

Summer School on Mathematical Control Theory

(3 - 28 September 2001)

Value Function in Optimal Control

Lecture 1

Hélele Frankowska

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

These are preliminary lecture notes, intended only for distribution to participants

VALUE FUNCTION IN OPTIMAL CONTROL

Hélène Frankowska, CNRS, FRANCE

TRIESTE, September 2001

VALUE FUNCTION FOR MAYER PROBLEM

Consider $T > 0$, a complete separable metric space U and a map $f : R^n \times U \mapsto R^n$. We associate with it the control system

$$(1) \quad x'(t) = f(x(t), u(t)), \quad u(t) \in U$$

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

$$\min \{g(x(T)) \mid x \text{ is a solution to (1), } x(0) = \xi_0\}$$

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

$$V(t_0, x_0) = \inf \{g(x(T)) \mid x \text{ solves (1), } x(t_0) = x_0\}$$

We replaced one problem by a family of problems.

In general V is nonsmooth, but still contains a lot of information about optimal solutions.

The concept of control can be described as the process of **influencing the behavior of a dynamical system** so as to achieve the desired goal:

to maximize a profit, to minimize the energy, to get from one point to another one, to remain at a given point with the minimum effort, etc.

“After correctly stating the problem of optimal control and having at hand some satisfactory existence theorems, augmented by necessary conditions for optimality, we can consider that we have sufficiently substantial basis to study some special problems, as for instance Moon Flight Problem”.

From a book on Optimal Control, 1969

“Control of nonlinear dynamical systems is an area that has seen some major theoretical developments in the last fifteen years. At the same time, it contains major unsolved mathematical problems, some of which relate to very practical application issues. In engineering practice, nonlinear systems are omnipresent; however, most of them have been designed by using traditional linear regulation techniques.”

“The vast bulk of the theory is linear. Moreover, the control paradigms involved are generally linear also. This leaves the subject in the strange position of **treating highly nonlinear physical systems being controlled by discrete combinatorial mechanisms (computers) using linear models**. Surely a case of looking, not where the penny is lost, but where the street lamp shines!”

From the Report of the Panel
“Future Directions in Control Theory:
A Mathematical Perspective”, 1988

OUTLINE

1. Differential Inclusions and Control Systems.
Basic theorems of Differential Inclusions.

2. Value Function and Optimal Feedback.
Necessary and sufficient conditions for optimality. Uniqueness of optimal solutions and differentiability of value function.

3. Hamilton-Jacobi-Bellman equation.
Viability Theorem, lower semicontinuous and viscosity solutions.

4. Value Function of Bolza problem.
Hamilton-Jacobi equation and characteristics.
Riccati equations and shocks of characteristics.
Smoothness of Value Function.

5. Hamilton-Jacobi-Bellman Equation for
constrained optimal control problems.
Neighbouring Feasible Trajectories Theorem.

Nonlinear Control Systems and Differential Inclusions

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 $U \subset \mathcal{Z}$. We associate with these data the control system

$$(2) \quad x' = f(t, x, u(t)), \quad u(t) \in U, \quad t \in [t_0, T]$$

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a solution to (2) if there exists a measurable map $u : [t_0, T] \mapsto U$, called *admissible control*, such that

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

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

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

Reduction to Differential Inclusion

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

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

and consider the differential inclusion

$$(3) \quad x'(t) \in F(t, x(t)) \text{ a.e. in } [t_0, T]$$

Clearly every solution x to control system (2) satisfies (3).

Solutions to Differential Inclusions

Consider a set-valued map F from $[t_0, T] \times R^n$ into subsets of R^n , i.e. for every (t, x) , $F(t, x) \subset R^n$. We associate with it the differential inclusion

$$(4) \quad x' \in F(t, x)$$

An absolutely continuous function $x : [t_0, T] \mapsto R^n$ is called a **solution** to (4) if

$$(5) \quad x'(t) \in F(t, x(t)) \text{ a.e. in } [t_0, T]$$

We denote by $\mathcal{S}_{[t_0, T]}(x_0)$ the set of solutions to the differential inclusion (4) starting at $x_0 \in R^n$ at time t_0 and defined on the time interval $[t_0, T]$:

$$\mathcal{S}_{[t_0, T]}(x_0) = \{x \mid x \text{ solves (4) on } [t_0, T], x(t_0) = x_0\}$$

The natural question arises whether (3) has the same solutions than the control system (2)? The answer is positive for a large class of maps f .

