united nations educational, scientific and cultural organization

international atomic energy agency

the
abdus salam
international centre for theoretical physics

SMR1327/5

Summer School on Mathematical Control Theory

(3 - 28 September 2001)

Value Function in Optimal Control

Hélene Frankowska
CNRS, CREA
Ecole Polytechnique
1, rue Descartes
75005 Paris
France

SUMMER SCHOOL ON MATHEMATICAL CONTROL THEORY
ICTP, TRIESTE, September 2001

Value Function in Optimal Control

Hélène Frankowska
CNRS, CREA,
ECOLE POLYTECHNIQUE
1, rue Descartes, 75005 PARIS, FRANCE

Contents

Chapter 1

Set-Valued Analysis

This chapter is concerned with the *differential inclusion* (*multivalued equation*):

$$x'(t) \in F(t, x(t)), \quad x(t_0) = x_0 \tag{1.1}$$

Its investigation was initiated in the thirties by the Polish and French mathematicians Zaremba in [42], [43] and Marchaud [31], [32].

Control theory motivated the renewal of the interest to the differential inclusion (1.1) in the earlier sixties. Filippov [14] and Ważewski [41] have shown that under very mild assumptions the control system

$$x' = f(t, x, u(t)), \quad u(t) \in U \text{ is measurable}, \quad x(t_0) = x_0 \tag{1.2}$$

can be reduced to differential inclusion (1.1). This placed control systems in the framework of ordinary differential "equations" with the difference that the right-hand side of these equations is multivalued.

However, very fortunately, the development of differential inclusions followed the same route that ODEs. There are existence results of *Peano* and *Cauchy-Lipschitz* type. When F is Lipschitz, then solutions depend on the initial condition in a Lipschitz way. We can as well differentiate solutions with respect to the initial condition (and to obtain *variational inclusions* instead of variational equations.) The only, but very important difference, is due to the fact that the solution to (1.1) is *a set* (of absolutely continuous functions $x(\cdot)$ starting at x_0 and satisfying $x'(t) \in F(t, x(t))$ almost everywhere.) For this reason the set-valued analysis arguments [5] have to be used in an essential way to investigate differential inclusions.

In Section 1 we recall Painlevé-Kuratowski limits, tangents to sets and generalized derivatives of functions and in Section 2 definitions concerning regularity and differentiation of set-valued maps that we shall use. We also gather some results on measurability and integration. The detailed study of these topics can be found for instance in [5].

Section 3 is devoted to differential inclusions. We start by the fundamental Filippov theorem and its applications. This is more than an existence theorem *à la Cauchy-Lipschitz*, but implies the same kind of consequences than the Gronwall inequality. In particular, we can compare solutions under perturbations of dynamics and/or initial conditions, and, in this respect, this theorem is particularly useful. We also discuss there a result due to Filippov and Ważewski which states that solutions to (1.1) are dense in solutions to the relaxed differential inclusion

$$x'(t) \in \overline{co}\, F(t, x(t)), \quad x(t_0) = x_0$$

This allows to extend the concept of infinitesimal generator to set-valued semigroups (reachable maps) and also to derive variational inclusions by differentiating solutions with respect to initial conditions.

Finally we state the very useful viability theorem for problems under state constraints. See [4] for many results of this theory.

A natural question do arise:

Can differential inclusion (1.1) *be reduced to control system* (1.2)?

This is not true in general and examples of "nonconvex" differential inclusions justify their study in the nonparametrized form. However, the answer is positive when F has convex images.

We state in Section 4 some theorems concerning parametrization of set-valued maps. Most of the results of this section are provided without proofs.

1.1 Preliminaries

1.1.1 Limits of Sets

Let X be a metric space supplied with a distance d. When K is a subset of X, we denote by

$$d_K(x) \;:=\; d(x,K) \;:=\; \inf_{y\in K} d(x,y)$$

the *distance from x to K*, where we set $d(x,\emptyset) := +\infty$. Limits of sets have been introduced by Painlevé in 1902, as it is reported by his student Zoretti. They have been popularized by Kuratowski in his famous book TOPOLOGIE and thus, often called *Kuratowski lower and upper limits* of sequences of sets.

Definition 1.1.1 *Let $(K_n)_{n\in\mathbf{N}}$ be a sequence of subsets of a metric space X. We say that the subset*

$$\mathrm{Limsup}_{n\to\infty} K_n \;:=\; \left\{x \in X \;\mid\; \liminf_{n\to\infty} d(x,K_n) = 0\right\}$$

is the upper limit *of the sequence K_n and that the subset*

$$\mathrm{Liminf}_{n\to\infty} K_n \;:=\; \{x \in X \;\mid\; \lim_{n\to\infty} d(x,K_n) = 0\}$$

is its lower limit. *A subset K is said to be the* limit *or the* set limit *of the sequence K_n if*

$$K \;=\; \mathrm{Liminf}_{n\to\infty} K_n \;=\; \mathrm{Limsup}_{n\to\infty} K_n \;=:\; \mathrm{Lim}_{n\to\infty} K_n$$

Lower and upper limits are obviously *closed.* We also see at once that

$$\mathrm{Liminf}_{n\to\infty} K_n \;\subset\; \mathrm{Limsup}_{n\to\infty} K_n$$

and that the upper limits and lower limits of the subsets K_n and of their closures $\overline{K_n}$ do coincide, since $d(x,K_n) = d(x,\overline{K_n})$.

Naturally, we can replace $\mathbf{N}$ by a metric (or even, topological) space X, and sequences of subsets $n \rightsquigarrow K_n$ by set-valued maps $x \rightsquigarrow F(x)$ (which associates with a point x a subset $F(x)$) and adapt the definition of upper and lower limits to this case, called the *continuous case*.

1.1.2 Tangent and Normal Cones to a Subset

We begin with a presentation of the contingent cones:

Definition 1.1.2 (Contingent Cones) *Let $K \subset X$ be a subset of a normed vector space X and $x \in \overline{K}$ belong to the closure of K. The* contingent[1] cone $T_K(x)$ *is defined by*

$$T_K(x) := \{v \mid \liminf_{h \to 0+} d_K(x+hv)/h = 0\} = \mathrm{Limsup}_{h \to 0+} \frac{K-x}{h}$$

It follows from the definition that $T_K(x)$ *is a closed cone.*

It is very convenient to have the following characterization of this cone in terms of sequences:

$$\begin{cases} v \in T_K(x) \text{ if and only if } \exists\, h_n \to 0+ \text{ and } \exists\, v_n \to v \\ \text{such that } \forall\, n, \ x + h_n v_n \in K \end{cases}$$

It implies that when K is convex, $T_K(x) = \overline{\bigcup_{\lambda \geq 0} \lambda(K-x)}$. We also observe that

$$\text{if } x \in \mathrm{Int}(K), \text{ then } T_K(x) = X$$

This situation may also happen when x does not belong to the interior of K (see Figure 1.1.)

We shall need the following very useful theorem.

Theorem 1.1.3 *Let X be a finite dimensional vector-space and K be a closed subset of X. Then for every $x \in K$*

$$\mathrm{Liminf}_{y \to_K x} T_K(y) = \mathrm{Liminf}_{y \to_K x} \overline{co}(T_K(y)) \subset T_K(x)$$

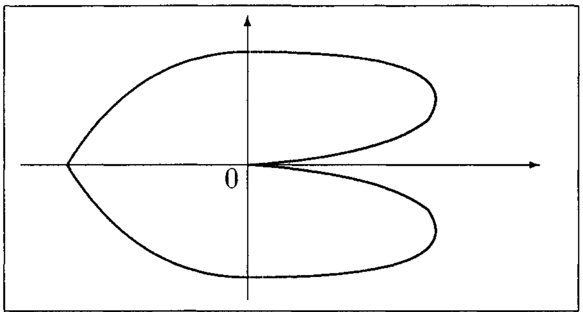
See for instance [5] for the proof.

Definition 1.1.4 (Subnormal Cones) *Let $K \subset X$ be a subset of a normed vector space X and $x \in \overline{K}$ belong to the closure of K. The* subnormal cone $N_K^0(x)$ *is defined by*

$$N_K^0(x) := \{p \in X^\star \mid \ < p, v > \leq 0 \ \ \forall\, v \in T_K(x)\}$$

[1] from the Latin *contingere*, to touch on all sides, introduced by G. Bouligand in the 30's. This term was already used by R. Descartes, in a 1638 letter to Mersenne criticizing P. de Fermat's method on tangents.

Figure 1.1: Contingent Cone at a Boundary Point may be the Whole Space



Subset K such that $T_K(0) = X$

1.1.3 Generalized Differentials of Nonsmooth Functions

Definition 1.1.5 *Let X be a normed vector space, $\varphi : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be an extended function and $x_0 \in X$ be such that $\varphi(x_0) \neq \pm\infty$.*

The superdifferential *of φ at x_0 is the closed convex set defined by:*

$$\partial_+\varphi(x_0) = \left\{ p \in \mathbf{R}^n \mid \limsup_{x \to x_0} \frac{\varphi(x) - \varphi(x_0) - < p, x - x_0 >}{\|x - x_0\|} \leq 0 \right\}$$

where $< \cdot, \cdot >$ denotes the scalar product.

The subdifferential *is defined in a similar way:*

$$\partial_-\varphi(x_0) = \left\{ p \in \mathbf{R}^n \mid \liminf_{x \to x_0} \frac{\varphi(x) - \varphi(x_0) - < p, x - x_0 >}{\|x - x_0\|} \geq 0 \right\}$$

We always have $\partial_+\varphi(x_0) = -\partial_-(-\varphi)(x_0)$.

The super and subdifferentials may also be characterized using *contingent epiderivatives*:

Definition 1.1.6 *Let X be a normed vector space, $\varphi : X \mapsto \mathbf{R} \cup \{\pm\infty\}$ be an extended function, $v \in X$ and $x_0 \in X$ be such that $\varphi(x_0) \neq \pm\infty$.*

The contingent epiderivative *of φ at x_0 in the direction v is given by*

$$D_\uparrow\varphi(x_0)(v) = \liminf_{h\to 0+,\, v'\to v} \frac{\varphi(x_0+hv')-\varphi(x_0)}{h}$$

and the contingent hypoderivative *of φ at x_0 in the direction v by*

$$D_\downarrow\varphi(x_0)(v) = \limsup_{h\to 0+,\, v'\to v} \frac{\varphi(x_0+hv')-\varphi(x_0)}{h}$$

Clearly

$$D_\uparrow\varphi(x_0) = -D_\downarrow(-\varphi)(x_0)$$

By a direct verification $D_\uparrow\varphi(x_0)$ is a lower semicontinuous map taking its values in $\mathbf{R}\cup\{\pm\infty\}$ whose epigraph is equal to the contingent cone to the epigraph of φ at $(x_0,\varphi(x_0))$.

When $\varphi:\mathbf{R}^n\mapsto\mathbf{R}$ is Lipschitz at x_0, then the contingent epi and hypoderivatives are reduced to the *Dini lower and upper derivatives*:

$$D_\uparrow\varphi(x_0)(v) = \liminf_{h\to 0+} \frac{\varphi(x_0+hv)-\varphi(x_0)}{h}$$

and

$$D_\downarrow\varphi(x_0)(v) = \limsup_{h\to 0+} \frac{\varphi(x_0+hv)-\varphi(x_0)}{h}$$

Proposition 1.1.7 [5] *Let $\varphi:\mathbf{R}^n\mapsto\mathbf{R}\cup\{\pm\infty\}$ be an extended function. Then*

$$\partial_-\varphi(x_0) = \{\, p\in\mathbf{R}^n \mid \forall\, v\in\mathbf{R}^n,\ D_\uparrow\varphi(x_0)(v) \ge \langle p,v\rangle \,\}$$

and

$$\partial_+\varphi(x_0) = \{\, p\in\mathbf{R}^n \mid \forall\, v\in\mathbf{R}^n,\ D_\downarrow\varphi(x_0)(v) \le \langle p,v\rangle \,\}$$

It is not difficult to show that φ is Fréchet differentiable at x_0 if and only if both super and subdifferentials of φ at x_0 are nonempty. Moreover in this case

$$\partial_+\varphi(x_0) = \partial_-\varphi(x_0) = \{\, \nabla\varphi(x_0) \,\}$$

Definition 1.1.8 *Let $\varphi : \mathbf{R}^n \mapsto \mathbf{R}$ be Lipschitz at x_0. We denote by $\partial^\star\varphi(x_0)$ the set of all cluster points of gradients $\nabla\varphi(x_n)$, when x_n converge to x_0 and φ is differentiable at x_n, i.e.,*

$$\partial^\star\varphi(x_0) \;=\; \mathrm{Limsup}_{x\to x_0}\,\{\;\nabla\varphi(x)\;\}$$

Proposition 1.1.9 (Clarke) *If $\partial^\star\varphi(x_0)$ is a singleton, then φ is differentiable at x_0.*

See [11, p.33] for the proof.

1.1.4 Semiconcave Functions

Definition 1.1.10 *Consider a convex subset K of $\mathbf{R}^n$. A function $\varphi : K \mapsto \mathbf{R}$ is called* semiconcave *if there exists $\omega : \mathbf{R}_+ \times \mathbf{R}_+ \mapsto \mathbf{R}_+$ such that*

$$\forall\, r \leq R,\ \forall\, s \leq S,\ \ \omega(r,s) \leq \omega(R,S)\ \ \&\ \ \lim_{s\to 0+}\omega(R,s) \;=\; 0 \qquad (1.3)$$

and for every $R > 0$, $\lambda \in [0,1]$ and all $x, y \in K \cap RB$

$$\lambda\varphi(x) + (1-\lambda)\varphi(y) \leq \varphi(\lambda x + (1-\lambda)y) + \lambda(1-\lambda)\|x-y\|\ \omega(R,\|x-y\|)$$

We say that φ is semiconcave at x_0 if there exists a neighborhood of x_0 in K such that the restriction of φ to it is semiconcave. We call the above function ω a modulus of semiconcavity of φ.

Observe that every concave[2] function $\varphi : K \mapsto \mathbf{R}$ is semiconcave (with ω equal to zero.)

Exercises

1. Let K be a convex subset of $\mathbf{R}^n$ and $\varphi : \mathbf{R}^n \mapsto \mathbf{R}$ be continuously differentiable on a neighborhood of K. Show that the restriction of φ to K is semi-concave.
2. Consider a subset K of $\mathbf{R}^n$ and define the function $\varphi : \mathbf{R}^n \mapsto \mathbf{R}_+$ by $\varphi(x) = \mathrm{dist}(x,K)^2$. Show that φ is semiconcave.

[2] Recall that a function $\varphi : K \mapsto \mathbf{R}$, where K is a convex subset of a vector space, is called *concave* if for all $x, y \in K$ and $\lambda \in [0,1]$, $\varphi(\lambda x + (1-\lambda)y) \geq \lambda\varphi(x) + (1-\lambda)\varphi(y)$.

3. Show that a continuous semiconcave function $\varphi : \mathbf{R}^n \mapsto \mathbf{R}$ is locally Lipschitz. □

In general a Lipschitz function does not have directional derivatives. Our next result implies in particular that for a semi-concave function, the directional derivatives exist.

Theorem 1.1.11 *Let $K \subset \mathbf{R}^n$ be a convex set, $x_0 \in K$ and let a function $\varphi : K \mapsto \mathbf{R}$ be Lipschitz and semiconcave at x_0. Then for every $v \in T_K(x_0)$*

$$\liminf_{\substack{v' \to v, h \to 0+ \\ x' \to_K x_0,\ x' + hv' \in K}} \frac{\varphi(x' + hv') - \varphi(x')}{h}$$

$$= \lim_{\substack{v' \to v, h \to 0+ \\ x_0 + hv' \in K}} \frac{\varphi(x_0 + hv') - \varphi(x_0)}{h}$$

In particular, if $x_0 \in \mathrm{Int}(K)$, then

$$\partial_+\varphi(x_0) \;=\; co\,(\partial^\star\varphi(x_0)) \tag{1.4}$$

(Clarke's generalized gradient of φ at x_0), where co states for the convex hull. Furthermore, setting $\varphi = -\infty$ outside of K, for all $x_0 \in K$

$$\mathrm{Limsup}_{x \to_{\mathrm{Int}(K)} x_0}\ \partial_+\varphi(x) \;\subset\; \partial_+\varphi(x_0)$$

Proof — It is enough to consider the case $\|v\| < 1$. Fix such v and let $\delta > 0$ be so that φ is semiconcave on $K \cap B_{2\delta}(x_0)$ with semiconcavity modulus $\omega(\cdot) := \omega(2\delta, \cdot)$. Let $x \in K \cap B_\delta(x_0)$. Then for all $0 < h_1 \le h_2 \le \delta$ such that $x + h_2 v \in K$ we have

$$\varphi(x + h_1 v) - \varphi(x) \;=\; \varphi\left(\tfrac{h_1}{h_2}(x + h_2 v) \;+\; \left(1 - \tfrac{h_1}{h_2}\right) x\right) \;-\; \varphi(x)$$

$$\ge\; \tfrac{h_1}{h_2}\varphi(x + h_2 v) \;-\; \tfrac{h_1}{h_2}\varphi(x) -\; h_1\left(1 - \tfrac{h_1}{h_2}\right)\|v\|\,\omega(h_2\,\|v\|)$$

Consequently,

$$\frac{\varphi(x + h_1 v) - \varphi(x)}{h_1} \ge \frac{\varphi(x + h_2 v) - \varphi(x)}{h_2} - \left(1 - \frac{h_1}{h_2}\right)\omega(h_2\,\|v\|)$$

and we proved that for every $x \in K \cap B_\delta(x_0)$ and all $0 < h' \leq h \leq \delta$,

$$\frac{\varphi(x + h'v) - \varphi(x)}{h'} \geq \frac{\varphi(x + hv) - \varphi(x)}{h} - \omega(h\,\|v\|) \tag{1.5}$$

Thus for every $0 < h \leq \delta$

$$\liminf_{\substack{h' \to 0+ \\ v' \to v \\ x + h'v' \in K}} \frac{\varphi(x + h'v') - \varphi(x)}{h'} \geq \frac{\varphi(x + hv) - \varphi(x)}{h} - \omega(h\,\|v\|)$$

Taking lim sup in the right-hand side of the above inequality when $x = x_0$, we deduce that

$$\lim_{\substack{h \to 0+, v' \to v \\ x_0 + hv' \in K}} \frac{\varphi(x_0 + hv') - \varphi(x_0)}{h}$$

does exist. Fix $\varepsilon > 0$ and $0 < \lambda < \delta$. From the Lipschitz continuity of φ it follows that there exists $0 < \alpha < \delta$ such that for all $x \in K \cap B_\alpha(x_0)$ and $v' \in B_\alpha(v)$

$$\frac{\varphi(x_0 + \lambda v) - \varphi(x_0)}{\lambda} \leq \frac{\varphi(x + \lambda v') - \varphi(x)}{\lambda} + \varepsilon$$

where $x_0 + \lambda v \in K$, $x + \lambda v' \in K$. Thus, using (1.5), we obtain that for all sufficiently small $\alpha > 0$,

$$\frac{\varphi(x_0 + \lambda v) - \varphi(x_0)}{\lambda}$$

$$\leq \inf_{\substack{x \in K \cap B_\alpha(x_0) \\ h \in]0, \lambda], v' \in B_\alpha(v) \\ x + hv' \in K}} \frac{\varphi(x + hv') - \varphi(x)}{h} + \omega(\lambda\|v'\|) + \varepsilon$$

Letting ε, α and λ converge to zero we end the proof of the first statement. The second one results from the alternative definition of

Clarke's generalized gradient, i.e. $p \in co\ (\partial^\star \varphi(x_0))$ if and only if for all v

$$\liminf_{\substack{v' \to v, h \to 0+ \\ x' \to x_0}} \frac{\varphi(x'+hv') - \varphi(x')}{h} \le < p, v >$$

To prove the last statement we set $\varphi = -\infty$ outside of K. Consider a sequence $x_m \in \mathrm{Int}(K)$ converging to x_0 and a sequence $p_m \in \partial_+\varphi(x_m)$ converging to some p. We have to show that $p \in \partial_+\varphi(x_0)$.

From (1.4) and the Carathéodory theorem, we deduce that there exist $\lambda_i^m \ge 0$ and $x_i^m \in \mathrm{Int}(K)$ converging to x_0 when $m \to \infty$ such that φ is differentiable at x_i^m and for all i the sequence $\nabla\varphi(x_i^m)$ converges to some p_i when $m \to \infty$, and for every m, $\sum_{i=0}^n \lambda_i^m = 1$,

$$\lim_{m\to\infty} \left(\sum_{i=0}^n \lambda_i^m \nabla\varphi(x_i^m) \right) \;=\; p$$

Taking a subsequence and keeping the same notations, we may assume that $(\lambda_0^m, ..., \lambda_n^m)$ converge to some $(\lambda_0, ..., \lambda_n)$. Thus $p = \sum_{i=0}^n \lambda_i p_i$. Since $\partial_+\varphi(x_0)$ is convex, the above yields that it is enough to prove our statement only in the case when φ is differentiable at x_m. Fix $v \in T_K(x_0)$ and consider $h_m \to 0+$ such that $x_m + h_m v \in K$ and

$$\frac{\varphi(x_m + h_m v) - \varphi(x_m)}{h_m} \;\le\; \langle \nabla\varphi(x_m),\; v\rangle \;+\; \frac{1}{m}$$

This and the first claim imply that

$$\limsup_{v'\to v,\; h\to 0+} \frac{\varphi(x_0 + hv') - \varphi(x_0)}{h} \;\le\; \langle p, v\rangle$$

Hence from Proposition 1.1.7 we deduce that $p \in \partial_+\varphi(x_0)$. □

Proposition 1.1.12 *Let $\varphi : \mathbf{R}^n \mapsto \mathbf{R}$ be Lipschitz and semiconcave at x_0. If $\partial_+\varphi(x_0)$ is a singleton, then φ is differentiable at x_0 and*

$$\partial^\star\varphi(x_0) \;=\; \{\; \nabla\varphi(x_0)\; \}$$

In particular, if $\partial_+\varphi(x)$ is a singleton for all x near x_0, then φ is continuously differentiable at x_0.

Proposition 1.1.13 *Let $\varphi : \mathbf{R}^n \mapsto \mathbf{R}$, $x_0 \in \mathbf{R}^n$. If φ is Lipschitz at x_0 and both φ and $-\varphi$ are semiconcave at x_0, then φ is continuously differentiable on a neighborhood of x_0.*

Proof — Since φ and $-\varphi$ are semiconcave at x_0, by Theorem 1.1.11, there exists a neighborhood $\mathcal{N}$ of x_0 such that for all $x \in \mathcal{N}$

$$\partial_+\varphi(x) = co(\partial^\star\varphi(x)), \quad \partial_-\varphi(x) = -\partial_+(-\varphi)(x) = -co(\partial^\star(-\varphi)(x))$$

Hence both $\partial_+\varphi(x)$ and $\partial_-\varphi(x)$ are nonempty. Therefore φ is differentiable on $\mathcal{N}$. The conclusion follows from Proposition 1.1.12. □

We investigate next closedness of the level sets of regularized lower derivatives.

Proposition 1.1.14 *Let $K \subset \mathbf{R}^n$ and $\varphi : K \mapsto \mathbf{R}$ be locally Lipschitz. Define the set-valued map $Q : K \rightsquigarrow \mathbf{R}^n$ by:*
for all $x \in K$, $Q(x)$ is equal to

$$\{v \mid \liminf_{\substack{v' \to v, h \to 0+ \\ x' \to_K x, x'+hv' \in K}} \frac{\varphi(x'+hv') - \varphi(x')}{h} \le 0\}$$

Then Q has closed nonempty images and Graph(Q) *is closed.*

Proof — Clearly for every x, $0 \in Q(x)$. It remains to show that for every sequence $(x_n, v_n) \in K \times \mathbf{R}^n$ converging to some $(x, v) \in K \times \mathbf{R}^n$ and satisfying $v_n \in Q(x_n)$, we have $v \in Q(x)$. Fix such a sequence and let $\varepsilon_n \to 0+$. Then there exist $h_n \to 0+$, $x'_n \to_K x$, $v'_n \to v$ such that for every n, $x'_n + h_n v'_n \in K$ and

$$\frac{\varphi(x'_n + h_n v'_n) - \varphi(x'_n)}{h_n} \le \varepsilon_n$$

Taking lim inf in the above inequality we end the proof. □

1.1.5 Subnormal Cones to the Epigraph

Recall that

$$\mathcal{E}p(D_\uparrow\varphi(x_0)) = T_{\mathcal{E}p(\varphi)}(x_0, \varphi(x_0)) \tag{1.6}$$

where $\mathcal{E}p$ denotes the epigraph.

The subnormal cone to $\mathcal{E}p(\varphi)$ at $(x_0, \varphi(x_0))$ is given by

$$N^0_{\mathcal{E}p(\varphi)}(x_0, \varphi(x_0)) := \left\{p \in \mathbf{R}^n \mid \forall\, v \in T_{\mathcal{E}p(\varphi)}(x_0, \varphi(x_0)),\ \langle p, v\rangle \le 0\right\}$$

Thus

Proposition 1.1.15 *Let $\varphi : \mathbf{R}^n \mapsto \mathbf{R} \cup \{\pm\infty\}$ and $x_0 \in \mathrm{Dom}(\varphi)$. Then the following statements are equivalent*

$$\begin{array}{ll} i) & p \in \partial_-\varphi(x_0) \\ ii) & \forall\, u \in \mathbf{R}^n,\ < p, u > \le D_\uparrow\varphi(x_0)(u) \\ iii) & (p, -1) \in N^0_{\mathcal{E}p(\varphi)}(x_0, \varphi(x_0)) \end{array}$$

We shall also need the following technical result.

Lemma 1.1.16 ([37]) *Consider an extended lower semicontinuous function $\varphi : \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ and $x_0 \in \mathrm{Dom}(\varphi)$. Let $p \in \mathbf{R}^n$ be such that*

$$(p, 0) \in N^0_{\mathcal{E}p(\varphi)}(x_0, \varphi(x_0)), \quad p \ne 0$$

Then for every $\varepsilon > 0$, there exist x_ε, p_ε in $\mathbf{R}^n$ and $q_\varepsilon < 0$ satisfying

$$\|x_\varepsilon - x_0\| \le \varepsilon,\ \ \|p_\varepsilon - p\| \le \varepsilon \quad \& \quad (p_\varepsilon, q_\varepsilon) \in N^0_{\mathcal{E}p(\varphi)}(x_\varepsilon, \varphi(x_\varepsilon))$$

1.2 Regularity of Set-Valued Maps

We recall next some definitions concerning set-valued maps. Let X, Y denote metric spaces and $F : X \rightsquigarrow Y$ be a set-valued map. For every $x \in X$ the subset $F(x)$ is called the *image* of F at x. The *domain* of F is the subset

$$\mathrm{Dom}(F) := \{x \in X \mid F(x) \ne \emptyset\}$$

and its *graph*

$$\mathrm{Graph}(F) := \{(x, y) \in X \times Y \mid y \in F(x)\}$$

Definition 1.2.1 *The map F is called* upper semicontinuous *at x if and only if for any neighborhood $\mathcal{U}$ of $F(x)$,*

$$\exists\, \eta > 0 \quad \textit{such that } \forall\, x' \in B_\eta(x), \;\; F(x') \subset \mathcal{U}$$

It is said to be upper semicontinuous *on a subset $K \subset X$ if and only if it is upper semicontinuous at any point $x \in K$.*

The map F is called lower semicontinuous *at x if and only if for any open subset $\mathcal{U} \subset Y$ such that $\mathcal{U} \cap F(x) \neq \emptyset$,*

$$\exists\, \eta > 0 \text{ such that } \forall\, x' \in B_\eta(x), \;\; F(x') \cap \mathcal{U} \neq \emptyset$$

It is said to be lower semicontinuous *on a subset $K \subset X$ if for every $x \in K$ and for any open subset $\mathcal{U} \subset Y$ with $\mathcal{U} \cap F(x) \neq \emptyset$,*

$$\exists\, \eta > 0 \text{ such that } \forall\, x' \in B_\eta(x) \cap K, \;\; F(x') \cap \mathcal{U} \neq \emptyset$$

We shall say that F is continuous *at x if it is* both *upper and lower semicontinuous at x, and that it is* continuous *on a subset $K \subset X$ if and only if it is upper and lower semicontinuous on K.*

Notice that if F is upper semicontinuous on X, then its domain is closed.

When $F(x)$ is compact, F is upper semicontinuous at x if and only if

$$\forall\, \varepsilon > 0, \; \exists\, \eta > 0 \text{ such that } \; \forall\, x' \in B_\eta(x), \;\; F(x') \subset \bigcup_{y \in F(x)} B_\varepsilon(y)$$

Proposition 1.2.2 [5] *The graph of an upper semicontinuous set-valued map $F : X \rightsquigarrow Y$ with closed images is closed. The converse is true if we assume that Y* is compact.

Definition 1.2.3 *When (X, d_X) is a metric space and Y is a normed space, we shall say that $F : X \rightsquigarrow Y$ is* Lipschitz *($L-$Lipschitz) on a subset $K \subset \mathrm{Dom}(F)$ if there exists $L \geq 0$ such that*

$$\forall\, x_1,\; x_2 \in K, \;\; F(x_1) \;\subset\; F(x_2) + L d_X(x_1, x_2) B$$

The set-valued map F is called locally Lipschitz around $x \in X$ if there exists a neighborhood $\mathcal{N}$ of x such that F is Lipschitz on $\mathcal{N}$.

We recall next definitions of derivatives of set-valued maps.

Definition 1.2.4 *Let X, Y be normed spaces, $F : X \rightsquigarrow Y$ be a set-valued map and $y \in F(x)$.*

The adjacent derivative *$dF(x,y)$ is the set-valued map from X to Y defined by*

$$\forall\, u \in X, \;\; v \in dF(x,y)(u) \iff \forall\, h_n \to 0+ \quad \exists\, u_n \to u$$

$$\text{such that } \lim_{n\to\infty} \text{dist}\left(v,\; \frac{F(x+h_n u_n) - y}{h_n}\right) = 0$$

If F is Lipschitz around x, then an equivalent definition is given by

$$\forall\, u \in X, \;\; dF(x,y)(u) = \text{Liminf}_{h\to 0+} \tfrac{F(x+hu)-y}{h} =$$

$$\lim_{h\to 0+} \text{dist}\left(v,\; \frac{F(x+hu) - y}{h}\right) = 0$$

We shall need the following proposition.

Proposition 1.2.5 [5] *Let us assume that the images of F are convex and that F is Lipschitz around x. Then for any $(x,y) \in$* Graph*(F) the images of the adjacent derivative $dF(x,y)$ are convex and*

$$dF(x,y)(0) = T_{F(x)}(y)$$

$$\forall u \in \text{Dom}(dF(x,y)), D^\flat F(x,y)(u) + dF(x,y)(0) = dF(x,y)(u)$$

Proof — Let v_1 and v_2 belong to $dF(x,y)(u)$. Then, for any sequence $h_n > 0$ converging to 0, there exist sequences u_{1n} and u_{2n} converging to u and sequences v_{1n} and v_{2n} converging to v_1 and v_2 respectively such that

$$\forall\, n, \;\; y + h_n v_{in} \in F(x + h_n u_{in}) \;\; (i = 1,2)$$

Since F is Lipschitz around x, there exists $l > 0$ such that for all n large enough,

$$y + h_n v_{2n} \in F(x + h_n u_{1n}) + l h_n \|u_{2n} - u_{1n}\|$$

so that we can find another sequence v_{3n} converging to v_2 such that

$$y + h_n v_{3n} \subset F(x + h_n u_{1n})$$

Now, $F(x + h_n u_{1n})$ being convex, we deduce that for all $\lambda \in [0, 1]$,

$$y + h_n(\lambda v_{1n} + (1 - \lambda)v_{3n}) \in F(x + h_n u_{1n})$$

Since $\lambda v_{1n} + (1 - \lambda)v_{3n}$ converges to $\lambda v_1 + (1 - \lambda)v_2$, this element belongs to $dF(x, y)(u)$.

Notice that $v \in dF(x, y)(0)$ if and only if $d(v, (F(x) - y)/h)$ converges to 0. Since $F(x)$ is convex, it coincides with the tangent cone.

Since $0 \in dF(x, y)(0)$ we obtain that

$$\forall\, u,\ dF(x, y)(u) \subset dF(x, y)(u) + dF(x, y)(0)$$

To prove the opposite inclusion fix

$$v \in dF(x, y)(u) \quad \& \quad w \in dF(x, y)(0)$$

Let $h_n \to 0+$, $v_n \to v$ be such that

$$\forall\, n,\ y + h_n v_n \in F(x + h_n u)$$

By convexity of $F(x)$, there exist $w_n \to w$ such that for n large enough, $y + \sqrt{h_n} w_n \in F(x)$. Then, by the Lipschitz continuity of F, for all large n and for some w'_n, we have

$$y + \sqrt{h_n} w'_n \in F(x + h_n u)\ ;\ \|w'_n - w_n\| \leq l\sqrt{h_n}\,\|u\|$$

Thus

$$\begin{cases} (1 - \sqrt{h_n})(y + h_n v_n) + \sqrt{h_n}(y + \sqrt{h_n} w'_n) \\ = y + h_n(v_n + w'_n) - \sqrt{h_n}\, h_n v_n = y + h_n(v + w) + h_n \varepsilon(h_n) \\ \in F(x + h_n u) \end{cases}$$

where $\varepsilon(h_n)$ converges to 0. Hence

$$\lim_{n \to \infty} \operatorname{dist}\left(v + w, \frac{F(x + h_n u) - y}{h_n}\right) = 0$$

This ends the proof. □

Let X be a complete separable metric space, $t_0 < T$ be real numbers and $U : [t_0, T] \rightsquigarrow X$ be a set-valued map with closed, possibly empty images. It is called (Lebesgue) *measurable* if for every open subset $\mathcal{O} \subset X$, the set

$$\{ t \in [t_0, T] \mid U(t) \cap \mathcal{O} \neq \emptyset \} \text{ is Lebesgue measurable}$$

or, equivalently, if for every closed subset $\mathcal{C} \subset X$, the set

$$\{ t \in [t_0, T] \mid U(t) \cap \mathcal{C} \neq \emptyset \} \text{ is Lebesgue measurable}$$

A measurable single-valued map $u : [t_0, T] \mapsto X$ satisfying

$$\forall\, t \in [t_0, T], \;\; u(t) \in U(t)$$

is called a *measurable selection* of $U(\cdot)$.

Measurable selections are dense:

Theorem 1.2.6 [5] *Let X be a complete separable metric space and $U : [t_0, T] \rightsquigarrow X$ be a set-valued map with closed nonempty images. Then the following two statements are equivalent:*

i) — U is measurable

ii) — There exist measurable selections $u_n(\cdot)$ of $U(\cdot)$, $n = 1, \ldots$ such that for every $t \in [t_0, T]$, $U(t) = \overline{\bigcup_{n \geq 1} u_n(t)}$.

Proposition 1.2.7 [5] *Let X be a complete separable metric space and $U_n : [t_0, T] \rightsquigarrow X$, $n = 1, \ldots$ be measurable set-valued maps with closed images. Then the set-valued maps*

$$t \rightsquigarrow \bigcap_{n \geq 1} U_n(t), \quad t \rightsquigarrow \overline{\bigcup_{n \geq 1} U_n(t)}$$

and

$$t \rightsquigarrow \operatorname{Liminf}_{n \to \infty} U_n(t), \quad t \rightsquigarrow \operatorname{Limsup}_{n \to \infty} U_n(t)$$

are measurable.

Corollary 1.2.8 *Let X, Y be complete separable metric spaces, $x : [t_0, T] \mapsto X$ be a measurable single-valued map and $F : [t_0, T] \times X \rightsquigarrow Y$ be a set-valued map with nonempty closed images satisfying the following assumptions:*

$$\begin{cases} i) & \forall\, x \in X \ \text{ the set-valued map } \ F(\cdot, x) \ \text{ is measurable} \\ ii) & \text{For almost every } \ t \in [t_0, T], \ F(t, \cdot) \ \text{ is continuous at } \ x(t) \end{cases}$$

Then the map $t \rightsquigarrow F(t, x(t))$ is measurable.

Proof — Since $x(\cdot)$ is measurable, there exist measurable maps $x_n : [t_0, T] \mapsto X$ assuming only finite number of values such that for almost every $t \in [t_0, T]$, $\lim_{n\to\infty} x_n(t) = x(t)$. From the assumption $i)$ we deduce that the map $t \rightsquigarrow F(t, x_n(t))$ is measurable and from the assumption $ii)$, that for almost all $t \in [t_0, T]$

$$F(t, x(t)) \ = \ \text{Liminf}_{n\to\infty} \ F(t, x_n(t))$$

Proposition 1.2.7 completes the proof. □

Let us denote by $B(x, \rho)$ the closed ball in X of center x and radius ρ. When $K \subset X$ and $y \in X$ we denote by $\Pi_K(y)$ the projection of y on K given by

$$\Pi_K(y) \ := \ \{ \ x \ \in \ K \mid \ d_X(x, y) \ = \ \text{dist}\ (y, K) \ \}$$

Of course it may happen that the set $\Pi_K(y)$ is empty. Denote by $\overline{co}$ the closed convex hull.

Proposition 1.2.9 [5] *Let X be a separable Banach space, $U : [t_0, T] \rightsquigarrow X$ be a measurable set-valued map with closed nonempty images and $g : [t_0, T] \mapsto X$, $k : [t_0, T] \mapsto \mathbf{R}_+$ be measurable single-valued maps. Then the maps*

$$t \ \rightsquigarrow \ \overline{co}\, U(t), \ \ t \ \rightsquigarrow \ B(g(t), k(t)), \ \ t \ \rightsquigarrow \ \Pi_{U(t)}(g(t))$$

and $t \mapsto \text{dist}(g(t), U(t))$ are measurable. Consequently, if

$$\{v \in U(t) \mid \|v - g(t)\| \ \le \ k(t)\} \ \ne \ \emptyset \ \text{ almost everywhere in } \ [t_0, T]$$

then there exists a measurable selection $u(t) \in U(t)$ such that for almost all $t \in [t_0, T]$, $\|u(t) - g(t)\| \le k(t)$.

Consider a metric space Y. We recall that a map $\varphi : [t_0, T] \times X \mapsto Y$ is called *Carathéodory*, if for every $x \in X$, $\varphi(\cdot, x)$ is measurable and for almost all $t \in [t_0, T]$, the map $\varphi(t, \cdot)$ is continuous.

Proposition 1.2.10 [5] *Consider complete separable metric spaces X, Y, a Carathéodory map $\varphi : [t_0, T] \times X \mapsto Y$ and a measurable set-valued map $U : [t_0, T] \rightsquigarrow X$ with closed nonempty images. Then for every measurable map $h : [t_0, T] \mapsto Y$ satisfying*

$$h(t) \in \varphi(t, U(t)) \text{ almost everywhere in } [t_0, T]$$

there exists a measurable selection $u(t) \in U(t)$ such that $h(t) = \varphi(t, u(t))$ for almost all $t \in [t_0, T]$.

Definition 1.2.11 *Consider metric spaces X, Y and a set-valued map $G : [t_0, T] \times X \rightsquigarrow Y$ with closed images. It is called a* Carathéodory set-valued map *if for every $x \in X$, the map $t \rightsquigarrow G(t, x)$ is measurable and for every $t \in [t_0, T]$, the map $x \rightsquigarrow G(t, x)$ is continuous.*

Theorem 1.2.12 (Direct Image [5]) *Let X be a complete separable metric space and $U : [t_0, T] \rightsquigarrow X$ a measurable set-valued map with closed images.*

Consider a Carathéodory set-valued map G from $[t_0, T] \times X$ to a complete separable metric space Y. Then, the map

$$[t_0, T] \ni t \rightsquigarrow \overline{G(t, U(t))}$$

is measurable.

Denote by $L^1(t_0, T; \mathbf{R}^n)$ the Banach space of (Lebesgue) integrable maps $u : [t_0, T] \mapsto \mathbf{R}^n$ with the norm

$$\|u\|_{L^1} = \int_{t_0}^{T} \|u(t)\| \, dt$$

Definition 1.2.13 *Consider a set-valued map $U : [t_0, T] \rightsquigarrow \mathbf{R}^n$ and denote by $\mathcal{U}$ the set of integrable selections of U, i.e.,*

$$\mathcal{U} := \{ u \in L^1(t_0, T; \mathbf{R}^n) \mid u(t) \in U(t) \text{ almost everywhere in } [t_0, T] \}$$

The integral of U on $[t_0, T]$ is defined by

$$\int_{t_0}^{T} U dt := \left\{ \int_{t_0}^{T} u(t) dt \mid u \in \mathcal{U} \right\}$$

We say that a set-valued map $U : [t_0, T] \rightsquigarrow \mathbf{R}^n$ is *integrably bounded* if there exists an integrable function $\psi : [t_0, T] \mapsto \mathbf{R}_+$ such that $U(t) \subset \psi(t)B$ almost everywhere in $[t_0, T]$.

Let K be a nonempty subset of a vector space Y. A point $x \in K$ is called *extremal* if for all $y, z \in K$ and $0 < \lambda < 1$ satisfying $x = \lambda y + (1 - \lambda)z$, we have $x = y = z$.

Theorem 1.2.14 (Aumann) *Let $U : [t_0, T] \rightsquigarrow \mathbf{R}^n$ be a measurable set-valued map with nonempty closed images. Then the integral $\int_{t_0}^T U dt$ is convex and extremal points of $\overline{co}\left(\int_{t_0}^T U dt\right)$ are contained in $\int_{t_0}^T U dt$. If in addition U is integrably bounded, then the integral of U is also compact and $\int_{t_0}^T U ds = \int_{t_0}^T \overline{co} U ds$.*

See for instance [5] for the proof.

