RENDICONTI
DEL SEMINARIO
MATEMATICO

Università e Politecnico di Torino

Control Theory and its Applications

CONTENTS

Volume 56, N. 4 **1998**

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

A. Agrachev

COMPACTNESS FOR SUB-RIEMANNIAN LENGTH-MINIMIZERS AND SUBANALYTICITY

Abstract.
We establish compactness properties for sets of length-minimizing admissible paths of a prescribed small length. This implies subanayticity of small sub-Riemannian balls for a wide class of real-analytic sub-Riemannian structures: for any structure without abnormal minimizers and for many structures without strictly abnormal minimizers.

1. Introduction

Let M be a C^∞ Riemannian manifold, $\dim M = n$. A distribution on M is a smooth linear subbundle Δ of the tangent bundle TM. We denote by Δ_q the fiber of Δ at $q \in M$; $\Delta_q \subset T_q M$. A number $k = \dim \Delta_q$ is the *rank* of the distribution. We assume that $1 < k < n$. The restriction of the Riemannian structure to Δ is a *sub-Riemannian structure*.

Lipschitzian integral curves of the distribution Δ are called *admissible paths*; these are Lipschitzian curves $t \mapsto q(t)$, $t \in [0, 1]$, such that $\dot{q}(t) \in \Delta_{q(t)}$ for almost all t.

We fix a point $q_0 \in M$ and study only admissible paths started from this point, i.e. we impose the initial condition $q(0) = q_0$. Sections of the linear bundle Δ are smooth vector fields; iterated Lie brackets of these vector fields define a flag

$$\Delta_{q_0} \subset \Delta^2_{q_0} \subset \cdots \subset \Delta^m_{q_0} \cdots \subset T_q M$$

in the following way:

$$\Delta^m_{q_0} = \operatorname{span}\{[X_1, [X_2, [\ldots, X_m]\ldots](q_0) : X_i(q) \in \Delta_q,\ i = 1, \ldots, m,\ q \in M\}.$$

A distribution Δ is *bracket generating* at q_0 if $\Delta^m_{q_0} = T_{q_0} M$ for some $m > 0$. If Δ is bracket generating, then according to a classical Rashevski-Chow theorem (see [15, 22]) there exist admissible paths connecting q_0 with any point of an open neighborhood of q_0. Moreover, applying a general existence theorem for optimal controls [16] one obtains that for any q_1 from a small enough neighborhood of q_0 there exists a shortest admissible path connecting q_0 with q_1. The length of this shortest path is the *sub-Riemannian* or *Carnot-Caratheodory distance* between q_0 and q_1.

For the rest of the paper we assume that Δ is bracket generating at the given initial point q_0. We denote by $\rho(q)$ the sub-Riemannian distance between q_0 and q. It follows from the Rashevsky-Chow theorem that ρ is a continuous function defined on a neighborhood of q_0. Moreover, ρ is Hölder-continuous with the Hölder exponent $\frac{1}{m}$, where $\Delta^m_{q_0} = T_{q_0} M$. A *sub-Riemannian sphere* $S(r)$ is the set of all points at sub-Riemannian distance r from q_0, $S(r) = \rho^{-1}(r)$.

In contrast to the Riemannian distance, the sub-Riemannian distance ρ is never smooth in a punctured neighborhood of q_0 (see Theorem 1) and the main motivation for this research is to understand regularity properties of ρ. In the Riemannian case, where all paths are available, the set of shortest paths connecting q_0 with the sphere of a small radius r is parametrized by the points of the sphere. This is not true for the set of shortests admissible paths connecting q_0 with the sub-Riemannian sphere $S(r)$. The structure of the last set may be rather complicated; we show that this set is at least compact in H^1-topology (Theorem 2). The situation is much simpler if no one among so called abnormal geodesics of length r connect q_0 with $S(r)$. In the last case, the mentioned set of shortests admissible paths can be parametrized by a compact part of a cylinder $S^{k-1} \times \mathbb{R}^{n-k}$ (Theorem 3). In Theorem 4 we recall an efficient necessary condition for a length r admissible path to be a shortest one. In Theorem 5 we state a result, which is similar to that of Theorem 3 but more efficient and admitting nonstrictly abnormal geodesics as well.

We apply all mentioned results to the case of real-analytic M and Δ. The main problem here is to know whether the distance function ρ is subanalytic. Positive results for some special classes of distributions were obtained in [8, 17, 19, 20, 23] and the first counterexample was described in [10] (see [13, 14] for further examples and for study of the "transcendence" of ρ).

Both positive results and the counterexamples gave an indication that the problem is intimately related to the existence of abnormal length-minimizers. Corollaries 2, 3, 4 below make this statement a well-established fact: they show very clear that only abnormal length-minimizers may destroy subanalyticity of ρ out of q_0.

What remains? The situation with subanalyticity in a whole neighborhood including q_0 is not yet clarified. This subanalyticity is known only for a rather special type of distributions (the best result is stated in [20]). Another problem is to pass from examples to general statements for sub-Riemannian structures with abnormal length-minimizers. Such length-minimizers are exclusive for rank $k \geq 3$ distributions (see discussion at the end of the paper) and typical for rank 2 distributions (see [7, 21, 24]). A natural conjecture is:
If $k = 2$ and $\Delta^2_{q_0} \neq \Delta^3_{q_0}$, then ρ is not subanalytic.

2. Geodesics

We are working in a small neighborhood O_{q_0} of $q_0 \in M$, where we fix an orthonormal frame $X_1, \ldots, X_k \in \operatorname{Vect} M$ of the sub-Riemannian structure under consideration. Admissible paths are thus solutions to the differential equations

$$\dot{q} = \sum_{i=1}^{k} u_i(t) X_i(q), \; q \in O_{q_0}, \; q(0) = q_0, \tag{1}$$

where $u = (u_1(\cdot), \ldots, u_k(\cdot)) \in L_2^k[0, 1]$.

Below $\|u\| = \left(\int_0^1 \sum_{i=0}^k u_i^2(t)\, dt \right)^{1/2}$ is the norm in $L_2^k[0, 1]$. We also set $\|q(\cdot)\| = \|u\|$, where $q(\cdot)$ is the solution to (1). Let

$$U_r = \{u \in L_2^k[0, 1] : \|u\| = r\}$$

be the sphere of radius r in $L_2^k[0, 1]$. Solutions to (1) are defined for all $t \in [0, 1]$, if u belongs to the sphere of a small enough radius r. In this paper we take u only from such spheres without special mentioning. The length $l(q(\cdot)) = \int_0^1 \left(\sum_{i=1}^k u_i^2(t) \right)^{1/2} dt$ is well-defined and satisfies the inequality

(2) $$l(q(\cdot)) \leq \|q(\cdot)\| = r\,.$$

The length doesn't depend on the parametrization of the curve while the norm $\|u\|$ depends. We say that u and $q(\cdot)$ are *normalized* if $\sum_{i=1}^{k} u_i^2(t)$ doesn't depend on t. For normalized u, and only for them, inequality (2) becoms equality.

REMARK 1. The notations $\|q(\cdot)\|$ and $l(q(\cdot))$ reflect the fact that these quantities do not depend on the choice of the orthonormal frame $X_1, \ldots, X_k$ and are characteristics of the *trajectory* $q(\cdot)$ rather than the *control* u. L_2-topology in the space of controls is H_1-topology in the space of trajectories.

We consider the endpoint mapping $f : u \mapsto q(1)$. This is a well-defined smooth mapping of a neighborhood of the origin of $L_2^k[0, 1]$ into M. We set $f_r = f\big|_{U_r}$. Critical points of the mapping $f_r : U_r \to M$ are called *extremal controls* and correspondent solutions to the equation (1) are called *extremal trajectories* or *geodesics*.

An extremal control u and the correspondent geodesic $q(\cdot)$ are *regular* if u is a regular point of f; otherwise they are *singular* or *abnormal*.

Let C_r be the set of normalized critical points of f_r; in other words, C_r is the set of normalized extremal controls of the length r. It is easy to check that $f_r^{-1}(S(r)) \subset C_r$. Indeed, among all admissible curves of the length no greater than r only geodesics of the length exactly r can reach the sub-Riemannian sphere $S(r)$. Controls $u \in f_r^{-1}(S(r))$ and correspondent geodesics are called *minimal*.

Let $D_u f : L_2^k[0, 1] \to T_{f(u)}M$ be the differential of f at u. Extremal controls (and only them) satisfy the equation

(3) $$\lambda D_u f = \nu u$$

with some "Lagrange multipliers" $\lambda \in T^*_{f(u)}M \setminus 0$, $\nu \in \mathbb{R}$. Here $\lambda D_u f$ is the composition of the linear mapping $D_u f$ and the linear form $\lambda : T_{f(u)}M \to \mathbb{R}$, i.e. $(\lambda D_u f) \in L_2^k[0, 1]^* = L_2^k[0, 1]$. We have $\nu \neq 0$ for regular extremal controls, while for abnormal controls ν can be taken 0. In principle, abnormal controls may admit Lagrange multipliers with both zero and nonzero ν. If it is not the case, then the control and the geodesic are called *strictly abnormal*.

Pontryagin maximum principle gives an efficient way to solve equation (3), i.e. to find extremal controls and Lagrange multipliers. A coordinate free formulation of the maximum principle uses the canonical symplectic structure on the cotangent bundle T^*M. The symplectic structure associates a Hamiltonian vector field $\vec{a} \in \operatorname{Vect} T^*M$ to any smooth function $a : T^*M \to \mathbb{R}$ (see [11] for the introduction to symplectic methods).

We define the functions h_i, $i = 1, \ldots, k$, and h on T^*M by the formulas

$$h_i(\psi) = \langle \psi, X_i(q) \rangle\,, \quad h(\psi) = \frac{1}{2}\sum_{i=1}^{k} h_i^2(\psi)\,, \quad \forall q \in M,\ \psi \in T_q^*M\,.$$

Pontryagin maximum principle implies the following

PROPOSITION 1. *A triple* (u, λ, ν) *satisfies equation (3) if and only if there exists a solution*

$\psi(t),\ 0 \le t \le 1$, *to the system of differential and pointwise equations*

$$\dot{\psi} = \sum_{i=1}^{k} u_i(t)\vec{h}_i(\psi)\,, \quad h_i(\psi(t)) = \nu u_i(t) \tag{4}$$

with boundary conditions $\psi(0) \in T^*_{q_0}M,\ \psi(1) = \lambda$.

Here $(\psi(t), \nu)$ are Lagrange multipliers for the extremal control $u_t : \tau \mapsto tu(t\tau)$; in other words, $\psi(t)D_{u_t}f = \nu u_t$.

Note that abnormal geodesics remain to be geodesics after an arbitrary reparametrization, while regular geodesics are automatically normalized. We say that a geodesic is *quasi-regular* if it is normalized and is not strictly abnormal. Setting $\nu = 1$ we obtain a simple description of all quasi-regular geodesics.

COROLLARY 1. *Quasi-regular geodesics are exactly projections to* M *of the solutions to the differential equation* $\dot{\psi} = \vec{h}(\psi)$ *with initial conditions* $\psi(0) \in T^*_{q_0}M$. *If* $h(\psi(0))$ *is small enough, then such a solution exists (i.e. is defined on the whole segment* $[0, 1]$*). The length of the geodesic equals* $\sqrt{2h(\psi(0))}$ *and the Lagrange multiplier* $\lambda = \psi(1)$.

The next result demonstrates a sharp difference between Riemannian and sub-Riemannian distance functions.

THEOREM 1. *Any neighbourhood of* q_0 *in* M *contains a point* $q \neq q_0$, *where the distance function* ρ *is not continuously differentiable.*

This theorem is a kind of folklore; everybody agrees it is true but I have never seen the proof. What follows is a sketch of the proof.

Suppose ρ is continuously differentiable out of q_0. Take a minimal geodesic $q(\cdot)$ of the length r. Then $\tau \mapsto q(t\tau)$ is a minimal geodesic of the length tr for any $t \in [0, 1]$ and we have $\rho(q(t)) \equiv rt$; hence $\langle d_{q(t)}\rho, \dot{q}(t)\rangle = r$. Since any point of a neighborhood of q_0 belongs to some minimal geodesic, we obtain that ρ has no critical points in the punctured neighborhood. In particular, the spheres $S(r) = \rho^{-1}(r)$ are C^1-hypersurfaces in M. Moreover, $S(r) = \partial f(U_r)$; hence $\left(d_{q(1)}\rho\right) D_u f_r = 0$ and we obtain the equality $\left(d_{q(1)}\rho\right) D_u f = \frac{1}{r}u$, where u is the extremal control associated with $q(\cdot)$. Hence $q(\cdot)$ is the projection to M of the solution to the equation $\dot{\psi} = \vec{h}(\psi)$ with the boundary condition $\psi(1) = rd_{q(1)}\rho$. Moreover, we easily conclude that $\psi(t) = rd_{q(t)}\rho$ and come to the equation

$$\dot{q}(t) = r\sum_{i=1}^{k}\langle d_{q(t)}\rho, X_i(q(t))\rangle X_i(q(t))\,.$$

For the rest of the proof we fix local coordinates in a neighborhood of q_0. We are going to prove that the vector field $V(q) = r\sum_{i=1}^{k}\langle d_q\rho, X_i(q)\rangle X_i(q)$, $q \neq q_0$, has index 1 at its isolated singularity q_0. Let $B_\varepsilon = \{q \in \mathbb{R}^n : |q - q_0| \le \varepsilon\}$ be a so small ball that $\rho(q) < \frac{r}{2}$, $\forall q \in B_\varepsilon$. Let $s \mapsto q(s; q_\varepsilon)$ be the solution to the equation $\dot{q} = V(q)$ with the initial condition $q(0; q_\varepsilon) = q_\varepsilon \in B_\varepsilon$. Then $q(\frac{r}{2}; q_\varepsilon) \notin B_\varepsilon$. In particular, the vector field W_ε on B_ε defined by the formula $W(q_\varepsilon) = q(\frac{r}{2}; q_\varepsilon) - q_\varepsilon$ looks “outward” and has index 1. The family of the fields $V_s(q_\varepsilon) = \frac{1}{s}(q(s; q_\varepsilon) - q_\varepsilon)$, $0 \le s \le \frac{r}{2}$ provides a homotopy of $V\big|_{B_\varepsilon}$ and $\frac{r}{2}W$, hence V has index 1 at q_0 as well.

On the other hand, the field V is a linear combination of $X_1, \dots, X_k$ and takes its values near the k-dimensional subspace $\mathrm{span}\{X_1(q_0), \dots, X_k(q_0)\}$. Such a field must have index 0 at q_0. This contradiction completes the proof.

Corollary 1 gives us a parametrization of the space of quasi-regular geodesics by the poins of an open subset Ψ of $T^*_{q_0}M$. Namely, Ψ consists of $\psi_0 \in T^*_{q_0}M$ such that the solution $\psi(t)$ to the equation $\dot{\psi} = \vec{h}(\psi)$ with the initial condition $\psi(0) = \psi_0$ is defined for all $t \in [0, 1]$. The composition of this parametrization with the endpoint mapping f is the *exponential mapping* $\mathcal{E} : \Psi \to M$. Thus $\mathcal{E}(\psi(0)) = \pi(\psi(1))$, where $\pi : T^*M \to M$ is the canonical projection.

The space of quasi-regular geodesics of a small enough length r are parametrized by the points of the manifold $H(r) = h^{-1}(\frac{r^2}{2}) \cap T^*_{q_0}M \subset \Psi$. Clearly, $H(r)$ is diffeomorphic to $\mathbb{R}^{n-k} \times S^{k-1}$ and $H(sr) = sH(r)$ for any nonnegative s.

All results about subanalyticity of the distance function ρ are based on the following statement. As usually, the distances r are assumed to be small enough.

PROPOSITION 2. *Let M and the sub-Riemannian structure be real-analytic. Suppose that there exists a compact $K \subset h^{-1}(\frac{1}{2}) \cap T^*_{q_0}M$ such that $S(r) \subset \mathcal{E}(rK)$, $\forall r \in (r_0, r_1)$. Then ρ is subanalytic on $\rho^{-1}((r_0, r_1))$.*

Proof. It follows from our assumptions and Corollary 1 that

$$\rho(q) = \min\{r : \psi \in K,\ \mathcal{E}(r\psi) = q\}, \quad \forall q \in \rho^{-1}((r_0, r_1)) .$$

The mapping $\mathcal{E}$ is analytic thanks to the analyticity of the vector field $\vec{h}$. The compact K can obviously be chosen semi-analytic. The proposition follows now from [25, Prop. 1.3.7].

□

3. Compactness

Let $\mathcal{O} \subset L_2^k[0, 1]$ be the domain of the endpoint mapping f. Recall that $\mathcal{O}$ is a neighborhood of the origin of $L_2^k[0, 1]$ and $f : \mathcal{O} \to M$ is a smooth mapping. We are going to use not only defined by the norm "strong" topology in the Hilbert space $L_2^k[0, 1]$, but also weak topology. We denote by $\mathcal{O}_{\mathrm{weak}}$ the topological space defined by weak topology restricted to $\mathcal{O}$.

PROPOSITION 3. *$f : \mathcal{O}_{weak} \to M$ is a continuous mapping.*

This proposition easily follows from some classical results on the continuous dependence of solutions to ordinary differential equations on the right-hand side. Nevertheless, I give an independent proof in terms of the chronological calculus (see [1, 5]) since it is very short. We have

$$\begin{aligned} f(u) &= q_0 \overrightarrow{\exp} \int_0^1 \sum_{i=1}^k u_i(t) X_i \, dt \\ &= q_0 + \sum_{i=1}^k q_0 \int_0^1 \left(u_i(t) \overrightarrow{\exp} \int_0^t \sum_{j=1}^k u_j(t) X_j \, d\tau \right) dt \circ X_i . \end{aligned}$$

The integration by parts gives:

$$\int_0^1 \left(u_i(t) \overrightarrow{\exp} \int_0^t \sum_{j=1}^k u_j(t) X_j \, d\tau \right) dt = \int_0^1 u_i(t)\, dt \, \overrightarrow{\exp} \int_0^1 \sum_{j=1}^k u_j(t) X_j \, dt$$

$$- \sum_{i=1}^k \int_0^1 \left(u_j(t) \int_0^t u_i(\tau)\, d\tau \, \overrightarrow{\exp} \int_0^t \sum_{j=1}^k u_j(t) X_j \, d\tau \right) dt \circ X_j \, .$$

It remains to mention that the mapping $u(\cdot) \mapsto \int_0^{\cdot} u(\tau)\, d\tau$ is a compact operator in $L_2^k[0,1]$. A detailed study of the continuity of $\overrightarrow{\exp}$ in various topologies see in [18].

THEOREM 2. *The set of minimal geodesics of a prescribed length r is compact in H_1-topology for any small enough r.*

Proof. We have to prove that $f_r^{-1}(S(r))$ is a compact subset of U_r. First of all, $f_r^{-1}(S(r)) = f^{-1}(S(r)) \cap \operatorname{conv} U_r$, where $\operatorname{conv} U_r$ is a ball in $L_2^k[0,1]$. This is just because $S(r)$ cannot be reached by trajectories of the length smaller than r. Then the continuity of ρ implies that $S(r) = \rho^{-1}(r)$ is a closed set and the continuity of f in weak topology implies that $f^{-1}(S(r))$ is weakly closed. Since $\operatorname{conv} U_r$ is weakly compact we obtain that $f_r^{-1}(S(r))$ is weakly compact. What remains is to note that weak topology resricted to the sphere U_r in the Hilbert space is equivalent to strong topology.

□

THEOREM 3. *Suppose that all minimal geodesics of the length r are regular. Then we have that $\mathcal{E}^{-1}(S(r)) \cap H(r)$ is compact.*

Proof. Denote by $u_{\psi(0)}$ the extremal control associated with $\psi(0) \in H(r)$ so that $\mathcal{E}(\psi(0)) = f(u_{\psi(0)})$. We have $u_{\psi(0)} = (h_1(\psi(\cdot)), \ldots, h_k(\psi(\cdot)))$ (see Proposition 1 and its Corollary). In particular, $u_{\psi(0)}$ continuously depends on $\psi(0)$.

Take a sequence $\psi_m(0) \in \mathcal{E}^{-1}(S(r)) \cap H(r)$, $m = 1, 2, \ldots$; the controls $u_{\psi_m(0)}$ are minimal, the set of minimal controls of the length r is compact, hence there exists a convergent subsequence of this sequence of controls and the limit is again a minimal control. To simplify notations, we suppose without losing generality that the sequence $u_{\psi_m(0)}$, $m = 1, 2, \ldots$, is already convergent, $\exists \lim_{m\to\infty} u_{\psi_m(0)} = \bar{u}$.

It follows from Proposition 1 that $\psi_m(1) D_{u_{\psi_m(0)}} f = u_{\psi_m(0)}$. Suppose that M is endowed with some Riemannian structure so that the length $|\psi_m(1)|$ of the cotangent vector $\psi_m(1)$ has a sense. There are two possibilities: either $|\psi_m(1)| \to \infty$ $(m \to \infty)$ or $\psi_m(1)$, $m = 1, 2, \ldots$, contains a convergent subsequence.

In the first case we come to the equation $\lambda D_{\bar{u}} f = 0$, where λ is a limiting point of the sequence $\frac{1}{|\psi_m(1)|}\psi_m(1)$, $|\lambda| = 1$. Hence $\bar{u}$ is an abnormal minimal control that contradicts the assumption of the theorem.

In the second case let $\psi_{m_l}(1)$, $l = 1, 2, \ldots$, be a convergent subsequence. Then $\psi_{m_l}(0)$, $l = 1, 2, \ldots$, is also convergent, $\exists \lim_{l\to\infty} \psi_{m_l}(0) = \bar{\psi}(0) \in H(r)$. Then $\bar{u} = u_{\bar{\psi}(0)}$ and we are done.

□

COROLLARY 2. *Let M and the sub-Riemannian structure be real-analytic. Suppose that all minimal geodesics of the length r_0 are regular for some $r_0 < r$. Then ρ is subanalytic on*

$\rho^{-1}((r_0, r])$.

Proof. According to Theorem 3, $K_0 = \mathcal{E}^{-1}(S(r_0)) \cap H(r_0)$ is a compact set and $\{u_{\psi(0)} : \psi(0) \in K_0\}$ is the set of all minimal extremal controls of the length r_0. The minimality of an extremal control $u_{\psi(0)}$ implies the minimality of the control $u_{s\psi(0)}$ for $s < 1$, since $u_{s\psi(0)}(\tau) = su_{\psi(0)}(\tau)$ and a reparametrized piece of a minimal geodesic is automatically minimal. Hence $S(r_1) \subset \mathcal{E}\left(\frac{r_1}{r_0} K_0\right)$ for $r_1 \geq r_0$ and the required subanalyticity follows from Proposition 2.

□

Corollary 2 gives a rather strong sufficient condition for subanalyticity of the distance function ρ out of q_0. In particular, the absence of abnormal minimal geodesics implies subanalyticity of ρ in a punctured neighborhood of q_0. This condition is not however quite satisfactory because it doesn't admit abnormal quasi-regular geodesics. Though being non generic, abnormal quasi-regular geodesics appear naturally in problems with symmetries. Moreover, they are common in so called nilpotent approximations of sub-Riemannian structures at (see [5, 12]). The nilpotent approximation (or nilpotenization) of a generic sub-Riemannian structure q_0 leads to a simplified quasi-homogeneous approximation of the original distance function. It is very unlikely that ρ loses subanalyticity under the nilpotent approximation, although the above sufficient condition loses its validity. In the next section we give chekable sufficient conditions for subanalyticity, wich are free of the above mentioned defect.

4. Second Variation

Let $u \in U_r$ be an extremal control, i.e. a critical point of f_r. Recall that the Hessian of f_r at u is a quadratic mapping

$$\mathrm{Hes}_u f_r : \ker D_u f_r \to \operatorname{coker} D_u f_r \, ,$$

an independent on the choice of local coordinates part of the second derivative of f_r at u. Let (λ, ν) be Lagrange multipliers associated with u so that equation (3) is satisfied. Then the covector $\lambda : T_{f(u)}M \to \mathbb{R}$ annihilates $\operatorname{im} D_u f_r$ and the composition

$$\lambda \mathrm{Hes}_u f_r : \ker D_u f_r \to \mathbb{R} \tag{5}$$

is well-defined.

Quadratic form (5) is the *second variation* of the sub-Riemannian problem at (u, λ, ν). We have

$$\lambda \mathrm{Hes}_u f_r(v) = \lambda D_u^2 f(v, v) - \nu |v|^2 \, , \quad v \in \ker D_u f_r \, .$$

Let $q(\cdot)$ be the geodesic associated with the control u. We set

$$\operatorname{ind}(q(\cdot), \lambda, \nu) = \operatorname{ind}_+(\lambda \mathrm{Hes}_u f_r) - \dim \operatorname{coker} D_u f_r \, , \tag{6}$$

where $\operatorname{ind}_+(\lambda \mathrm{Hes}_u f_r)$ is the positive inertia index of the quadratic form $\lambda \mathrm{Hes}_u f_r$. Decoding some of the symbols we can re-write:

$$\begin{aligned} \operatorname{ind}(q(\cdot), \lambda, \nu) \;\; = \;\; & \sup\{\dim V : V \subset \ker D_u f_r, \ \lambda D_u^2 f(v, v) > \nu |v|^2, \ \forall v \in V \setminus 0\} \\ & - \dim\{\lambda' \in T^*_{f(u)} M : \lambda' D_u f_r = 0\} \, . \end{aligned}$$

The value of $\operatorname{ind}(q(\cdot), \lambda, \nu)$ may be an integer or $+\infty$.

REMARK 2. Index (5) doesn't depend on the choice of the orthonormal frame $X_1, \dots, X_k$ and is actually a characteristic of the geodesic $q(\cdot)$ and the Lagrange multipliers (λ, ν). Indeed, a change of the frame leads to a smooth transformation of the Hilbert manifold U_r and to a linear transformation of variables in the quadratic form $\lambda \mathrm{Hes}_u f_r$ and linear mapping $D_u f_r$. Both terms in the right-hand side of (5) remain unchanged.

PROPOSITION 4. $(u, \lambda, \nu) \mapsto \mathrm{ind}(q(\cdot), \lambda, \nu)$ *is a lower semicontinuous function on the space of solutions of (3).*

Proof. We have $\dim \operatorname{coker} D_u f_r = \operatorname{codim} \ker D_u f_r$. Here $\ker D_u f_r = \ker D_u f \cap \{u\}^{\perp} \subset L_2^k[0,1]$ is a subspace of finite codimension in $L_2^k[0,1]$. The multivalued mapping $u \longmapsto (\ker D_u f_r) \cap U_r$ is upper semicontinuous in the Hausdorff topology, just because $u \mapsto D_u f$ is continuous.

Take (u, λ, ν) satisfying (3). If u' is close enough to u, then $\ker D_{u'} f_r$ is arbitraryly close to a subspace of codimension

$$\dim \operatorname{coker} D_u f_r - \dim \operatorname{coker} D_{u'} f_r$$

in $D_u f_r$. Suppose $V \subset \ker D_u f_r$ is a finite-dimensional subspace such that $\lambda \mathrm{Hes}_u f_r\big|_V$ is a positive definite quadratic form. If u' is sufficiently close to u, then $\ker D_{u'} f_r$ contains a subspace V' of dimension

$$\dim V - (\dim \operatorname{coker} D_u f_r - \dim \operatorname{coker} D_{u'} f_r)$$

that is arbitrarily close to a subspace of V. If λ' is sufficiently close to λ, then the quadratic form $\lambda' \mathrm{Hes}_{u'} f_r\big|_{V'}$ is positive definite.

