On the Strong Invariance Property for Non-Lipschitz Dynamics*

Mikhail Krastanov† Michael Malisoff‡ Peter Wolenski‡

Abstract

We provide a new sufficient condition for strong invariance for differential inclusions, under very general conditions on the dynamics, in terms of a Hamiltonian inequality. In lieu of the usual Lipschitzness assumption on the multifunction, we assume a feedback realization condition that can in particular be satisfied for measurable dynamics that are neither upper nor lower semicontinuous.

1 Introduction

Topics in flow invariance theory provide the foundation for considerable current research in control theory and optimization (cf. [6, 7, 8, 10, 18, 21, 22]). The setting is that a multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defining dynamics and a closed set $S \subseteq \mathbb{R}^n$ determining state constraints are given, and the theory then contains sufficient conditions under which some, or all, of the trajectories of F that start in S remain in S. More particularly, we say that (F, S) is *weakly invariant in* $\mathbb{R}^n$ provided for each $\bar{x} \in S$, *there exists* a trajectory $t \mapsto \phi(t)$ of F starting at $\bar{x}$ that remains in S for all t before its escape time $\mathrm{Esc}(\phi, \mathbb{R}^n)$ (precise definitions are given in section 2 below). Weak invariance of a set is also termed *viability*, and sufficient (and necessary) conditions for weak invariance have been developed under very general conditions on the dynamics, both in terms of tangential-type inclusions and Hamiltonian inequalities.

A more restrictive invariance property is as follows: We say that (F, S) is *strongly invariant in* $\mathbb{R}^n$ provided for each $\bar{x} \in S$, *each trajectory* $t \mapsto \phi(t)$ of F starting at $\bar{x}$ remains in S for all t before its escape time $\mathrm{Esc}(\phi, \mathbb{R}^n)$. In contrast with weak invariance theorems, which merely require the dynamics to have locally bounded values and closed graph, sufficient conditions for strong invariance usually invoke a Lipschitz condition on the dynamics (cf. section 3.2 for a survey of results in this direction). For example, if F is locally Lipschitz and nonempty and compact-convex valued with linear growth, then it is well known (cf. [7, Chapter 4]) that (F, S) is strongly invariant in $\mathbb{R}^n$ if and only if $F(x) \subseteq T^C_S(x)$ for all $x \in S$, where T^C_S denotes the Clarke tangent cone.

However, this cone characterization can fail if F is non-Lipschitz, as illustrated in the following simple example: Take $n = 1$, $S = \{0\}$, $F(0) = [-1, +1]$, and $F(x) = \{-\mathrm{sign}(x)\}$ for $x \neq 0$. Then $T^C_S(0) = \{0\}$, even though (F, S) is strongly invariant in $\mathbb{R}$. This example satisfies our dynamic assumptions (cf. Example 2.2 below). It is also covered by our main theorem (see section 3).

Starting from strong invariance and its Hamiltonian characterizations, one can develop uniqueness results and regularity theory for proximal solutions of Hamilton-Jacobi-Bellman equations, stability theory, infinitesimal characterizations of monotonicity, and many other applications (cf. [1, 7, 8, 14, 21]). On the other hand, it is well appreciated that many important dynamics are non-Lipschitz, and may even be discontinuous, and therefore are beyond the scope of the usual strong invariance characterizations. Therefore, the development of conditions guaranteeing strong invariance under less restrictive assumptions is a problem that is of considerable ongoing research interest.

---

*Corresponding Author: Peter Wolenski. Date: January 31, 2004.

†Department of Biomathematics; Institute of Mathematics and Informatics; Bulgarian Academy of Sciences; Acad. G. Bonchev str., bl 8; 1113 Sofia; BULGARIA; krast@math.bas.bg

‡Department of Mathematics; 304 Lockett Hall; Louisiana State University and A. & M. College; Baton Rouge, LA 70803-4918; USA; {malisoff, wolenski}@math.lsu.edu

This motivates the search for sufficient conditions for strong invariance for non-Lipschitz differential inclusions, which is the focus of this note. Donchev, Rios and Wolenski [10, 11] recently developed necessary and sufficient conditions for strong invariance for so-called *one-sided Lipschitz* differential inclusions. See also [18] for an autonomous normal type characterization of strong invariance for certain systems with a discontinuous component. These works apply under special conditions on the structure of the dynamics (cf. section 3 for further details).

In this note, we pursue a very different approach. Rather than restricting the structure of the dynamics, we provide a sufficient condition for strong invariance under an appropriate feedback realization hypothesis. This hypothesis is related to Sussmann's 'unique limiting' property that was introduced in [20] in the context of exit time optimal control problems with continuous dynamics, and to Malisoff's "Lipschitz upper envelope" condition from [16, 17]. Roughly speaking, our realization property states that each trajectory ϕ of the dynamics F is also a unique trajectory of a nonautonomous singleton-valued dynamics f for which $f(t,x) \in \text{cone}\{F(x)\}$ for all t and all x near $\phi(0)$ (cf. section 2 below for a precise statement of our hypothesis). This is a less restrictive assumption than those of the known strong invariance characterizations because it can be satisfied by important classes of differential inclusions with measurable, but possibly neither upper nor lower semicontinuous, right-hand sides (cf. section 2 for examples).

In section 2, we state our realization hypothesis precisely and provide the necessary background on differential inclusions and nonsmooth analysis. We also illustrate the applicability of our hypothesis to a broad class of discontinuous dynamics that are beyond the scope of the well known strong invariance results. In section 3, we announce our strong invariance result and discuss its relationship to the known theorems in strong invariant system theory. Section 4 contains the proof of our strong invariance criterion.

2 Assumptions and Preliminaries

2.1 Basic Hypothesis

Our main object of study in this note is an autonomous differential inclusion $\dot{x} \in F(x)$. In this subsection, we state our hypothesis on F and illustrate its relevance using several applications. Our novel feature is the requirement that each trajectory of F be realizable as the unique solution to a nonautonomous local feedback selection of F. On the other hand, we will not require the Lipschitz property or other structural assumptions on F that are generally invoked in strong invariant system theory (cf. [7, 9, 10, 14, 18]).

To make our realization hypothesis precise, we require the following definitions and notation. By a *trajectory* of $\dot{x} \in F(x)$ on an interval $[0,T]$ starting at a point $x_o \in \mathbb{R}^n$, we mean an absolutely continuous function $\phi : [0,T] \to \mathbb{R}^n$ for which $\phi(0) = x_o$ and $\dot{\phi}(t) \in F(\phi(t))$ for (Lebesgue) almost all (a.a.) $t \in [0,T]$. We let $\text{Traj}_T(F,x)$ denote the set of all trajectories $\phi : [0,T] \to \mathbb{R}^n$ for F starting at x on all possible intervals $[0,T]$, and we set $\text{Traj}(F,x) := \cup_{T\geq 0}\text{Traj}_T(F,x)$ and $\text{Traj}(F) := \cup_{x\in\mathbb{R}^n}\text{Traj}(F,x)$.