Theorem 1.2.15 *Let $U : [t_0, T] \rightsquigarrow \mathbf{R}^n$ be a measurable set-valued map with closed images having at least one integrable selection.*

Then for every $\varepsilon > 0$ and integrable selection $\overline{u}(t) \in \overline{co}\, U(t)$ there exists an integrable selection $u(t) \in U(t)$ such that

$$\sup_{t \in [t_0, T]} \left\| \int_{t_0}^t u(s) ds - \int_{t_0}^t \overline{u}(s) ds \right\| \leq \varepsilon$$

In particular this yields that

$$\overline{\int_{t_0}^T \overline{co}\, U dt} = \overline{\int_{t_0}^T U dt}$$

Proof — Fix $\varepsilon > 0$, an integrable selection $\overline{u}(t) \in \overline{co}\ U(t)$ and let $u_0(\cdot)$ be an integrable selection of $U(\cdot)$. Define measurable set-valued maps $U_n : [t_0, T] \rightsquigarrow X$ with closed nonempty images by

$$\forall\, t \in [t_0, T], \;\; U_n(t) = u_0(t) \cup (U(t) \cap nB)$$

and set $\varepsilon_n(t) := \text{dist}(\overline{u}(t), \overline{co}(U_n(t)))$. By Proposition 1.2.9, $\varepsilon_n(\cdot)$ is measurable for each n. Furthermore, the sequence $\{\varepsilon_n(\cdot)\}_{n \geq 1}$ is integrably bounded and $\lim_{n \to \infty} \varepsilon_n(t) = 0$ for $t \in [t_0, T]$. Using again Proposition 1.2.9, we deduce that for every $n \geq 1$ there exists a measurable selection $u_n(t) \in \overline{co}\, U_n(t)$ such that

$$\|u_n(t) - \overline{u}(t)\| \leq \varepsilon_n(t) \text{ almost everywhere in } [t_0, T]$$

Therefore, by the Lebesgue dominated convergence theorem, the sequence u_n converges to $\overline{u}$ in $L^1(t_0, T; \mathbf{R}^n)$ and for all n large enough

$$\sup_{t\in[t_0,T]} \left\| \int_{t_0}^{t} u_n(s)ds - \int_{t_0}^{t} \overline{u}(s)ds \right\| \leq \int_{t_0}^{T} \|u_n(s) - \overline{u}(s)\| \, ds \leq \frac{\varepsilon}{2}$$

It remains to show that for every $n \geq 1$ there exists an integrable selection $u(t) \in U_n(t) \subset U(t)$ such that

$$\sup_{t\in[t_0,T]} \left\| \int_{t_0}^{t} u(s)ds - \int_{t_0}^{t} u_n(s)ds \right\| \leq \frac{\varepsilon}{2}$$

Fix $n \geq 1$ and let $\psi : [t_0, T] \mapsto \mathbf{R}_+$ be an integrable function such that $U_n(t) \subset \psi(t)B$ for $t \in [t_0, T]$. Let $i \geq 1$ be so large, that for any measurable subset $I \subset [t_0, T]$ of the Lebesgue measure less than $(T - t_0)/i$ we have $\int_I \psi(s)ds \leq \varepsilon/4$. We denote by I_j the interval

$$I_j = \left[t_0 + \frac{j-1}{i}(T - t_0), \; t_0 + \frac{j}{i}(T - t_0) \right], \quad j = 1, ..., i$$

By Theorem 1.2.14,

$$\forall \, j = 1, ..., i, \quad \int_{I_j} \overline{co} \, U_n(s)ds = \int_{I_j} U_n(s)ds$$

This yields that for every $1 \leq j \leq i$ there exists a measurable selection $f_j(t) \in U_n(t)$ such that

$$\int_{I_j} f_j(s)ds = \int_{I_j} u_n(s)ds$$

Let u be a selection of U_n equal to f_j on the interior of I_j for every $j = 1, ..., i$. Then for every $t \in [t_0, T]$, there exists j such that $t \in I_j$ and

$$\left\| \int_{t_0}^{t} (u - u_n)(s)ds \right\| \leq \left\| \sum_{r=1}^{j-1} \int_{I_r} (u - u_n)(s)ds \right\| + \int_{I_j} \|u - u_n\| (s)ds$$

$$\leq \int_{I_j} \left(\|u(s)\| + \|u_n(s)\| \right) ds \leq 2 \int_{I_j} \psi(s)ds \leq \varepsilon/2 \quad \square$$

1.3 Differential Inclusions

Consider $t_0 < T$ and denote by $\mathcal{C}(t_0, T; \mathbf{R}^n)$ the Banach space of continuous maps from $[t_0, T]$ into $\mathbf{R}^n$ with the norm

$$\|x\|_{\mathcal{C}} = \sup_{t \in [t_0, T]} \|x(t)\|$$

We first define what we call a solution to differential inclusions.

In the case of differential equations, there is no ambiguity since the derivative $x'(\cdot)$ of a solution $x(\cdot)$ to a differential equation $x'(t) = f(t, x(t))$ inherits the properties of the map f and of the function $x(\cdot)$. It is continuous whenever f is continuous and measurable whenever f is continuous with respect to x and measurable with respect to t.

This is no longer the case with differential inclusions[3].

We recall that a function $x \in \mathcal{C}(t_0, T; \mathbf{R}^n)$ is called *absolutely continuous* if for for almost all $t \in [t_0, T]$ the derivative $x'(t)$ exists, $x' \in L^1(t_0, T; \mathbf{R}^n)$ and

$$\forall\, t \in [t_0, T], \;\; x(t) = x(t_0) + \int_{t_0}^{t} x'(s)ds$$

Let $W^{1,1}(t_0, T; \mathbf{R}^n)$ denote the Banach space of absolutely continuous

[3]The extension of Peano's Theorem to differential inclusions is due to Marchaud and Zaremba who proved independently in the thirties the existence of respectively *contingent* and *paratingent* solutions to differential inclusions (called *champs de demi-cônes* at the time. The generalization of the concept of derivative to the notion of contingent derivative is due to B. Bouligand, who wrote: "... Nous ferons tout d'abord observer ... que la notion de contingent éclaire celle de différentielle".) Then Ważewski proposed at the beginning of the sixties to look for solutions among *absolutely continuous* functions. He wrote: "... I learned the results of Zaremba's dissertation before the second world war, since I was a referee of that paper. Then a few years ago I came across with some results on optimal control and I have noticed a close connection between the optimal control problem and the theory of Marchaud-Zaremba." The author learned that this "coming across" happened during a seminar talk of C. Olech on a paper by LaSalle at Ważewski's seminar.

Ważewski proved that one can replace the contingent or paratingent derivatives of functions by derivatives of absolutely continuous functions defined almost everywhere in the definition of a solution to a differential inclusion, that he called *orientor field*.

functions from $[t_0, T]$ to $\mathbf{R}^n$ with the norm[4]

$$\|x\|_{W^{1,1}} = \|x(t_0)\| + \int_{t_0}^{T} \|x'(t)\| dt$$

Consider a set-valued map F from $[t_0, T] \times \mathbf{R}^n$ into subsets of $\mathbf{R}^n$. We associate with it the differential inclusion

$$x' \in F(t, x) \tag{1.7}$$

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a *solution* to (1.7) if

$$x'(t) \in F(t, x(t)) \text{ almost everywhere in } [t_0, T] \tag{1.8}$$

1.3.1 Filippov's Theorem

We investigate here some properties of solutions to differential inclusion (1.7) in the case when F is Lipschitz with respect to x.

We denote by $\mathcal{S}_{[t_0,T]}(x_0)$ the set of solutions to (1.7) starting at $x_0 \in \mathbf{R}^n$ and defined on the time interval $[t_0, T]$:

$$\mathcal{S}_{[t_0,T]}(x_0) = \{x \mid x \text{ is a solution to (1.7) on } [t_0, T],\ x(t_0) = x_0\}$$

and set $L^1(t_0, T) = L^1(t_0, T; \mathbf{R}_+)$ (the set of nonnegative integrable functions.)

Let $y \in W^{1,1}(t_0, T; \mathbf{R}^n)$ be an absolutely continuous function. Filippov's theorem provides an estimate of the distance from y to the set $\mathcal{S}_{[t_0,T]}(x_0) \subset W^{1,1}(t_0, T; \mathbf{R}^n)$ under the following assumptions on F:

$$\left\{ \begin{array}{ll} i) & \forall\, (t, x) \in [t_0, T] \times \mathbf{R}^n,\ \ F(t, x) \text{ is closed} \\ \\ ii) & \forall\, x \in \mathbf{R}^n \text{ the set-valued map } F(\cdot, x) \text{ is measurable} \\ \\ iii) & \exists\, \beta > 0,\ k \in L^1(t_0, T) \text{ such that for almost all} \\ & t \in [t_0, T],\ F(t, x) \text{ is nonempty for } x \in y(t) + \beta B \\ & \text{the map } F(t, \cdot) \text{ is } k(t) - \text{Lipschitz on } y(t) + \beta B \end{array} \right. \tag{1.9}$$

[4]It is called *Sobolev space*.

Theorem 1.3.1 *Consider a set-valued map* $F : [t_0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ *and an absolutely continuous function* $y \in W^{1,1}(t_0, T; \mathbf{R}^n)$. *Assume that (1.9) holds true and that the function*

$$t \mapsto \gamma(t) := \operatorname{dist}(y'(t), F(t, y(t)))$$

is integrable. Let $\delta \geq 0$ *and set*

$$\eta(t) = e^{\int_{t_0}^t k(\tau)d\tau}\delta + \int_{t_0}^t \gamma(s)\, e^{\int_s^t k(\tau)d\tau} ds$$

If $\eta(T) \leq \beta$, *then for all* $x_0 \in \mathbf{R}^n$ *with* $\|x_0 - y(t_0)\| \leq \delta$, *there exists* $x \in \mathcal{S}_{[t_0,T]}(x_0)$ *such that*

$$\forall\, t \in [t_0,\, T],\ \ \|x(t) - y(t)\| \leq \eta(t)$$

and

$$\|x'(t) - y'(t)\| \leq k(t)\eta(t) + \gamma(t) \text{ a.e. in } [t_0, T]$$

Remark — From Corollary 1.2.8 and Proposition 1.2.9 follows that under assumptions (1.9) the function $t \mapsto \operatorname{dist}(y'(t), F(t, y(t)))$ is always measurable. □

The proof can be found in [14], [1]. The above result can be extended to the whole half line:

Theorem 1.3.2 *Consider a set-valued map* $F : \mathbf{R}_+ \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ *and an absolutely continuous function* $y \in W^{1,1}(0, \infty; \mathbf{R}^n)$. *Assume that (1.9) holds true with the time interval* $[t_0, T]$ *replaced by* $\mathbf{R}_+$ *and that the function* $t \mapsto \gamma(t) := \operatorname{dist}(y'(t), F(t, y(t)))$ *is integrable on* $[0, \infty[$. *Let* $\delta \geq 0$ *and set*

$$\eta(t) = e^{\int_0^t k(\tau)d\tau}\delta + \int_0^t \gamma(s)\, e^{\int_s^t k(\tau)d\tau} ds$$

If $\limsup_{t\to\infty} \eta(t) \leq \beta$, *then for all* $x_0 \in \mathbf{R}^n$ *with* $\|x_0 - y(0)\| \leq \delta$, *there exists* $x \in \mathcal{S}_{[0,\infty[}(x_0)$ *such that*

$$\forall\, t \geq 0,\ \ \|x(t) - y(t)\| \leq \eta(t)$$

and

$$\|x'(t) - y'(t)\| \leq k(t)\eta(t) + \gamma(t) \text{ a.e. in } [0, \infty[$$

Proof — Theorem 1.3.1 yields an estimate on the finite interval $[0,1]$. Hence there exists a solution $x(\cdot) \in \mathcal{S}_{[0,1]}(x_0)$ satisfying the required estimates on the interval $[0,1]$ and in particular

$$\|x(1) - y(1)\| \le e^{\int_0^1 k(\tau)d\tau}\delta + \int_0^1 \gamma(s)e^{\int_s^1 k(\tau)d\tau}ds$$

This and Theorem 1.3.1 imply that there exists a solution $z(\cdot) \in \mathcal{S}_{[1,2]}(x(1))$ satisfying the required estimates on $[1,2]$. Hence we can extend $x(\cdot)$ on the interval $[0,2]$ by concatenating it with $z(\cdot)$ and we reiterate this process. □

The above theorems yield the following corollaries.

Corollary 1.3.3 *Consider a set-valued map* $F : [t_0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ *and a point* $x_0 \in \mathbf{R}^n$. *We assume that* F *satisfies (1.9) with* $y \equiv x_0$ *and is lower semicontinuous at* (t_0, x_0). *Then for every* $u \in F(t_0, x_0)$ *there exist* $t_1 > t_0$ *and a solution* $x(\cdot) \in \mathcal{S}_{[t_0,t_1]}(x_0)$ *with* $x'(t_0) = u$.

Proof — Fix $u \in F(t_0, x_0)$. It is enough to consider the absolutely continuous function

$$\forall\, t \in [t_0, T],\;\; y(t) = x_0 + (t - t_0)u$$

Then for every $t \in [t_0, T]$ such that $(t - t_0)\|u\| \le \beta$ we have

$$\text{dist}\,(u, F(t, y(t))) \le \text{dist}\,(u, F(t, x_0)) + k(t)\|y(t) - x_0\|$$

$$= \text{dist}(u, F(t, x_0)) + k(t)(t - t_0)\|u\|$$

By Theorem 1.3.1 there exist $t_1 > t_0$ and a solution $x \in \mathcal{S}_{[t_0,t_1]}(x_0)$ such that

$$\|x(t) - y(t)\| \le \int_{t_0}^t (\text{dist}(u, F(s, x_0)) + k(s)(s - t_0)\|u\|)\, e^{\int_s^t k(\tau)d\tau}ds$$

$$\le e^{\int_{t_0}^t k(s)ds}\left(\int_{t_0}^t \text{dist}(u, F(s, x_0))ds + (t - t_0)\|u\| \int_{t_0}^t k(s)ds\right)$$

for all $t \in [t_0, t_1]$. Thus

$$\forall\, t \in [t_0, t_1],\;\; \|x(t) - x_0 - (t - t_0)u\| = o(t - t_0)$$

and the result follows. □

Corollary 1.3.4 *Let $y_0 \in \mathbf{R}^n$, $y \in \mathcal{S}_{[t_0,T]}(y_0)$ and assume that F, y satisfy (1.9). Then there exists $\delta > 0$ depending only on $k(\cdot)$ such that for all $x_0 \in B(y_0, \delta)$ we have*

$$\inf_{x \in \mathcal{S}_{[t_0,T]}(x_0)} \|x - y\|_C \;\leq\; e^{\int_{t_0}^T k(s)ds} \|x_0 - y_0\|$$

1.3.2 Relaxation Theorems

Let $x_0 \in \mathbf{R}^n$. In this section we compare solutions to the differential inclusion

$$\begin{cases} x'(t) & \in & F(t, x(t)) \text{ almost everywhere in } [t_0, T] \\ x(t_0) & = & x_0 \end{cases} \tag{1.10}$$

and of the convexified (relaxed) differential inclusion:

$$\begin{cases} x'(t) & \in & \overline{co}\, F(t, x(t)) \text{ almost everywhere in } [t_0, T] \\ x(t_0) & = & x_0 \end{cases} \tag{1.11}$$

Observe that if F satisfies (1.9), then so does the set-valued map $(t, x) \rightsquigarrow \overline{co}(F(t, x))$.

Theorem 1.3.5 *Let $y : [t_0, T] \mapsto \mathbf{R}^n$ be a solution to the relaxed inclusion (1.11). Assume that F and y satisfy (1.9) and that the set-valued map $[t_0, T] \ni t \rightsquigarrow F(t, y(t))$ has at least one integrable selection (or, equivalently, that the map $t \mapsto dist(0, F(t, y(t)))$ is integrable.)*

Then for every $\varepsilon > 0$ there exists a solution x to (1.10) such that $\|x - y\|_C \leq \varepsilon$.

Proof — By Corollary 1.2.8 and assumptions (1.9) the set-valued map $t \rightsquigarrow F(t, y(t))$ is measurable and has closed images.

Fix $\varepsilon > 0$ so small that $\varepsilon < \beta - \varepsilon$. By Theorem 1.2.15 there exists an integrable selection $u(s) \in F(s, y(s))$ such that

$$\sup_{t \in [t_0, T]} \left\| \int_{t_0}^t (u - y')(s) ds \right\| \;\leq\; \varepsilon\, e^{-\int_{t_0}^T k(s)ds} \left(1 + \int_{t_0}^T k(s) ds\right)^{-1}$$

Define the absolutely continuous function $\overline{y} : [t_0, T] \mapsto \mathbf{R}^n$ by

$$\forall\, t \in [t_0, T], \;\; \overline{y}(t) = x_0 + \int_{t_0}^{t} u(s)ds$$

Then $\overline{y}(t_0) = x_0$ and

$$\forall\, t \in [t_0, T], \;\; \|\overline{y}(t) - y(t)\| \leq \varepsilon$$

Thus $F(t, \cdot)$ is $k(t)$–Lipschitz on the ball $B(\overline{y}(t), \beta - \varepsilon)$. Furthermore, for almost all $t \in [t_0, T]$,

$$\mathrm{dist}(\overline{y}'(t), F(t, \overline{y}(t))) \leq k(t) \|\overline{y}(t) - y(t)\| = k(t) \left\| \int_{t_0}^{t} (u - y')(s)ds \right\|$$

Set

$$\eta(t) := \int_{t_0}^{t} \mathrm{dist}(\overline{y}'(s), F(s, \overline{y}(s)))\, e^{\int_s^t k(\tau)d\tau} ds$$

Then, by the choice of u and ε, $\eta(T) \leq \varepsilon \leq \beta - \varepsilon$. Theorem 1.3.1 ends the proof. □

Theorem 1.3.6 (Relaxation) *Let $F : [t_0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ be a set-valued map with closed nonempty images and $x_0 \in \mathbf{R}^n$. Assume that there exists $k \in L^1(t_0, T)$ such that for almost every $t \in [t_0, T]$, $F(t, \cdot)$ is $k(t)$-Lipschitz and that the map $t \rightsquigarrow F(t, 0)$ has at least one integrable selection.*

Then solutions to differential inclusion (1.10) are dense in solutions to the relaxed inclusion (1.11) in the metric of uniform convergence.

Proof — It is enough to observe that for every $y \in \mathcal{C}(t_0, T; \mathbf{R}^n)$ we have $F(t, 0) \subset F(t, y(t)) + k(t) \|y(t)\| B$. Since $t \rightsquigarrow F(t, 0)$ has an integrable selection, from Proposition 1.2.9 we infer that so does the set-valued map $t \rightsquigarrow F(t, y(t))$. Theorem 1.3.5 ends the proof. □

Theorem 1.3.7 *Let $x_0 \in \mathbf{R}^n$ and $\mathcal{S}^{co}_{[t_0,T]}(x_0)$ denote the set of solutions to the relaxed inclusion (1.11). Under all assumptions of Theorem 1.3.6 suppose that the set-valued map $t \rightsquigarrow F(t, 0)$ is integrably bounded.*

Then the closure of $\mathcal{S}_{[t_0,T]}(x_0)$ in the metric of uniform convergence is compact and is equal to $\mathcal{S}^{co}_{[t_0,T]}(x_0)$.

Proof — We first show that $\mathcal{S}_{[t_0,T]}(x_0)$ is relatively compact in $\mathcal{C}(t_0, T; \mathbf{R}^n)$ (i.e., its closure is compact.) Indeed consider a sequence $x_n(\cdot) \in \mathcal{S}_{[t_0,T]}(x_0)$ and let $\psi(\cdot) \in L^1(t_0, T)$ be such that $F(t, 0) \subset \psi(t)B$ almost everywhere in $[t_0, T]$. Then for almost all $t \in [t_0, T]$ and for all $n \geq 1$ we have

$$\|x_n'(t)\| \leq \sup_{e \in F(t,0)} \|e\| + k(t)\|x_n(t)\| \leq \psi(t) + k(t)\|x_n(t)\|$$

Thus

$$\forall\, t \in [t_0, T], \ \|x_n(t)\| \leq \|x_0\| + \int_{t_0}^{t} \psi(s)ds + \int_{t_0}^{t} k(s)\|x_n(s)\|\, ds$$

This and Gronwall's lemma[5] imply that there exists $M > 0$ such that

$$\forall\, t \in [t_0, T], \ \forall\, n \geq 1, \ \|x_n(t)\| \leq M$$

Thus the sequence $x_n'(\cdot)$ is integrably bounded and thereby the sequence $x_n(\cdot)$ is equicontinuous. By the Dunford-Pettis criterion[6] a subsequence $\{x_{n_k}'\}$ converges weakly in $L^1(t_0, T; \mathbf{R}^n)$ to an integrable map $g : [t_0, T] \mapsto \mathbf{R}^n$.

Using Ascoli's theorem, taking a subsequence and keeping the same notations, we may also assume that $x_{n_k}(\cdot)$ converge uniformly

---

[5] Which states that if continuous functions u and α from $[t_0, T]$ into $\mathbf{R}_+$ satisfy

$$\forall\, t \in [t_0, T], \ u(t) \leq \alpha(t) + \int_{t_0}^{t} k(s)u(s)ds$$

for some integrable function $k \in L^1(t_0, T; \mathbf{R}_+)$, then

$$\forall\, t \in [t_0, T], \ u(t) \leq \alpha(t) + \int_{t_0}^{t} k(s)\alpha(s)\, e^{\int_s^t k(\tau)d\tau} ds$$

(see [1].)

[6] Which states that a bounded subset $K \subset L^1(t_0, T; \mathbf{R}^n)$ is weakly sequentially precompact if and only if

$$\lim_{\mu(E) \to 0+} \int_E f(s)ds = 0 \text{ uniformy for } f \text{ in } K$$

where μ denotes the Lebesgue measure on $[t_0, T]$.

to a continuous map $x : [t_0, T] \mapsto \mathbf{R}^n$. Since for every $n \geq 1$, $x_n(t) = x_0 + \int_{t_0}^{t} x_n'(s)ds$, taking the limit we obtain that

$$\forall\, t \in [t_0, T], \;\; x(t) \; = \; x_0 + \int_{t_0}^{t} g(s)ds$$

Thus $x(\cdot)$ is absolutely continuous and $x' = g$. Since

$$x_n'(t) \in \overline{co}F(t, x_n(t)) \subset \overline{co}F(t, x(t)) + k(t)\|x(t) - x_n(t)\|B$$

Mazur's theorem yields that $x(\cdot)$ is a solution to the differential inclusion (1.11). Theorem 1.3.6 ends the proof. □

1.3.3 Infinitesimal Generator of Reachable Map

Consider $T > 0$, a set-valued map $F : [0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ and let $x_0 \in \mathbf{R}^n$, $\rho > 0$ be given. In this section we assume that

$$\left\{ \begin{array}{ll} i) & \forall\, (t, x) \in [0, T] \times \mathbf{R}^n, \;\; F(t, x) \;\text{ is closed} \\ \\ ii) & \forall\, x \in \mathbf{R}^n, \;\; F(\cdot, x) \;\text{ is measurable} \\ \\ iii) & \forall\, (t, x) \in [0, T] \times B_\rho(x_0), \;\; F(t, x) \neq \emptyset \\ \\ iv) & \exists\, L > 0 \;\text{ such that for every }\; t \in [0, T], \\ & \forall\, x, y \in B_\rho(x_0), \;\; F(t, x) \;\subset\; F(t, y) + L\,\|x - y\|\,B \end{array} \right. \tag{1.12}$$

For all $0 \leq t_0 \leq t_1 \leq T$ and $\xi \in \mathbf{R}^n$ set

$$R(t_1, t_0)\xi \;:=\; \{\, x(t_1) \mid \; x \in \mathcal{S}_{[t_0, t_1]}(\xi) \,\}$$

This is the so-called *reachable set* of the inclusion

$$x' \in F(t, x) \tag{1.13}$$

from (t_0, ξ) at time t_1.

We first observe that the set-valued map R enjoys the following semigroup properties:

$$\left\{ \begin{array}{l} \forall\, 0 \leq t_1 \leq t_2 \leq t_3 \leq T, \; \forall\, \xi \in \mathbf{R}^n, \;\; R(t_3, t_2)R(t_2, t_1)\xi = R(t_3, t_1)\xi \\ \\ \forall\, 0 \leq t \leq T, \;\;\; \forall\, \xi \in \mathbf{R}^n, \;\; R(t, t)\xi \;=\; \xi \end{array} \right.$$

When F is sufficiently regular, the set-valued map $\overline{co}F(\cdot,\cdot)$ is the infinitesimal generator of the semigroup $R(\cdot,\cdot)$ in the sense that the difference quotions $(R(t+h,t)\xi-\xi)/h$ converge to $\overline{co}F(t,\xi)$:

Theorem 1.3.8 *Assume that (1.12) holds true and let* $t_0 \in [0,T[$.
If F is lower semicontinuous at (t_0,x_0), *then*

$$\overline{co}\, F(t_0,x_0) \;\subset\; \operatorname{Liminf}_{h\to 0+}\frac{R(t_0+h,t_0)x_0-x_0}{h}$$

If F is upper semicontinuous at (t_0,x_0) and $F(t_0,x_0)$ is bounded, then

$$\operatorname{Limsup}_{h\to 0+}\frac{R(t_0+h,t_0)x_0-x_0}{h} \;\subset\; \overline{co}\, F(t_0,x_0)$$

Consequently, if F is continuous at (t_0,x_0) and $F(t_0,x_0)$ is bounded, then

$$\operatorname{Lim}_{h\to 0+}\frac{R(t_0+h,t_0)x_0-x_0}{h} \;=\; \overline{co}\, F(t_0,x_0)$$

Proof — The set-valued map $(t,x) \rightsquigarrow \overline{co}\, F(t,x)$ is lower semicontinuous at (t_0,x_0) if so is F. Fix $u \in \overline{co}F(t_0,x_0)$. By Corollary 1.3.3, there exist $t_1 > t_0$ and a solution $x(\cdot)$ to the relaxed inclusion (1.11) with T replaced by t_1 such that $x'(t_0) = u$. Using Theorem 1.3.5, we deduce that for every sufficiently small $h > 0$, there exists $x_h(\cdot) \in \mathcal{S}_{[t_0,t_0+h]}(x_0)$ such that $\|x_h(t_0+h) - x(t_0+h)\| \le h^2$. Hence

$$u \;\in\; \operatorname{Liminf}_{h\to 0+}\frac{R(t_0+h,t_0)x_0-x_0}{h}$$

Since u is an arbitrary point in $\overline{co}F(t_0,x_0)$, the first statement follows.

To prove the second one we first observe that our assumptions imply that for some $\varepsilon > 0$, $M > 0$ and all $t \in [t_0,t_0+\varepsilon]$, $x \in B_\varepsilon(x_0)$ we have $F(t,x) \subset MB$. This yields that for some $t_1 > t_0$ and all $x \in \mathcal{S}_{[t_0,t_1]}(x_0)$

$$\forall\, t \;\in\; [t_0,t_1],\;\; \|x(t)-x_0\| \;\le\; M(t-t_0)$$

Fix $v \in \operatorname{Limsup}_{h\to 0+}[R(t_0+h,t_0)x_0 - x_0]/h$ and consider a sequence $h_n > 0$ converging to zero and $x_n(\cdot) \in \mathcal{S}_{[t_0,t_0+h_n]}(x_0)$ such that

$$v \;=\; \lim_{n\to\infty}\frac{x_n(t_0+h_n)-x_n(t_0)}{h_n}$$

Since F is upper semicontinuous at (t_0, x_0), there exist $\varepsilon_n \to 0+$ such that

$$\forall\, t \in [t_0, t_0 + h_n],\; F(t, x_0) \subset F(t_0, x_0) + \varepsilon_n B$$

Since for all large n

$$x_n(t_0 + h_n) - x_n(t_0) \in \int_{t_0}^{t_0+h_n} F(t, x_n(t))dt$$

$$\subset \int_{t_0}^{t_0+h_n} F(t, x_0)dt + \left(\int_{t_0}^{t_0+h_n} L\,\|x_n(t) - x_0\|\,dt\right) B$$

$$\subset \int_{t_0}^{t_0+h_n} F(t_0, x_0)dt + \left(\int_{t_0}^{t_0+h_n} (\varepsilon_n + LM(t - t_0))dt\right) B$$

$$\subset h_n \overline{co}(F(t_0, x_0)) + \left(\varepsilon_n h_n + LMh_n^2\right) B$$

dividing by h_n and taking the limit we get $v \in \overline{co}(F(t_0, x_0))$.

1.3.4 Variational Inclusions

This section is devoted to differentiability of solutions to differential inclusion (1.7) with respect to the initial condition.

We denote by $d_x F(t, \overline{x}, \overline{y})$ the adjacent derivative of $F(t, \cdot, \cdot)$ (with respect to x) of the set-valued map $F(t, \cdot)$ at $(\overline{x}, \overline{y}) \in \text{Graph}(F(t, \cdot))$.

Theorem 1.3.9 (Adjacent variational inclusion) [5] *We consider the solution map $\mathcal{S}_{[t_0,T]}(\cdot)$ as the set-valued map from $\mathbf{R}^n$ to $W^{1,1}(t_0, T; \mathbf{R}^n)$ and a solution $y(\cdot)$ to differential inclusion (1.10). Assume that* (1.9) *holds true, $u \in \mathbf{R}^n$ and let $w \in W^{1,1}(t_0, T; \mathbf{R}^n)$ be a solution to the linearized inclusion.*

$$\left\{ \begin{array}{lll} w'(t) & \in & d_x F(t, y(t), y'(t))(w(t)) \text{ a.e. in } [t_0, T] \\ w(t_0) & = & u \end{array} \right. \tag{1.14}$$

Then for all $u_h \in \mathbf{R}^n$ converging to u when $h \to 0+$ and for all small $h > 0$, there exists $x_h \in \mathcal{S}_{[t_0,T]}(x_0 + hu_h)$ such that the difference quotions $(x_h - x)/h$ converge to w in $W^{1,1}(t_0, T; \mathbf{R}^n)$ when $h \to 0+$.

In particular, $w \in d\,\mathcal{S}(x_0, y(\cdot))(u)$.

The above result was proved in [5] in the case when $u_h = u$. Corollary 1.3.4 allows to extend it to an arbitrary sequence u_h.

Theorem 1.3.10 (Convex adjacent variational inclusion) *We consider the solution map $\mathcal{S}_{[t_0,T]}(\cdot)$ as the set-valued map from $\mathbf{R}^n$ to $\mathcal{C}(t_0,T;\mathbf{R}^n)$. Let y be a solution to the differential inclusion (1.10).*

Assume that (1.9) *holds true, $u \in \mathbf{R}^n$ and let w be a solution to the inclusion*

$$\begin{cases} w'(t) & \in & d_x(\overline{co}\, F)(t, y(t), y'(t))(w(t)) \text{ a.e. in } [t_0, T] \\ w(t_0) & = & u \end{cases}$$

Then for all $u_h \in \mathbf{R}^n$ converging to u when $h \to 0+$ and for all small $h > 0$, there exists $x_h \in \mathcal{S}_{[t_0,T]}(x_0 + hu_h)$ such that the difference quotions $(x_h - x)/h$ converge to w in $\mathcal{C}(t_0,T;\mathbf{R}^n)$ when $h \to 0+$.

Proof — It is enough to apply Theorems 1.3.5 and 1.3.9. □

1.3.5 Viability Theorem

We recall here some definitions and the statement of Viability Theorem.

Let $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ be a set-valued map and $K \subset \mathrm{Dom}(F)$ be a nonempty subset.

The subset K enjoys the *viability property* for the differential inclusion

$$x' \in F(x) \tag{1.15}$$

if for any initial state $x_0 \in K$, there exists at least one solution $x(\cdot)$ to (1.15) starting at x_0 which is viable in K in the sense that $x(t) \in K$ for all $t \geq 0$. The viability property is said to be *local* if for any initial state $x_0 \in K$, there exist $T(x_0) > 0$ and a solution starting at x_0 which is viable in K on the interval $[0, T(x_0)]$ in the sense that for every $t \in [0, T(x_0)]$, $x(t) \in K$.

We say that K is a *viability domain* of F if

$$\forall\, x \in K, \quad R(x) := F(x) \cap T_K(x) \neq \emptyset$$

Theorem 1.3.11 (Viability Theorem) *If F is upper semicontinuous with nonempty compact convex images, then a locally compact set K enjoys the local viability property if and only if it is a viability domain of F. In this case, if for some $c > 0$, we have*

$$\forall\, x \in K, \ \|R(x)\| := \inf_{u \in R(x)} \|u\| \leq c(\|x\| + 1)$$

and if K is closed, then K enjoys the viability property.

We refer to [4, Aubin] for the proof and many applications of viability theory.

The following result provides a very useful *duality* characterization of viability domains:

Proposition 1.3.12 (Ushakov, [27]) *Assume that the set-valued map $F : K \rightsquigarrow \mathbf{R}^n$ is upper semicontinuous with convex compact values. Then the following three statements are equivalent:*

$$\begin{array}{ll} i) & \forall\, x \in K, \;\; F(x) \cap T_K(x) \neq \emptyset \\ ii) & \forall\, x \in K, \;\; F(x) \cap \overline{co}\,(T_K(x)) \neq \emptyset \\ iii) & \forall\, x \in K, \;\; \forall\, p \in N_K^0(x), \;\; \sigma(F(x), -p) \geq 0 \end{array} \tag{1.16}$$

where $\sigma(F(x), \cdot)$ denotes the support function of $F(x)$.

(see for instance [5] for the proof).

1.4 Parametrization of Set-Valued Maps

We recall here few results concerning parametrization of set-valued maps. Their proofs can be found in [5, Chapter 9]. Theorems comparing solutions to differential inclusion and solutions to the corresponding parametrized system will be provided in the next chapter.

Consider a metric space X, reals $t_0 < T$ and a set-valued map $F : [t_0, T] \times X \rightsquigarrow \mathbf{R}^n$.

Definition 1.4.1 *Consider subsets $C(t) \subset X$, where $t \in [t_0, T]$. The set-valued map F is called* measurable/Lipschitz *on $\{C(t)\}_{t \in [t_0,T]}$ if for every $t \in [t_0, T]$, there exists $k(t) \geq 0$ such that*

$$\left\{ \begin{array}{l} \forall\, x \in X, \; F(\cdot, x) \text{ is measurable} \\ \forall\, t \in [t_0, T], \forall\, x \in C(t), \;\; F(t, x) \neq \emptyset \text{ and is closed} \\ \forall\, t \in [t_0, T], \;\; F(t, \cdot) \text{ is } k(t) - \text{Lipschitz on } C(t) \end{array} \right.$$

Definition 1.4.2 *Let U be a metric space and $C(t) \subset X$, $t \in [t_0, T]$ be given nonempty subsets of X. We say that a single-valued map*

$$f : [t_0, T] \times X \times U \mapsto \mathbf{R}^n$$

is a measurable/Lipschitz parametrization *of F on $\{C(t)\}_{t \in [t_0,T]}$ with the constants $k(t)$, $t \in [t_0, T]$ if*

$$\begin{cases} i) & \forall\, (t,x) \in [t_0,T] \times X, \;\; F(t,x) = f(t,x,U) \\ ii) & \forall\, (x,u) \in X \times U, \;\; f(\cdot,x,u) \text{ is measurable} \\ iii) & \forall\, (t,u) \in [t_0,T] \times U, \;\; f(t,\cdot,u) \text{ is } k(t)\text{-Lipschitz on } C(t) \\ iv) & \forall\, (t,x) \in [t_0,T] \times X, \;\; f(t,x,\cdot) \text{ is continuous} \end{cases}$$

Theorem 1.4.3 (Parametrization of Unbounded Maps) *Consider a metric space X and a set-valued map $F : [t_0, T] \times X \rightsquigarrow \mathbf{R}^n$ with closed convex images.*

Assume that F is measurable/Lipschitz on $\{C(t)\}_{t \in [t_0,T]}$ and let $k(t)$, $t \in [t_0, T]$ denote the corresponding Lipschitz constants.

Then there exists a measurable/Lipschitz parametrization f of F on $\{C(t)\}_{t \in [t_0,T]}$ with $U = \mathbf{R}^n$ such that:

$$\begin{cases} \forall\, (t,u) \in [t_0,T] \times \mathbf{R}^n, \;\; f(t,\cdot,u) \text{ is } ck(t) - \text{Lipschitz on } C(t) \\ \forall\, (t,x) \in [t_0,T] \times X, \;\; f(t,x,\cdot) \text{ is } c - \text{Lipschitz on } \mathbf{R}^n \end{cases}$$

with c independent of F. Furthermore if F is continuous, so is f.

Theorem 1.4.4 (Parametrization of Bounded Maps) *Under the assumptions of Theorem 1.4.3 suppose that the images of F are compact.*

Then there exists a measurable/Lipschitz parametrization f of F on the family of sets $\{C(t)\}_{t \in [t_0,T]}$ with U equal to the closed unit ball B in $\mathbf{R}^n$ such that:

$$\begin{cases} i) & \forall\, (t,u) \in [t_0,T] \times B, \;\; f(t,\cdot,u) \text{ is } ck(t) - \text{Lipschitz on } C(t) \\ ii) & \forall\, t \in [t_0,T], \; \forall\, x \in X, \; \forall\, u, v \in B \\ & \|f(t,x,u) - f(t,x,v)\| \le c\left(\max_{y \in F(t,x)} \|y\|\right) \|u - v\| \end{cases}$$

with c independent of F. Furthermore if F is continuous, so is f.

Chapter 2

Control Systems and Differential Inclusions

In this chapter we discuss several types of control systems and their relations to differential inclusions. Namely, we shall single out

- Explicit control systems
- State dependent control systems
- Implicit control systems

The explicit control system

$$x' = f(t, x, u(t)), \quad u(t) \in U(t)$$

is the most investigated in the literature. It is well adapted to the techniques of Ordinary Differential Equations and can be seen as a parametrized family of ODE's. Indeed let us define the set of admissible controls $\mathcal{U}$ as the set of all measurable selections $u(t) \in U(t)$ and with every $u(\cdot) \in \mathcal{U}$, let us associate $\varphi_u(t, x) = f(t, x, u(t))$. Then the above control system may be replaced by ordinary differential equations

$$x' = \varphi_u(t, x), \quad u \in \mathcal{U}$$

So questions of existence, uniqueness and differentiability of solutions with respect to initial conditions may still be investigated using classical results. Another possible approach is to define the set-valued map F by $F(t,x) = f(t,x,U(t))$ and to consider the differential inclusion

$$x' \in F(t,x) \tag{2.1}$$

In Section 1 we show that under quite mild assumptions on the maps f and U, these two problems are equivalent. We apply this fact and variational inclusions from Chapter 1 to characterize variations of solutions. This will be used in Chapters 3 and 5 to prove necessary conditions for optimality.

State dependent control systems

$$x' = f(t,x,u(t)), \quad u(t) \in U(t,x)$$

present additional difficulties: we can no longer choose controls independently of the state. A possible solution to this would be to pick first a selection $u(t,x) \in U(t,x)$ and then to consider the differential equation

$$x' = f(t,x,u(t,x))$$

However we have to use classical existence theorems to guarantee existence of a solution to such equation and, thereby, to assume at least continuity of u with respect to the state variable x. This would exclude a quite large number of solutions, because it is not possible to associate with every of them such regular selection u. This is why it is more natural in this case to use differential inclusion (2.1) with the set-valued map

$$F(t,x) = f(t,x,U(t,x)) = \{ f(t,x,u) \mid u \in U(t,x) \}$$

In Section 2 we show that this new system has the same solution set and prove some results about variations of solutions.

Linear implicit system (*descriptor system*)

$$Ex' = Ax + Bu(t), \quad u(t) \in U$$

where E, A, B are possibly rectangular matrices, arises in models of electrical networks. When E, A are square, the above system is sometimes called *singular* because E may be noninvertible. Solutions to such system are usually understood in the distributional sense. Here we restict our attention to *absolutely continuous solutions* only and prove in Section 3 that this implicit system may be reduced to the explicit one (in the sense that the sets of solutions are the same):

$$x' = Dx + v(t), \quad v(t) \in V, \quad x \in Q$$

where $V \subset Q$ are subspaces obtained using E, A, B and D is a linear operator from Q into itself whose range is orthogonal to V.

Nonlinear implicit control systems

$$f(x, x', u(t)) = 0, \quad u(t) \in U$$

appear often in different models. To investigate them, we shall use differential inclusion (2.1) with the set-valued map $F(t,x) = \{v \mid 0 \in f(x,v,U)\}$ for answering in Section 4 the same type of questions: comparison of solution sets and variations of solutions.

Although the nature of these systems appear to be different, the differential inclusion formulation allows to develop a unified approach to all of them. However one should always keep in mind that differential inclusions being rather an abstract representation, their investigation would remain unsatisfactory as long as the results are not translated in terms of the original systems. This is why we are also computing derivatives and variations of set-valued maps F defined in the above examples.

2.1 Nonlinear Control Systems

Consider a complete separable metric space $\mathcal{Z}$, real numbers $t_0 < T$ and a map (describing the dynamics)

$$f : [t_0, T] \times \mathbf{R}^n \times \mathcal{Z} \mapsto \mathbf{R}^n$$

Let $U : [t_0, T] \rightsquigarrow \mathcal{Z}$ be a set-valued map (of controls) with nonempty images. We associate with these data the control system

$$x' = f(t, x, u(t)), \quad u(t) \in U(t), \quad t \in [t_0, T] \tag{2.2}$$

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a solution to (2.2) if there exists a measurable map $u : [t_0, T] \mapsto \mathcal{Z}$, called *admissible control*, such that

$$x'(t) = f(t, x(t), u(t)), \quad u(t) \in U(t) \text{ almost everywhere in } [t_0, T]$$

2.1.1 Reduction to Differential Inclusion

Define the set-valued map from $[t_0, T] \times \mathbf{R}^n$ to $\mathbf{R}^n$ by

$$F(t,x) \; = \; f(t,x,U(t))$$

and consider the differential inclusion

$$x'(t) \; \in \; F(t,x(t)) \;\text{ almost everywhere in }\; [t_0,T] \qquad (2.3)$$

Clearly every solution x to control system (2.2) satisfies (2.3). Hence x is also a solution to differential inclusion (2.3).

The natural question arises whether (2.3) has the same solutions than the control system (2.2)? The answer is positive for a quite large class of maps f.

We impose the following assumptions on f and U:

$$\begin{cases} \forall\,(x,u) \;\in\; \mathbf{R}^n \times \mathcal{Z}, \quad f(\cdot,x,u) \;\text{ is measurable} \\ \\ \forall\, t \;\in\; [t_0,T], \quad f(t,\cdot,\cdot) \;\text{ is continuous} \\ \\ U(\cdot) \;\text{ is measurable and has closed nonempty images} \end{cases} \qquad (2.4)$$

Theorem 2.1.1 *Assume that (2.4) holds true. Then the set of solutions to control system (2.2) coincide with the set of solutions to differential inclusion (2.3).*

Proof — Fix a solution $x(\cdot)$ to differential inclusion (2.3). By Theorem 1.2.12 and our assumptions, the map $(t,u) \mapsto f(t,x(t),u)$ is Carathéodory. So the proof follows from Proposition 1.2.10. □

The images of the set-valued map F defined above in general are not closed, while most Theorems of Chapter 1 deal only with closed valued maps. We provide next two results concerning "closure" of F.

Proposition 2.1.2 *Assume (2.4) and define the set-valued map clF by*

$$\forall\,(t,x) \;\in\; [t_0,T] \times \mathbf{R}^n, \quad clF(t,x) \;=\; \overline{f(t,x,U(t))}$$

Then $clF(\cdot,x)$ is measurable for every $x \in \mathbf{R}^n$. Furthermore if for some $x_0 \in \mathbf{R}^n$, $\varepsilon > 0$, $\bar{t} \in [t_0,T]$ and all $u \in U(\bar{t})$, $f(\bar{t},\cdot,u)$ is $k(\bar{t})$-Lipschitz on $B_\varepsilon(x_0)$, then so is $clF(\bar{t},\cdot)$. Finally, if $U(\cdot)$ has compact images, then so does F and, consequently, $clF = F$.