We come to the inequality $\mathrm{ind}(q'(\cdot), \lambda', \nu') \geq \mathrm{ind}(q(\cdot), \lambda, \nu)$ for any solution (u', λ', ν') of (3) close enough to (u, λ, ν); here $q'(\cdot)$ is the geodesic associated to the control u'.

□

THEOREM 4. *If* $q(\cdot)$ *is minimal geodesic, then there exist associated with* $q(\cdot)$ *Lagrange multipliers* λ, ν *such that* $\mathrm{ind}(q(\cdot), \lambda, \nu) < 0$.

This theorem is a direct corollary of a general result announced in [2] and proved in [3]; see also [8] for the updated proof of exactly this corollary.

THEOREM 5. *Suppose that* $\mathrm{ind}(q(\cdot), \lambda, 0) \geq 0$ *for any abnormal geodesic* $q(\cdot)$ *of the length* r *and associated Lagrange multipliers* $(\lambda, 0)$. *Then there exists a compact* $K_r \subset H(r)$ *such that* $S(r) = \mathcal{E}(K_r)$.

Proof. We use notations introduced in the first paragraph of the proof of Theorem 3. Let $q_{\psi(0)}$ be the geodesic associated to the control $u_{\psi(0)}$. We set

$$K_r = \{\psi(0) \in H(r) \cap \mathcal{E}^{-1}(S(r)) : \mathrm{ind}(q_{\psi(0)}, \psi(1), 1) < 0\}. \tag{7}$$

It follows from Theorem 4 and the assumption of Theorem 5 that $\mathcal{E}(K_r) = S(r)$. What remains is to prove that K_r is compact.

Take a sequence $\psi_m(0) \in K_r$, $m = 1, 2, \dots$; the controls $u_{\psi_m(0)}$ are minimal, the set of minimal controls of the length r is compact, hence there exists a convergent subsequence of this sequence of controls and the limit is again a minimal control. To simplify notations, we suppose without losing generality that the sequence $u_{\psi_m(0)}$, $m = 1, 2, \ldots$, is already convergent, $\exists \lim_{m\to\infty} u_{\psi_m(0)} = \bar{u}$.

It follows from Proposition 1 that $\psi_m(1) D_{u_{\psi_m(0)}} f = u_{\psi_m(0)}$. There are two possibilities: either $|\psi_m(1)| \to \infty$ $(m \to \infty)$ or $\psi_m(1)$, $m = 1, 2, \ldots$, contains a convergent subsequence.

In the first case we come to the equation $\bar{\lambda} D_{\bar{u}} f = 0$, where $\bar{\lambda}$ is a limiting point of the sequence $\frac{1}{|\psi_m(1)|}\psi_m(1)$, $|\bar{\lambda}| = 1$. Lower semicontinuity of $\mathrm{ind}(q(\cdot), \lambda, \nu)$ implies the inequality $\mathrm{ind}(\bar{q}(\cdot), \bar{\lambda}, 0) < 0$, where $\bar{q}(\cdot)$ is the geodesic associated with the control $\bar{u}$. We come to a contradiction with the assumption of the theorem.

In the second case let $\psi_{m_l}(1)$, $l = 1, 2, \ldots$, be a convergent subsequence. Then $\psi_{m_l}(0)$, $l = 1, 2, \ldots$, is also convergent, $\exists \lim_{l\to\infty} \psi_{m_l}(0) = \bar{\psi}(0) \in H(r)$. Then $\bar{u} = u_{\bar{\psi}(0)}$ and $\mathrm{ind}(\bar{q}(\cdot), \bar{\psi}(1), 1) < 0$ because of lower semicontinuity of $\mathrm{ind}(q(\cdot), \lambda, \nu)$. Hence $\bar{\psi}(0) \in K_r$ and we are done.

□

COROLLARY 3. *Let M and the sub-Riemannian structure be real-analytic. Suppose $r_0 < r$ is such that* $\mathrm{ind}(q(\cdot), \lambda, 0) \geq 0$ *for any abnormal geodesic $q(\cdot)$ of the length r_0 and associated Lagrange multipliers* $(\lambda, 0)$. *Then ρ is subanalytic on* $\rho^{-1}((r_0, r])$.

Proof. Let K_{r_0} be defined as in (7). Then K_{r_0} is compact and $\{u_{\psi(0)} : \psi(0) \in K_{r_0}\}$ is the set of all minimal extremal controls of the length r_0. The minimality of an extremal control $u_{\psi(0)}$ implies the minimality of the control $u_{s\psi(0)}$ for $s < 1$, since $u_{s\psi(0)}(\tau) = s u_{\psi(0)}(\tau)$ and a reparametrized piece of a minimal geodesic is automatically minimal. Hence $S(r_1) \subset \mathcal{E}\left(\frac{r_1}{r_0} K_{r_0}\right)$ for $r_1 \geq r_0$ and the required subanalyticity follows from Proposition 2.

□

Among 2 terms in expression (6) for $\mathrm{ind}(q(\cdot), \lambda, \nu)$ only the first one, the inertia index of the second variation, is nontrivial to evaluate. Fortunately, there is an efficient way to compute this index for both regular and singular (abnormal) geodesics, as well as a good supply of conditions that garantee the finiteness or infinity of the index (see [2, 4, 6, 9]). The simplest one is the *Goh condition* (see [6]):

If $\mathrm{ind}(q(\cdot), \psi(1), 0) < +\infty$, then $\psi(t)$ annihilates $\Delta^2_{q(t)}$, $\forall t \in [0, 1]$.

Recall that $\psi(t)$ annihilates $\Delta_{q(t)}$, $0 \leq t \leq 1$, for any Lagrange multiplier $(\psi(1), 0)$ associated with $q(\cdot)$. We say that $q(\cdot)$ is a *Goh geodesic* if there exist Lagrange multipliers $(\psi(1), 0)$ such that $\psi(t)$ annihilates $\Delta^2_{q(t)}$, $\forall t \in [0, 1]$. In particuar, strictly abnormal minimal geodesics must be Goh geodesics. Besides that, the Goh condition and Corollary 3 imply

COROLLARY 4. *Let M and the sub-Riemannian structure be real-analytic and $r_0 < r$. If there are no Goh geodesics of the length r_0, then ρ is subanalytic on* $\rho^{-1}((r_0, r])$.

I'll finish the paper with a brief analysis of the Goh condition. Suppose that $q(\cdot)$ is an abnormal geodesic with Lagrange multipliers $(\psi(1), 0)$, and $k = 2$. Differentiating the identities $h_1(\psi(t)) = h_2(\psi(t)) = 0$ with respect to t, we obtain $u_2(t)\{h_2, h_1\}(\psi(t)) = u_1(t)\{h_1, h_2\}(\psi(t)) = 0$, where $\{h_1, h_2\}(\psi(t)) = \langle \psi(t), [X_1, X_2](q(t)) \rangle$ is the Poisson bracket. In other words, the Goh condition is automatically satisfied by any abnormal geodesic.

The situation changes dramatically if $k > 2$. In order to understand why, we need some notation. Take $\lambda \in T^*M$ and set

$$b_0(\lambda) = \big(\{h_1, h_2\}(\lambda), \{h_1, h_3\}(\lambda), \dots, \{h_{k-1}, h_k\}(\lambda)\big),$$

a vector in $\mathbb{R}^{\frac{k(k-1)}{2}}$ whose coordinates are numbers $\{h_i, h_j\}(\lambda)$, $1 \le i < j \le k$, with lexicographically ordered indeces (i, j). Set also $\beta_0 = \frac{k(k-1)}{2}$. The Goh condition for $q(\cdot), \psi(1)$ implies the identity $b_0(\psi(t)) = 0, \forall t \in [0, 1]$. The differentiation of this identity with respect to t in virtue of (4) gives the equality

$$\sum_{i=1}^{k} u_i(t)\{h_i, b_0\}(\psi(t)) = 0, \quad 0 \le t \le 1. \tag{8}$$

Consider the space $\bigwedge^k \mathbb{R}^{\beta_0}$, the k-th exterior power of $\mathbb{R}^{\beta_0}$. The standard lexicographic basis in $\bigwedge^k \mathbb{R}^{\beta_0}$ gives the identification $\bigwedge^k \mathbb{R}^{\beta_0} \cong \mathbb{R}^{\binom{\beta_0}{k}}$. We set $\beta_1 = \beta + \binom{\beta_0}{k}$ and

$$b_1(\lambda) = (b_0(\lambda), \{h_1, b_0\}(\lambda) \wedge \cdots \wedge \{h_k, b_0\}(\lambda)) \in \mathbb{R}^{\beta_1}.$$

Equality (8) implies: $b_1(\psi(t)) = 0, 0 \le t \le 1$.

Now we set by induction $\beta_{i+1} = \beta_i + \binom{\beta_i}{k}$, $i = 0, 1, 2, \dots$, and fix identifications $\mathbb{R}^{\beta_i} \times \mathbb{R}^{\binom{\beta_i}{k}} \cong \mathbb{R}^{\beta_{i+1}}$. Finaly, we define

$$b_{i+1}(\lambda) = (b_i(\lambda), \{h_1, b_i\}(\lambda) \wedge \cdots \wedge \{h_k, b_i\}(\lambda)) \in \mathbb{R}^{\beta_{i+1}}, \ i = 1, 2, \dots.$$

Successive differentiations of the Goh condition give the equations $b_i(\psi(t)) = 0$, $i = 1, 2, \dots$. It is easy to check that the equation $b_{i+1}(\lambda) = 0$ is not, in general, a consequence of the equation $b_i(\lambda) = 0$ and we indeed impose more and more restrictive conditions on the locus of Goh geodesics.

A natural conjecture is that admitting Goh geodesics distributions of rank $k > 2$ form a set of infinite codimension in the space of all rank k distributions, i.e. they do not appear in generic smooth families of distributions parametrized by finite-dimensional manifolds. It may be not technically easy, however, to turn this conjecture into the theorem.

Anyway, Goh geodesics are very exclusive for the distributions of rank greater than 2. Yet they may become typical under a priori restictions on the growth vector of the distribution (see [6]).

Note in proof. An essential progress was made while the paper was waiting for the publication. In particular, the conjecture on Goh geodesics has been proved as well as the conjecture stated at the end of the Introduction. These and other results will be included in our joined paper with Jean Paul Gauthier, now in preparation.

References

[1] AGRACHEV A. A., GAMKRELIDZE R. V., *The exponential representation of flows and chronological calculus*, Matem. Sbornik **107** (1978), 467–532; English transl. in: Math. USSR Sbornik **35** (1979), 727–785.

[2] AGRACHEV A. A., GAMKRELIDZE R. V., *The index of extremality and quasi-extremal controls*, Dokl. AN SSSR **284** (1985); English transl. in: Soviet Math. Dokl. **32** (1985), 478–481.

[3] AGRACHEV A. A., GAMKRELIDZE R. V., *Quasi-extremality for control systems*, Itogi Nauki i Tekhn, VINITI, Moscow. Ser. Sovremennye Problemy Matematiki, Novejshie Dostizheniya **35** (1989), 109–134; English transl. in: J. Soviet Math. (Plenum Publ. Corp.) (1991), 1849–1864.

[4] AGRACHEV A. A., *Quadratic mappings in geometric control theory*, Itogi Nauki i Tekhn, VINITI, Moscow. Ser. Problemy Geometrii **20** (1988), 111–205; English transl. in: J. Soviet Math. (Plenum Publ. Corp.) **51** (1990), 2667–2734.

[5] AGRACHEV A. A., GAMKRELIDZE R. V., SARYCHEV A. V., *Local invariants of smooth control systems*, Acta Applicandae Mathematicae **14** 1989, 191–237.

[6] AGRACHEV A. A., SARYCHEV A. V., *Abnormal sub-Riemannian geodesics: Morse index and rigidity*, Annales de l'Institut Henri Poincaré-Analyse non linéaire **13** (1996), 635–690.

[7] AGRACHEV A. A., SARYCHEV A. V., *Strong minimality of abnormal geodesics for 2-distributions*, J. Dynamical and Control Systems **1** (1995), 139–176.

[8] AGRACHEV A. A., SARYCHEV A. V., *Sub-Riemannian metrics: minimality of abnormal geodesics versus subanalyticity*, Preprint Univ. Bourgogne, Lab. Topologie, October 1998, 30 p.

[9] AGRACHEV A. A., *Feedback invariant optimal control theory, II. Jacobi curves for singular extremals*, J. Dynamical and Control Systems **4** (1998), 583–604.

[10] AGRACHEV A. A., BONNARD B., CHYBA M., KUPKA I., *Sub-Riemannian sphere in Martinet flat case*, J. ESAIM: Control, Optimisation and Calculus of Variations **2** (1997), 377–448.

[11] ARNOL'D V. I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York-Berlin 1978.

[12] BELLAÏCHE A., *The tangent space in sub-Riemannian geometry*, in the book: "Sub-Riemannian geometry", Birkhäuser 1996, 1–78.

[13] BONNARD B., CHYBA M., *Méthodes géométriques et analytique pour étudier l'application exponentielle, la sphère et le front d'onde en géometrie SR dans le cas Martinet*, J. ESAIM: Control, Optimisation and Calculus of Variations, submitted.

[14] BONNARD B., LAUNAY G., TRÉLAT E., *The transcendence we need to compute the sphere and the wave front in Martinet SR-geometry*, Proceed. Int. Confer. Dedicated to Pontryagin, Moscow, Sept.'98, to appear.

[15] CHOW W-L., *Über Systeme von linearen partiellen Differentialgleichungen ester Ordnung*, Math. Ann. **117** (1939), 98–105.

[16] FILIPPOV A. F., *On certain questions in the theory of optimal control*, Vestnik Moskov. Univ., Ser. Matem., Mekhan., Astron. **2** (1959), 25–32.

[17] ZHONG GE, *Horizontal path space and Carnot-Caratheodory metric*, Pacific J. Mathem. **161** (1993), 255–286.

[18] SARYCHEV A. V., *Nonlinear systems with impulsive and generalized functions controls*, in the book: "Nonlinear synthesis", Birkhäuser 1991, 244–257.

[19] JAQUET S., *Distance sous-riemannienne et sous analycité*, Thèse de doctorat, 1998.

[20] JACQUET S., *Subanalyticity of the sub-Riemannian distance*, J. Dynamical and Control Systems, submitted.

[21] MONTGOMERY R., *A Survey on singular curves in sub-Riemannian geometry*, J. of Dynamical and Control Systems **1** (1995), 49–90.

[22] RASHEVSKY P. K., *About connecting two points of a completely nonholonomic space by admissible curve*, Uch. Zapiski Ped. Inst. Libknechta **2** (1938), 83-94.

[23] SUSSMANN H. J., *Optimal control and piecewise analyticity of the distance function*, in: Ioffe A., Reich S., Eds., Pitman Research Notes in Mathematics, Longman Publishers 1992, 298–310.

[24] SUSSMANN H. J., LIU W., *Shortest paths for sub-Riemannian metrics on rank 2 distributions*, Mem. Amer. Math. Soc. **564** (1995), 104 p.

[25] TAMM M., *Subanalytic sets in the calculus of variations*, Acta mathematica **46** (1981), 167–199.

AMS Subject Classification: ???.

Andrei AGRACHEV
Steklov Mathematical Institute,
ul. Gubkina 8,
Moscow 117966, Russia
& S.I.S.S.A.
Via Beirut 4,
Trieste 34014, Italy

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

M. Bardi* – S. Bottacin

ON THE DIRICHLET PROBLEM FOR NONLINEAR DEGENERATE ELLIPTIC EQUATIONS AND APPLICATIONS TO OPTIMAL CONTROL

Abstract.

We construct a generalized viscosity solution of the Dirichlet problem for fully nonlinear degenerate elliptic equations in general domains by the Perron-Wiener-Brelot method. The result is designed for the Hamilton-Jacobi-Bellman-Isaacs equations of time-optimal stochastic control and differential games with discontinuous value function. We study several properties of the generalized solution, in particular its approximation via vanishing viscosity and regularization of the domain. The connection with optimal control is proved for a deterministic minimum-time problem and for the problem of maximizing the expected escape time of a degenerate diffusion process from an open set.

Introduction

The theory of viscosity solutions provides a general framework for studying the partial differential equations arising in the Dynamic Programming approach to deterministic and stochastic optimal control problems and differential games. This theory is designed for scalar fully nonlinear PDEs

(1) $$F(x, u(x), Du(x), D^2u(x)) = 0 \text{ in } \Omega ,$$

where Ω is a general open subset of $\mathbb{R}^N$, with the monotonicity property

(2) $$F(x, r, p, X) \leq F(x, s, p, Y) \quad \text{if } r \leq s \text{ and } X - Y \text{ is positive semidefinite,}$$

so it includes $1st$ order Hamilton-Jacobi equations and $2nd$ order PDEs that are degenerate elliptic or parabolic in a very general sense [18, 5].

The Hamilton-Jacobi-Bellman (briefly, HJB) equations in the theory of optimal control of diffusion processes are of the form

(3) $$\sup_{\alpha \in A} \mathcal{L}^{\alpha} u = 0 ,$$

** Partially supported by M.U.R.S.T., projects “Problemi nonlineari nell’analisi e nelle applicazioni fisiche, chimiche e biologiche” and “Analisi e controllo di equazioni di evoluzione deterministiche e stocastiche”, and by the European Community, TMR Network “Viscosity solutions and their applications”.

where α is the control variable and, for each α, $\mathcal{L}^\alpha$ is a linear nondivergence form operator

$$\mathcal{L}^\alpha u := -a_{ij}^\alpha \frac{\partial^2 u}{\partial x_i \partial x_j} + b_i^\alpha \frac{\partial u}{\partial x_i} + c^\alpha u - f^\alpha , \tag{4}$$

where f and c are the running cost and the discount rate in the cost functional, b is the drift of the system, $a = \frac{1}{2}\sigma\sigma^T$ and σ is the variance of the noise affecting the system (see Section 3.2). These equations satisfy (2) if and only if

$$a_{ij}^\alpha(x)\xi_i\xi_j \geq 0 \text{ and } c^\alpha(x) \geq 0, \text{ for all } x \in \Omega,\ \alpha \in A,\ \xi \in \mathbb{R}^N , \tag{5}$$

and these conditions are automatically satisfied by operators coming from control theory. In the case of deterministic systems we have $a_{ij}^\alpha \equiv 0$ and the PDE is of $1st$ order. In the theory of two-person zero-sum deterministic and stochastic differential games the Isaacs' equation has the form

$$\sup_{\alpha \in A} \inf_{\beta \in B} \mathcal{L}^{\alpha,\beta} u = 0 , \tag{6}$$

where β is the control of the second player and $\mathcal{L}^{\alpha,\beta}$ are linear operators of the form (4) and satisfying assumptions such as (5).

For many different problems it was proved that the value function is the unique continuous viscosity solution satisfying appropriate boundary conditions, see the books [22, 8, 4, 5] and the references therein. This has a number of useful consequences, because we have PDE methods available to tackle several problems, such as the numerical calculation of the value function, the synthesis of approximate optimal feedback controls, asymptotic problems (vanishing noise, penalization, risk-sensitive control, ergodic problems, singular perturbations ...). However, the theory is considerably less general for problems with *discontinuous* value function, because it is restricted to deterministic systems with a single controller, where the HJB equation is of first order with convex Hamiltonian in the p variables. The pioneering papers on this issue are due to Barles and Perthame [10] and Barron and Jensen [11], who use different definitions of non-continuous viscosity solutions, see also [27, 28, 7, 39, 14], the surveys and comparisons of the different approaches in the books [8, 4, 5], and the references therein.

For cost functionals involving the exit time of the state from the set Ω, the value function is discontinuous if the noise vanishes near some part of the boundary and there is not enough controllability of the drift; other possible sources of discontinuities are the lack of smoothness of $\partial\Omega$, even for nondegenerate noise, and the discontinuity or incompatibility of the boundary data, even if the drift is controllable (see [8, 4, 5] for examples). For these functionals the value should be the solution of the Dirichlet problem

$$\begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Omega , \\ u = g & \text{on } \partial\Omega , \end{cases} \tag{7}$$

where $g(x)$ is the cost of exiting Ω at x and we assume $g \in C(\partial\Omega)$. For $2nd$ order equations, or $1st$ order equations with nonconvex Hamiltonian, there are no local definitions of weak solution and weak boundary conditions that ensure existence and uniqueness of a possibly discontinuous solution. However a global definition of generalized solution of (7) can be given by the following variant of the classical Perron-Wiener-Brelot method in potential theory. We define

$$\begin{aligned} \mathcal{S} &:= \{w \in BUSC(\overline{\Omega}) \text{ subsolution of (1)},\ w \leq g \text{ on } \partial\Omega\} \\ \mathcal{Z} &:= \{W \in BLSC(\overline{\Omega}) \text{ supersolution of (1)},\ W \geq g \text{ on } \partial\Omega\} , \end{aligned}$$

where $BUSC(\overline{\Omega})$ (respectively, $BLSC(\overline{\Omega})$) denote the sets of bounded upper (respectively, lower) semicontinuous functions on $\overline{\Omega}$, and we say that $u : \overline{\Omega} \to \mathbb{R}$ is a generalized solution of (7) if

$$u(x) = \sup_{w \in \mathcal{S}} w(x) = \inf_{W \in \mathcal{Z}} W(x)\,. \tag{8}$$

With respect to the classical Wiener's definition of generalized solution of the Dirichlet problem for the Lapalce equation in general nonsmooth domains [45] (see also [16, 26]), we only replace sub- and superharmonic functions with viscosity sub- and supersolutions. In the classical theory the inequality $\sup_{w \in \mathcal{S}} w \le \inf_{W \in \mathcal{Z}} W$ comes from the maximum principle, here it comes from the *Comparison Principle* for viscosity sub- and supersolutions; this important result holds under some additional assumptions that are very reasonable for the HJB equations of control theory, see Section 1.1; for this topic we refer to Jensen [29] and Crandall, Ishii and Lions [18]. The main difference with the classical theory is that the PWB solution for the Laplace equation is harmonic in Ω and can be discontinuous only at boundary points where $\partial\Omega$ is very irregular, whereas here u can be discontinuous also in the interior and even if the boundary is smooth: this is because the very degenerate ellipticity (2) neither implies regularizing effects, nor it guarantees that the boundary data are attained continuously. Note that if a continuous viscosity solution of (7) exists it coincides with u, and both the sup and the inf in (8) are attained.

Perron's method was extended to viscosity solutions by Ishii [27] (see Theorem 1), who used it to prove general existence results of continuous solutions. The PWB generalized solution of (7) of the form (8) was studied indipendently by the authors and Capuzzo-Dolcetta [4, 1] and by M. Ramaswamy and S. Ramaswamy [38] for some special cases of equations of the form (1), (2). In [4] this notion is called *envelope solution* and several properties are studied, in particular the equivalence with the generalized minimax solution of Subbotin [41, 42] and the connection with deterministic optimal control. The connection with pursuit-evasion games can be found in [41, 42] within the Krasovskii-Subbotin theory, and in our paper with Falcone [3] for the Fleming value; in [3] we also study the convergence of a numerical scheme.

The purposes of this paper are to extend the existence and basic properties of the PWB solution in [4, 1, 38] to more general operators, to prove some new continuity properties with respect to the data, in particular for the vanishing viscosity method and for approximations of the domain, and finally to show a connection with stochastic optimal control. For the sake of completeness we give all the proofs even if some of them follow the same argument as in the quoted references.

Let us now describe the contents of the paper in some detail. In Subsection 1.1 we recall some known definitions and results. In Subsection 1.2 we prove the existence theorem under an assumption on the boundary data g that is reminiscent of the compatibility conditions in the theory of $1st$ order Hamilton-Jacobi equations [34, 4]; this condition implies that the PWB solution is either the minimal supersolution or the maximal subsolution (i.e., either the inf or the sup in (8) is attained), and it is verified in time-optimal control problems. We recall that the classical Wiener Theorem asserts that for the Laplace equation any continuous boundary function g is *resolutive* (i.e., the PWB solution of the corresponding Dirichlet problem exists), and this was extended to some quasilinear nonuniformly elliptic equations, see the book of Heinonen, Kilpeläinen and Martio [25]. We do not know at the moment if this result can be extended to some class of fully nonlinear degenerate equations; however we prove in Subsection 2.1 that the set of resolutive boundary functions in our context is closed under uniform convergence as in the classical case (cfr. [26, 38]).

In Subsection 1.3 we show that the PWB solution is consistent with the notions of generalized solution by Subbotin [41, 42] and Ishii [27], and it satisfies the Dirichlet boundary condition in the weak viscosity sense [10, 28, 18, 8, 4]. Subsection 2.1 is devoted to the stability of the PWB solution with respect to the uniform convergence of the boundary data and the operator F. In Subsection 2.2 we consider merely local uniform perturbations of F, such as the vanishing viscosity, and prove a kind of stability provided the set Ω is simultaneously approximated from the interior.

In Subsection 2.3 we prove that for a nested sequence of open subsets Ω_n of Ω such that $\bigcup_n \Omega_n = \Omega$, if u_n is the PWB solution of the Dirichlet problem in Ω_n, the solution u of (7) satisfies

$$u(x) = \lim_n u_n(x)\,, \quad x \in \Omega\,. \tag{9}$$

This allows to approximate u with more regular solutions u_n when $\partial\Omega$ is not smooth and Ω_n are chosen with smooth boundary. This approximation procedure goes back to Wiener [44] again, and it is standard in elliptic theory for nonsmooth domains where (9) is often used to *define* a generalized solution of (7), see e.g. [30, 23, 12, 33]. In Subsection 2.3 we characterize the boundary points where the data are attained continuously in terms of the existence of suitable local barriers.

The last section is devoted to two applications of the previous theory to optimal control. The first (Subsection 3.1) is the classical minimum time problem for deterministic nonlinear systems with a closed target. In this case the lower semicontinuous envelope of the value function is the PWB solution of the homogeneous Dirichlet problem for the Bellman equation. The proof we give here is different from the one in [7, 4] and simpler. The second application (Subsection 3.2) is about the problem of maximizing the expected discounted time that a controlled degenerate diffusion process spends in Ω. Here we prove that the value function itself is the PWB solution of the appropriate problem. In both cases $g \equiv 0$ is a subsolution of the Dirichlet problem, which implies that the PWB solution is also the minimal supersolution.

It is worth to mention some recent papers using related methods. The thesis of Bettini [13] studies upper and lower semicontinuous solutions of the Cauchy problem for degenerate parabolic and $1st$ order equations with applications to finite horizon differential games. Our paper [2] extends some results of the present one to boundary value problems where the data are prescribed only on a suitable part of $\partial\Omega$. The first author, Goatin and Ishii [6] study the boundary value problem for (1) with Dirichlet conditions in the viscosity sense; they construct a PWB-type generalized solution that is also the limit of approximations of Ω from the outside, instead of the inside. This solution is in general different from ours and it is related to control problems involving the exit time from $\overline{\Omega}$, instead of Ω.

