0, 0) \) equals \([0, e^{t-1}] \) minus the set \( \{(r, r); 1 \leq r \leq e^{t-1}\} \). Hence \( R^{(t)}(0, 0) \) is not closed in \( X \) when \( t > 1 \).

**Example 4.5.** Here is an example where semigroups can behave differently than reachable sets after escape times.
Let $X = \mathbb{R}^1 \setminus \{1\}$ and $F(x) = [0, 1]$ for all $x \in X$. Then

$$R^{(t)}(\xi) = \begin{cases} [\xi, \xi + t] & \text{if } \xi + t < 1 \text{ or } \xi > 1 \\ [\xi, 1) & \text{if } \xi + t \geq 1 \text{ and } \xi < 1, \end{cases}$$

Define

$$G^{(t)}(\xi) = \begin{cases} [\xi, \xi + t] & \text{if } \xi + t < 1 \text{ or } \xi > 1 \\ [\xi, \xi + t) \setminus \{1\} & \text{if } \xi + t \geq 1 \text{ and } \xi < 1. \end{cases}$$

One can easily verify that (a)–(d) of Theorem 3.1 holds for $\{G^{(t)}\}_{t \geq 0}$ and $\{T_0(\xi)\}_{\xi \in X}$, where $T_0(\xi) := T_x(\xi)$.

**Example 4.6.** Finally, we offer two examples where semigroups can behave differently at infinity. Let $X = \mathbb{R}^1$ and $F(x) = [|x|, 1 + x^2]$ for $x \in X$. Then

$$R^{(t)}(\xi) = \begin{cases} \left[\frac{e^{it}}{\sin(t + \tan^{-1}(\xi))}, \tan(t + \tan^{-1}(\xi))\right] & \text{if } t < \pi/2 - \tan^{-1}\xi \\ \left[\frac{e^{it}}{\sin(t + \tan^{-1}(\xi))}, \infty\right) & \text{otherwise.} \end{cases}$$

Let

$$G^{(t)}(\xi) = \begin{cases} R^{(t)}(\xi) & \text{if } t \leq \pi/2 - \tan^{-1}\xi \\ (\infty, \tan(t + \tan^{-1}(\xi)) \cup R^{(t)}(\xi)) & \text{otherwise.} \end{cases}$$

Alternatively, if $\xi'$ is any point of $\mathbb{R}^1$, we could set

$$G^{(t)}(\xi) = \begin{cases} R^{(t)}(\xi) & \text{if } t < \pi/2 - \tan^{-1}\xi \\ R^{(t)}(\xi) \cup R^{(t - \pi/2 + \tan^{-1}(\xi'))(\xi')} & \text{otherwise.} \end{cases}$$

In both of these examples $G^{(t)}$ satisfies (a)–(d) of Theorem 3.1 but they differ with $R^{(t)}$ after escape times.

**5. Some Preparatory Lemmas**

In this section, the groundwork is laid for the proof of Theorem 3.1. The following technical lemmas provide detailed information on the behavior of reachable sets. Recall that the basic assumptions given in Section 3 are in force.

**Lemma 5.1.** Let $K \subseteq X$ be compact and $\delta > 0$ so that $K + \delta B \subseteq X$. Define $r = \sup \{|v| : v \in F(K + \delta B)\}$ and let $\varepsilon > 0$ be arbitrary. Then for $0 \leq T \leq (1/r) \min\{\delta, \varepsilon\}$ and $\xi \in K$, we have

$$R^{(T)}(\xi) \subseteq \xi + \varepsilon B. \quad (5.1)$$
Proof. Without loss of generality, we assume \( \varepsilon \leq \delta \). Let \( \xi \in K \) and \( 0 < T \leq \varepsilon / r \). Fix \( x(\cdot) \in S^{(T)}(\xi) \) and define \( t_0 = \inf \{ t: 0 < t < T \} \) with \( x(t) \notin \xi + \varepsilon B \). Note that \( t_0 > 0 \) since \( x(\cdot) \) is continuous. Also if \( 0 < t < t_0 \), then \( x(s) \in \xi + \varepsilon B \subseteq K + \delta B \), and hence \( |\dot{x}(s)| \leq r \) a.e. \( s \in [0, t_0] \). Therefore

\[
|x(t) - \xi| \leq \int_0^{t_0} |\dot{x}(s)| \, ds \leq rt_0 \leq rT \leq \varepsilon. \tag{5.2}
\]

If one of the inequalities in (5.2) is strict, it follows from the continuity of \( x(\cdot) \) and the definition of \( t_0 \) that \( t_0 = T \). However, if the inequalities in (5.2) are in fact equalities, we also have \( t_0 = T \). Since \( T \) can be taken equal to \( \varepsilon / r \) and \( x(\cdot) \) is arbitrary, it follows that (5.1) holds.