Proof — Measurability follows from Theorem 1.2.12. The proof of the last two statements is obvious. □

Theorem 2.1.3 *Assume (2.4) and let $\overline{x}(\cdot)$ be a solution to the differential inclusion*

$$x'(t) \in clF(t, x(t)) \text{ almost everywhere in } [t_0, T] \tag{2.5}$$

Further assume that there exist $\rho > 0$ and $k \in L^1(t_0, T)$ such that for for almost every $t \in [t_0, T]$ and all $u \in U(t)$, the map $f(t, \cdot, u)$ is $k(t)$-Lipschitz on $B_\rho(\overline{x}(t))$.

Then for all $\varepsilon > 0$ there exists a solution $x(\cdot)$ to (2.2) such that $x(t_0) = \overline{x}(t_0)$ and $\|x - \overline{x}\|_{W^{1,1}} \leq \varepsilon$.

Proof — By Theorem 1.2.12 and (2.4) the map $(t, u) \mapsto f(t, \overline{x}(t), u)$ is Carathéodory. Fix $\varepsilon > 0$, $N \geq 1$. By Proposition 1.2.10 there exists a measurable selection $u(t) \in U(t)$ such that

$$\|\overline{x}'(t) - f(t, \overline{x}(t), u(t))\| \leq \varepsilon/N$$

Consider the system

$$x' = f(t, x, u(t)), \quad x(t_0) = \overline{x}(t_0)$$

Choosing N large enough and using Filippov's Theorem 1.3.1 with $F(t, x) = f(t, x, u(t))$ and $y = \overline{x}$ we end the proof. □

Theorem 2.1.4 *Assume that (2.4) holds true and for some $\gamma \in L^1(t_0, T)$ and for almost all $t \in [t_0, T]$*

$$\forall\, x \in \mathbf{R}^n, \quad \sup_{u \in U(t)} \|f(t, x, u)\| \leq \gamma(t)(1 + \|x\|)$$

Further assume that for every $R > 0$ there exists $k_R \in L^1(t_0, T)$ such that for almost all $t \in [t_0, T]$ and for every $u \in U(t)$, $f(t, \cdot, u)$ is $k_R(t)$-Lipschitz on $B_R(0)$.

If the sets $f(t, x, U(t))$ are closed and convex, then the set of solutions to control system (2.2) starting at x_0 is compact in $C(t_0, T; \mathbf{R}^n)$.

Proof — It is enough to apply Theorems 2.1.1 and 1.3.7. □

2.1.2 Linearization

Consider a solution z to control system (2.2) and let $\overline{u}$ be a corresponding control. We associate with it the following linearization of (2.2) along the solution-control pair $(z, \overline{u})$:

$$\begin{cases} w'(t) \ = \ \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t))w(t) \ + \ v(t) \\ \\ v(t) \in V(t) := T_{\overline{co}f(t,z(t),U(t))}(f(t, z(t), \overline{u}(t))) \text{ a.e.} \end{cases} \tag{2.6}$$

where $T_{\overline{co}f(t,z(t),U(t))}(f(t, z(t), \overline{u}(t)))$ denotes the tangent cone to the convex set $\overline{co}f(t, z(t), U(t))$ at $f(t, z(t), \overline{u}(t))$.

We assume that

$$\begin{cases} \text{The derivative } \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t)) \text{ exists a.e. in } [t_0, T] \\ \\ \text{For some } \varepsilon > 0, \ k \in L^1(t_0, T) \text{ and for a.e. } t \in [t_0, T] \\ \forall\, u \in U(t), \ \ f(t, \cdot, u) \text{ is } k(t) - \text{Lipschitz on } B_\varepsilon(z(t)) \end{cases} \tag{2.7}$$

Recall that the solution $w(\cdot)$ to (2.6) starting at w_0 and corresponding to an integrable selection $v(s) \in V(s)$ is given by

$$\forall\, t \in [t_0, T], \ \ w(t) \ = \ X(t)w_0 \ + \ \int_{t_0}^{t} X(t)X(s)^{-1}v(s)ds$$

where $X(\cdot)$ denotes the fundamental solution to the linear system

$$X'(t) \ = \ \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t))X(t), \quad X(t_0) \ = \ Id \tag{2.8}$$

Theorem 2.1.5 *Assume that (2.4) and (2.7) hold true. Then for every solution $w(\cdot)$ to linearized system (2.6) and elements $\{w_h\}_{h>0}$ in $\mathbf{R}^n$ satisfying $\lim_{h\to 0+} w_h = w(t_0)$, there exist solutions $\{x_h\}_{h>0}$ to (2.2) such that*

$$x_h(t_0) \ = \ z(t_0) + hw_h \ \text{ for all } \ h > 0 \ \text{ small enough}$$

and the difference quotients $(x_h - z)/h$ converge uniformly to w when h goes to zero.

Proof — By Theorem 2.1.3 we may replace control system (2.2) by differential inclusion (2.5). Propositions 2.1.2, 1.2.5 allow to apply the variational inclusion (Theorem 1.3.10) and to deduce the result after observing that

$$\forall\, w, \quad \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t))w \;\in\; d_x F(t, z(t), z'(t))(w) \quad \text{a.e. in } [t_0, T] \quad \square$$

2.2 State Dependent Control Systems

In the previous section we have considered the map of controls $U(\cdot)$ depending only on time. When it also depends on the states, then the control system is called a *state dependent* control system.

Let $\mathcal{Z}$ be a complete separable metric space and let

$$U : [t_0, T] \times \mathbf{R}^n \rightsquigarrow \mathcal{Z}$$

be a given set-valued map. Consider the control system

$$x' \;=\; f(t, x, u), \quad u \;\in\; U(t, x), \quad t \;\in\; [t_0, T] \tag{2.9}$$

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a solution to (2.9) if for some measurable selection $u(t) \in U(t, x(t))$ we have

$$x'(t) \;=\; f(t, x(t), u(t)) \text{ almost everywhere in } [t_0, T]$$

2.2.1 Reduction to Differential Inclusion

We introduce the set-valued map $F : [t_0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ defined by

$$F(t, x) \;=\; f(t, x, U(t, x)) \;=\; \{f(t, x, v) \mid v \;\in\; U(t, x)\}$$

and replace (2.9) by the differential inclusion

$$x'(t) \;\in\; F(t, x(t)) \text{ almost everywhere in } [t_0, T] \tag{2.10}$$

We impose the following assumptions:

$$\begin{cases} \forall\, (x, u) \;\in\; \mathbf{R}^n \times \mathcal{Z}, \;\; f(\cdot, x, u) \text{ is measurable} \\ \\ \forall\, t \;\in\; [t_0, T], \;\; f(t, \cdot, \cdot) \text{ is continuous} \\ \\ U \text{ is Carathéodory and has closed nonempty images} \end{cases} \tag{2.11}$$

Theorem 2.2.1 *If (2.11) holds true, then the sets of solutions to control system (2.9) and differential inclusion (2.10) do coincide.*

Proof — Clearly every solution to (2.9) solves also (2.10). Conversely, consider a solution x to differential inclusion (2.10). By Theorem 1.2.12 the set-valued map $t \rightsquigarrow U(t, x(t))$ is measurable and the map $(t, u) \mapsto f(t, x(t), u)$ is Carathéodory. Applying Proposition 1.2.10 we can find a measurable selection $u(t) \in U(t, x(t))$ such that $x'(t) = f(t, x(t), u(t))$ almost everywhere in $[t_0, T]$. □

Hence we can rewrite dynamical system (2.9) in the differential inclusion formulation (2.10). In general F does not have closed values. However, using arguments comparable to those from the proof of Theorem 2.1.3 we get

Proposition 2.2.2 *Assume that (2.11) holds true. Then the set-valued map $clF : [t_0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ defined by*

$$clF(t, x) \; = \; \overline{f(t, x, U(t, x))}$$

is measurable with respect to t. Furthermore if for some $t \in [t_0, T]$, $k(t) \geq 0$, $l(t) \geq 0$, $x_0 \in \mathbf{R}^n$, $\rho > 0$, the map $f(t, \cdot, \cdot)$ is $k(t)-$Lipschitz on $B_\rho(x_0) \times \mathcal{Z}$ and the set-valued map $U(t, \cdot)$ is $l(t)-$Lipschitz on $B_\rho(x_0)$, then $clF(t, \cdot)$ is $k(t)(1 + l(t))-$Lipschitz on $B_\rho(x_0)$.

Let $\overline{x}(\cdot)$ be a solution to the differential inclusion

$$x'(t) \; \in \; clF(t, x(t)) \text{ almost everywhere in } [t_0, T] \tag{2.12}$$

Further assume that there exist $\rho > 0$ and $k \in L^1(t_0, T)$ such that for for almost every $t \in [t_0, T]$ and all $u \in U(t, x)$, the map $f(t, \cdot, u)$ is $k(t)$-Lipschitz on $B_\rho(\overline{x}(t))$.

Then for all $\varepsilon > 0$ there exists a solution $x(\cdot)$ to (2.9) such that $x(t_0) = \overline{x}(t_0)$ and $\|x - \overline{x}\|_{W^{1,1}} \leq \varepsilon$.

2.2.2 Linearization

In this section we assume that $\mathcal{Z}$ is a separable Banach space. Consider a solution z to (2.9) and let $\overline{u}(t) \in U(t, z(t))$ be a corresponding control. We associate to it the following linearization of (2.9) along the pair $(z, \overline{u})$:

$$\begin{cases} w'(t) \in A(t)w(t) + B(t)d_xU(t, z(t), \overline{u}(t))(w(t)) + v(t) \\ v(t) \in T_{\overline{co}f(t,z(t),U(t,z(t)))}(f(t, z(t), \overline{u}(t))) \text{ a.e. in } [t_0, T] \end{cases} \tag{2.13}$$

where

$$A(t) = \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t)), \quad B(t) = \frac{\partial f}{\partial u}(t, z(t), \overline{u}(t))$$

and d_xU denotes the (partial) adjacent derivative of U with respect to the state variable x.

We impose the following assumptions

$$\begin{cases} \text{The derivative } \frac{\partial f}{\partial(x,u)}(t, z(t), \overline{u}(t)) \text{ exists a.e. in } [t_0, T] \\ \exists\, \varepsilon > 0 \text{ and functions } k,\ l : [t_0, T] \mapsto \mathbf{R}_+ \text{ such that} \\ f(t, \cdot, \cdot) \text{ is } k(t) - \text{Lipschitz on } B_\varepsilon(z(t)) \times \mathcal{Z} \text{ and} \\ U(t, \cdot) \text{ is } l(t) - \text{Lipschitz on } B_\varepsilon(z(t)) \text{ for a.e. } t \in [t_0, T] \end{cases} \tag{2.14}$$

Theorem 2.2.3 *Assume that $\mathcal{Z}$ is a separable Banach space, that (2.11), (2.14) hold true and the map $t \mapsto k(t)(1 + l(t))$ is integrable.*

If at least one of the following two conditions holds true:

$$\begin{cases} i) & \forall\, (t, x) \in [t_0, T] \times \mathbf{R}^n, \ f(t, x, U(t, x)) \text{ is closed} \\ ii) & \forall\, (t, x) \in [t_0, T] \times \mathbf{R}^n, \ U(t, x) \text{ is convex} \end{cases}$$

then for every solution $w(\cdot)$ to (2.13) and elements $\{w_h\}_{h>0}$ in $\mathbf{R}^n$ satisfying $\lim_{h \to 0+} w_h = w(t_0)$, there exists a family $\{x_h(\cdot)\}_{h>0}$ of solutions to (2.9) such that

$$x_h(t_0) = z(t_0) + hw_h \text{ for all small } h > 0$$

and the difference quotients $(x_h - z)/h$ converge uniformly to w when h goes to zero.

Proof — We apply the variational inclusion (Theorem 1.3.10) to deduce the result from the following relation:

$$\forall\, v \in \mathbf{R}^n,\ A(t)v + B(t)d_xU(t, z(t), \overline{u}(t))v \subset d_xF(t, z(t), z'(t))v$$

for almost all $t \in [t_0, T]$. □

2.3 Linear Implicit Control Systems

Let $E, A \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ be linear operators from $\mathbf{R}^n$ into $\mathbf{R}^m$, U be a finite dimensional vector space and $B \in \mathcal{L}(U, \mathbf{R}^m)$. Consider the implicit control system

$$Ex' = Ax + Bu, \quad u \in U \tag{2.15}$$

When $n = m$ this system is sometimes called *singular*, because E may be noninvertible.

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a solution to (2.15) corresponding to a measurable control $u : [t_0, T] \mapsto U$ if

$$Ex'(t) = Ax(t) + Bu(t) \text{ almost everywhere in } [t_0, T]$$

Our aim is to reduce (2.15) to the explicit system

$$x' = Dx + v, \quad v \in V, \quad x \in Q \tag{2.16}$$

where $V \subset Q$ are subspaces of $\mathbf{R}^n$ and D is a linear operator from Q into Q.

Let us denote by $\mathcal{B}$ the range of B and for every $y \in \mathbf{R}^m$ set

$$E^{-1}(y) = \{ x \in \mathbf{R}^n \mid Ex = y \}$$

We introduce a *decreasing family* of subspaces:

$$K_0 = \mathbf{R}^n, \ldots, K_{k+1} = A^{-1}(EK_k + \mathcal{B}), \quad k \geq 0$$

Since they are subspaces of $\mathbf{R}^n$, we obtain

$$Q := \bigcap_{k \geq 1} K_k = K_j$$

for some $j \leq n - 1$. Furthermore $A^{-1}(EQ + \mathcal{B}) = Q$ and therefore the set-valued map $\mathcal{F} : Q \rightsquigarrow Q$ given by

$$\forall x \in Q, \quad \mathcal{F}(x) := E^{-1}(Ax + \mathcal{B}) \cap Q$$

has nonempty images. It is also clear that for every $x \in Q$, $\mathcal{F}(x)$ is an affine subspace of Q.

Theorem 2.3.1 *Every solution $x(\cdot)$ to (2.15) defined on the time interval $[t_0, T]$ satisfies $x(t) \in Q$ for all $t \in [t_0, T]$.*

Proof — Fix a solution $x : [t_0, T] \mapsto \mathbf{R}^n = K_0$. Assume that we already know that for some $0 \leq k < n-1$, $x(t) \in K_k$ for all t. Then $x'(t) \in K_k$ almost everywhere and, consequently,

$$x(t) \in A^{-1}(Ex'(t) + \mathcal{B}) \subset A^{-1}(EK_k + \mathcal{B}) = K_{k+1} \text{ for a.e. } t \in [t_0, T]$$

Continuity of $x(\cdot)$ yields that $x(t) \in K_{k+1}$ for all $t \in [t_0, T]$ and the proof ends by the induction argument. □

Let the map $\mathcal{D} : Q \mapsto Q$ and the subspace $V \subset Q$ be defined by

$$\forall\, x \in Q,\ \ \mathcal{D}x \in \mathcal{F}(x),\ \|\mathcal{D}x\| = \min_{y \in \mathcal{F}(x)} \|y\|\,,\ V = E^{-1}(\mathcal{B}) \cap Q = \mathcal{F}(0)$$

Proposition 2.3.2 *The map $\mathcal{D}$ defined above is a linear operator from Q into itself. Furthermore for every $x \in Q$, $\mathcal{F}(x) = \mathcal{D}x + V$ and $\mathcal{D}x$ is orthogonal to V.*

Proof — Since Graph($\mathcal{F}$) is a subspace,

$$\mathcal{D}x + V \ \subset\ \mathcal{F}(x) + V \ =\ \mathcal{F}(x) + \mathcal{F}(0) \ \subset\ \mathcal{F}(x)$$

To prove the equality, consider $y \in \mathcal{F}(x) \subset Q$. Then $Ey \in Ax + \mathcal{B}$, $E\mathcal{D}x \in Ax + \mathcal{B}$ and therefore $y - \mathcal{D}x \in Q$ and $E(y - \mathcal{D}x) \in \mathcal{B}$. Hence $y - \mathcal{D}x \in V$ and $\mathcal{F}(x) = \mathcal{D}x + V$.

It remains to show that $\mathcal{D}$ is linear. The element $\mathcal{D}x$ being the orthogonal projection of zero onto $\mathcal{D}x + V$, we deduce that $\mathcal{D}(Q) \subset V^{\perp}$ (orthogonal to V in Q). Fix x, $y \in Q$. Then

$$-E\mathcal{D}x \in -Ax + \mathcal{B},\ -E\mathcal{D}y \in -Ay + \mathcal{B},\ E\mathcal{D}(x+y) \in Ax + Ay + \mathcal{B}$$

Adding these inclusions, we get $E(\mathcal{D}(x+y) - \mathcal{D}x - \mathcal{D}y) \in \mathcal{B}$. Thus

$$\mathcal{D}(x+y) - \mathcal{D}x - \mathcal{D}y \ \in\ V \cap V^{\perp} \quad \Longrightarrow \quad \mathcal{D}(x+y) \ =\ \mathcal{D}x + \mathcal{D}y$$

Finally $\mathcal{D}$ is homogeneous, because

$$\mathcal{F}(\lambda x) \ =\ E^{-1}(\lambda Ax + \mathcal{B}) \cap Q \ =\ \lambda \mathcal{F}(x) \ \square$$

Theorem 2.3.3 *Solutions to (2.15) and (2.16) do coincide. Denote by B^+ the orthogonal right inverse*[1] *of B. Then the map*

$$u(x) = \begin{cases} B^+(E\mathcal{D}x - Ax) & \text{if } x \in Q \\ \emptyset & \text{if not} \end{cases}$$

is a regulation law for (2.15): for every $x_0 \in Q$ there exists a C^∞−solution to the singular system

$$Ex' = Ax + Bu(x), \; x(0) = x_0 \tag{2.17}$$

defined on $[0, \infty[$. It is unique if and only if $\ker(E) \cap Q = \{0\}$.

Proof — By Theorem 2.3.1 and Proposition 2.3.2 every solution to (2.15) solves (2.16). Conversely every solution $x : [t_0, T] \mapsto \mathbf{R}^n$ to (2.16) satisfies $Ex'(t) \in Ax(t) + \mathcal{B}$ almost everywhere. Hence, by Proposition 1.2.10, $x(\cdot)$ solves (2.15).

Fix $x_0 \in Q$ and consider the solution $x(\cdot) \in C^\infty$ to the linear system

$$x' = \mathcal{D}x, \; x(0) = x_0$$

It is defined on $[0, +\infty[$ and is also a solution to (2.15). To prove the latter statement of our theorem, observe that (2.17) may be written as:

$$Ex' = E\mathcal{D}x, \; x(0) = x_0 \tag{2.18}$$

So the solution to (2.17) is unique if and only if the solution to (2.18) is unique. But this happens whenever zero is the only solution to the differential inclusion

$$x' \in \mathcal{D}x + \ker(E) \cap Q, \; x(0) = 0$$

Consequently uniqueness is equivalent to $\ker(E) \cap Q = \{0\}$. □

Observe that the above results allow to study implicit system (2.15) even in the case when the solution corresponding to a given control and a given initial state is not unique. We investigate next necessary and sufficient conditions for uniqueness.

[1]That is B^+ is the linear operator from $\mathcal{B}$ into U with B^+x equal to the orthogonal projection of zero onto $B^{-1}(x)$. Clearly for every $x \in \mathbf{R}^m$, $BB^+x = x$.

We say that (2.15) *enjoys uniqueness* if to every measurable control $\overline{u} : [0,T] \mapsto U$, $T > 0$ and every initial state $x_0 \in \mathbf{R}^n$ corresponds at most one solution to control system (2.15) starting at x_0. Set $(A^{-1}E)^0(\mathbf{R}^n) = \mathbf{R}^n$ and define recursively

$$\forall\, k \geq 0,\ (A^{-1}E)^{k+1}(\mathbf{R}^n) \ = \ A^{-1}E\left((A^{-1}E)^k(\mathbf{R}^n)\right)$$

Theorem 2.3.4 *Consider the subspace $P = (A^{-1}E)^{n-1}(\mathbf{R}^n)$. The following statements are equivalent :*
i) System (2.15) enjoys uniqueness
ii) $\ker E \cap P = \{\ 0\ \}$

Proof — Observe that *i)* is equivalent to: $\overline{x}(\cdot) \equiv 0$ is the only solution to the linear system $Ex' = Ax$ starting at zero. Hence, by Theorem 2.3.3 applied with $B = 0$, (2.15) enjoys uniqueness if and only if *ii)* holds true. □

2.4 Nonlinear Implicit Control Systems

Let $\mathcal{Z}$ be a complete separable metric space, $f : \mathbf{R}^n \times \mathbf{R}^n \times \mathcal{Z} \mapsto \mathbf{R}^m$ be a continuous map and $U \subset \mathcal{Z}$ be a given closed set. Consider the implicit control system

$$f(x,x',u(t)) \ = \ 0, \quad u(t) \in U,\ \ t \ \in \ [t_0,T] \tag{2.19}$$

An absolutely continuous function $x : [t_0,T] \mapsto \mathbf{R}^n$ is called a solution to (2.19) if there exists a measurable map $u : [t_0,T] \mapsto U$ such that $f(x(t),x'(t),u(t)) = 0$ almost everywhere in $[t_0,T]$.

2.4.1 Reduction to Differential Inclusion

Define the set-valued map $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ by

$$F(x) \ = \ \{v \ \in \ \mathbf{R}^n \mid \exists\, u \ \in \ U \ \text{ with } \ f(x,v,u) \ = \ 0\}$$

and consider the differential inclusion

$$x'(t) \ \in \ F(x(t)) \ \text{ almost everywhere in } \ [t_0,T] \tag{2.20}$$

Clearly every solution to (2.19) solves (2.20). The following lemma implies the converse statement.

Lemma 2.4.1 *If f is continuous, then the solution sets of (2.20) and (2.19) do coincide.*

Proof — Fix a solution x to (2.20). The existence of a measurable selection $u(t) \in U$ with $f(x(t), x'(t), u(t)) = 0$ almost everywhere in $[t_0, T]$ follows from Proposition 1.2.10. □

The introduced set-valued map F has a closed graph whenever f is continuous and U is compact. If moreover for all $\overline{x} \in \mathbf{R}^n$

$$\exists\, \varepsilon > 0 \text{ such that } \liminf_{\|v\| \to \infty} \min_{x \in B_\varepsilon(\overline{x}),\, u \in U} \|f(x, v, u)\| > 0 \qquad (2.21)$$

then

$$\exists\, R > 0 \text{ such that } \forall\, x \in B_\varepsilon(\overline{x}),\ F(x) \subset B_R(0) \qquad (2.22)$$

This and Proposition 1.2.2 yield that if (2.21) holds true and f is continuous, then F is upper semicontinuous on its domain of definition and has compact images.

Another sufficient condition for the upper semicontinuity of F is given by

Proposition 2.4.2 *Assume that $\mathcal{Z}$ is a finite dimensional vector space, f is continuous and for every $\overline{x} \in \mathbf{R}^n$*

$$\exists\, \varepsilon > 0 \text{ such that } \liminf_{\substack{\|v\| \to \infty \\ \|u\| \to \infty}} \inf_{x \in B_\varepsilon(\overline{x}),\, u \in U} \|f(x, v, u)\| > 0 \qquad (2.23)$$

Then (2.22) holds true, F is upper semicontinuous on its domain of definition and has compact images.

In general the images of F are not convex and for this reason inclusion (2.20) is not easy to investigate even when F is upper semicontinuous.

Our next aim is to provide a sufficient condition for Lipschitz continuity of F. We assume that

$$\begin{cases} f(\cdot,\cdot,\cdot) \text{ is continuous} \\ \forall\, u \in U,\ f(\cdot,\cdot,u) \text{ is differentiable} \\ \dfrac{\partial f}{\partial v}(\cdot,\cdot,\cdot) \text{ is continuous} \end{cases} \qquad (2.24)$$

Theorem 2.4.3 *Assume (2.24) and that for an open subset* $\mathcal{N} \subset \mathbf{R}^n$ *the following holds true*

$$\forall\ (x,v,u) \in f^{-1}(0) \text{ with } x \in \mathcal{N},\ u \in U,\ \frac{\partial f}{\partial v}(x,v,u) \text{ is surjective}$$

Further assume that at least one of the following two conditions is satisfied:

i) U *is compact and for every* $\overline{x} \in \mathcal{N}$ *(2.21) is valid*

ii) $\mathcal{Z}$ *is finite dimensional and every* $\overline{x} \in \mathcal{N}$ *satisfies (2.23),*

Then the set $\mathrm{Dom}(F) \cap \mathcal{N}$ *is open and* F *is locally Lipschitz on it. Furthermore for all* $(x,v,u) \in f^{-1}(0)$ *with* $(x,u) \in \mathcal{N} \times U$, *we have*

$$\ker\left(\frac{\partial f}{\partial(x,v)}(x,v,u)\right) \subset \mathrm{Graph}(dF(x,v))$$

The proof results from the inverse mapping theorems [22].

2.4.2 Linearization of Implicit Systems

Consider a solution z to (2.19) and let $\overline{u}$ be a corresponding control. We associate with it the following linear time dependent implicit system

$$\begin{cases} A(t)w(t) \ + \ B(t)(w'(t) - v(t)) \ = \ 0 \\ \\ v(t) \ \in \ T_{\overline{co}F(z(t))}(z'(t)) \ \text{ a.e. in } [t_0, T] \end{cases} \tag{2.25}$$

where

$$A(t) \ = \ \frac{\partial f}{\partial x}(z(t), z'(t), \overline{u}(t)), \quad B(t) \ = \ \frac{\partial f}{\partial v}(z(t), z'(t), \overline{u}(t))$$

Theorem 2.4.4 *Assume that all hypothezis of Theorem 2.4.3 hold true with* $\mathcal{N} = z([0,T]) + \rho B$, *where* $\rho > 0$. *Let* w *be a solution to (2.25).*

Then for all elements $\{w_h\}_{h>0}$ *in* $\mathbf{R}^n$ *satisfying* $\lim_{h\to 0+} w_h = w(t_0)$, *there exist solutions* x_h *to implicit system (2.19) such that*

$$x_h(t_0) \ = \ z(t_0) + hw_h \ \text{ for all small } h > 0$$

and the difference quotients $(x_h - z)/h$ *converge uniformly to* w *when* h *goes to zero.*

Proof — To prove this result we apply Lemma 2.4.1, Theorems 2.4.3, 1.2.5 and variational inclusion (Theorem 1.3.10). □

Chapter 3

Value Function of Mayer's Problem

In this chapter we address the Mayer problem arising in control theory. We start with the free end point case:

$$\text{minimize } g(x(T))$$

over all solutions to the control system

$$x' = f(t, x, u(t)), \quad u(t) \in U(t) \tag{3.1}$$

satisfying the initial condition

$$x(0) = \xi_0 \tag{3.2}$$

By a simple change of variables the classical Bolza problem

$$\text{minimize}\left\{ g(x(T)) + \int_0^T L(t, x(t), u(t))dt \right\}$$

over all state-control solutions (x, u) of (3.1), (3.2) may be reduced to the Mayer problem.

The basic objective of optimal control theory is to find necessary and sufficient conditions for optimality and to construct optimal solutions. Necessary conditions are available in the form of Pontriagin's maximum principle. It implements the Fermat rule in the case of optimal control problems.

When applied to an abstract minimization problem: $\min_{x\in K}\varphi(x)$, where K is a subset of a normed space, Fermat rule states that if $\overline{x}\in K$ is a minimizer, then

$$\forall\, v \;\in\; T_K(\overline{x}),\;\; \langle \nabla\varphi(\overline{x}),\, v\rangle \;\geq\; 0$$

In the above $T_K(\overline{x})$ denotes the contingent cone to K at $\overline{x}$.

In the same way as the Euler-Lagrange equation is a consequence of the Fermat rule in Calculus of Variations, the maximum principle can be deduced from the above rule: If the state-control pair $(z,\overline{u})$ is optimal, then the solution $p(\cdot)$ (called the co-state) to the adjoint system

$$p'(t) \;=\; -\left(\frac{\partial f}{\partial x}(t,z(t),\overline{u}(t))\right)^{\star} p(t),\;\; p(T) \;=\; -\nabla g(z(T))$$

satisfies the transversality condition

$$p(t) \;\in\; N^0_{R(t)}(z(t)) \;\text{ almost everywhere in }\; [0,T]$$

where $N^0_{R(t)}(z(t))$ denotes the subnormal cone to the reachable set $R(t)$ of (3.1) from ξ_0 at time t. This last inclusion implies the maximum principle:

$$\langle p(t),\; f(t,z(t),\overline{u}(t))\rangle \;=\; \max_{u\in U(t)} \langle p(t),\; f(t,z(t),u)\rangle \;\text{ a.e. in }\; [0,T]$$

In Section 3 we complete these conditions to obtain sufficient ones by using the value function

$$V(t_0,x_0) \;=\; \inf\{\, g(x(T)) \mid x \text{ is a solution to (3.1)},\; x(t_0)=x_0 \,\}$$

and the Hamiltonian H of the control system (3.1):

$$H(t,x,p) \;=\; \sup_{u\in U(t)} \langle p, f(t,x,u)\rangle$$

In general V and H are nonsmooth functions and we have to use notions of superdifferentials from Chapter 1.

The value function allows to single out optimal solutions. Indeed, it is nondecreasing along solutions to (3.1) and is constant along optimal solutions.

Sufficient conditions that we prove are of the following type: for almost every $t \in [0,T]$, there exists $p(t)$ such that

$$(\langle p(t), z'(t)\rangle\,,\; -p(t)) \in \partial_+ V(t, z(t))$$

where $\partial_+ V$ denotes the superdifferential of V. We also show that the co-state of the maximum principle verifies the above relations.

To find the value function from its definition at first glance seems to be an impossible task, because it amounts to solving as many optimization problems as there are initial points (t_0, x_0). But very fortunately, under quite general assumptions, the value function is the *unique solution* to the Hamilton-Jacobi equation:

$$-\frac{\partial V}{\partial t}(t,x) \;+\; H\left(t, x, -\frac{\partial V}{\partial x}(t,x)\right) \;=\; 0, \quad V(T,\cdot) \;=\; g(\cdot)$$

However, since even in very regular situations the value function is merely Lipschitz, solutions of the above Hamilton-Jacobi equation have to be understood in a generalized sense, where derivatives are replaced by subdifferentials (see Chapter 4.)

We also investigate what are the regularity properties of the system which are inherited by the value function (Lipschitz continuity in Section 1, semiconcavity in Section 4 and lower semicontinuity in Chapter 4.) In Section 3 we show that differentiability of V is related to uniqueness of optima and is preserved along each optimal solution.

When the Hamiltonian H is smooth enough and the value function is differentiable at $(0, \xi_0)$, then the following necessary and sufficient condition for optimality holds true:

Let $x(\cdot)$, $p(\cdot)$ solve the Hamiltonian system

$$\begin{cases} x'(t) &=& \frac{\partial H}{\partial p}(t, x(t), p(t)) \\ p'(t) &=& -\frac{\partial H}{\partial x}(t, x(t), p(t)), \quad t \;\in\; [0,T] \end{cases}$$

Then x is optimal if and only if $x(0) = \xi_0$, $p(0) = -\frac{\partial V}{\partial x}(0, \xi_0)$.

The value function can be also used to construct the optimal feedback map:

$$G(t,x) \;=\; \left\{ v \;\in\; f(t, x, U(t)) \;\mid\; \frac{\partial V}{\partial (1,v)}(t,x) \;=\; 0 \right\}$$

Namely the following property holds true: a solution $\overline{x}$ to (3.1) is optimal for our minimization problem if and only if it is a solution to the differential inclusion

$$x' \in G(t,x), \; x(0) = \xi_0 \tag{3.3}$$

To investigate the regularity of the set-valued map G, we show in Section 4 that for sufficiently smooth f and g, the value function is semiconcave. As a consequence, we obtain that the feedback map G is upper semicontinuous on $[0,T[\times X$ and has nonempty compact images. In particular whenever G is single-valued, it is continuous and optimal solutions are continuously differentiable.

If the data are convex, then the value function is convex, G has convex values and inclusion (3.3) fits the well investigated framework of upper semicontinuous convex valued maps (see [1].)

3.1 Value Function

3.1.1 Mayer and Bolza Problems

Consider $T > 0$, a complete separable metric space $\mathcal{Z}$, a set-valued map $U : [0,T] \rightsquigarrow \mathcal{Z}$ and a map $f : [0,T] \times \mathbf{R}^n \times \mathcal{Z} \mapsto \mathbf{R}^n$. We associate with it the control system

$$x'(t) = f(t,x(t),u(t)), \; u(t) \in U(t) \tag{3.4}$$

Let an extended function $g : \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ and $\xi_0 \in \mathbf{R}^n$ be given. Consider the minimization problem, called *Mayer's optimal control problem* :

$$\min \{g(x(T)) \mid x \text{ is a solution to (3.4)}, \; x(0) = \xi_0\} \tag{3.5}$$

The value function associated with this problem is defined by: for all $(t_0, x_0) \in [0,T] \times \mathbf{R}^n$

$$V(t_0,x_0) = \inf\{g(x(T)) \mid x \text{ is a solution to (3.4)}, \; x(t_0) = x_0\} \tag{3.6}$$

We impose the following assumptions

$$\begin{cases} \forall\,(x,u)\;\in\;\mathbf{R}^n\times\mathcal{Z}, \quad f(\cdot,x,u)\;\text{ is measurable} \\ \text{For a.e. }\; t\;\in\;[0,T], \quad f(t,\cdot,\cdot)\;\text{ is continuous} \\ U(\cdot)\;\text{ is measurable and has closed nonempty images} \end{cases} \tag{3.7}$$

and

$$\begin{cases} i) & \exists\, k\in L^1(0,T)\;\text{ such that for a.e. }\; t\in[0,T], \\ & \forall\, u\in U(t), f(t,\cdot,u)\;\text{ is }\; k(t)-\text{Lipschitz} \\ ii) & \exists\,\gamma\in L^1(0,T)\;\text{ such that for a.e. }\; t\in[0,T], \\ & \sup_{u\in U(t)}\|f(t,0,u)\|\;\leq\;\gamma(t) \\ iii) & \text{g is locally Lipschitz} \end{cases} \tag{3.8}$$

These assumptions imply that V is actually equal to the value function of the relaxed problem in which system (3.4) is replaced by the differential inclusion

$$x'(t)\;\in\;\overline{co}\,(f(t,x(t),U(t)))\;\text{ almost everywhere} \tag{3.9}$$

Hence for all $(t_0,x_0)\in[0,T]\times\mathbf{R}^n$,

$$V(t_0,x_0)\;=\;\inf\{g(x(T))\mid x\text{ solves (3.9)},\; x(t_0)=x_0\} \tag{3.10}$$

The *Bolza problem* has the same nature, but its cost involves the integral functional: Consider in addition a function $L:[0,T]\times\mathbf{R}^n\times\mathcal{Z}\mapsto\mathbf{R}$ and the following minimization problem:

$$\text{minimize}\left\{\;g(x(T))\;+\;\int_0^T L(t,x(t),u(t))dt\;\right\} \tag{3.11}$$

over all solution-control pairs (x,u) to (3.4) with $x(0)=\xi_0$.

We denote by $\widehat{x}=(x^0,x)$ elements of $\mathbf{R}^{n+1}$ and we set

$$\forall\, t\in[0,T],\;\widehat{x}=(x^0,x),\;u\in\mathcal{Z},\;\widehat{f}(t,\widehat{x},u):=(L(t,x,u),f(t,x,u))$$

Then it is not difficult to realize that a solution-control pair $(z, \overline{u})$ of (3.4) is optimal for problem (3.11) if and only if the map

$$t \mapsto \widehat{z}(t) := \left(\int_0^t L(s, z(s), \overline{u}(s))ds,\ z(t)\right)$$

solves the problem

$$\text{minimize}\left(g(x(T)) + x^0(T)\right)$$

over all solutions to the control system

$$\widehat{x}'(t) = \widehat{f}(t, \widehat{x}(t), u(t)),\quad u(t) \in U(t),\quad \widehat{x}(0) = (0, \xi_0)$$

This new problem is of Mayer's type.

3.1.2 Lipschitz Continuity of the Value Function

More generally consider an extended function $g : \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$, a set-valued map $F : [0,T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$, $\xi_0 \in \mathbf{R}^n$ and the differential inclusion

$$x'(t) \in F(t, x(t)) \text{ almost everywhere} \tag{3.12}$$

We investigate the minimization problem

$$\min\{g(x(T)) \mid x \text{ is a solution to (3.12)},\ x(0) = \xi_0\}$$

The corresponding value function is given by:

For all $(t_0, x_0) \in [0,T] \times \mathbf{R}^n$,

$$V(t_0, x_0) = \inf\{g(x(T)) \mid x \text{ solves (3.12)},\ x(t_0) = x_0\} \tag{3.13}$$

Let $\mathcal{S}_{[t_0,T]}(x_0)$ denote the set of solutions to (3.12) starting at x_0 at time t_0 and defined on the time interval $[t_0, T]$. The value function is nondecreasing along solutions to (3.12):

$$\forall\, x \in \mathcal{S}_{[t_0,T]}(x_0),\ \forall\, t_0 \le t_1 \le t_2 \le T,\ V(t_1, x(t_1)) \le V(t_2, x(t_2))$$

and satisfies the following *dynamic programming principle*:

$$\forall\, t \in [t_0, T],\ V(t_0, x_0) = \inf\left\{V(t, x(t)) \mid x \in \mathcal{S}_{[t_0,T]}(x_0)\right\} \tag{3.14}$$

Furthermore $x \in \mathcal{S}_{[t_0,T]}(x_0)$ is optimal for problem (3.13) if and only if $V(t, x(t)) \equiv g(x(T))$.

We impose the following assumptions on F and g

$$\left\{\begin{array}{ll} i) & F \text{ has closed nonempty images} \\ ii) & \forall\, x \in \mathbf{R}^n,\ \ F(\cdot, x) \text{ is measurable} \\ iii) & \exists\, k \in L^1(0,T),\ \forall\, t \in [0,T],\ \forall\, x, y \in \mathbf{R}^n, \\ & F(t,x) \subset F(t,y) + k(t)\,\|x-y\|\,B \\ iv) & \exists\, \gamma \in L^1(0,T),\ \forall\, t \in [0,T],\ \ F(t,0) \subset \gamma(t)B \\ v) & g \text{ is locally Lipschitz} \end{array}\right. \tag{3.15}$$

and observe that if the map f from Section 3.1.1 satisfies (3.7) and (3.8), then, by Chapters 1,2, assumptions (3.15) hold true for the set-valued map $F(t,x) := \overline{co}\,(f(t,x,U(t)))$. This and (3.10) yield that results of this section may be applied as well to the Mayer problem considered in Section 3.1.1.

We recall that the directional derivative of a function $\varphi : \mathbf{R}^m \mapsto \mathbf{R}$ at $x_0 \in \mathbf{R}^m$ in the direction $v \in \mathbf{R}^m$ (when it exists) is defined by

$$\frac{\partial \varphi}{\partial v}(x_0) = \lim_{h \to 0+} \frac{\varphi(x_0 + hv) - \varphi(x_0)}{h}$$

Theorem 3.1.1 *Assume (3.15). Then for every $R > 0$, there exists $L_R > 0$ such that*

i) For all $(t_0, x_0) \in [0,T] \times B_R(0)$ and every solution $x \in \mathcal{S}_{[t_0,T]}(x_0)$

$$\forall\, t \in [t_0, T],\ \ \|x(t)\| \le L_R$$

and the map $[t_0, T] \ni t \mapsto V(t, x(t))$ is absolutely continuous.

Furthermore for almost every $t \in [t_0, T]$, the directional derivative

$$\frac{\partial V}{\partial(1, x'(t))}(t, x(t))$$

does exist.

ii) For all $t \in [0,T]$, $V(t,\cdot)$ is L_R-Lipschitz on $B_R(0)$

Finally, if for all $R > 0$, there exists $c_R \geq 0$ such that

$$\text{For a.e. } t \in [0,T],\ \ \forall\, x \in B_R(0),\ \ \sup_{y\in F(t,x)} \|y\| \ \leq\ c_R \tag{3.16}$$

then for every $R > 0$, there exists $C_R > 0$ such that

$$\forall\, x\ \in\ B_R(0),\ \ \ V(\cdot, x)\ \text{ is }\ C_R - \text{Lipschitz}$$

Proof — Consider any solution $x \in \mathcal{S}_{[t_0,T]}(x_0)$ to differential inclusion (3.12). Then for almost all $t \in [t_0, T]$

$$x'(t)\ \in\ F(t, x(t))\ \subset\ F(t,0)\ +\ k(t)\,\|x(t)\|\,B$$

Thus

$$\forall\, t\ \in\ [t_0, T],\ \ \|x(t)\|\ \leq\ \|x_0\| + \int_{t_0}^{t} \gamma(s)ds + \int_{t_0}^{t} k(s)\,\|x(s)\|\,ds$$

This and Gronwall's lemma yield the first statement. Since φ is locally Lipschitz we deduce *ii)* from Filippov's theorem.

Let $x_1 \in \mathcal{S}_{[t_0,T]}(x_0)$. We claim that the map $t \mapsto V(t, x_1(t))$ is absolutely continuous. Indeed fix $t_0 \leq t_1 < t_2 \leq T$. By (3.14), there exists $x_2 \in \mathcal{S}_{[t_1,T]}(x_1(t_1))$ such that

$$V(t_2, x_2(t_2))\ \leq\ V(t_1, x_1(t_1)) + |t_2 - t_1|$$

Then from *i)* we deduce that for $i = 1, 2$

$$\begin{aligned}\|x_i(t_2) - x_i(t_1)\| \ &\leq\ \textstyle\int_{t_1}^{t_2} \gamma(s)ds + \int_{t_1}^{t_2} k(s)\,\|x_i(s)\|\,ds \\ &\leq\ \textstyle\int_{t_1}^{t_2} \gamma(s)ds + L_{\|x_0\|} \int_{t_1}^{t_2} k(s)ds\end{aligned}$$

Thus, by *ii)*, for a constant L depending only on $\|x_0\|$

$$\begin{aligned} 0\ &\leq V(t_2, x_1(t_2)) - V(t_1, x_1(t_1)) \\ &\leq V(t_2, x_1(t_2)) - V(t_2, x_2(t_2)) + |t_2 - t_1| \\ &\leq L\,\|x_1(t_2) - x_2(t_2)\| + |t_2 - t_1| \qquad (3.17) \\ &\leq L(\|x_1(t_2) - x_1(t_1)\| + \|x_2(t_2) - x_1(t_1)\|) + |t_2 - t_1| \\ &\leq 2L \textstyle\int_{t_1}^{t_2} \gamma(s)ds + 2L_{\|x_0\|} L \int_{t_1}^{t_2} k(s)ds + |t_2 - t_1| \end{aligned}$$

Recall the following characterization of absolutely continuous maps:
A function $f : [a,b] \mapsto \mathbf{R}$ is absolutely continuous if and only if

$$\begin{cases} i) & \exists\, v(f) > 0,\ \forall\, a = a_1 \le b_1 \le \ldots \le a_m \le b_m = b, \\ & \sum_{i=1}^m |f(b_i) - f(a_i)| \le v(f) \\ ii) & \forall\, \varepsilon > 0,\ \exists\, \delta > 0 \text{ such that } \forall\, a \le a_i < b_i \le b,\ i = 1, \ldots, m \\ & \text{satisfying }]a_i, b_i[\cap]a_j, b_j[= \emptyset \text{ for } i \ne j,\ \sum_{i=1}^m (b_i - a_i) \le \delta \\ & \text{we have } \sum_{i=1}^m |f(b_i) - f(a_i)| \le \varepsilon \end{cases}$$

Thus, by (3.17), the map $t \mapsto \varphi(t) := V(t, x_1(t))$ is absolutely continuous.