1. Generalized solutions of the Dirichlet problem

1.1. Preliminaries

Let F be a continuous function

$$F : \Omega \times \mathbb{R} \times \mathbb{R}^N \times S(N) \to \mathbb{R}\,,$$

where Ω is an open subset of $\mathbb{R}^N$, $S(N)$ is the set of symmetric $N \times N$ matrices equipped with its usual order, and assume that F satisfies (2). Consider the partial differential equation

$$F(x, u(x), Du(x), D^2u(x)) = 0 \ \text{ in } \Omega\,, \tag{10}$$

where $u : \Omega \to \mathbb{R}$, Du denotes the gradient of u and D^2u denotes the Hessian matrix of second derivatives of u. From now on subsolutions, supersolutions and solutions of this equation will be understood in the viscosity sense; we refer to [18, 5] for the definitions. For a general subset E of $\mathbb{R}^N$ we indicate with $USC(E)$, respectively $LSC(E)$, the set of all functions $E \to \mathbb{R}$ upper, respectively lower, semicontinuous, and with $BUSC(E)$, $BLSC(E)$ the subsets of functions that are also bounded.

DEFINITION 1. *We will say that equation (10) satisfies the* Comparison Principle *if for all subsolutions $w \in BUSC(\overline{\Omega})$ and supersolutions $W \in BLSC(\overline{\Omega})$ of (10) such that $w \leq W$ on $\partial\Omega$, the inequality $w \leq W$ holds in $\overline{\Omega}$.*

We refer to [29, 18] for the strategy of proof of some comparison principles, examples and references. Many results of this type for first order equations can be found in [8, 4].

The main examples we are interested in are the Isaacs equations:

$$\sup_{\alpha} \inf_{\beta} \mathcal{L}^{\alpha,\beta} u(x) = 0 \tag{11}$$

and

$$\inf_{\beta} \sup_{\alpha} \mathcal{L}^{\alpha,\beta} u(x) = 0\,, \tag{12}$$

where

$$\mathcal{L}^{\alpha,\beta} u(x) = -a_{ij}^{\alpha,\beta}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + b_i^{\alpha,\beta}(x) \frac{\partial u}{\partial x_i} + c^{\alpha,\beta}(x) u - f^{\alpha,\beta}(x)\,.$$

Here F is

$$F(x, r, p, X) = \sup_{\alpha} \inf_{\beta} \{-\text{trace}\,(a^{\alpha,\beta}(x) X) + b^{\alpha,\beta}(x) \cdot p + c^{\alpha,\beta}(x) r - f^{\alpha,\beta}(x)\}\,.$$

If, for all $x \in \overline{\Omega}$, $a^{\alpha,\beta}(x) = \frac{1}{2}\sigma^{\alpha,\beta}(x)(\sigma^{\alpha,\beta}(x))^T$, where $\sigma^{\alpha,\beta}(x)$ is a matrix of order $N \times M$, ^{T} denotes the transpose matrix, $\sigma^{\alpha,\beta}$, $b^{\alpha,\beta}$, $c^{\alpha,\beta}$, $f^{\alpha,\beta}$ are bounded and uniformly continuous in $\overline{\Omega}$, uniformly with respect to α, β, then F is continuous, and it is proper if in addition $c^{\alpha,\beta} \geq 0$ for all α, β.

Isaacs equations satisfy the Comparison Principle if Ω is bounded and there are positive constants K_1, K_2, and C such that

$$F(x, t, p, X) - F(x, s, q, Y) \leq \max\{K_1 \text{trace}\,(Y - X),\ K_1(t - s)\} + K_2|p - q|\,, \tag{13}$$

for all $Y \leq X$ and $t \leq s$,

$$\|\sigma^{\alpha,\beta}(x) - \sigma^{\alpha,\beta}(y)\| \leq C|x - y|,\ \text{for all } x, y \in \overline{\Omega} \text{ and all } \alpha, \beta \tag{14}$$

$$|b^{\alpha,\beta}(x) - b^{\alpha,\beta}(y)| \leq C|x - y|,\ \text{for all } x, y \in \overline{\Omega} \text{ and all } \alpha, \beta\,, \tag{15}$$

see Corollary 5.11 in [29]. In particular condition (13) is satisfied if and only if

$$\max\{\lambda^{\alpha,\beta}(x), c^{\alpha,\beta}(x)\} \geq K > 0 \text{ for all } x \in \overline{\Omega},\ \alpha \in A,\ \beta \in B\,,$$

where $\lambda^{\alpha,\beta}(x)$ is the smallest eigenvalue of $A^{\alpha,\beta}(x)$. Note that this class of equations contains as special cases the Hamilton-Jacobi-Bellman equations of optimal stochastic control (3) and linear degenerate elliptic equations with Lipschitz coefficients.

Given a function $u : \Omega \to [-\infty, +\infty]$, we indicate with u^* and u_*, respectively, the upper and the lower semicontinuous envelope of u, that is,

$$\begin{aligned} u^*(x) &:= \lim_{r\searrow 0} \sup\{u(y) : y \in \Omega,\ |y-x| \le r\}, \\ u_*(x) &:= \lim_{r\searrow 0} \inf\{u(y) : y \in \Omega,\ |y-x| \le r\}. \end{aligned}$$

PROPOSITION 1. *Let S (respectively Z) be a set of functions such that for all $w \in S$ (respectively $W \in Z$) w^* is a subsolution (respectively W_* is a supersolution) of (10). Define the function*

$$u(x) := \sup_{w\in S} w(x),\ x \in \Omega,\ (\textit{respectively } u(x) := \inf_{W\in Z} W(x)).$$

If u is locally bounded, then u^* is a subsolution (respectively u_* is a supersolution) of (10).

The proof of Proposition 1 is an easy variant of Lemma 4.2 in [18].

PROPOSITION 2. *Let $w_n \in BUSC(\Omega)$ be a sequence of subsolutions (respectively $W_n \in BLSC(\Omega)$ a sequence of supersolutions) of (10), such that $w_n(x) \searrow u(x)$ for all $x \in \Omega$ (respectively $W_n(x) \nearrow u(x)$) and u is a locally bounded function. Then u is a subsolution (respectively supersolution) of (10).*

For the proof see, for instance, [4]. We recall that, for a generale subset E of $\mathbb{R}^N$ and $\hat{x} \in E$, the second order superdifferential of u at $\hat{x}$ is the subset $J_E^{2,+}u(\hat{x})$ of $\mathbb{R}^N \times S(N)$ given by the pairs (p, X) such that

$$u(x) \le u(\hat{x}) + p \cdot (x - \hat{x}) + \frac{1}{2} X(x-\hat{x}) \cdot (x-\hat{x}) + o(|x-\hat{x}|^2)$$

for $E \ni x \to \hat{x}$. The opposite inequality defines the second order subdifferential of u at $\hat{x}$, $J_E^{2,-}u(\hat{x})$.

LEMMA 1. *Let u^* be a subsolution of (10). If u_* fails to be a supersolution at some point $\hat{x} \in \Omega$, i.e. there exist $(p, X) \in J_\Omega^{2,-}u_*(\hat{x})$ such that*

$$F(\hat{x}, u_*(\hat{x}), p, X) < 0,$$

then for all $k > 0$ small enough, there exists $U_k : \Omega \to \mathbb{R}$ such that U_k^* is subsolution of (10) and

$$\begin{cases} U_k(x) \ge u(x), \ \ \sup_\Omega (U_k - u) > 0, \\ U_k(x) = u(x) \textit{ for all } x \in \Omega \textit{ such that } |x-\hat{x}| \ge k. \end{cases}$$

The proof is an easy variant of Lemma 4.4 in [18]. The last result of this subsection is Ishii's extension of Perron's method to viscosity solutions [27].

THEOREM 1. *Assume there exists a subsolution u_1 and a supersolution u_2 of (10) such that $u_1 \le u_2$, and consider the functions*

$$\begin{aligned} U(x) &:= \sup\{w(x) : u_1 \le w \le u_2,\ w^* \textit{ subsolution of } (10)\}, \\ W(x) &:= \inf\{w(x) : u_1 \le w \le u_2,\ w_* \textit{ supersolution of } (10)\}. \end{aligned}$$

Then U^*, W^* are subsolutions of (10) and U_*, W_* are supersolutions of (10).

1.2. Existence of solutions by the PWB method

In this section we present a notion of weak solution for the boundary value problem

$$(16) \qquad \begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Omega\,, \\ u = g & \text{on } \partial\Omega\,, \end{cases}$$

where F satisfies the assumptions of Subsection 1.1 and $g : \partial\Omega \to \mathbb{R}$ is continuous. We recall that $\mathcal{S}$, $\mathcal{Z}$ are the sets of all subsolutions and all supersolutions of (16) defined in the Introduction.

DEFINITION 2. *The function defined by*

$$\underline{H}_g(x) := \sup_{w \in \mathcal{S}} w(x)\,,$$

is the lower envelope viscosity solution*, or* Perron-Wiener-Brelot lower solution*, of (16). We will refer to it as the* lower e-solution*. The function defined by*

$$\overline{H}_g(x) := \inf_{W \in \mathcal{Z}} W(x)\,,$$

is the upper envelope viscosity solution*, or* PWB upper solution*, of (16), briefly* upper e-solution*. If $\underline{H}_g = \overline{H}_g$, then*

$$H_g := \underline{H}_g = \overline{H}_g$$

is the envelope viscosity solution *or* PWB solution *of (16), briefly* e-solution*. In this case the data g are called* resolutive.

Observe that $\underline{H}_g \leq \overline{H}_g$ by the Comparison Principle, so the e-solution exists if the inequality $\geq$ holds as well. Next we prove the existence theorem for e-solutions, which is the main result of this section. We will need the following notion of global barrier, that is much weaker than the classical one.

DEFINITION 3. *We say that w is a* lower *(respectively,* upper*)* barrier *at a point $x \in \partial\Omega$ if $w \in \mathcal{S}$ (respectively, $w \in \mathcal{Z}$) and*

$$\lim_{y \to x} w(y) = g(x)\,.$$

THEOREM 2. *Assume that the Comparison Principle holds, and that $\mathcal{S}$, $\mathcal{Z}$ are nonempty.*

i) If there exists a lower barrier at all points $x \in \partial\Omega$, then $H_g = \min_{W \in \mathcal{Z}} W$ is the e-solution of (16).

ii) If there exists an upper barrier at all points $x \in \partial\Omega$, then $H_g = \max_{w \in \mathcal{S}} w$ is the e-solution of (16).

Proof. Let w be the lower barrier at $x \in \partial\Omega$, then by definition $w \leq \underline{H}_g$. Thus

$$(\underline{H}_g)_*(x) = \liminf_{y \to x} \underline{H}_g(y) \geq \liminf_{y \to x} w(y) = g(x)\,.$$

By Theorem 1 $(\underline{H}_g)_*$ is a supersolution of (10), so we can conclude that $(\underline{H}_g)_* \in \mathcal{Z}$. Then $(\underline{H}_g)_* \geq \overline{H}_g \geq \underline{H}_g$, so $\underline{H}_g = \overline{H}_g$ and $\underline{H}_g \in \mathcal{Z}$.

□

EXAMPLE 1. Consider the problem

$$\left\{\begin{array}{ll} -a_{ij}(x)u_{x_i x_j}(x) + b_i(x)u_{x_i}(x) + c(x)u(x) = 0 & \text{in } \Omega\,, \\ u(x) = g(x) & \text{on } \partial\Omega\,, \end{array}\right. \tag{17}$$

with the matrix $a_{ij}(x)$ such that $a_{11}(x) \geq \mu > 0$ for all $x \in \Omega$. In this case we can show that all continuous functions on $\partial\Omega$ are resolutive. The proof follows the classical one for the Laplace equation, the only hard point is checking the superposition principle for viscosity sub- and supersolutions. This can be done by the same methods and under the same assumptions as the Comparison Principle.

1.3. Consistency properties and examples

Next results give a characterization of the e-solution as pointwise limit of sequences of sub and supersolutions of (16). If the equation (10) is of first order, this property is essentially Subbotin's definition of (generalized) minimax solution of (16) [41, 42].

THEOREM 3. *Assume that the Comparison Principle holds, and that $\mathcal{S}$, $\mathcal{Z}$ are nonempty.*

i) If there exists $\underline{u} \in \mathcal{S}$ continuous at each point of $\partial\Omega$ and such that $\underline{u} = g$ on $\partial\Omega$, then there exists a sequence $w_n \in \mathcal{S}$ such that $w_n \nearrow H_g$.

ii) If there exists $\overline{u} \in \mathcal{Z}$ continuous at each point of $\partial\Omega$ and such that $\overline{u} = g$ on $\partial\Omega$, then there exists a sequence $W_n \in \mathcal{Z}$ such that $W_n \searrow H_g$.

Proof. We give the proof only for i), the same proof works for ii). By Theorem 2 $H_g = \min_{W\in\mathcal{Z}} W$. Given $\epsilon > 0$ the function

$$u_\epsilon(x) := \sup\{w(x) : w \in \mathcal{S},\ w(x) = \underline{u}(x) \text{ if } \operatorname{dist}(x, \partial\Omega) < \epsilon\}\,, \tag{18}$$

is bounded, and $u_\delta \leq u_\epsilon$ for $\epsilon < \delta$. We define

$$V(x) := \lim_{n\to\infty} (u_{1/n})_*(x)\,,$$

and note that, by definition, $H_g \geq u_\epsilon \geq (u_\epsilon)_*$, and then $H_g \geq V$. We claim that $(u_\epsilon)_*$ is supersolution of (10) in the set

$$\Omega_\epsilon := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \epsilon\}\,.$$

To prove this claim we assume by contradiction that $(u_\epsilon)_*$ fails to be a supersolution at $y \in \Omega_\epsilon$. Note that, by Proposition 1, $(u_\epsilon)^*$ is a subsolution of (10). Then by Lemma 1, for all $k > 0$ small enough, there exists U_k such that U_k^* is subsolution of (10) and

$$\sup_\Omega (U_k - u_\epsilon) > 0, \quad U_k(x) = u_\epsilon(x) \text{ if } |x - y| \geq k\,. \tag{19}$$

We fix $k \leq \operatorname{dist}(y, \partial\Omega) - \epsilon$, so that $U_k(x) = u_\epsilon(x) = \underline{u}(x)$ for all x such that $\operatorname{dist}(x, \partial\Omega) < \epsilon$. Then $U_k^*(x) = \underline{u}(x)$, so $U_k^* \in \mathcal{S}$ and by the definition of u_ϵ we obtain $U_k^* \leq u_\epsilon$. This gives a contradiction with (19) and proves the claim.

By Proposition 2 V is a supersolution of (10) in Ω. Moreover if $x \in \partial\Omega$, for all $\epsilon > 0$, $(u_\epsilon)_*(x) = g(x)$, because $u_\epsilon(x) = \underline{u}(x)$ if $\operatorname{dist}(x, \partial\Omega) < \epsilon$ by definition, $\underline{u}$ is continuous and $\underline{u} = g$ on $\partial\Omega$. Then $V \geq g$ on $\partial\Omega$, and so $V \in \mathcal{Z}$.

To complete the proof we define $w_n := (u_{1/n})^*$, and observe that this is a nondecreasing sequence in $\mathcal{S}$ whose pointwise limit is $\geq V$ by definition of V. On the other hand $w_n \leq H_g$ by definition of H_g, and we have shown that $H_g = V$, so $w_n \nearrow H_g$.

□

COROLLARY 1. *Assume the hypotheses of Theorem 3. Then H_g is the e-solution of (16 if and only if there exist two sequences of functions $w_n \in \mathcal{S}$, $W_n \in \mathcal{Z}$, such that $w_n = W_n = g$ on $\partial\Omega$ and for all $x \in \overline{\Omega}$*

$$w_n(x) \to H_g(x), \quad W_n(x) \to H_g(x) \text{ as } n \to \infty .$$

REMARK 1. It is easy to see from the proof of Theorem 3, that in case $i)$, the e-solution H_g satisfies

$$H_g(x) = \sup_{\epsilon} u_\epsilon(x) \quad x \in \overline{\Omega},$$

where

$$u_\epsilon(x) := \sup\{w(x) : w \in \mathcal{S},\ w(x) = \underline{u}(x) \text{ for } x \in \overline{\Omega} \setminus \overline{\Theta}_\epsilon\}, \tag{20}$$

and Θ_ϵ, $\epsilon \in]0,1]$, is any family of open sets such that $\overline{\Theta}_\epsilon \subseteq \Omega$, $\Theta_\epsilon \supseteq \Theta_\delta$ for $\epsilon < \delta$ and $\bigcup_\epsilon \Theta_\epsilon = \Omega$.

EXAMPLE 2. Consider the Isaacs equation (11) and assume the sufficient conditions for the Comparison Principle.

- If

  $$g \equiv 0 \text{ and } f^{\alpha,\beta}(x) \geq 0 \text{ for all } x \in \overline{\Omega},\ \alpha \in A,\ \beta \in B ,$$

  then $\underline{u} \equiv 0$ is subsolution of the PDE, so the assumption $i)$ of Theorem 3 is satisfied.
- If the domain Ω is bounded with smooth boundary and there exist $\overline{\alpha} \in A$ and $\mu > 0$ such that

  $$a_{ij}^{\overline{\alpha},\beta}(x)\xi_i\xi_j \geq \mu|\xi|^2 \text{ for all } \beta \in B,\ x \in \overline{\Omega},\ \xi \in \mathbb{R}^N ,$$

  then there exists a classical solution $\underline{u}$ of

  $$\begin{cases} \inf_{\beta\in B} \mathcal{L}^{\overline{\alpha},\beta}\underline{u} = 0 & \text{in } \Omega , \\ \underline{u} = g & \text{on } \partial\Omega , \end{cases}$$

  see e.g. Chapt. 17 of [24]. Then $\underline{u}$ is a supersolution of (11), so the hypothesis $ii)$ of Theorem 3 is satisfied.

Next we compare e-solutions with Ishii's definitions of non-continuous viscosity solution and of boundary conditions in viscosity sense. We recall that a function $u \in BUSC(\overline{\Omega})$ (respectively $u \in BLSC(\overline{\Omega})$) is a *viscosity subsolution* (respectively a *viscosity supersolution*) *of the boundary condition*

$$u = g \text{ or } F(x, u, Du, D^2u) = 0 \text{ on } \partial\Omega , \tag{21}$$

if for all $x \in \partial\Omega$ and $\phi \in C^2(\overline{\Omega})$ such that $u-\phi$ attains a local maximum (respectively minimum) at x, we have

$$(u-g)(x) \leq 0 \text{ (resp. } \geq 0) \text{ or } F(x, u(x), D\phi(x), D^2\phi(x)) \leq 0 \text{ (resp. } \geq 0)\,.$$

An equivalent definition can be given by means of the semijets $J^{2,+}_{\overline{\Omega}}u(x)$, $J^{2,-}_{\overline{\Omega}}u(x)$ instead of the test functions, see [18].

PROPOSITION 3. *If $\underline{H}_g : \overline{\Omega} \to \mathbb{R}$ is the lower e-solution (respectively, $\overline{H}_g$ is the upper e-solution) of (16), then $\underline{H}_g^*$ is a subsolution (respectively, $\overline{H}_{g*}$ is a supersolution) of (10) and of the boundary condition (21).*

Proof. If $\underline{H}_g$ is the lower e-solution, then by Proposition 1, $\underline{H}_g^*$ is a subsolution of (10). It remains to check the boundary condition.

Fix an $y \in \partial\Omega$ such that $\underline{H}_g^*(y) > g(y)$, and $\phi \in C^2(\overline{\Omega})$ such that $\underline{H}_g^* - \phi$ attains a local maximum at y. We can assume, without loss of generality, that

$$\underline{H}_g^*(y) = \phi(y), \quad (\underline{H}_g^* - \phi)(x) \leq -|x-y|^3 \text{ for all } x \in \overline{\Omega} \cap B(y,r)\,.$$

By definition of $\underline{H}_g^*$, there exists a sequence of points $x_n \to y$ such that

$$(\underline{H}_g - \phi)(x_n) \geq -\frac{1}{n} \text{ for all } n\,.$$

Moreover, since $\underline{H}_g$ is the lower e-solution, there exists a sequence of functions $w_n \in S$ such that

$$\underline{H}_g(x_n) - \frac{1}{n} < w_n(x_n) \text{ for all } n\,.$$

Since the function $w_n - \phi$ is upper semicontinuous, it attains a maximum at $y_n \in \overline{\Omega} \cap \overline{B}(y,r)$, such that, for n big enough,

$$-\frac{2}{n} < (w_n - \phi)(y_n) \leq -|y_n - y|^3.$$

So as $n \to \infty$

$$y_n \to y, \quad w_n(y_n) \to \phi(y) = \underline{H}_g^*(y) > g(y)\,.$$

Note that $y_n \notin \partial\Omega$, because $y_n \in \partial\Omega$ would imply $w_n(y_n) \leq g(y_n)$, which gives a contradiction to the continuity of g at y. Therefore, since w_n is a subsolution of (10), we have

$$F(y_n, w_n(y_n), D\phi(y_n), D^2\phi(y_n)) \leq 0\,,$$

and letting $n \to \infty$ we get

$$F(y, \underline{H}_g^*(y), D\phi(y), D^2\phi(y)) \leq 0\,,$$

by the continuity of F.

□

REMARK 2. By Proposition 3, if the e-solution H_g of (16) exists, it is a non-continuous viscosity solution of (10) (21) in the sense of Ishii [27]. These solutions, however, are not unique in general. An e-solution satisfies also the Dirichlet problem in the sense that it is a non-continuous solution of (10) in Ishii's sense and $H_g(x) = g(x)$ for all $x \in \partial\Omega$, but neither this property characterizes it. We refer to [4] for explicit examples and more details.

REMARK 3. Note that, by Proposition 3, if the e-solution H_g is continuous at all points of $\partial\Omega_1$ with $\Omega_1 \subset \Omega$, we can apply the Comparison Principle to the upper and lower semicontinuous envelopes of H_g and obtain that it is continuous in Ω_1. If the equation is uniformly elliptic in Ω_1 we can also apply in Ω_1 the local regularity theory for continuous viscosity solutions developed by Caffarelli [17] and Trudinger [43].

2. Properties of the generalized solutions

2.1. Continuous dependence under uniform convergence of the data

We begin this section by proving a result about continuous dependence of the e-solution on the boundary data of the Dirichlet Problem. It states that the set of resolutive data is closed with respect to uniform convergence. Throughout the paper we denote with $\rightrightarrows$ the uniform convergence.

THEOREM 4. *Let $F : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^N \times S(N) \to \mathbb{R}$ be continuous and proper, and let $g_n : \partial\Omega \to \mathbb{R}$ be continuous. Assume that $\{g_n\}_n$ is a sequence of resolutive data such that $g_n \rightrightarrows g$ on $\partial\Omega$. Then g is resolutive and $H_{g_n} \rightrightarrows H_g$ on Ω.*

The proof of this theorem is very similar to the classical one for the Laplace equation [26]. We need the following result:

LEMMA 2. *For all $c > 0$, $\underline{H}_{(g+c)} \leq \underline{H}_g + c$ and $\overline{H}_{(g+c)} \leq \overline{H}_g + c$.*

Proof. Let

$$\mathcal{S}_c := \{w \in BUSC(\overline{\Omega}) : w \text{ is subsolution of (10)},\ w \leq g + c \text{ on } \partial\Omega\}\,.$$

Fix $u \in \mathcal{S}_c$, and consider the function $v(x) = u(x) - c$. Since F is proper it is easy to see that $v \in \mathcal{S}$. Then

$$\underline{H}_{(g+c)} := \sup_{u\in\mathcal{S}_c} u \leq \sup_{v\in\mathcal{S}} v + c := \underline{H}_g + c\,.$$

□

of Theorem 4. Fix $\epsilon > 0$, the uniform convergence implies $\exists m : \forall n \geq m$: $g_n - \epsilon \leq g \leq g_n + \epsilon$. Since g_n is resolutive by Lemma 2, we get

$$H_{g_n} - \epsilon \leq \underline{H}_{(g_n-\epsilon)} \leq \underline{H}_g \leq \underline{H}_{(g_n+\epsilon)} \leq H_{g_n} + \epsilon\,.$$

Therefore $H_{g_n} \rightrightarrows \underline{H}_g$. The proof that $H_{g_n} \rightrightarrows \overline{H}_g$, is similar.

□

Next result proves the continuous dependence of e-solutions with respect to the data of the Dirichlet Problem, assuming that the equations F_n are strictly decreasing in r, uniformly in n.

THEOREM 5. *Let $F_n : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^N \times S(N) \to \mathbb{R}$ is continuous and proper, $g : \partial\Omega \to \mathbb{R}$ is continuous. Suppose that $\forall n,\ \forall \delta > 0\ \exists \epsilon$ such that*

$$F_n(x, r - \delta, p, X) + \epsilon \le F_n(x, r, p, X)$$

for all $(x, r, p, X) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^N \times S(N)$, and $F_n \rightrightarrows F$ on $\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^N \times S(N)$. Suppose g is resolutive for the problems

$$\begin{cases} F_n(x, u, Du, D^2u) = 0 & \text{in } \Omega\,, \\ u = g & \text{on } \partial\Omega\,. \end{cases} \tag{22}$$

Suppose $g_n : \partial\Omega \to \mathbb{R}$ is continuous, $g_n \rightrightarrows g$ on $\partial\Omega$ and g_n is resolutive for the problem

$$\begin{cases} F_n(x, u, Du, D^2u) = 0 & \text{in } \Omega\,, \\ u = g_n & \text{on } \partial\Omega\,. \end{cases} \tag{23}$$

Then g is resolutive for (16) and $H^n_{g_n} \rightrightarrows H_g$, where $H^n_{g_n}$ is the e-solution of (23).

Proof. Step 1. For fixed $\delta > 0$ we want to show that there exists m such that for all $n \ge m$: $|\underline{H}^n_g - \underline{H}_g| \le \delta$, where H^n_g is the e-solution of (22).

We claim that there exists m such that $\underline{H}^n_g - \delta \le \underline{H}_g$ and $\overline{H}_g \le \overline{H}^n_g + \delta$ for all $n \ge m$. Then

$$\underline{H}^n_g - \delta \le \underline{H}_g \le \overline{H}_g \le \overline{H}^n_g + \delta = \underline{H}^n_g + \delta\,.$$

This proves in particular $H^n_g \rightrightarrows \underline{H}_g$ and $H^n_g \rightrightarrows \overline{H}_g$, and then $\overline{H}_g = \underline{H}_g$, so g is resolutive for (16).

It remains to prove the claim. Let

$$\mathcal{S}^n_g := \{v \text{ subsolution of } F_n = 0 \text{ in } \Omega,\ v \le g \text{ on } \partial\Omega\}\,.$$

Fix $v \in \mathcal{S}^n_g$, and consider the function $u = v - \delta$. By hypothesis there exists an ϵ such that $F_n(x, u(x), p, X) + \epsilon \le F_n(x, v(x), p, X)$, for all $(p, X) \in J^{2,+}_{\Omega} u(x)$. Then using uniform convergence of F_n at F we get

$$F(x, u(x), p, X) \le F_n(x, u(x), p, X) + \epsilon \le F_n(x, v(x), p, X) \le 0\,,$$

so v is a subsolution of the equation $F_n = 0$ because $J^{2,+}_{\Omega} v(x) = J^{2,+}_{\Omega} u(x)$.