A multifunction $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to have *linear growth* provided there exist positive constants c_1 and c_2 such that $||v|| \leq c_1 + c_2||x||$ for all $v \in G(x)$ and $x \in \mathbb{R}^n$, where $||\cdot||$ denotes the Euclidean norm. For any interval I, a function $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ is said to have *linear growth (on I)* provided $x \mapsto G(x) := \{f(t,x) : t \in I\}$ has linear growth. For any sets $D, M \subseteq \mathbb{R}^n$ and $\eta \in \mathbb{R}$, we set $M + \eta D := \{m + \eta d : m \in M, d \in D\}$, and $\text{cone}\{D\} := \cup\{\eta D : \eta \geq 0\}$. Also, $\mathcal{B}_n(p) := \{x \in \mathbb{R}^n : \|x - p\| \leq 1\}$ for all $p \in \mathbb{R}^n$ and $\mathcal{B}_n := \mathcal{B}_n(0)$. A function $\omega(\cdot) : [0,\infty) \to [0,\infty)$ is called a *modulus* provided it is nondecreasing and continuous with $\omega(0) = 0$. For each $T \geq 0$, we let $\mathcal{C}[0,T]$ denote the set of all functions $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ that satisfy

(C_1) For each $x \in \mathbb{R}^n$, the map $t \mapsto f(t,x)$ is measurable;

(C_2) For each compact set $K \subseteq \mathbb{R}^n$, there exists a modulus $\omega_{f,K}(\cdot)$ such that, for all $t \in [0,T]$ and x_1, $x_2 \in K$, $\|f(t,x_1) - f(t,x_2)\| \leq \omega_{f,K}(\|x_1 - x_2\|)$; and

(C_3) f has linear growth on [0,T].

It is noteworthy that $\omega_{f,K}(\cdot)$ in the previous definition is independent of $t \in [0,T]$. For each $\bar{x} \in \mathbb{R}^n$, denote by $\mathcal{C}_F([0,T],\bar{x})$ those $f \in \mathcal{C}[0,T]$ that are also selections of the cone of F for almost all $t \in [0,T]$ and all $x \in \mathbb{R}^n$ sufficiently near $\bar{x}$; that is,

$$\mathcal{C}_F([0,T],\bar{x}) := \{f \in \mathcal{C}[0,T] : \exists \gamma > 0 \text{ s.t. } f(t,x) \in \text{cone}\,\{F(x)\} \text{ for a.a. } t \in [0,T] \text{ and all } x \in \gamma \mathcal{B}_n(\bar{x})\}.$$

Notice that while elements $f \in \mathcal{C}_F([0,T],\bar{x})$ are defined on all of $[0,T] \times \mathbb{R}^n$, they are only required to satisfy $f(t,x) \in \text{cone}\,\{F(x)\}$ on *part of* their domain. We also let $\mathcal{C}_F[0,T]$ denote those $f \in \mathcal{C}[0,T]$ such that $f(t,x) \in \text{cone}\,\{F(x)\}$ for almost all $t \in [0,T]$ and all $x \in \mathbb{R}^n$. We will assume the following:

(U) For each $\bar{x} \in \mathbb{R}^n$, $T \geq 0$, and $\phi \in \text{Traj}_T(F,\bar{x})$, there exists $f \in \mathcal{C}_F([0,T],\bar{x})$ for which ϕ is the unique solution of the initial value problem $\dot{y}(t) = f(t,y(t))$, $y(0) = \bar{x}$ on $[0,T]$.

Notice that hypothesis (U) is weaker than requiring a continuous selection from the dynamics F that realizes the trajectory. This is because f is allowed to depend on time as well as the state, and need only be a *local* selection. Moreover, f is allowed to depend on the choice of the trajectory ϕ, and need not be continuous. In practice, hypothesis (U) can be checked using open or closed loop controls, and may be satisfied for non-Lipschitz dynamics. The following examples illustrate these points and also show how to use cones to check condition (U).

Example 2.1. Assume $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is Lipschitz and nonempty and compact-convex valued. We claim that F satisfies condition (U). To see why, let $\bar{x} \in \mathbb{R}^n$, $T > 0$, and $\phi \in \text{Traj}_T(F,\bar{x})$ be given, and set

$$f(t,x) = \text{proj}_{F(x)}(\dot{\phi}(t))$$

(i.e., $f(t,x)$ is the closest point to $\dot{\phi}(t)$ in $F(x)$, which is well defined by the convexity of $F(x)$). Then $f \in \mathcal{C}_F[0,T]$ satisfies the requirement. If on the other hand $F : \mathbb{R} \rightrightarrows \mathbb{R}$ is defined by $F(x) = \{1\}$ for $x < 0$, $F(0) = \{0\} \cup [1,2]$, and $F(x) = [0,2]$ for $x > 0$, and if $\phi \in \text{Traj}(F)$, then $f(t,x) \equiv \dot{\phi}(t) \in \text{cone}\{F(x)\}$ for almost all t and all $x \in \mathbb{R}^n$. Therefore, condition (U) is again satisfied, even though F is neither upper nor lower semicontinuous nor convex valued.

Example 2.2. Consider the example from the introduction, namely, $n = 1$, $F(0) = [-1,+1]$, and $F(x) = \{-\text{sign}(x)\}$ for $x \neq 0$. We claim that (U) is again satisfied. To see why, let $T > 0$, $\bar{x} \in \mathbb{R}$, and $\phi \in \text{Traj}_T(F,\bar{x})$ be given. Note that $(F,\{0\})$ is strongly invariant in $\mathbb{R}$. Therefore, either (i) ϕ starts at some $\bar{x} \neq 0$ and then moves to 0 at unit speed and then stays at 0 or (ii) $\phi \equiv 0$. If $\bar{x} \neq 0$, then the requirement is satisfied using

$$f(t,x) \equiv -\text{sign}(\bar{x})\beta(t),$$

where $\beta(t) = 1$ if $t \in [0,|\bar{x}|]$ and 0 otherwise. In this case, $f(t,x) \in \text{cone}\{F(x)\}$ for all $t \in [0,T]$ and $x \in (|\bar{x}|/2)\mathcal{B}_1(\bar{x})$. If instead $\bar{x} = 0$, then the requirement is satisfied with $f(t,x) \equiv 0 \in \text{cone}\{F(x)\}$ for all $t \in [0,T]$ and $x \in \mathbb{R}$.

Example 2.3. Assume $F(x) = g(x,A)U(x)$ where $A \subseteq \mathbb{R}^m$ is compact, and $g : \mathbb{R}^n \times A \to \mathbb{R}^n$ is continuous and satisfies

(H) For each compact set $K \subseteq \mathbb{R}^n$, there exists $L_K > 0$ such that for all $x_1, x_2 \in K$ and $a \in A$, $(g(x_1,a) - g(x_2,a)) \cdot (x_1 - x_2) \leq L_K \|x_1 - x_2\|^2$. Also, $x \mapsto g(x,A)$ has linear growth.

and $U : \mathbb{R}^n \rightrightarrows \mathbb{R}$ is bounded, measurable, closed and nonempty valued, and satisfies $U(x) \cap (0,\infty) \neq \emptyset$ and $U(x) \cap (-\infty,0) \neq \emptyset$ for all $x \in \mathbb{R}^n$. (The argument we are about to give still applies if instead of assuming that $U(x) \cap (0,\infty) \neq \emptyset$ and $U(x) \cap (-\infty,0) \neq \emptyset$ for all $x \in \mathbb{R}^n$, we instead assume that either $U : \mathbb{R}^n \rightrightarrows (0,\infty)$ or $U : \mathbb{R}^n \rightrightarrows (-\infty,0)$.) One can easily check (cf. [2]) that condition (H) guarantees existence of at most one solution $\phi : [0,T] \to \mathbb{R}^n$ of $\dot{x} = g(x,\alpha)\beta$ for each initial condition, $T > 0$, and bounded measurable functions $\alpha : [0,T] \to A$ and $\beta : [0,T] \to \mathbb{R}$. To check condition ($U$), let $\phi \in \text{Traj}(F)$. Consider the multifunction

$$G(t) := \{(a,b) \in A \times U(\phi(t)) : \dot{\phi}(t) = g(\phi(t),a)b\}.$$

Notice that $t \mapsto A \times U(\phi(t))$ is closed valued and measurable, and that $(t,(a,b)) \mapsto g(\phi(t),a)b$ is measurable. Applying the Filippov selection lemma (cf. [21, p. 72]) to G, we find a measurable pair (α,β) for which $\alpha(t) \in A$, $\beta(t) \in U(\phi(t))$, and $\dot{\phi}(t) = g(\phi(t),\alpha(t))\beta(t)$ for almost all t. We now show that condition (U) holds with the choice

$$f(t,x) := g(x,\alpha(t))\beta(t).$$

In general, we will not have $\beta(t) \in U(x)$ for all t and x. In fact, it could be that $U(\phi(t)) \cap U(x) = \emptyset$ for some t and x, so we may not have $f(t,x) \in F(x)$ for a.a. t and x. On the other hand, one can easily check that $\beta(t) \in \text{cone}\{U(x)\}$ for all t and x, so $f(t,x) \in g(x,A)\text{cone}\{U(x)\} = \text{cone}\{F(x)\}$ for a.a. t and all x, and this gives the desired result.