Fix $t \in [t_0, T]$ such that φ and x_1 are differentiable at t. Then from the local Lipschitz continuity of V with respect to the second variable

$$\lim_{h \to 0+} \frac{V(t+h, x_1(t) + hx_1'(t)) - V(t, x_1(t))}{h} = \lim_{h \to 0+} \frac{\varphi(t+h) - \varphi(t)}{h}$$

To prove the last statement of our theorem, observe that (3.16) and $i)$ imply that for all $R > 0$, there exists l_R such that every $x \in \mathcal{S}_{[t_0,T]}(x_0)$ is l_R-Lipschitz whenever $x_0 \in B_R(0)$. Fix $0 \le t_0 < t_1 \le T$, $x_0 \in B_R(0)$. By (3.14) there exists $x \in \mathcal{S}_{[t_0,T]}(x_0)$ such that $V(t_1, x(t_1)) \le V(t_0, x_0) + |t_1 - t_0|$. Then

$$\begin{aligned} & |V(t_1, x_0) - V(t_0, x_0)| \\ \le\ & |V(t_1, x(t_1)) - V(t_0, x_0)| + |V(t_1, x(t_1)) - V(t_1, x_0)| \\ \le\ & |t_1 - t_0| + L_R \|x(t_1) - x_0\| \le (L_R l_R + 1)|t_1 - t_0| \qquad \square \end{aligned}$$

3.1.3 Optimal Feedback

When the value function is directionally differentiable, it has many properties related to dynamics of the system.

Proposition 3.1.2 *Assume (3.15). If for some* $(t_0, x_0) \in [0, T[\times \mathbf{R}^n$, *$F$ is lower semicontinuous at* (t_0, x_0) *and for some* $v \in \overline{co}\,(F(t_0, x_0))$, *the directional derivative of V at* (t_0, x_0) *in the direction* $(1, v)$ *exists, then this directional derivative is nonnegative.*

Proof — Consider a solution $x(\cdot)$ to differential inclusion (3.12) satisfying $x(t_0) = x_0$, $x'(t_0) = v$. Since $V(t, \cdot)$ is Lipschitz on a neighborhood of x_0 with the Lipschitz constant independent of t and since V is nondecreasing along solutions to (3.12),

$$\lim_{h\to 0+} \frac{V(t_0 + h, x_0 + hv) - V(t_0, x_0)}{h}$$

$$= \lim_{h\to 0+} \frac{V(t_0 + h, x(t_0 + h)) - V(t_0, x_0)}{h} \geq 0 \quad \square$$

To characterize optimal solutions, we introduce the following feedback map $G : [0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ defined by

$$\forall\ (t, x) \in [0, T] \times \mathbf{R}^n,\ \ G(t, x) = \left\{ v \ \in \ F(t, x) \ \mid \ \frac{\partial V}{\partial(1, v)}(t, x) = 0 \right\}$$

(notice that the sets $G(t, x)$ may be empty.)

Theorem 3.1.3 *Assume (3.15) and let* $t_0 \in [0, T]$. *Then the following two statements are equivalent:*

i) x is a solution to the differential inclusion

$$x'(t) \ \in \ G(t, x(t)) \text{ almost everywhere in } [t_0, T] \tag{3.18}$$

ii) x is a solution to differential inclusion (3.12) defined on the time-interval $[t_0, T]$ *and for every* $t \in [t_0, T]$, $V(t, x(t)) = g(x(T))$.

Proof — Fix a solution x to (3.12) defined on $[t_0, T]$ and set $\varphi(t) = V(t, x(t))$. By Theorem 3.1.1, φ is absolutely continuous and for almost all $t \in [t_0, T]$

$$\varphi'(t) \ = \ \frac{\partial V}{\partial(1, x'(t))}(t, x(t))$$

Assume that *i*) holds true. Hence, for almost every $t \in [t_0, T]$, the set $G(t, x(t))$ is nonempty and $\varphi'(t) = 0$ almost everywhere in $[t_0, T]$. Consequently $\varphi \equiv V(T, x(T)) = g(x(T))$.

Assume next that *ii*) is verified. Then, differentiating the map $t \mapsto \varphi(t)$, we obtain that for every $t_0 < t < T$, $\varphi'(t) = 0$. Therefore for almost all $t \in [t_0, T]$, $x'(t) \in G(t, x(t))$. □

Corollary 3.1.4 *Assume (3.15). Then, a solution $x \in \mathcal{S}_{[t_0,T]}(x_0)$ is optimal for problem (3.13) if and only if it is a solution to differential inclusion (3.18) satisfying the initial condition $x(t_0) = x_0$.*

Theorem 3.1.5 *Assume (3.15) and that the images of F are convex. Then for every $t_0 \in [0, T]$ and $x_0 \in \mathbf{R}^n$, inclusion (3.18) has at least one solution $x \in \mathcal{S}_{[t_0,T]}(x_0)$.*

Proof — By Theorem 1.3.7, problem (3.13) has at least one optimal solution $\overline{x}$. Furthermore $V(t, \overline{x}(t)) \equiv g(\overline{x}(T))$. Theorem 3.1.3 ends the proof.

3.2 Maximum Principle for Free End Point Problems

3.2.1 Adjoint System

Consider a complete separable metric space $\mathcal{Z}$, real numbers $t_0 < T$ and

$$f : [t_0, T] \times \mathbf{R}^n \times \mathcal{Z} \mapsto \mathbf{R}^n$$

Let $U : [t_0, T] \rightsquigarrow \mathcal{Z}$ be a set-valued map and consider the control system

$$x' = f(t, x, u(t)), \;\; u(t) \in U(t), \;\; t \in [t_0, T] \tag{3.19}$$

Fix a state-control solution $(z, \overline{u})$ to control system (3.19). We assume that (3.7) holds true and

$$\left\{ \begin{array}{l} \text{The derivative } \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t)) \text{ exists a.e. in } [t_0, T] \\ \\ \text{For some } \varepsilon > 0, \; k \in L^1(t_0, T) \text{ and for a.e. } t \in [t_0, T] \\ \forall\, u \in U(t), \;\; f(t, \cdot, u) \text{ is } k(t) - \text{Lipschitz on } B_\varepsilon(z(t)) \end{array} \right. \tag{3.20}$$

Denote by $X(\cdot)$ the fundamental solution to the linear system

$$X'(t) = \frac{\partial f}{\partial x}(t, z(t), \overline{u}(t))X(t), \quad X(t_0) = \text{Id} \tag{3.21}$$

We recall the following property of the fundamental solution.

Proposition 3.2.1 *Let $X(t)^\star$ denote the transposed matrix. Then every solution p to the adjoint system*

$$-p' = \left(\frac{\partial f}{\partial x}(s, z(s), \overline{u}(s))\right)^\star p \tag{3.22}$$

verifies $p(t) = (X(t)^\star)^{-1} X(T)^\star p(T)$ for all $t \in [t_0, T]$.

Proof — Set $A(s) = \frac{\partial f}{\partial x}(s, z(s), \overline{u}(s))$. Then, differentiating the identity $X(t)X(t)^{-1} = \text{Id}$, we obtain that for almost all $t \in [t_0, T]$

$$0 = X'(t)X(t)^{-1} + X(t)(X^{-1})'(t)$$
$$= A(t)X(t)X(t)^{-1} + X(t)(X^{-1})'(t) = A(t) + X(t)(X^{-1})'(t)$$

Hence $(X^{-1})'(t) = -X(t)^{-1}A(t)$ and therefore $(X(\cdot)^\star)^{-1}$ is the fundamental solution to

$$Y'(t) = -A(t)^\star Y(t), \quad Y(t_0) = \text{Id}$$

Thus the solution $p(\cdot)$ to (3.22) verifies $p(t) = (X(t)^\star)^{-1} p(t_0)$ for all $t \in [t_0, T]$. So, $p(t_0) = X(T)^\star p(T)$ and the proof follows. □

Let us associate with control system (3.19) the Hamiltonian $H : [0, T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ defined by

$$H(t, x, p) = \sup_{u \in U(t)} < p, f(t, x, u) >$$

Under assumptions (3.7), (3.8), H is measurable with respect to t, locally Lipschitz with respect to (x, p) and convex with respect to the third variable p.

Denote by $R(t)$ the reachable set of (3.19) at time t from $z(t_0)$:

$$R(t) = \{x(t) \mid x \text{ is a solution to } (3.19), \ x(t_0) = z(t_0)\}$$

by $T_{R(t)}(z(t))$ the contingent cone to $R(t)$ at $z(t)$ and by

$$N^0_{R(t)}(z(t)) \;:=\; \left(T_{R(t)}(z(t))\right)^-$$

the *subnormal cone* to $R(t)$ at $z(t)$ (negative polar cone of the contingent cone to $R(t)$ at $z(t)$.)

Lemma 3.2.2 *Assume that (3.7) and (3.20) hold true. If p is a solution to adjoint system* (3.22) *such that* $p(T) \in N^0_{R(T)}(z(T))$, *then*

$$\forall\, t \;\in\; [t_0, T], \quad p(t) \;\in\; N^0_{R(t)}(z(t)) \tag{3.23}$$

and the following maximum principle *holds true:*

$$\langle p(t), f(t, z(t), \overline{u}(t))\rangle \;=\; H(t, z(t), p(t)) \text{ a.e. in } [t_0, T] \tag{3.24}$$

Proof — From Theorem 1.3.9 we deduce that for all $v \in T_{R(t)}(z(t))$, $X(T)X(t)^{-1}v \in T_{R(T)}(z(T))$. Thus $\langle p(T), X(T)X(t)^{-1}v\rangle \leq 0$ and, using Proposition 3.2.1, we obtain

$$< p(t), v > = \left\langle (X(t)^\star)^{-1} X(T)^\star p(T),\; v\right\rangle \;\leq\; 0$$

If w solves the linearized system (2.6) and $t_0 = 0$, $w(0) = 0$, then $w(T) \in T_{R(T)}(z(T))$ and $w(T) = \int_0^T X(T)X(s)^{-1}v(s)ds$. Thus

$$0 \geq < p(T), w(T) > = < \int_0^T < (X(s)^\star)^{-1} X(T)^\star p(T), v(s) > ds$$

Thus $< \int_0^T < p(s), v(s)ds > \leq 0$ for all integrable selection $v(s) \in T_{\overline{co}f(s,z(s),U(s))}(z'(s))$. Since

$$0 \in \overline{co}f(s, z(s), U(s)) - z'(s) \subset T_{\overline{co}f(s,z(s),U(s))}(f(s, z(s), \overline{u}(s)))$$

we deduce, using results on measurable set-valued maps from Chapter 1 that

$$0 \geq < \int_0^T \sup_{v \in \overline{co}f(s,z(s),U(s)) - z'(s)} < p(s), v > ds \geq 0$$

Which yields (3.24) □

3.2.2 Maximum Principle

Let $g : \mathbf{R}^n \mapsto \mathbf{R}$ be a differentiable function and $x_0 \in \mathbf{R}^n$ be given. Consider the problem

$$\text{minimize } g(x(T))$$

over all solutions to control system (3.19) satisfying $x(t_0) = x_0$.

Theorem 3.2.3 *If a state-control solution $(z, \overline{u})$ solves the above problem and (3.7), (3.20) hold true, then the solution p to adjoint system* (3.22) *such that*

$$p(T) = -\nabla g(z(T))$$

satisfies (3.23) *and maximum principle* (3.24).

Remark — The map $p(\cdot)$ in the above theorem is called *co-state* or *adjoint variable* associated with $(z, \overline{u})$. □

Proof — Since z is optimal, we have

$$\min_{y \in R(T)} g(y) = g(z(T))$$

Let $v \in T_{R(T)}(z(T))$ and $h_n \to 0+, v_n \to v$ be such that $z(T)+h_n v_n \in R(T)$. Then $g(z(T)+h_n v_n)-g(z(T)) \geq 0$. Dividing by h_n and taking the limit we obtain $< \nabla g(z(T)), v > \geq 0$. Hence $-\nabla g(z(T)) \in N^0_{R(T)}(z(T))$. Lemma 3.2.2 yields the conclusion. □

The above theorem can be applied to derive necessary optimality conditions for various minimization problems:

Exercises

1. **Minimization with respect to both End Points.** Consider a differentiable function $\varphi : \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ and the problem

$$\text{minimize } \varphi(x(t_0), x(T))$$

over all solutions to control system (3.19). Assume that a state-control solution $(z, \overline{u})$ is optimal and (3.7), (3.20) hold true.

Show that the solution p to adjoint system (3.22) such that

$$p(T) \;=\; -\frac{\partial \varphi}{\partial x_2}(z(t_0), z(T)) \tag{3.25}$$

satisfies maximum principle (3.24) and

$$p(t_0) \;=\; \frac{\partial \varphi}{\partial x_1}(z(t_0), z(T))$$

2. **Problem with Initial Point Constraints.** Consider the same problem as in Exercise 1 with an additional restriction

$$x(t_0) \;\in\; K \tag{3.26}$$

where K is a given subset of $\mathbf{R}^n$. Assume that a state-control solution $(z, \overline{u})$ is optimal and (3.7), (3.20) hold true.

Show that the solution p to adjoint system (3.22), (3.25) satisfies maximum principle (3.24) and transversality condition

$$p(t_0) \;\in\; \frac{\partial \varphi}{\partial x_1}(z(t_0), z(T)) \;+\; N_K^0(z(t_0)) \tag{3.27}$$

3. **Bolza Problem.** Let φ, K be as in Exercise 2. Consider a function $L : [t_0, T] \times \mathbf{R}^n \times \mathcal{Z} \mapsto \mathbf{R}$ and the optimization problem

$$\text{minimize} \left\{ \varphi(x(t_0), x(T)) \;+\; \int_{t_0}^{T} L(t, x(t), u(t))dt \right\}$$

over all state-control solutions (x, u) to system (3.19) satisfying (3.26). Assume that a state-control solution $(z, \overline{u})$ is optimal and that the map (f, L) satisfy assumptions (3.7), (3.20).

Show that the solution p to the equation

$$-p' \;=\; \left(\frac{\partial f}{\partial x}(t, z(t), \overline{u}(t))\right)^{\star} p \;-\; \frac{\partial L}{\partial x}(t, z(t), \overline{u}(t))$$

satisfying (3.25) verifies transversality condition (3.27) and the maximum principle:

$$\begin{cases} \langle p(t), f(t, z(t), \overline{u}(t))\rangle \;-\; L(t, z(t), \overline{u}(t)) \\ \\ = \;\max_{u\in U(t)}(\langle p(t), f(t, z(t), u)\rangle \;-\; L(t, z(t), u)) \;\text{ a.e. in } [t_0, T] \end{cases}$$

4. **Free Time Mayer Problem.** Let g be a function as in Section 5.2.2 and consider the problem

$$\text{minimize } g(x(t))$$

over all solutions to control system (3.19) satisfying $x(t_0) = x_0$. Let $(z, \overline{u})$ be an optimal state-control solution, $T > 0$ be the corresponding optimal time and assume that (3.7), (3.20) hold true. Let p denote the corresponding co-state. Assume in addition that $f(\cdot, x, \cdot)$ is continuous, $U(\cdot)$ is lower semicontinuous and for all $x \in \mathrm{R}^n$, the set $f(T, x, U(T))$ is bounded. Show that

$$\max_{u \in U(T)} < p(T),\ f(T, z(T), u) > = 0 \quad \Box$$

3.3 Necessary and Sufficient Conditions for Optimality

We begin this section with a sufficient condition for optimality involving the superdifferential of the value function.

3.3.1 Sufficient Conditions

Theorem 3.3.1 *Assume that (3.7), (3.8) hold true and let $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$. Consider a solution $z : [t_0, T] \mapsto \mathbf{R}^n$ to control system (3.4) with $z(t_0) = x_0$ and let $\overline{u}$ be a corresponding control. If for almost every $t \in [t_0, T]$, there exists $p(t) \in \mathbf{R}^n$ such that*

$$(\langle p(t), z'(t) \rangle,\ -p(t)) \in \partial_+ V(t, z(t)) \tag{3.28}$$

then z is optimal for problem (3.6).

Proof— By Theorem 3.1.1 the map $\psi(t) := V(t, z(t))$ is absolutely continuous. Let $t \in [t_0, T]$ be such that the derivatives $\psi'(t)$ and $z'(t)$ do exist and (3.28) holds true. Then, using Theorem 3.1.1 and Proposition 1.1.7

$$0 \;=\; \langle (< p(t), z'(t) >, -p(t)), (1, z'(t)) \rangle \geq D_\downarrow V(t, z(t))(1, z'(t))$$

$$\geq\; \limsup_{h \to 0+} \frac{V(t+h, z(t+h)) - V(t, z(t))}{h} \;=\; \psi'(t)$$

This yields that ψ is nonincreasing. Since the value function is also nondecreasing along solutions to control system (3.4), we deduce that the map $t \mapsto V(t, z(t))$ is constant. So z is optimal. □

3.3.2 Necessary and Sufficient Conditions

Theorem 3.3.2 *Assume (3.7), (3.8), that f is differentiable with respect to x and g is differentiable. A state-control solution $(z, \overline{u})$ to control system (3.4) with $z(t_0) = x_0$ is optimal for problem (3.6) if and only if the solution $p : [t_0, T] \mapsto \mathbf{R}^n$ to the adjoint system*

$$-p'(t) \;=\; \left(\frac{\partial f}{\partial x}(t, z(t), \overline{u}(t))\right)^{\star} p(t), \;\; p(T) \;=\; -\nabla g(z(T)) \qquad (3.29)$$

satisfies the maximum principle

$$< p(t), f(t, z(t), \overline{u}(t)) > \;=\; H(t, z(t), p(t)) \;\text{ a.e. in } [t_0, T] \qquad (3.30)$$

and the generalized transversality conditions

$$(H(t, z(t), p(t)), -p(t)) \;\in\; \partial_+ V(t, z(t)) \;\text{ a.e. in } [t_0, T] \qquad (3.31)$$

$$-p(t) \;\in\; \partial_+ V_x(t, z(t)) \;\text{ for every } t \in [t_0, T] \qquad (3.32)$$

where $\partial_+ V_x(t, z(t))$ denotes the superdifferential of $V(t, \cdot)$ at $z(t)$.

Furthermore, if V is semiconcave and H is continuous, then (3.31) holds true everywhere in $[t_0, T]$.

Proof — Sufficiency is a straightforward consequence of Theorem 3.3.1 and (3.30), (3.31). The fact that (3.29) and (3.30) are necessary follows from Theorem 3.2.3.

Fix $t \in [t_0, T]$, $v \in \mathbf{R}^n$ and consider the solution $w(\cdot)$ to the linear system

$$\begin{cases} w'(s) &= \frac{\partial f}{\partial x}(s, z(s), \overline{u}(s)) w(s), \quad s \in [t, T] \\ w(t) &= v \end{cases}$$

Then $w(T) = X(T)X(t)^{-1}v$, where $X(\cdot)$ is the fundamental solution to (3.21). For every $h > 0$, let x_h be the solution to the differential equation

$$\begin{cases} x'(s) &= f(s, x(s), \overline{u}(s)), \quad s \in [t, T] \\ x(t) &= z(t) + hv \end{cases}$$

From the variational equation we know that the difference quotients $(x_h - z)/h$ converge uniformly to w.

To prove (3.32) it is enough to observe that

$$\langle -p(t), v\rangle = \langle (X(t)^\star)^{-1} X(T)^\star \nabla\varphi(z(T)), v\rangle = \langle \nabla\varphi(z(T)), w(T)\rangle$$

$$\geq \limsup_{h\to 0+} (V(t, z(t) + hv) - V(t, z(t)))/h$$

To prove the necessity of (3.31) fix $t \in [0, T[$ such that $z'(t) = f(t, z(t), \overline{u}(t))$ and equality (3.30) holds true, $v \in \mathbf{R}^n$, $\alpha \in \mathbf{R}$. Then from (3.30), using that $V(t, \cdot)$ is locally Lipschitz, that V is nondecreasing along solutions to (4.3) and is constant along z, we deduce

$$\limsup_{h\to 0+} \left(V(t + \alpha h, z(t) + h(\alpha z'(t) + v)) - V(t, z(t))\right)/h$$

$$= \limsup_{h\to 0+} \left(V(t + \alpha h, z(t + \alpha h) + hw(t + \alpha h)) - V(t, z(t))\right)/h$$

$$= \limsup_{h\to 0+} \left(V(t + \alpha h, x_h(t + \alpha h)) - V(t, z(t))\right)/h$$

$$\leq \limsup_{h\to 0+} \left(\varphi(x_h(T)) - \varphi(z(T))\right)/h = \langle \nabla\varphi(z(T)), w(T)\rangle$$

$$= \langle \nabla\varphi(z(T)), X(T)X(t)^{-1}v\rangle = \left\langle (X(t)^\star)^{-1} X(T)^\star \nabla\varphi(z(T)), v\right\rangle$$

$$= \langle -p(t), v\rangle = \langle -p(t), -\alpha z'(t)\rangle + \langle -p(t), \alpha z'(t) + v\rangle$$

$$= \alpha H(t, z(t), p(t)) + \langle -p(t), \alpha z'(t) + v\rangle$$

Hence we deduce that for every $\alpha \in \mathbf{R}$ and $v_1 \in \mathbf{R}^n$

$$D_\downarrow V(t, z(t))(\alpha, v_1) \leq \alpha H(t, z(t), p(t)) + \langle -p(t), v_1\rangle$$

and (3.31) follows from Proposition 1.1.7.

When V is semiconcave, then, from Theorem 1.1.11, its superdifferential is upper semicontinuous. Thus the last statement results from (3.31) and continuity of $H(\cdot)$, $p(\cdot)$, $z(\cdot)$. □

For autonomous systems, that is with f and U independent of t, the Hamiltonian is constant along any optimal state/co-state pair (z,p). Indeed, recalling (3.30), (3.29), for almost all $s \in [t_0,T]$ we obtain

$$\begin{cases} H(z(t),p(t)) - H(z(s),p(s)) \\ \geq \langle p(t), f(z(t),\overline{u}(s)) - f(z(s),\overline{u}(s))\rangle + \langle p(t)-p(s), f(z(s),\overline{u}(s))\rangle \\ = o(|t-s|) \end{cases}$$

for all $t \in [t_0,T]$. Since the above argument is symmetric,

$$|H(z(t),p(t)) - H(z(s),p(s))| \leq o(|t-s|)$$

Using that $t \mapsto H(z(t),p(t))$ is absolutely continuous, we deduce that $H(z(t),p(t))$ is constant.

When the Hamiltonian $H(t,\cdot,\cdot)$ is differentiable at $(z(t),p(t))$ for all $t \in [t_0,T]$, then z and the co-state p satisfy the *Hamiltonian system*

$$\begin{cases} z'(t) = \frac{\partial H}{\partial p}(t,z(t),p(t)) \\ p'(t) = -\frac{\partial H}{\partial x}(t,z(t),p(t)) \text{ a.e. in } [t_0,T] \end{cases}$$

This follows from Theorem 3.3.2 and

Proposition 3.3.3 *Let $(t,z,\overline{p}) \in [t_0,T] \times \mathbf{R}^n \times \mathbf{R}^n$ and $\overline{u} \in U(t)$ be such that $\langle \overline{p}, f(t,z,\overline{u})\rangle = H(t,z,\overline{p})$. Then*

i) If $H(t,\cdot,\overline{p})$ is differentiable at z, then

$$\frac{\partial H}{\partial x}(t,z,\overline{p}) = \left(\frac{\partial f}{\partial x}(t,z,\overline{u})\right)^{\star}\overline{p}$$

ii) If $H(t,z,\cdot)$ is differentiable at $\overline{p}$, then

$$\frac{\partial H}{\partial p}(t,z,\overline{p}) = f(t,z,\overline{u})$$

Proof — It is enough to observe that for every $v \in \mathbf{R}^n$

$$\lim_{h\to 0+} \frac{H(t,z+hv,\overline{p}) - H(t,z,\overline{p})}{h} \geq \lim_{h\to 0+} \frac{\langle \overline{p}, f(t,z+hv,\overline{u}) - f(t,z,\overline{u})\rangle}{h}$$

$$= \left\langle \overline{p}, \frac{\partial f}{\partial x}(t,z,\overline{u})v\right\rangle = \left\langle \left(\frac{\partial f}{\partial x}(t,z,\overline{u})\right)^{\star}\overline{p}, v\right\rangle$$

and

$$\lim_{h\to 0+} \frac{H(t,z,\overline{p}+hv)-H(t,z,\overline{p})}{h} \geq \lim_{h\to 0+} \frac{\langle \overline{p}+hv,\, f(t,z,\overline{u})\rangle - \langle \overline{p},\, f(t,z,\overline{u})\rangle}{h}$$

$$= \langle v, f(t,z,\overline{u})\rangle \quad \square$$

3.3.3 Co-state and Superdifferentials of Value

Proposition 3.3.4 *Assume (3.15) and let z be an optimal solution to problem (3.13). Then for almost every $t \in [t_0, T]$,*

$$\forall\, (p_t, p_x) \in \partial_+ V(t, z(t)), \quad -p_t \; - \; \langle p_x, z'(t)\rangle \; = \; 0 \tag{3.33}$$

Furthermore, if F is lower semicontinuous, then for almost every $t \in [t_0, T]$,

$$\forall\, (p_t, p_x) \in \partial_+ V(t, z(t)), \quad -p_t \; + \; H(t, z(t), -p_x) \; = \; 0 \tag{3.34}$$

If (3.16) holds true and F is continuous, then (3.34) is satisfied for all $t_0 < t < T$.

Proof — Let $t \in]t_0, T[$ be such that $z'(t) \in F(t, z(t))$. Then for all $(p_t, p_x) \in \partial_+ V(t, z(t))$

$$0 \geq \limsup_{s\to t+} \frac{V(s,z(s)) - V(t,z(t)) - p_t(s-t) - \langle p_x, z(s) - z(t)\rangle}{|s-t| + \|z(s) - z(t)\|}$$

Since $V(\cdot, z(\cdot))$ is constant, the above estimate yields

$$0 \; \geq \; \lim_{s\to t+} \frac{-p_t(s-t) - \langle p_x, z(s) - z(t)\rangle}{|s-t|} \; = \; -p_t \; - \; \langle p_x, z'(t)\rangle$$

By the same arguments, taking $s \to t-$ we derive $0 \geq p_t + \langle p_x, z'(t)\rangle$ and so (3.33) is proved. Assume next that F is lower semicontinuous and fix $v \in F(t_0, x_0), t_0 < T$. Consider $x \in \mathcal{S}_{[t,T]}(x_0)$ satisfying $x'(t_0) = v$. Since for all small $h > 0$, $V(t_0, x_0) \leq V(t_0 + h, x(t_0 + h))$, we deduce that $D_\downarrow V(t_0, x_0)(1, v) \geq 0$. Consequently, for all $(p_t, p_x) \in \partial_+ V(t_0, x_0)$, $p_t + < p_x, v > \geq 0$. But this yields $-p_t + H(t_0, x_0, -p_x) \leq 0$. From the last inequality and (3.33) we get (3.34).

If (3.16) holds true and F is continuous, then z is Lipschitz. Fix $t_0 < t < T$. Then for a sequence $h_n \to 0+$ and some $v \in \overline{co}\,(F(t, z(t)))$

$$\lim_{n\to\infty} \frac{z(t-h_n) - z(t)}{h_n} = -v$$

Fix $(p_t, p_x) \in \partial_+ V(t, z(t))$. Applying exactly the same arguments as before, we deduce that $p_t + \langle p_x, v\rangle \leq 0$ yielding equality (3.34) at t. □

Theorem 3.3.5 *Assume (3.7), (3.8), that f is differentiable with respect to x and g is differentiable. Suppose further that $V(t_0, \cdot)$ is differentiable at x_0 and let $(z, \overline{u})$ be an optimal state-control solution to problem (3.6). Then the co-state $p : [t_0, T] \mapsto \mathbf{R}^n$ corresponding to $(z, \overline{u})$ and given by Theorem 3.3.2 verifies*

$$\{-p(t)\} = \partial_+ V_x(t, z(t)) \text{ for all } t \in [t_0, T]$$

Hence if $V(t, \cdot)$ is semiconcave, then $\frac{\partial V}{\partial x}(t, z(t)) = -p(t)$.

Remark — In Section 4 below, we show that under some additional regularity assumptions on f, V is semiconcave. □

Proof — We already know from Theorem 3.3.2 that

$$-p(t) \in \partial_+ V_x(t, z(t)) \text{ for all } t \in [t_0, T]$$

Thus $p(t_0) = -\frac{\partial V}{\partial x}(t_0, x_0)$.

Fix $v \in \mathbf{R}^n$ and let w, x_h have the same meaning as in the proof of Theorem 3.3.2 with t replaced by t_0. Then, since V is nondecreasing along solutions to control system (4.3) and constant along z,

$$\begin{aligned}\langle -p(t_0), v\rangle &= \left\langle \frac{\partial V}{\partial x}(t_0, x_0), v\right\rangle = \lim_{h\to 0+} \frac{V(t_0, x_0 + hv) - V(t_0, x_0)}{h} \\ &\leq \limsup_{h\to 0+} \frac{V(t, x_h(t)) - V(t, z(t))}{h} \\ &= \limsup_{h\to 0+} \frac{V(t, z(t) + hw(t)) - V(t, z(t))}{h}\end{aligned}$$

for all $t \in [t_0, T]$. Hence for every $q \in \partial_+ V_x(t, z(t))$ we have

$$\langle -p(t_0), v\rangle \;\leq\; \langle q, w(t)\rangle \;=\; \langle q, X(t)v\rangle \;=\; \langle X(t)^\star q, v\rangle$$

where X denotes the fundamental solution to (3.21).

Since $v \in \mathbf{R}^n$ is arbitrary, $p(t_0) = -X(t)^\star q$. On the other hand, $p(\cdot)$ being a solution to (3.29), we know that $p(t_0) = X(t)^\star p(t)$ and we deduce that $-p(t) = q$. This yields that $\partial_+ V_x(t, z(t))$ is a singleton and ends the proof. □

3.3.4 Hamiltonian System

Whenever H happens to be more regular we can prove the following theorem concerning optimal design. For every $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ define

$$\partial^\star V_x(t_0, x_0) \;=\; \partial^\star W(x_0)$$

where W is given by $W(x) = V(t_0, x)$.

Theorem 3.3.6 *Assume (3.7), (3.8), that f is differentiable with respect to x, g is differentiable and that $H(t,\cdot,\cdot)$ is differentiable*[1] *on $\mathbf{R}^n \times (\mathbf{R}^n \backslash \{0\})$ for almost every $t \in [t_0, T]$.*

Further assume that the sets $f(t, x, U(t))$ are convex and compact and for every $R > 0$, there exists a nonnegative integrable function $l_R \in L^1(0, T)$ such that for all $x, y \in RB$ and $p, q \in RB \backslash \frac{1}{R}B$

$$\left\{\begin{array}{l} \left\|\frac{\partial H}{\partial x}(t,x,p) - \frac{\partial H}{\partial x}(t,y,q)\right\| \;+\; \left\|\frac{\partial H}{\partial p}(t,x,p) - \frac{\partial H}{\partial p}(t,y,q)\right\| \\ \qquad\qquad \leq\; l_R(t)(\|x-y\| + \|p-q\|) \end{array}\right. \tag{3.35}$$

Let $(t_0, x_0) \in [t_0, T] \times \mathbf{R}^n$ and $p_0 \neq 0$ be such that $-p_0 \in \partial^\star V_x(t_0, x_0)$. Then the Hamiltonian system

$$\left\{\begin{array}{lllll} x'(t) & = & \frac{\partial H}{\partial p}(t, x(t), p(t)), & & x(t_0) \;=\; x_0 \\ p'(t) & = & -\frac{\partial H}{\partial x}(t, x(t), p(t)), & & p(t_0) \;=\; p_0 \\ p(t) & \neq & 0 \text{ for all } t \in [t_0, T] & & \end{array}\right. \tag{3.36}$$

[1] It is well known that $H(t, x, \cdot)$ is not differentiable at zero when $f(t, x, U(t))$ is not a singleton. For this reason we exclude zero in our differentiability assumptions.

has a unique solution $(z(\cdot), \overline{p}(\cdot))$ *defined on* $[t_0, T]$. *Moreover* $z(\cdot)$ *is optimal for problem (3.6).*

Consequently, problem (3.6) has at least as many optimal solutions as there are elements in the set $\partial^\star V_x(t_0, x_0)\backslash\{0\}$.

Furthermore, if $\nabla g(\cdot)$ *is continuous at* $z(T)$, *then* $\overline{p}(\cdot)$ *is the co-state corresponding to* $z(\cdot)$ *given by Theorem 3.3.2.*

Remark — A typical example of a nonlinear control system with closed convex images is the affine system:

$$x' = f(x) + \sum_{i=1}^{m} u_i g_i(x), \quad u_i \in [a_i, b_i]$$

where f and g_i are maps from $\mathbf{R}^n$ to itself and $a_i \leq b_i$ are given numbers. □

Proof — From the very definition of $\partial^\star V_x(t_0, x_0)$ it follows that there exists a sequence x_k converging to x_0 such that $V(t_0, \cdot)$ is differentiable at x_k and

$$-p_0 = \lim_{k\to\infty} \frac{\partial V}{\partial x}(t_0, x_k)$$

Let (z_k, u_k) be an optimal state-control solution for problem (3.6) with x_0 replaced by x_k. By Theorem 3.3.2 (applied with x_0 replaced by x_k), for every k there exists an absolutely continuous function $\overline{p}_k : [t_0, T] \mapsto \mathbf{R}^n$ such that

$$\begin{cases} -\overline{p}_k'(t) = \left(\frac{\partial f}{\partial x}(t, z_k(t), u_k(t))\right)^\star \overline{p}_k(t), \text{ a.e. in } [t_0, T] \\ -\overline{p}_k(t_0) = \frac{\partial V}{\partial x}(t_0, x_k), \quad \overline{p}_k(T) = -\nabla g(z_k(T)) \end{cases} \tag{3.37}$$

Therefore, $p_k(t) \neq 0$ for all $t \in [t_0, T]$ and sufficiently large k. By Proposition 3.3.3, for every $t \in [t_0, T]$,

$$\begin{cases} z_k(t) = x_0 + \displaystyle\int_{t_0}^{t} \frac{\partial H}{\partial p}(s, z_k(s), \overline{p}_k(s))ds \\ \overline{p}_k(t) = p_0 - \displaystyle\int_{t_0}^{t} \frac{\partial H}{\partial x}(s, z_k(s), \overline{p}_k(s))ds \end{cases} \tag{3.38}$$

Recalling assumptions (3.8) and Theorem 3.1.1, we conclude that z_k, $k = 1, \dots$ are equicontinuous and equibounded. Furthermore, from (3.8) and (3.37) it follows that $\overline{p}_k$ are also equicontinuous and equibounded, because the maps $t \mapsto \frac{\partial f}{\partial x}(t, z_k(t), u_k(t))$ are integrably bounded on $[t_0, T]$. So, taking a subsequence and keeping the same notation, we may assume that $(z_k, \overline{p}_k)$ converge uniformly to some $(z, \overline{p})$ and $\frac{\partial f}{\partial x}(\cdot, z_k(\cdot), u_k(\cdot))$ converge weakly in $L^1(t_0, T; \mathbf{R}^n \times \mathbf{R}^n)$ to some $A(\cdot)$. In particular $\overline{p}(t_0) = p_0 \neq 0$ and $\overline{p}$ solves the linear system

$$-\overline{p}'(t) \;=\; A(t)^{\star}\overline{p}(t), \;\; \text{almost everywhere in } [t_0, T]$$

Thus $\overline{p}(t) \neq 0$ for all $t \in [t_0, T]$. Fix $R > 1$ so that

$$\forall\, s \in [t_0, T], \;\; \frac{2}{R} \;\leq\; \|\overline{p}(s)\| \;\leq\; \frac{R}{2}$$

Then, for all sufficiently large k and all $s \in [t_0, T]$, we have

$$\frac{1}{R} \;\leq\; \|p_k(s)\| \;\leq\; R$$

So, using (3.35) and taking the limit in (3.38), we deduce $(z, \overline{p})$ is a solution to Hamiltonian system (3.36).

Since $\overline{p}$ never vanishes, assumption (3.35) implies that $(z, \overline{p})$ is the only solution to (3.36). On the other hand

$$V(t_0, x_0) \;=\; \lim_{k \to \infty} V(t_0, z_k(t_0)) \;=\; \lim_{k \to \infty} g(z_k(T)) \;=\; g(z(T))$$

and therefore z is optimal for problem (3.6).

If ∇g is continuous at $z(T)$, then from (3.37) it follows that $\overline{p}(T) = -\nabla g(z(T)) \neq 0$. Let p_1 be a co-state corresponding to the optimal solution z given by Theorem 3.3.2. Then $p_1(t) \neq 0$ for all $t \in [t_0, T]$ and, by Proposition 3.3.3 it solves the problem

$$\begin{cases} -p'(t) \;=\; \frac{\partial H}{\partial x}(t, z(t), p(t)), \;\; \text{a.e. in } [t_0, T] \\ \\ p(T) \;=\; -\nabla g(z(T)) \end{cases}$$

Since $\overline{p}$ is also a solution to this system, $p_1 = \overline{p}$ by uniqueness. □

3.3.5 Uniqueness of Optimal Solution and Differentiability of Value Function

Theorem 3.3.6 yields that if $\partial^\star V_x(t_0, x_0)\backslash\{0\}$ is not a singleton, then optimal solution to (3.6) is not unique. We prove a similar statement under less restrictive regularity assumptions on $H(t, x, \cdot)$.

Theorem 3.3.7 *Assume (3.7) (3.8), that g is continuously differentiable, f is differentiable with respect to x, $f(t, x, U(t))$ are convex and compact and that for every $t \in [0, T]$, $\frac{\partial H}{\partial x}(t, \cdot, \cdot)$ is continuous.*

Further assume that for every $R > 0$, there exists a nonnegative integrable function $l_R \in L^1(0, T)$ such that

$$\forall\, x, y, p \in RB, \ \left\| \frac{\partial H}{\partial x}(t, x, p) - \frac{\partial H}{\partial x}(t, y, p) \right\| \le l_R(t)\, \|x - y\|$$

If problem (3.6) has a unique optimal solution z, then for all $t \in [t_0, T]$, $\partial^\star V_x(t, z(t))$ is a singleton and, consequently, $V(t, \cdot)$ is differentiable at $z(t)$.

Proof — Observe that for every $t \in [t_0, T]$, problem (3.6) has a unique optimal solution with (t_0, x_0) replaced by $(t, z(t))$. For this reason we prove the result only for $V(t_0, \cdot)$.

By Proposition 1.1.9, it suffices to show that $\partial^\star V_x(t_0, x_0)$ is a singleton. Let $p_1, p_2 \in \partial^\star V_x(t_0, x_0)$ and consider sequences $\{x_k^1\}$ and $\{x_k^2\}$ converging to x_0, such that

$$\lim_{k\to+\infty} \frac{\partial V}{\partial x}(t_0, x_k^i) \ = \ p_i, \ \ i = 1, 2$$

Let z_k^i be optimal solutions to problem (3.6) with x_0 replaced by x_k^i, $i = 1, 2$ and denote by p_k^i the corresponding co-states given by Theorem 3.3.2. Then, by Proposition 3.3.3,

$$\begin{cases} (p_k^i)'(t) & = & -\frac{\partial H}{\partial x}(t, z_k^i(t), p_k^i(t)) \ \text{a.e. in } [t_0, T] \\ p_k^i(T) & = & -\nabla g(z_k^i(T)), \ p_k^i(t_0) \ = \ -\frac{\partial V}{\partial x}(t_0, x_k^i) \end{cases}$$

By Theorem 3.1.1, z_k^i are bounded, equicontinuous and $V(T, z_k^i(T)) = g(z_k^i(T))$. Since the solution to (3.6) is unique by our assumptions, we deduce that z_k^i converge uniformly to z for $i = 1, 2$. Taking subsequences and keeping the same notations, we may assume that p_k^i converge uniformly to the unique solution p to the system

$$p'(t) \;=\; -\frac{\partial H}{\partial x}(t, z(t), p(t)), \quad p(T) \;=\; -\nabla g(z(T))$$

Thus, $p_1 = p(t_0) = p_2$. □

Theorem 3.3.8 *We posit all hypothezis of Theorem 3.3.6 and we assume that g is continuously differentiable. Then $V(t_0, \cdot)$ is differentiable at x_0 with the derivative different from zero if and only if there exists a unique optimal solution z to problem (3.6) satisfying $\nabla g(z(T)) \neq 0$.*

Proof — Assume that $\frac{\partial V}{\partial x}(t_0, x_0) \neq 0$. Let z be optimal for problem (3.6). By Theorem 3.3.2, $\nabla g(z(T)) \neq 0$. By Proposition 3.3.3, every optimal state/co-state pair solves Hamiltonian system (3.36) with $p_0 = -\frac{\partial V}{\partial x}(t_0, x_0)$. This and Theorem 3.3.6 yield uniqueness of optimal solution.