**Lemma 5.2.** Let \( \xi \in X \) and \( T > 0 \). Suppose there exists \( \delta > 0 \) so that \( R(\xi + \delta B) \subseteq X \). Then the multifunction on \( [0, T] \) defined by \( t \mapsto R^{(t)}(\xi) \) is compact valued and continuous in the Hausdorff metric on \( [0, T] \).

Proof. The results of this lemma are well known. The compactness result is an immediate consequence of [2, Theorem 3.1.7]. The continuity assertion is another exercise in applying [2, Theorem 3.1.7].

**Lemma 5.3.** For each \( \xi \in X \), we have \( T_x(\xi) \geq \delta / r \), where \( \delta \) and \( r \) are chosen so that \( \xi + \delta B \subseteq X \) and \( r > \sup \{ |v|: v \in F(\xi + \delta B) \} \).

Proof. Let \( \xi \in X \) and choose \( \delta > 0 \) so that \( \xi + \delta B \subseteq X \). Set \( r = \sup \{ |v|: v \in F(\xi + \delta B) \} \). For all \( 0 < t \leq \delta / r \), we have by Lemma 5.1 that \( R^{(t)}(\xi) \subseteq \xi + \delta B \). This implies that the hypotheses of Lemma 5.2 are satisfied (with perhaps a different choice of \( \delta \)). Since the image of a compact set under a continuous compact-valued multifunction is compact, it follows that \( R^{(t \leq t)}(\xi) \) is compact for all \( 0 \leq t \leq \delta / r \). Hence \( T_x(\xi) \geq \delta / r > 0 \).

**Lemma 5.4.** Let \( K \subseteq X \) be compact and \( \delta > 0 \) so that \( K + \delta B \subseteq X \). Let \( \lambda > 0 \) be a Lipschitz constant for \( F \) on \( K + \delta B \). Define \( r = \sup \{ |v|: v \in F(K + \delta B) \} \). Let \( \varepsilon > 0 \). Set \( T = (1/r) \min \{ \varepsilon / \lambda, \delta \} \). Then for \( \xi \in K \) and \( v \in F(\xi) \), there exists \( x(\cdot) \in S^{(T)}(\xi) \) so that

\[
\left| \frac{x(t) - \xi}{t} - v \right| < \varepsilon \quad \text{for all} \quad 0 < t \leq T. \tag{5.3}
\]

Proof. Without loss of generality, we assume \( \varepsilon \leq \delta \). For the construction \( x(\cdot) \), see Aubin and Cellina [1, pp. 115–117]. (In the notation of [1], \( x(\cdot) \) is \( y_{\bullet \cdot} \).) Actually \( x(\cdot) \) is \( C^1 \) with \( \dot{x}(0) = v \). From the top of [1, p. 117],
one can deduce that the modulus of continuity of $\dot{x}(\cdot)$ is $\leq \varepsilon/\lambda r = T$. So now if $0 < t \leq T$, then

$$\left| \frac{x(t) - x(0)}{t} - v \right| = \frac{1}{t} \int_0^t |\dot{x}(s) - \dot{x}(0)| \, ds < \varepsilon. \quad (5.4)$$

Since (5.4) holds for any $\xi \in K$ and $v \in F(\xi)$, this completes the proof.  

6. PROOF OF THEOREM 3.1

(a) Suppose $\xi \in X$ and $0 \leq T < T_x(\xi)$. We have already pointed out in Lemma 5.2 that $R^{(T)}(\xi)$ is compact. That $R^{(T)}(\xi)$ is nonempty is also well known; it can also be deduced directly from Lemma 5.4.

(b) Let $K \subseteq X$ be compact. Choose $\delta > 0$ so that $K + \delta B \subseteq X$ and $r = \sup \{|v|: v \in F(K + \delta B)\}$. A direct consequence of Lemma 5.3 is $\inf_{\xi \in K} T_x(\xi) \geq \delta/r > 0$.

(c) The semigroup property is an immediate consequence of the definitions, as noticed in (1.3).