Example 2.4. Let $R \subseteq \mathbb{R}^n$ be compact and $F(x) = \{\lambda(x) + U(x)\delta(x)\}R$, where (i) $\delta, \lambda : \mathbb{R}^n \to \mathbb{R}$ are continuous functions such that $|\lambda(x)| \geq 2|\delta(x)|$ for all x, (ii) $g(x,A) := \lambda(x) + A\delta(x)$ satisfies condition (H) above with $n = 1$ and $A = [-1,+1]$, and (iii) $U : \mathbb{R}^n \to [-1,+1]$ is measurable. To check condition (U), let $\phi \in \text{Traj}(F)$. Applying the Filippov lemma as in the previous example, we find measurable functions u and r such that $\dot{\phi}(t) = [\lambda(\phi(t)) + u(t)\delta(\phi(t))]r(t)$ for almost all t, with $u(t) \equiv U(\phi(t))$. Notice that

$$\lambda(x) + U(y)\delta(x) \in (\min\{\lambda(x)/2, 3\lambda(x)/2\}, \max\{\lambda(x)/2, 3\lambda(x)/2\})$$

for all x and y, by our stated conditions. This implies that $\lambda(x) + U(y)\delta(x) \in \text{cone}\{\lambda(x) + U(x)\delta(x)\}$ for all $x, y \in \mathbb{R}^n$, so we can satisfy condition (U) using $f(t,x) := [\lambda(x) + u(t)\delta(x)]r(t)$. In this case $f \in \mathcal{C}_F[0,T]$.

2.2 Preliminaries in Nonsmooth Analysis

The principal nonsmooth objects used in this note are the proximal subgradient and normal cone, and here we review these concepts; see [7] for a complete treatment. Let $S \subseteq \mathbb{R}^n$ be closed and $s \in S$. A vector $\zeta \in \mathbb{R}^n$ is called a *proximal normal* vector of S at s provided there exists $\sigma = \sigma(\zeta, s) > 0$ so that

$$\langle \zeta, s' - s \rangle \leq \sigma ||s' - s||^2 \quad \text{for all } s' \in S. \tag{1}$$

The set of all proximal normals of S at s is denoted by $N_S^P(s)$ and is a convex cone. One can show (cf. [7, p. 25]) that for each $\delta > 0$ and $s \in S$, $\zeta \in N_S^P(s)$ if and only if there exists $\sigma = \sigma(\zeta, s) > 0$ so that

$$\langle \zeta, s' - s \rangle \leq \sigma ||s' - s||^2 \quad \text{for all } s' \in S \cap \delta\mathcal{B}_n(s). \tag{2}$$

Recall that the *distance function* $d_S(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is given by $d_S(x) := \min\{\|x - s\| : s \in S\}$.

For the related functional concept, assume $f : \mathbb{R}^n \to (-\infty, \infty]$ is lower semicontinuous and let $x \in \text{domain}(f) := \{x' : f(x') < \infty\}$. Then $\zeta \in \mathbb{R}^n$ is called a *proximal subgradient* for f at x provided there exist $\sigma > 0$ and $\eta > 0$ so that

$$f(x') \geq f(x) + \langle \zeta, x' - x \rangle - \sigma \|x' - x\|^2 \quad \text{for all } x' \in \eta\mathcal{B}_n(x). \tag{3}$$

The set of all proximal subgradients for f at x is denoted by $\partial_P f(x)$. This set could be empty at some points, even for C^1 functions (e.g., $\partial_P f(0) = \emptyset$ if $f(x) = -|x|^{3/2}$).

We next state the version of the Clarke-Ledyaev Mean Value Inequality needed for our strong invariance results. Let $[x, Y]$ denote the closed convex hull of $x \in \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^n$.

Theorem 1. *Assume $x \in \mathbb{R}^n$, $Y \subseteq \mathbb{R}^n$ is compact and convex, and $\Psi : \mathbb{R}^n \to (\infty, +\infty]$ is lower semicontinuous. Then for any $\delta < \min_{y \in Y} \Psi(y) - \Psi(x)$ and $\lambda > 0$, there exist $z \in [x, Y] + \lambda\mathcal{B}_n$ and $\zeta \in \partial_P \Psi(z)$ so that $\delta < \langle \zeta, y - x \rangle$ for all $y \in Y$.*

For the proof, see [7, p. 117]; an infinite dimensional version also holds (see [5]), but is not needed here.

2.3 Background in Differential Inclusions

In this subsection, we review invariant systems theory and a standard result on compactness of trajectories for discontinuous dynamics. The following definition of escape times was introduced in [22]:

Definition 2.5. Let $\mathcal{G} \in \mathbb{R}^n$ be open, $x_0 \in \mathcal{G}$, and $x(\cdot)$ be a trajectory of a differential inclusion $\dot{x} \in F(x)$ with $x(0) = x_0$ defined on a half–open interval $[0, T)$, where $0 < T \leq \infty$. Then T is called an *escape time* of $x(\cdot)$ from $\mathcal{G}$ provided at least one of the following conditions hold:

(E_1) $T = \infty$ and $x(t) \in \mathcal{G}$ for all $t \geq 0$;
(E_2) $x(t) \in \mathcal{G}$ for all $t \in [0, T)$ and $\|x(t)\| \to \infty$ as $t \uparrow T$; or
(E_3) $T < \infty$, $x(t) \in \mathcal{G}$ for all $t \in [0, T)$, and $d_{\mathcal{G}^c}(x(t)) \to 0$ as $t \uparrow T$.

We next define strong and weak invariance. Assume $\mathcal{G} \subseteq \mathbb{R}^n$ is open and $x_0 \in \mathcal{G}$. The set of all trajectories of F originating from x_0 that remain in $\mathcal{G}$ over a maximal interval is denoted by $\Upsilon_{(F,\mathcal{G})}(x_0)$. Therefore, $\Upsilon_{(F,\mathcal{G})}(x_0)$ consists of those absolutely continuous functions $x(\cdot)$ defined on a half–open interval $[0, T)$ for which $x(0) = x_0$ and $\dot{x}(t) \in F(x(t))$ is satisfied for almost all $t \in [0, T]$, where $T = \mathrm{Esc}\,(x(\cdot); \mathcal{G})$.

Definition 2.6. Assume $E \subseteq \mathbb{R}^n$ is closed, and $\mathcal{G} \subseteq \mathbb{R}^n$ is open with $E \cap \mathcal{G} \neq \emptyset$. Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

(a) (F, E) is called *weakly invariant in* $\mathcal{G}$ provided that for every $x_0 \in E \cap \mathcal{G}$, there exists a trajectory $x(\cdot) \in \Upsilon_{(F,\mathcal{G})}(x_0)$ that satisfies $x(t) \in E$ for all $t \in [0, \mathrm{Esc}\,(x(\cdot); \mathcal{G}))$.
(b) (F, E) is called *strongly invariant in* $\mathcal{G}$ provided for every $x_0 \in E$, every trajectory $x(\cdot) \in \Upsilon_{(F,\mathcal{G})}(x_0)$ satisfies $x(t) \in E$ for all $t \in [0, \mathrm{Esc}\,(x(\cdot); \mathcal{G}))$.