We have shown that for all $v \in \mathcal{S}^n_g$ there exists $u \in \mathcal{S}$ such that $v = u + \delta$, and this proves the claim.

Step 2. Using the argument of proof of Theorem 4 with the problem

$$\begin{cases} F_m(x, u, Du, D^2u) = 0 & \text{in } \Omega\,, \\ u = g_n & \text{on } \partial\Omega\,, \end{cases} \tag{24}$$

we see that fixing $\delta > 0$, there exists p such that for all $n \ge p$: $|\underline{H}^m_{g_n} - \underline{H}^m_g| \le \delta$ for all m.

Step 3. Using again arguments of proof of Theorem 4, we see that fixing $\delta > 0$ there exists q such that for all $n, m \ge q$: $|\underline{H}^m_{g_n} - \underline{H}^m_{g_m}| \le \delta$.

Step 4. Now take $\delta > 0$, then there exists p such that for all $n, m \geq p$:

$$|\underline{H}^m_{g_m} - \underline{H}_g| \leq |\underline{H}^m_{g_m} - \underline{H}^m_{g_n}| + |\underline{H}^m_{g_n} - \underline{H}^m_g| + |\underline{H}^m_g - \underline{H}_g| \leq 3\delta\,.$$

Similarly $|\overline{H}^m_{g_m} - \overline{H}_g| \leq 3\delta$. But $\underline{H}^m_{g_m} = \overline{H}^m_{g_m}$, and this complete the proof.

□

2.2. Continuous dependence under local uniform convergence of the operator

In this subsection we study the continuous dependence of e-solutions with respect to perturbations of the operator, depending on a parameter h, that are not uniform over all $\Omega \times \mathbb{R} \times \mathbb{R}^N \times S(N)$ as they were in Theorem 5, but only on compact subsets of $\Omega \times \mathbb{R} \times \mathbb{R}^N \times S(N)$. A typical example we have in mind is the vanishing viscosity approximation, but similar arguments work for discrete approximation schemes, see [3]. We are able to pass to the limit under merely local perturbations of the operator by approximating Ω with a nested family of open sets Θ_ϵ, solving the problem in each Θ_ϵ, and then letting $\epsilon,\ h$ go to 0 "with h linked to ϵ" in the following sense.

DEFINITION 4. *Let v^ϵ_h, $u : Y \to \mathbb{R}$, for $\epsilon > 0$, $h > 0$, $Y \subseteq \mathbb{R}^N$. We say that v^ϵ_h* converges to u as $(\epsilon, h) \searrow (0,0)$ with h linked to ϵ *at the point x, and write*

$$\lim_{\substack{(\epsilon,h)\searrow(0,0)\\ h\leq h(\epsilon)}} v^\epsilon_h(x) = u(x) \tag{25}$$

if for all $\gamma > 0$, there exist a function $\tilde{h} :]0,+\infty[\to]0,+\infty[$ and $\overline{\epsilon} > 0$ such that

$$|v^\epsilon_h(y) - u(x)| \leq \gamma,\ \textit{for all } y \in Y : |x - y| \leq \tilde{h}(\epsilon)$$

for all $\epsilon \leq \overline{\epsilon}$, $h \leq \tilde{h}(\epsilon)$.

To justify this definition we note that:

$i)$ it implies that for any x and $\epsilon_n \searrow 0$ there is a sequence $h_n \searrow 0$ such that $v^{\epsilon_n}_{h_n}(x_n) \to u(x)$ for any sequence x_n such that $|x - x_n| \leq h_n$, e.g. $x_n = x$ for all n, and the same holds for any sequence $h'_n \geq h_n$;

$ii)$ if $\lim_{h\searrow 0} v^\epsilon_h(x)$ exists for all small ϵ and its limit as $\epsilon \searrow 0$ exists, then it coincides with the limit of Definition 4, that is,

$$\lim_{\substack{(\epsilon,h)\searrow(0,0)\\ h\leq h(\epsilon)}} v^\epsilon_h(x) = \lim_{\epsilon\searrow 0}\lim_{h\searrow 0} v^\epsilon_h(x)\,.$$

REMARK 4. If the convergence of Definition 4 occurs on a compact set K where the limit u is continuous, then (25) can be replaced, for all $x \in K$ and redefining $\tilde{h}$ if necessary, with

$$|v^\epsilon_h(y) - u(y)| \leq 2\gamma,\ \text{ for all } y \in K : |x - y| \leq \tilde{h}(\epsilon)\,,$$

and by a standard compactness argument we obtain the uniform convergence in the following sense:

DEFINITION 5. *Let K be a subset of $\mathbb{R}^N$ and v^ϵ_h, $u : K \to \mathbb{R}$ for all $\epsilon, h > 0$. We say that v^ϵ_h* converge uniformly *on K to u as $(\epsilon, h) \searrow (0,0)$ with h* linked to ϵ *if for any $\gamma > 0$ there are $\overline{\epsilon} > 0$ and $\tilde{h} :]0,+\infty[\to]0,+\infty[$ such that*

$$\sup_K |v^\epsilon_h - u| \leq \gamma$$

for all $\epsilon \le \overline{\epsilon}$, $h \le \tilde{h}(\epsilon)$.

The main result of this subsection is the following. Recall that a family of functions $v_h^\epsilon : \Omega \to \mathbb{R}$ is locally uniformly bounded if for each compact set $K \subseteq \Omega$ there exists a constant C_K such that $\sup_K |v_h^\epsilon| \le C_K$ for all $h, \epsilon > 0$. In the proof we use the weak limits in the viscosity sense and the stability of viscosity solutions and of the Dirichlet boundary condition in viscosity sense (21) with respect to such limits.

THEOREM 6. *Assume the Comparison Principle holds,* $\mathcal{Z} \neq \emptyset$ *and let* $\underline{u}$ *be a continuous subsolution of (16) such that* $\underline{u} = g$ *on* $\partial\Omega$. *For any* $\epsilon \in]0,1]$, *let* Θ_ϵ *be an open set such that* $\overline{\Theta_\epsilon} \subseteq \Omega$, *and for* $h \in]0,1]$ *let* v_h^ϵ *be a non-continuous viscosity solution of the problem*

$$
(26) \qquad \begin{cases} F_h(x, u, Du, D^2u) = 0 & \text{in } \Theta_\epsilon\,, \\ u(x) = \underline{u}(x) \text{ or } F_h(x, u, Du, D^2u) = 0 & \text{on } \partial\Theta_\epsilon\,, \end{cases}
$$

where $F_h : \overline{\Theta_\epsilon} \times \mathbb{R} \times \mathbb{R}^N \times S(N) \to \mathbb{R}$ *is continuous and proper. Suppose* $\{v_h^\epsilon\}$ *is locally uniformly bounded,* $v_h^\epsilon \ge \underline{u}$ *in* $\overline{\Omega}$, *and extend* $v_h^\epsilon := \underline{u}$ *in* $\overline{\Omega} \setminus \Theta_\epsilon$. *Finally assume that* F_h *converges uniformly to* F *on any compact subset of* $\Omega \times \mathbb{R} \times \mathbb{R}^N \times S(N)$ *as* $h \searrow 0$, *and* $\Theta_\epsilon \supseteq \Theta_\delta$ *if* $\epsilon < \delta$, $\bigcup_{0<\epsilon\le 1} \Theta_\epsilon = \Omega$.

Then v_h^ϵ *converges to the e-solution* H_g *of (16) with* h *linked to* ϵ, *that is, (25) holds for all* $x \in \overline{\Omega}$; *moreover the convergence is uniform (as in Def. 5) on any compact subset of* $\overline{\Omega}$ *where* H_g *is continuous.*

Proof. Note that the hypotheses of Theorem 3 are satisfied, so the e-solution H_g exists. Consider the weak limits

$$
\begin{aligned}
\underline{v}_\epsilon(x) &:= \liminf_{h\searrow 0}{}_{*}\, v_h^\epsilon(x) := \sup_{\delta>0} \inf\{v_h^\epsilon(y) : |x-y| < \delta,\ 0 < h < \delta\}\,, \\
\overline{v}_\epsilon(x) &:= \limsup_{h\searrow 0}{}^{*}\, v_h^\epsilon(x) := \inf_{\delta>0} \sup\{v_h^\epsilon(y) : |x-y| < \delta,\ 0 < h < \delta\}\,.
\end{aligned}
$$

By a standard result in the theory of viscosity solutions, see [10, 18, 8, 4], $\underline{v}_\epsilon$ and $\overline{v}_\epsilon$ are respectively supersolution and subsolution of

$$
(27) \qquad \begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Theta_\epsilon\,, \\ u(x) = \underline{u}(x) \text{ or } F(x, u, Du, D^2u) = 0 & \text{on } \partial\Theta_\epsilon\,. \end{cases}
$$

We claim that $\overline{v}_\epsilon$ is also a subsolution of (16). Indeed $v_h^\epsilon \equiv \underline{u}$ in $\Omega \setminus \Theta_\epsilon$, so $\overline{v}_\epsilon \equiv \underline{u}$ in the interior of $\Omega \setminus \Theta_\epsilon$ and then in this set it is a subsolution. In Θ_ϵ we have already seen that $\overline{v}_\epsilon = (\overline{v}_\epsilon)^*$ is a subsolution. It remains to check what happens on $\partial\Theta_\epsilon$. Given $\hat{x} \in \partial\Theta_\epsilon$, we must prove that for all $(p, X) \in J_\Omega^{2,+}\overline{v}_\epsilon(\hat{x})$ we have

$$
(28) \qquad F_h(\hat{x}, \overline{v}_\epsilon(\hat{x}), p, X) \le 0\,.
$$

1st Case: $\overline{v}_\epsilon(\hat{x}) > \underline{u}(\hat{x})$. Since $\overline{v}_\epsilon$ satisfies the boundary condition on $\partial\Theta_\epsilon$ of problem (27), then for all $(p, X) \in J_{\overline{\Theta_\epsilon}}^{2,+}\overline{v}_\epsilon(\hat{x})$ (28) holds. Then the same inequality holds for all $(p, X) \in J_\Omega^{2,+}\overline{v}_\epsilon(\hat{x})$ as well, because $J_\Omega^{2,+}\overline{v}_\epsilon(\hat{x}) \subseteq J_{\overline{\Theta_\epsilon}}^{2,+}\overline{v}_\epsilon(\hat{x})$.

2nd Case: $\overline{v}_\epsilon(\hat{x}) = \underline{u}(\hat{x})$. Fix $(p, X) \in J_\Omega^{2,+}\overline{v}_\epsilon(\hat{x})$, by definition

$$
\overline{v}_\epsilon(x) \le \overline{v}_\epsilon(\hat{x}) + p \cdot (x - \hat{x}) + \frac{1}{2} X(x - \hat{x}) \cdot (x - \hat{x}) + o(|x - \hat{x}|^2)
$$

for all $x \to \hat{x}$. Since $\overline{v}_\epsilon \geq \underline{u}$ and $\overline{v}_\epsilon(\hat{x}) = \underline{u}(\hat{x})$, we get

$$\underline{u}(x) \leq \underline{u}(\hat{x}) + p \cdot (x - \hat{x}) + \frac{1}{2} X(x - \hat{x}) \cdot (x - \hat{x}) + o(|x - \hat{x}|^2)\,,$$

that is $(p, X) \in J^{2,+}_{\overline{\Omega}} \underline{u}(\hat{x})$. Now, since $\underline{u}$ is a subsolution, we conclude

$$F(\hat{x}, \overline{v}_\epsilon(\hat{x}), p, X) = F(\hat{x}, \underline{u}(\hat{x}), p, X) \leq 0\,.$$

We now claim that

$$u_\epsilon \leq \underline{v}_\epsilon \leq \overline{v}_\epsilon \leq H_g \text{ in } \overline{\Omega}\,, \tag{29}$$

where u_ϵ is defined by (20). Indeed, since $\underline{v}_\epsilon$ is a supersolution in Θ_ϵ and $\underline{v}_\epsilon \geq \underline{u}$, by the Comparison Principle $\underline{v}_\epsilon \geq w$ in Θ_ϵ for any $w \in \mathcal{S}$ such that $w = \underline{u}$ on $\partial\Theta_\epsilon$. Moreover $\underline{v}_\epsilon \equiv \underline{u}$ on $\Omega \backslash \Theta_\epsilon$, so we get $\underline{v}_\epsilon \geq u_\epsilon$ in $\overline{\Omega}$. To prove the last inequality we note that H_g is a supersolution of (16) by Theorem 3, which implies $\overline{v}_\epsilon \leq H_g$ by Comparison Principle.

Now fix $x \in \overline{\Omega}$, $\epsilon > 0$, $\gamma > 0$ and note that, by definition of lower weak limit, there exists $\overline{h} = \overline{h}(x, \epsilon, \gamma) > 0$ such that

$$\underline{v}_\epsilon(x) - \gamma \leq v^\epsilon_h(y)$$

for all $h \leq \overline{h}$ and $y \in \overline{\Omega} \cap B(x, \overline{h})$. Similarly there exists $\overline{k} = \overline{k}(x, \epsilon, \gamma) > 0$ such that

$$v^\epsilon_h(y) \leq \overline{v}_\epsilon(x) + \gamma$$

for all $h \leq \overline{k}$ and $y \in \overline{\Omega} \cap B(x, \overline{k})$. From Remark 1, we know that $H_g = \sup_\epsilon u_\epsilon$, so there exists $\overline{\epsilon}$ such that

$$H_g(x) - \gamma \leq u_\epsilon(x), \text{ for all } \epsilon \leq \overline{\epsilon}\,.$$

Then, using (29), we get

$$H_g(x) - 2\gamma \leq v^\epsilon_h(y) \leq H_g(x) + \gamma$$

for all $\epsilon \leq \overline{\epsilon}$, $h \leq \tilde{h} := \min\{\overline{h}, \overline{k}\}$ and $y \in \overline{\Omega} \cap B(x, \tilde{h})$, and this completes the proof.

□

REMARK 5. Theorem 6 applies in particular if v^ϵ_h are the solutions of the following vanishing viscosity approximation of (10)

$$\begin{cases} -h\Delta v + F(x, v, Dv, D^2 v) = 0 & \text{in } \Theta_\epsilon\,, \\ v = \underline{u} & \text{on } \partial\Theta_\epsilon\,. \end{cases} \tag{30}$$

Since F is degenerate elliptic, the PDE in (30) is uniformly elliptic for all $h > 0$. Therefore we can choose a family of nested Θ_ϵ with smooth boundary and obtain that the approximating v^ϵ_h are much smoother than the e-solution of (16). Indeed (30) has a classical solution if, for instance, either F is smooth and $F(x, \cdot, \cdot, \cdot)$ is convex, or the PDE (10 is a Hamilton-Jacobi-Bellman equation (3 where the linear operators $\mathcal{L}^\alpha$ have smooth coefficients, see [21, 24, 31]. In the nonconvex case, under some structural assumptions, the continuity of the solution of (30) follows from a barrier argument (see, e.g., [5]), and then it is twice differentiable almost everywhere by a result in [43], see also [17].

2.3. Continuous dependence under increasing approximation of the domain

In this subsection we prove the continuity of the e-solution of (16) with respect to approximations of the domain Ω from the interior. Note that, if $v_h^\epsilon = v^\epsilon$ for all h in Theorem 6, then $v^\epsilon(x) \to H_g(x)$ for all $x \in \Omega$ as $\epsilon \searrow 0$. This is the case, for instance, if v^ϵ is the unique e-solution of

$$\begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Theta_\epsilon\,, \\ u = \underline{u} & \text{on } \partial\Theta_\epsilon\,, \end{cases}$$

by Proposition 3. The main result of this subsection extends this remark to more general approximations of Ω from the interior, where the condition $\overline{\Theta}_\epsilon \subseteq \Omega$ is dropped. We need first a monotonicity property of e-solutions with respect to the increasing of the domain.

LEMMA 3. *Assume the Comparison Principle holds and let $\Omega_1 \subseteq \Omega_2 \subseteq \mathbb{R}^N$, H_g^1, respectively H_g^2, be the e-solution in Ω_1, respectively Ω_2, of the problem*

$$\begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Omega_i\,, \\ u = g & \text{on } \partial\Omega_i\,, \end{cases} \tag{31}$$

with $g : \overline{\Omega}_2 \to \mathbb{R}$ continuous and subsolution of (31) with $i = 2$. If we define

$$\tilde{H}_g^1(x) = \begin{cases} H_g^1(x) & \text{if } x \in \overline{\Omega}_1 \\ g(x) & \text{if } x \in \Omega_2 \setminus \overline{\Omega}_1\,, \end{cases}$$

then $H_g^2 \geq \tilde{H}_g^1$ in Ω_2.

Proof. By definition of e-solution $H_g^2 \geq g$ in Ω_2, so H_g^2 is also supersolution of (31) in Ω_1. Therefore $H_g^2 \geq H_g^1$ in Ω_1 because H_g^1 is the smallest supersolution in Ω_1, and this completes the proof.

□

THEOREM 7. *Assume that the hypotheses of Theorem 3 i) hold with $\underline{u}$ continuous and Ω bounded. Let $\{\Omega_n\}$ be a sequence of open subsets of Ω, such that $\Omega_n \subseteq \Omega_{n+1}$ and $\bigcup_n \Omega_n = \Omega$. Let u_n be the e-solution of the problem*

$$\begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Omega_n\,, \\ u = \underline{u} & \text{on } \partial\Omega_n\,. \end{cases} \tag{32}$$

If we extend $u_n := \underline{u}$ in $\Omega \setminus \Omega_n$, then $u_n(x) \nearrow H_g(x)$ for all $x \in \Omega$, where H_g is the e-solution of (16).

Proof. Note that for all n there exists an $\epsilon_n > 0$ such that $\Omega_{\epsilon_n} = \{x \in \Omega : \text{dist}\,(x, \partial\Omega) \geq \epsilon_n\} \subseteq \Omega_n$. Consider the e-solution u_{ϵ_n} of problem

$$\begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Omega_{\epsilon_n}\,, \\ u = \underline{u} & \text{on } \partial\Omega_{\epsilon_n}\,. \end{cases}$$

If we set $u_{\epsilon_n} \equiv \underline{u}$ in $\Omega \setminus \Omega_{\epsilon_n}$, by Theorem 6 we get $u_{\epsilon_n} \to H_g$ in Ω, as remarked at the beginning of this subsection. Finally by Lemma 3 we have $H_g \geq u_n \geq u_{\epsilon_n}$ in Ω, and so $u_n \to H_g$ in Ω.

□

REMARK 6. If $\partial\Omega$ is not smooth and F is uniformly elliptic Theorem 7 can be used as an approximation result by choosing Ω_n with smooth boundary. In fact, under some structural assumptions, the solution u_n of (32) turns out to be continuous by a barrier argument (see, e.g., [5]), and then it is twice differentiable almost everywhere by a result in [43], see also [17]. If, in addition, F is smooth and $F(x, \cdot, \cdot, \cdot)$ is convex, or the PDE (10) is a HJB equation (3) where the linear operators $\mathcal{L}^\alpha$ have smooth coefficients, then u_n is of class C^2, see [21, 24, 31, 17] and the references therein. The Lipschitz continuity of u_n holds also if F is not uniformly elliptic but it is coercive in the p variables.

2.4. Continuity at the boundary

In this section we study the behavior of the e-solution at boundary points and characterize the points where the boundary data are attained continuously by means of barriers.

PROPOSITION 4. *Assume that hypothesis $i)$ (respectively $ii)$) of Theorem 2 holds. Then the e-solution H_g of (16) takes up the boundary data g continuously at $x_0 \in \partial\Omega$, i.e.* $\lim_{x\to x_0} H_g(x) = g(x_0)$, *if and only if there is an upper (respectively lower) barrier at x_0 (see Definition 3).*

Proof. The necessity is obvious because Theorem 2 $i)$ implies that $H_g \in \mathcal{Z}$, so H_g is an upper barrier at x if it attains continuously the data at x.

Now we assume W is an upper barrier at x. Then $W \geq H_g$, because $W \in \mathcal{Z}$ and H_g is the minimal element of $\mathcal{Z}$. Therefore

$$g(x) \leq H_g(x) \leq \liminf_{y\to x} H_g(y) \leq \limsup_{y\to x} H_g(y) \leq \lim_{y\to x} W(y) = g(x)\,,$$

so $\lim_{y\to x} H_g(y) = g(x) = H_g(x)$.

□

In the classical theory of linear elliptic equations, local barriers suffice to characterize boundary continuity of weak solutions. Similar results can be proved in our fully nonlinear context. Here we limit ourselves to a simple result on the Dirichlet problem with homogeneous boundary data for the Isaacs equation

$$\begin{cases} \sup_\alpha \inf_\beta \{-a_{ij}^{\alpha,\beta} u_{x_i x_j} + b_i^{\alpha,\beta} u_{x_i} + c^{\alpha,\beta} u - f^{\alpha,\beta}\} = 0 & \text{in } \Omega\,, \\ u = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{33}$$

DEFINITION 6. *We say that $W \in BLSC(\overline{B(x_0, r) \cap \Omega})$ with $r > 0$ is an* upper local barrier *for problem (33) at $x_0 \in \partial\Omega$ if*

$i)$ *$W \geq 0$ is a supersolution of the PDE in (33) in $B(x_0, r) \cap \Omega$,*

$ii)$ *$W(x_0) = 0$, $W(x) \geq \mu > 0$ for all $|x - x_0| = r$,*

$iii)$ *W is continuous at x_0.*

PROPOSITION 5. *Assume the Comparison Principle holds for (33), $f^{\alpha,\beta} \geq 0$ for all α, β, and let H_g be the e-solution of problem (33). Then H_g takes up the boundary data continuously at $x_0 \in \partial\Omega$ if and only if there exists an upper local barrier W at x_0.*

Proof. We recall that H_g exists because the function $u \equiv 0$ is a lower barrier for all points $x \in \partial\Omega$ by the fact that $f^{\alpha,\beta} \geq 0$, and so we can apply Theorem 2. Consider a supersolution w of (33). We claim that the function V defined by

$$V(x) = \begin{cases} \rho W(x) \wedge w(x) & \text{if } x \in \overline{B(x_0, r) \cap \Omega}\,, \\ w(x) & \text{if } x \in \Omega \setminus \overline{B}(x_0, r)\,, \end{cases}$$

is an upper barrier at x_0 for $\rho > 0$ large enough. It is easy to check that ρW is a supersolution of (33) in $B(x_0, r) \cap \Omega$, so V is a supersolution in $B(x_0, r) \cap \Omega$ (by Proposition 1) and in $\Omega \setminus \overline{B}(x_0, r)$. Since w is bounded, by property $ii)$ in Definition 6, we can fix ρ and $\epsilon > 0$ such that $V(x) = w(x)$ for all $x \in \Omega$ satisfying $r - \epsilon < |x - x_0| \leq r$. Then V is supersolution even on $\partial B(x_0, r) \cap \Omega$. Moreover it is obvious that $V \geq 0$ on $\partial\Omega$ and $V(x_0) = 0$. We have proved that V is supersolution of (33) in Ω.

It remains to prove that $\lim_{x\to x_0} V(x) = 0$. Since the constant 0 is a subsolution of (33) and w is a supersolution, we have $w \geq 0$. Then we reach the conclusion by $ii)$ and $iii)$ of Definition 6.

□

EXAMPLE 3. We construct an upper local barrier for (33) under the assumptions of Proposition 5 and supposing in addition

$$\partial\Omega \text{ is } C^2 \text{ in a neighbourhood of } x_0 \in \partial\Omega\,,$$

there exists an α^* such that for all β either

$$a_{ij}^{\alpha^*,\beta}(x_0)n_i(x_0)n_j(x_0) \geq c > 0 \tag{34}$$

or

$$-a_{ij}^{\alpha^*,\beta}(x_0)d_{x_i x_j}(x_0) + b_i^{\alpha^*,\beta}(x_0)n_i(x_0) \geq c > 0 \tag{35}$$

where n denotes the exterior normal to Ω and d is the *signed distance* from $\partial\Omega$

$$d(x) = \begin{cases} \operatorname{dist}(x, \partial\Omega) & \text{if } x \in \Omega\,, \\ -\operatorname{dist}(x, \partial\Omega) & \text{if } x \in \mathbb{R}^N \setminus \Omega\,. \end{cases}$$

Assumptions (34) and (35) are the natural counterpart for Isaacs equation in (33) of the conditions for boundary regularity of solutions to linear equations in Chapt. 1 of [37]. We claim that

$$W(x) = 1 - e^{-\delta(d(x)+\lambda|x-x_0|^2)}$$

is an upper local barrier at x_0 for a suitable choice of $\delta, \lambda > 0$. Indeed it is easy to compute

$$\begin{aligned} -a_{ij}^{\alpha,\beta}(x_0)W_{x_i x_j}(x_0) + b_i^{\alpha,\beta}(x_0)W_{x_i}(x_0) + c^{\alpha,\beta}(x_0)W(x_0) - f^{\alpha,\beta}(x_0) = \\ -\delta a_{ij}^{\alpha,\beta}(x_0)d_{x_i x_j}(x_0) + \delta^2 a_{ij}^{\alpha,\beta}(x_0)d_{x_i}(x_0)d_{x_j}(x_0) + \delta b_i^{\alpha,\beta}(x_0)d_{x_i}(x_0) \\ -2\delta\lambda \operatorname{Tr}[a^{\alpha,\beta}(x_0)] - f^{\alpha,\beta}(x_0)\,. \end{aligned}$$

Next we choose α^* as above and assume first (34). In this case, since the coefficients are bounded and continuous and d is C^2, we can make W a supersolution of the PDE in (33) in a neighborhood of x_0 by taking δ large enough. If, instead, (35) holds, we choose first λ small and then δ large to get the same conclusion.

3. Applications to Optimal Control

3.1. A deterministic minimum-time problem

Our first example of application of the previous theory is the time-optimal control of nonlinear deterministic systems with a closed and nonempty target $\Gamma \subset \mathbb{R}^N$. For this minimum-time problem we prove that the lower semicontinuous envelope of the value function is the e-solution of the associated Dirichlet problem for the Bellman equation. This result can be also found in [7] and [4], but we give here a different and simpler proof. Consider the system

$$\text{(36)} \qquad \begin{cases} y\prime(t) = f(y(t), a(t)) \quad t > 0\,, \\ y(0) = x\,, \end{cases}$$

where $a \in \mathcal{A} := \{a : [0, \infty) \to A \text{ measurable}\}$ is the set of admissible controls, with

$$\text{(37)} \qquad \begin{array}{l} A \text{ a compact space, } f : \mathbb{R}^N \times A \to \mathbb{R}^N \text{ continuous} \\ \exists L > 0 \text{ such that } (f(x, a) - f(y, a)) \cdot (x - y) \le L|x - y|^2\,, \end{array}$$

for all $x, y \in \mathbb{R}^N$, $a \in A$. Under these assumptions, for any $a \in \mathcal{A}$ there exists a unique trajectory of the system (36) defined for all t, that we denote $y_x(t, a)$ or $y_x(t)$. We also define the minimum time for the system to reach the target using the control $a \in \mathcal{A}$:

$$t_x(a) := \begin{cases} \inf\{t \ge 0 : y_x(t, a) \in \Gamma\}\,, & \text{if } \{t \ge 0 : y_x(t, a) \in \Gamma\} \neq \emptyset\,, \\ +\infty & \text{otherwise}\,. \end{cases}$$

The *value function* for this problem, named *minimum time* function, is

$$T(x) = \inf_{a \in \mathcal{A}} t_x(a), \quad x \in \mathbb{R}^N\,.$$

Consider now the Kruzkov transformation of the minimum time

$$v(x) := \begin{cases} 1 - e^{-T(x)}\,, & \text{if } T(x) < \infty\,, \\ 1\,, & \text{otherwise}\,. \end{cases}$$

The new unknown v is itself the value function of a time-optimal control problem with a discount factor, and from its knowledge one recovers immediately the minimum time function T. We remark that in general v has no continuity properties without further assumptions; however, it is lower semicontinuous if $f(x, A)$ is a convex set for all x, so in such a case $v = v_*$ (see, e.g., [7, 4]).