Conversely, assume that (3.6) has a unique optimal solution z and $\nabla g(z(T)) \neq 0$. By Theorem 3.3.7, $V(t_0, \cdot)$ is differentiable at x_0. Theorem 3.3.2 implies that $\frac{\partial V}{\partial x}(t_0, x_0) \neq 0$. □

3.4 Semiconcavity of Value Function

We provide next a sufficient condition for semi-concavity of the value function on $[0, T] \times \mathbf{R}^n$. Throughout the whole section we assume the following

$$\left\{\begin{array}{l} \exists\, \omega : \mathbf{R}_+ \times \mathbf{R}_+ \mapsto \mathbf{R}_+ \text{ such that (1.3) holds true and} \\ \\ \forall\, \lambda \in [0,1],\ R > 0,\ x_0, x_1 \in B_R(0),\ t \in [0,T],\ u \in U(t) \\ \|\lambda f(t, x_0, u) + (1-\lambda) f(t, x_1, u) - f(t, x_\lambda, u)\| \\ \leq\ \lambda(1-\lambda)\,\|x_1 - x_0\|\,\omega(R, \|x_1 - x_0\|), \\ \text{where } x_\lambda = \lambda x_0 + (1-\lambda) x_1 \\ \\ g : \mathbf{R}^n \mapsto \mathbf{R} \text{ is semiconcave} \end{array}\right. \tag{3.39}$$

Remark — Assumptions (3.39) hold true in particular when g is continuously differentiable and f is continuously differentiable with respect to x uniformly in (t, u). More precisely, if we assume that there exists a function $\omega : \mathbf{R}_+ \times \mathbf{R}_+ \mapsto \mathbf{R}_+$ satisfying (1.3) such that

$$\left\| \frac{\partial f}{\partial x}(t, x_1, u) - \frac{\partial f}{\partial x}(t, x_2, u) \right\| \leq \omega(R, \|x_1 - x_2\|)$$

for all $t \in [0, T]$, $u \in U(t)$ and $x_1, x_2 \in B_R(0)$.

2) Vice versa, Proposition 1.1.13 implies that, if f satisfies (3.39), then f is continuously differentiable with respect to x. □

Theorem 3.4.1 *Assume (3.7), (3.8) and (3.39). Then there exists* $\overline{\omega} : \mathbf{R}_+ \times \mathbf{R}_+ \mapsto \mathbf{R}_+$ *satisfying (1.3) such that for all* $t \in [0, T]$, $\lambda \in [0, 1]$, $R > 0$

$$\forall x_0, x_1 \in B_R(0),\ \lambda V(t, x_1) + (1 - \lambda)V(t, x_0) - V(t, \lambda x_1 + (1 - \lambda)x_0)$$

$$\leq \lambda(1 - \lambda) \|x_1 - x_0\| \overline{\omega}(R, \|x_1 - x_0\|)$$

Consequently for every $t \in [0, T]$, $V(t, \cdot)$ *is semiconcave.*

Proof — For every $t \in [0, T]$ and control $u(s) \in U(s)$ (admissible control), we denote by $y(\cdot; t, x, u)$ the solution to the system

$$\begin{cases} y'(s) &= f(s, y(s), u(s)),\quad s \in [t, T] \\ y(t) &= x \end{cases}$$

By Theorem 3.1.1 for every $R > 0$ there exists L_R such that

$$\forall\, x \in B_R(0),\ \ \forall\, s \in [t, T],\ \ \|y(s; t, x, u)\| \leq L_R \tag{3.40}$$

Moreover, by the Gronwall lemma, for all $t \in [0, T]$, $s \in [t, T]$, $x_0, x_1 \in \mathbf{R}^n$ and all admissible control $u(\cdot)$, we have

$$\|y(s; t, x_1, u) - y(s; t, x_0, u)\| \leq e^{\int_t^s k(\tau)d\tau} \|x_1 - x_0\| \tag{3.41}$$

Step 1. We claim that there exists $\omega_1 : \mathbf{R}_+ \times \mathbf{R}_+ \mapsto \mathbf{R}_+$ satisfying (1.3) such that for all $0 \leq t \leq s \leq T$, $R > 0$, $x_0, x_1 \in B_R(0)$, $\lambda \in [0, 1]$

and admissible control $u(\cdot)$, we have

$$\|\lambda y(s;t,x_1,u)+(1-\lambda)y(s;t,x_0,u)-y(s;t,\lambda x_0+(1-\lambda)x_1,u)\|$$

$$\leq \lambda(1-\lambda)\,\|x_1-x_0\|\,\omega_1(R,\|x_1-x_0\|)$$

Indeed set $x_\lambda=\lambda x_0+(1-\lambda)x_1$ and define

$$y_\lambda(\tau)=\lambda y(\tau;t,x_1,u)+(1-\lambda)y(\tau;t,x_0,u)-y(\tau;t,x_\lambda,u)$$

Then $y_\lambda(t)=0$ and

$$\begin{aligned} y'_\lambda(\tau) &= \lambda f(\tau,y(\tau;t,x_1,u),u(\tau))+ \\ &+(1-\lambda)f(\tau,y(\tau;t,x_0,u),u(\tau))-f(\tau,y(\tau;t,x_\lambda,u),u(\tau)) \end{aligned}$$

From assumptions (3.8), (3.39) we obtain

$$\begin{aligned} \|y'_\lambda(\tau)\| &\leq k(\tau)\,\|y_\lambda(\tau)\|+\lambda(1-\lambda)\,\|y(\tau;t,x_1,u)-y(\tau;t,x_0,u)\|\times \\ &\times\,\omega(L_R,\|y(\tau;t,x_1,u)-y(\tau;t,x_0,u)\|) \end{aligned}$$

Our claim results from (3.41) and the Gronwall lemma.

Step 2. Fix $\varepsilon>0$ and a control u_ε such that

$$V(t,x_\lambda)\;>\;g(y(T;t,x_\lambda,u_\varepsilon))-\varepsilon$$

Let ω_g denote a modulus of semiconcavity of g and C_R a Lipschitz constant of g on the ball of radius L_R. Then from (3.41) and Step 1,

$$\begin{aligned} &\lambda V(t,x_1)\;+\;(1-\lambda)V(t,x_0)\;-\;V(t,x_\lambda)\;< \\ &\lambda g(y(T;t,x_1,u_\varepsilon))+(1-\lambda)g(y(T;t,x_0,u_\varepsilon))-g(y(T;t,x_\lambda,u_\varepsilon))+\varepsilon \\ &\leq \lambda(1-\lambda)\,\|y(T;t,x_1,u_\varepsilon)-y(T;t,x_0,u_\varepsilon)\|\times \\ &\times\,\omega_g(L_R,\|y(T;t,x_1,u_\varepsilon)-y(T;t,x_0,u_\varepsilon)\|) \\ &+C_R\,\|\lambda y(T;t,x_1,u_\varepsilon)+(1-\lambda)y(T;t,x_0,u_\varepsilon)-y(T;t,x_\lambda,u_\varepsilon)\|+\varepsilon \\ &\leq e^{\int_t^T k(s)ds}\lambda(1-\lambda)\,\|x_1-x_0\|\,\omega_g\left(L_R,\;e^{\int_t^T k(s)ds}\,\|x_1-x_0\|\right) \\ &+C_R\lambda(1-\lambda)\,\|x_1-x_0\|\;\omega_1(R,\|x_1-x_0\|)\;+\;\varepsilon \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary the proof follows. □

If we impose some additional assumptions on f, then V is semi-concave also with respect to t.

Theorem 3.4.2 *Assume (3.7), (3.8) and (3.39), that k and U are* time independent *and for every $R > 0$, there exists $k_R > 0$ such that for all $x \in B_R(0)$ and $u \in U$, $f(\cdot, x, u)$ is k_R-Lipschitz.*

Further assume that for all $R > 0$, there exists $c_R \geq 0$ such that

$$\forall\,(t,u) \in [0,T] \times U, \quad \forall\, x \in B_R(0), \quad \|f(t,x,u)\| \;\leq\; c_R \tag{3.42}$$

Then the value function is semi-concave on $[0,T] \times \mathbf{R}^n$.

Proof — Consider $0 \leq t_1 < t_0 \leq T$, $R > 0$ and let $x_0, x_1 \in B_R(0)$, $\lambda \in [0,1]$. Define

$$x_\lambda \;=\; \lambda x_1 + (1-\lambda)x_0, \quad t_\lambda \;=\; \lambda t_1 + (1-\lambda)t_0$$

and pick any $\varepsilon > 0$. By (3.14) there exists an admissible control u_ε such that

$$V(t_0, y(t_0; t_\lambda, x_\lambda, u_\varepsilon)) \;<\; V(t_\lambda, x_\lambda) \;+\; \varepsilon$$

Define

$$\tau(s) \;=\; \begin{cases} \lambda s \;+\; (1-\lambda)t_0, & \text{if } t_1 \;\leq\; s \;\leq\; t_0 \\ s & \text{otherwise} \end{cases} \tag{3.43}$$

Since the value function is nondecreasing along solutions to our control system, we have

$$\lambda V(t_1, x_1) + (1-\lambda)V(t_0, x_0) - V(t_\lambda, x_\lambda) \;\leq\; (1-\lambda)V(t_0, x_0) \;+$$

$$\lambda V(t_0, y(t_0; t_1, x_1, u_\varepsilon \circ \tau)) - V(t_0, y(t_0; t_\lambda, x_\lambda, u_\varepsilon)) + \varepsilon$$

Define L_R as in the proof of Theorem 3.4.1 and set

$$y_1(s) \;=\; y(s; t_1, x_1, u_\varepsilon \circ \tau), \quad y_\lambda(s) \;=\; y(s; t_\lambda, x_\lambda, u_\varepsilon)$$

Let K_R denote the Lipschitz constant of V on $[0,T] \times L_R B$ (which exists by Theorem 3.1.1.) From Theorem 3.4.1 and the latter inequality we obtain

$$\begin{aligned} &\lambda V(t_1,x_1) \;+\; (1-\lambda)V(t_0,x_0) \;-\; V(t_\lambda,x_\lambda) \\ &\le \lambda(1-\lambda)\,\|y_1(t_0)-x_0\|\,\overline{\omega}(L_R,\|y_1(t_0)-x_0\|) \qquad (3.44) \\ &+ K_R\,\|\lambda y_1(t_0) \;+\; (1-\lambda)x_0 \;-\; y_\lambda(t_0)\| + \varepsilon \end{aligned}$$

On the other hand from assumption (3.42) it follows that

$$\forall\, s \in [t_1,t_0],\;\; \|y_1(s)-x_0\| \;\le\; \|x_1-x_0\| + M_R(t_0-t_1) \qquad (3.45)$$

where $M_R = c_{L_R}$. Set

$$z(s) \;=\; \lambda y_1\left(\tau^{-1}(s)\right) \;+\; (1-\lambda)x_0 \;-\; y_\lambda(s)$$

and notice that

$$z(t_\lambda) \;=\; 0, \quad z(t_0) \;=\; \lambda y_1(t_0) + (1-\lambda)x_0 - y_\lambda(t_0)$$

On the other hand, by assumptions of theorem there exists $L > 0$ such that for every $t \in [0,T]$ and $u \in U$, $f(t,\cdot,u)$ is L-Lipschitz on $\mathbf{R}^n$. Hence, using (3.39), we obtain the following estimates

$$\begin{aligned} &\|z'(s)\| \;=\; \|f\left(\tau^{-1}(s), y_1\circ\tau^{-1}(s), u_\varepsilon(s)\right) \;-\; f(s,y_\lambda(s),u_\varepsilon(s))\| \\ &\le\; C_R|\tau^{-1}(s)-s| + L\,\|y_1\circ\tau^{-1}(s) \;-\; y_\lambda(s)\| \\ &\le\; L\,\|z(s)\| \;+\; L(1-\lambda)\,\|y_1\circ\tau^{-1}(s) \;-\; x_0\| \;+\; C_R\,\tfrac{1-\lambda}{\lambda}(t_0-s) \end{aligned}$$

where $C_R := k_{L_R}$. Therefore from the Gronwall inequality and (3.45) we deduce that for some $c > 0$ depending only on L and c_R

$$\left\{\begin{aligned} &\|z(t_0)\| \le c(1-\lambda)\int_{t_\lambda}^{t_0}\left(\|y_1\circ\tau^{-1}(s)-x_0\| + \tfrac{t_0-s}{\lambda}\right)ds \\ &\le c\lambda(1-\lambda)(t_0-t_1)\left(\|x_1-x_0\| + (\tfrac{1}{2}+M_R)(t_0-t_1)\right) \end{aligned}\right. \qquad (3.46)$$

Since $\varepsilon > 0$ is arbitrary, inequalities (3.44), (3.45), (3.46) imply the conclusion. □

3.4.1 Differentiability along Optimal Solutions

In this section we provide further results concerning differentiability of V along optimal solutions.

Theorem 3.4.3 *Under all assumptions of Theorems 3.3.7 and 3.4.2, suppose that problem (3.6) has a unique optimal solution z. Then V is differentiable at $(t, z(t))$ for all $t \in [t_0, T]$.*

The proof of this statement is left as an **exercise**.

Theorem 3.3.8 yields

Corollary 3.4.4 *Under hypothezis of Theorems 3.3.6 and 3.4.2, assume that g is continuously differentiable. Then $V(\cdot,\cdot)$ is differentiable at (t_0, x_0) with the partial derivative $\frac{\partial V}{\partial x}(t_0, x_0)$ different from zero if and only if there is a unique optimal solution z to problem (3.6) satisfying $\nabla g(z(T)) \neq 0$.*

Usually the value function is not everywhere differentiable. However this is always the case for "convex" problems, as we prove below.

Theorem 3.4.5 *Assume (3.7), (3.8), (3.39), that g is convex and*

$$\forall\, t \in [0, T], \quad \text{Graph}(f(t, \cdot, U(t))) \text{ is closed and convex} \tag{3.47}$$

Then $V(t, \cdot)$ is convex and continuously differentiable on $\mathbf{R}^n$.

Moreover, if all assumptions of Theorem 3.4.2 are verified, then V is continuously differentiable on $[0, T] \times \mathbf{R}^n$.

Proof — Assumptions (3.47) and (3.8) yield that for every $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ there exists a solution z to control system

$$x' = f(t, x(t), u(t)), \quad u(t) \in U(t), \quad x(t_0) = x_0$$

satisfying $V(t_0, x_0) = g(z(T))$.

Fix $t_0 \in [0, T]$, $x_0, x_1 \in \mathbf{R}^n$, $\lambda \in [0, 1]$ and consider solutions $x : [t_0, T] \mapsto \mathbf{R}^n$ and $y : [t_0, T] \mapsto \mathbf{R}^n$ to (3.4) such that

$$V(t_0, x_0) = g(x(T)), \quad V(t_0, x_1) = g(y(T))$$

Define the map $z : [t_0, T] \mapsto \mathbf{R}^n$ by $z(t) = \lambda x(t) + (1-\lambda)y(t)$. From (3.47), we deduce that z is a solution to control system (3.4) satisfying $z(t_0) = \lambda x_0 + (1-\lambda)x_1$. Thus, from convexity of g,

$$V(t_0, \lambda x_0 + (1-\lambda)x_1) \leq g(z(T)) \leq \lambda V(t_0, x_0) + (1-\lambda)V(t_0, x_1)$$

and therefore $V(t_0, \cdot)$ is convex.

Next, as $V(t, \cdot)$ is both convex and semiconcave for all $t \in [0, T]$, Proposition 1.1.13 yields that $V(t, \cdot)$ is continuously differentiable on $\mathbf{R}^n$. The last statement follows from Proposition 1.1.12. □

3.4.2 Regularity of Optimal Feedback

One of the major issues of optimal control theory is to find an "equation" for optimal solutions. Theorem 3.1.3 provides an inclusion formulation. However, in general, the set-valued map G is not regular enough to make us able to approximate solutions to (3.18) using, say, Euler's scheme. This is one of the reasons why we have to investigate regularity of the set-valued map G.

Theorem 3.4.6 *Under all assumptions of Theorem 3.4.2, suppose that the sets $f(t, x, U)$ are closed. Then G has compact nonempty images and is upper semicontinuous on $[0, T[\times\mathbf{R}^n$.*

Proof — From Theorems 3.4.2 and 1.1.11 we know that for every $(t, x) \in [0, T[\times\mathbf{R}^n$ and every $v \in \mathbf{R}^n$ the directional derivative $\frac{\partial V}{\partial(1,v)}(t, x)$ exists. Define the set-valued map

$$\widehat{Q} : [0, T[\times\mathbf{R}^n \rightsquigarrow \mathbf{R}^n$$

by: for every $(t, x) \in [0, T[\times\mathbf{R}^n$, $\widehat{Q}(t, x)$ is equal to

$$\{v \in \mathbf{R}^n \mid \liminf_{\substack{x' \to x,\, h \to 0+ \\ t' \to t,\ t' \geq 0}} \frac{V(t'+h,\ x'+hv) - V(t', x')}{h} \leq 0\}$$

From Proposition 1.1.14 follows that Graph$(\widehat{Q})$ is closed in $[0, T[\times\mathbf{R}^n \times\mathbf{R}^n$. By Proposition 3.1.2, for all $v \in \overline{co}(f(t, x, U))$, $\frac{\partial V}{\partial(1,v)}(t, x) \geq 0$. Thus

$$G(t, x) = \widehat{Q}(t, x) \cap f(t, x, U)$$

Consequently, Graph(G) is closed in $[0,T[\times\mathbf{R}^n \times \mathbf{R}^n$. From Proposition 1.2.2 we deduce that G is upper semicontinuous on $[0,T[\times\mathbf{R}^n$. □

Corollary 3.4.7 *Under all assumptions of Theorem 3.4.2 suppose that the sets $f(t,x,U)$ are closed. If G is single-valued on a subset $K \subset [0,T[\times\mathbf{R}^n$, then the map $K \ni (t,x) \mapsto G(t,x)$ is continuous.*

Theorem 3.4.8 *We posit all assumptions of Theorems 3.4.2, 3.4.5 and suppose that g is convex. Then G has convex compact images and is upper semicontinuous. Furthermore, if for every (t,x) the set $f(t,x,U)$ is strictly convex[2], then G is single valued and continuous on the set*

$$\left\{(t,x) \in [0,T[\times\mathbf{R}^n \mid \frac{\partial V}{\partial x}(t,x) \neq 0\right\}$$

Proof — By Theorem 3.4.5, we know that V is continuously differentiable. This and convexity of $f(t,x,U)$ yield that for all $(t,x) \in [0,T[\times\mathbf{R}^n$

$$G(t,x) = f(t,x,U) \cap \{v \in \mathbf{R}^n \mid \langle \nabla V(t,x), (1,v)\rangle = 0\}$$

is convex. Theorem 3.4.6 ends the proof of the first statement. From Proposition 3.1.2 it follows that for all $(t,x) \in [0,T[\times\mathbf{R}^n$

$$\begin{cases} v \in G(t,x) \iff \\ v \in f(t,x,U) \ \& \ \sup_{u\in U}\left\langle -\frac{\partial V}{\partial x}(t,x),\ f(t,x,u)\right\rangle = \left\langle -\frac{\partial V}{\partial x}(t,x),\ v\right\rangle \end{cases}$$

This and strict convexity of $f(t,x,U)$ imply that G is single valued on $\left\{(t,x) \in [0,T[\times\mathbf{R}^n \mid \frac{\partial V}{\partial x}(t,x) \neq 0\right\}$. Corollary 3.4.7 completes the proof. □

[2] A subset $K \subset \mathbf{R}^n$ is strictly convex, if for all $x_0, x_1 \in K$ and $0 < \lambda < 1$ we have $\lambda x_0 + (1-\lambda)x_1 \in \text{Int}(K)$ whenever $x_0 \neq x_1$.

Chapter 4

Hamilton-Jacobi-Bellman Equation

Introduction

Consider the Hamilton-Jacobi-Bellman equation:

$$-\frac{\partial V}{\partial t}(t,x) \; + \; H\left(t,x,-\frac{\partial V}{\partial x}(t,x)\right) \; = \; 0, \quad V(T,\cdot) \; = \; g_K(\cdot) \quad (4.1)$$

associated to the Mayer problem with end point constraints:

$$\text{minimize} \; \{ \; g(x(T)) \; | \; x(T) \in K \; \}$$

over all solutions to control system

$$x' \; = \; f(t,x,u(t)), \;\; u(t) \in U \qquad (4.2)$$

satisfying the initial condition $x(0) = \xi_0$, where K is a given subset (called target) and g_K denotes the restriction of g to K.

In the above equation (4.1), the Hamiltonian H is given by:

$$H(t,x,p) \; = \; \sup_{u \in U} \langle p, f(t,x,u) \rangle$$

The value function for the constrained Mayer problem is defined by

$$V(t_0,x_0) = \inf\{g(x(T)) \; | \; x \text{ solves (4.2)}, \; x(t_0) = x_0, \; x(T) \in K\}$$

In general V is merely lower semicontinuous and is equal to $+\infty$ at all points from which it is impossible to reach the target K. In fact one can even avoid using the target K in the definition of minimization problem by setting $g(x) = +\infty$ whenever $x \notin K$.

The value function is nondecreasing along solutions to (4.2) and is constant along optimal solutions. These two properties and the final value $V(T, \cdot) = g(\cdot)$ characterize the value function.

There have been several concepts of "generalized" solutions to Hamilton-Jacobi equation (4.1): *viscosity solutions*, *contingent solutions*, *lower semicontinuous solutions*. Under quite general assumptions we shall prove that all these concepts of solutions are equivalent and that the value function is the unique solution.

The outline is as follows: In Section 1 we state several characterizations of the value function, which are proved in the subsequent sections. Section 2 is devoted to equivalence between contingent solutions and semicontinuous solutions and to the monotone behavior of contingent solutions. Then we show that the value function is the only solution to (4.1). A comparison to continuous viscosity solutions is provided in Section 3.

4.1 Solutions to Hamilton-Jacobi Equation

Consider $T > 0$, a complete separable metric space U and a map $f : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}^n$. We associate with it the control system

$$x'(t) = f(t, x(t), u(t)), \quad u(t) \in U \tag{4.3}$$

Let an extended function $g : \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ and $\xi_0 \in \mathbf{R}^n$ be given. Consider the minimization problem (*Mayer's problem*):

$$\min \{g(x(T)) \mid x \text{ is a solution to } (4.3),\ x(0) = \xi_0\} \tag{4.4}$$

The value function $V : [0, T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ associated with it is defined by: for all $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$

$$V(t_0, x_0) = \inf\{g(x(T)) \mid x \text{ is a solution to } (4.3),\ x(t_0) = x_0\}$$

We impose the following assumptions

$$\left\{\begin{array}{ll} i) & \forall\, R>0,\ \exists\, c_R \in L^1(0,T) \text{ such that for almost all } t \\ & \forall\, u \in U,\ f(t,\cdot,u) \text{ is } c_R(t)-\text{Lipschitz on } B_R(0) \\ ii) & \exists\, k \in L^1(0,T) \text{ such that for almost all } t \in [0,T], \\ & \forall\, x \in \mathbf{R}^n,\ \sup_{u\in U} \|f(t,x,u)\| \le k(t)(1+\|x\|) \\ iii) & \forall\, (t,x) \in [0,T]\times\mathbf{R}^n,\ f(t,x,U) \text{ is convex, compact} \\ iv) & f \text{ is continuous and for all } (t,x) \in [0,T]\times\mathbf{R}^n, \\ & \lim_{(t',x')\to(t,x)} \sup_{u\in U} \|f(t,x,u)-f(t',x',u)\| = 0 \\ v) & \text{g is lower semicontinuous} \end{array}\right. \tag{4.5}$$

By these assumptions the control system (4.3) may be replaced by the differential inclusion

$$x'(t) \in F(t,x(t)) \text{ almost everywhere} \tag{4.6}$$

where $F(t,x) = f(t,x,U)$. Furthermore, F satisfies the following conditions:

$$\left\{\begin{array}{ll} a) & F \text{ is continuous and has nonempty convex compact} \\ & \text{images} \\ b) & \exists\, k \in L^1(0,T) \text{ such that for almost all } t \in [0,T], \\ & \forall\, x \in \mathbf{R}^n,\ \sup_{v\in F(t,x)} \|v\| \le k(t)(1+\|x\|) \\ c) & \forall\, R>0,\ \exists\, c_R \in L^1(0,T) \text{ such that for a.e. } t \in [0,T] \\ & F(t,\cdot) \text{ is } c_R(t)\text{-Lipschitz on } B_R(0) \end{array}\right. \tag{4.7}$$

Recall that $\mathcal{S}_{[t_0,T]}(x_0)$ denotes the set of absolutely continuous solutions to (4.6) defined on $[t_0,T]$ and satisfying the initial condition $x(t_0) = x_0$.

We show next that

$$V(t_0,x_0) = \min\left\{ g(x(T)) \mid x \in \mathcal{S}_{[t_0,T]}(x_0) \right\} \tag{4.8}$$

Theorem 4.1.1 *Consider an extended lower semicontinuous function* $g : \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *and assume (4.5).*

Then V *takes its values in* $\mathbf{R} \cup \{+\infty\}$ *and is lower semicontinuous.*

This Theorem follows from

Theorem 4.1.2 *Consider a set-valued map* $F : [0,T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ *and an extended lower semicontinuous function* $g : \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$.

Assume (4.7) and define $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{\pm\infty\}$ *by*

$$V(t_0, x_0) \;=\; \inf \{g(x(T)) \mid x \text{ solves } (4.6),\; x(t_0) = x_0\}$$

Then V *is lower semicontinuous taking its values in* $\mathbf{R} \cup \{+\infty\}$ *and (4.8) holds true.*

Proof — Fix $(t_0, x_0) \in [0,T] \times \mathbf{R}^n$. Since the set of solutions $\mathcal{S}_{[t_0,T]}(x_0)$ is compact and g is lower semicontinuous we deduce that $V(t_0, x_0) > -\infty$ and (4.8). To prove that V is lower semicontinuous consider a sequence $(t_n, x_0^n) \to (t_0, x_0)$ such that $t_n \in [0,T]$

$$\liminf_{(t,x)\to(t_0,x_0),\; t\in[0,T]} V(t,x) \;=\; \lim_{n\to\infty} V(t_n, x_0^n)$$

and let $x_n \in \mathcal{S}_{[t_n,T]}(x_0^n)$ be such that

$$g(x_n(T)) \;\le\; V(t_n, x_0^n) \;+\; \frac{1}{n}$$

For every $n \ge 1$ such that $t_n > t_0$ consider a solution y_n to inclusion

$$x'(s) \;\in\; -F(t_n - s, x(s)),\;\; x(0) \;=\; x_0^n$$

defined on $[0, t_n - t_0]$. Set

$$z_n(t) \;=\; \begin{cases} x_n(t) & \text{if } t \ge t_n \\ y_n(t_n - t) & \text{otherwise} \end{cases}$$

Then $z_n \in \mathcal{S}_{[t_0,T]}(y_n(t_n - t_0))$. We also observe that (4.7) b) yields that $y_n(t_n - t_0)$ converge to x_0. By Fillipov's theorem there exist $\bar{z}_n \in \mathcal{S}_{[t_0,T]}(x_0)$ such that $z_n - \bar{z}_n$ converge uniformly to zero. The set $\mathcal{S}_{[t_0,T]}(x_0)$ being compact in the space of continuous functions, there exists a subsequence $\bar{z}_{n_k}$ converging to some $z \in \mathcal{S}_{[t_0,T]}(x_0)$. But then also z_{n_k} converge to z. Since g is lower semicontinuous, $g(z(T)) \le \lim_{n\to\infty} V(t_n, x_0^n)$. On the other hand, $V(t_0, x_0) \le g(z(T))$ and the proof follows. □

The Hamiltonian associated to control system (4.3) is the function $H : [0,T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ defined by

$$H(t,x,p) = \max_{u\in U} \langle p, f(t,x,u)\rangle$$

Consider the Hamilton-Jacobi equation

$$-\frac{\partial V}{\partial t}(t,x) + H\left(t,\ x,\ -\frac{\partial V}{\partial x}(t,x)\right) = 0, \quad V(T,\cdot) = g(\cdot) \qquad (4.9)$$

In the result stated below we use notions of super/subdifferentials and epi/hypoderivatives introduced in Chapter 1.

Definition 4.1.3 *An extended lower semicontinuous function* $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *is called a* lower semicontinuous solution *to (4.9) if it satisfies the following conditions:*

$$\begin{cases} V(T,\cdot) = g(\cdot) \text{ and for all } (t,x) \in]0,T[\times \mathbf{R}^n, \\ \forall\, (p_t,p_x) \in \partial_- V(t,x), \quad -p_t + H(t,x,-p_x) = 0 \\ \forall\, (p_t,p_x) \in \partial_- V(0,x), \quad -p_t + H(0,x,-p_x) \ge 0 \\ \forall\, (p_t,p_x) \in \partial_- V(T,x), \quad -p_t + H(T,x,-p_x) \le 0 \end{cases}$$

Definition 4.1.4 *An extended lower semicontinuous function* $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *is called a* viscosity supersolution *to (4.9) if for all* $t \in]0,T[$ *and* $x \in \mathbf{R}^n$ *such that* $(t,x) \in \mathrm{Dom}(V)$ *we have*

$$\forall\, (p_t,p_x) \in \partial_- V(t,x), \quad -p_t + H(t,x,-p_x) \ge 0$$

An extended upper semicontinuous function $V : [0,T] \times \mathbf{R}^n \to \mathbf{R} \cup \{-\infty\}$ *is called a* viscosity subsolution *to (4.9) if for all* $0 < t < T$ *and* $x \in \mathbf{R}^n$ *such that* $(t,x) \in \mathrm{Dom}(V)$ *we have*

$$\forall\, (p_t,p_x) \in \partial_+ V(t,x), \quad -p_t + H(t,x,-p_x) \le 0$$

Let $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R}$ *be a continuous function. It is called a* viscosity solution *to (4.9) if for all* $t \in]0,T[$ *and* $x \in \mathbf{R}^n$

$$\begin{array}{l} \forall\, (p_t,p_x) \in \partial_- V(t,x), \quad -p_t + H(t,x,-p_x) \ge 0 \\ \forall\, (p_t,p_x) \in \partial_+ V(t,x), \quad -p_t + H(t,x,-p_x) \le 0 \end{array}$$

Theorem 4.1.5 *Assume (4.7) and let* $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *be an extended lower semicontinuous function.*

Then the following four statements are equivalent:

i) V is the value function given by (4.8)

ii) V is a lower semicontinuous solution to (4.9)

iii) V is a contingent solution to (4.9) in the sense that $V(T,\cdot) = g(\cdot)$ and for all $(t,x) \in \mathrm{Dom}(V)$,
$0 \le t < T \implies \inf_{v \in F(t,x)} D_\uparrow V(t,x)(1,v) \le 0$
$0 < t \le T \implies \sup_{v \in F(t,x)} D_\uparrow V(t,x)(-1,-v) \le 0$

iv) $V(T,\cdot) = g(\cdot)$ and for all $(t,x) \in]0,T[\times \mathbf{R}^n$,
$\forall\, (p_t, p_x) \in \partial_- V(t,x), \quad -p_t + H(t,x,-p_x) = 0$
$\forall\, \overline{x} \in \mathbf{R}^n, \quad V(0,\overline{x}) = \liminf_{t \to 0+,\, x \to \overline{x}} V(t,x)$
$\forall\, \overline{x} \in \mathbf{R}^n, \quad g(\overline{x}) = \liminf_{t \to T-,\, x \to \overline{x}} V(t,x)$

Finally, if V *is continuous on* $[0,T] \times \mathbf{R}^n$ *then the above statements are equivalent to:*

v) V *is a viscosity solution to (4.9).*

4.2 Lower Semicontinuous Solutions

4.2.1 Lower Semicontinuous & Contingent Solutions

The equivalence between statements *ii)* and *iii)* of Theorem 4.1.5 follows from Theorems 4.2.1 and 4.2.2 proved below.

Let $T > 0$. Consider a set-valued map $F : [0,T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ with nonempty bounded images and define the Hamiltonian $H : [0,T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ by

$$H(t,x,p) \;=\; \sup_{v \in F(t,x)} < p, v > \tag{4.10}$$

Then $H(t,x,\cdot)$ is convex and positively homogeneous. Furthermore, if F is continuous, then so is H.

Theorem 4.2.1 *Consider an extended lower semicontinuous function* $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$. *Assume that* F *is upper semicontinuous and has nonempty convex compact images on* $\mathrm{Dom}(V)$.

Then the following four statements are equivalent :
i) For all $(t, x) \in \mathrm{Dom}(V)$ *such that* $t < T$ *and for every* $(p_t, p_x, q) \in N^0_{\mathcal{E}p(V)}(t, x, V(t, x))$

$$-p_t + H(t, x, -p_x) \;\geq\; 0 \tag{4.11}$$

ii) For all $(t, x) \in \mathrm{Dom}(V)$ *such that* $t < T$ *and all* $y \geq V(t, x)$

$$(\{1\} \times F(t, x) \times \{0\}) \cap T_{\mathcal{E}p(V)}(t, x, y) \;\neq\; \emptyset$$

iii) For all $(t, x) \in \mathrm{Dom}(V)$ *such that* $t < T$

$$\inf_{v \in F(t,x)} D_\uparrow V(t, x)(1, v) \leq 0$$

iv) For all $(t, x) \in \mathrm{Dom}(V)$ *such that* $t < T$

$$\forall\, (p_t, p_x) \;\in\; \partial_- V(t, x), \quad -p_t + H(t, x, -p_x) \geq 0$$

Proof — Fix $(t, x) \in \mathrm{Dom}(V)$ such that $t < T$ and observe that

$$\forall\, y \;\geq\; V(t, x), \quad T_{\mathcal{E}p(V)}(t, x, V(t, x)) \;\subset\; T_{\mathcal{E}p(V)}(t, x, y) \tag{4.12}$$

Since the contingent cone to the epigraph is the epigraph of the epiderivative, *ii)* is equivalent to *iii)*.

Assume that *iii)* holds true. Fix $(t, x) \in \mathrm{Dom}(V)$ such that $t < T$ and $(p_t, p_x, q) \in [T_{\mathcal{E}p(V)}(t, x, V(t, x))]^-$. Consider $v \in F(t, x)$ satisfying

$$D_\uparrow V(t, x)(1, v) \leq 0$$

or, equivalently, $(1, v, 0) \in T_{\mathcal{E}p(V)}(t, x, V(t, x))$. Hence $p_t + \langle p_x, v\rangle \leq 0$ and *i)* follows.

Assume next that *i)* holds true. We claim that

$$(\{1\} \times F(t, x) \times \{0\}) \cap \overline{co}\left(T_{\mathcal{E}p(V)}(t, x, y)\right) \;\neq\; \emptyset \tag{4.13}$$

for all $(t, x) \in \mathrm{Dom}(V)$ such that $t < T$ and $y \geq V(t, x)$.

By (4.12) it is enough to prove (4.13) with $y = V(t, x)$. Otherwise, by the separation theorem, there exists

$$(p_t, p_x, q) \;\in\; N^0_{\mathcal{E}p(V)}(t, x, V(t, x))$$

such that

$$p_t + \inf_{v \in F(t,x)} \langle p_x, v\rangle > 0$$

Consequently $-p_t + H(t,x,-p_x) < 0$, which contradicts *i)* and proves (4.13).

Finally, since F is upper semicontinuous and has convex compact images, Proposition 1.3.12 and relation (4.13) imply *ii)*.

By Proposition 1.1.15, *i)* yields *iv)*.

Assume next that *iv)* is verified. Fix $(t,x) \in \mathrm{Dom}(V)$ such that $t < T$ and $(p_t, p_x, q) \in N^0_{\mathcal{E}p(V)}(t,x,V(t,x))$. Since

$$\{0\} \times \{0\} \times \mathbf{R}_+ \subset T_{\mathcal{E}p(V)}(t,x,V(t,x))$$

we have $q \leq 0$. If $q < 0$, then

$$\left(\frac{p_t}{|q|}, \frac{p_x}{|q|}, -1\right) \in N^0_{\mathcal{E}p(V)}(t,x,V(t,x))$$

From Proposition 1.1.15 and *iv)* we deduce that

$$-\frac{p_t}{|q|} + H\left(t, x, -\frac{p_x}{|q|}\right) \geq 0$$

Multiplying by $|q|$ the above relations we derive (4.11).

It remains to consider the case $q = 0$ and $(p_t, p_x) \neq 0$. For this aim it is enough to apply Lemma 1.1.16 and to use the same arguments as above. □

Theorem 4.2.2 *Consider an extended lower semicontinuous function* $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *and assume that* F *is lower semicontinuous and has nonempty compact images on* $\mathrm{Dom}(V)$.

Then the following four statements are equivalent :

i) For all $(t,x) \in \mathrm{Dom}(V)$ *such that* $t > 0$ *and for every* $(p_t, p_x, q) \in N^0_{\mathcal{E}p(V)}(t,x,V(t,x))$

$$-p_t + H(t,x,-p_x) \leq 0$$

ii) For all $(t,x) \in \mathrm{Dom}(V)$ *such that* $t > 0$ *and all* $y \geq V(t,x)$

$$\{-1\} \times (-F(t,x)) \times \{0\} \subset T_{\mathcal{E}p(V)}(t,x,y)$$

iii) For all $(t,x) \in \mathrm{Dom}(V)$ *such that* $t > 0$

$$\sup_{v \in F(t,x)} D_\uparrow V(t,x)(-1,-v) \leq 0$$

iv) For all $(t,x) \in \mathrm{Dom}(V)$ *such that* $t > 0$

$$\forall\ (p_t, p_x)\ \in\ \partial_- V(t,x),\quad -p_t + H(t,x,-p_x) \leq 0$$

Proof — We deduce using (4.12) that *ii)* is equivalent to *iii)*. Clearly, *ii)* yields *i)*. We next claim that *i)* implies that

$$\{-1\} \times (-F(t,x)) \times \{0\}\ \subset\ \overline{co}\left(T_{\mathcal{E}p(V)}(t,x,y)\right) \tag{4.14}$$

for all $(t,x) \in \mathrm{Dom}(V)$ such that $t > 0$ and $y \geq V(t,x)$.

Indeed, by (4.12) it is enough to consider the case $y = V(t,x)$. If (4.14) is not satisfied then, by the separation theorem, there exist $v \in F(t,x)$ and $(p_t, p_x, q) \in N^0_{\mathcal{E}p(V)}(t,x,V(t,x))$ such that $-p_t + \langle -p_x, v\rangle > 0$. Consequently $-p_t + H(t,x,-p_x) > 0$, which contradicts *i)* and so inclusion (4.14) follows.

Since F is lower semicontinuous, (4.14) and Theorem 1.1.3 yield that for all $(t,x) \in \mathrm{Dom}(V)$ such that $t > 0$ and all $y \geq V(t,x)$,

$$\begin{array}{l}\{-1\} \times (-F(t,x)) \times \{0\} \\ \subset \mathrm{Liminf}_{(t',x',y') \to_{\mathcal{E}p(V)} (t,x,y)}\ \overline{co}\left(T_{\mathcal{E}p(V)}(t',x',y')\right) \subset T_{\mathcal{E}p(V)}(t,x,y)\end{array}$$

Arguments similar to those of the proof of Theorem 4.2.1 yield that *i)* is equivalent to *iv)*.

4.2.2 Monotone Behavior of Contingent Solutions

Consider a set-valued map $F : [0,T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ and the differential inclusion

$$x'(t)\ \in\ F(t,x(t)) \text{ almost everywhere} \tag{4.15}$$

In this section we investigate a relationship between monotone behavior of a function V along solutions to (4.15) and contingent inequalities *iii)* of Theorem 4.1.5.

Theorem 4.2.3 *Let $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ be an extended lower semicontinuous function. Assume that F is upper semicontinuous, that $F(t,x)$ is nonempty convex and compact for all $(t,x) \in \mathrm{Dom}(V)$ and that for some $k \in L^1(0,T)$*

$$\forall\, (t,x) \in \mathrm{Dom}(V), \quad \sup_{v \in F(t,x)} \|v\| \le k(t)(1+\|x\|)$$

Then the following two statements are equivalent

i) $\forall\, (t,x) \in \mathrm{Dom}(V)$ with $t < T$, $\inf_{v \in F(t,x)} D_{\uparrow}V(t,x)(1,v) \le 0$

ii) For every $(t_0,x_0) \in [0,T] \times \mathbf{R}^n$, there exists $\overline{x} \in \mathcal{S}_{[t_0,T]}(x_0)$ such that $V(t,\overline{x}(t)) \le V(t_0,x_0)$ for all $t \in [t_0,T]$.

Proof — Assume that *i)* holds true and fix $(t_0,x_0) \in \mathrm{Dom}(V)$. Define the upper semicontinuous set-valued map

$$\widehat{F} : \mathbf{R}_+ \times \mathbf{R}^n \times \mathbf{R} \rightsquigarrow \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}$$

by

$$\widehat{F}(t,x,z) = \begin{cases} \{1\} \times F(t,x) \times \{0\} & \text{when } t < T \\ [0,1] \times \overline{co}\,(F(T,x) \cup \{0\}) \times \{0\} & \text{when } t \ge T \end{cases}$$

and consider the viability problem

$$\begin{cases} (t,x,z)' & \in & \widehat{F}(t,x,z) \\ (t,x,z)(t_0) & = & (t_0,x_0,V(t_0,x_0)) \\ (t,x,z) & \in & \mathcal{E}p(V) \end{cases} \tag{4.16}$$

By Theorem 4.2.1, for all $(t,x,z) \in \mathcal{E}p(V)$ we have

$$\widehat{F}(t,x,z) \cap T_{\mathcal{E}p(V)}(t,x,z) \neq \emptyset$$

Since $\widehat{F}$ is upper semicontinuous and has convex compact nonempty images and linear growth on the closed set $\mathcal{E}p(V)$, the Viability Theorem 1.3.11 yields that problem (4.16) has a solution

$$[t_0,T] \ni t \mapsto (t,\overline{x}(t),z(t)) \in \mathcal{E}p(V)$$

Thus $V(t, \overline{x}(t)) \le z(t) = V(t_0, x_0)$ for all $t \in [t_0, T]$ and *ii)* follows.

Conversely, assume that *ii)* is satisfied. Fix $(t_0, x_0) \in \mathrm{Dom}(V)$ with $t_0 < T$ and let $\overline{x}$ be as in *ii)*. Since F is bounded on a neighborhood of (t_0, x_0), we deduce that $\overline{x}(\cdot)$ is Lipschitz at t_0. Let $h_n \to 0+$ be such that $[x(t_0 + h_n) - x(t_0)]/h_n$ converge to some v. Theorem 1.3.8 yields that $v \in F(t_0, x_0)$. On the other hand

$$D_\uparrow V(t_0, x_0)(1, v) \le \liminf_{n \to \infty} \frac{V(t_0 + h_n, x(t_0 + h_n)) - V(t_0, x_0)}{h_n} \le 0 \quad \square$$

Theorem 4.2.4 *Let* $V : [0, T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *be an extended lower semicontinuous function. If* F *verifies (4.7), then the following two statements are equivalent:*

i) $\forall\, (t, x) \in \mathrm{Dom}(V)$ with $t > 0, \quad \sup_{v \in F(t,x)} D_\uparrow V(t, x)(-1, -v) \le 0$

ii) For every $x \in \mathcal{S}_{[t_0,T]}(x_0)$ *and all* $t \in [t_0, T]$, $V(t_0, x_0) \le V(t, x(t))$.