(d) Let $K \subseteq X$ and $\varepsilon > 0$. We first fix some notation. Let $\delta > 0$, $\lambda \geq 1$, and $r > 0$ so that $K + \delta B \subseteq X$, $F$ is Lipschitz of order $\lambda$ on $K + \delta B$, and $r = \sup \{|v|: v \in F(K + \delta B)\}$. Without loss of generality, we assume that $\varepsilon \leq \delta$. Set $T = \varepsilon/\lambda r$.

Now let $\xi \in D$ and $0 < t \leq T$. If $x(\cdot) \in S^{(T)}(\xi)$, then

$$\frac{1}{t} (x(t) - \xi) = \frac{1}{t} \int_0^t \dot{x}(s) \, ds$$

$$\leq \frac{1}{t} \int_0^t F(x(s)) \, ds$$

$$\leq \frac{1}{t} \int_0^t F\left( \xi + \frac{\varepsilon}{\lambda} B \right) \, ds \quad \text{(by Lemma 5.1)}$$

$$\leq \frac{1}{t} \int_0^t (F(\xi) + \varepsilon B) \, ds \quad \text{(by the Lipschitz property)}$$

$$= F(\xi) + \varepsilon B$$

(by the Mean Value Theorem and convexity of $F(\xi)$).

From this it now follows that for all $0 < t \leq T$, we have

$$\frac{1}{t} \left( R^{(t)}(\xi) - \xi \right) \leq F(\xi) + \varepsilon B. \quad (6.1)$$
To obtain the reverse inclusion, let \( \xi \in K \) and \( v \in F(\xi) \). By Lemma 5.4, there exists \( x(\cdot) \in S^{(T)}(\xi) \) so that

\[
\frac{|x(t) - \xi|}{t} - v < \varepsilon \quad \text{for all } 0 < t \leq T.
\] (6.2)

Since \( v \) was any element in \( F(\xi) \), a consequence of (6.2) is that for all \( 0 < t \leq T \),

\[
F(\xi) \subseteq \frac{1}{t} (R^{(t)}(\xi) - \xi) + \varepsilon B.
\] (6.3)

Combining (6.2) and (6.3) reveals that

\[
\text{dist}_{H} \left( F(\xi), \frac{1}{t} (R^{(t)}(\xi) - \xi) \right) < \varepsilon
\]

for all \( 0 < t \leq T \). Since \( T \) does not depend on the particular choice of \( \xi \) in \( K \), the proof of (d) is complete.

7. TWO RESULTS BY FILIPPOV

The proof of Theorem 3.2 will require two theorems from Filippov [3]. In this section we state these results in forms that are readily applicable to our purpose. The original results in [3] are somewhat more general.

The first result is a staple of differential inclusion theory. Again the basic assumptions set forth in Section 3 are still in force here. If \( x(\cdot) \in AC[0, T] \) has range entirely within \( X \), we define \( \rho(x) = \int_0^T \text{dist}(\dot{x}(t), F(x(t))) \, dt \).

**THEOREM 7.1** (Filippov [3]). Suppose \( x(\cdot) \in AC[0, T] \) and \( \delta > 0 \) so that the set \( K := \{ \xi : |\xi - x(t)| < \delta \text{ for some } t \in [0, T] \} \) is contained in \( X \). Let \( \lambda > 0 \) be a Lipschitz constant for \( F \) on \( K \). Assume further that \( \rho(x) \leq \delta e^{-\lambda T} \). Then there exists \( \tilde{x}(\cdot) \in S^{(T)}(x(0)) \) with \( \|x - \tilde{x}\| < \rho(x) e^{\lambda T} \).

A straightforward proof can be found, for example, in Clarke [2, p. 115].

The second of Filippov's results is perhaps lesser known than Theorem 7.1. It will allow us to get a better handle on elements in the reachable set.

**THEOREM 7.2** (Filippov [3, Theorem 6]). Let \( \xi \in X, T > 0, \) and \( \varepsilon > 0 \). Then for each \( x(\cdot) \in S^{(T)}(\xi) \), there exists \( \tilde{x}(\cdot) \in S^{(T)}(\xi) \cap C^1[0, T] \) so that \( \|x - \tilde{x}\| < \varepsilon \).