The Dirichlet problem associated to v by the Dynamic Programming method is

$$\text{(38)} \qquad \begin{cases} v + H(x, Dv) = 0\,, & \text{in } \mathbb{R}^N \setminus \Gamma\,, \\ v = 0\,, & \text{in } \partial\Gamma\,, \end{cases}$$

where

$$H(x, p) := \max_{a \in A}\{-f(x, a) \cdot p\} - 1\,.$$

A Comparison Principle for this problem can be found, for instance, in [4].

THEOREM 8. *Assume (37). Then v_* is the e-solution and the minimal supersolution of (38).*

Proof. Note that by (37) and the fact that $w \equiv 0$ is a subsolution of (38), the hypotheses of Theorem 3 are satisfied, so the e-solution exists and it is a supersolution. It is well known that v_* is a supersolution of $v + H(x, Dv) = 0$ in $\mathbb{R}^N \setminus \Gamma$, see, e.g., [28, 8, 4]; moreover $v_* \geq 0$ on $\partial\Gamma$, so v_* is a supersolution of (38). In order to prove that v_* is the lower e-solution we construct a sequence of subsolutions of (38) converging to v_*.

Fix $\epsilon > 0$, and consider the set

$$\Gamma_\epsilon := \{x \in \mathbb{R}^N : \operatorname{dist}(x, \partial\Gamma) \leq \epsilon\},$$

let T_ϵ be the minimum time function for the problem with target Γ_ϵ, and v_ϵ its Kruzkov transformation. By standard results [28, 8, 4] v_ϵ is a non-continuous viscosity solution of

$$\begin{cases} v + H(x, Dv) = 0\,, & \text{in } \mathbb{R}^N \setminus \Gamma_\epsilon\,, \\ v = 0 \text{ or } v + H(x, Dv) = 0\,, & \text{in } \partial\Gamma_\epsilon\,. \end{cases}$$

With the same argument we used in Theorem 6, we can see that v_ϵ^* is a subsolution of (38). We define

$$u(x) := \sup_\epsilon v_\epsilon^*(x)$$

and will prove that $u = v_*$.

By the Comparison Principle $v_\epsilon^* \leq v_*$ for all $\epsilon > 0$, then $u(x) \leq v_*(x)$. To prove the opposite inequality we observe it is obvious in Γ and assume by contradiction there exists a point $\hat{x} \notin \Gamma$ such that:

$$\sup_\epsilon v_\epsilon(\hat{x}) \leq \sup_\epsilon v_\epsilon^*(\hat{x}) < v_*(\hat{x})\,. \tag{39}$$

Consider first the case $v_*(\hat{x}) < 1$, that is, $T_*(\hat{x}) < +\infty$. Then there exists $\delta > 0$ such that

$$T_\epsilon^*(\hat{x}) < T_*(\hat{x}) - \delta < +\infty, \text{ for all } \epsilon > 0\,. \tag{40}$$

By definition of minimum time, for all ϵ there is a control a_ϵ such that

$$t_{\hat{x}}^\epsilon(a_\epsilon) \leq T_\epsilon(\hat{x}) + \frac{\delta}{2} < +\infty\,. \tag{41}$$

Let $z_\epsilon \in \Gamma_\epsilon$ be the point reached at time $t_{\hat{x}}^\epsilon(a_\epsilon)$ by the trajectory starting from $\hat{x}$, using control a_ϵ. By standard estimates on the trajectories, we have for all ϵ

$$|z_\epsilon| = |y_{\hat{x}}(t_{\hat{x}}^\epsilon(a_\epsilon))| \leq \left(|\hat{x}| + \sqrt{2MT(\hat{x})}\right) e^{MT(\hat{x})}\,,$$

where $M := L + \sup\{|f(0, a)| : a \in A\}$. So, for some $R > 0$, $z_\epsilon \in \overline{B}(0, R)$ for all ϵ. Then we can find subsequences such that

$$z_{\epsilon_n} \to z \in \partial\Gamma, \quad t_n := t_{\hat{x}}^{\epsilon_n}(a_{\epsilon_n}) \to \bar{t}, \text{ as } n \to \infty\,. \tag{42}$$

From this, (40) and (41) we get

$$\bar{t} < T_*(\hat{x}) - \frac{\delta}{2}\,. \tag{43}$$

Let $\overline{y}_{\epsilon_n}$ be the solution of the system

$$\begin{cases} y\prime = f(y, a_{\epsilon_n}) & t < t_n\,, \\ y(t_{\hat{x}}(a_{\epsilon_n})) = z\,, \end{cases}$$

that is, the trajectory moving backward from z using control a_{ϵ_n}, and set $x_n := \overline{y}_{\epsilon_n}(0)$. In order to prove that $x_n \to \hat{x}$ we consider the solution y_{ϵ_n} of

$$\begin{cases} y\prime = f(y, a_{\epsilon_n}) & t < t_n\,, \\ y(t_n) = z_{\epsilon_n}\,, \end{cases}$$

that is, the trajectory moving backward from z_{ϵ_n} and using control a_{ϵ_n}. Note that $y_{\epsilon_n}(0) = \hat{x}$. By differentiating $|y_{\epsilon_n} - \overline{y}_{\epsilon_n}|^2$, using (37) and then integrating we get, for all $t < t_n$,

$$|y_{\epsilon_n}(t) - \overline{y}_{\epsilon_n}(t)|^2 \le |z_{\epsilon_n} - z|^2 + \int_t^{t_n} 2L|y_{\epsilon_n}(s) - \overline{y}_{\epsilon_n}(s)|^2 ds\,.$$

Then by Gronwall's lemma, for all $t < t_n$,

$$|y_{\epsilon_n}(t) - \overline{y}_{\epsilon_n}(t)| \le |z_{\epsilon_n} - z| e^{L(t_n - t)}\,,$$

which gives, for $t = 0$,

$$|\hat{x} - x_n| \le |z_{\epsilon_n} - z| e^{Lt_n}\,.$$

By letting $n \to \infty$, we get that $x_n \to \hat{x}$.

By definition of minimum time $T(x_n) \le t_n$, so letting $n \to \infty$ we obtain $T_*(\hat{x}) \le \overline{t}$, which gives the desired contradiction with (43).

The remaining case is $v_*(\hat{x}) = 1$. By (39) $T^*_\epsilon(\hat{x}) \le K < +\infty$ for all ϵ. By using the previous argument we get (42) with $\overline{t} < +\infty$ and $T_*(\hat{x}) \le \overline{t}$. This is a contradiction with $T_*(\hat{x}) = +\infty$ and completes the proof.

□

3.2. Maximizing the mean escape time of a degenerate diffusion process

In this subsection we study a stochastic control problem having as a special case the problem of maximizing the expected discounted time spent by a controlled diffusion process in a given open set $\Omega \subseteq \mathbb{R}^N$. A number of engineering applications of this problem are listed in [19], where, however, a different cost criterion is proposed and a nondegeneracy assumption is made on the diffusion matrix. We consider a probability space $(\Omega', \mathcal{F}, P)$ with a right-continuous increasing filtration of complete sub-σ fields $\{\mathcal{F}_t\}$, a Brownian motion B_t in $\mathbb{R}^M$ $\mathcal{F}_t$-adapted, a compact set A, and call $\mathcal{A}$ the set of progressively measurable processes α_t taking values in A. We are given bounded and continuous maps σ from $\mathbb{R}^N \times A$ into the set of $N \times M$ matrices and $b : \mathbb{R}^N \times A \to \mathbb{R}^N$ satisfying (14), (15) and consider the controlled stochastic differential equation

$$(SDE) \begin{cases} dX_t = \sigma^{\alpha_t}(X_t) dB_t - b^{\alpha_t}(X_t) dt\,, & t > 0\,, \\ X_0 = x\,. \end{cases}$$

For any $\alpha_. \in \mathcal{A}$ (SDE) has a pathwise unique solution X_t which is $\mathcal{F}_t$-progressively measurable and has continuous sample paths. We are given also two bounded and uniformly continuous maps $f, c : \mathbb{R}^N \times A \to \mathbb{R}$, $c^\alpha(x) \ge c_0 > 0$ for all x, α, and consider the payoff functional

$$J(x, \alpha_.) := E\left(\int_0^{t_x(\alpha_.)} f^{\alpha_t}(X_t) e^{-\int_0^t c^{\alpha_s}(X_s)\,ds} dt\right),$$

where E denotes the expectation and

$$t_x(\alpha_.) := \inf\{t \ge 0 : X_t \notin \Omega\},$$

where, as usual, $t_x(\alpha_.) = +\infty$ if $X_t \in \Omega$ for all $t \ge 0$. We want to maximize this payoff, so we consider the value function

$$v(x) := \sup_{\alpha_. \in \mathcal{A}} J(x, \alpha_.).$$

Note that for $f = c \equiv 1$ the problem becomes the maximization of the mean discounted time $E(1 - e^{-t_x(\alpha_.)})$ spent by the trajectories of (SDE) in Ω.

The Hamilton-Jacobi-Bellman operator and the Dirichlet problem associated to v by the Dynamic Programming method are

$$F(x, u, Du, D^2u) := \min_{\alpha \in A}\{-a_{ij}^\alpha(x) u_{x_i x_j} + b^\alpha(x) \cdot Du + c^\alpha(x) u - f^\alpha(x)\},$$

where the matrix (a_{ij}) is $\frac{1}{2}\sigma\sigma^T$, and

$$\text{(44)} \qquad \begin{cases} F(x, u, Du, D^2u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

see, for instance, [40, 35, 36, 22, 32] and the references therein. The proof that the value function satisfies the Hamilton-Jacobi-Bellman PDE is based on the Dynamic Programming Principle

$$\text{(45)} \qquad v(x) = \sup_{\alpha_. \in \mathcal{A}} E\left(\int_0^{\theta \wedge t_x} f^{\alpha_t}(X_t) e^{-\int_0^t c^{\alpha_s}(X_s)\,ds} dt + v(X_{\theta \wedge t_x}) e^{-\int_0^{\theta \wedge t_x} c^{\alpha_s}(X_s)\,ds}\right),$$

where $t_x = t_x(\alpha_.)$, for all $x \in \overline{\Omega}$ and all $\mathcal{F}_t$-measurable stopping times θ. Although the DPP (45) is generally believed to be true under the current assumptions (see, e.g., [35]), we were able to find its proof in the literature only under some additional conditions, such as the convexity of the set

$$\{(a^\alpha(x), b^\alpha(x), f^\alpha(x), c^\alpha(x)) : \alpha \in A\}$$

for all $x \in \overline{\Omega}$, see [20] (this is true, in particular, when relaxed controls are used), or the independence of the variance of the noise from the control [15], i.e., $\sigma^\alpha(x) = \sigma(x)$ for all x, or the continuity of v [35]. As recalled in Subsection 1.1 a Comparison Principle for (44) can be found in [29], see also [18] and the references therein.

In order to prove that v is the e-solution of (44), we approximate Ω with a nested family of open sets with the properties

$$\text{(46)} \qquad \Theta_\epsilon \subseteq \Omega,\ \epsilon \in]0, 1],\ \Theta_\epsilon \supseteq \overline{\Theta}_\delta \text{ for } \epsilon < \delta,\ \bigcup_\epsilon \Theta_\epsilon = \Omega.$$

For each $\epsilon > 0$ we call v_ϵ the value function of the same control problem with t_x replaced with

$$t_x^\epsilon(\alpha_.) := \inf\{t \ge 0 : X_t \notin \Theta_\epsilon\}$$

in the definition of the payoff J. In the next theorem we assume that each v_ϵ satisfies the DPP (45) with t_x replaced with t_x^ϵ.

Finally, we make the additional assumption

$$f^\alpha(x) \geq 0 \text{ for all } x \in \Omega,\ \alpha \in A\,. \tag{47}$$

which ensures that $\underline{u} \equiv 0$ is a subsolution of (44). The main result of this subsection is the following.

THEOREM 9. *Under the previous assumptions the value function v is the e-solution and the minimal supersolution of (44), and*

$$v = \sup_{0<\epsilon\leq 1} v_\epsilon = \lim_{\epsilon\searrow 0} v_\epsilon\,.$$

Proof. Note that v_ϵ is nondecreasing as $\epsilon \searrow 0$, so $\lim_{\epsilon\searrow 0} v_\epsilon$ exists and equals the sup. By Theorem 3 with $g \equiv 0$, $\underline{u} \equiv 0$, there exists the e-solution H_0 of (44). We consider the functions u_ϵ defined by (20) and claim that

$$u_{2\epsilon} \leq (v_\epsilon)_* \leq v_\epsilon^* \leq H_0\,.$$

Then

$$H_0 = \sup_{0<\epsilon\leq 1} v_\epsilon\,, \tag{48}$$

because $H_0 = \sup_\epsilon u_{2\epsilon}$ by Remark 1. We prove the claim in three steps.

Step 1. By standard methods [35, 9], the Dynamic Programming Principle for v_ϵ implies that v_ϵ is a non-continuous viscosity solution of the Hamilton-Jacobi-Bellman equation $F = 0$ in Θ_ϵ and v_ϵ^* is a viscosity subsolution of the boundary condition

$$u = 0 \text{ or } F(x, u, Du, D^2u) = 0 \text{ on } \partial\Theta_\epsilon\,, \tag{49}$$

as defined in Subsection 1.3.

Step 2. Since $(v_\epsilon)_*$ is a supersolution of the PDE $F = 0$ in Θ_ϵ and $(v_\epsilon)_* \geq 0$ on $\partial\Theta_\epsilon$, the Comparison Principle implies $(v_\epsilon)_* \geq w$ for any subsolution w of (44) such that $w = 0$ on $\partial\Theta_\epsilon$. Since $\partial\Theta_\epsilon \subseteq \Omega \setminus \overline{\Theta}_{2\epsilon}$ by (46), we obtain $u_{2\epsilon} \leq v_{\epsilon *}$ by the definition (20) of $u_{2\epsilon}$.

Step 3. We claim that v_ϵ^* is a subsolution of (44). In fact we noted before that it is a subsolution of the PDE in Θ_ϵ, and this is true also in $\Omega \setminus \overline{\Theta}_\epsilon$ where $v_\epsilon^* \equiv 0$ by (47), whereas the boundary condition is trivial. It remains to check the PDE at all points of $\partial\Theta_\epsilon$. Given $\hat{x} \in \partial\Theta_\epsilon$, we must prove that for all $\phi \in C^2(\Omega)$ such that $v_\epsilon^* - \phi$ attains a local maximum at $\hat{x}$, we have

$$F(\hat{x}, v_\epsilon^*(\hat{x}), D\phi(\hat{x}), D^2\phi(\hat{x})) \leq 0\,. \tag{50}$$

$1st$ Case: $v_\epsilon^*(\hat{x}) > 0$. Since v_ϵ^* satisfies (49), for all $\phi \in C^2(\overline{\Theta}_\epsilon)$ such that $v_\epsilon^* - \phi$ attains a local maximum at $\hat{x}$ (50) holds. Then the same inequality holds for all $\phi \in C^2(\Omega)$ as well.

$2nd$ Case: $v_\epsilon^*(\hat{x}) = 0$. Since $v_\epsilon^* - \phi$ attains a local maximum at $\hat{x}$, for all x near $\hat{x}$ we have

$$v_\epsilon^*(x) - v_\epsilon^*(\hat{x}) \leq \phi(x) - \phi(\hat{x})\,.$$

By Taylor's formula for ϕ at $\hat{x}$ and the fact that $v_\epsilon^*(x) \geq 0$, we get

$$D\phi(\hat{x}) \cdot (x - \hat{x}) \geq o(|x - \hat{x}|)\,,$$

and this implies $D\phi(\hat{x}) = 0$. Then Taylor's formula for ϕ gives also

$$(x - \hat{x}) \cdot D^2\phi(\hat{x})(x - \hat{x}) \geq o(|x - \hat{x}|^2),$$

and this implies $D^2\phi(\hat{x}) \geq 0$, as it is easy to check. Then

$$F(\hat{x}, v_\epsilon^*(\hat{x}), D\phi(\hat{x}), D^2\phi(\hat{x})) = F(\hat{x}, 0, 0, D^2\phi(\hat{x})) \leq 0$$

because $a^\alpha \geq 0$ and $f^\alpha \geq 0$ for all x and α. This completes the proof that v_ϵ^* is a subsolution of (44). Now the Comparison Principle yields $v_\epsilon^* \leq H_0$, since H_0 is a supersolution of (44).

It remains to prove that $v = \sup_{0<\epsilon\leq 1} v_\epsilon$. To this purpose we take a sequence $\epsilon_n \searrow 0$ and define

$$J_n(x, \alpha_.) := E\left(\int_0^{t_x^{\epsilon_n}(\alpha_.)} f^{\alpha_t}(X_t)e^{-\int_0^t c^{\alpha_s}(X_s)\,ds}dt\right).$$

We claim that

$$\lim_n J_n(x, \alpha_.) = \sup_n J_n(x, \alpha_.) = J(x, \alpha_.) \text{ for all } \alpha_. \text{ and } x\,.$$

The monotonicity of $t_x^{\epsilon_n}$ follows from (46) and it implies the monotonicity of J_n by (47). Let

$$\tau := \sup_n t_x^{\epsilon_n}(\alpha_.) \leq t_x(\alpha_.),$$

and note that $t_x(\alpha_.) = +\infty$ if $\tau = +\infty$. In the case $\tau < +\infty$, $X_{t_x^{\epsilon_n}} \in \partial\Theta_{\epsilon_n}$ implies $X_\tau \in \partial\Omega$, so $\tau = t_x(\alpha_.)$ again. This and (47) yield the claim by the Lebesgue monotone convergence theorem. Then

$$v(x) = \sup_{\alpha_.}\sup_n J_n(x, \alpha_.) = \sup_n\sup_{\alpha_.} J_n(x, \alpha_.) = \sup_n v_{\epsilon_n} = \sup_\epsilon v_\epsilon\,,$$

so (48) gives $v = H_0$ and completes the proof.

□

REMARK 7. From Theorem 9 it is easy to get a *Verification theorem* by taking the supersolutions of (44) as verification functions. We consider a presynthesis $\alpha^{(x)}$, that is, a map $\alpha^{(\cdot)} : \Omega \to \mathcal{A}$, and say it is optimal at x_o if $J(x_o, \alpha^{(x_o)}) = v(x_o)$. Then Theorem 9 gives immediately the following sufficient condition of optimality: *if there exists a verification function* W *such that* $W(x_o) \leq J(x_o, \alpha^{(x_o)})$, *then* $\alpha^{(\cdot)}$ *is optimal at* x_o; moreover, a characterization of global optimality is the following: $\alpha^{(\cdot)}$ *is optimal in* Ω *if and only if* $J(\cdot, \alpha^{(\cdot)})$ *is a verification function*.

REMARK 8. We can combine Theorem 9 with the results of Subsection 2.2 to approximate the value function v with smooth value functions. Consider a Brownian motion $\tilde{B}_t$ in $\mathbb{R}^N$ $\mathcal{F}_t$-adapted and replace the stochastic differential equation in (SDE) with

$$dX_t = \sigma^{\alpha_t}(X_t)dB_t - b^{\alpha_t}(X_t)dt + \sqrt{2h}\,d\tilde{B}_t, \quad t > 0\,,$$

for $h > 0$. For a family of nested open sets with the properties (46) consider the value function v_h^ϵ of the problem of maximizing the payoff functional J with t_x replaced with t_x^ϵ. Assume for simplicity that $a^\alpha, b^\alpha, c^\alpha, f^\alpha$ are smooth (otherwise we can approximate them by mollification). Then v_h^ϵ is the classical solution of (30), where F is the HJB operator of this subsection and $\underline{u} \equiv 0$, by the results in [21, 24, 36, 31], and it is possible to synthesize an optimal Markov control policy for the problem with $\epsilon, h > 0$ by standard methods (see, e.g., [22]). By Theorem 6 v_h^ϵ converges to v as $\epsilon, h \searrow 0$ with h linked to ϵ.

References

[1] BARDI M., BOTTACIN S., *Discontinuous solution of degenerate elliptic boundary value problems*, Preprint 22, Dip. di Matematica P. e A., Università di Padova, 1995.

[2] BARDI M., BOTTACIN S., *Characteristic and irrelevant boundary points for viscosity solutons of nonlinear degenerate elliptic equations*, Preprint 25, Dip. di Matematica P. e A., Università di Padova, 1998.

[3] BARDI M., BOTTACIN S., FALCONE M., *Convergence of discrete schemes for discontinuous value functions of pursuit-evasion games*, in "New Trends in Dynamic Games and Applications", G .J. Olsder ed., Birkhäuser, Boston 1995, 273–304.

[4] BARDI M., CAPUZZO-DOLCETTA I., *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston 1997.

[5] BARDI M., CRANDALL M., EVANS L. C., SONER H. M., SOUGANIDIS P. E., *Viscosity solutions and applications*, I. Capuzzo Dolcetta and P.-L. Lions eds., Springer Lecture Notes in Mathematics 1660, Berlin 1997.

[6] BARDI M., GOATIN P., ISHII H., *A Dirichlet type problem for nonlinear degenerate elliptic equations arising in time-optimal stochastic control*, Preprint 1, Dip. di Matematica P. e A., Università di Padova, 1998, to appear in Adv. Math. Sci. Appl.

[7] BARDI M., STAICU V., *The Bellman equation for time-optimal control of noncontrollable, nonlinear systems*, Acta Appl. Math. **31** (1993), 201–223.

[8] BARLES G., *Solutions de viscosité des equations de Hamilton-Jacobi*, Springer-Verlag, 1994.

[9] BARLES G., BURDEAU J., *The Dirichlet problem for semilinear second-order degenerate elliptic equations and applications to stochastic exit time control problems*, Comm. Partial Differential Equations **20** (1995), 129–178.

[10] BARLES G., PERTHAME B., *Discontinuous solutions of deterministic optimal stopping time problems*, RAIRO Modél. Math. Anal. Numér **21** (1987), 557–579.

[11] BARRON E. N., JENSEN R., *Semicontinuous viscosity solutions of Hamilton-Isaacs equation with convex Hamiltonians*, Comm. Partial Differential Equations **15** (1990), 1713–1742.

[12] BERESTYCKI H., NIRENBERG L., VARADHAN S. R. S., *The Principal Eigenvalue and Maximum Principle for Second-Order Elliptic Operators in General Domains*, Comm. Pure App. Math. **XLVII** (1994), 47–92.

[13] BETTINI P., *Problemi di Cauchy per equazioni paraboliche degeneri con dati discontinui e applicazioni al controllo e ai giochi differenziali*, Thesis, Università di Padova, 1998.

[14] BLANC A.-PH., *Deterministic exit time control problems with discontinuous exit costs*, SIAM J. Control Optim. **35** (1997), 399–434.

[15] BORKAR V. S., *Optimal control of diffusion processes*, Pitman Research Notes in Mathematics Series **203**, Longman, Harlow 1989.

[16] BRELOT M., *Familles de Perron et problème de Dirichlet*, Acta Litt. Sci. Szeged **9** (1939), 133–153.

[17] CAFFARELLI L. A., CABRÉ X., *Fully nonlinear elliptic equations*, Amer. Math. Soc., Providence, RI, 1995.

[18] CRANDALL M. C., ISHII H., LIONS P. L., *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. **27** (1992), 1–67.

[19] DUPUIS P., MCENEANEY W. M., *Risk-sensitive and robust escape criteria*, SIAM J. Control Optim. **35** (1997), 2021–2049.

[20] EL KAROUI N., HÙÙ NGUYEN D., JEANBLANC-PICQUÉ M., *Compactification methods in the control of degenerate diffusions: existence of an optimal control*, Stochastics **20** (1987), 169–219.

[21] EVANS L. C., *Classical solutions of the Hamilton-Jacobi-Bellman equation for uniformally elliptic operators*, Trans. Amer. Math. Soc. **275** (1983), 245–255.

[22] FLEMING W. H., SONER H. M., *Controlled Markov Process and Viscosity Solutions*, Springer-Verlag, New York 1993.

[23] FREIDLIN M., *Functional integration and partial differential equations*, Princeton University Press, Princeton 1985.

[24] GILBARG D., TRUDINGER N. S., *Elliptic Partial Differential Equations of Second Order*, 2nd ed. Springer-Verlag, Berlin 1983.

[25] HEINONEN J., KILPELÄINEN T., MARTIO O., *Nonlinear Potential Theory of Degenerate Elliptic Equations*, Oxford Science Publications, Clarendon Press, Oxford 1993.

[26] HELMS L. L., *Introduction to Potential Theory*, Wiley-Interscience, New York 1969.

[27] ISHII H., *Perron's method for Hamilton-Jacobi equations*, Duke Math. J. **55** (1987), 369–384.

[28] ISHI H., *A boundary value problem of the Dirichlet type for Hamilton-Jacobi equations*, Ann. Scuola Norm. Sup. Pisa **16** (1986), 105–135.

[29] JENSEN R., *Uniqueness criteria for viscosity solutions of fully nonlinear elliptic partial differential equations*, Indiana Univ. Math. J. **38** (1989), 629–667.

[30] KELLOGG O. D., *Foundations of Potential Theory*, Verlag-Springer, Berlin 1929.

[31] KRYLOV N. V., *Nonlinear Elliptic and Parabolic Equations of the Second Order*, D. Reidel Publishing Company, Dordrecht 1987.

[32] KRYLOV N. V., *Smoothness of the payoff function for a controllable process in a domain*, Math. USSR-Izv. **34** (1990), 65–95.

[33] KRYLOV N. V., *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Mathematics, 12, American Mathematical Society, Providence, RI, 1996.

[34] LIONS P. L., *Generalized solutions of Hamilton-Jacobi equations*, Pitman, Boston 1982.

[35] LIONS P. L., *Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 1: The dynamic programming principle and applications, Part 2: Viscosity solutions and uniqueness*, Comm. Partial. Differential Equations **8** (1983), 1101–1174 and 1229–1276.

[36] LIONS P.-L., *Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. III - Regularity of the optimal cost function*, in "Nonlinear partial differential equations and applications", Collège de France seminar, Vol. V (Paris 1981/1982), 95–205, Res. Notes in Math., 93, Pitman, Boston, Mass.-London 1983.