We impose the following assumptions on f and U :

$$(6) \quad \begin{cases} \forall (x, u), f(\cdot, x, u) \text{ is measurable} \\ \forall t \in [t_0, T], f(t, \cdot, \cdot) \text{ is continuous} \\ U \text{ is nonempty and closed} \end{cases}$$

Theorem 1.1 *Assume that (6) holds true. Then the set of solutions to control system (2) coincide with the set of solutions to differential inclusion (3).*

State Dependent Control Systems

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

$$U : \mathbf{R}^n \rightsquigarrow \mathcal{Z}$$

be a given set-valued map, i.e. $U(x) \subset \mathcal{Z}$. Consider the control system

$$(7) \quad x' = f(t, x, u), \quad u \in U(x), \quad t \in [t_0, T]$$

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a solution to (7) if for some measurable selection $u(t) \in U(x(t))$ we have

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

Reduction to Differential Inclusion

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

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

and replace (7) by the differential inclusion

$$(8) \quad x'(t) \in F(t, x(t)) \text{ a.e. in } [t_0, T]$$

We impose the following assumptions:

$$(9) \quad \begin{cases} \forall (x, u), f(\cdot, x, u) \text{ is measurable} \\ \forall t \in [t_0, T], f(t, \cdot, \cdot) \text{ is continuous} \\ U \text{ is } \mathbf{continuous} \text{ with closed images} \end{cases}$$

Theorem 1.2 *If (9) holds true, then the sets of solutions to control system (7) and differential inclusion (8) do coincide.*

Nonlinear Implicit Control Systems

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

$$(10) \quad f(x, x', u(t)) = 0, \quad u(t) \in U, \quad t \in [t_0, T]$$

An absolutely continuous function $x : [t_0, T] \mapsto \mathbf{R}^n$ is called a solution to (10) 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 $[t_0, T]$.

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

$$(11) \quad x'(t) \in F(x(t)) \quad \text{a.e. in } [t_0, T]$$

Clearly every solution to (10) solves (11).

Lemma 1.3 *If f is continuous, then the solution sets of (11) and (10) do coincide.*

REGULARITY OF SET-VALUED MAPS

Let X, Y be 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 .

A set-valued map F is called **upper semicontinuous** at x if and only if for any neighborhood \mathcal{U} of $F(x)$,

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

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

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

If F is single-valued both notions coincide.

F is called **continuous** if it is both lower and upper semicontinuous.

Definition 1.4 *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 $K \subset X$ if there exists $L \geq 0$ such that*

$$\forall x_1, x_2 \in K, \quad F(x_1) \subset F(x_2) + Ld_X(x_1, x_2)B$$

where B denotes the closed unit ball in Y .

FILIPPOV'S THEOREM

Consider an absolutely continuous

$$y : [t_0, T] \mapsto \mathbb{R}^n$$

Filippov's theorem provides an estimate of the distance from y to the set $\mathcal{S}_{[t_0, T]}(x_0)$ under the following assumptions on F :

- $i) \quad \forall x \in \mathbb{R}^n, F(x)$ is closed
- $ii) \quad \exists \beta > 0, \forall t, F$ is k -Lipschitz on $y(t) + \beta B$

Theorem 1.5 *Let $\delta \geq 0$ and set*

$$\gamma(s) = \text{dist}(y'(s), F(y(s)))$$

$$\eta(t) = e^{k(t-t_0)}\delta + \int_{t_0}^t \gamma(s) e^{k(t-s)} ds$$

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

$$\|x(t) - y(t)\| \leq \eta(t) \text{ for all } t \in [t_0, T]$$

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

LIPSCHITZ DEPENDENCE ON INITIAL CONDITIONS

Corollary 1.6 *Let $y_0 \in R^n$, $y \in \mathcal{S}_{[t_0, T]}(y_0)$. Assume that F has closed images and is k -Lipschitz on a neighborhood of $y([t_0, T])$. Then there exists $\delta > 0$ depending only on k such that for all $x_0 \in B(y_0, \delta)$ we have*

$$\inf_{x \in \mathcal{S}_{[t_0, T]}(x_0)} \|x - y\|_{W^{1,1}} \leq e^{k(T-t_0)} \|x_0 - y_0\|$$

Filippov's theorem allows to estimate the effect on the solutions of **perturbations** of dynamics F or initial conditions. It is as useful as Gronwall's lemma...