For Hamiltonian characterizations of strong invariance for *locally Lipschitz dynamics*, see [7]. See also [10] for a characterization of strong invariance for systems satisfying appropriate one-sided Lipschitzness and dissipativity conditions, and [11] for general one-sided Lipschitz dynamics (with a modified Hamiltonian). Our main contribution will be a new sufficient condition for strong invariance for dynamics satisfying the realizability condition (U), including cases where F is neither lower nor upper semicontinuous and not tractable by the known strong invariance results. Our condition is a Hamiltonian inequality involving a lower semicontinuous verification function. However, necessity is not true generally, and it is not clear what or if a modification of the inequality can be made to ensure a complete characterization.

The following is a variant of the well known "compactness of trajectories" lemma. This result says more than just that a bounded set of solutions is relatively compact. Rather, a stronger conclusion holds in that approximate trajectories have subsequences that converge to a trajectory. The proof is a special case of the compactness of trajectories proof in [7].

Lemma 2.7. Let $\bar{x} \in \mathbb{R}^n$, $T > 0$, $\tilde{f} \in \mathcal{C}[0, T]$ be also continuous in t, and $\{y_i : [0, T] \to \mathbb{R}^n\}$ be a sequence of uniformly bounded absolutely continuous functions satisfying $y_i(0) = \bar{x}$ for all i. Assume

$$\dot{y}_i(t) \in \tilde{f}(\tau_i(t), y_i(t) + r_i(t)) + \delta_i(t)\mathcal{B}_n \text{ for a.a. } t \in [0, T] \tag{4}$$

for all i, where $\{\delta_i(\cdot)\}$ is a sequence of nonnegative measurable functions that converges to 0 in L^2 as $i \to \infty$, $\{r_i(\cdot)\}$ is a sequence of measurable functions converging uniformly to 0 as $i \to \infty$, and $\{\tau_i(\cdot)\}$ is a sequence of measurable functions converging uniformly to t on $[0, T]$ as $i \to \infty$. Then there exists a trajectory y of $\dot{y} = \tilde{f}(t, y)$, $y(0) = \bar{x}$ such that a subsequence of y_i converges to y uniformly on $[0, T]$.

We will apply Lemma 2.7 to continuous mollifications of our feedback maps $f \in \mathcal{C}[0, T]$. More precisely, set

$$\eta(t) = \begin{cases} C \exp\left(\frac{1}{t^2-1}\right), & |t| < 1 \\ 0, & |t| \geq 1 \end{cases}$$

where the constant $C > 0$ is chosen so that $\int_{\mathbb{R}} \eta(s)ds = 1$. For each $\varepsilon > 0$ and $t \in \mathbb{R}$, set

$$\eta_\varepsilon(t) := \frac{1}{\varepsilon}\eta\left(\frac{t}{\varepsilon}\right).$$

Notice for future use that

$$\int_{\mathbb{R}} \eta_\varepsilon(t)dt = 1 \ \ \forall \varepsilon > 0. \tag{5}$$

Define the following convolutions of $f \in \mathcal{C}[0,T]$ in the t-variable:

$$f_\varepsilon(t,x) := \int_{\mathbb{R}} f(s,x)\eta_\varepsilon(t-s)ds \ \left(= \int_{\mathbb{R}} f(t-s,x)\eta_\varepsilon(s)ds \right) \tag{6}$$

with the convention that $f(s,x) = 0$ for $s \notin [0,T]$. Then $f_\varepsilon \in \mathcal{C}[0,T]$ and is continuous for all $\varepsilon > 0$. (In fact, f_ε is a C^∞ function of t for each $x \in \mathbb{R}^n$, but we will not need this fact. See [12, 13] for the well known theory of convolutions and mollifiers.) We will apply Lemma 2.7 to a sequence $\tilde{f} := f_{\varepsilon(i)}$ with $\varepsilon(i) > 0$ converging to zero. In this case, we will use ideas from the standard proof that

$$f_{\varepsilon(i)}(\cdot,x) \to f(\cdot,x) \ \text{ in } \ L_1[0,T] \ \ \forall x \in \mathbb{R}^n$$

to build trajectories of $f \in \mathcal{C}[0,T]$ that respect the state constraint.

Remark 2.8. Note for later use that if $\tau_i(t) \equiv t$ in Lemma 2.7, then the conclusions of the lemma remain true even if the t-continuity hypothesis on $f \in \mathcal{C}[0,T]$ is omitted. This follows from the proof of the compactness of trajectories lemma in [7].

3 Strong Invariance Theorem

3.1 Statement of Theorem and Remarks

Let $H_F : \mathbb{R}^n \times \mathbb{R}^n \to [-\infty, +\infty]$ denote the *(upper) Hamiltonian* for our dynamics F; i.e.,

$$H_F(x,p) := \sup_{v \in F(x)} \langle v, p \rangle.$$

For any subset $D \subseteq \mathbb{R}^n$, we write $H_F(x,D) \le 0$ to mean that $H_F(x,d) \le 0$ for all $d \in D$. By definition, this inequality holds vacuously if $D = \emptyset$, e.g., if D is the empty set of proximal subdifferentials of a function at some point.

Theorem 2. *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfy (U), let $\Psi : \mathbb{R}^n \to (-\infty, +\infty)$ be lower semicontinuous, and set $\mathcal{S} = \{x \in \mathbb{R}^n : \Psi(x) \le 0\}$. If there exists an open set $\mathcal{U} \subseteq \mathbb{R}^n$ containing $\mathcal{S}$ for which $H_F(x, \partial_P \Psi(x)) \le 0$ for all $x \in \mathcal{U}$, then $(F, \mathcal{S})$ is strongly invariant in $\mathbb{R}^n$.*

We defer the proof of this theorem to section 4. Theorem 2 differs from the usual strong invariance statements in the manner in which the set S is described, but it allows for some interplay between constraint and data assumptions. Note that we require the Hamiltonian inequality in a *neighborhood* $\mathcal{U}$ of $\mathcal{S}$, for the result is not true in general if the Hamiltonian condition is placed only on $\mathcal{S}$, even if Ψ and F are smooth. For example, take $n = 1$, $\Psi(x) = x^2$, and $F(x) \equiv \{1\}$. In this case, $\mathcal{S} = \{0\}$ and $H_F(0, \partial_P \Psi(0)) = 0$, but $(F, \mathcal{S})$ is not strongly invariant. On the other hand, Example 2.2 is covered by Theorem 2, once we choose the verification function $\Psi(x) = x^2$. In this case, the Hamiltonian condition reads $H_F(x, \Psi'(x)) = -2x \operatorname{sign}(x) = -2|x| \le 0$ for all $x \in \mathbb{R}$, so our sufficient condition for strong invariance is satisfied.

Theorem 2 contains the usual sufficient condition for strong invariance for an arbitrary closed set $S \subseteq \mathbb{R}^n$ by letting Ψ be the characteristic function I_S of F; that is, $I_S(x) = 0$ if $x \in S$ and is 1 otherwise. Then $\partial_P \Psi(x) = \{0\}$ for all $x \notin \text{boundary}\,(S)$, and $\partial_P \Psi(x) = N_S^P(x)$ for all $x \in \text{boundary}\,(S)$. This implies the following special case of Theorem 2:

Corollary 3.1. Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfy (U) and $S \subseteq \mathbb{R}^n$ be closed. If $H_F(x, N_S^P(x)) \le 0$ for all $x \in \text{boundary}\,(S)$, then (F, S) is strongly invariant in $\mathbb{R}^n$.

The converse of Corollary 3.1 does not hold, as illustrated by the simple example given in the introduction. This means that the converse of Theorem 2 does not hold.

Remark 3.2. Theorem 2 remains true (by the same proof) if its Hamiltonian inequality is replaced by: $\langle f(t,x), p\rangle \le 0$ for all $T \ge 0$, $\bar{x} \in \mathbb{R}^n$, $f \in \mathcal{C}_F([0,T],\bar{x})$, $t \in [0,T]$, $x \in \mathcal{U}$, and $p \in \partial_P \Psi(x)$.