[37] OLEINIK O. A., RADKEVIC E. V., *Second order equations with nonnegative characteristic form*, Plenum Press, New York 1973.

[38] RAMASWAMY M., RAMASWAMY S., *Perron's method and barrier functions for the viscosity solutions of the Dirichlet problem for some nonlinear partial differential equations*, Z. Anal. Anwendungen **13** (1994), 199–207.

[39] SORAVIA P., *Discontinuous viscosity solutions to Dirichlet problems for Hamilton-Jacobi equations with convex Hamiltonians*, Comm. Partial Diff. Equations **18** (1993), 1493–1514.

[40] STROOCK D., VARADHAN S. R. S., *On Degenerate Elliptic-Parabolic Operators of Second Order and Their Associated Diffusions*, Comm. Pure App. Math. **XXV** (1972), 651–713.

[41] SUBBOTIN A. I., *Discontinuous solutions of a Dirichlet type boundary value problem for first order P.D.E.*, Russian Numer. Anal. Math. Modelling J. **8** (1993), 145–164.

[42] SUBBOTIN A. I., *Generalized solutions of first order PDEs: The Dynamic Optimization Perspective*, Birkhäuser, Boston 1995.

[43] TRUDINGER N. S., *On the twice differentiability of viscosity solutions of nonlinear elliptic equations*, Bull. Austral. Math. Soc. **39** (1989), 443–447.

[44] WIENER N., *Certain notions in Potential Theory*, J. Math. Phys. (M.I.T.) **3** (1924), 24–51.

[45] WIENER N., *Note on a paper of O. Perron*, J. Math. Phys. (M.I.T.) **4** (1925), 21–32.

AMS Subject Classification: ???.

Martino BARDI, Sandra BOTTACIN
Dipartimento di Matematica P. e A.
Università di Padova
via Belzoni 7, I-35131 Padova, Italy

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

R. M. Bianchini

HIGH ORDER NECESSARY OPTIMALITY CONDITIONS

Abstract.
In this paper we present a method for determining some variations of a singular trajectory of an affine control system. These variations provide necessary optimality conditions which may distinguish between maximizing and minimizing problems. The generalized Legendre-Clebsch conditions are an example of these type of conditions.

1. Introduction

The variational approach to Majer minimization control problems can be roughly summarized in the following way: let x^* be a solution on the interval $[t_i, t_e]$ relative to the control u^*; if the pair (x^*, u^*) is optimal, then the cone of tangent vectors to the reachable set at $x^*(t_e)$ is contained in the subspace where the cost increase. If there are constraints on the end-points, then the condition is no more necessary; nevertheless in [1] it has been proved that particular subcones of tangent vectors, the regular tangent cones, have to be contained in a cone which depends on the cost and on the constraints. Tangent vectors whose collection is a regular tangent cone are named good trajectory variations, see [8].

The aim of this paper is to construct good trajectory variations of a trajectory of an affine control process which contains singular arcs, i.e. arcs of trajectory relative to the drift term of the process. It is known, [2], that the optimal trajectory of an affine control process may be of this type; however the pair $(x^*, 0)$ may satisfy the Pontrjagin Maximum Principle without being optimal. Therefore it is of interest in order to single out a smaller number of candidates to the optimum, to know as many good trajectory variations as we can.

In [5] good trajectory variations of the pair $(x^*, 0)$ have been constructed by using the relations in the Lie algebra associated to the system at the points of the trajectory. The variations constructed in that paper are of bilateral type, i.e. both the directions $+v$ and $-v$ are good variations. In this paper I am going to find conditions which single out unilateral variations, i.e. only one direction need to be a variation. Unilateral variations are of great interest because, contrary to the bilateral ones, they distinguish between maximizing and minimizing problems.

2. Notations and preliminary results

To each family $\mathbf{f} = (f_0,\ f_1,\ \ldots,\ f_m)$ of C^∞ vector fields on a finite dimensional manifold M we associate the affine control process $\Sigma_{\mathbf{f}}$ on M

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) \qquad |u_i| \leq \alpha \tag{1}$$

where the control $u = (u_1, \dots, u_m)$ is a piecewise constant map whose values belong to the hypercube $|u_i| \le \alpha$. We will denote by $S_{\mathbf{f}}(t, t_0, y, u)$ the value at time t of the solution of $\Sigma_{\mathbf{f}}$ relative to the control u, which at time t_0 is equal to y. We will omit t_0 if it is equal to 0, so that $S_{\mathbf{f}}(t, y, u) = S_{\mathbf{f}}(t, 0, y, u)$; we will also use the exponential notation for constant control map, for example $\exp t f_0 \cdot y = S_{\mathbf{f}}(t, y, 0)$.

We want to construct some variations of the trajectory $t \mapsto x_{\mathbf{f}}(t) = \exp t f_0 \cdot x_0$, $t \in [t_0, t_1]$ at time $\tau \in [t_0, t_1]$. We will consider trajectory variations produced by needle-like control variations concentrated at τ. The definition is the following:

DEFINITION 1. *A vector $v \in T_{x_{\mathbf{f}}(\tau)}M$ is a right (left) trajectory variation of $x_{\mathbf{f}}$ at τ if for each $\epsilon \in [0, \overline{\epsilon}]$ there exists a control map $u(\epsilon)$ defined on the interval $[0, a(\epsilon)]$, $\lim_{\epsilon \to 0^+} a(\epsilon) = 0$, such that $u(\epsilon)$ depends continuously on ϵ in the L^1 topology and the map $\epsilon \mapsto \exp(-a(\epsilon) f_0) \cdot S_{\mathbf{f}}(a(\epsilon), x_{\mathbf{f}}(\tau), u(\epsilon))$, ($\epsilon \mapsto S_{\mathbf{f}}(a(\epsilon), x_{\mathbf{f}}(\tau - a(\epsilon)), u(\epsilon))$) has v as tangent vector at $\epsilon = 0$.*

The variations at τ indicates the controllable directions of the reference trajectory from $x_{\mathbf{f}}(\tau)$; they are local objects at $x_{\mathbf{f}}(\tau)$ and in any chart at $x_{\mathbf{f}}(\tau)$ they are characterized by the property

$$S_{\mathbf{f}}(a(\epsilon), x_{\mathbf{f}}(\tau), u(\epsilon)) = x_{\mathbf{f}}(\tau + a(\epsilon)) + \epsilon\, v + o(\epsilon) \in R(\tau + a(\epsilon), x_{\mathbf{f}}(t_0)) \tag{2}$$

where $R(t, x)$ is the set of points reachable in time t from x.

The transport along the reference flow generated by the 0 control from time τ to time t_1 of a variation at τ is a tangent vector to the reachable set at time t_1 in the point $x_{\mathbf{f}}(t_1)$. The transport of particular trajectory variations, the good ones, gives rise to tangent vectors whose collection is a regular tangent cone. The definition of good variations is the following:

DEFINITION 2. *A vector $v \in T_{x^*(\tau)}M$ is a good* right variation (left variation) at τ of order *k if there exists positive numbers $\overline{c}$, $\overline{\epsilon}$ and for each $\epsilon \in [0, \overline{\epsilon}]$ a family of admissible control maps, $u_\epsilon(c)$, $c \in [0, \overline{c}]$ with the following properties:*

1. *$u_\epsilon(c)$ is defined on the interval $[0, a\epsilon^k]$*
2. *for each ϵ, $c \mapsto u_\epsilon(c)$ is continuous in the L^1 topology*
3. *$\exp[-(1 + a)\epsilon^k] f_0 \cdot S_{\mathbf{f}}(a\epsilon^k, x_{\mathbf{f}}(\tau + \epsilon^k), u_\epsilon(c)) = x_{\mathbf{f}}(\tau) + \epsilon c\, v + o(\epsilon)$ ($\exp \epsilon^k f_0 \cdot S_{\mathbf{f}}(a\epsilon^k, x_{\mathbf{f}}(\tau - (1 + a)\epsilon^k), u_\epsilon(c)) = x_{\mathbf{f}}(\tau) + \epsilon c\, v + o(\epsilon)$) uniformly w.r.t. c.*

The good trajectory variations will be simpler named g-variations. Standing the definitions, the variations of a trajectory are more easily found than its g-variations. However a property proved in [8] allows to find g-variations as limit points of trajectory ones. More precisely the following Proposition holds:

PROPOSITION 1. *Let I be an interval contained in $[t_0, t_1]$ and let $g \in L^1(I)$ be such that $g(t)$ is a right (left) trajectory variation at t for each t in the set L^+, (L^-), of right (left) Lebesgue points of g. For each $t \in L^+$, ($t \in L^-$), let $[0, a_t(\epsilon)]$ be the interval as in Definition 1 relatively to the variation $g(t)$. If there exists positive numbers N and s such that for each $\tau \in L^+$, ($\tau \in L^-$), $0 < a_\tau(\epsilon) \le (N\epsilon)^s$, then for each $t \in L^+$ $g(t)$ is a right variation, (for each $t \in L^-$, $g(t)$ is a left variation), at t of order s.*

Let τ be fixed; to study the variations at τ we can suppose without loss of generality that M is an open neighborhood of $0 \in \mathbb{R}^n$. Moreover by Corollary 3.3 in [5], we can substitute to the family $\mathbf{f}$ the family ϕ where ϕ_i is the Taylor polynomial of f_i of order sufficiently large. We can therefore suppose that $\mathbf{f}$ is an analytic family of vector fields on $\mathbb{R}^n$.

Let me recall some properties of analytic family of vectors fields. Let $\mathbf{X} = \{X_0, X_1, \ldots, X_m\}$ be $(m+1)$ indeterminate; $L(\mathbf{X})$ is the Lie algebra generated by $\mathbf{X}$ with Lie bracket defined by

$$[S, T] = ST - TS .$$

$\hat{L}(\mathbf{X})$ denotes the set of all formal series, $\sum_{k=1}^{\infty} P_k$, each P_k homogeneous Lie polynomial of degree k. For each $S \in \hat{L}(\mathbf{X})$ we set

$$\exp S = \sum_{k=0}^{\infty} \frac{S^k}{k!}$$

and

$$\log(Id + Z) = \sum_{k=0}^{\infty} \frac{(-1)^{k+1} Z^k}{k} .$$

The following identities hold

$$\exp(\log Z) = Z \qquad\qquad \log(\exp S) = S .$$

Formula di Campbell-Hausdorff [9]
For each $P,\ Q \in \hat{L}(\mathbf{X})$ there exists an unique $Z \in \hat{L}(\mathbf{X})$ such that

$$\exp P \cdot \exp Q = \exp Z$$

and Z is given by

$$Z = P + Q + \frac{1}{2}[P, Q] + \cdots$$

Let u be an admissible piecewise constant control defined on the interval $[0, T(u)]$; by the Campbell Hausdorff formula we can associate to u an element of $\hat{L}(\mathbf{X})$, $\log u$, in the following way: if $u(t) = (\omega_1^i, \ldots, \omega_m^i)$ in the interval (t_{i-1}, t_i) then

$$\exp \log u = \exp(T(u) - t_{k-1}) \left(X_0 + \sum_{i=1}^{m} \omega_i^k X_i \right) \cdot \cdots \cdot \exp t_1 \left(X_0 + \sum_{i=1}^{m} \omega_i^1 X_i \right)$$

If $\mathbf{f}$ is an analytic family, then $\log u$ is linked to $S_{\mathbf{f}}(T(u), y, u)$ by the following proposition [6]

PROPOSITION 2. *If* $\mathbf{f}$ *is an analytic family of vector fields on an analytic manifold* M *then for each compact* $K \subset M$ *there exist* T *such that, if* $\log_{\mathbf{f}} u$ *denotes the serie of vector fields obtained by substituting in* $\log u$, f_i *for* X_i, *then* $\forall y \in K$ *and* $\forall u,\ T(u) < T$, *the serie* $\exp \log_{\mathbf{f}} u \cdot y$ *converges uniformly to* $S_{\mathbf{f}}(T(u), y, u)$.

In the sequel we will deal only with right variations. The same ideas can be used to construct left variations.

To study the right trajectory variations it is useful to introduce $\operatorname{Log} u$ defined by

$$\exp(\operatorname{Log} u) = \exp -T(u) X_0 \cdot \exp(\log u) .$$

By definition it follows that if y belongs to a compact set and $T(u)$ is sufficiently small, then $\exp(-T(u)) f_0 \cdot S_{\mathbf{f}}(T(u), y, u)$ is defined and it is the sum of the serie $\exp(\operatorname{Log}_{\mathbf{f}} u) y$; notice that

$\exp(\mathrm{Log}_{\mathbf{f}} u)y$ is the value at time $T(u)$ of the solution of the pullback system introduced in [4] starting at y.

Let $u(\epsilon)$ be a family of controls which depend continuously on ϵ and such that $T(u(\epsilon)) = o(1)$. Such a family will be named control variation if

$$\mathrm{Log}\, u(\epsilon) = \sum \epsilon^{j_i}\, Y^i \tag{3}$$

with $Y^i \in \mathrm{Lie}\,\mathbf{X}$ and $j_i < j_{i+1}$. Let j_i be the smallest integer for which $Y^s_{\mathbf{f}}(x_{\mathbf{f}}(\tau)) \neq 0$; Y^i is named $\mathbf{f}$-*leading term* of the control variation at τ because it depends on the family $\mathbf{f}$ and on the time τ.

The definition of exp and Proposition 2 imply that if Y_i is an $\mathbf{f}$-leading term of a control variation, then

$$\exp(-T(u(\epsilon)) \cdot S_{\mathbf{f}}(T(u(\epsilon)),\ x_{\mathbf{f}}(\tau),\ u(\epsilon)) = x_{\mathbf{f}}(\tau) + \epsilon^{j_i} Y^i_{\mathbf{f}}(x_{\mathbf{f}}(\tau)) + o(\epsilon^{j_i})\,;$$

therefore by Definition 1, $Y^j_{\mathbf{f}}(x(\tau))$ is a variation of $x_{\mathbf{f}}$ at τ of order $1/j_i$. Since the set of variation is a cone, we have:

PROPOSITION 3. *Let Θ be an element of* Lie $\mathbf{X}$*; if a positive multiple of Θ is the* $\mathbf{f}$*-leading element at τ of a control variation, then $\Theta_{\mathbf{f}}(x_{\mathbf{f}}(\tau))$ is a variation at τ.*

3. General Result

The results of the previous section can be improved by using the relations in Lie $\mathbf{f}$ at $x_{\mathbf{f}}(\tau)$. The idea is that these relations allow to modify the leading term of a given control variation and therefore one can obtain more than one trajectory variation from a control variation.

Let us recall some definitions given by Susmann, [6], [7].

DEFINITION 3. *An* admissible weight *for the process (1) is a set of positive numbers, $\mathbf{l} = (l_0,\ l_1,\ \ldots, l_m)$, which verify the relations $l_0 \leq l_i$, $\forall i$.*

By means of an admissible weight, one can give a weight to each bracket in Lie $\mathbf{X}$, [6]. Let Λ be a bracket in the indeterminate $X_i's$; $|\Lambda|_i$ is the number of times that X_i appears in Λ.

DEFINITION 4. *Let $\mathbf{l} = \{l_0, l_1, \ldots, l_m\}$ be an admissible weight, the* $\mathbf{l}$*-weight of a bracket Φ is given by*

$$\|\Phi\|_{\mathbf{l}} = \sum_{i=0}^{m} l_i |\Phi|_i\ .$$

An element $\Theta \in$ Lie $\mathbf{X}$ *is said* $\mathbf{l}$*-homogeneous if it is a linear combination of brackets with the same* $\mathbf{l}$*-weight, which we name the* $\mathbf{l}$*-weight of the element.*

The weight of a bracket, Φ, with respect to the standard weight $\mathbf{l} = \{1, 1, \ldots, 1\}$ coincides with its length and it is denoted by $\|\Phi\|$.

The weight introduce a partial order relation in Lie $\mathbf{X}$.

DEFINITION 5. *Let $\Theta \in$* Lie $\mathbf{X}$*; following Susmann [7] we say that Θ is* $\mathbf{l}\,\mathbf{f}$-neutralized *at a point y if the value at y of $\Theta_{\mathbf{f}}$ is a linear combination of the values of brackets with less* $\mathbf{l}$*-weight, i.e. $\Theta_{\mathbf{f}}(y) = \sum \alpha_j\, \Phi^j_{\mathbf{f}}(y)$, $\|\Phi^j\|_{\mathbf{l}} < \|\Theta\|_{\mathbf{l}}$. The number* $\max \|\Phi_j\|$ *is the* order of the neutralization.

Let N be a positive integer; with S_N we denote the subspace of $\mathrm{Lie}\,\mathbf{X}$ spanned by the brackets whose length is not greater than N and with Q_N we denote the subspace spanned by the brackets whose length is greater than N. $\mathrm{Lie}\,\mathbf{X}$ is direct sum of S_N and Q_N.

DEFINITION 6. *Let u be any control;* $\log_N u$ *and* $\mathrm{Log}_N u$ *are the projections of* $\log u$ *and* $\mathrm{Log}\, u$ *respectively, on S_N.*

DEFINITION 7. *An element $\Phi \in S_N$ is a N-*good *element if there exists a neighborhood V of* 0 *in S_N and a C^1 map $u : V \to L^1$, such that $u(V)$ is contained in the set of admissible controls and*

$$\mathrm{Log}_N\, u(\Theta) = \Phi + \Theta\,.$$

Notice that there exist N-good elements whatever is the natural N.

We are going to present a general result.

THEOREM 1. *Let Z be an N-good element and let $\mathbf{l}$ be an admissible weight. $Z = \sum Y^i$, Y^i $\mathbf{l}$-homogeneous element such that if $b_i = \|Y^i\|_{\mathbf{l}}$, then $b_i \le b_j$ if $i < j$. If there exists j such that for each $i < j$, Y^i is $\mathbf{l}\,\mathbf{f}$-neutralized at τ with order not greater than N and $b_j < b_{j+1}$, then*

1. *$Y^j_{\mathbf{f}}(x_{\mathbf{f}}(\tau))$ is a variation at τ of order $\|Y^j\|_{\mathbf{l}}$;*
2. *if Φ is a bracket contained in S_N, $\|\Phi\|_{\mathbf{l}} < b_j$, then $\pm\Phi_{\mathbf{f}}(x_{\mathbf{f}}(\tau))$ is a variation at τ of order $\|\Phi\|_{\mathbf{l}}$.*

Proof. We are going to provide the proof in the case in which there is only one element which is $\mathbf{l}\,\mathbf{f}$-neutralized at τ. The proof of the general case is analogous. By hypothesis there exist $\mathbf{l}$-homogeneous elements W^j, $c_j = \|W^j\|_{\mathbf{l}} < \|Y^1\|_{\mathbf{l}}$, such that:

$$Y^1_{\mathbf{f}}(x_{\mathbf{f}}(\tau)) = \sum \alpha_j W^j_{\mathbf{f}}(x_{\mathbf{f}}(\tau))\,. \tag{4}$$

Let u be an admissible control; the control defined in $[0, \epsilon^{l_0} T(u)]$ by

$$\delta_\epsilon\, u(t) = (\epsilon^{l_1 - l_0} u_1(t/\epsilon^{l_0}), \ldots, \epsilon^{l_m - l_0} u_m(t/\epsilon^{l_0}))$$

is an admissible control; such control will be denoted by $\delta_\epsilon u$. The map $\epsilon \mapsto \delta_\epsilon u$ is continuous in the L^1 topology and $T(\delta_\epsilon u)$ goes to 0 with ϵ.

Let Y be any element of $\hat{L}(\mathbf{X})$; $\delta_\epsilon(Y)$ is the element obtained by multiplying each indeterminate X_i in Y by ϵ^{l_i}. The definition of $\delta_\epsilon u$ implies:

$$\mathrm{Log}\,\delta_\epsilon u = \delta_\epsilon\, \mathrm{Log}\, u\,.$$

$\delta_\epsilon Y^1 = \epsilon^{b_1} Y_1$ and $\delta_\epsilon(\sum \alpha_j W^j) = \sum \alpha_j \epsilon^{c_j} W^j$; therefore

$$\delta_\epsilon(Y^1 - \sum \alpha_j \epsilon^{(b_1 - c_j)} W^j)_{\mathbf{f}}$$

vanishes at $x_{\mathbf{f}}(\tau)$. By hypothesis there exists a neighborhood V of $0 \in S_N$ and a continuous map $u : V \to L^1$ such that

$$\mathrm{Log}_N\, u(\Phi) = Z + \Phi\,.$$

Set $\Theta(\epsilon) = -\sum \alpha_j \epsilon^{(b_1 - c_j)} W^j$; $\Theta(\epsilon)$ depends continuously from ϵ and since $(b_1 - c_j) < 0$, $\Theta(\epsilon) \in V$ if ϵ is sufficiently small. Therefore the control variation $\delta_\epsilon u(\Theta(\epsilon))$ proves the first assertion.

Let Φ satisfies the hypothesis; if σ and ϵ are sufficiently small $\sigma\Phi + \Theta(\epsilon) \in V$ and

$$\delta_\epsilon u(\Theta(\epsilon) + \sigma\Phi)$$

ia a control variation which has **f**-leading term equal to $\delta\Phi$. The second assertion is proved.

□

For the previous result to be applicable, we need to know how the N-good controls are made. The symmetries of the system give some information on this subject.

Let me recall some definitions introduced in [6] and in [4].

DEFINITION 8. *The* bad brackets *are the brackets in* Lie $\mathbf{X}$ *which contain* X_0 *an odd number of times and each* X_i *an even number of times. Let* $\mathcal{B}$ *be the set of bad brackets*

$$\mathcal{B} = \{\Lambda\ ,\ \ |\Lambda|_0 \textit{ is odd } \ |\Lambda|_i \textit{ is even } i = 1, \dots, m\}\,.$$

The set of the obstructions *is the set*

$$\mathcal{B}^* = \mathrm{Lie}\,(X_0, \mathcal{B}) \setminus \{aX_0\ ,\ a \in \mathbb{R}\}\,.$$

PROPOSITION 4. *For each integer* N *there exists a* N*-good element which belongs to* $\mathcal{B}^*$.

Proof. In [6] Sussmann has proved that there exists an element $\Psi \in \mathcal{B}$ and a C^1 map, $\overline{u}$, from a neighborhood of $0 \in S_N$ in L^1 such that the image of $\overline{u}$ is contained in the set of admissible controls and

$$\log_N \overline{u}(\Theta) = \Psi + \Theta\,.$$

This result is obtained by using the symmetries of the process. Standard arguments imply that it is possible to construct a C^1 map u from a neighborhood of $0 \in S_N$ to L^1 such that the image of u is contained in the set of admissible controls and

$$\mathrm{Log}_N\, u(\Theta) = \Xi + \Theta$$

with $\Xi \in \mathcal{B}^*$.

□

The previous proposition together with Theorem 1 imply that the trajectory variations are linked to the neutralization of the obstruction.

Theorem 1 can be used to find g-variations if the **f l**-neutralization holds on an interval containing τ.

4. Explicit optimality conditions for the single input case

In this section I am going to construct g-variations of a trajectory of an affine control process at any point of an interval in which the reference control is constantly equal to 0. It is known that if x^* is a solution of a sufficiently regular control process such that

$$g_0(x^*(t_1)) = \min_{y \in R(t_1, x^*(t_0)) \cap S} g_0(y)\,,$$

then there exists an adjoint variable $\lambda(t)$ satisfying the Pontrjagin Maximum Principle and such that for each τ and for each g-variation, v, of x^* at τ

$$\lambda(\tau)v \le 0\,.$$

Therefore the g-variations I will obtain, provide necessary conditions of optimality for the singular trajectory.

For simplicity sake I will limit myself to consider an affine single input control process

$$\dot{x} = f_0(x) + u f_1(x)$$

and I suppose that the control which generate the reference trajectory, x^*, which we want to test, is constantly equal to 0 on an interval I containing τ so that $x^*(t) = x_{\mathbf{f}}(t)$, $\forall t \in I$. The $\mathbf{f}\,\mathbf{l}$-neutralization of the obstructions on I provides g-variations at τ. In [4] it has been proved that if each bad bracket, Θ, $\|\Theta\|_{\mathbf{l}} \leq p$ is $\mathbf{f}\,\mathbf{l}$-neutralized on $x_{\mathbf{f}}(I)$, then all the obstructions whose $\mathbf{l}$-weight is not greater than p are $\mathbf{f}\,\mathbf{l}$-neutralized on $x_{\mathbf{f}}(I)$. Moreover if Φ is a bracket which is $\mathbf{f}\,\mathbf{l}$-neutralized on I, then $[X_0, \Phi]$ is $\mathbf{f}\,\mathbf{l}$-neutralized on I. Therefore to know which obstructions are $\mathbf{f}\,\mathbf{l}$-neutralized on I it is sufficient to test those bad brackets whose first element is equal to X_1. Let $\mathbf{l}$ be an admissible weight; $\mathbf{l}$ induces an increasing filtration in Lie $\mathbf{X}$

$$\{0\} = Y_{\mathbf{l}}^0 \subset Y_{\mathbf{l}}^1 \subset \cdots \subset Y_{\mathbf{l}}^n \subset \ldots$$

$Y^i = \text{span}\,\{\Phi : \|\Phi\|_{\mathbf{l}} \leq p_i\}$, $p_i < p_{i+1}$. Let p_j be such that each bad bracket whose weight is less than or equal to p_j is $\mathbf{f}\,\mathbf{l}$-neutralized on an interval containing τ. We know that $Y_{\mathbf{f}}^j(x_{\mathbf{f}}(\tau))$ is a subspace of g-variation at τ, which are obviously bilateral variations. Unilateral g-variation can be contained in the set of $\mathbf{l}$-homogeneous elements belonging to $Y_{\mathbf{f}}^{j+1}(x_{\mathbf{f}}(\tau))$. Notice that each subspace $Y_{\mathbf{f}}^i(x_{\mathbf{f}}(\tau))$ is finite dimensional and that the sequence $\{Y_{\mathbf{f}}^i(x_{\mathbf{f}}(\tau))\}$ become stationary for i sufficiently large. Therefore we are interested only in the elements of S_N with N sufficiently large. Let N be such that each $Y_{\mathbf{f}}^i(x_{\mathbf{f}}(\tau))$ is spanned by brackets whose length is less than N. The following Lemma proves that it is possible to modify the weight $\mathbf{l}$ in order to obtain a weight $\bar{\mathbf{l}}$ with the following properties:

1. each bracket which is $\mathbf{f}\,\mathbf{l}$-neutralized at τ is $\mathbf{f}\,\bar{\mathbf{l}}$-neutralized at τ;
2. the $\bar{\mathbf{l}}$-homogeneous elements are linear combination of brackets which contain the same numbers both of X_0 than of X_1.