Proof — Assume that *i)* holds true. Since *i)* does not involve T, it is enough to prove the inequality in *ii)* for $t = T$. By Theorem 4.2.2, for all $0 \le t < T$ and $x \in \mathbf{R}^n$ such that $(T-t, x) \in \mathrm{Dom}(V)$ and for all $z \ge V(T - t, x)$,

$$\{-1\} \times (-F(T - t, x)) \times \{0\} \subset T_{\mathcal{E}p(V)}(T - t, x, z) \qquad (4.17)$$

Let U denote the closed unit ball in $\mathbf{R}^n$. From Theorem 1.4.4 there exists a continuous function $f : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}^n$ and $\alpha > 0$ such that

$$\begin{cases} \forall\, (t, x) \in [0, T] \times \mathbf{R}^n, \ F(t, x) = f(t, x, U) \\ \forall\, u \in U, \ f(t, \cdot, u) \text{ is } \alpha c_R(t) - \text{Lipschitz on } B_R(0) \text{ a.e. in } [0, T] \\ \forall\, (t, x) \in [0, T] \times \mathbf{R}^n \text{ and for all } u, v \in U, \text{ we have} \\ \|f(t, x, u) - f(t, x, v)\| \le \alpha(\sup_{y \in F(t,x)} \|y\|)\, \|u - v\| \end{cases}$$

Fix $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ and $x \in \mathcal{S}_{[t_0,T]}(x_0)$. It is enough to consider the case $V(T, x(T)) < \infty$.

Consider a measurable map $u : [t_0, T] \mapsto U$ such that

$$x'(t) = f(t, x(t), u(t))$$

almost everywhere. Consider a sequence of continuous maps $u_k : [t_0, T] \mapsto U$ converging to u in $L^1(t_0, T; U)$ and let x_k denote the solution to

$$x_k'(t) = f(t, x_k(t), u_k(t)), \ t \in [t_0, T], \ x_k(T) = x(T)$$

The Gronwall lemma implies that x_k converge uniformly to x. On the other hand, the map $t \mapsto (T - t, x_k(T - t), V(T, x(T)))$ is the only solution to

$$\left\{ \begin{array}{lll} \gamma'(t) & = & -1 \\ y'(t) & = & -f(T - t, y(t), u_k(T - t)) \\ z'(t) & = & 0 \\ \gamma(0) & = & T, \ y(0) = x(T), \ z(0) = V(T, x(T)) \end{array} \right. \tag{4.18}$$

By (4.17) we know that

$$\forall \ (\gamma, x, z) \in \mathcal{E}p(V), \ (-1, -f(\gamma, x, u_k(\gamma)), 0) \in T_{\mathcal{E}p(V)}(\gamma, x, z)$$

On the other hand the map

$$(t, x) \rightsquigarrow \{-f(T - t, x, u_k(T - t))\}$$

being continuous, from Viability Theorem we deduce that problem (4.18) has at least one solution

$$[0, T - t_0] \ni t \mapsto (\gamma(t), y(t), z(t)) \in \mathcal{E}p(V)$$

Consequently,

$$\forall \, 0 \le t \le T - t_0, \ (T - t, x_k(T - t), V(T, x(T))) \in \mathcal{E}p(V)$$

and therefore

$$\forall \, 0 \le t \le T - t_0, \ V(T, x(T)) \ge V(T - t, x_k(T - t))$$

In particular, $V(t_0, x_k(t_0)) \le V(T, x(T))$. Taking the limit when $k \to \infty$ and using that V is lower semicontinuous, we deduce *ii)* for $t = T$.

Conversely, assume that *ii)* is verified. Let $(t_0, x_0) \in \text{Dom}(V)$ be such that $t_0 > 0$. Fix $v \in F(t_0, x_0)$. Corollary 1.3.3 implies that for some $\overline{h} > 0$ there exist $y_0 \in \mathbf{R}^n$ and $y \in \mathcal{S}_{[t_0-\overline{h},t_0]}(y_0)$ such that $y(t_0) = x_0$ and

$$\lim_{h\to 0+} \frac{y(t_0-h) - x_0}{h} = -v$$

On the other hand, by *ii*),

$$\forall\, h \in [0,\overline{h}],\quad V(t_0-h, y(t_0-h)) \leq V(t_0,x_0)$$

Consequently $D_\uparrow V(t_0,x_0)(-1,-v) \leq 0$. Since $v \in F(t_0,x_0)$ is arbitrary *i*) follows.

4.2.3 Value Function & Contingent Solutions

We prove here equivalence of *i*) and *iii*) of Theorem 4.1.5.

The Proposition below yields the implication *i*) $\Longrightarrow$ *iii*).

Proposition 4.2.5 *Assume (4.7) and let V be defined by (4.8). Then for all $(t_0,x_0) \in \mathrm{Dom}(V)$,*

$$\begin{cases} t_0 < T & \Longrightarrow & \inf_{v\in F(t_0,x_0)} D_\uparrow V(t_0,x_0)(1,v) \leq 0 \\ t_0 > 0 & \Longrightarrow & \sup_{v\in F(t_0,x_0)} D_\uparrow V(t_0,x_0)(-1,-v) \leq 0 \end{cases}$$

Proof — Fix (t_0,x_0) as above. Then, there exists $x \in \mathcal{S}_{[t_0,T]}(x_0)$ such that $V(t,x(t)) \equiv g(x(T))$. Theorem 4.2.3 ends the proof of the first statement. The second one follows from Theorem 4.2.4. □

The implication *iii*) $\Longrightarrow$ *i*) follows from

Theorem 4.2.6 *Assume that (4.7) holds true. Then the function V defined by (4.8) is the only lower semicontinuous function from $[0,T] \times \mathbf{R}^n$ into $\mathbf{R} \cup \{+\infty\}$ satisfying*

$$\begin{cases} V(T,\cdot) = g(\cdot) \text{ and for all } (t,x) \in \mathrm{Dom}(V), \\ 0 \leq t < T \Longrightarrow \inf_{v\in F(t,x)} D_\uparrow V(t,x)(1,v) \leq 0 \\ 0 < t \leq T \Longrightarrow \sup_{v\in F(t,x)} D_\uparrow V(t,x)(-1,-v) \leq 0 \end{cases} \tag{4.19}$$

Proof — Proposition 4.2.5 implies that V verifies (4.19). Conversely, consider an extended lower semicontinuous $W : [0,T]\times\mathbf{R}^n \mapsto$

$\mathbf{R} \cup \{+\infty\}$ satisfying (4.19). Fix $(t_0, x_0) \in \mathrm{Dom}(W)$ with $t_0 < T$. By Theorem 4.2.3, there exists $\overline{x} \in S_{[t_0,T]}(x_0)$ such that

$$V(t_0, x_0) \leq g(\overline{x}(T)) = W(T, \overline{x}(T)) \leq W(t_0, x_0)$$

Therefore $W \geq V$. To prove the opposite inequality, consider $(t_0, x_0) \in \mathrm{Dom}(V)$ and $\overline{x} \in \mathcal{S}_{[t_0,T]}(x_0)$ such that $V(t_0, x_0) = g(\overline{x}(T))$. Thus, by Theorem 4.2.4,

$$W(t_0, x_0) \leq W(T, \overline{x}(T)) = g(\overline{x}(T)) = V(t_0, x_0)$$

Hence $W \leq V$ and the proof is complete.

4.2.4 Regularity of Value Function at Boundary Points

We observe that *i)* yields *iv)*. To prove that *iv)* yields *i)*, by using the equivalence proved in the preceeding section, it is enough to show that *iv)* implies *iii)*.

Consider a set-valued map $F : [0, T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$.

Theorem 4.2.7 *Assume (4.7). If an extended lower semicontinuous function* $V : [0, T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *satisfies*

$$\begin{cases} \forall\, (t, x) \in]0, T[\times \mathbf{R}^n, \forall\, (p_t, p_x) \in \partial_- V(t, x), -p_t + H(t, x, -p_x) = 0 \\ \forall\, \overline{x} \in \mathbf{R}^n, \quad V(0, \overline{x}) = \liminf_{t \to 0+,\, x \to \overline{x}} V(t, x) \\ \forall\, \overline{x} \in \mathbf{R}^n, \quad V(T, \overline{x}) = \liminf_{t \to T-,\, x \to \overline{x}} V(t, x) \end{cases}$$

then for all $(t, x) \in \mathrm{Dom}(V)$,

$$\begin{cases} 0 < t \leq T \implies \sup_{v \in F(t,x)} D_\uparrow V(t, x)(-1, -v) \leq 0 \\ 0 \leq t < T \implies \inf_{v \in F(t,x)} D_\uparrow V(t, x)(1, v) \leq 0 \end{cases} \tag{4.20}$$

Proof — From the proofs of Theorems 4.2.1 and 4.2.2 we deduce that for all $(t, x) \in \mathrm{Dom}(V)$ with $0 < t < T$ we have

$$\inf_{v \in F(t,x)} D_\uparrow V(t, x)(1, v) \leq 0, \quad \sup_{v \in F(t,x)} D_\uparrow V(t, x)(-1, -v) \leq 0$$

This and Theorems 4.2.3 and 4.2.4 yield that for all $0 < t_1 \leq t_2 < T$

$$\forall\, x \in \mathcal{S}_{[t_1,t_2]}(x_1), \quad V(t_1, x_1) \leq V(t_2, x(t_2)) \tag{4.21}$$

and

$$\forall\ (t_1, x_1),\ \exists\ x \in \mathcal{S}_{[t_1,t_2]}(x_1),\ V(t_1, x_1) = V(t_2, x(t_2)) \tag{4.22}$$

Fix $\overline{x} \in \mathrm{Dom}(V(T,\cdot))$ and let $t_i \to T-,\ x_i \to \overline{x}$ be such that

$$\lim_{i\to\infty} V(t_i, x_i) = V(T, \overline{x})$$

Consider any $x_0 \in \mathbf{R}^n$ and $x \in \mathcal{S}_{[0,T]}(x_0)$ satisfying $x(T) = \overline{x}$. Then we can find $\overline{y}_i$ and $y_i \in \mathcal{S}_{[0,T]}(\overline{y}_i)$ such that $y_i(t_i) = x_i$ and y_i converge to x uniformly on $[0,T]$. Then for all arbitrary, but fixed $0 < t < T$ and all i large enough,

$$V(t, y_i(t)) \le V(t_i, x_i)$$

Since V is lower semicontinuous,

$$V(t, x(t)) \le \liminf_{i\to\infty} V(t, y_i(t)) \le \lim_{i\to\infty} V(t_i, x_i) = V(T, \overline{x})$$

This, (4.21) and Theorem 4.2.4 yield the first inequality in (4.20).

To prove the second one fix $(0, \overline{x}) \in \mathrm{Dom}(V)$ and consider $t_i \to 0+,\ \overline{x}_i \to \overline{x}$ such that

$$V(0, \overline{x}) = \lim_{i\to\infty} V(t_i, \overline{x}_i)$$

Let $\overline{y}_i \in \mathbf{R}^n$ and $x_i \in \mathcal{S}_{[0,T]}(\overline{y}_i)$ be such that $x_i(t_i) = \overline{x}_i$ and

$$\forall\ t_i \le t \le T - \frac{1}{i} \ \text{ we have } \ V(t_i, \overline{x}_i) = V(t, x_i(t))$$

Taking a subsequence and keeping the same notations, we may assume that x_i converge uniformly to some $x \in \mathcal{S}_{[0,T]}(\overline{x})$. Then for all $0 < t < T$,

$$V(0, \overline{x}) = \lim_{i\to\infty} V(t, x_i(t)) \ge V(t, x(t))$$

This, (4.22) and Theorem 4.2.3 imply the second inequality in (4.20).

4.3 Viscosity Solutions

In this Section we prove that statements *i)* and *v)* of Theorem 4.1.5 are equivalent, whenever V is continuous.

Let $F : [0,T] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$ be a set-valued map with nonempty compact images and H be defined by (4.10).

Consider the Hamilton-Jacobi-Bellman equation

$$-\frac{\partial V}{\partial t}(t,x) + H\left(t,x,-\frac{\partial V}{\partial x}(t,x)\right) = 0 \qquad (4.23)$$

Clearly, any V satisfying *ii)* of Theorem 4.1.5 is a viscosity supersolution.

Theorem 4.3.1 *Let* $V : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R} \cup \{+\infty\}$ *be an extended lower semicontinuous function. Assume that* F *is upper semicontinuous and has convex compact nonempty images.*

Then the following two statements are equivalent:

i) V *is a viscosity supersolution of (4.23)*

ii) For all $0 < t < T$ *and* $x \in \mathbf{R}^n$ *such that* $V(t,x) \neq +\infty$, *we have*

$$\inf_{v \in F(t,x)} D_{\uparrow}V(t,x)(1,v) \leq 0 \qquad (4.24)$$

Proof — This follows by the same arguments as in the proof of Theorem 4.2.1. □

Notice next that

$$T_{\mathcal{H}yp(\varphi)}(x_0,\varphi(x_0)) = \mathcal{H}yp\ (D_{\downarrow}\varphi(x_0))$$

where $\mathcal{H}yp$ denotes for the hypograph.

In particular, $p \in \partial_+\varphi(x_0)$ if and only if

$$\forall\, u \in \mathbf{R}^n,\ \ D_{\downarrow}\varphi(x_0)(u) \leq \langle p,u\rangle \qquad (4.25)$$

Thus

$$p \in \partial_+\varphi(x_0) \iff (-p,\ +1) \in N^0_{\mathcal{H}yp\,(\varphi)}(x_0,\varphi(x_0)) \qquad (4.26)$$

Theorem 4.3.2 *Let $V : [0,T] \times \mathbf{R}^n \to \mathbf{R}$ be continuous. Assume that F satisfies (4.7).*

Then the following two statements are equivalent

i) V is a viscosity subsolution of (4.23)

ii) For all $0 < t < T$, x , $\sup_{v \in F(t,x)} D_{\uparrow} V(t,x)(-1,-v) \leq 0$

Proof — Assume that *ii)* holds true. Fix $0 < t_0 < T$. By Theorem 4.2.4, for every $t_0 \leq t_1 < T$ and every $x \in \mathcal{S}_{[t_0,t_1]}(x_0)$ the following holds true:

$$\forall\, t \in [t_0,t_1], \quad V(t_0,x_0) \leq V(t,x(t))$$

Fix $v \in F(t_0,x_0)$. By Corollary 1.3.3 there exist $t_0 < t_1 < T$ and $x \in \mathcal{S}_{[t_0,t_1]}(x_0)$ such that $x'(t_0) = v$. The above inequality yields that $0 \leq D_{\downarrow} V(t_0,x_0)(1,v)$. Consequently,

$$\forall\, (p_t,p_x) \in \partial_+ V(t_0,x_0), \quad 0 \leq p_t + \langle p_x, v\rangle$$

Since $v \in F(t_0,x_0)$ is arbitrary, V is a viscosity subsolution.

Assume *i)*. We claim that for all (t,x) such that $0 < t < T$ and all $z \leq V(t,x)$ we have

$$\forall\, (q_t,q_x,q) \in N^0_{\mathcal{H}yp\,(V)}(t,x,z), \quad q_t + H(t,x,q_x) \leq 0 \tag{4.27}$$

Indeed it is enough to consider the case $z = V(t,x)$. Fix such (q_t,q_x,q). Clearly $q \geq 0$. If $q > 0$ then

$$\left(\frac{q_t}{q},\, \frac{q_x}{q},\, +1\right) \in N^0_{\mathcal{H}yp\,(V)}(t,x,V(t,x))$$

Hence, by (4.26) and *i)*,

$$\frac{q_t}{q} + H\left(t,\, x,\, \frac{q_x}{q}\right) \leq 0$$

and therefore $q_t + H(t,x,q_x) \leq 0$. If $q = 0$, applying Lemma 1.1.16 to the extended lower semicontinuous function $(s,y) \mapsto -V(s,y)$, we can find a sequence $(t_i,x_i) \to (t,x)$ and a sequence

$$\left(q_t^i,\, q_x^i,\, q^i\right) \in N^0_{\mathcal{H}yp\,(V)}(t,x,V(t,x))$$

such that $q^i > 0$ and (q_t^i, q_x^i) converge to (q_t, q_x). This and continuity of H yield (4.27).

We next deduce from (4.27) and the separation theorem that for all (t,x) such that $0 < t < T$ and all $z \leq V(t,x)$

$$\{1\} \times F(t,x) \times \{0\} \subset \overline{co}\left(T_{\mathcal{H}yp\,(V)}(t,x,z)\right)$$

This, Theorem 1.1.3 and lower semicontinuity of F imply that for all (t,x) satisfying $0 < t < T$

$$\begin{array}{l} \{1\} \times F(t,x) \times \{0\} \\ \subset \operatorname{Liminf}_{\substack{(t',x',z') \to (t,x,V(t,x)) \\ (t',x',z') \in \mathcal{H}yp\,(V)}} \overline{co}\left(T_{\mathcal{H}yp\,(V)}(t',x',z')\right) \\ \subset T_{\mathcal{H}yp\,(V)}(t,x,V(t,x)) \;=\; \mathcal{H}yp\,(D_{\downarrow}V(t,x)) \end{array}$$

Thus for all (t,x) satisfying $0 < t < T$,

$$\inf_{v \in F(t,x)} D_{\downarrow}V(t,x)(1,v) \;\geq\; 0$$

Define $W(t,x) = -V(T-t,x)$. Then for all (t,x) such that $0 < t < T$ and for all $v \in F(T-t,x)$, we have

$$D_{\uparrow}W(t,x)(-1,v) = -D_{\downarrow}V(T-t,x)(1,v) \;\leq\; 0$$

Applying Theorem 4.2.4 to W and the set-valued map

$$\widehat{F}(t,x) = -F(T-t,x)$$

we deduce that for every solution y to the inclusion

$$y'(t) \in \widehat{F}(t,y(t)) \;\text{ a.e. in } [t_0,t_1]$$

where $0 < t_0 \leq t_1 < T$ we have

$$\forall\, t \,\in\, [t_0,t_1], \quad W(t_0,x_0) \leq W(t,y(t))$$

Fix any $v \in F(t_0,x_0)$ and consider a solution $y(\cdot)$ to the differential inclusion

$$\left\{\begin{array}{l} y' \,\in\, \widehat{F}(t,y) \\ y(T-t_0) \;=\; x_0, \;\; y'(T-t_0) \;=\; -v \end{array}\right.$$

Then for all small $s > 0$,

$$W(T - t_0, x_0) \leq W(T - t_0 + s, y(T - t_0 + s))$$

and therefore for a sequence $v_s \to v$ we have

$$V(t_0 - s, x_0 - sv_s) \leq V(t_0, x_0)$$

This yields that $D_{\uparrow}V(t_0, x_0)(-1, -v) \leq 0$. Since $v \in F(t_0, x_0)$ is arbitrary, *ii*) follows. □

Let $V : [0, T] \times \mathbf{R}^n \mapsto \mathbf{R}$ be a continuous *viscosity solution* to (4.23). Then, by Theorems 4.3.1, 4.3.2 and Proposition 1.1.15, V verifies *iv*) of Theorem 4.1.5.

This completes the proof of Theorem 4.1.5.

Chapter 5

Value Function of Bolza Problem and Riccati Equations

Introduction

This chapter is concerned with the characteristics of the Hamilton-Jacobi equation

$$-\frac{\partial V}{\partial t} + H\left(t, x, -\frac{\partial V}{\partial x}\right) = 0, \quad V(T, \cdot) = g(\cdot) \tag{5.1}$$

i.e. solutions to the *Hamiltonian system*

$$\begin{cases} x'(t) = \dfrac{\partial H}{\partial p}(t, x(t), p(t)), & x(T) = x_T \\ \\ -p'(t) = \dfrac{\partial H}{\partial x}(t, x(t), p(t)), & p(T) = -\nabla g(x_T) \end{cases} \tag{5.2}$$

In Chapter 3 such system arises as "extremal equations" in optimal control, since the Pontryagin maximum principle states that if $x : [t_0, T] \mapsto \mathbf{R}^n$ is optimal for the Mayer problem and $\nabla g(x(T)) \neq 0$, then there exists $p : [t_0, T] \mapsto \mathbf{R}^n$ such that (x, p) solves (5.2) with $x_T = x(T)$. This is not however a sufficient condition for optimality because it may happen that to a given $x_0 \in \mathbf{R}^n$ correspond two distinct solutions (x_i, p_i), $i = 1, 2$ of (5.2) satisfying

$$x_i(t_0) = x_0 \tag{5.3}$$

and with one of x_i being not optimal. If the solution of (5.2) is unique for every $x_T \in \mathbf{R}^n$, then

$$p_1(t_0) \neq p_2(t_0) \tag{5.4}$$

Whenever (5.3) and (5.4) hold true for some solutions (x_i, p_i), $i = 1, 2$ of (5.2), we say that the system (5.2) has a *shock* at time t_0.

It was already observed in Chapter 3 that for the Mayer problem the Hamiltonian $H(t, x, \cdot)$ is not differentiable at zero. For this reason the system (5.2) is not well defined. In this chapter we study the Bolza problem:

$$\text{minimize} \int_{t_0}^{T} L(t, x(t), u(t))dt \ + \ g(x(T))$$

over solution-control pairs (x, u) of the control system

$$\begin{cases} x'(t) &= f(t, x(t), u(t)), \quad u(t) \in U \\ x(t_0) &= x_0 \end{cases}$$

The Hamiltonian for this problem is given by

$$H(t, x, p) = \sup_{u \in U}(\langle p, f(t, x, u)\rangle - L(t, x, u))$$

For a class of nonlinear control problems $H(t, \cdot, \cdot)$ is everywhere differentiable. We provide in Section 1 an example of such situation.

If shocks never occur on the time interval $[0, T]$, then the solution of (5.1) can be constructed by simply setting

$$V(t_0, x(t_0)) = g(x(T)) + \int_{t_0}^{T} L(t, x(t), u(t))dt$$

where x solves (5.2) for some $p(\cdot)$ and the control $u(t) \in U$ is so that $x'(t) = f(t, x(t), u(t))$ almost everywhere in $[t_0, T]$. Then V

is the value function of the above Bolza optimal control problem. Furthermore in this case V is continuously differentiable and

$$\frac{\partial V}{\partial x}(t, x(t)) = -p(t) \quad \& \quad \frac{\partial V}{\partial t}(t, x(t)) = H(t, x(t), p(t))$$

It is well known that shocks do happen. This is the very reason why the value function is not smooth and why one should not expect smooth solutions to the Hamilton-Jacobi-Bellman equation (5.1). Also it was shown in Chapter 4 that for the Mayer problem the value function is not smooth at some point (t_0, x_0), where the co-state is nondegenerate, if and only if the optimal trajectory is not unique. We shall show under what circumstances a similar statement holds true everywhere for the Bolza problem.

If we could guarantee that on some time interval $[t_0, T]$ there is no shocks, then the value function would be continuously differentiable on $[t_0, T] \times \mathbf{R}^n$ solution of (5.1). In the same time we have the uniqueness of optimal trajectories and obtain the optimal feedback low $G : [t_0, T] \times \mathbf{R}^n \rightsquigarrow U$ by setting

$$G(t, x) = \left\{ u \mid H(t, x, -\frac{\partial V}{\partial x}(t, x)) = \langle -\frac{\partial V}{\partial x}(t, x), f(t, x, u) \rangle - L(t, x, u) \right\}$$

Then the closed loop control system

$$x' = f(t, x, u(t, x)), \quad u(t, x) \in G(t, x), \quad x(t_0) = x_0$$

has exactly one solution and it is optimal for the Bolza problem.

Actually, when the data is smooth, the shocks would not occure till time t_0 if for every (x, p) solving (5.2) on $[t_0, T]$ the matrix Riccati equation

$$\left\{ \begin{array}{l} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t, x(t), p(t))P \; + \; P \, \dfrac{\partial^2 H}{\partial x \partial p}(t, x(t), p(t)) \; + \\[2ex] + \; P \, \dfrac{\partial^2 H}{\partial p^2}(t, x(t), p(t))P \; + \; \dfrac{\partial^2 H}{\partial x^2}(t, x(t), p(t)) \; = \; 0 \\[2ex] P(T) = -g''(x(T)) \end{array} \right. \tag{5.5}$$

has a solution on $[t_0, T]$.

In Section 1 we show that the existence of global solutions to Riccati equations (5.5) implies the absence of shocks. Section 2 is devoted to comparison theorems for solutions of (5.5). In Section 3 we relate the nonexistence of shocks to smoothness of the value function and uniqueness of optimal solutions and then apply the above results to problems with concave-convex Hamiltonians.

5.1 Matrix Riccati Equations and Shocks

In this section we relate the absence of shocks of the Hamilton-Jacobi-Bellman equation (5.1) with the existence of solutions to matrix Riccati equations (5.2).

Consider $H : [0,T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ and $\psi : \mathbf{R}^n \mapsto \mathbf{R}^n$. We assume that $H(t,\cdot,\cdot)$ is differentiable and associate to this data the Hamiltonian system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t,x(t),p(t)), \quad x(T) = x_T \\ \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t,x(t),p(t)), \quad p(T) = \psi(x_T) \end{cases} \tag{5.6}$$

Definition 5.1.1 *The system (5.6) has a* shock *at time t_0 if there exist two solutions $(x_i,p_i)(\cdot)$, $i = 1,2$ of (5.6) such that*

$$x_1(t_0) = x_2(t_0) \quad \& \quad p_1(t_0) \neq p_2(t_0)$$

Definition 5.1.2 *The* Hamiltonian system *(5.6) is called* complete *if for every x_T, the solution of (5.6) is defined on $[0,T]$ and depends continuously on the "initial" state in the following sense:*

Let (x_i,p_i) be solutions of (5.6) satisfying $x_i(t_i) \to x_0$, $p_i(t_i) \to p_0$ for some $t_i \to t_0$, $x_0 \in \mathbf{R}^n$, $p_0 \in \mathbf{R}^n$. Then (x_i,p_i) converge uniformly to the solution (x,p) of (5.6) such that $x(t_0) = x_0$ and $p(t_0) = p_0$.

Remark —

a) If the Hamiltonian system (5.6) is complete, then for all $t_0 \in [0,T]$, $x_0 \in \mathbf{R}^n$, $p_0 \in \mathbf{R}^n$ it has at most one solution (x,p) satisfying $x(t_0) = x_0$, $p(t_0) = p_0$.

b) The Hamiltonian system (5.6) is complete for instance if for all $r > 0$ there exists $k_r \in L^1(0,T)$ such that the mapping $\frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot)$ is $k_r(t)$-Lipschitz on $B_r(0)$ and has a linear growth: for some $\gamma \in L^1(0,T)$

$$\forall\, x,\ p \in \mathbf{R}^n, \quad \left\| \frac{\partial H}{\partial(x,p)}(t,x,p) \right\| \leq \gamma(t)\,(\|x\| + \|p\| + 1) \quad \square$$

Example — Consider

$$f : [0,T]\times\mathbf{R}^n \mapsto \mathbf{R}^n,\ g : [0,T]\times\mathbf{R}^n \mapsto L(U,\mathbf{R}^n),\ l : [0,T]\times\mathbf{R}^n \mapsto \mathbf{R}$$

where U is a finite dimensional space and let $R(t) \in L(U,U)$ be self-adjoint and positive for every $t \in [0,T]$. Define

$$H(t,x,p) = \langle p, f(t,x)\rangle + \sup_{u\in U}\left(\langle p, b(t,x)u\rangle - \frac{1}{2}\langle R(t)u,u\rangle \right) - l(t,x)$$

Then it is not difficult to check that

$$H(t,x,p) = \langle p, f(t,x)\rangle + \left\langle R(t)^{-1}b(t,x)^\star p,\ b(t,x)^\star p \right\rangle - l(t,x)$$

Thus an appropriate smoothness of $f(t,\cdot)$, $b(t,\cdot)$ and $l(t,\cdot)$ implies differentiability of $H(t,\cdot,\cdot)$ and completeness of the associated Hamiltonian system. $\square$

Theorem 5.1.3 *Assume that ψ is locally Lipschitz, that $H(t,\cdot,\cdot)$ is twice continuously differentiable and that for every $r > 0$, there exists $k_r \in L^1(0,T)$ satisfying*

$$\frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \text{ is } k_r(t) - \text{Lipschitz on } B_r(0)$$

Further assume the Hamiltonian system (5.6) is complete and define for every $t \in [0,T]$ the set

$$M_t = \{(x(t),p(t)) \mid (x,p) \text{ solves (5.6) for some } x_T \in \mathbf{R}^n\}$$

Then the following two statements are equivalent:

i) $\forall\, t \in [0,T]$, M_t is the graph of a locally Lipschitz function *from an open set* $\mathcal{D}(t)$ *into* $\mathbf{R}^n$

ii) $\forall\,(x,p)$ *solving (5.6) on* $[0,T]$ *and* $P_T \in \partial^\star\psi(x(T))$, *the matrix Riccati equation*

$$\begin{cases} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t,x(t),p(t))P \;+\; P\,\dfrac{\partial^2 H}{\partial x \partial p}(t,x(t),p(t)) \;+ \\[2ex] +\;P\,\dfrac{\partial^2 H}{\partial p^2}(t,x(t),p(t))P \;+\; \dfrac{\partial^2 H}{\partial x^2}(t,x(t),p(t)) \;=\; 0 \\[2ex] P(T) = P_T \end{cases} \tag{5.7}$$

has a solution on $[0,T]$.

Furthermore, if i) (or equivalently ii)) holds true, then

$$\psi \text{ is differentiable } \implies M_t \text{ is the graph of a differentiable function}$$

$$\psi \in C^1 \implies M_t \text{ is the graph of a } C^1 - \text{function}$$

Corollary 5.1.4 *Under all assumptions of Theorem 5.1.3, suppose that for every* (x,p) *solving (5.6) on* $[0,T]$ *and* $P_T \in \partial^\star\psi(x(T))$, *the matrix Riccati equation (5.7) has a solution on* $[0,T]$. *Then the Hamiltonian system (5.6) has no shocks on* $[0,T]$.

To prove the above theorem the following lemma is needed.

Lemma 5.1.5 *Assume that the Hamiltonian system (5.6) is complete and for every* $r > 0$, *there exists* $k_r \in L^1(0,T)$ *such that*

$$\frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \text{ is } k_r(t) - \text{Lipschitz on } B_r(0)$$

Let $K \subset \mathbf{R}^n$ *be a compact set. Consider a locally Lipschitz function* $\psi : \mathbf{R}^n \mapsto \mathbf{R}^n$ *and the subsets* $M_t(K)$, $t \in [0,T]$ *defined by*

$$M_t(K) = \{(x(t),p(t)) \mid (x,p) \text{ solves } (5.6),\; x_T \in K\}$$

Then there exists $\delta > 0$ *such that for all* $t \in [T-\delta, T]$, $M_t(K)$ *is the graph of a Lipschitz function.*

Proof — From the completeness of (5.6) we deduce that the subsets $M_t(K)$ are compact and contained in the ball $B_r(0)$ for some $r > 0$. Set $k(t) = k_r(t)$

We proceed by a contradiction argument. Assume for a moment that there exist $t_i \to T-$ such that $M_{t_i}(K)$ is not the graph of a Lipschitz function. Then for every i we can find two distinct solutions (x_j^i, p_j^i), $j = 1, 2$ of the Hamiltonian system (5.6) such that

$$\varepsilon_i \;:=\; \frac{\|x_1^i(t_i) - x_2^i(t_i)\|}{\|p_1^i(t_i) - p_2^i(t_i)\|} \;\to\; 0 \;\text{ as }\; i \to +\infty$$

Since for every $s \in [t_i, T]$ we have

$$\left\|x_1^i(s) - x_2^i(s)\right\| \leq$$

$$\varepsilon_i\|p_1^i(t_i) - p_2^i(t_i)\| + \textstyle\int_{t_i}^s k(\tau)(\left\|x_1^i(\tau) - x_2^i(\tau)\right\| + \left\|p_1^i(\tau) - p_2^i(\tau)\right\|)d\tau$$

the Gronwall lemma implies that for some $C > 0$ independent from i and for all $s \in [t_i, T]$

$$\left\|x_1^i(s) - x_2^i(s)\right\| \leq C(\varepsilon_i\|p_1^i(t_i) - p_2^i(t_i)\| + \int_{t_i}^s k(\tau)\|p_1^i(\tau) - p_2^i(\tau)\|d\tau)$$

Hence for some $C_1 > 0$ and all i large enough and $s \in [t_i, T]$,

$$\left\|p_1^i(s) - p_2^i(s)\right\| \leq$$

$$\left\|p_1^i(t_i) - p_2^i(t_i)\right\| + \textstyle\int_{t_i}^s k(\tau)(\left\|x_1^i(\tau) - x_2^i(\tau)\right\| + \left\|p_1^i(\tau) - p_2^i(\tau)\right\|)d\tau$$

$$\leq C_1 \left\|p_1^i(t_i) - p_2^i(t_i)\right\| + C_1 \textstyle\int_{t_i}^s k(\tau) \left\|p_1^i(\tau) - p_2^i(\tau)\right\| d\tau$$

From the Gronwall lemma we deduce that for some $L > 0$ independent from i and all $s \in [t_i, T]$,

$$\left\|p_1^i(s) - p_2^i(s)\right\| \leq L \left\|p_1^i(t_i) - p_2^i(t_i)\right\|$$

This implies that

$$\overline{\varepsilon}_i \;:=\; \sup_{s\in[t_i,T]} \frac{\left\|x_1^i(s) - x_2^i(s)\right\|}{\left\|p_1^i(t_i) - p_2^i(t_i)\right\|} \;\text{ converge to zero} \qquad (5.8)$$

We next observe that for all $s \in [t_i, T]$,

$$\|p_1^i(s) - p_2^i(s)\| \leq$$

$$\leq \|p_1^i(T) - p_2^i(T)\| + \int_s^T k(\tau)(\|x_1^i(\tau) - x_2^i(\tau)\| + \|p_1^i(\tau) - p_2^i(\tau)\|)d\tau$$

$$\leq \|p_1^i(T) - p_2^i(T)\| + \int_s^T k(\tau)(\|p_1^i(\tau) - p_2^i(\tau)\| + \bar{\varepsilon}_i \|p_1^i(t_i) - p_2^i(t_i)\|)d\tau$$

Applying again the Gronwall lemma and taking i large enough we get

$$\left\|p_1^i(t_i) - p_2^i(t_i)\right\| \leq L_1 \left\|p_1^i(T) - p_2^i(T)\right\| + \frac{1}{2}\left\|p_1^i(t_i) - p_2^i(t_i)\right\|$$

for some L_1 independent from i. Hence for all large i

$$\left\|p_1^i(t_i) - p_2^i(t_i)\right\| \leq 2L_1 \left\|p_1^i(T) - p_2^i(T)\right\|$$

and therefore, by (5.8),

$$\frac{\|x_1^i(T) - x_2^i(T)\|}{\|p_1^i(T) - p_2^i(T)\|} = \frac{\|x_1^i(T) - x_2^i(T)\|}{\|p_1^i(t_i) - p_2^i(t_i)\|} \times \frac{\|p_1^i(t_i) - p_2^i(t_i)\|}{\|p_1^i(T) - p_2^i(T)\|} \to 0$$

Thus

$$\frac{\|\psi(x_1^i(T)) - \psi(x_2^i(T))\|}{\|x_1^i(T) - x_2^i(T)\|} = \frac{\|p_1^i(T) - p_2^i(T)\|}{\|x_1^i(T) - x_2^i(T)\|} \to +\infty$$

which contradicts the Lipschitz continuity of ψ on K. □

Proof of Theorem 5.1.3 — Assume first that for all $t \in [0, T]$, M_t is the graph of a locally Lipschitz function. Consider a solution (x, p) of (5.6) and the linear system

$$\begin{cases} U' = \dfrac{\partial^2 H}{\partial x \partial p}(t, x(t), p(t))U + \dfrac{\partial^2 H}{\partial p^2}(t, x(t), p(t))V \\ -V' = \dfrac{\partial^2 H}{\partial x^2}(t, x(t), p(t))U + \dfrac{\partial^2 H}{\partial p \partial x}(t, x(t), p(t))V \\ U(T) = Id, \ V(T) = P_T \end{cases}$$

where $U,\ V : [0,T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$ are matrix functions and $P_T \in \partial^\star \psi(x(T))$. Let (x_n, p_n) be solutions of (5.6) such that

$$\lim_{n\to\infty} x_n(T) = x(T) \quad \& \quad \lim_{n\to\infty} \psi'(x_n(T)) = P_T$$

By completeness of (5.6), (x_n, p_n) converge uniformly to (x, p).

The variational equation implies that for any $(w(\cdot), q(\cdot))$ solving

$$\left\{\begin{array}{lll} w' & = \dfrac{\partial^2 H}{\partial x \partial p}(t, x_n(t), p_n(t))w \ + \ \dfrac{\partial^2 H}{\partial p^2}(t, x_n(t), p_n(t))q \\ \\ -q' & = \dfrac{\partial^2 H}{\partial x^2}(t, x_n(t), p_n(t))w \ + \ \dfrac{\partial^2 H}{\partial p \partial x}(t, x_n(t), p_n(t))q \\ \\ w(T) & = w_T, \ \ q(T) = \psi'(x_n(T))w_T \end{array}\right. \tag{5.9}$$

we have $(w(t), q(t)) \in T_{M_t}(x_n(t), p_n(t))$ (contingent cone to M_t at $(x_n(t), p_n(t))$). Because M_t is the graph of a locally Lipschitz function, for some l_t independent from n, $\|q(t)\| \le l_t \|w(t)\|$.

Taking the limit in (5.9) we deduce that every solution (w, q) of

$$\left\{\begin{array}{lll} w' & = & \dfrac{\partial^2 H}{\partial x \partial p}(t, x(t), p(t))w \ + \ \dfrac{\partial^2 H}{\partial p^2}(t, x(t), p(t))q \\ \\ -q' & = & \dfrac{\partial^2 H}{\partial x^2}(t, x(t), p(t))w \ + \ \dfrac{\partial^2 H}{\partial p \partial x}(t, x(t), p(t))q \\ \\ w(T) & = & w_T, \ \ q(T) = P_T w_T \end{array}\right. \tag{5.10}$$

satisfies $\|q(t)\| \le l_t \|w(t)\|$.

Thus, by uniqueness of solution to (5.10), if $w_T \neq 0$, then $w(\cdot)$ never vanishes. Since

$$w(t) = U(t)w_T \quad \& \quad q(t) = V(t)w_T$$

this implies that $U(t)$ is not singular for all $t \in [0, T]$. Setting

$$P(t) = V(t)U(t)^{-1}$$

we check that P satisfies (5.7).

Conversely let (5.7) have a solution on $[0,T]$ for all (x,p) solving (5.6). For every $r > 0$, $t \in [0,T]$ consider the compact sets

$$\Pi_{rt} = \{(x(t), p(t)) \mid (x,p) \text{ solves } (5.6),\ x(T) \in B_r(0)\}$$

We first claim that for every $r > 0$ and $t_0 \in [0,T]$, Π_{rt_0} is the graph of a Lipschitz function. Indeed fix r, t_0 as above and assume for a moment that Π_{rt_0} is not the graph of a Lipschitz function.

By Lemma 5.1.5 for all s near T, Π_{rs} is still the graph of a Lipschitz function. Define

$$\bar{t} = \inf_{t\in[t_0,T]} \{\ \forall\, s \in [t,T],\ \Pi_{rs} \text{ is the graph of a Lipschitz function}\}$$

Then $t_0 \leq \bar{t} < T$ and $\Pi_{r\bar{t}}$ is not the graph of a Lipschitz function, because otherwise, by Lemma 5.1.5, we could make $\bar{t}$ smaller which would contradict its choice. Define the sets

$$D_r(s) = \{x(s) \mid (x,p) \text{ solves } (5.6),\ \|x(T)\| < r\}$$

Observe that for all $r > 0$ and $s \in]\bar{t}, T]$, $D_r(s)$ is open. Its closure is equal to the set

$$\overline{D_r(s)} = \{x(s) \mid (x,p) \text{ solves } (5.6),\ x(T) \in B_r(0)\}$$

by completeness of (5.6).

Define next the Lipschitz function $\Phi_{rs} : \overline{D_r(s)} \mapsto \mathbf{R}^n$ by

$$\text{Graph}(\Phi_{rs}) = \Pi_{rs}$$

The Rademacher theorem yields Φ_{rs} is differentiable almost everywhere on $D_r(s)$.

Fix a sequence $t_n \to \bar{t}+$ and observe that the family $\{\Phi_{rt_n}\}_{n\geq 1}$ can not be equilipschitz, because otherwise, using that

$$\Pi_{r\bar{t}} = \text{Lim}_{n\to\infty}\Pi_{rt_n}$$

we would deduce that $\Pi_{r\bar{t}}$ is the graph of a Lipschitz function. Thus there exists a sequence $\overline{x}_n \in D_r(t_n)$ such that $\Phi'_{rt_n}(\overline{x}_n) \to \infty$. Hence

$$\exists\ (u_n, v_n) \in \mathbf{R}^n \times \mathbf{R}^n \text{ satisfying } \Phi'_{t_n}(\overline{x}_n)u_n = v_n,\ \|v_n\| = 1,\ \|u_n\| \to 0$$

Let (x_n, p_n) be a solution of (5.6) such that $x_n(t_n) = \overline{x}_n$ and $p_n(t_n) = \Phi_{rt_n}(\overline{x}_n)$. Since Φ_{rt_n} is differentiable at $\overline{x}_n$, using variational equation, we deduce that ψ is differentiable at $x_n(T)$. Taking a subsequence and keeping the same notations, by completeness of (5.6), we may assume that (x_n, p_n) converge uniformly to a solution (x, p) of (5.6) and for some $P_T \in \partial^\star \psi(x(T))$

$$v_n \to v, \quad \psi'(x_n(T)) \to P_T$$

Consider next the solutions (w_n, q_n) of

$$\begin{cases} w' &= \dfrac{\partial^2 H}{\partial x \partial p}(t, x_n(t), p_n(t))w \;+\; \dfrac{\partial^2 H}{\partial p^2}(t, x_n(t), p_n(t))q \\[2ex] -q' &= \dfrac{\partial^2 H}{\partial x^2}(t, x_n(t), p_n(t))w \;+\; \dfrac{\partial^2 H}{\partial p \partial x}(t, x_n(t), p_n(t))q \\[2ex] w(t_n) &= u_n, \;\; q(t_n) = v_n \end{cases}$$

The variational equation yields $q_n(T) = \psi'(x_n(T))w_n(T)$.

Since $\lim_{n\to\infty}(u_n, v_n) = (0, v)$, passing to the limit in the above system, we deduce that (5.10) has a solution (w, q) satisfying

$$w(\overline{t}) = 0, \quad q(\overline{t}) \neq 0, \quad q(T) = P_T w(T)$$

In particular $w(T) \neq 0$ and $U(\overline{t})w(T) = 0$. On the other hand, by the previous arguments, $P(t) = V(t)U(t)^{-1}$ solves (5.7) on $]\overline{t}, T]$. If P is well defined on $[\overline{t}, T]$, then $V(\overline{t}) = P(\overline{t})U(\overline{t})$ and $q(\overline{t}) = V(\overline{t})w(T) = 0$, which leads to a contradiction and proves our claim.