Under our basic assumptions of \( F \), a straightforward proof of Theorem 7.2 can be found in Wolenski [8].
8. Proof of Theorem 3.2(i)

We are given real numbers \( \{ T_{\xi}(\xi) \} \in X \) and multifunctions \( \{ G^{(i)} \} \) satisfying (a)–(d) of Theorem 3.1, where \( T_X \) and \( R^{(i)} \) are replaced by \( T_G \) and \( G^{(i)} \). Let \( T > 0 \) and \( \xi \in X \). In this section we show that the inclusion \( R^{(T)}(\xi) \subseteq \text{cl} \, G^{(T)}(\xi) \) holds.

Let \( x(\cdot) \in S^{(T)}(\xi) \cap C^1[0, T] \). From Theorem 7.2, it immediately follows that to prove \( R^{(T)}(\xi) \subseteq \text{cl} \, G^{(T)}(\xi) \), it suffices only to show \( x(T) \in \text{cl} \, G^{(T)}(\xi) \).

We first fix some notation. Choose \( \delta > 0 \) so that \( \delta : = x[0, T] + \delta B = \{ \eta : \text{there exist } t \in [0, T] \text{ so that } |x(t) - \eta| \leq \delta \} \) is contained in \( X \). Let \( \lambda > 0 \) be a Lipschitz constant for \( F \) on \( K \), and let \( r = \sup \{ |v| : v \in F(K) \} \). For large \( N \), set \( h = T/N \) and \( t_j = jh \) for \( j = 0, 1, \ldots, N \). Define \( e_N \) by

\[
e_N = \max \left\{ \frac{hr}{1 - h}, \sup \text{dist}_{\eta \in K} \left( F(\eta), \frac{G^{(h)}(\eta) - \eta}{h} \right), \sup_{0 \leq j \leq N} \left| \frac{x(t_{j+1}) - x(t_j)}{h} - \ddot{x}(t_j) \right| \right\}.
\]

From (d) and \( x(\cdot) \in C^1[0, T] \), one has that \( e_N \to 0 \) as \( N \to \infty \). We assume \( N \) is large enough so that \( e_N < \frac{1}{2} \min \{ \delta, \lambda \delta/e^{i T} \} \).

Now set \( y_0 = x(t_0) \) and \( u_0 = \ddot{x}(t_0) \). Inductively, suppose \( y_j \) and \( u_j \) have been chosen so that

\[
y_j \in G^{(h)}(y_{j-1}) \cap K
\]

\[
u_j \in F(y_j) \quad \text{with} \quad |u_j - \ddot{x}(t_j)| \leq \lambda |y_j - x(t_j)|
\]

\[
|y_j - x(t_j)| \leq 2he_N \left( \frac{1 - x'}{1 - \alpha} \right), \quad \text{where} \quad \alpha = 1 + h\lambda.
\]

When \( j = 0 \), (8.2) is vacuous and (8.3) and (8.4) are trivial. Note that (8.4) expanded becomes

\[
|y_j - x(t_j)| \leq e_N \frac{2}{\lambda} \left( \left( 1 + \frac{\lambda T}{N} \right)^j - 1 \right) \leq \frac{2e_N}{\lambda} e^{i T}.
\]

Since \( y_j \in K \), we have by (8.1) that

\[
F(y_j) \subseteq \frac{G^{(h)}(y_j) - y_j}{h} + e_N B,
\]

and so \( y_{j+1} \in G^{(h)}(y_j) \) can be chosen satisfying

\[
|y_{j+1} - y_j - hu_j| \leq he_N.
\]
Choose $u_{j+1} \in F(y_{j+1})$ to be the nearest element to $\dot{x}(t_{j+1})$. To apply the Lipschitz property of $F$, we first must show $y_{j+1} \in K$. This is done by noting
\begin{equation}
|y_{j+1} - x(t_j)| \leq |y_{j+1} - y_j - hu_j| + h|u_j| + |y_j - x(t_j)|
\leq h\varepsilon_N + hr + \frac{2\varepsilon_N}{\lambda} e^{\lambda T} \quad \text{(by (8.7), (8.3), (8.5))}
\leq \delta \quad \text{(by (8.1)).}
\end{equation}

Hence $y_{j+1} \in K$, and so from the Lipschitz property of $F$ on $K$, we have
\begin{equation}
|u_{j+1} - \dot{x}(t_{j+1})| \leq \lambda |y_{j+1} - x(t_{j+1})|. \tag{8.8}
\end{equation}