3.2 Relationship to Known Strong Invariance Results

Theorem 2 improves on the known strong invariance results because it does not require the usual Lipschitz or other structural assumptions on the dynamics. The papers [4, 6, 14, 15] provide strong invariance results for locally Lipschitz dynamics (see also [7, Chapter 4]). In [4], Clarke showed that strong invariance of (F, S) in $\mathbb{R}^n$ is equivalent to

$$F(x) \subseteq T_S^C(x) \quad \forall x \in S \tag{7}$$

where T_S^C is the Clarke tangent cone (cf. [7]). Recall that $v \in T_S^C(x)$ if and only if for all sequences $x_i \in S$ converging to x and all sequences $t_i > 0$ decreasing to 0, there exists a sequence $v_i \in \mathbb{R}^n$ converging to v such that $x_i + t_i v_i \in S$ for all i. In particular, if $S = \{0\}$, then $T_S^C(0) = \{0\}$. Later, Krastanov [15] gave an infinitesimal characterization of normal-type, by showing strong invariance is equivalent to the following: $H_F(x, N_S^P(x)) \le 0$ for all $x \in S$. See [3, 6] for Hilbert space versions, and [14, 21] for other strong invariance results for Lipschitz dynamics and nonautonomous versions.

Donchev [9] extended these characterizations beyond the autonomous Lipschitz case to "almost continuous, one-sided Lipschitz" multifunctions. Rios and Wolenski [18] proved an autonomous normal-type characterization that allows for a discontinuous component. Donchev, Rios, and Wolenski [10] proved a necessary and sufficient condition for strong invariance for a discontinuous nonautonomous differential inclusion $F : \mathbb{R}^n \times \mathcal{I} \rightrightarrows \mathbb{R}^n$ whose right-hand side is the sum of an almost upper semicontinuous dynamic $D(t,x)$ with nonempty compact convex values that is dissipative in x, and an almost lower semicontinuous multifunction $G(t,x)$ that is one-sided Lipschitz in x. In terms of the nonautonomous Hamiltonians defined for any dynamics R by

$$H_R(t,x,\zeta) := \sup_{v \in R(t,x)} \langle v, \zeta\rangle,$$

the main result of [10] says: If $S \subseteq \mathbb{R}^n$ is closed, then $(D+G, S)$ is strongly invariant in $\mathbb{R}^n$ if and only if there exists a subset $I \subseteq \mathcal{I}$ of full measure in $\mathcal{I}$ such that

$$H_G(t,x,N_S^P(x)) - H_D(t,x,-N_S^P(x)) \;\le\; 0 \quad \forall (t,x) \in I \times S.$$

This result applies to cases where the Clarke tangency condition (7) is not satisfied, and covers the example in the introduction. They have gone further in [11] to provide a characterization of the general one-sided Lipschitz case, in which the Hamiltonian is replaced by a limiting condition.

On the other hand, Theorem 2 does not make any structural assumptions on the dynamics. Moreover, our feedback realizability hypothesis (U) can be satisfied for dynamics that are not tractable by the well known strong invariance results. For instance, see the examples in section 2.

4 Proof of Strong Invariance Theorem

This section is devoted to the proof of Theorem 2.

Fix $T > 0$ and $\bar{x} \in \mathcal{S}$. We first develop some properties that hold for all $f \in \mathcal{C}_F([0,T],\bar{x})$. Fixing $f \in \mathcal{C}_F([0,T],\bar{x})$ and $\varepsilon > 0$, and fixing $\gamma > 0$ such that $f(t,x) \in \text{cone}\{F(x)\}$ for all $x \in \gamma\mathcal{B}_n(\bar{x})$ and almost all $t \in [0,T]$, set

$$G_f^\varepsilon[t,x,k] \;=\; \overline{co}\left\{ f_\varepsilon(t,y) : \|y-x\| \le \frac{1}{k} \right\} \;\subseteq\; \mathbb{R}^n \tag{8}$$

for each $t \in [0,T]$, $x \in \mathbb{R}^n$ and $k \in \mathbb{N}$, where f_ε is the regularization of f defined by (6) and $\overline{co}$ denotes the closed convex hull. This is well defined because $f \in \mathcal{C}[0,T]$. By reducing $\gamma > 0$ as necessary, we can assume that $\gamma\mathcal{B}_n(\bar{x}) \subseteq \mathcal{U}$. We also set

$$g_f^\varepsilon[t,x,k] = 1 + \max\{\|p\| : p \in G_f^\varepsilon[t,x,k]\}$$

for all $t \in [0,T]$, $x \in \mathbb{R}^n$, and $k \in \mathbb{N}$. Note that

$$g_f^\varepsilon[t,x,k] \le g_f[t,x,k] := 1 + c_1 + c_2\left(||x|| + \frac{1}{k}\right) \qquad \forall t \in [0,T], x \in \mathbb{R}^n, k \in \mathbb{N} \tag{9}$$

where c_1 and c_2 are the constants from the linear growth requirement on f, so the sets $G_f^\varepsilon[t,x,k]$ are compact. The following estimate is based on Theorem 1 from section 2:

Claim 4.1. If $x \in \frac{\gamma}{2}\mathcal{B}_n(\bar{x})$, $t \ge 0$, $k \in \mathbb{N}$, and $h > 0$ are such that

$$0 < h \le \frac{1}{2k\, g_f[t,x,k]} \quad \text{ and } \quad x + hg_f[t,x,k]\mathcal{B}_n \subseteq \frac{2\gamma}{3}\mathcal{B}_n(\bar{x}), \tag{10}$$

then

$$\Psi(x + hv) \le \Psi(x) + \frac{h}{k} \tag{11}$$

holds for some $v \in G_f^\varepsilon[t,x,k]$.

Proof. Suppose the contrary. Fix $x \in \frac{\gamma}{2}\mathcal{B}_n(\bar{x})$, $t \ge 0$, $k \in \mathbb{N}$, and $h > 0$ satisfying (10) but such that

$$\Psi(x + hv) > \Psi(x) + \frac{h}{k} \quad \forall v \in G_f^\varepsilon[t,x,k]. \tag{12}$$

It follows that

$$\delta := \frac{h}{2k} < \min_{y \in Y} \Psi(y) - \Psi(x), \tag{13}$$

where $Y = x + hG_f^\varepsilon[t,x,k]$. This is because Ψ is lower semicontinuous. Let $\lambda \in (0, \frac{1}{2k})$ be such that

$$x + hg_f[t,x,k]\mathcal{B}_n + \lambda\mathcal{B}_n \subseteq \gamma\mathcal{B}_n(\bar{x}). \tag{14}$$

Next we apply Theorem 1 with the choices $Y = x + hG_f^\varepsilon[t,x,k]$ and δ defined by (13). It follows that there exist $z \in [x, Y] + \lambda\mathcal{B}_n$ and $\zeta \in \partial_P\Psi(z)$ for which

$$\delta < \min_{y \in Y}\langle \zeta, y - x\rangle = \min_{v \in G_f^\varepsilon[t,x,k]} \langle \zeta, hv\rangle. \tag{15}$$

Note that $z \in \gamma\mathcal{B}_n(\bar{x}) \subseteq \mathcal{U}$, by (14). Since $z \in [x, Y] + \lambda\mathcal{B}_n$, (10) combined with the choice of λ gives

$$\|z - x\| \le hg_f[t,x,k] + \lambda \le \frac{1}{k}.$$

Therefore $f_\varepsilon(t,z) \in G_f^\varepsilon[t,x,k]$, by the definition (8) of $G_f^\varepsilon[t,x,k]$. Since $f(s,z) \in \text{cone}\{F(z)\}$ for a.a. $s \in [0,T]$ (by our choice of $\gamma > 0$), our Hamiltonian hypothesis gives $\langle \zeta, f(s,z)\rangle \le 0$ for almost all $s \in [0,T]$. Therefore, (15) gives

$$0 \;<\; \delta \;\le\; h\langle \zeta, f_\varepsilon(t,z)\rangle \;=\; h\int_{\mathbb{R}} \eta_\varepsilon(t-s)\langle \zeta, f(s,z)\rangle ds \;\le\; 0. \tag{16}$$