Observe next that for every $s \in [0, T]$ the sequence of open subsets $\{D_r(s)\}_{r>0}$ is nondecreasing. Define the open set

$$\mathcal{D}(s) = \bigcup_{k>0} D_k(s)$$

Then

$$\mathcal{D}(s) = \{x \mid \exists\, p \text{ such that } (x, p) \in M_s\}$$

Since $\{\Pi_{rs}\}_{r>0}$ is a nondecreasing sequence of graphs of Lipschitz functions, $M_s = \bigcup_{r>0} \Pi_{rs}$ is the graph of a function from $\mathcal{D}(s)$ into $\mathbf{R}^n$.

We next show that M_s is the graph of a locally Lipschitz function. Indeed fix $\overline{x} \in \mathcal{D}(s)$, $r > 0$ such that $B_r(\overline{x}) \subset \mathcal{D}(s)$. Since $B_r(\overline{x})$ is compact and the family of open sets $D_r(s)$ is nondecreasing, for some $k > 0$, $B_r(\overline{x}) \subset D_k(s)$. But we already know that $M_s \cap D_k(s) \times \mathbf{R}^n$ is the graph of a Lipschitz function.

The last two statements follow from the variational equation.

5.2 Matrix Riccati Equations

We investigate here matrix differential equations of the following type

$$P' + A(t)^\star P + PA(t) + PE(t)P + D(t) = 0, \quad P(T) = P_T$$

5.2.1 Comparison Theorems

The aim of this section is to provide two comparison properties for solutions of Riccati equations.

Theorem 5.2.1 *Let A, E_i, $D_i : [0,T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$, $i = 1, 2$ be integrable. We assume that $E_1(t)$ and $D_1(t)$ are self-adjoint for almost every $t \in [0,T]$ and*

$$D_1(t) \leq D_2(t), \quad E_1(t) \leq E_2(t) \text{ a.e. in } [0,T] \tag{5.11}$$

Consider self-adjoint operators $P_{iT} \in L(\mathbf{R}^n, \mathbf{R}^n)$ such that

$$P_{1T} \leq P_{2T}$$

and solutions $P_i(\cdot) : [t_0, T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$ to the matrix equations

$$P' + A(t)^\star P + PA(t) + PE_i(t)P + D_i(t) = 0, \quad P_i(T) = P_{iT} \tag{5.12}$$

for $i = 1, 2$. If P_2 is self-adjoint, then $P_1 \leq P_2$ on $[t_0, T]$.

Proof — From uniqueness of solution to (5.12), using that $E_1(t)$ and $D_1(t)$ are self-adjoint, it is not difficult to deduce that P_1 is self-adjoint. For all $t \in [t_0, T]$, set

$$Z = P_2 - P_1, \quad \mathcal{A}(t) = A(t) + \frac{1}{2}E_1(t)(P_1(t) + P_2(t))$$

Then

$$\mathcal{A}(t)^\star Z(t) + Z(t)\mathcal{A}(t) =$$

$$= A(t)^\star Z(t) + Z(t)A(t) - P_1(t)E_1(t)P_1(t) + P_2(t)E_1(t)P_2(t)$$

Therefore Z solves the Riccati equation

$$Z' + \mathcal{A}(t)^\star Z + Z\mathcal{A}(t) + P_2(t)(E_2(t) - E_1(t))P_2(t) + D_2(t) - D_1(t) = 0$$

Denote by $X(\cdot, t)$ the solution to

$$X' = -\mathcal{A}(s)^\star X, \quad X(t,t) = Id$$

A direct verification yields

$$Z(t) = X(t,T)(P_{2T} - P_{1T})X(t,T)^\star +$$

$$+\int_t^T X(t,s)(D_2(s) - D_1(s) + P_2(s)(E_2(s) - E_1(s))P_2(s))X(t,s)^\star ds$$

This and assumptions (5.11) imply $Z \geq 0$ on $[t_0, T]$. □

Theorem 5.2.2 *Let $A, E_i, D_i : [0,T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$, $i = 1, 2$ be integrable. We assume that $E_1(t)$, $D_1(t)$ are self-adjoint for almost all $t \in [0,T]$ and*

$$D_1(t) \leq D_2(t), \;\; 0 \leq E_1(t) \leq E_2(t) \;\; \text{a.e. in} \;\; [0,T]$$

Consider self-adjoint operators $P_{iT} \in L(\mathbf{R}^n, \mathbf{R}^n)$ such that $P_{1T} \leq P_{2T}$ and solutions $P_i(\cdot) : [t_i, T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$ to the matrix equations

$$P' + A(t)^\star P + PA(t) + PE_i(t)P + D_i(t) = 0, \quad P_i(T) = P_{iT}$$

where $i = 1, 2$. If P_2 is self-adjoint, then the solution P_1 is defined at least on $[t_2, T]$ and $P_1 \leq P_2$.

Proof — Consider the square root $B(t)$ of $E_1(t)$, i.e. for almost every $t \in [0,T]$, $E_1(t) = B(t)B(t)^\star$ and set

$$t_0 = \inf_{t \in [0,T]} \{P_1 \;\text{ is defined on }\; [t,T]\}$$

Thus either the solution P_1 exists on $[0,T]$ or $\|P_1(t)\| \to \infty$ when $t \to t_0+$. It is enough to show that if $t_2 \le t_0$, then P_1 is bounded on $]t_0,T]$. So let us assume that $t_2 \le t_0$. By Theorem 5.2.1 for every $t_0 < t \le T$ we have $P_1(t) \le P_2(t)$. Pick any $x \in \mathbf{R}^n$ of norm one. Since $P_1 = P_1^\star$ we get

$$\langle B(t)^\star P_1(t)x, B(t)^\star P_1(t)x\rangle =$$

$$-\langle P_1'(t)x,x\rangle - \langle A(t)^\star P_1(t)x,x\rangle - \langle P_1(t)A(t)x,x\rangle - \langle D_1(t)x,x\rangle$$

Therefore for every $x \in \mathbf{R}^n$ of norm one and all $t_0 < t \le T$

$$\int_t^T \|B(s)^\star P_1(s)x\|^2\, ds \le$$

$$\le -\int_t^T < P_1'(s)x,x > +2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds + \|D_1\|_{L^1(t,T)}$$

$$\le \ \langle P_1(t)x,x\rangle + \|P_{1T}\| + 2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds + \|D_1\|_{L^1(t,T)}$$

$$\le \ \|P_2(t)\| + 2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds + \|P_{1T}\| + \|D_1\|_{L^1(t,T)}$$

$$\le c + 2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds$$

for some c independent from t, because P_2 is bounded on $[t_2,T]$.

On the other hand for any $y \in \mathbf{R}^n$ of norm one

$$-\langle P_1'(t)x,y\rangle \ = \ \langle P_1(t)B(t)B(t)^\star P_1(t)x,y\rangle \ + \langle A(t)^\star P_1(t)x,y\rangle +$$

$$+\langle P_1(t)A(t)x,y\rangle + \langle D_1(t)x,y\rangle$$

Integrating on $[t,T]$ and using the latter inequality and the Hölder inequality, we obtain

$$\langle P_1(t)x,y\rangle \le \|P_{1T}\| + \|B^\star(\cdot)P_1(\cdot)x\|_{L^2(t,T)}\,\|B^\star(\cdot)P_1(\cdot)y\|_{L^2(t,T)} +$$

$$+2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds + \|D_1\|_{L^1(t,T)}$$

$$\le \ c_1 + 2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds + \left[\left(c + 2\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds\right)^{1/2}\right]^2$$

for some c_1 independent from t. Since this holds true for all x and $y \in \mathbf{R}^n$ of norm one,

$$\forall\, t_0 < t \leq T, \quad \|P_1(t)\| \ \leq \ c + c_1 + 4\int_t^T \|A(s)\|\,\|P_1(s)\|\,ds$$

Applying the Gronwall lemma we deduce that $\|P_1(t)\|$ is bounded on $]t_0, T]$ by a constant independent from t.

5.2.2 Existence of Solutions

We deduce from the previous section sufficient conditions for existence of solutions to the matrix Riccati equations.

Theorem 5.2.3 *Let $A, E, D : [0,T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$ be integrable. We assume that $E(t)$, $D(t)$ are self-adjoint and $E(t) \geq 0$ for almost every $t \in [0,T]$. Consider a self-adjoint operator $P_T \in L(\mathbf{R}^n, \mathbf{R}^n)$ and assume that there exists an absolutely continuous $P : [t_0, T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$ such that for every $t \in [t_0, T]$, $P(t)$ is self-adjoint and*

$$P'(t) + A(t)^\star P(t) + P(t)A(t) + P(t)E(t)P(t) + D(t) \leq 0 \text{ a.e. in } [t_0, T]$$

and $P_T \leq P(T)$. Then the solution $\overline{P}$ to the equation

$$P' + A(t)^\star P + PA(t) + PE(t)P + D(t) = 0, \quad P(T) = P_T \qquad (5.13)$$

is defined at least on $[t_0, T]$ and $\overline{P} \leq P$ on $[t_0, T]$.

Proof — Set

$$\Gamma(t) = P'(t) + A(t)^\star P(t) + P(t)A(t) + P(t)E(t)P(t) + D(t)$$

Then $\Gamma(t) \leq 0$ and is self-adjoint and P solves the Riccati equation

$$P' + A(t)^\star P + PA(t) + PE(t)P + D(t) - \Gamma(t) = 0$$

where $D(t) - \Gamma(t) \geq D(t)$. By Theorem 5.2.2, $\overline{P}$ is defined at least on $[t_0, T]$ and $\overline{P} \leq P$. □

Corollary 5.2.4 *Under all assumptions on A, E, D of Theorem 5.2.3 consider a self-adjoint nonpositive $P_T \in L(\mathbf{R}^n, \mathbf{R}^n)$. If for almost all $t \in [0,T]$, $D(t) \leq 0$, then the solution $\overline{P}$ to the matrix Riccati equation (5.13) is well defined on $[0,T]$ and $\overline{P} \leq 0$.*

Proof — We apply Theorem 5.2.3 with $P(\cdot) \equiv 0$. □

5.3 Value Function of Bolza Problem

Consider the minimization problem

$$(P) \qquad \text{minimize} \int_{t_0}^{T} L(t, x(t), u(t))dt \; + \; g(x(T))$$

over solution-control pairs (x, u) of the control system

$$\begin{cases} x'(t) &= f(t, x(t), u(t)), \quad u(t) \in U \\ x(t_0) &= x_0 \end{cases} \tag{5.14}$$

where $t_0 \in [0, T]$, $x_0 \in \mathbf{R}^n$, U is a complete separable metric space,

$$g : \mathbf{R}^n \mapsto \mathbf{R}, \;\; L : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}, \;\; f : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}^n$$

The *Hamiltonian* $H : [0, T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ is defined by

$$H(t, x, p) = \sup_{u \in U} (\langle p, f(t, x, u) \rangle - L(t, x, u))$$

We denote by $\mathcal{U}$ the set of all measurable controls $u : [0, T] \mapsto U$ and by $x(\cdot; t_0, x_0, u)$ the solution of (5.14) starting at time t_0 from the initial condition x_0 and corresponding to the control $u(\cdot) \in \mathcal{U}$. Of course not to every $u \in \mathcal{U}$ corresponds a solution $x(\cdot; t_0, x_0, u)$ of (5.14).

For all $(t_0, x_0, u) \in [0, T] \times \mathbf{R}^n \times \mathcal{U}$ set

$$\Phi(t_0, x_0, u) \;=\; \int_{t_0}^{T} L(t, x(t; t_0, x_0, u), u(t))dt + g(x(T; t_0, x_0, u))$$

if this expression is well defined and $\Phi(t_0, x_0, u) = +\infty$ otherwise.

The value function associated to the Bolza problem (P) is defined by

$$V(t_0, x_0) \;=\; \inf_{u \in \mathcal{U}} \Phi(t_0, x_0, u)$$

when (t_0, x_0) range over $[0, T] \times \mathbf{R}^n$.

Proposition 5.3.1 *Assume that $H(t, \cdot, \cdot)$ is differentiable. Then*

$$\frac{\partial H}{\partial p}(t, x, p) = \{f(t, x, u) \mid \langle p, f(t, x, u) \rangle - L(t, x, u) = H(t, x, p)\}$$

and

$$\frac{\partial H}{\partial x}(t,x,p) = \left\{\frac{\partial f}{\partial x}(t,x,u)^\star p - \frac{\partial L}{\partial x}(t,x,u) \mid \langle p, f(t,x,u)\rangle - L(t,x,u) = H(t,x,p)\right\}$$

Proof — By Proposition 3.3.3 applied to the Hamiltonian

$$\mathcal{H}(t,x,(p,q)) = \sup_{u\in U} \langle (f(t,x,u),\ L(t,x,u)),\ (p,q)\rangle$$

at $(p,q) = (p,-1)$ we get

$$\frac{\partial H}{\partial p}(t,x,p) = \frac{\partial \mathcal{H}}{\partial p}(t,x,(p,-1)) =$$

$$= \{f(t,x,u) \mid \langle p, f(t,x,u)\rangle - L(t,x,u) = H(t,x,p)\}$$

and

$$\frac{\partial H}{\partial x}(t,x,p) = \frac{\partial \mathcal{H}}{\partial x}(t,x,(p,-1)) =$$

$$\left\{\tfrac{\partial f}{\partial x}(t,x,u)^\star p - \tfrac{\partial L}{\partial x}(t,x,u) \mid \langle p, f(t,x,u)\rangle - L(t,x,u) = H(t,x,p)\right\}$$

□

Consider the Hamiltonian system

$$\begin{cases} x'(t) = \dfrac{\partial H}{\partial p}(t,x(t),p(t)), & x(T) = x_T \\[2ex] -p'(t) = \dfrac{\partial H}{\partial x}(t,x(t),p(t)), & p(T) = -\nabla g(x_T) \end{cases} \tag{5.15}$$

Throughout the whole section we impose the following hypothesis:

$\mathbf{H}_1$) $f,\ L$ are continuous and $\forall\, r > 0,\ \exists\, k_r \in L^1(0,T)$ such that $\forall\, u \in U,\ \ (f(t,\cdot,u), L(t,\cdot,u))$ is $k_r(t)$ – Lipschitz on $B_r(0)$

$\mathbf{H}_2$) $f(t,\cdot,u),\ L(t,\cdot,u)$ are differentiable and $g \in \mathcal{C}^1$

$\mathbf{H}_3$) H and $\frac{\partial H}{\partial p}$ are continuous on $[0,T]\times \mathbf{R}^n \times \mathbf{R}^n$

$\mathbf{H}_4$) The Hamiltonian system (5.15) is complete

$\mathbf{H}_5$) For all $(t,x) \in [0,T]\times\mathbf{R}^n$, the set $\{(f(t,x,u), L(t,x,u)+r) \mid u \in U,\ r \geq 0\}$ is closed and convex

5.3.1 Maximum Principle

As in Chapter 3 to study differentiability of the value function we shall use the maximum principle:

Theorem 5.3.2 *Assume $H_1), H_2)$ and let $(\overline{x}, \overline{u})$ be an optimal solution-control pair of (P) for some $(t_0, x_0) \in [0,T] \times \mathbf{R}^n$. If $H(t,\cdot,\cdot)$ is differentiable, then there exists $p : [t_0, T] \mapsto \mathbf{R}^n$ such that $(\overline{x}, p)$ solves the Hamiltonian system*

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \quad x(t_0) = x_0 \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \quad p(T) = -\nabla g(\overline{x}(T)) \\ p(t_0) &\in -\partial_+ V_x(t_0, x_0) \end{cases} \tag{5.16}$$

where $\partial_+ V_x(t_0, x_0)$ denotes the superdifferential of $V(t_0, \cdot)$ at x_0.
Consequently for almost all $t \in [t_0, T]$,

$$H(t, \overline{x}(t), p(t)) = \langle p(t), \overline{x}'(t)\rangle - L(t, \overline{x}(t), \overline{u}(t))$$

Proof — Fix $v \in \mathbf{R}^n$ and let $h_k \to 0+, v_k \to v$ be such that

$$\begin{aligned} D_\downarrow V_x(t_0, x_0)(v) &:= \limsup_{h\to 0+,\, v'\to v} \frac{V(t_0, x_0 + hv') - V(t_0, x_0)}{h} \\ &= \lim_{k\to\infty} \frac{V(t_0, x_0 + h_k v_k) - V(t_0, x_0)}{h_k} \end{aligned}$$

For all k large enough consider the solution $x_k(\cdot)$ of the system

$$\begin{cases} x'(t) &= f(t, x(t), \overline{u}(t)) \\ x(t_0) &= x_0 + h_k v_k \end{cases}$$

The variational equation implies that the sequence $(x_k - \overline{x})/h_k$ converges to the solution $w(\cdot)$ of the linear system

$$w'(t) = \frac{\partial f}{\partial x}(t, \overline{x}(t), \overline{u}(t))w(t), \quad w(t_0) = v$$

Let $X(\cdot)$ denote the fundamental solution of

$$X'(t) = \frac{\partial f}{\partial x}(t, \overline{x}(t), \overline{u}(t))X(t), \quad X(t_0) = Id$$

Then $w(t) = X(t)v$ for all $t \in [t_0, T]$. Thus

$$D_{\downarrow}V_x(t_0, x_0)(v) \leq$$

$$\limsup_{k\to\infty} \frac{\int_{t_0}^T (L(t, x_k(t), \overline{u}(t)) - L(t, \overline{x}(t), \overline{u}(t)))dt + g(x_k(T)) - g((\overline{x}(T))}{h_k}$$

$$= \left\langle \int_{t_0}^T X(t)^\star \frac{\partial L}{\partial x}(t, \overline{x}(t), \overline{u}(t))dt + X(T)^\star \nabla g(\overline{x}(T)), v \right\rangle$$

Consider the solution $p(\cdot)$ to the adjoint system

$$\begin{cases} -p' &= \frac{\partial f}{\partial x}(t, \overline{x}(t), \overline{u}(t))^\star p - \frac{\partial L}{\partial x}(t, \overline{x}(t), \overline{u}(t)) \\ p(T) &= -\nabla g(\overline{x}(T)) \end{cases}$$

Then

$$p(t) = -X(t)^{\star^{-1}} \left(X(T)^\star \nabla g((\overline{x}(T)) + \int_t^T X(s)^\star \frac{\partial L}{\partial x}(s, \overline{x}(s), \overline{u}(s))ds \right)$$

Consequently, for all $v \in \mathbf{R}^n$,

$$D_{\downarrow}V_x(t_0, x_0)(v) \leq \langle -p(t_0), v\rangle$$

and so $p(t_0) \in -\partial_+ V_x(t_0, x_0)$. By the maximum principle for a.e. $t \in [t_0, T]$,

$$\langle p(t), f(t, \overline{x}(t), \overline{u}(t))\rangle - L(t, \overline{x}(t), \overline{u}(t)) \;=\; H(t, \overline{x}(t), p(t))$$

Since $H(t, \cdot, \cdot)$ is differentiable, we deduce from Proposition 5.3.1 that $(\overline{x}, p)$ solves the Hamiltonian system (5.16).

5.3.2 Differentiability of Value Function and Uniqueness of Optimal Solutions

We shall need the following consequence of the maximum principle.

Theorem 5.3.3 *Assume $H_1) - H_5)$, that V is locally Lipschitz and for every $(t_0, x_0) \in [0,T] \times \mathbf{R}^n$ the problem (P) has an optimal solution. Then for every*

$$\overline{p} \in \partial_x^\star V(t_0, x_0) \; := \; \mathrm{Limsup}_{x_i \to x_0,\, t_i \to t_0} \left\{ \frac{\partial V}{\partial x}(t_i, x_i) \right\}$$

there exists a solution (x, p) of (5.15) satisfying

$$x(t_0) = x_0 \;\; \& \;\; p(t_0) = \overline{p}$$

and x is optimal for problem (P).

In particular if (P) has a unique optimal trajectory $z(\cdot)$, then the set $\partial_x^\star V(t_0, x_0)$ is a singleton. Consequently, $V(t_0, \cdot)$ is differentiable at x_0.

Remark — Various sufficient conditions for local Lipschitz continuity of the value function and for the existence of optimal controls for (P) may be found in many books. They can also be deduced from results of chapter 1. We shall not dwell on this question in this chapter. □

Proof — Let $\overline{p} \in \partial_x^\star V(t_0, x_0)$ and $(t_k, x_k) \to (t_0, x_0)$ be such that

$$\lim_{k \to \infty} \frac{\partial V}{\partial x}(t_k, x_k) \; = \; \overline{p}$$

Consider optimal solution-control pairs (z_k, u_k) of (P) with (t_0, x_0) replaced by (t_k, x_k). From Theorem 5.3.2 it follows that there exist absolutely continuous functions p_k such that for all k, (z_k, p_k) solves the following problem

$$\begin{cases} x'(t) \;=\; \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \;\; x(t_k) = x_k, \;\; p(t_k) = -\dfrac{\partial V}{\partial x}(t_k, x_k) \\[2ex] -p'(t) \;=\; \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \;\; p(T) = -\nabla g((z_k(T)) \end{cases}$$

We extend (z_k, p_k) on the time interval $[0, t_k]$ as the solution to the Hamiltonian system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \quad x(t_k) = x_k \\ \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \quad p(t_k) = p_k(t_k) \end{cases}$$

By completeness of (5.15), (z_k, p_k) converge uniformly to the unique solution (z, p) of the Hamiltonian system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \quad x(t_0) = x_0 \\ \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \quad p(t_0) = \overline{p} \end{cases}$$

By Proposition 5.3.1 for all $k \geq 1$ and almost all $t \in [t_k, T]$,

$$H(t, z_k(t), p_k(t)) = \langle p_k(t), z_k'(t) \rangle - L(t, z_k(t), u_k(t))$$

and from H_3) it follows that $\{z_k'(\cdot)\}$ is bounded in $L^\infty(0,T)$.

We extend $L(\cdot, z_k(\cdot), u_k(\cdot))$ on $[0, t_k[$ by zero function and deduce from the above equality and H_3) that $\{|L(\cdot, z_k(\cdot), u_k(\cdot))|\}_{k \geq 1}$ is bounded in $L^\infty(0,T)$.

Taking a subsequence and keeping the same notations we may assume that

$(z_k'(\cdot),\ L(\cdot, z_k(\cdot), u_k(\cdot)))$ converges weakly in $L^1(0,T)$ to $(y(\cdot), \alpha(\cdot))$

Since for every $t \in [t_k, T]$, $z_k(t) = x_k + \int_{t_k}^t z_k'(s)ds$, taking the limit, we obtain $z(t) = x_0 + \int_{t_0}^T y(s)ds$. Consequently $z'(\cdot) = y(\cdot)$. On the other hand,

$$V(t_k, x_k) = g((z_k(T)) + \int_{t_k}^T L(s, z_k(s), u_k(s))ds$$

Hence, by continuity of V, passing to the limit, we obtain

$$V(t_0, x_0) = g(z(T)) + \int_{t_0}^T \alpha(s)ds$$

By Mazur's theorem and H_1), H_5) for almost all $t \in [t_0, T]$,

$$(y(t), \alpha(t)) \in \{(f(t, z(t), u), L(t, z(t), u) + r) \mid u \in U,\ r \geq 0\}$$

Hence, applying the measurable selection theorem, we can find $\overline{u} \in \mathcal{U}$ and a mesurable $r(\cdot) : [t_0, T] \mapsto \mathbf{R}_+$ such that for almost all t,

$$y(t) = f(t, z(t), \overline{u}(t)) \quad \& \quad \alpha(t) = L(t, z(t), \overline{u}(t)) + r(t)$$

This implies that z corresponds to the control $\overline{u} \in \mathcal{U}$. Finally, since $r(t) \geq 0$,

$$V(t_0, x_0) \geq g(z(T)) + \int_{t_0}^{T} L(s, z(s), \overline{u}(s))ds$$

and therefore $(z, \overline{u})$ is optimal for (P).

To prove the last statement fix $\overline{p}_i \in \partial_x^\star V(t_0, x_0)$, $i = 1, 2$ and let (z_i, p_i), $i = 1, 2$ be solutions of (5.15) such that $p_i(t_0) = \overline{p}_i$. From the uniqueness of optimal trajectory z and the first claim of our theorem, we deduce that $z^1 = z^2 = z$. Consequently, (z, p_i) solve the Hamiltonian system (5.15) with $x_T = z(T)$ for $i = 1, 2$. So, by uniqueness, $p_1(t_0) = p_2(t_0)$.

5.3.3 Smoothness of the Value Function

Differentiablility of the value function is related to solutions of (5.15) in the following way.

Theorem 5.3.4 *Assume $H_1) - H_5)$, that V is locally Lipschitz and for every $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ the problem (P) has an optimal solution.*

Then the following four statements are equivalent:

i) The value function V is continuously differentiable
ii) For every $t_0 \in [0, T]$, $V(t_0, \cdot)$ is continuously differentiable
iii) $\forall$ $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ the optimal trajectory of (P) is unique
iv) For the Hamiltonian system (5.15) the set

$$M_t := \{(x(t), p(t)) \mid (x, p) \text{ solves } (5.15) \text{ on } [t, T]\}$$

is the graph of a continuous function $\pi_t : \mathbf{R}^n \mapsto \mathbf{R}^n$.

Furthermore, iv) yields that $\pi_t(\cdot) = -\frac{\partial V}{\partial x}(t,\cdot)$ *and every solution* (x,p) *of (5.15) restricted to* $[t_0,T]$ *satisfies:* x *is optimal for (P) with* $x_0 = x(t_0)$ *and* $p(t) = -\frac{\partial V}{\partial x}(t,x(t))$ *for all* $t \in [0,T]$.

Before proving the above theorem, we shall state few corollaries.

Corollary 5.3.5 *Under all assumptions of Theorem 5.3.4, suppose* U *is a finite dimensional space, that for some* $\overline{f} : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R}^n$, $b : [0,T] \times \mathbf{R}^n \mapsto L(U,\mathbf{R}^n)$ *we have*

$$\forall\ (t,x) \in [0,T] \times \mathbf{R}^n, \quad f(t,x,u) = \overline{f}(t,x) + b(t,x)u$$

and for every (t,x), $\frac{\partial L}{\partial u}(t,x,\cdot)$ *is bijective. Then the (equivalent) statements* $i) - iv)$ *of Theorem 5.3.4 are equivalent to*

v) For every $(t_0,x_0) \in [0,T] \times \mathbf{R}^n$ *there exists a unique optimal control* $\overline{u}(\cdot)$ *solving the problem (P). Furthermore, if* z *denotes the corresponding optimal trajectory, then for all* $t \in [t_0,T]$,

$$\overline{u}(t) = \left(\frac{\partial L}{\partial u}(t,z(t),\cdot)\right)^{-1}\left(-b(t,z(t))^\star \frac{\partial V}{\partial x}(t,z(t))\right)$$

The above corollary follows from *iii)* of Theorem 5.3.4 and the fact that $\overline{u}$ verifies

$$H(t,z(t),p(t)) = \langle p(t), z'(t)\rangle - L(t,z(t),\overline{u}(t)) \text{ a.e. in } [t_0,T]$$

where $p(\cdot)$ is the co-state of the maximum principle (see Theorem 5.3.2).

Our next corollary links results of Section 1 and Theorem 5.3.4.

Corollary 5.3.6 *Under all assumptions of Theorem 5.3.4, suppose that* $\nabla g(\cdot)$ *is locally Lipschitz,* $H(t,\cdot,\cdot)$ *is twice continuously differentiable and*

$$\forall\, r > 0,\ \exists\, k_r \in L^1(0,T),\ \frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \text{ is } k_r(t) - \text{Lipschitz on } B_r(0)$$

Then the following two statements are equivalent:

i) $\forall\, t \in [0,T]$, $\frac{\partial V}{\partial x}(t,\cdot)$ *is locally Lipschitz* *ii)* $\forall\,(x,p)$ *solving (5.15) on* $[0,T]$ *and every* $P_T \in \partial^\star(\nabla g)(x(T))$, *the matrix Riccati equation*

$$\begin{cases} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t,x(t),p(t))P \; + \; P\, \dfrac{\partial^2 H}{\partial x \partial p}(t,x(t),p(t)) \; + \\[2ex] + \; P\, \dfrac{\partial^2 H}{\partial p^2}(t,x(t),p(t))P \; + \; \dfrac{\partial^2 H}{\partial x^2}(t,x(t),p(t)) \; = \; 0 \\[2ex] P(T) = P_T \end{cases}$$

has a solution on $[0,T]$.

Furthermore, if i) (or equivalently ii)) holds true, then

$$\nabla g \;\text{ is differentiable } \;\Longrightarrow\; \frac{\partial V}{\partial x}(t,\cdot) \;\text{ is differentiable}$$

and for every (x,p) *solving (5.15),* $P(t) = -\frac{\partial^2 V}{\partial x^2}(t,x(t))$. *If moreover* $g \in C^2$, *then* $V(t,\cdot) \in C^2$.

Proof — Let M_t be defined as in Theorem 5.3.4. If *i)* holds true, then, by Theorem 5.3.4, M_t is the graph of a locally Lipschitz function π_t. By the maximum principle (Theorem 5.3.2), $\pi_t(\cdot) = -\frac{\partial V}{\partial x}(t,\cdot)$. Applying Theorem 5.1.3, we deduce *ii)*. Conversely, assume that *ii)* is verified. Thus, by Theorem 5.1.3, M_t is the graph of a locally Lipschitz function from an open set $\mathcal{D}(t) \subset \mathbf{R}^n$ into $\mathbf{R}^n$. By the maximum principle, $M_t = \text{Graph}(-\frac{\partial V}{\partial x}(t,\cdot))$. Hence *i)*. The last statement follows from Theorem 5.1.3, because $P(t)$ describes the evolution of tangent space to M_t at $(x(t),p(t))$. □

To prove Theorem 5.3.4 we need the following lemma.

Lemma 5.3.7 *Under all assumptions of Theorem 5.3.4 consider* $(t_0,x_0) \in]0,T[\, \times \mathbf{R}^n$ *such that* V *is differentiable at* (t_0,x_0). *Then*

$$-\frac{\partial V}{\partial t}(t_0,x_0) + H\left(t_0, x_0, -\frac{\partial V}{\partial x}(t_0,x_0)\right) = 0$$

i.e., V *satisfies the Hamilton-Jacobi-Bellman equation almost everywhere in* $[0,T] \times \mathbf{R}^n$.

Proof — Fix any $\overline{u} \in U$ and consider a solution x to

$$\begin{cases} x'(t) &= f(t, x, \overline{u}) \\ x(t_0) &= x_0 \end{cases}$$

Observe that for all small $h > 0$ it is defined on $[t_0, t_0 + h]$ and, by the very definition of the value function,

$$V(t_0 + h, x(t_0 + h)) + \int_{t_0}^{t_0+h} L(s, x(s), \overline{u})ds - V(t_0, x_0) \geq 0$$

Dividing by $h > 0$ and taking the limit we prove

$$\forall\, \overline{u} \in U,\ \ \frac{\partial V}{\partial t}(t_0, x_0) + \left\langle \frac{\partial V}{\partial x}(t_0, x_0), f(t_0, x_0, \overline{u}) \right\rangle + L(t_0, x_0, \overline{u}) \geq 0$$

Consider next an optimal solution-control pair $(z, \overline{u})$ of the Bolza problem (P). Then

$$V(t_0 + h, z(t_0 + h)) + \int_{t_0}^{t_0+h} L(s, z(s), u)ds - V(t_0, x_0) = 0 \quad (5.17)$$

By Theorem 5.3.2, $z(\cdot)$ solves the Hamiltonian system (5.15) with $x_T = z(T)$. Hence, by H_3), $z(\cdot) \in C^1$ $(z(t_0 + h) - z(t_0))/h$ converge to some v when $h \to 0+$. By (5.17), for some $\sigma \in \mathbf{R}$,

$$\lim_{h \to 0+} \frac{1}{h} \int_{t_0}^{t_0+h} L(s, z(s), \overline{u}(s))ds = \sigma$$

By H_5)

$$(v, \sigma) \in \{(f(t_0, x_0, u), L(t_0, x_0, u) + r) \mid u \in U,\ r \geq 0\}$$

Thus for some $u_0 \in U$ and $r_0 > 0$

$$(v, \sigma) = (f(t_0, x_0, u_0), L(t_0, x_0, u_0) + r_0)$$

Dividing (5.17) by h and taking the limit yields

$$\frac{\partial V}{\partial t}(t_0, x_0) + \left\langle \frac{\partial V}{\partial x}(t_0, x_0), f(t_0, x_0, u_0) \right\rangle + L(t_0, x_0, u_0) + r_0 = 0$$

So we proved the existence of $u_0 \in U$ such that

$$\frac{\partial V}{\partial t}(t_0, x_0) + \left\langle \frac{\partial V}{\partial x}(t_0, x_0), f(t_0, x_0, u_0) \right\rangle + L(t_0, x_0, u_0) \leq 0$$

The two inequalities derived above imply the result. □

Proof of Theorem 5.3.4 — Clearly $i) \Longrightarrow ii)$. Assume next that $ii)$ holds true. Fix $0 \leq t_0 < T$, $x_0 \in \mathbf{R}^n$ and let $\overline{x}$ be an optimal solution to problem (P). Then, by Theorem 5.3.2, there exists $p(\cdot)$ such that $(\overline{x}, p)$ solves the system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \quad x(t_0) = x_0 \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \quad p(t_0) = -\dfrac{\partial V}{\partial x}(t_0, x_0) \end{cases}$$

Since the solution to such system is unique, we deduce $iii)$.

Conversely assume that $iii)$ holds true. Fix $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$. Then, by Lemma 5.3.3, $\partial_x^\star V(t_0, x_0)$ is a singleton. We claim that the set

$$\partial^\star V(t_0, x_0) := \text{Limsup}_{(t,x)\to(t_0,x_0)}\{\nabla V(t, x)\}$$

is a singleton.

Indeed let $(p_t, p_x) \in \partial^\star V(t_0, x_0)$ and $(t_i, x_i) \to (t_0, x_0)$ be such that $\nabla V(t_i, x_i) \to (p_t, p_x)$. Then $\{p_x\} = \partial_x^\star V(t_0, x_0)$ and, by Lemma 5.3.7, V satisfies at (t_i, x_i) the Hamilton-Jacobi-Bellman equation

$$-\frac{\partial V}{\partial t}(t_i, x_i) + H\left(t_i, x_i, -\frac{\partial V}{\partial x}(t_i, x_i)\right) = 0$$

Taking the limit we get

$$p_t = H(t, x, -p_x)$$

So p_t is uniquely defined and, thus, $\partial^\star V(t_0, x_0)$ is a singleton implying that V is differentiable at (t_0, x_0). Since $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ is arbitrary, we deduce that V is continuously differentiable on $[0, T] \times \mathbf{R}^n$. Hence we proved $iii) \Longrightarrow i)$.

Assume next that *iv)* holds true. Fix $t_0 \in [0, T]$ and $x_0 \in \mathbf{R}^n$. By Lemma 5.3.3,

$$(x_0,\ \partial_x^\star V(t_0, x_0)) \ \subset \ \mathrm{Graph}(\pi_{t_0})$$

Thus $\partial_x^\star V(t_0, x_0)$ is a singleton. In particular

$$\mathrm{Limsup}_{x \to x_0} \left\{ \frac{\partial V}{\partial x}(t_0, x) \right\} \ \text{ is a singleton}$$

and therefore, *ii)* is verified.

It remains to show that *ii)* yields *iv)*. For this aim fix $t_0 \in [0, T]$ and define the continuous mapping $\Psi : \mathbf{R}^n \mapsto \mathbf{R}^n$ in the following way:

For all $x_0 \in \mathbf{R}^n$ consider the solution (x, p) to the system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \ \ x(t_0) = x_0 \\ \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \ \ p(t_0) = -\dfrac{\partial V}{\partial x}(t_0, x_0) \end{cases}$$

and set $\Psi(x_0) = x(T)$. By the maximum principle (Theorem 5.3.2) we know that $p(T) = -\nabla g(x(T))$. Thus $(x(T), p(T)) \in \mathrm{Graph}(-\nabla g)$. In particular this yields Ψ is one-one. By the Invariance of Domain Theorem, $\Psi(\mathbf{R}^n)$ is open. Thus the set

$$\{(\Psi(x_0), -\nabla g(\Psi(x_0))) \mid x_0 \in \mathbf{R}^n\} \text{ is open and closed in Graph}(-\nabla g)$$

So it coincides with $\mathrm{Graph}(-\nabla g)$. Hence, by uniqueness of solution to the Hamiltonian system (5.15), $M_t = \mathrm{Graph}(-\frac{\partial V}{\partial x}(t_0, \cdot))$. The proof is complete.

5.3.4 Problems with Concave-Convex Hamiltonians

Observe that in general one has

$$\frac{\partial^2 H}{\partial p^2}\ (t, x(t), p(t)) \ \geq \ 0$$

for every solution (x,p) of the Hamiltonian system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t,x(t),p(t)), \quad x(T) = x_T \\ \\ -p'(t) &= \dfrac{\partial H}{\partial x}(t,x(t),p(t)), \quad p(T) = -\nabla g(x_T) \end{cases} \tag{5.18}$$

and that whenever in addition $H(t,\cdot,p(t))$ is concave for all $t \in [0,T]$, then

$$\frac{\partial^2 H}{\partial x^2}\,(t,x(t),p(t)) \;\leq\; 0$$

If g is convex, then every matrix from the generalized Jacobian $\partial^\star g(x(T))$ is nonnegative. From Corollary 5.2.4 we deduce that for every $P_T \in \partial^\star(\nabla g)(x(T))$, the solution $P(\cdot)$ of the matrix Riccati equation

$$\begin{cases} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t,x(t),p(t))P \;+\; P\,\dfrac{\partial^2 H}{\partial x \partial p}(t,x(t),p(t)) \;+ \\ \\ +\,P\,\dfrac{\partial^2 H}{\partial p^2}(t,x(t),p(t))P \;+\; \dfrac{\partial^2 H}{\partial x^2}(t,x(t),p(t)) \;=\; 0 \\ \\ P(T) = -P_T \end{cases} \tag{5.19}$$

exists on $[0,T]$. By Theorem 5.1.3, no shocks of (5.18) can occur backward in time. Hence we deduce from Theorem 5.3.4 and Corollary 5.3.6 the following results.

Theorem 5.3.8 *Assume $H_1) - H_5)$, that V is locally Lipschitz and for every $(t_0,x_0) \in [0,T] \times \mathbf{R}^n$ the problem (P) has an optimal solution.*

Further assume that $\nabla g(\cdot)$ is locally Lipschitz, $H(t,\cdot,\cdot)$ is twice continuously differentiable and

$$\forall\, r > 0,\ \exists\, k_r \in L^1(0,T),\ \frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \text{ is } k_r(t) - \text{Lipschitz on } B_r(0)$$

If for every solution (x,p) of (5.18), $H(t,\cdot,p(t))$ is concave and g is convex, then $V \in C^1$ and $\frac{\partial V}{\partial x}(t,\cdot)$ is locally Lipschitz.

Moreover, every solution (x, p) *of (5.18) is an optimal trajectory-co-state pair. If in addition* $g \in C^2$, *then* $V(t, \cdot) \in C^2$ *and, in this case,* $P(t) = -\frac{\partial^2 V}{\partial x^2}(t, x(t))$ *solves the matrix Riccati equation (5.19) with* $P_T = -g''(x(T))$.

Chapter 6

Hamilton-Jacobi-Bellman Equation for Problems under State-Constraints

Introduction

Consider the optimal control problem

$$
\text{(P)} \qquad \begin{cases} \text{Minimize } g(x(1)) \\ \text{over } \ x \in W^{1,1}([0,1];\mathbf{R}^n) \text{ satisfying} \\ x'(t) \in F(t,x(t)) \quad \text{a.e. } t \in [0,1], \\ x(t) \in K \quad \forall\, t \in [0,1], \\ x(0) = x_0, \end{cases}
$$

the data for which comprise: a function $g : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, a set-valued map $F : [0,1] \times \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$, a closed set $K \subset \mathbf{R}^n$ and a vector $x_0 \in \mathbf{R}^n$. Solutions of the above differential inclusion satisfying the constraints of (P), are called *feasible* arcs (for (P)).

Note that, since g is extended valued, (P) incorporates the end-point constraint:

$$x(1) \in C$$

where $C := \operatorname{dom} g$.

Denote by $V : [0,1] \times K \to \mathbf{R} \cup \{+\infty\}$ the value function for (P): for each $(t,x) \in [0,1] \times K$, $V(t,x)$ is defined to be the infimum cost for the problem

$$
(P_{t,x}) \quad \begin{cases} \text{Minimize } g(y(1)) \\ \text{over } y \in W^{1,1}([t,1];\mathbf{R}^n) \text{ satisfying} \\ y'(s) \in F(s,y(s)) \quad \text{a.e. } s \in [t,1], \\ y(s) \in K \quad \forall\, s \in [t,1], \\ y(t) = x. \end{cases}
$$

Thus

$$
V(t,x) = \inf(P_{t,x}).
$$

(If $(\mathrm{P}_{t,x})$ has no feasible arcs, we set $V(t,x) = +\infty$.)

In this chapter we explore the relationship between the value function and the Hamilton-Jacobi Equation:

$$
(HJE) \begin{cases} -\frac{\partial V}{\partial t} + H(t,x,-\frac{\partial V}{\partial x}) \;=\; 0 \text{ for } (t,x) \in (0,1) \times \operatorname{int}K \\ V(1,x) \;=\; g(x) \text{ for } x \in K. \end{cases}
$$

To get uniqueness of solutions to the above PDE in the constrained case we are lead to impose some kind of constraint qualification on the dynamic constraint at boundary points of the state constraint set.

In [9], Capuzzo-Dolcetta and Lions showed that the value function is continuous and is the unique viscosity solution to (HJE) under hypotheses which include the "inward-pointing" constraint qualification:

$$
\min_{v \in F(t,x)} n_x \cdot v < 0 \quad \forall\, x \in \operatorname{bdy} K
$$

where n_x denotes the unit outward normal vector at the point $x \in \operatorname{bdy} K$. Hypotheses of this nature were introduced by Soner [40] to ensure continuity of the value function and to provide a characterization of the value function in terms of viscosity solutions of the relevant Hamilton-Jacobi equation, for an infinite horizon problem.

When the "inward pointing" constraint qualification fails, or when the terminal cost function g is chosen to take account of an endpoint constraint, we can expect that the value function will be discontinuous.

We restrict attention to a special class of state constraints sets, namely a finite intersection of smooth manifolds. (Nonetheless, this is a framework which allows state constraints sets with nonsmooth boundaries, and covers state constraints encountered in most engineering applications.) A key element is an extension of Filippov's theorem to the constrained case.