We turn next to showing that (8.4) holds for $j+1$. By the triangle inequality,
\begin{equation}
|y_{j+1} - x(t_{j+1})| \leq |y_{j+1} - y_j - hu_j| + |y_j - x(t_j)|
+ h|u_j - \dot{x}(t_j)| + |x(t_j) + hx(t_j) - x(t_{j+1})|. \tag{8.9}
\end{equation}

Each term on the right side of (8.9) can now be estimated. By (8.7) the first term is $\leq h\varepsilon_N$. By (8.4) the second terms is $\leq 2h\varepsilon_N((1 - \alpha)/(1 - \alpha))$. By (8.3) and (8.4) the third term is $\leq 2\lambda h^2 \varepsilon_N((1 - \alpha)/(1 - \alpha))$. And by (8.1) the fourth term is $\leq h\varepsilon_N$. Adding these estimates together gives us
\begin{equation}
|y_{j+1} - x(t_{j+1})| \leq 2h\varepsilon_N \left( 1 + \frac{1 - \alpha}{1 - \alpha} + h\lambda \left( \frac{1 - \alpha}{1 - \alpha} \right) \right)
= 2h\varepsilon_N \left( \frac{1 - \alpha^{j+1}}{1 - \alpha} \right). \tag{8.10}
\end{equation}

The induction is now completed since (8.2)–(8.4) hold for $j+1$.

At stage $N$, (8.5) becomes
\begin{equation}
|y_N - x(T)| \leq \frac{2\varepsilon_N}{\lambda} e^{\lambda T}. \tag{8.10}
\end{equation}

Also the semigroup property (b) and the choices of $y_j$ in (8.2) reveal that
\begin{equation}
y_N \in G^{(h)}(G^{(h)}(...G^{(h)}(\xi)...)) = G^{(T)}(\xi). \tag{8.11}
\end{equation}

Finally, since $\varepsilon_N \to 0$ as $N \to \infty$, (8.10) and (8.11) imply that $x(T) \in \text{cl } G^{(T)}(\xi)$. 
9. Proof of Theorem 3.2(ii)

We now turn to the proof of part (ii). Let $\xi \in X$ and $0 \leq T < T_X(\xi)$. Consider the following condition:

there exists $\delta > 0$ so that $K := \text{cl} \left( \bigcup_{0 \leq t \leq T} G^{(t)}(\xi) + \delta B \right)$ is compact in $X$. (*)

We first prove $G^{(T)}(\xi) \subseteq R^{(T)}(\xi)$ under the added assumption (*), then we will later remove this assumption and also show that $G^{(T)}(\xi)$ is closed.

Let $\varepsilon > 0$ and $\gamma \in G^{(T)}(\xi)$. Assume (*) holds. Without loss of generality, we assume $\varepsilon \leq \delta/2$. By the uniform convergence property (d) of the $G^{(t)}$'s, there exists $t_0 > 0$ so that for all $0 < t \leq t_0$ and $\eta \in K$, we have

$$\frac{G^{(t)}(\eta) - \eta}{t} \subseteq F(\eta) + \varepsilon B.$$ (9.1)

From assumption (b), we can shrink $t_0$ if necessary so that $0 < t_0 < \inf_{\eta \in K} T_{G}(\eta)$ also holds. In particular, $G^{(t)}(\eta) \neq \emptyset$ for all $\eta \in K$ and $0 < t \leq t_0$. Let $\lambda > 1$ be a Lipschitz constant for $F$ on $K$ and $r := \sup \{ |v| : v \in F(K) \}$.

Now fix a large positive integer $N$ with $T/N < \min \{ t_0, \varepsilon/\lambda(r + \varepsilon) \}$. Set $h = T/N$ and $t_j = j h$ for $j = 0, 1, \ldots, N$. By the semigroup property (c) of $\{ G^{(t)} \}$, there exist $y_0 = \xi, y_1, \ldots, y_N = \gamma$ so that

$$y_{j+1} \in G^{(h)}(y_j), \quad j = 0, 1, \ldots, N - 1.$$ (9.2)

From (*) we have that each $y_j$ is contained in $K$, $j = 0, \ldots, N - 1$. Consequently, combining (9.1) (with $\eta = y_j$ and $t = h$) and (9.2) gives

$$y_{j+1} \in y_j + hF(y_j) + \varepsilon hB, \quad j = 0, 1, \ldots, N - 1.$$ (9.3)