The contradiction (16) concludes the proof of Claim 4.1. □

Now set

$$D := \frac{\gamma}{2}\mathcal{B}_n(\bar{x}) \subseteq \mathcal{U}. \tag{17}$$

Let $\omega_{f,K}$ be a modulus of continuity for $x \mapsto f(t,x)$ on $K := D + \mathcal{B}_n$ for all $t \in [0,T]$. Such a modulus exists by condition (C_2). Then $\omega_{f,K}$ is also a modulus of continuity of $K \ni x \mapsto f_\varepsilon(t,x)$ for all $t \in [0,T]$ and $\varepsilon > 0$. The following estimate follows from Carathéodory's Lemma (cf. [19, p. 55]):

Claim 4.2. Let $(t,x,k) \in [0,T] \times D \times \mathbb{N}$ and $v \in G_f^\varepsilon[t,x,k]$. Then $\|v - f_\varepsilon(t,x)\| \le \omega_{f,K}(1/k) + 1/k$.

Proof. Using the Carathéodory Lemma and the definition of $G_f^\varepsilon[t,x,k]$, we can write

$$v = \Delta + \sum_{j=0}^{n} \alpha_j f_\varepsilon(t,x_j),$$

where

$$\alpha_j \in [0,1] \;\; \forall j, \quad \sum_{j=0}^{n} \alpha_j = 1, \quad \|x_j - x\| \le 1/k \; \forall j, \quad ||\Delta|| \le 1/k.$$

In particular, $x_j \in K$ for all j. This gives

$$||v - f_\varepsilon(t,x)\| \;\le\; \left\| v - \sum_{j=0}^{n} \alpha_j f_\varepsilon(t,x_j) \right\| \;+\; \left\| \sum_{j=0}^{n} \alpha_j \{ f_\varepsilon(t,x_j) - f_\varepsilon(t,x) \} \right\| \;\le\; \frac{1}{k} + \omega_{f,K}\left(\frac{1}{k}\right),$$

as desired. □

Next define

$$\delta(D) := 1 + c_1 + c_2 + c_2 \sup\{\|v\| : v \in D\}.$$

It follows from the estimate (9) that

$$G_f^\varepsilon[t,x,k] \subseteq \delta(D)\mathcal{B}_n \;\; \forall\, t \in [0,T], \;\; x \in D, \;\; k \in \mathbb{N}. \tag{18}$$

Next set

$$\tilde{T} := \min\left\{T, \frac{\gamma}{8\delta(D)}\right\}, \quad \text{and} \quad h_k := \frac{\gamma}{4k\delta(D)} \qquad \forall k \in \mathbb{N}. \tag{19}$$

Choose $N > 2$ such that

$$D + h_k\delta(D)\mathcal{B}_n \;\subseteq\; \frac{2\gamma}{3}\mathcal{B}_n(\bar{x}) \quad \forall k \ge N. \tag{20}$$

By the choices of γ and $\delta(D)$,

$$0 \;<\; h_k \;\le\; \frac{1}{2kg_f[t,x,k]} \qquad \forall t \in [0,T], \; x \in D, \; k \in \mathbb{N}. \tag{21}$$

Set $c(k) = \text{Ceiling}(\tilde{T}/h_k)$, i.e., $c(k)$ is the smallest integer $\ge \tilde{T}/h_k$. For each $k \ge N$, we then define a partition

$$\pi(k) : 0 = t_{0,k} < t_{1,k} < \ldots < t_{c(k),k} = \tilde{T}$$

by setting $t_{i,k} = t_{i-1,k} + h_k$ for $i = 1, 2, \ldots, c(k) - 1$. Since

$$\tilde{T} - h_k \le [c(k) - 1]h_k \le \tilde{T}$$

for all k, it follows that $t_{c(k),k} - t_{c(k)-1,k} \le h_k$ for all k. We also define sequences

$$x_{0,k}, \; x_{1,k}, \; x_{2,k}, \ldots, \; x_{c(k),k} \;\in\; \mathbb{R}^n \tag{22}$$

for $k \geq N$ as follows. We set $x_{0,k} = \bar{x}$ and $x_{1,k} = \bar{x} + (t_{1,k} - t_{0,k})v_{o,k}$, where $v = v_{o,k} \in G_f^\varepsilon[0, \bar{x}, k]$ satisfies the requirement from Claim 4.1 for the pair $(t_{0,k}, x_{0,k}) = (0, \bar{x})$. By (18), we get

$$\|x_{1,k} - \bar{x}\| \;\leq\; h_k \delta(D) \;=\; \frac{\gamma}{4k}, \tag{23}$$

so $x_{1,k} \in D$. If $c(k) \geq 2$, then we set $x_{2,k} = x_{1,k} + (t_{2,k} - t_{1,k})v_{1,k}$, where $v_{1,k} \in G_f^\varepsilon[t_{1,k}, x_{1,k}, k]$ satisfies the requirement from Claim 4.1 for the pair $(t_{1,k}, x_{1,k})$. Reapplying (18) gives

$$\|x_{2,k} - x_{1,k}\| \leq h_k \delta(D) = \frac{\gamma}{4k},$$

so (23) gives

$$\|x_{2,k} - \bar{x}\| \leq \frac{\gamma}{2k}.$$

Therefore $x_{2,k} \in D$. We now repeat this process using $x_{2,k} \in D$ instead of $x_{1,k}$. Proceeding inductively gives sequences

$$v_{i,k} \in G_f^\varepsilon[t_{i,k}, x_{i,k}, k]$$

and sequences (22) that satisfy

$$x_{i+1,k} = x_{i,k} + (t_{i+1,k} - t_{i,k})v_{i,k}$$

for $i = 0, 1, \ldots, c(k) - 1$. In this case,

$$\frac{c(k)\gamma}{4k} \leq \left(\frac{\tilde{T}}{h_k} + 1\right) \frac{\gamma}{4k} \;=\; \tilde{T}\delta(D) + \frac{\gamma}{4k} \leq \frac{\gamma}{2},$$

by the choices of $\tilde{T}$ and $k \geq 2$. Therefore,

$$\|x_{i,k} - \bar{x}\| \;\leq\; \frac{c(k)\gamma}{4k} \;\leq\; \frac{\gamma}{2}$$

for all i and k. It follows that the sequences $\{x_{i,k}\}$ lie in D.

For each $k \geq N$, we then choose $x_{\pi(k)}$ to be the polygonal arc satisfying $x_{\pi(k)}(0) = \bar{x}$ and

$$\dot{x}_{\pi(k)}(t) = f_\varepsilon(\tau_k(t), x_{\pi(k)}(t) + r_k(t)) + z_k(\tau_k(t)) \tag{24}$$

for all $t \in [0, \tilde{T}] \setminus \pi(k)$, where $\tau_k(t)$ is the partition point $t_{i,k} \in \pi(k)$ immediately preceding t for each $t \in [0, \tilde{T}]$,

$$z_k(t_{i,k}) := v_{i,k} - f_\varepsilon(t_{i,k}, x_{\pi(k)}(t_{i,k})) \;\; \forall i, k \tag{25}$$

the $v_{i,k} \in G_f^\varepsilon[t_{i,k}, x_{\pi(k)}(t_{i,k}), k]$ satisfy the conclusions from Claim 4.1, and

$$r_k(t) := x_{\pi(k)}(\tau_k(t)) - x_{\pi(k)}(t) \tag{26}$$

for all $t \in [0, \tilde{T}]$ and k. Then $x_{\pi(k)}$ is the polygonal arc connecting the points $x_{i,k}$ for $i = 0, 1, 2, \ldots, c(k)$. In particular, $x_{i,k} \equiv x_{\pi(k)}(t_{i,k})$.