6.1 Constrained Hamilton-Jacobi-Bellman Equation

The following theorem provides two characterizations of the value function for optimal control problems with endpoint and state constraints, in terms of lower semicontinuous solutions of the Hamilton-Jacobi equation and one in terms of epiderivative solutions.

It is assumed that the state constraint set K is expressible as

$$K \;=\; \cap_{j=1}^{r}\{x : h_j(x) \leq 0\}$$

for a finite family of $C^{1,1}$ functions $\{h_j : \mathbf{R}^n \to \mathbf{R}\}_{j=1}^{r}$. ($C^{1,1}$ denotes the class of C^1 functions with locally Lipschitz continuous gradients.) The "active set" of index values $I(x)$, at a point $x \in \text{bdy}\, K$, is

$$I(x) \;:=\; \{j \in (1, \ldots, r) : h_j(x) = 0\}.$$

Recall the notations $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$ for all real numbers a, b. We write

$$h^+(x) := \left(\max_{j=1,2,\ldots,r} h_j(x)\right) \vee 0.$$

$W^{1,1}([a, b]; \mathbf{R}^n)$ denotes the space of absolutely continuous n-vector valued functions on $[a, b]$, with norm

$$||x||_{W^{1,1}} = ||x(a)|| + \int_a^b ||x'(t)||dt.$$

Theorem 6.1.1 *Take a function* $V : [0, 1] \times K \to \mathbf{R} \cup \{+\infty\}$. *Assume that the following hypotheses are satisfied:*

(H1) F is a continuous set-valued map, which takes values in the space of non-empty, closed, convex sets,

(H2) There exists $c > 0$ such that

$$F(t,x) \subset c(1+\|x\|)B \quad \forall\,(t,x) \in [0,1] \times \mathbf{R}^n,$$

(H3) There exists $k \in L^1$ such that

$$F(t,x) \subset F(t,x_1)) + k(t)\|x - x_1\|B \quad \forall\, t \in [0,1],\ x, x_1 \in \mathbf{R}^n \times \mathbf{R}^n,$$

(H4) g is lower semicontinuous.

Assume furthermore that

(CQ) For all $x \in K$ and $t \in [0,1]$ there exists $v \in F(t,x)$ such that

$$\forall j \in I(x), \quad \nabla h_j(x) \cdot v > 0.$$

Then assertions (a)-(c) below are equivalent:

(a) V is the value function for (P).

(b) V is lower semicontinuous and

(i) $\forall\,(t,x) \in ([0,1[\times K) \cap \operatorname{dom} V$

$$\inf_{v \in F(t,x)} D_{\uparrow}V(t,x)(1,v) \;\leq\; 0$$

(ii) $\forall\,(t,x) \in]0,1] \times \operatorname{int} K) \cap \operatorname{dom} V$

$$\sup_{v \in F(t,x)} D_{\uparrow}V(t,x)(-1,-v) \;\leq\; 0$$

(iii) $\forall\, x \in K$

$$\liminf_{\{(t',x') \to (1,x) : t' < 1, x' \in \operatorname{int} K\}} V(t',x') = V(1,x) = g(x)$$

(c) V is lower semicontinuous and

(i) $\forall\,(t,x) \in (]0,1[\times \operatorname{int} K) \cap \operatorname{dom} V,\ (p_t, p_x) \in \partial_{-}V(t,x)$

$$-p_t + H(t,x,-p_x) = 0.$$

(ii) $\forall\ (t,x) \in (]0,1[\times \text{bdy}\, K) \cap \text{dom}\, V,\ (p_t, p_x) \in \partial_{-} V(t,x)$

$$-p_t + H(t,x,-p_x) \geq 0$$

(iii) $\forall\ x \in K$,

$$\liminf_{\{(t',x')\to(0,x):t'>0\}} V(t',x') = V(0,x)$$

and

$$\liminf_{\{(t',x')\to(1,x):t'<1,x'\in \text{int}\, K\}} V(t',x') = V(1,x) = g(x).$$

Example Consider

$$\begin{cases} \text{Minimize } g(x(1)) \\ x'(t) \in F(t,x(t)) \\ x(t) \in K \\ x(0) = x_0, \end{cases}$$

in which $n = 1$, $g(x) = x$, $F(t,x) = \{1\}$, $K = \{x : x \leq 0\}$, $x_0 = 0$.

By inspection

$$V(t,x) = \begin{cases} +\infty & \text{if } x > -(1-t) \\ x + (1-t) & \text{if } x \leq -(1-t) \end{cases}$$

The hypotheses for application of Thm. 6.1.1 are satisfied, including the outward-pointing constraint qualification (CQ). Thm. 6.1.1 therefore tells us that V is the unique solution of (HJE) (in the sense specified).

Notice that $V(t,x) = +\infty$ at some points in $[0,1] \times K$, despite the fact that g is everywhere finite valued (no endpoint constraints).

6.2 A Neighbouring Feasible Trajectories Theorem

A key role in the proof of the Main Theorem is played by an estimate governing the distance of the set of trajectories satisfying a given state constraint from a given trajectory which violates the constraint. This estimate is provided by the following Existence of Feasible Neighbouring Trajectories (EFNT) Theorem, which can be regarded as a kind of refined viability theorem, in which the information that a 'viable' solution exists whenever viability condition holds true is supplemented by information about where it is located, in relation to a given solution when a 'strict' viability condition holds true.

As before, we limit attention to state constraint sets K associated with a family of functional inequalities:

$$K = \cap_{j=1}^{r} \{x : h_j(x) \leq 0\},$$

in which the h_j's are given $C^{1,1}$ functions.

Theorem 6.2.1 *Fix $r_0 > 0$. Assume that for some $c > 0$, $\alpha > 0$ and $k(\cdot) \in L^1$, the following hypotheses are satisfied:*

(i) F takes values in the space of non-empty, closed sets and $F(\cdot, x)$ is measurable for each $x \in \mathbf{R}^n$.

(ii) $F(t,x) \subset c(1 + \|x\|)B \quad \forall\, (t,x) \in [0,1] \times \mathbf{R}^n$.

(iii) $F(t,x) \subset F(t,x')) + k(t)\|x - x'\|B \quad \forall\, t \in [0,1],\ x, x' \in \mathbf{R}^n$.

Assume furthermore that there exists some $\alpha > 0$ such that

$$(CQ)' \qquad \min_{v \in F(t,x)} \max_{j \in I(x)} \nabla h_j(x) \cdot v < -\alpha \quad ,$$
$$x \in B(0, e^c(r_0 + c)) \cap \operatorname{bdy} K,\ t \in [0,1].$$

Then there exists a constant ϑ (which depends on r_0, c, α and $k \in L^1$) with the following property: given any $t_0 \in [0,1]$ and any $\hat{x} \in \mathcal{S}_{[t_0,1]}$ such that $\hat{x}(t_0) \in B(0,r_0) \cap K$, an $x \in \mathcal{S}_{[t_0,1]}(\hat{x}(t_0))$ can be found such that

$$x(t) \in K \quad \forall\, t \in [t_0, 1]$$

and

$$||x - \hat{x}||_{W^{1,1}([t_0,1];\mathbf{R}^n)} \leq \vartheta \max_{t \in [t_0,1]} h^+(\hat{x}(t)).$$

The need to introduce into (CQ) ′ the positive parameter α a rised because it is not hypothesized that F is a continuous multifunction. In the case F is continuous, then (CQ) ′ is implied by the condition

$$\min_{v\in F(t,x)} \max_{j\in I(x)} \nabla h_j(x)\cdot v < 0 \quad \forall\, x \in B(0, e^c(r_0+c))\cap \text{bdy}\, K,\ t\in[0,1].$$

Proof. Set $R = e^c(r_0 + c)$ and $c_0 = c(1 + R)$. Let k_h be a common Lipschitz constant for the h_j's on $B(0, R)$ and let κ be a common Lipschitz constant for the ∇h_j's on $B(0, R)$.

Note that for any $[t', t''] \subset [0, 1]$ and any solution $y : [t', t''] \to \mathbf{R}^n$ to our differential inclusion such that $y(t') \in B(0, r_0)$, we have $y(t) \in B(0, R)$. This follows from Gronwall's lemma.

Let $\omega : R^+ \to \mathbf{R}^n$ be a modulus of continuity for $t \to \int_0^t k(s)ds$, i.e. ω is monotone increasing, $\lim_{\sigma\downarrow 0}\omega(\sigma) = 0$, and

$$\omega(t'-t) \geq \int_t^{t'} k(s)ds \quad \forall\, [t,t'] \subset [0,1].$$

Define

$$I_\beta(\xi) \ := \ \{j \in \{1,\ldots,r\} : 0 \geq h_j(\xi) \geq -\beta\}.$$

Under the hypotheses, there exists $\beta > 0$ and $\alpha > 0$ such that

$$\forall \xi \in K\cap B(0,R), t\in[0,1] \qquad \min_{v\in F(t,\xi)} \max_{j\in I_\beta(\xi)} \nabla h_j(\xi)\cdot v \ < \ -\alpha.$$

Fix $\alpha' \in (0, \alpha)$. Choose $\eta \in (0, 1)$ such that $N := \eta^{-1}$ is an integer and the following conditions are satisfied:

$$\eta \ < \ (\alpha-\alpha')(c_0^2\kappa)^{-1} \tag{6.1}$$

$$\omega(\eta) \ < \ \log\left(\frac{\alpha-\alpha'}{8c_0k_h}+1\right) \tag{6.2}$$

$$\eta(k_hc_0+\alpha') \ < \ \beta, \tag{6.3}$$

and

$$6\frac{c_0^2}{\alpha'}\kappa e^{\omega(\eta)}\eta + 6\frac{c_0}{\alpha'}k_h(e^{\omega(\eta)}-1)2c_03/\alpha' \ < \ 1. \tag{6.4}$$

Set

$$\vartheta' \ := \ \max\left\{\frac{6c_0}{\alpha'}\exp\left(\int_0^1 k(s)ds\right), \frac{6}{\alpha'\eta}c_0\right\}. \tag{6.5}$$

Step 1: We show that, for every $\xi \in K$ and $\tau \in (0, 1-\eta]$ there exists a solution $\tilde{x} : [\tau, \tau+\eta] \to \mathbf{R}^n$ such that $\tilde{x}(\tau) = \xi$ and

$$h_j(\tilde{x}(t)) \le -\alpha'(t-\tau) \quad \forall j, \forall t \in [\tau, \tau+\eta].$$

Fix $\xi \in K$ and $\tau \in (0, 1-\eta]$. Consider a measurable function $v : [\tau, \tau+\eta] \to \mathbf{R}^n$ such that $v(t) \in F(t, \xi)$ a.e. $t \in [\tau, \tau+\eta]$ and

$$j \in I_\beta(\xi) \text{ implies } \nabla h_j(\xi) \cdot v(t) \ < \ -\alpha \text{ for a.e. } t \in [\tau, \tau+\eta].$$

Set $z(t) = \xi + \int_\tau^t v(s)ds$. We have, for all $j \in I_\beta(\xi)$ and $t \in [\tau, \tau+\eta]$,

$$\begin{aligned} h_j(z(t)) &= h_j(\xi) + \int_\tau^t \nabla h_j(z(s)) \cdot v(s)ds \\ &\le 0 + \int_\tau^t \nabla h_j(z(s)) \cdot v(s)ds \\ &\le \int_\tau^t \nabla h_j(\xi) \cdot v(s)ds + \int_\tau^t \|\nabla h_j(z(t)) - \nabla h_j(\xi)\| \cdot \|v(s)\|ds \\ &\le -\alpha(t-\tau) + \kappa c_0^2(t-\tau)^2/2 \\ &\le (-\alpha + (\alpha-\alpha')/2)(t-\tau) \le -\left(\frac{\alpha+\alpha'}{2}\right)(t-\tau). \end{aligned}$$

(We have used (6.1).)

Fix $j \in I_\beta(\xi)$. By Filippov's Theorem, applied to the reference trajectory z, there exists a solution $\tilde{x} : [\tau, \tau+\eta] \to \mathbf{R}^n$ such that $\tilde{x}(\tau) = \xi$ and, for all $t \in [\tau, \tau+\eta]$,

$$\begin{aligned} \|\tilde{x}(t) - z(t)\| &\le \int_\tau^t d_{F(s,z(s))}(z'(s))\exp\left(\int_s^t k(\sigma)d\sigma\right) ds \\ &\le c_0(t-\tau)\int_\tau^t k(s)\exp\left(\int_s^t k(\sigma)d\sigma\right) ds \\ &\le c_0(t-\tau)(\exp\left(\int_\tau^t k(s)ds)\right) - 1) \\ &\le c_0(t-\tau)\left(e^{\omega(t-\tau)} - 1\right) \le \frac{\alpha-\alpha'}{8k_h}(t-\tau). \end{aligned}$$

(We have used (6.2). But then, for all $t \in [\tau, \tau+\eta]$,

$$\begin{aligned} h_j(\tilde{x}(t)) &\le k_h\|\tilde{x}(t) - z(t)\| + h_j(z(t)) \\ &\le \left(\frac{\alpha-\alpha'}{8} - \frac{\alpha+\alpha'}{2}\right)(t-\tau) \le -\alpha'(t-\tau). \end{aligned}$$

On the other hand, if $h_j(\xi) \leq -\beta$, then, for all $t \in [\tau, \tau + \eta]$,

$$h_j(\tilde{x}(t)) \;\leq\; -\beta + (t - \tau)k_h c_0 \leq -\alpha'\eta \leq -\alpha'(t - \tau).$$

We see that, in this case too, the inequality is satisfied. Step 1 is complete.

Step 2: Take any $\tau \in [0, 1 - \eta]$ and any solution $x_1 : [0,1] \to \mathbf{R}^n$ such that $x_1(t) \in K$ for all $t \in [0, \tau]$. We shall show that there exists an solution $x_2 : [0,1] \to \mathbf{R}^n$ such that $x_2(t) = x_1(t)$ for all $t \in [0, \tau]$,

$$x_2(t) \in K \quad \forall\, t \in [\tau, \tau + \eta]$$

and

$$||x_1 - x_2||_{W^{1,1}([0,1];\mathbf{R}^n)} \;\leq\; \vartheta' \max_{t \in [0,1]} h^+(x_1(t)).$$

Set

$$\Delta \;=\; \max_{t \in [0,1]} h^+(x_1(t)).$$

Suppose that $\Delta \geq \alpha'\eta/3$. By Step 1, there exists a solution $x_1 : [0, \tau + \eta] \to \mathbf{R}^n$ such that $x_2(t) = x_1(t)$ for all $t \in [0, \tau]$ and $x_2(t) \in K$ for $t \in [0, \tau + \eta]$. We have

$$||x_1 - x_2||_{W^{1,1}([0,1];\mathbf{R}^n)} \leq 2c_0 \leq \vartheta' \left(\frac{\alpha'\eta}{3}\right) \leq \vartheta'\Delta,$$

by (6.5), as required. We can therefore assume that

$$\Delta \;<\; \alpha'\eta/3.$$

Set

$$\eta' = 3\Delta/\alpha'.$$

Notice that $\eta' \leq \eta$ and $\tau + \eta' \leq 1$. By the results of Step 1, there exists a solution $z : [0, \tau + \eta'] \to \mathbf{R}^n$ such that $z(t) = x_1(t)$ for all $t \in [0, \tau]$,

$$z(t) \in K \quad \forall t \in [0, \tau + \eta']$$

and

$$h_j(z(\tau + \eta')) \leq -\alpha'\eta' = -3\Delta \quad \forall j.$$

By Filippov's Theorem, there exists a solution $y : [\tau + \eta', 1] \to \mathbf{R}^n$ such that $y(\tau + \eta') = z(\tau + \eta')$ and, for all $t \in [\tau + \eta', 1]$,

$$\begin{aligned}
||y(t) - x_1(t)|| &\leq \exp\left(\int_{\tau+\eta'}^{t} k(s)ds\right) ||z(\tau + \eta') - x_1(\tau + \eta')|| \\
&\leq \exp\left(\int_{\tau+\eta'}^{t} k(s)ds\right) 2c_0\eta', \qquad (6.6) \\
||y'(t) - x_1'(t)|| &\leq k(t)\exp\left(\int_{\tau+\eta'}^{t} k(s)ds\right) ||z(\tau + \eta') - x_1(\tau + \eta')|| \\
&\leq k(t)\exp\left(\int_{\tau+\eta'}^{t} k(s)ds\right) 2c_0\eta'. \qquad (6.7)
\end{aligned}$$

Now concatenate $z : [0, \tau + \eta'] \to \mathbf{R}^n$ and $y : [\tau + \eta', 1] \to \mathbf{R}^n$ to form the solution $x_2 : [0, 1] \to \mathbf{R}^n$.

Since $x_1(0) = x_2(0)$,

$$\begin{aligned}
||x_1 - x_2||_{W^{1,1}} &= ||x_1' - x_2'||_{L^1([0,\tau+\eta'];\mathbf{R}^n)} + ||x_1' - x_2'||_{L^1([\tau+\eta',1];\mathbf{R}^n)} \\
&\leq 2c_0\eta' + 2c_0\eta'(\exp\left(\int_{\tau+\eta'}^{1} (k(s)ds\right) - 1) \\
&= 2c_0\eta'\exp\left(\int_{\tau+\eta'}^{1} k(s)ds\right) \\
&= (6c_0/\alpha')\exp\left(\int_{\tau+\eta'}^{1} k(s)ds\right) \Delta \leq \vartheta'\Delta.
\end{aligned}$$

(We have used (6.7).) It remains to show that

$$x_2(t) \in K \text{ for all } t \in [0, \tau + \eta].$$

The condition is clearly satisfied for any $t \in [0, \tau + \eta']$. On the other hand, for any $t \in [\tau + \eta', \tau + \eta]$ and j, we have from (6.6) and (6.7)

$$\begin{aligned}
h_j(x_2(t)) &= h_j(x_2(\tau + \eta')) + \int_{\tau+\eta'}^{t} \nabla h_j(x_2(s)) \cdot x_2'(s)ds \\
&= h_j(x_2(\tau + \eta')) + \int_{\tau+\eta'}^{t} \nabla h_j(x_1(s)) \cdot x_1'(s)ds \\
&\quad + \int_{\tau+\eta'}^{t} (\nabla h_j(x_2(s)) - \nabla h_j(x_1(s))) \cdot x_1'(s))ds \\
&\quad + \int_{\tau+\eta'}^{t} \nabla h_j(x_2(s)) \cdot (x_2'(s) - x_1'(s))ds
\end{aligned}$$

$$\leq \quad -3\Delta + 2\Delta + (6c_0^2\kappa/\alpha')\eta e^{\omega(\eta)}\Delta + k_h\left(e^{\omega(\eta)} - 1\right)2c_0(3/\alpha')\Delta$$
$$\leq \quad \left(-1 + (6c_0^2\kappa/\alpha')\eta e^{\omega(\eta)} + k_h(e^{\omega(\epsilon')-1})6c_0/\alpha'\right)\Delta \; \leq \; 0.$$

(We have used (6.4).) Step 2 is complete.

Take any solution $\hat{x} : [0,1] \to \mathbf{R}^n$ such that $\hat{x}(0) \in B(0, r_0)$. Recall that $N = \eta^{-1}$ is an integer.

Set $x_0 = \hat{x}$. Use the results of Step 2 recursively to generate a finite sequence of solutions $x_0, \ldots, x_N$ on $[0,1]$ with the properties

$$x_i(t) \in K \quad \forall\, t \in [0, i/N]$$

and

$$||x_i - x_{i-1}||_{W^{1,1}([0,1];\mathbf{R}^n)} \;\leq\; \vartheta' \max_{t\in[0,1]} h^+(x_{i-1}(t)).$$

Now set $x = x_N$. Clearly

$$x(t) \in K \quad \forall\, t \in [0,1].$$

It is a routine exercise to show, using the results of Step 2 and the triangle inequality, that

$$||x - \hat{x}||_{W^{1,1}([t_0,1];\mathbf{R}^n)} \;\leq\; \vartheta \max_{t\in[t_0,1]} h^+(\hat{x}(t)), \tag{6.8}$$

in which $t_0 = 0$ and

$$\vartheta \;:=\; k_h^{-1}[(1 + k_h\vartheta')^N - 1].$$

A special case of the theorem has been proved, in which $t_0 = 0$.

Take any $t_0 \in [0,1]$. Suppose that $\hat{x} : [t_0, 1] \to \mathbf{R}^n$ is a solution such that $\hat{x}(t_0) \in K \cap B(0, r_0)$. Define

$$\bar{F}(t,x) \;=\; \begin{cases} F(t,x) & t \geq t_0 \\ F(t_0,x) \cup \{0\} & t < t_0. \end{cases}$$

and

$$\hat{y}(t) \;=\; \begin{cases} \hat{x}(t) & t \geq t_0 \\ \hat{x}(t_0) & t < t_0. \end{cases}$$

Now apply the earlier construction to $\bar{F}$ and $\hat{y}$ to obtain a solution $x_0 : [0,1] \to \mathbf{R}^n$ to $y' \in \bar{F}(t,y)$. It is a a simple matter to check that the solution x obtained by restricting x_0 to $[t_0, 1]$ satisfies $x(t_0) = \hat{x}(t_0)$, $x(t) \in K$ for all $t \in [t_0, 1]$ and also inequality (6.8) (with the same constant ϑ). The theorem is proved.

6.3 Proof of the Main Theorem

We isolate in the following lemma the steps in the proof of Thm. 6.1.1 requiring the constraint qualification (CQ).

Lemma 6.3.1 *(i) Take any point $x_1 \in K$. Then there exists $\delta \in]0,1[$ and a solution $y : [1-\delta, 1] \to \mathbf{R}^n$ such that $y(1) = x_1$ and*

$$y(t) \in \operatorname{int} K \quad \forall\, t \in [1-\delta, 1[.$$

(ii) Take any $t_0 \in [0,1[$ and any solution $x : [t_0, 1] \to K$. Take also a sequence of points $\{(\tau_i, \xi_i)\}$ in $[t_0, 1[\times \operatorname{int} K$ such that $(\tau_i, \xi_i) \to (1, x(1))$. Then there exists a sequence of solutions $\{x_i : [t_0, \tau_i] \to \mathbf{R}^n\}$ such that $x_i(\tau_i) = \xi_i$

$$x_i(t) \in \operatorname{int} K \quad \forall\, t \in [t_0, \tau_i], i = 1, 2, \ldots$$

and

$$||x_i - x||_{L^\infty([t_0,\tau_i];\mathbf{R}^n)} \to 0 \text{ as } i \to \infty.$$

Proof. According to (CQ), there exists $v \in F(1, x_1)$ and $\alpha > 0$ such that

$$\nabla h_j(x_1) \cdot v > \alpha \quad \forall\, j \in I(x_1).$$

For some $\delta \in]0, 1-t_0]$, whose magnitude will be set presently, define

$$z(t) = x_1 - (1-t)v \quad \text{for } t \in [1-\delta, 1].$$

By Filippov's Theorem, there exists a solution $x : [1-\delta, 1] \to \mathbf{R}^n$ such that $x(1) = x_1$ and

$$\|x(t) - z(t)\| \leq \exp\{\int_0^1 k(t)dt\} \int_t^1 d_{F(s,z(s))}(v)ds$$

for all $t \in [1-\delta, 1]$. We deduce from the continuity of $(t,x) \rightsquigarrow F(t,x)$ and the continuous differentiability of the h_j's that there exists a function $\eta : R^+ \to R^+$ such that $\eta(\theta) \downarrow 0$ as $\theta \downarrow 0$,

$$\|x(1-s) - (x_1 - sv)\| \leq \eta(s)s \quad \text{for } s \in [0, \delta]$$

and

$$h_j(x(1-s)) \leq h_j(x_1) + \nabla h_j(x_1) \cdot (x(1-s) - x_1) + \eta(s)s$$

for all $s \in [0, \delta]$ and $j \in I(x_1)$. But then, since $h_j(x_1) = 0$ for all $j \in I(x_1)$, there exists M (M does not depend on s) such that

$$h_j(x(1-s)) \leq -s\nabla h_j(x_1) \cdot v + M\eta(s)s$$

for all $j \in I(x_1)$. Hence

$$s^{-1}h_j(x(1-s)) \leq -\alpha + M\eta(s) \quad \forall\, s \in [0, \delta[,\ j \in I(x_1).$$

It follows that, if we now choose δ such that $M\eta(\delta) < \alpha$, then $h_j(x(t)) < 0$ for all $j \in I(x_1)$. Since $h_j(x_1) < 0$ for all $j \notin I(x_1)$, we can arrange, by a further reduction in the size of δ, that

$$\max_{j \in \{1,\ldots,r\}} h_j(x(t)) < 0 \quad \forall\, t \in [1-\delta, 1[.$$

(ii) Define the sequence of positive numbers

$$\gamma_i := \left(-\max_{j=1,\ldots,r} h_j(\xi_i)\right) \wedge (i^{-1}) \quad \text{for } i = 1, 2, \ldots$$

Since $\{\xi_i\} \subset \operatorname{int} K$ and (CQ) holds true, it follows that $\gamma_i > 0$ for all i. Clearly $\gamma_i \downarrow 0$. For each i define

$$h_j^i(x) := h_j(x) + \gamma_i.$$

Apply Filippov's Theorem to $x' \in F(t, x)$, taking as reference trajectory x restricted to $[t_0, \tau_i]$. This yields a solution $y_i : [t_0, \tau_i] \to \mathbf{R}^n$ satisfying $y_i(\tau_i) = \xi_i$ and

$$||y_i - x||_{L^\infty([t_0,\tau_i];\mathbf{R}^n)} \leq \exp\{\int_0^1 k(t)dt\}\|x(\tau_i) - \xi_i\|.$$

Since $(x(\tau_i) - \xi) \to 0$ as $i \to \infty$, we can conclude that

$$||y_i - x||_{L^\infty([t_0,\tau_i])} \to 0 \quad \text{as } i \to \infty. \tag{6.9}$$

By the comments following the statement of Theorem 6.2.1, (CQ) yields (CQ) $'$. So we deduce from Thm.6.2.1 applied to the set-valued map $-F$ that there exists $\vartheta > 0$ and a sequence of solutions $\{x_i : [t_0, \tau_i] \to \mathbf{R}^n\}$ such that $x_i(\tau_i) = \xi_i$ and

$$||y_i - x_i||_{L^\infty([t_0,\tau_i];\mathbf{R}^n)} \leq \vartheta \left[\max_{t\in[t_0,\tau_i]} \max_j h_j(y_i(t)) + \gamma_i\right]^+$$

$$h_j(x_i(t)) + \gamma_i \leq 0 \quad \forall\, t \in [t_0, \tau_i],\ j \in I(x_i(t)),\ i = 1, 2, \ldots$$

This means that

$$x_i(t) \in \operatorname{int} K \quad \forall\, t \in [t_0, \tau_i],\ i = 1, 2, \ldots$$

Since $h_j(x(t)) \leq 0$ for all $t \in [0, 1]$, we deduce from (6.9) that

$$||x_i - x||_{L^\infty([t_0,\tau_i])} \to 0 \quad \text{as } i \to \infty.$$

In the next lemma, reference is made to the δ-tube about $\bar{x} : [t_0, t_1] \to \mathbf{R}^n$:

$$T_\delta(\bar{x}) \ := \ \{(t, x) \in [t_0, t_1] \times \mathbf{R}^n : \|x - \bar{x}(t)\| < \delta\}.$$

Lemma 6.3.2 *Take* $[t_0, t_1] \subset [0, 1]$ *such that* $t_0 < t_1$, *a solution* $\bar{x} : [t_0, t_1] \to \mathbf{R}^n$, $\delta > 0$ *and a lower semicontinuous function* $V : [t_0, t_1] \times \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ *such that for all* $(t, x) \in T_\delta(\bar{x})$ *with* $t < t_1$,

$$\forall\ (p_t, p_x) \in \partial_- V(t, x),\ -p_t + H(t, x, -p_x) \leq 0 \tag{6.10}$$

Then, for any $t_0 \leq t' \leq t'' < t_1$,

$$V(t', \bar{x}(t')) \leq V(t'', \bar{x}(t'')).$$

Proof. We deduce in the same way as in Chapter 4, proofs of Theorems 4.2.2 and 4.2.4 that

$$V(t', \bar{x}(t')) \leq V(t'', \bar{x}(t'')).$$

The fact that $t'' < t_1$ (strict inequality) is important here, since no regularity hypotheses have been imposed on $t \to V(t, \cdot)$ at $t = t_1$.

Proof of Thm. 6.1.1

(a) $\Rightarrow$ (b). The value function V is lower semicontinuous by the same arguments as those used in Chapter 4.

Under the hypotheses, $(t, x) \in \operatorname{dom} V$ implies that $(\mathrm{P}_{t,x})$ has a solution. It is a straightforward matter to show that, if y is a minimizer for $(\mathrm{P}_{t,x})$, then $s \to V(s, y(s))$ is constant on $[t, 1]$; b(i) can be deduced from this property.

It can also be shown that, if $y : [t, 1] \to \mathbf{R}^n$ is a solution satisfying the constraints of $(\mathrm{P}_{t,x})$, then $s \to V(s, y(s))$ is non-decreasing on $[t, 1]$; b(ii) can be deduced from this latter property.

Since V is lower semicontinuous, it remains only to verify

$$\liminf_{\{(t',x') \to (1,x) : t' < 1, x' \in \operatorname{int} K\}} V(t', x') \leq V(1, x) \quad \forall\, x \in K.$$

Lemma 6.3.1 tells us that there exists $\delta \in]0, 1[$ and a solution $y : [1 - \delta, 1] \to \mathbf{R}^n$ such that $y(1) = x$ and

$$y(t) \in \operatorname{int} K \quad \forall\, t \in [1 - \delta, 1[.$$

But $V(t, y(t)) \leq V(1, x)$, a basic monotonicity property of the value function. Since y is continuous,

$$\liminf_{\{(t',x') \to (1,x) : t' < 1, x' \in \operatorname{int} K\}} V(t', x') \leq \limsup_{t \uparrow 1} V(t, y(t)) \leq V(1, x).$$

as required.

(b) $\Rightarrow$ (c). This implication is a consequence duality relationships between $\partial_- V$ and $D_\uparrow V$.

(c) $\Rightarrow$ (a). Assume that V satisfies (c). Take any $x_0 \in K$ and $t_0 \in [0, 1]$.

Step 1: We show that

$$V(t_0, x_0) \; \geq \; \inf(P_{t_0,x_0}). \tag{6.11}$$

This inequality is automatically satisfied if $V(t_0, x_0) = +\infty$. So we assume that $V(t_0, x_0) < +\infty$.

Notice that, since $\operatorname{dom} V \subset K$, conditions c(i) and c(ii) imply

$$\forall (t,x) \in]0,1[\times \mathbf{R}^n,\ (p_t, p_x) \in \partial_- V(t,x)$$
$$-p_t + H(t,x,-p_x) \leq 0$$

and

$$\liminf_{\{(t',x')\to(0,x): t'>0\}} V(t',x') = V(0,x) \quad \forall\, x \in \mathbf{R}^n,$$

(We here regard V as a function on $[0,1] \times \mathbf{R}^n$ which takes value $+\infty$ at points $(t,x) \notin [0,1] \times K$.) But then we deduce by applying the same arguments as in Chapter 4 the existence of a solution $x : [t_0, 1] \to \mathbf{R}^n$ such that $x(t_0) = x_0$ and

$$V(t_0, x_0) \geq V(t, x(t)) \quad \forall\, t \in [t_0, 1].$$

This inequality implies that $V(t, x(t)) < +\infty$ for all $t \in [t_0, 1]$. Since $\operatorname{dom} V \subset K$, we conclude that $x(\cdot)$ satisfies the state constraint. It also implies that

$$V(t_0, x_0) \geq V(1, x(1)) = g(x(1)) \geq \inf(\mathrm{P}_{t,x}).$$

This is the required inequality.

Step 2: We show that

$$V(t_0, x_0) \ \leq\ \inf(P_{t_0,x_0}). \tag{6.12}$$

This will complete the proof, since (6.12) combines with (6.11) to give $V(t_0, x_0) = \inf(P_{t_0,x_0})$.

(6.12) is automatically satisfied if $\inf(P_{t_0,x_0}) = +\infty$. So we assume that it is finite. In this case, $\inf(P_{t_0,x_0})$ is the infimum of $g(x(1))$ over all feasible arcs of (P_{t_0,x_0}). It therefore suffices to show that

$$V(t_0, x_0) \leq g(\bar{x}(1)),$$

where $\bar{x} \in W^{1,1}([t_0, 1]; \mathbf{R}^n)$ is an arbitrary feasible arc of (P_{t_0,x_0}).

By hypothesis,

$$g(\bar{x}(1)) = \liminf_{\{(\tau,\xi)\to(1,\bar{x}(1)): \tau<1, \xi \in \operatorname{int} K\}} V(\tau, \xi).$$

There exists, therefore, a sequence $\{(\tau_i, \xi_i)\}$ in $[t_0, 1) \times \operatorname{int} K$ such that $\xi_i \to \bar{x}(1)$ and

$$V(\tau_i, \xi_i) \to g(\bar{x}(1)). \tag{6.13}$$

Lemma 6.3.1(ii) asserts the existence of a sequence of solutions $x_i : [t_0, \tau_i] \to \mathbf{R}^n$ such that $x_i(\tau_i) = \xi_i$,

$$x_i(t) \in \operatorname{int} K \quad \forall\, t \in [t_0, \tau_i]$$

and

$$||x_i - \bar{x}||_{L^\infty([t_0,\tau_i];\mathbf{R}^n)} \to 0 \quad \text{as } i \to \infty. \tag{6.14}$$

Filippov's Theorem tells us that x_i can be extended to all of $[t_0, 1]$ (we write the extension also x_i) as a solution to our differential inclusion. Choose $\sigma_i \in]\tau_i, 1[$ and $\epsilon_i > 0$ such that

$$x_i(t) + \epsilon_i B \subset \operatorname{int} K \quad \forall\, t \in [t_0, \sigma_i].$$

Now apply Lemma 6.3.2 with $\sigma_i = t_1$ and $\bar{x} = x_i$ to conclude that

$$V(t_0, x_i(t_0)) \leq V(\tau_i, \xi_i).$$

It follows from (6.13), (6.14) and the lower semicontinuity of V that

$$V(t_0, x_0) = V(t_0, \bar{x}(t_0)) \leq \liminf_i V(t_0, x_i(t_0)) \leq \lim_i V(\tau_i, \xi_i) = g(\bar{x}(1))$$

as required.

Exercises.

We impose all the assumptions of Theorem 6.2.1.

1. Assume that g is continuous on K. Show that the value function of the problem

$$\begin{cases} \text{Minimize } g(y(1)) \\ \text{over } y \in W^{1,1}([t,1];\mathbf{R}^n) \text{ satisfying} \\ y'(s) \in F(s, y(s)) \quad \text{a.e. } s \in [t,1], \\ y(s) \in K \quad \forall\, s \in [t,1], \\ y(t) = x. \end{cases}$$

coincides with the value function of the relaxed problem

$$\begin{cases} \text{Minimize } g(y(1)) \\ \text{over } y \in W^{1,1}([t,1];\mathbf{R}^n) \text{ satisfying} \\ y'(s) \in \overline{co}F(s,y(s)) \quad \text{a.e. } s \in [t,1], \\ y(s) \in K \quad \forall\, s \in [t,1], \\ y(t) = x. \end{cases}$$

and that V is continuous on $[0,1] \times K$.

2. Assuming that g is locally Lipschitz on K, show that in this case V is locally Lipschitz on $[0,1] \times K$.

3. Show that if g is continuous on K, then the value function V satisfies the following properties

$$\text{(i) } \forall\, (t,x) \in (]0,1[\times \text{int}\, K) \cap \text{dom}\, V,\ (p_t,p_x) \in \partial_+ V(t,x)$$

$$-p_t + H(t,x,-p_x) \le 0.$$

$$\text{(ii) } \forall\, (t,x) \in (]0,1[\times K) \cap \text{dom}\, V,\ (p_t,p_x) \in \partial_- V(t,x)$$

$$-p_t + H(t,x,-p_x) \ge 0$$

4. Show that if W is continuous on $[0,1] \times K$, satisfies the boundary condition $W(1,\cdot) = g$ and the above properties (i), (ii), then W is the value function.

5. State and prove a relaxation theorem under state constraints.

Bibliography

[1] AUBIN J.-P. & CELLINA A. (1984) DIFFERENTIAL INCLUSIONS, Springer-Velag, Grundlehren der math. Wiss. # 264

[2] AUBIN J.-P. (1981) *Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions*, Advances in Mathematics, Supplementary Studies, Ed. Nachbin L., 160-232.

[3] AUBIN J.-P. (1990) *A survey of viability theory*, SIAM J. Control Optim. **28**, 749-788.

[4] AUBIN J.-P. (1991) VIABILITY THEORY, Birkhäuser, Boston.

[5] AUBIN J.-P. & FRANKOWSKA H. (1990) SET-VALUED ANALYSIS, Birkhäuser, Boston.

[6] BARRON E.N. & JENSEN R. (1990) *Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex Hamiltonian*, Comm. Partial Differential Equations **15**, 1713-1742.

[7] BUTTAZZO G. (1989) SEMICONTINUITY, RELAXATION AND INTEGRAL REPRESENTATION PROBLEMS IN THE CALCULUS OF VARIATIONS, Pitman Res. Notes Math. Ser., Longman, Harlow.

[8] CANNARSA P. & FRANKOWSKA H. (1991) *Some characterizations of optimal trajectories in control theory*, SIAM J. on Control

[9] I. CAPUZZO-DOLCETTA and P.-L. LIONS, "Hamilton Jacobi Equations with State Constraints", *Trans. Am. Math. Soc.*, 318 (1990), pp.643-685.

[10] CESARI L. (1983) OPTIMIZATION THEORY AND APPLICATIONS. PROBLEMS WITH ORDINARY DIFFERENTIAL EQUATIONS, Appl. Math. 17, Springer-Verlag, Berlin.

[11] CLARKE F.H. (1983) OPTIMIZATION AND NONSMOOTH ANALYSIS, Wiley-Interscience

[12] CRANDALL M.G., EVANS L.C. & LIONS P.L. (1984) *Some properties of viscosity solutions of Hamilton-Jacobi equation* , Trans. Amer. Math. Soc., 282(2), 487-502

[13] CRANDALL M.G. & LIONS P.L. (1983) *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277, 1-42

[14] FILIPPOV A.F. (1958) *On some problems of optimal control theory*, Vestnik Moskowskogo Universiteta, Math. no.2, 25-32

[15] FILIPPOV A.F. (1967) *Classical solutions of differential equations with multivalued right hand side*, SIAM J. on Control, 5, 609-621

[16] FLEMING W. H. & RISHEL R. W. (1975) DETERMINISTIC AND STOCHASTIC OPTIMAL CONTROL, Springer-Verlag

[17] FRANKOWSKA H. (1989) *Contingent cones to reachable sets of control systems*, SIAM J. on Control and Optimization, 27, 170-198

[18] FRANKOWSKA H. (1989) *Hamilton-Jacobi equation: viscosity solutions and generalized gradients*, J. of Math. Analysis and Appl. 141, 21-26)

[19] FRANKOWSKA H. (1989) *Non smooth solutions to an Hamilton-Jacobi equation*, Proceedings of the International Conference Bellman Continuum, Antibes, France,

June 13-14, 1988, Lecture Notes in Control and Information Sciences, Springer Verlag

[20] FRANKOWSKA H. (1989) *Optimal trajectories associated to a solution of contingent Hamilton-Jacobi equations*, Applied Mathematics and Optimization, 19, 291-311

[21] FRANKOWSKA H. (1989) *Set-valued analysis and some control problems*, Proceedings of the International Conference 30 YEARS OF MODERN CONTROL THEORY, Kingston, June 3-6, 1988 E. Roxin Editor, Marcel Dekker

[22] FRANKOWSKA H. (1990) *Some inverse mapping theorems*, Ann. Inst. Henri Poincar, Analyse Non Linaire, 3, 183-234

[23] FRANKOWSKA H. (1991) *Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equations*, Proceedings of IEEE CDC Conference, Brighton, England, December 1991.

[24] FRANKOWSKA H. (1993) *Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **31**, 257-272.

[25] FRANKOWSKA H., PLASKACZ S. & RZEŻUCHOWSKI T. (1995) *Measurable viability theorems and Hamilton-Jacobi-Bellman equation*, J. Differential Equations **116**, 265-305.

[26] FRANKOWSKA H. & RAMPAZZO F., (2000) *Filippov's and Filippov-Ważewski's Theorems on Closed Domains*, J.D.E.

[27] GUSEINOV H. G., SUBBOTIN A. I. & USHAKOV V. N. (1985) *Derivatives for multivalued mappings with applications to game theoretical problems of control*, Problems of Control and Information Theory, Vol.14, 155-167

[28] HADDAD G. (1981) *Monotone trajectories of differential inclusions with memory*, Isr. J. Math., 39, 83-100

[29] IOFFE A.D. (1977) *On lower semicontinuity of integral functionals*, SIAM J. Control Optim. **15**, 521-521 and 991-1000.

[30] LIONS P.-L. (1982) GENERALIZED SOLUTIONS OF HAMILTON-JACOBI EQUATIONS, Pitman, Boston

[31] MARCHAUD H. (1934) *Sur les champs de demi-cônes et les équations différentielles du premier ordre*, Bull. Sc. Math., 62, 1-38

[32] MARCHAUD H. (1936) *Sur les champs continus de demi-cônes convexes et leurs intégrales*, Compos. Math. 1, 89-127

[33] NAGUMO M. (1942) *Uber die lage der integralkurven gewhnlicher differentialgleichungen*, Proc. Phys. Math. Soc. Japan, 24, 551-559

[34] OLECH C. (1969) *Existence theorems for optimal control problems involving multiple integrals*, J. Diff. Eq., 6, 512-526

[35] OLECH C. (1969) *Existence theorems for optimal control problems with vector-valued cost functions*, Trans. Am. Math. Soc., 136, 157-180

[36] OLECH C. (1976) *Weak lower semicontinuity of integral functionals*, J. Optim. Theory Appl. **19**, 3-16

[37] ROCKAFELLAR T. (1981) *Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization*, Math. Oper. Res. **6**, 424-436.

[38] ROCKAFELLAR T. & WETS R. (1998) VARIATIONAL ANALYSIS, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin.

[39] SUBBOTIN A. I. (1980) *A generalization of the basic equation of the theory of the differential games*, Soviet. Math. Dokl., 22, 358-362.

[40] SONER H.M. (1986) *Optimal Control with State-Space Constraints*, SIAM J. Control and Optimization, 24, pp. 552-561.

[41] WAŻEWSKI T. (1963) *On an optimal control problem*, Proc. Conference DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS, PRAGUE, 1962, 229-242

[42] ZAREMBA S.C. (1934) *Sur une extension de la notion d'équation diférentielle*, COMPTES RENDUS ACAD. SCIENCES 199, 545-548

[43] ZAREMBA S.C. (1936) *Sur les équations au paratingent*, Bull. Sc. Math., 60, 139-160