EXAMPLE OF APPLICATION OF FILIPPOV'S THEOREM

Corollary 1.7 *We assume that F is locally Lipschitz at x_0 and has closed images. Then for every $u \in F(x_0)$ 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 $u \in F(x_0)$. It is enough to consider the absolutely continuous function

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

Then for all $t \in [t_0, T]$ such that $(t - t_0) \|u\| \leq \beta$

$$d(u, F(y(t))) \leq d(u, F(x_0)) + k \|y(t) - x_0\|$$

$$= k(t - t_0) \|u\|$$

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

$$\|x(t) - y(t)\| \leq \int_{t_0}^t k(s - t_0) \|u\| e^{k(t-s)} ds$$

$$\leq \frac{1}{2} e^{k(t-t_0)} k(t - t_0)^2 \|u\|$$

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

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

RELAXATION THEOREMS

Let $x_0 \in R^n$. We compare next solutions to the differential inclusion

$$(12) \quad \begin{cases} x'(t) \in F(x(t)) & \text{a.e. in } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

and to the convexified (relaxed) differential inclusion:

$$(13) \quad \begin{cases} x'(t) \in \overline{\text{co}} F(x(t)) & \text{a.e. in } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

Observe that if F is Lipschitz on a set K , then so does the set-valued map $x \rightsquigarrow \overline{\text{co}}F(x)$, where $\overline{\text{co}}$ denotes the closed convex hull.

Theorem 1.8 *Let $y : [t_0, T] \mapsto R^n$ be a solution to the relaxed inclusion (13). Assume that F has closed images and is Lipschitz on a neighborhood of $y([t_0, T])$.*

Then for every $\varepsilon > 0$ there exists a solution x to (13) such that $\|x - y\|_C \leq \varepsilon$.

Example Consider two control systems

$$x' = u, \quad u \in \{-1, 1\}, \quad x(0) = x_0$$

$$x' = u, \quad u \in [-1, 1], \quad x(0) = x_0$$

The goal is not to move from x_0 .

In the second case it is enough to apply the control $u \equiv 0$.

In the first case there is no solution, but it is possible to remain in any small neighborhood of x_0 by switching controls between -1 and +1 often enough.

Theorem 1.9 (Relaxation) *Let $F : R^n \rightsquigarrow R^n$ be a locally Lipschitz set-valued map with closed images and $x_0 \in R^n$.*

Then solutions to the differential inclusion

$$\begin{cases} x'(t) \in F(x(t)) \text{ a.e. in } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

are dense (in $C(t_0, T; R^n)$) in solutions to the relaxed inclusion

$$\begin{cases} x'(t) \in \overline{\text{co}} F(x(t)) \text{ a.e. in } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

COMPACTNESS OF THE SOLUTION SET

Theorem 1.10 *Let $F : R^n \rightsquigarrow R^n$ be a Lipschitz set-valued map with compact images and $x_0 \in R^n$ and $\mathcal{S}_{[t_0, T]}^{co}(x_0)$ denote the set of solutions to the relaxed inclusion. Then the closure of $\mathcal{S}_{[t_0, T]}(x_0)$ in the metric of uniform convergence is **compact** and is equal to $\mathcal{S}_{[t_0, T]}^{co}(x_0)$.*

Corollary 1.11 *Assume that for every x , $f(x, U)$ is convex and compact and for some $k \geq 0$ $f(\cdot, u)$ is k -Lipschitz for all $u \in U$. If g is lower semicontinuous, then the Mayer problem has an optimal solution and the value function is lower semicontinuous.*

Corollary 1.12 *Assume that for every x , $f(x, U)$ is compact and for some $k \geq 0$ $f(\cdot, u)$ is k -Lipschitz for all $u \in U$. If g is continuous, then the value functions of the Mayer problem and the relaxed Mayer problem coincide and are continuous.*

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 .

Definition 1.13 Let $(K_n)_{n \in \mathbf{N}}$ be a sequence of subsets of a metric space X . We say that

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

is the upper limit of the sequence K_n and

$$\text{Liminf}_{n \rightarrow \infty} K_n := \left\{ x \in X \mid \lim_{n \rightarrow \infty} d(x, K_n) = 0 \right\}$$

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

$$K = \text{Liminf}_{n \rightarrow \infty} K_n = \text{Limsup}_{n \rightarrow \infty} K_n$$

In this case we write $K = \text{Lim}_{n \rightarrow \infty} K_n$.

Naturally, we can replace \mathbf{N} by a metric 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*.