The definition of $r$ and the choice of $N$ turns (9.3) into the estimate

$$|y_{j+1} - y_j| \leq h(r + \varepsilon) \leq \varepsilon/\lambda.$$ (9.4)

Let $x(\cdot) : [0, T] \to \mathbb{R}^n$ be the piecewise linear interpolation of the $y_j$'s equally spaced on $[0, T]$. That is,

$$x(t) = y_j + \frac{t - t_j}{h} (y_{j+1} - y_j) \quad \text{if} \quad t_j \leq t < t_{j+1}.$$ (9.5)

From (9.4) and (9.5), it follows that

$$|x(t) - y_j| \leq |y_{j+1} - y_j| \leq \varepsilon/\lambda \quad \text{if} \quad t_j \leq t \leq t_{j+1}.$$ (9.6)
We have taken $\lambda \geq 1$ and $\varepsilon \leq \delta$, thus (9.6) and (*) imply that the range of $x(\cdot)$ lies within $K$. This allows us to apply the Lipschitz property of $F$:

$$\text{dist}_H(F(x(t)), F(y_j)) \leq \lambda |x(t) - y_j| \leq \varepsilon \quad \text{(by (9.6)).} \tag{9.7}$$

Next we estimate $\rho(x)$:

$$\rho(x) = \int_0^T \text{dist}(x(t), F(x(t))) \, dt$$

$$\leq \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \text{dist} \left( y_{j+1} - y_j \over h, F(y_j) \right) + \text{dist}_H(F(x(t)), F(y_j)) \, dt$$

$$\leq 2\varepsilon T \quad \text{(by (9.3) and (9.7)).} \tag{9.8}$$

According to Filippov's Theorem 7.1, there exist $\tilde{x}(\cdot) \in S^{(T)}(\xi)$ such that

$$\|x - \tilde{x}\| \leq e^{tT} \rho(x) \leq 2\varepsilon Te^{tT} \quad \text{(by (9.8)).} \tag{9.9}$$

Since $\varepsilon$ is arbitrarily small and $R^{(T)}(\xi)$ is compact, we conclude from (9.9) that $y \in R^{(T)}(\xi)$. Hence we have shown $G^{(T)}(\xi) \subseteq R^{(T)}(\xi)$ under the additional assumption (*).

We now abandon the previous choices of $T$, $K$, $\lambda$, etc., and begin anew. Let $\xi \in X$. Define

$$T_1 := \inf \{ T : G^{(T)}(\xi) \neq R^{(T)}(\xi) \}. \tag{9.10}$$

The uniform convergence property (d) and the bounded values of $F$ combine to imply that condition (*) holds for small values of $T > 0$. Moreover $G^{(T)}(\xi)$ is closed for $T < T_G(\xi)$ by assumption (b). Therefore by what was shown above in conjunction with the result of (i) (whose proof is given in Section 8), we have that $G^{(T)}(\xi) = R^{(T)}(\xi)$ for all small $T$. That is, $T_1 > 0$.

Suppose $T_1 < T_X(\xi)$. Let $K := \bigcup_{0 \leq T < T_1} R^{(T)}(\xi)$, a compact subset of $X$. By Lemma 5.2, we have $K = \text{cl} \bigcup_{0 \leq T < T_1} R^{(T)}(\xi)$, which in turn equals $\text{cl} \bigcup_{0 \leq T < T_1} G^{(T)}(\xi)$ by (9.10). We now introduce some further notation. Let $\delta > 0$ be such that $K + \delta B \subseteq X$; let $\lambda \geq 1$ so that $F$ is Lipschitz of order $\lambda$ on $K + \delta B$; and set $r = \sup \{|v| : v \in F(K + \delta B)\}$.

By assumption (d), there exists $t_1 > 0$ so that for all $\eta \in K$ and $0 < t \leq t_1$, we have

$$G^{(T)}(\eta) \subseteq \eta + tF(\eta) + \frac{\delta}{2} B.$$
Shrinking $t_1$ if necessary, we assume

$$0 < t_1 < \min \left\{ \frac{\delta}{2r}, \inf_{\eta \in K} T_\sigma(\eta), T_1 \right\}.$$

The same reasoning used above to show $T_1 > 0$ can be applied here to $\eta \in K$ in place of $\xi$, and consequently one has

$$G^{(\tau)}(\eta) = R^{(\tau)}(\eta) \quad \text{for} \quad \eta \in K, \quad 0 < t \leq t_1. \tag{9.11}$$