LEMMA 1. *Let $\mathbf{l}$ be an admissible weight; for each integer N there exists an admissible weight $\bar{\mathbf{l}}$ with the following properties: if Φ and Θ are brackets whose length is not greater than N, then*

1. $\|\Phi\|_{\mathbf{l}} < \|\Theta\|_{\mathbf{l}}$ *implies* $\|\Phi\|_{\bar{\mathbf{l}}} < \|\Theta\|_{\bar{\mathbf{l}}}$
2. $\|\Phi\|_{\mathbf{l}} = \|\Theta\|_{\mathbf{l}}$ *and* $\|\Phi\|_1 < \|\Theta\|_1$, *implies* $\|\Phi\|_{\bar{\mathbf{l}}} < \|\Theta\|_{\bar{\mathbf{l}}}$
3. $\|\Phi\|_{\mathbf{l}} = \|\Theta\|_{\mathbf{l}}$, $\|\Phi\|_1 = \|\Theta\|_1$ *and* $\|\Phi\|_0 < \|\Theta\|_0$, *implies* $\|\Phi\|_{\bar{\mathbf{l}}} < \|\Theta\|_{\bar{\mathbf{l}}}$

Proof. The set of brackets whose weight is not greater than N is a finite set therefore if ϵ_0, ϵ_1 are positive numbers sufficiently small, then $\bar{\mathbf{l}} = \{l_0 - \epsilon_0,\ l_1 + \epsilon_1\}$ is an admissible weight for which the properties 1), 2) and 3) hold.

□

In order to simplify the notation, we will use the multiplicative notation for the brackets:

$$XY = [X, Y], \quad ZXY = [Z, XY], \quad \text{etc.}$$

It is known, [10], that each bracket, Φ, in Lie $\mathbf{X}$ is linear combination of brackets right normed, i.e. of the following type:

$$X_0^{i_1} X_1^{i_2} \dots X_1^{i_s}\,, \quad i_j \in \{0, 1, \dots\}\,,$$

which contains both X_0 than X_1 the same number of times of Φ; therefore it is sufficient to test the neutralization of right normed bad brackets.

Any N-good element, Z, of Lie $\mathbf{X}$ can be written as:

$$Z = \sum a_i \Phi_i$$

Φ_i right normed bracket; we name a_i *coefficient* of Φ_i in Z.

LEMMA 2. *Let $N > 2n + 3$. The coefficient of $X_1 X_0^{2n+1} X_1$ in any N-good element is positive if n is even and negative if n is odd; the coefficient of $X_1^{2n-1} X_0 X_1$ is always positive.*

Proof. Let Z be a N-good element and let us consider the control process

$$\begin{cases} \dot{x}_1 = u \\ \dot{x}_2 = x_1 \\ \dots \\ \dot{x}_{n+1} = x_n \\ \dot{x}_{n+2} = x_{n+1}^2\,. \end{cases}$$

Take as reference trajectory $x_{\mathbf{f}}(t) \equiv 0$. The reachable set $R(t, 0)$ is contained, for any positive t, in the half space $x_{n+1} \geq 0$ and hence $-\frac{\partial}{\partial x_{n+1}}$ cannot be a variation at any time. The only elements in Lie $\mathbf{f}$ which are different from 0 in 0 are:

$$\begin{aligned} (X_1)_{\mathbf{f}} &= \frac{\partial}{\partial x_1}\,, \\ (X_0^i X_1)_{\mathbf{f}} &= (-1)^i \frac{\partial}{\partial x_{i+1}}\,,\ i = 1, \dots, n \\ (X_1 X_0^{2n+1} X_1)_{\mathbf{f}} &= (-1)^n 2 \frac{\partial}{x_{n+1}}\,. \end{aligned}$$

If the coefficient of $X_1 X_0^{2n+1} X_1$ in Z were equal to 0, then 0 will be locally controllable [6], which is an absurd. Therefore it is different from 0; its sign has to be equal to $(-1)^n$ because otherwise $-\frac{\partial}{\partial x_{n+1}}$ would be a variation. The first assertion is proved.

The second assertion is proved by using similar arguments applied to the system:

$$\begin{cases} \dot{x}_1 = u \\ \dot{x}_2 = x_1^{2n} \end{cases}$$

□

Let us now compute explicitly some g-variations. We recall that

$$\mathrm{ad}_X Y = [X, Y]\,, \quad \mathrm{ad}_X^{n+1} Y = [X, \mathrm{ad}_X^n Y]\,.$$

THEOREM 2. *If there exists an admissible weight* $\mathbf{l}$ *such that each bad bracket of* $\mathbf{l}$*-weight less than* $(2n+1)l_0 + 2l_1$ *is* $\mathbf{f}\,\mathbf{l}$*-neutralized on an interval containing* τ*, then*

$$(-1)^n [X_1, \mathrm{ad}_{X_0}^{2n+1} X_1]_{\mathbf{f}}(x^*(\tau))$$

is a g-variation at τ*.*

Proof. We can suppose that the weight $\mathbf{l}$ has the properties 1), 2) and 3) as in Lemma 1. Therefore the brackets of $\mathbf{l}$-weight equal to $2l_1 + (2n+1)l_0$ contains 2 X_1 and (2n+1) X_0. The brackets which have as first element X_0 are the adjoint with respect to X_0 of brackets which by hypothesis are $\mathbf{f}\,\mathbf{l}$-neutralized on the interval I and therefore are $\mathbf{f}\,\mathbf{l}$-neutralized. Since the only bad bracket of $\mathbf{l}$-weight $(2n+1)l_0 + 2l_1$ which has as first element X_1, is $X_1 X_0^{2n+1} X_1$ the theorem is a consequence of Theorem 1 and of Lemma 2.

□

Notice that the theorem contains as particular case the well known generalized Legendre-Clebsch conditions. In fact it is possible to choose σ such that each bracket which is $\mathbf{f}$-neutralized on I with respect to the weight $(0, 1)$ is $\mathbf{f}$-neutralized with respect to the weight $\mathbf{l} = (\sigma, 1)$; moreover bearing in mind that only a finite set of brackets are to be considered, we can suppose that if two brackets contain a different number of X_1, then the one which contains less X_1 has less $\mathbf{l}$-weight and that two brackets have the same $\mathbf{l}$-weight if and only if they contain the same number both of X_0 than of X_1. Since each bad bracket contains at least two X_1, then the only bad bracket whose weight is less than $(2n+1)\sigma + 2$ contains two X_1 and $(2i+1)$ X_0, $i = 0, \ldots, (n-1)$; among these the only ones which we have to consider are $X_1 X_0^{2i+1} X_1$. Set

$$S^i = \mathrm{span}\,\{\Phi;\ \text{which contains } i \text{ times } X_1\}\,.$$

If

$$(X_1 X_0^{2i+1} X_1)_{\mathbf{f}}(x_{\mathbf{f}}(I)) \in S^1_{\mathbf{f}}(x_{\mathbf{f}}(I))\,, \quad i = (1, \ldots, (n-1))$$

Theorem 2 implies that $(-1)^n (X_1 X_0^{2n+1} X_1)_{\mathbf{f}}(x_{\mathbf{f}}(\tau))$ is a g-variation of $x_{\mathbf{f}}$ at τ; therefore if x^* is optimal, then the adjoint variable can be chosen in such a way that

$$(-1)^n \lambda(t)(X_1 X_0^{2i+1} X_1)_{\mathbf{f}}(x^*(\tau)) \le 0\,, \qquad t \in I$$

condition which is known as generalized Legendre-Clebsch condition.

The following example shows that by using Theorem 2 one can obtain further necessary conditions which can be added to the Legendre-Clebsch ones.

EXAMPLE 1. Let:

$$\begin{aligned} f_0 &= \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_3} + x_3 \frac{\partial}{\partial x_4} + \left(\frac{x_3^2}{2} - \frac{x_2^3}{6} \right) \frac{\partial}{\partial x_5} + \frac{x_4^2}{2} \frac{\partial}{\partial x_6} \\ f_1 &= \frac{\partial}{\partial x_2}\,. \end{aligned}$$

The generalized Legendre-Clebsch condition implies that $-(X_1 X_0^3 X_1)_{\mathbf{f}}(x_{\mathbf{f}}(\tau)) = \frac{\partial}{\partial x_5}$ is a g-variation at τ. Let us apply Theorem 2 with the weight $\mathbf{l} = (1, 1)$; the bad brackets of $\mathbf{l}$-weight less than 7 are:

$$X_1 X_0 X_1,\ X_1 X_0^3 X_1,\ X_1^3 X_0 X_1,\ X_1^2 X_0^2 X_1 X_0 X_1,\ X_1^2 X_0 X_1 X_0^2 X_1,$$

the only one different from 0 along the trajectory is $(X_1 X_0^3 X_1)_{\mathbf{f}}$ which is at $x_{\mathbf{f}}(I)$ a multiple $(X_1^2 X_0 X_1)_{\mathbf{f}}(x_{\mathbf{f}}(I))$. Therefore it is $\mathbf{f}\,\mathbf{l}$-neutralized. Theorem 1 implies that $\pm(X_1 X_0^3 X_1)_{\mathbf{f}}(x_{\mathbf{f}}(\tau)) = \pm\frac{\partial}{\partial x_5}$, and $(X_1 X_0^5 X_1)_{\mathbf{f}}(x_{\mathbf{f}}(\tau)) = \frac{\partial}{\partial x_6}$ are g-variations.

Another necessary optimality condition can be deduced from Theorem 1 and Lemma 2.

THEOREM 3. *If there exists an admissible weight* $\mathbf{l}$ *such that all the bad brackets whose* $\mathbf{l}$*-weight is less than* $l_0 + 2n\, l_1$ *are* $\mathbf{f}\,\mathbf{l}$*-neutralized on an interval containing* τ*, then*

$$(X_1^{2n-1} X_0 X_1)_{\mathbf{f}}(x_{\mathbf{f}}(\tau))$$

is a g-variation at τ*.*

Proof. We can suppose that the weight $\mathbf{l}$ is such that the brackets with the same $\mathbf{l}$-weight contain the same number both of X_1 than of X_0. Since there is only one bracket, $X_1^{2n-1} X_0 X_1$ which contain $2n$ X_1 and 1 X_0, the theorem is a consequence of Theorem 1 and of Lemma 2.

□

Notice that this condition is active also in the case in which the degree of singularity is $+\infty$.

References

[1] BIANCHINI R. M., *Variational Approach to some Optimization Control Problems*, in Geometry in Nonlinear Control and Differential Inclusions, Banach Center Publication **32**, Warszawa 1995, 83–94.

[2] BRESSAN A., *A high order test for optimality of bang-bang controls*, SIAM Journal on Control and Optimization **23** (1985), 38–48.

[3] KRENER A., *The High Order Maximal Principle and its Applications to Singular Extremals*, SIAM Journal on Control and Optimization **15** (1977), 256–292.

[4] BIANCHINI R. M., STEFANI G., *Controllability along a Reference Trajectory: a Variational Approach*, SIAM Journal on Control and Optimization **31** (1993), 900–927.

[5] BIANCHINI R. M., STEFANI G., *Graded Approximations and Controllability Along a Trajectory*, SIAM Journal on Control and Optimization **28** (1990), 903–924.

[6] SUSSMANN H., *Lie Brackets and Local Controllability: a Sufficient Condition for Single Input system*, SIAM Journal on Control and Optimization **21** (1983), 686–713.

[7] SUSSMANN H., *A General Theorem on Local Controllability*, SIAM Journal on Control and Optimization **25** (1987), 158–194.

[8] BIANCHINI R. M., *Good needle-like variations*, in Proceedings of Symposia in Pure Mathematics **64**.

[9] MIKHALEV A. A., ZOLOTYKH A. A., *Combinatorial Aspects of Lie Superalgebras*, CRC Press, Boca Raton 1995.

[10] POSTNIKOV M., *Lie Groups and Lie Algebras*, (English translation) MIR Publishers 1986.

AMS Subject Classification: 49B10, 49E15.

R. M. BIANCHINI
Department of Mathematics "U. Dini"
University of Florence
Viale Morgagni 67/A, 50134 Firenze, Italy
e-mail: `rosabian@udini.math.unifi.it`

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

U. Boscain – B. Piccoli

GEOMETRIC CONTROL APPROACH TO SYNTHESIS THEORY

1. Introduction

In this paper we describe the approach used in geometric control theory to deal with optimization problems. The concept of synthesis, extensively discussed in [20], appears to be the right mathematical object to describe a solution to general optimization problems for control systems.

Geometric control theory proposes a precise procedure to accomplish the difficult task of constructing an optimal synthesis. We illustrate the strength of the method and indicate the weaknesses that limit its range of applicability.

We choose a simple class of optimal control problems for which the theory provides a complete understanding of the corresponding optimal syntheses. This class includes various interesting controlled dynamics appearing in Lagrangian systems of mathematical physics. In this special case the structure of the optimal synthesis is completely described simply by a couple of integers, (cfr. Theorem 3). This obviously provides a very simple classification of optimal syntheses. A more general one, for generic plane control-affine systems, was developed in [18, 10].

First we give a definition of optimal control problem. We discuss the concepts of solution for this problem and compare them. Then we describe the geometric control approach and finally show its strength using examples.

2. Basic definitions

Consider an optimal control problem $(\mathcal{P})$ in Bolza form:

$$\dot{x} = f(x,u), \qquad x \in M,\ u \in U$$
$$\min\left(\int L(x,u)\,dt + \varphi(x_{term})\right)$$
$$x_{in} = x_0, \qquad x_{term} \in N \subset M$$

where M is a manifold, U is a set, $f : M \times U \to TM$, $L : M \times U \to \mathbb{R}$, $\varphi : M \to \mathbb{R}$, the minimization problem is taken over all admissible trajectory-control pairs (x,u), x_{in} is the initial point and x_{term} the terminal point of the trajectory $x(\cdot)$. A solution to the problem $(\mathcal{P})$ can be given by an open loop control $u : [0,T] \to U$ and a corresponding trajectory satisfying the boundary conditions.

One can try to solve the problem via a feedback control, that is finding a function $u : M \to U$ such that the corresponding ODE $\dot{x} = f(x,u(x))$ admits solutions and the solutions to the Cauchy problem with initial condition $x(0) = x_0$ solve the problem $(\mathcal{P})$. Indeed, one explicits the dependence of $(\mathcal{P})$ on x_0, considers the family of problems $\mathcal{P} = (\mathcal{P}(x_0))_{x_0 \in M}$ and tries to solve them via a unique function $u : M \to U$, that is to solve the family $\mathcal{P}$ of problems all together.

A well-known approach to the solution of $\mathcal{P}$ is also given by studying the value function, that is the function $V : M \to \mathbb{R}$ assuming at each x_0 the value of the minimum for the corresponding problem $\mathcal{P}(x_0)$, as solution of the Hamilton-Jacobi-Bellmann equation, see [5, 13]. In general V is only a weak solution to the HJB equation but can be characterized as the unique "viscosity solution" under suitable assumptions.

Finally, one can consider a family Γ of pairs trajectory-control $(\gamma_{x_0}, \eta_{x_0})$ such that each of them solves the corresponding problem $\mathcal{P}(x_0)$. This last concept of solution, called synthesis, is the one used in geometric control theory and has the following advantages with respect to the other concepts:

1) Generality

2) Solution description

3) Systematic approach

Let us now describe in details the three items.

1) Each feedback gives rise to at least one synthesis if there are solutions to the Cauchy problems. The converse is not true, that is a synthesis is not necessarily generated by a feedback even if in most examples one is able to reconstruct a piecewise smooth control.

If one is able to define the value function this means that each problem $\mathcal{P}(x_0)$ has a solution and hence there exists at least one admissible pair for each $\mathcal{P}(x_0)$. Obviously, in this case, the converse is true that is every optimal synthesis defines a value function. However, to have a viscosity solution to the HJB equation one has to impose extra conditions.

2) Optimal feedbacks usually lack of regularity and generate too many trajectories some of which can fail to be optimal. See [20] for an explicit example. Thus it is necessary to add some structure to feedbacks to describe a solution. This is exactly what was done in [11, 22].

Given a value function one knows the value of the minimum for each problem $\mathcal{P}(x_0)$. In applications this is not enough, indeed one usually needs to drive the system along the optimal trajectory and hence to reconstruct it from the value function. This is a hard problem, [5]. If one tries to use the concept of viscosity solutions then in the case of Lagrangians having zeroes the solution is not unique. Various interesting problems (see for example [28]) present Lagrangians with zeroes. Recent results to deal with this problem can be found in [16].

3) Geometric control theory proposes a systematic way towards the construction of optimal syntheses, while there are not general methods to construct feedbacks or viscosity solutions for the HJB equation. We describe in the next session this systematic approach.

3. Optimal synthesis

The approach to construct an optimal synthesis can be summarized in the following way:

$$\begin{array}{ll} a) & \text{MP + geometric techniques} \\ & \Downarrow \\ b) & \text{Properties of extremal trajectories} \\ & \Downarrow \\ c) & \text{Construction of extremal synthesis} \\ & \Downarrow \\ d) & \text{Optimal synthesis} \end{array}$$

We now explain each item of the picture for a complete understanding of the scheme.

a) The Maximum Principle remains the most powerful tool in the study of optimal control problems forty years after its first publication, see [21]. A big effort has been dedicated to generalizations of the MP in recent years, see [6], [23], and references therein.

Since late sixties the study of the Lie algebra naturally associated to the control system has proved to be an efficient tool, see [15]. The recent approach of simplectic geometry proposed by Agrachev and Gamkrelidze, see [1, 4], provides a deep insight of the geometric properties of extremal trajectories, that is trajectories satisfying the Maximum Principle.

b) Making use of the tools described in **a)** various results were obtained. One of the most famous is the well known Bang-Bang Principle. Some similar results were obtained in [8] for a special class of systems. For some planar systems every optimal trajectory is not bang-bang but still a finite concatenation of special arcs, see [19, 24, 25].

Using the theory of subanalytic sets Sussmann proved a very general results on the regularity for analytic systems, see [26]. The regularity, however, in this case is quite weak and does not permit to drive strong conclusions on optimal trajectories.

Big improvements were recently obtained in the study of Sub-Riemannian metrics, see [2, 3]. In particular it has been showed the link between subanalyticity of the Sub-Riemannian sphere and abnormal extremals.

c) Using the properties of extremal trajectories it is possible in some cases to construct an extremal synthesis. This construction is usually based on a finite dimensional reduction of the problem: from the analysis of **b)** one proves that all extremal trajectories are finite concatenations of special arcs. Again, for analytic systems, the theory of subanalytic sets was extensively used: [11, 12, 22, 27].

However, even simple optimization problems like the one proposed by Fuller in [14] may fail to admit such a kind of finite dimensional reduction. This phenomenon was extensively studied in [17, 28].

d) Finally, once an extremal synthesis has been constructed, it remains to prove its optimality. Notice that no regularity assumption property can ensure the optimality (not even local) of a single trajectory. But the contrary happens for a synthesis. The classical results of Boltianskii and Brunovsky, [7, 11, 12], were recently generalized to be applicable to a wider class of systems including Fuller's example (see [20]).

4. Applications to second order equations

Consider the control system:

$$\dot{x} = F(x) + uG(x),\ x \in \mathbb{R}^2,\ F, G \in \mathcal{C}^3(\mathbb{R}^2, \mathbb{R}^2),\ F(0) = 0,\ |u| \le 1 \tag{1}$$

and let $\mathcal{R}(\tau)$ be the reachable set within time τ from the origin. In the framework of [9, 19, 18], we are faced with the problem of reaching from the origin (under generic conditions on F and G) in minimum time every point of $\mathcal{R}(\tau)$. Given a trajectory $\gamma : [a, b] \to \mathbb{R}^2$, we define the time along γ as $T(\gamma) = b - a$.
A trajectory γ of (1) is *time optimal* if, for every trajectory γ' having the same initial and terminal points, one has $T(\gamma') \ge T(\gamma)$. A *synthesis* for the control system (1) at time τ is a family $\Gamma = \{(\gamma_x, u_x)\}_{x\in\mathcal{R}(\tau)}$ of trajectory-control pairs s.t.:

(a) for each $x \in \mathcal{R}(\tau)$ one has $\gamma_x : [0, b_x] \to \mathbb{R}^2$, $\gamma_x(0) = 0$, $\gamma_x(b_x) = x$;

(b) if $y = \gamma_x(t)$, where t is in the domain of γ_x, then $\gamma_y = \gamma_x|_{[0,t]}$.

A synthesis for the control system (1) is *time optimal* if, for each $x \in \mathcal{R}(\tau)$, one has $\gamma_x(T(x)) = x$, where T is the minimum time function $T(x) := \inf\{\tau :\ x \in \mathcal{R}(\tau)\}$. We indicate by Σ a control system of the type (1) and by $Opt(\Sigma)$ the set of optimal trajectories. If γ_1, γ_2 are two trajectories then $\gamma_1 * \gamma_2$ denotes their concatenation. For convenience, we define also the vector fields: $X = F - G$, $Y = F + G$. We say that γ is an X-trajectory and we write $\gamma \in \text{Traj}\,(X)$ if it corresponds to the constant control -1. Similarly we define Y-trajectories. If a trajectory γ is a concatenation of an X-trajectory and a Y-trajectory, then we say that γ is a $Y * X$-trajectory. The time t at which the two trajectories concatenate is called X-Y switching time and we say that the trajectory has a X-Y switching at time t. Similarly we define trajectories of type $X * Y$, $X * Y * X$, etc.
In [19] it was shown that, under generic conditions, the problem of reaching in minimum time every point of the reachable set for the system (1) admits a regular synthesis. Moreover it was shown that $\mathcal{R}(\tau)$ can be partitioned in a finite number of embedded submanifolds of dimension 2, 1 and 0 such that the optimal synthesis can be obtained from a feedback $u(x)$ satisfying:

- on the regions of dimension 2, we have $u(x) = \pm 1$,
- on the regions of dimension 1, called frame curves (in the following FC), $u(x) = \pm 1$ or $u(x) = \varphi(x)$ (where $\varphi(x)$ is a feedback control that depends on F, G and on their Lie bracket $[F, G]$, see [19]). The frame curves that correspond to the feedback φ are called *turnpikes*. A trajectory that corresponds to the control $u(t) = \varphi(\gamma(t))$ is called a Z-trajectory.

The submanifolds of dimension 0 are called frame points (in the following FP). In [18] it was provided a complete classification of all types of FP and FC.
Given a trajectory $\gamma \in \Gamma$ we denote by $n(\gamma)$ the smallest integer such that there exist $\gamma_i \in \text{Traj}\,(X) \cup \text{Traj}\,(Y) \cup \text{Traj}\,(Z)$, $(i = 1, \dots, n(\gamma))$, satisfying $\gamma = \gamma_{n(\gamma)} * \cdots * \gamma_1$.

The previous program can be used to classify the solutions of the following problem.

Problem: *Consider an autonomous ODE in* $\mathbb{R}$*:*

$$\ddot{y} = f(y, \dot{y}), \tag{2}$$

$$f \in \mathcal{C}^3(\mathbb{R}^2), \qquad f(0,0) = 0 \tag{3}$$

that describes the motion of a point under the action of a force that depends on the position and on the velocity (for instance due to a magnetic field or a viscous fluid). Then let apply an external force, that we suppose bounded (e.g. $|u| \leq 1$):

$$\ddot{y} = f(y, \dot{y}) + u\,. \tag{4}$$

We want to reach in minimum time a point in the configuration space (y_0, v_0) from the rest state $(0, 0)$.

First of all observe that if we set $x_1 = y$, $x_2 = \dot{y}$, (4) becomes:

$$\dot{x}_1 = x_2 \tag{5}$$

$$\dot{x}_2 = f(x_1, x_2) + u\,, \tag{6}$$

that can be written in our standard form $\dot{x} = F(x) + uG(x)$, $x \in \mathbb{R}^2$ by setting $x = (x_1, x_2)$, $F(x) = (x_2, f(x))$, $G(x) = (0, 1) := G$.
A deep study of those systems was performed in [9, 10, 19, 18]. From now on we make use of notations introduced in [19]. A key role is played by the functions Δ_A, Δ_B:

$$\Delta_A(x) = \det(F(x), G(x)) = x_2 \tag{7}$$

$$\Delta_B(x) = \det(G(x), [F(x), G(x)]) = 1\,. \tag{8}$$

From these it follows:

$$\Delta_A^{-1}(0) = \{x \in \mathbb{R}^2 : x_2 = 0\} \tag{9}$$

$$\Delta_B^{-1}(0) = \emptyset\,. \tag{10}$$

The analysis of [19] has to be completed in the following way.
Lemma 4.1 of [19] has to be replaced by the following (see [19] for the definition of $Bad(\tau)$ and $\tan_A$):

LEMMA 1. *Let $x \in Bad\,(\tau)$ and $G(x) \neq 0$ then:*

A. $x \in (\Delta_A^{-1}(0) \cap \Delta_B^{-1}(0)) \quad \Rightarrow \quad x \in \tan_A$*;*

B. $x \in \tan_A, \quad X(x), Y(x) \neq 0 \quad \Rightarrow \quad x \in (\Delta_A^{-1}(0) \cap \Delta_B^{-1}(0))$.

Proof. The proof of **A.** is exactly as in [19]. Let us prove **B**. Being $G(x) \neq 0$ we can choose a local system of coordinates such that $G \equiv (1, 0)$. Then, with the same computations of [19], we have:

$$\Delta_B(x) = -\partial_1 F_2(x)\,. \tag{11}$$

From $x \in \tan_A$ it follows $x \in \Delta_A^{-1}(0)$, hence $F(x) = \alpha G$ ($\alpha \in \mathbb{R}$). Assume that $X(x)$ is tangent to $\Delta_A^{-1}(0)$, being the other case entirely similar. This means that $\nabla\Delta_A(x) \cdot X(x) = (\alpha - 1)\nabla\Delta_A(x) \cdot G = 0$. From $X(x) \neq 0$ we have that $\alpha \neq 1$, hence $\nabla\Delta_A \cdot G = 0$. This implies $\partial_1 F_2 = 0$ and using (11) we obtain $\Delta_B(x) = 0$, i.e. $x \in \Delta_B^{-1}(0)$.

□

Now the proof of Theorem 4.2 of [19] is completed considering the following case:

(4) $\quad G(x) \neq 0, \qquad X(x) = 0$ or $Y(x) = 0\,.$

Note that (4) implies $x \in \tan_A$. In this case we assume the generic condition ($(P_1), \ldots, (P_8)$ were introduced in [19]):

$$(P_9) \qquad \Delta_B(x) \neq 0\,.$$

Suppose $X(x) = 0$ and $Y(x) \neq 0$. The opposite case is similar. Choose a new local system of coordinates such that x is the origin, $Y = (0, -1)$ and $\Delta_A^{-1}(0) = \{(x_1, x_2) : x_2 = 0\}$.
Take $U = B(0, r)$, the ball of radius r centered at 0, and choose r small enough such that:

- 0 is the only bad point in U;
- $\Delta_B(x) \neq 0$ for every $x \in U$;
- for every $x \in U$ we have:

$$|X(x)| \ll 1\,. \tag{12}$$

Let $U_1 = U \cap \{(x_1, x_2) : x_2 > 0\}$, $U_2 = U \cap \{(x_1, x_2) : x_2 < 0\}$. We want to prove the following:

THEOREM 1. *If $\gamma \in Opt\,(\Sigma)$ and $\{\gamma(t) : t \in [b_0, b_1]\} \subset U$ then we have a bound on the number of arcs, that is $\exists N_x \in \mathbb{N}$ s.t. $n(\gamma|_{[b_0,b_1]}) \leq N_x$.*

In order to prove Theorem 1 we will use the following Lemmas.