Since $x_{i,k} \in D$ for all i and k, we conclude from Claim 4.2 that

$$\sup\{\|z_k(\tau_k(t))\| : t \in [0, \tilde{T}]\} \leq \omega_{f,K}(1/k) + 1/k \to 0$$

as $k \to +\infty$. Since $||\dot{x}_{\pi(k)}(t)|| = ||v_{i,k}|| \leq \delta(D)$ for all $t \in (t_{i,k}, t_{i+1,k})$ and all i and k, and since

$$|\tau_k(t) - t| \leq h_k \to 0 \;\text{ as }\; k \to +\infty \;\; \forall t \in [0, \tilde{T}] \tag{27}$$

we get

$$\sup\{\|r_k(t)\| : t \in [0, \tilde{T}]\} \to 0 \;\text{ as }\; k \to +\infty.$$

Since (24) has the form (4) from our compactness of trajectories lemma and f_ε is continuous, we can find a subsequence of $\{x_{\pi(k)}(\cdot)\}$ that converges uniformly to a trajectory y_ε of

$$\dot{y} = f_\varepsilon(t, y), \quad y(0) = \bar{x}.$$

By possibly passing to a subsequence without relabelling, we can assume that

$$x_{\pi(k)} \to y_\varepsilon \ \text{ uniformly } \ \text{on } \ [0, \tilde{T}]. \tag{28}$$

Moreover, since $x_{i+1,k} = x_{i,k} + (t_{i+1,k} - t_{i,k})v_{i,k} \in D$ for all $i = 0, 1, \dots, c(k) - 1$ and $k \geq N$, conditions (20) and (21) along with Claim 4.1 give

$$\begin{array}{rcl} \Psi(x_{1,k}) & \leq & \Psi(\bar{x}) + \frac{h_k}{k} \\ \Psi(x_{2,k}) & \leq & \Psi(x_{1,k}) + \frac{h_k}{k} \\ \vdots & & \\ \Psi(x_{c(k),k}) & \leq & \Psi(x_{c(k)-1,k}) + \frac{h_k}{k}. \end{array}$$

Summing these inequalities and recalling that $h_k \leq \gamma$ gives

$$\begin{array}{rcl} \Psi(x_{i,k}) & \leq & \Psi(\bar{x}) + \frac{c(k)h_k}{k} \\ & \leq & \Psi(\bar{x}) + \left(\frac{\tilde{T}}{h_k} + 1\right) \frac{h_k}{k} \\ & \leq & \Psi(\bar{x}) + \frac{1}{k}(\tilde{T} + \gamma) \end{array}$$

for all i and k . Hence,

$$\Psi(x_{\pi(k)}(\tau_k(t))) \leq \Psi(\bar{x}) + \frac{1}{k}(\tilde{T} + \gamma) \tag{29}$$

for all $t \in [0, \tilde{T}]$. It follows from (27) and (28) that

$$x_{\pi(k)}(\tau_k(t)) \to y_\varepsilon(t) \ \ \forall t \in [0, \tilde{T}] \ \text{ as } \ k \to +\infty. \tag{30}$$

Since Ψ is lower semicontinuous, it follows from (29)-(30) that $\Psi(y_\varepsilon(t)) \leq \Psi(\bar{x})$ for all $t \in [0, \tilde{T}]$.

Now we consider a *sequence* $\{\varepsilon(i)\}$ of positive numbers converging to zero. Let $y_i := y_{\varepsilon(i)} : [0, \tilde{T}] \to \mathbb{R}^n$ be the trajectories obtained by the preceding argument for $\varepsilon = \varepsilon(i)$ for all $i \in \mathbb{N}$. Note that $y_i(t) \in D$ for all i and t, because each of the polygonal arcs $x_{\pi(k)}$ constructed above joins points in D and D is closed and convex. Moreover,

$$\dot{y}_i(t) = f(t, y_i(t)) + \left[f_{\varepsilon(i)}(t, y_i(t)) - f(t, y_i(t))\right] \tag{31}$$

for all i and almost all $t \in [0, \tilde{T}]$. Since $||\dot{x}_{\pi(k)}(t)|| \leq \delta(D)$ for all k and a.a. $t \in [0, \tilde{T}]$ for all the polygonal arcs $x_{\pi(k)}$ defined above, we get

$$||y_i(t) - y_i(s)|| \leq \delta(D)(t - s) \ \text{ for } \ 0 \leq s \leq t \leq \tilde{T} \ \text{ for all } \ i. \tag{32}$$

Since the y_i are uniformly bounded and equicontinuous, we can assume (possibly by passing to a subsequence without relabelling) that there is a continuous function $y : [0, \tilde{T}] \to D$ such that

$$y_i \to y \ \text{ uniformly on } [0, \tilde{T}]. \tag{33}$$

We next show that y is a trajectory of f. To this end, we prove the following claim:

Claim 4.3. $f_{\varepsilon(i)}(t, y_i(t)) - f(t, y_i(t)) \to 0$ in $L^2[0, \tilde{T}]$ as $i \to \infty$.

Proof. Since $y_i \to y$ uniformly on $[0, \tilde{T}]$ and f is locally bounded and locally uniformly continuous in the x variable, it suffices to show that

$$\tilde{\delta}_i(t) := f_{\varepsilon(i)}(t, y(t)) - f(t, y(t)) \to 0 \text{ in } L^1[0, \tilde{T}] \tag{34}$$

as $i \to \infty$. We do this by adapting a standard mollification argument (see for example [13, Chapter 8]) as follows. We first extend y to all of $\mathbb{R}$ by defining $y(t) \equiv y(0)$ for all $t \leq 0$ and $y(t) \equiv y(\tilde{T})$ for all $t \geq \tilde{T}$. Recall that we also defined $f(s, x) \equiv 0$ for all $s \notin [0, T]$. It follows from (32) and (33) that

$$\mu(i) := 2\omega_{f,D} \left(\sup \{||y(t - \varepsilon(i)z) - y(t)|| : |z| \leq 1, t \in \mathbb{R}\}\right) \to 0 \tag{35}$$

as $i \to \infty$. Moreover, for $0 \leq t \leq \tilde{T}$, we can change variables to get

$$\begin{aligned} \|\tilde{\delta}_i(t)\| &\leq \int_{\mathbb{R}} \|f(t-s, y(t)) - f(t, y(t))\| \eta_{\varepsilon(i)}(s) ds \ \text{ (by (5))} \\ &= \int_{-1}^{1} \|f(t - \varepsilon(i)z, y(t)) - f(t, y(t))\| \eta(z) dz \\ &\leq \int_{-1}^{1} \|f(t - \varepsilon(i)z, y(t - \varepsilon(i)z)) - f(t, y(t))\| \eta(z) dz + \mu(i) \\ &= \int_{-1}^{1} \|g^{\varepsilon(i)z}(t) - g(t)\| \eta(z) dz + \mu(i) \end{aligned}$$

where

$$g(t) := f(t, y(t)), \quad g^{\varepsilon(i)z}(t) := g(t - \varepsilon(i)z).$$

Notice that $g \in L^1[0, \tilde{T}]$, and that

$$||g^{\varepsilon(i)z} - g||_1 \leq 2||g||_1 \quad \forall i.$$

Applying Minkowski's inequality for $L^1[0, \tilde{T}]$ therefore gives

$$||\tilde{\delta}_i||_1 \leq \int_{-1}^{1} ||g^{\varepsilon(i)z} - g||_1 \eta(z) dz + \tilde{T}\mu(i).$$

Moreover,

$$||g^{\varepsilon(i)z} - g||_1 \to 0$$

as $i \to \infty$ for each $z \in [-1, +1]$, by continuity of translation in the L^1 norm (see [13, Proposition 8.5]). The desired convergence (34) therefore follows from (35) and the dominated convergence theorem. □

It therefore follows from Remark 2.8 that a subsequence of $\{y_i\}$ converges to a trajectory of f uniformly on $[0, \tilde{T}]$. This must be the aforementioned function y, as desired. Again using the lower semicontinuity of Ψ, we therefore get

$$\Psi(y(t)) \leq \liminf_{i \to \infty} \Psi(y_{\varepsilon(i)}(t)) \leq \Psi(\bar{x})$$

for all $t \in [0, \tilde{T}]$.