DERIVATIVES OF SET-VALUED MAPS

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

The **adjacent derivative** $D^b F(x, y)$ is defined by: for all $u \in X$

$$D^b F(x, y)(u) := \operatorname{Liminf}_{h \rightarrow 0^+} \frac{F(x + hu) - y}{h}$$

Proposition 1.14 *If the images of F are convex and F is Lipschitz around x , then for any $y \in F(x)$, $D^b F(x, y)$ has convex images and $D^b F(x, y)(0) = T_{F(x)}(y) := \overline{\cup_{\lambda \geq 0} \lambda(F(x) - y)}$*

$$D^b F(x, y)(u) + D^b F(x, y)(0) = D^b F(x, y)(u)$$

Example $F(x) = f(x, U)$. Assume that F has convex images. Fix $\bar{x}, \bar{u} \in U$ and set $\bar{y} = f(\bar{x}, \bar{u})$. Then for all w

$$\frac{\partial f}{\partial x}(\bar{x}, \bar{u})(w) \in D^b F(\bar{x}, \bar{y})(w)$$

$$T_{F(\bar{x})}(\bar{y}) = D^b F(\bar{x}, \bar{y})(0)$$

$$\frac{\partial f}{\partial x}(\bar{x}, \bar{u})(w) + T_{F(\bar{x})}(\bar{y}) \subset D^b F(\bar{x}, \bar{y})(w)$$

VARIATIONAL INCLUSIONS

Theorem 1.15 *Consider the solution map $\mathcal{S}_{[t_0, T]}(\cdot)$ as the set-valued map from R^n to the (Sobolev) space $W^{1,1}(t_0, T; R^n)$ and a solution $y(\cdot)$ to the differential inclusion*

$$\begin{cases} x'(t) \in F(x(t)) \text{ a.e. in } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

Assume that F has closed images and is Lipschitz around $y([t_0, T])$. Let $v \in R^n$ and $w \in W^{1,1}(t_0, T; R^n)$ be a solution to the linearized inclusion

$$(14) \begin{cases} w'(t) \in D^b F(y(t), y'(t))(w(t)) \text{ a.e.} \\ w(t_0) = v \end{cases}$$

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

In particular, $w \in D^b \mathcal{S}(x_0, y(\cdot))(v)$.

Asking Less Convergence Allows to Relax the Differential Inclusion

Theorem 1.16 *Consider the solution map $\mathcal{S}_{[t_0, T]}(\cdot)$ as the set-valued map from R^n to $\mathcal{C}(t_0, T; R^n)$. Let y be a solution to the differential inclusion*

$$\begin{cases} x'(t) \in F(x(t)) \text{ a.e. in } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

Assume that F has closed images and is Lipschitz around $y([t_0, T])$, $v \in R^n$. Let w be a solution to the inclusion

$$\begin{cases} w'(t) \in D^b(\overline{\text{co}} F)(y(t), y'(t))(w(t)) \text{ a.e. in } [t_0, T] \\ w(t_0) = v \end{cases}$$

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

INFINITESIMAL GENERATOR of REACHABLE MAP

Consider a set-valued map $F : \mathbf{R}^n \rightsquigarrow \mathbf{R}^n$. For all $0 \leq t_0 \leq t_1$ 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 differential inclusion from (t_0, ξ) at time t_1 .

The set-valued map R enjoys the following semi-group properties: for all $0 \leq t_1 \leq t_2 \leq t_3$

$$\begin{cases} \forall \xi \in \mathbf{R}^n, & R(t_3, t_2)R(t_2, t_1)\xi = R(t_3, t_1)\xi \\ \forall \xi \in \mathbf{R}^n, & R(t, t)\xi = \xi \end{cases}$$

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

Theorem 1.17 *Let $x_0 \in \mathbf{R}^n$. Assume that F is locally Lipschitz around x_0 and has nonempty, compact images. Then*

$$\overline{co} F(x_0) = \text{Lim}_{h \rightarrow 0^+} \frac{R(h, 0)x_0 - x_0}{h}$$

In other words $R(h, 0)x_0 = x_0 + h\overline{co}F(x_0) + o(h)$.

PARAMETRIZATION OF SET-VALUED MAPS

If it was always the case, one could get back to single-valued maps and to avoid the curved arrows.

Consider a metric space X and $F : X \rightsquigarrow R^n$.

Definition 1.18 *Let U be a metric space. We say that a single-valued map*

$$f : X \times U \mapsto R^n$$

*is a **Lipschitz parametrization** of F if*

$$\begin{cases} i) \quad \forall x \in X, F(x) = f(x, U) \\ ii) \quad \exists k > 0, \forall u \in U, f(\cdot, u) \text{ is } k\text{-Lipschitz} \\ iv) \quad \forall x \in X, f(x, \cdot) \text{ is continuous} \end{cases}$$

Theorem 1.19 *Assume that F is k -Lipschitz and has nonempty compact convex images. Then there exist $5n > c > 0$ and a Lipschitz parametrization f of F with U equal to the closed unit ball B in R^n such that: $\forall (x, u) \in R^n \times U, v \in U$*

$$\begin{cases} f(\cdot, u) \text{ is } ck - \text{Lipschitz} \\ \|f(x, u) - f(x, v)\| \leq c \left(\max_{y \in F(x)} \|y\| \right) \|u - v\| \end{cases}$$