So now if $0 < t < t_1$, then we have

$$G^{(T_1 + t/2)}(\xi) = G^{(t_1 - t/2)}(G^{(T_1 + t - t_1)}(\xi)) \quad \text{(semigroup property)}
= G^{(t_1 - t/2)}(R^{(T_1 + t - t_1)}(\xi)) \quad \text{(by (9.10))}
= R^{(t_1 - t/2)}(R^{(T_1 + t - t_1)}(\xi)) \quad \text{(by (9.11))}
= R^{(T_1 + t/2)}(\xi). \tag{9.12}$$

However, (9.12) says that $G^{(T)}(\xi) = R^{(T)}(\xi)$ for all $0 \leq T < T_1 + t_1/2$, a contradiction to the choice of $T_1$. We conclude that $T_1 \geq T(\xi)$, which finishes the proof of (ii).

10. THE NONAUTONOMOUS VERSION

In this last section, we state the analogues of Theorems 3.1 and 3.2 allowing explicit time dependence on the data $F$.

Suppose $X \subseteq \mathbb{R}^n$ is nonempty and open and $F: [0, \infty) \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$ satisfies

for all $t \in [0, \infty)$, $\text{dom } F(t, \cdot) = X$

$F(t, \xi)$ is compact and convex for all $(t, \xi) \in [0, \infty) \times X$

for all $\xi \in X$, $t \mapsto F(t, \xi)$ is continuous

for all $T > 0$ and $0 \leq t \leq T$, $F(t, \cdot)$ is locally Lipschitz on $X$ with the Lipschitz constant independent of $t \in [0, T]$.

Let $0 \leq t_0 < t_1$ and $\xi \in \mathbb{R}^n$. Consider the differential inclusion

$$x(\cdot) \in AC[t_0, t_1]$$

$$\dot{x}(t) \in F(t, x(t)) \quad \text{a.e.} \quad t \in [t_0, t_1], \quad \tag{10.1}$$

$$x(t_0) = \xi.$$
We must reset our notation:

\[ R(t_0, t_1, \xi) = \{ x(t_1) : x(\cdot) \text{ satisfies (10.1)} \} \]

\[ R(t_0 \leq t_1, \xi) = \bigcup_{t_0 \leq t \leq t_1} R(t_0, t_1, \xi) \]

\[ T_{X, t_0}(\xi) = \sup \{ t_1 : R(t_0 \leq t_1, \xi) \text{ is compact} \}. \]

Now the nonautonomous version can be stated as

**Theorem 10.1.** The following hold:

(a) for each \( \xi \in X \) and \( 0 \leq t_0 < t_1 < T_{X, t_0}(\xi) \), \( R(t_0, t_1, \xi) \) is nonempty and compact,

(b) for each \( T > 0 \) and compact \( K \subseteq X \), we have \( \inf \{ T_{X, t_0}(\xi) : 0 \leq t_0 \leq T, \xi \in K \} > 0 \),

(c) for all \( 0 \leq t_0 < t_1 < t_2 \) and \( \xi \in X \), \( R(t_0, t_2, \xi) = R(t_1, t_2, R(t_0, t_1, \xi)) \),

(d) for all \( T > 0 \), the multifunctions \( (t, \xi) \mapsto (1/h) \{ R(t, t + h, \xi) - \xi \} \) parametrized by \( h > 0 \) converge to \( F \) as \( h \downarrow 0 \) uniformly on compact subsets of \([0, T] \times X\).

**Theorem 10.2 (Uniqueness Theorem).** Suppose real numbers \( \{ T_{G, t}(\xi) \}_{\xi \in X} \) and multifunctions \( \{ G(t_0, t_1, \cdot) \}_{t_0 > t_0 \geq 0} \) are given so that (a)-(d) of Theorem 10.1 hold where \( T_{G, t} \) and \( G \) replace \( T_{X, t} \) and \( F \), respectively. Then for all \( \xi \in X \), we have

(i) for all \( t_1 > t_0 \geq 0 \), \( R(t_0, t_1, \xi) \subseteq \text{cl} G(t_0, t_1, \xi), \) and

(ii) for all \( 0 \leq t_0 < t_1 < T_{X, t_0}(\xi) \), \( R(t_0, t_1, \xi) = G(t_0, t_1, \xi). \)

**Acknowledgment**

I thank Professor R. T. Rockafellar for suggesting this semigroup approach and for his invaluable help in obtaining the results.

**References**