The strong invariance asserted in the theorem is now immediate. Indeed, let $x_o \in \mathcal{S}$, $T \geq 0$, and $\phi \in \text{Traj}_T(F, x_o)$ be given. We next show that

$$\bar{t} \ := \ \sup \{t \geq 0 : \Psi(\phi(s)) \leq \Psi(x_o) \text{ for } 0 \leq s \leq t\} \ = \ T, \tag{36}$$

which would imply that ϕ remains in $\mathcal{S}$ on $[0, T]$. To this end, note that if this supremum were some time $\bar{t} \in [0, T)$, then the lower semicontinuity of Ψ would give

$$\Psi(\phi(\bar{t})) \ \leq \ \Psi(x_o). \tag{37}$$

In particular, $\bar{x} := \phi(\bar{t}) \in \mathcal{S}$. Let $f \in \mathcal{C}_F([0,T],\bar{x})$ satisfy the requirement (U) for F and the trajectory

$$[0, T-\bar{t}] \ni t \mapsto y(t) := \phi(t+\bar{t}),$$

and let $\gamma > 0$ be such that $f(t,x) \in \text{cone}\{F(x)\}$ for almost all $t \in [0, T-\bar{t}]$ and all $x \in \gamma\mathcal{B}_n(\bar{x})$. By reducing $\gamma > 0$ as necessary, we can assume that $\gamma\mathcal{B}_n(\bar{x}) \subseteq \mathcal{U}$.

By uniqueness of solutions of the initial value problem

$$\dot{y} = f(t,y), \quad y(0) = \phi(\bar{t})$$

on $[0, T-\bar{t}]$, the first part of the proof applied to f and the initial value $\bar{x} = \phi(\bar{t}) \in \mathcal{S}$ would give $\tilde{t} \in (0, T-\bar{t}\,)$ such that

$$\Psi(\phi(\bar{t}+t\,)) - \Psi(\phi(\bar{t}\,)) \quad \leq \quad 0 \quad \forall t \in [0, \tilde{t}]. \tag{38}$$

Here we use the fact that the trajectory on $[0, \tilde{T}]$ constructed above for f starting at $\bar{x}$ can be extended to $[0, T-\bar{t}]$, by the linear growth assumption (C_3), and therefore coincides with y by our uniqueness assumption in (U). Since ϕ remains in $\mathcal{S}$ on $[0, \bar{t}]$, summing (37)-(38) would then contradict the definition of the supremum $\bar{t}$. This establishes (36) and proves the theorem.

Acknowledgement: This work has been funded in part by the US National Academy of Sciences under the Collaboration in Basic Science and Engineering (COBASE) Program, supported by Contract No. INT-0002341 from the US National Science Foundation. The contents of this publication do not necessarily reflect the views or policies of the US National Academy of Sciences or the US National Science Foundation, nor does mention of trade names, commercial products or organizations imply endorsement by the US National Academy of Sciences or the US National Science Foundation. The first author (Krastanov) was also supported by Swiss NSF Contract 7 IP 65642 (Program SCOPES) and by the Bulgarian Ministry of Science and Higher Education - National Fund for Science Research under contracts MM-807/98 and MM-1104/01. The second author (Malisoff) was also supported by the Louisiana Board of Regents Support Fund Grant LEQSF(2003-06)-RD-A-12.

References

[1] D. Angeli and E. Sontag, *Monotone Control Systems.* IEEE Trans. Autom. Control **48** (2003), 1684–1698.

[2] M. Bardi and I. Capuzzo–Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations.* Birkhäuser, Boston, 1997.

[3] H. Brezis, *On a characterization of flow-invariant sets.* Commun. Pure Appl. Math. **23** (1970), 261–263.

[4] F. Clarke, *Generalized gradients and applications.* Trans. Amer. Math. Soc. **205** (1975), 247–262.

[5] F. Clarke and Y. Ledyaev, *Mean value inequalities in Hilbert space.* Trans. Amer. Math. Soc. **344** (1994), 307–324.

[6] F. Clarke, Y. Ledyaev and M. Radulescu, *Approximate invariance and differential inclusions in Hilbert spaces.* J. Dynam. Control Syst. **3** (1997), 493–518.

[7] F. Clarke, Y. Ledyaev, R. Stern and P. Wolenski, *Nonsmooth Analysis and Control Theory.* Springer-Verlag, New York, 1998.

[8] F. Clarke and R. Stern, *State constrained feedback stabilization.* SIAM J. Control Optim. **42** (2003), 422–441.

[9] T. Donchev, *Properties of the reachable set of control systems.* Systems and Control Letters **46** (2002), 379–386.

[10] T. Donchev, V. Rios and P. Wolenski, *A characterization of strong invariance for perturbed dissipative systems.* Optimal Control, Stabilization, and Nonsmooth Analysis (M. de Queiroz, M. Malisoff and P. Wolenski, Eds.), Lecture Notes in Control and Information Sciences, Springer-Verlag, Heidelberg, to appear.

[11] T. Donchev, V. Rios and P. Wolenski, *Strong invariance and the one-sided Lipschitz property*, submitted.

[12] L. Evans, *Partial Differential Equations. Graduate Studies in Mathematics, 19.* American Mathematical Society, Providence, RI, 1998.

[13] G. Folland, *Real Analysis. Modern Techniques and their Applications.* John Wiley and Sons, New York, 1984.

[14] H. Frankowska, S. Plaskacz and T. Rzezuchowski, *Measurable viability theorems and the Hamilton-Jacobi-Bellman equation.* Journal of Differential Equations **116** (1995), 265-305.

[15] M. Krastanov, *Forward invariant sets, homogeneity and small-time local controllability.* Geometry in Nonlinear Control and Differential Inclusions, Banach Center Publ. 32, Polish Acad. Sci., Warsaw, 1995, 287–300.

[16] M. Malisoff, *A remark on the Bellman equation for optimal control problems with exit times and noncoercing dynamics.* Proceedings of the 38th IEEE Conference on Decision and Control (Phoenix, AZ, December 1999), 877–881.

[17] M. Malisoff, *Viscosity solutions of the Bellman equation for exit time optimal control problems with non–Lipschitz dynamics.* ESAIM: Control, Optimisation and Calculus of Variations **6** (2001), 415–441.

[18] V. Rios and P. Wolenski, *A characterization of strongly invariant systems for a class of non-Lipschitz multifunctions.* Proceedings of the 42nd IEEE Conference on Decision and Control (Maui, HI, December 2003), to appear.

[19] R. Rockafellar and R. Wets, *Variational analysis.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] Vol. 317, Springer-Verlag, Berlin, 1998.

[20] H. Sussmann, *Uniqueness results for the value function via direct trajectory-construction methods.* Proceedings of the 42nd IEEE Conference on Decision and Control (Maui, HI, December 2003), to appear.

[21] R. Vinter, *Optimal Control.* Birkhäuser, Boston, 2000.

[22] P. Wolenski and Y. Zhuang, *Proximal analysis and the minimal time function.* SIAM J. Control Optim. **36** (1998), 1048–1072.