LEMMA 2. *Let $\gamma \in Opt\,(\Sigma)$ and assume that γ has a switching at time $t_1 \in \mathrm{Dom}\,(\gamma)$ and that $\Delta_A(\gamma(t_1)) = 0$. Then $\Delta_A(\gamma(t_2)) = 0$, $t_2 \in \mathrm{Dom}\,(\gamma)$, iff t_2 is a switching time for γ.*

Proof. The proof is contained in [10].

□

LEMMA 3. *Let $\gamma : [a, b] \to U$ be an optimal trajectory such that $\gamma([a, b]) \subset U_1$ or $\gamma([a, b]) \subset U_2$, then $n(\gamma) \leq 2$.*

Proof. It is a consequence of Lemma 3.5 of [19] and of the fact that every point of U_1 (respectively U_2) is an ordinary point i.e. $\Delta_A(x) \cdot \Delta_B(x) \neq 0$.

□

LEMMA 4. *Consider $\gamma \in Opt\,(\Sigma)$, $\{\gamma(t) : t \in [b_0, b_1]\} \subset U$. Assume that there exist a X-Y switching time $\bar{t} \in (b_0, b_1)$ for γ and $\gamma(\bar{t}) \in U_1$. Then $\gamma|_{[\bar{t},b_1]}$ is a Y-trajectory.*

Proof. Assume by contradiction that γ switches at time $t' \in (b_0, b_1)$, $t' > \bar{t}$. If $\gamma(t') \in U_1$ then this contradicts the conclusion of Lemma 3. If $\gamma(t') \in \Delta_A^{-1}(0)$ then this contradicts Lemma 2. Assume $\gamma(t') \in U_2$. From $\mathrm{sgn}\,\Delta_A(\gamma(\bar{t})) = -\mathrm{sgn}\,\Delta_A(\gamma(t'))$ we have that $\frac{1}{2}X(\gamma(\bar{t})) \wedge Y = -\frac{1}{2}(X(\gamma(t')) \wedge Y)$. This means that:

$$\mathrm{sgn}\,(X_2(\gamma(\bar{t})) = -\mathrm{sgn}\,(X_2(\gamma(t'))\,, \tag{13}$$

where X_2 is the second component of X. Choose $t_0 \in (b_0, \bar{t})$ and define the trajectory $\bar{\gamma}$ satisfying $\bar{\gamma}(b_0) = \gamma(b_0)$ and corresponding to the control $\bar{u}(t) = -1$ for $t \in [b_0, t_0]$ and $\bar{u}(t) = 1$ for $t \in [t_0, b_1]$. From (13) there exists $t_1 > t_0$ s.t. $\bar{\gamma}(t_1) = \gamma(t) \in U_2$. Using (12) it is easy to prove that $t_1 < t$. This contradicts the optimality of γ (see fig. 1).

□

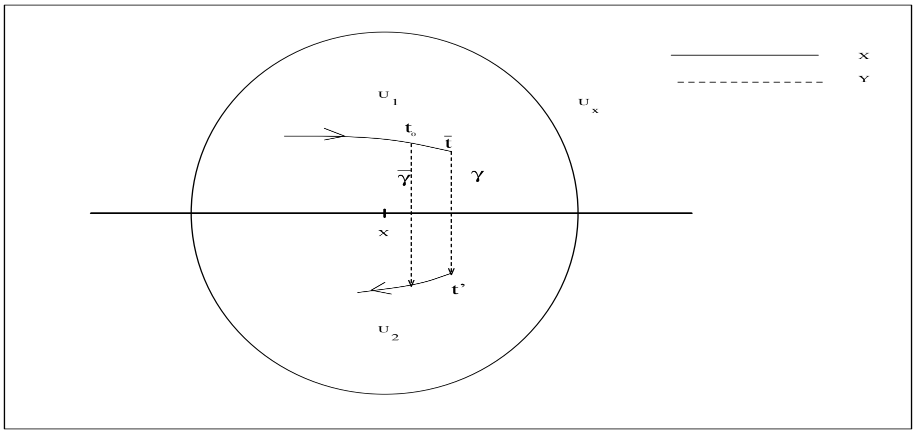

Figure 1:

of Theorem 1. For sake of simplicity we will write γ instead of $\gamma|_{[b_0,b_1]}$.

Assume first that no switching happens on $\Delta_A^{-1}(0)$. We have the following cases (see fig. 2 for some of these):

(A) γ has no switching; $n(\gamma) = 1$;

(B) for some $\varepsilon > 0$, $\gamma|_{[b_0,b_0+\varepsilon]}$ is an X-trajectory, $\gamma(b_0) \in U_1$, $n(\gamma) > 1$;

- **(B1)** if γ switches to Y in U_1, by Lemma 4, $n(\gamma) = 2$;
- **(B2)** if γ crosses $\Delta_A^{-1}(0)$ and switches to Y in U_2, by Lemma 3, γ does not switch anymore. Hence $n(\gamma) = 2$;

(C) for some $\varepsilon > 0$, $\gamma|_{[b_0,b_0+\varepsilon]}$ is an X-trajectory, $\gamma(b_0) \in U_2$, $n(\gamma) > 1$;

- **(C1)** if γ switches to Y before crossing $\Delta_A^{-1}(0)$ then, by Lemma 3, $n(\gamma) = 2$;
- **(C2)** if γ reaches U_1 without switching, then we are in the **(A)** or **(B)** case, thus $n(\gamma) \leq 2$;

(D) for some $\varepsilon > 0$, $\gamma|_{[b_0,b_0+\varepsilon]}$ is a Y-trajectory, $\gamma(b_0) \in U_1$, $n(\gamma) > 1$;

- **(D1)** if γ switches to X in U_1 and never crosses $\Delta_A^{-1}(0)$ then by Lemma 3 $n(\gamma) = 2$;
- **(D2)** if γ switches to X in U_1 (at time $t_0 \in [b_0, b_1]$) and then it crosses $\Delta_A^{-1}(0)$, then $\gamma|_{[t_0,b_1]}$ satisfies the assumptions of **(A)** or **(C)**. Hence $n(\gamma) \leq 3$;
- **(D3)** if γ switches to X in U_2 at $t_0 \in [b_0, b_1]$ and then it does not cross $\Delta_A^{-1}(0)$, we have $n(\gamma) = 2$;
- **(D4)** if γ switches to X in U_2 and then it crosses $\Delta_A^{-1}(0)$ we are in cases **(A)** or **(B)** and $n(\gamma) \leq 3$;

(E) for some $\varepsilon > 0$, $\gamma|_{[b_0,b_0+\varepsilon]}$ is a Y-trajectory, $\gamma(b_0) \in U_2$, $n(\gamma) > 1$.

- **(E1)** if γ switches to X in U_2 and it does not cross $\Delta_A^{-1}(0)$ then $n(\gamma) = 2$;

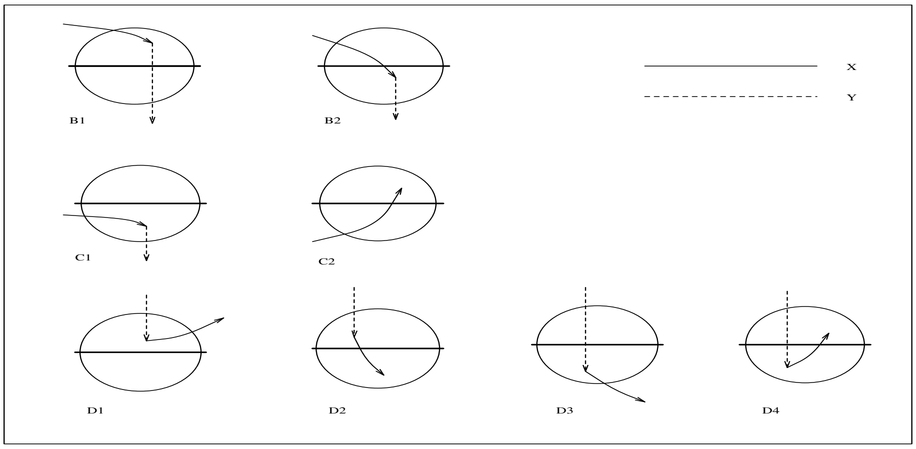

Figure 2:

(E2) if γ switches to X in U_2 and then it crosses $\Delta_A^{-1}(0)$, we are in case **(A)** or **B**. Hence $n(\gamma) \leq 3$.

If γ switches at $\Delta_A^{-1}(0)$, by Lemma 2 all the others switchings of γ happen on the set $\Delta_A^{-1}(0)$. Moreover, if γ switches to Y it has no more switchings. Hence $n(\gamma) \leq 3$.

The Theorem is proved with $N_x = 3$.

□

By direct computations it is easy to see that the *generic* conditions $P_1, \ldots, P_9$, under which the construction of [19] holds, are satisfied under the condition:

$$f(x_1, 0) = \pm 1 \quad \Rightarrow \quad \partial_1 f(x_1, 0) \neq 0 \tag{14}$$

that obviously implies $f(x_1, 0) = 1$ or $f(x_1, 0) = -1$ only in a finite number of points.
In the framework of [9, 19, 18] we will prove that, for our problem (5), (6), with the condition (14), the "shape" of the optimal synthesis is that shown in fig. 3. In particular the partition of the reachable set is described by the following

THEOREM 2. *The optimal synthesis of the control problem (5) (6) with the condition (14), satisfies the following:*

1. *there are no turnpikes;*

2. *the trajectory $\gamma^\pm$ (starting from the origin and corresponding to constant control ± 1) exits the origin with tangent vector $(0, \pm 1)$ and, for an interval of time of positive measure, lies in the set $\{(x_1, x_2) : x_1, x_2 \geq 0\}$ respectively $\{(x_1, x_2) : x_1, x_2 \leq 0\}$;*

3. *$\gamma^\pm$ is optimal up to the first intersection (if it exists) with the x_1-axis. At the point in which γ^+ intersects the x_1-axis it generates a switching curve that lies in the half plane $\{(x_1, x_2) : x_2 \geq 0\}$ and ends at the next intersection with the x_1-axis (if it exists). At that*

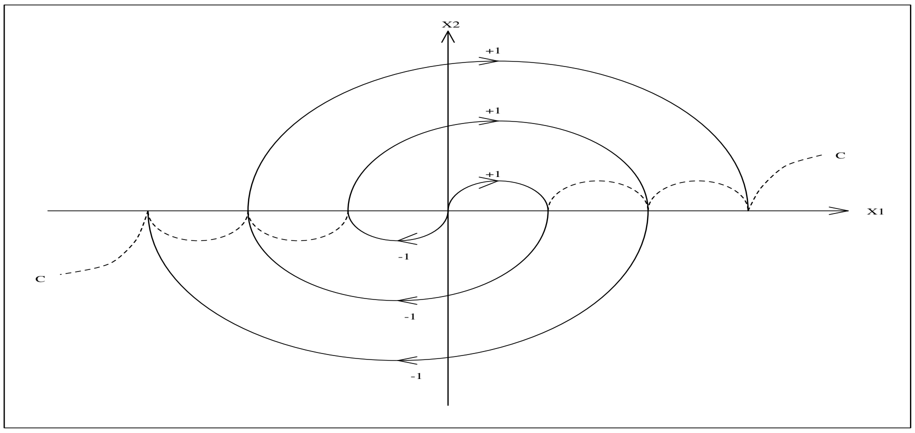

Figure 3: The shape of the optimal synthesis for our problem.

point another switching curve generates. The same happens for γ^- and the half plane $\{(x_1, x_2) : x_2 \leq 0\}$;

4. *let y_i, for $i = 1, \ldots, n$ (n possibly $+\infty$) (respectively z_i, for $i = 1, \ldots, m$ (m possibly $+\infty$)) be the set of boundary points of the switching curves contained in the half plane $\{(x_1, x_2) : x_2 \geq 0\}$ (respectively $\{(x_1, x_2) : x_2 \leq 0\}$) ordered by increasing (resp. decreasing) first components. Under generic assumptions, y_i and z_i do not accumulate. Moreover:*

- *For $i = 2, \ldots, n$, the trajectory corresponding to constant control $+1$ ending at y_i starts at z_{i-1};*
- *For $i = 2, \ldots, m$, the trajectory corresponding to constant control -1 ending at z_i starts at y_{i-1}.*

REMARK 1. The union of $\gamma^\pm$ with the switching curves is a one dimensional $\mathcal{C}^0$ manifold M. Above this manifold the optimal control is $+1$ and below is -1.

REMARK 2. The optimal trajectories turn clockwise around the origin and switch along the switching part of M. They stop turning after the last y_i or z_i and tend to infinity with $x_1(t)$ monotone after the last switching.

From **4.** of Theorem 2 it follows immediately the following:

THEOREM 3. *To every optimal synthesis for a control problem of the type (5) (6) with the condition (14), it is possible to associate a couple $(n, m) \in (\mathbb{N} \cup \infty)^2$ such that one of the following cases occurs:*

A. *$n = m$, n finite;*

B. *$n = m + 1$, n finite;*

C. *$n = m - 1$, n finite;*

D. *$n = \infty$, $m = \infty$.*

Moreover, if Γ_1, Γ_2 are two optimal syntheses for two problems of kind (5), (6), (14), and (n_1, m_1) (resp. (n_2, m_2)) are the corresponding couples, then Γ_1 is equivalent to Γ_2 iff $n_1 = n_2$ and $m_1 = m_2$.

REMARK 3. In Theorem 3 the equivalence between optimal syntheses is the one defined in [9].

of Theorem 3. Let us consider the synthesis constructed by the algorithm described in [9]. The stability assumptions (SA1), ... ,(SA6) holds. The optimality follows from Theorem 3.1 of [9].

1. By definition a turnpike is a subset of $\Delta_B^{-1}(0)$. From (8) it follows the conclusion.

2. We leave the proof to the reader.

3. Let $\gamma_2^{\pm}(t) = \pi_2(\gamma^{\pm}(t))$, where $\pi_2 : \mathbb{R}^2 \to \mathbb{R}$, $\pi_2(x_1, x_2) = x_2$, and consider the adjoint vector field $v : \mathbb{R}^2 \times \text{Dom}\,(\gamma^{\pm}) \times \text{Dom}\,(\gamma^{\pm}) \to \mathbb{R}^2$ associated to $\gamma^{\pm}$ that is the solution of the Cauchy problem:

$$(15) \qquad \begin{cases} \dot{v}(v_0, t_0; t) &= (\nabla F \pm \nabla G)(\gamma^{\pm}(t)) \cdot v(v_0, t_0; t) \\ v(v_0, t_0; t_0) &= v_0\,, \end{cases}$$

We have the following:

LEMMA 5. *Consider the $\gamma^{\pm}$ trajectories for the control problem (5), (6). We have that $v(G, t; 0)$ and G are parallel iff $\Delta_A(\gamma^{\pm}(t)) = 0$ (i.e. $\gamma_2^{\pm}(t) = 0$).*

Proof. Consider the curve γ^{+}, the case of γ^{-} being similar. From (9) we know that $\Delta_A(\gamma^{+}(t)) = 0$ iff $\gamma_2^{+}(t) = 0$. First assume $\Delta_A(\gamma^{+}(t)) = 0$. We have that G and $(F + G)(\gamma^{+}(t))$ are collinear that is $G = \alpha(F + G)(\gamma^{+}(t))$ with $\alpha \in \mathbb{R}$. For fixed t_0, t the map:

$$(16) \qquad f_{t_0,t} : v_0 \mapsto v(v_0, t_0; t)$$

is clearly linear and injective, then using (15) and $\dot{\gamma}^{+}(t) = (F + G)(\gamma^{+}(t))$, we obtain $v(G, t; 0) = \alpha v\left((F + G)(\gamma^{+}(t)), t; 0\right) = \alpha(F + G)(0) = \alpha G$.
Viceversa assume $v(G, t; 0) = \alpha G$, then (as above) we obtain $v(G, t; 0) = \alpha v((F + G)(\gamma^{+}(t)), t; 0)$. From the linearity and the injectivity of (16) we have $G = \alpha(F + G)(\gamma^{+}(t))$ hence $\Delta_A(\gamma^{+}(t)) = 0$.

□

LEMMA 6. *Consider the trajectory γ^{+} for the control problem (5), (6). Let $\bar{t} > 0$ (possibly $+\infty$) be the first time such that $\gamma_2^{+}(\bar{t}) = 0$. Then γ^{+} is extremal exactly up to time $\bar{t}$. And similarly for γ^{-}.*

Proof. In [19] it was defined the function $\theta(t) = \arg\left(G(0), v\left(G(\gamma^{+}(t)), t, 0\right)\right)$. This function has the following properties:

(i) $\text{sgn}\,(\dot{\theta}(t)) = \text{sgn}(\Delta_B(\gamma(t))$, that was proved in Lemma 3.4 of [19]. From (8) we have that $\text{sgn}\,(\dot{\theta}(t)) = 1$ so $\theta(t)$ is strictly increasing;

(ii) γ^+ is extremal exactly up to the time in which the measure of the range of θ is π i.e. up to the time:

(17) $$t^+ = \min\{t \in [0, \infty] : |\theta(s_1) - \theta(s_2)| = \pi, \text{ for some } s_1, s_2 \in [0, t^+]\},$$

under the hypothesis $\dot\theta(t^+) \neq 0$. This was proved in Proposition 3.1 of [9].

From Lemma 5 we have that $\Delta_A(\gamma^+(t)) = 0$ iff there exists $n \in \mathbb{N}$ satisfying:

(18) $$\theta(G, v(G, t, 0)) = n\pi .$$

In particular (18) holds for $t = \bar t$ and some n. From the fact that $\bar t$ is the first time in which $\gamma_2^+(\bar t) = 0$ and hence the first time in which $\Delta_A(\gamma^+(\bar t)) = 0$, we have that $n = 1$.

From $\theta(\bar t) = \pi$ and sgn $(\dot\theta(\bar t)) = 1$ the Theorem is proved with $t^+ = \bar t$.

□

From Lemma 6, $\gamma^\pm$ are extremal up to the first intersection with the x_1-axis.

Let $\underline{t}$ be the time such that $\gamma^-(\underline{t}) = z_1$, defined in **4** of Theorem 2. The extremal trajectories that switch along the C-curve starting at y_1 (if it exists), are the trajectories that start from the origin with control -1 and then, at some time $t' < \underline{t}$, switch to control $+1$. Since the first switching occurs in the orthant $\{(x_1, x_2) : x_1, x_2 < 0\}$, by a similar argument to the one of Lemma 6, the second switch has to occur in the half space $\{(x_1, x_2) : x_2 > 0\}$, because otherwise between the two switches we have meas (range $(\theta(t))) > \pi$. This proves that the switching curves never cross the x_1-axis.

4. The two assertions can be proved separately. Let us demonstrate only the first, being the proof of the second similar. Define $y_0 = z_0 = (0, 0)$. By definition the $+1$ trajectory starting at z_0 reaches y_1. By Lemma 2 we know that if an extremal trajectory has a switching at a point of the x_1-axis, then it switches iff it intersects the x_1-axis again. This means that the extremal trajectory that switches at y_i has a switching at z_j for some j. By induction one has $j \geq i - 1$. Let us prove that $j = i - 1$. By contradiction assume that $j > i - 1$, then there exists an extremal trajectory switching at z_{i-1} that switches on the C curve with boundary points y_{i-1}, y_i. This is forbidden by Lemma 2.

□

EXAMPLES 1. In the following we will show the qualitative shape of the synthesis of some physical systems coupled with a control. More precisely we want to determine the value of the couple (m, n) of Theorem 3.

Duffin Equation

The Duffin equation is given by the formula $\ddot y = -y - \varepsilon(y^3 + 2\mu\dot y)$, $\varepsilon, \mu > 0$, ε small. By introducing a control term and transforming the second order equation in a first order system, we have:

$$\begin{aligned} \dot x_1 &= x_2 && (19)\\ \dot x_2 &= -x_1 - \varepsilon(x_1^3 + 2\mu x_2) + u . && (20)\end{aligned}$$

From this form it is clear that $f(x) = -x_1 - \varepsilon(x_1^3 + 2\mu x_2)$.
Consider the trajectory γ^+. It starts with tangent vector $(0, 1)$, then, from (19), we see that it

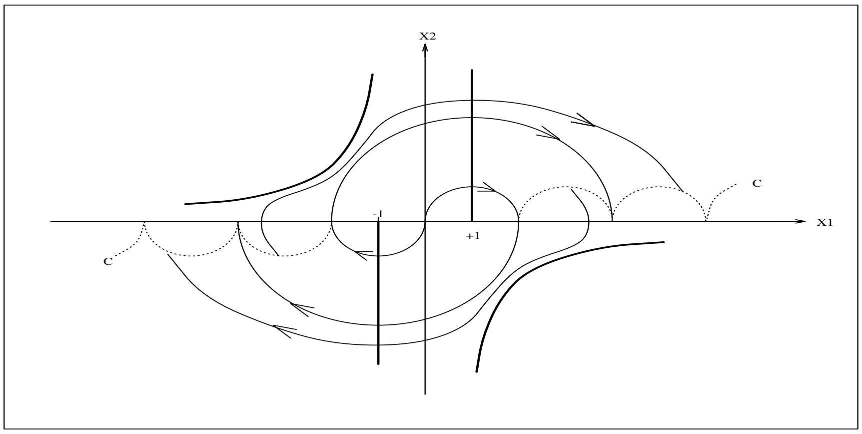

Figure 4: The synthesis for the Van Der Pol equation.

moves in the orthant $\Omega := \{(x_1, x_2),\ x_1, x_2 > 0\}$. To know the shape of the synthesis we need to know where $(F+G)_2(x) = 0$. If we set $a = \frac{1}{2\varepsilon\mu}$, this happens where

$$x_2 = a(1 - x_1 - \varepsilon x^3)\,. \tag{21}$$

From (19) and (20) we see that, after meeting this curve, the trajectory moves with $\dot{\gamma}_1^+ > 0$ and $\dot{\gamma}_2^+ < 0$. Then it meets the x_1-axis because otherwise if $\gamma^+(t) \in \Omega$ we necessarily have (for $t \to \infty$) $\gamma_1^+ \to \infty$, $\dot{\gamma}_2^+ \to 0$, that is not permitted by (20). The behavior of the trajectory γ^- is similar.

In this case, the numbers (n, m) are clearly (∞, ∞) because the $+1$ trajectory that starts at z_1 meets the curve (21) exactly one time and behaves like γ^+. So the C-curve that starts at y_1 meets again the x_1 axis. The same happens for the -1 curve that starts at y_1. In this way an infinite sequence of y_i and z_i is generated.

Van der Pol equation

The Van der Pol equation is given by the formula $\ddot{y} = -y + \epsilon(1-y^2)\dot{y} + u$, $\varepsilon > 0$ and small. The associated control system is: $\dot{x}_1 = x_2$, $\dot{x}_2 = -x_1 + \varepsilon(1 - x_1^2)x_2 + u$. We have $(F+G)_2(x) = 0$ on the curves $x_2 = -\frac{1}{\varepsilon(x_1+1)}$ for $x_1 \neq \pm 1$, $x_1 = -1$. After meeting these curves, the γ^+ trajectory moves with $\dot{\gamma}_1^+ > 0$ and $\dot{\gamma}_2^+ < 0$ and, for the same reason as before, meets the x_1-axis. Similarly for γ^-. As in the Duffin equation, we have that m and n are equal to $+\infty$. But here, starting from the origin, we cannot reach the regions: $\left\{(x_1, x_2) : x_1 < -1, x_2 \geq -\frac{1}{\varepsilon(x_1+1)}\right\}$, $\left\{(x_1, x_2) : x_1 > -1, x_2 \leq -\frac{1}{\varepsilon(x_1-1)}\right\}$ (see fig. 4).

Another example

In the following we will study an equation whose synthesis has $n, m < \infty$. Consider the equa-

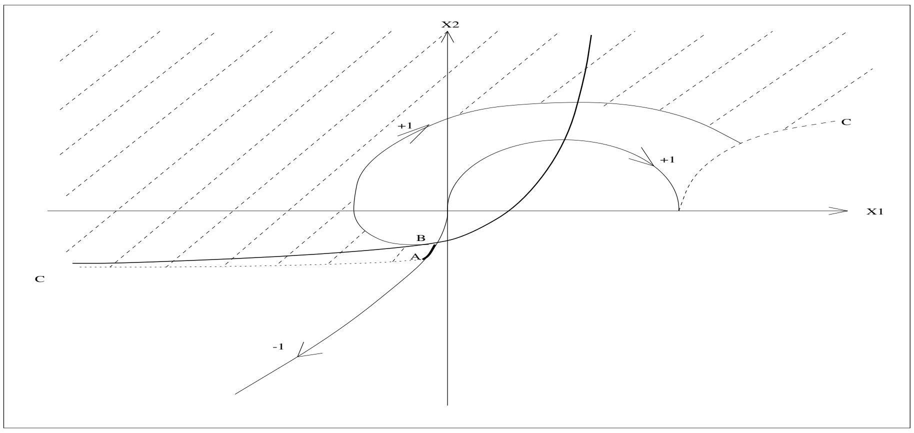

Figure 5: The synthesis for the control problem (22), (23). The sketched region is reached by curves that start from the origin with control -1 and then switch to $+1$ control between the points A and B.

tion: $\ddot{y} = -e^y + \dot{y} + 1$. The associated control system is:

$$\dot{x}_1 = x_2 \tag{22}$$

$$\dot{x}_2 = -e^{x_1} + x_2 + 1 + u \tag{23}$$

We have $\dot{\gamma}_2^+ = 0$ on the curve $x_2 = e^{x_1} - 2$. After meeting this curve, the γ^+ trajectory meets the x_1-axis.

Now the synthesis has a different shape because the trajectories corresponding to control -1 satisfy $\dot{\gamma}_2 = 0$ on the curve:

$$x_2 = e^{x_1} \tag{24}$$

that is contained in the half plane $\{(x_1, x_2) : x_2 > 0\}$. Hence γ^- never meets the curve given by (24) and this means that $m = 0$. Since we know that n is at least 1, by Remark 4, we have $n = 1$, $m = 0$. The synthesis is drawn in fig. 5.

5. Optimal syntheses for Bolza Problems

Quite easily we can adapt the previous program to obtain information about the optimal syntheses associated (in the previous sense) to second order differential equations, but for more general minimizing problems.

We have the well known:

LEMMA 7. *Consider the control system:*

$$\dot{x} = F(x) + uG(x), \quad x \in \mathbb{R}^2, \quad F, G \in \mathcal{C}^3(\mathbb{R}^2, \mathbb{R}^2), \quad F(0) = 0, \quad |u| \leq 1\,. \tag{25}$$

Let $L : \mathbb{R}^2 \to \mathbb{R}$ be a $\mathcal{C}^3$ bounded function such that there exists $\delta > 0$ satisfying $L(x) > \delta$ for any $x \in \mathbb{R}^2$.
Then, for every $x_0 \in \mathbb{R}^2$, the problem: $\min \int_0^\tau L(x(t))\,dt$ *s.t.* $x(0) = 0$, $x(\tau) = x_0$, *is equivalent to the* minimum time problem *(with the same boundary conditions) for the control system* $\dot{x} = F(x)/L(x) + uG(x)/L(x)$.

By this lemma it is clear that if we have a second order differential equation with a bounded-external force $\ddot{y} = f(y, \dot{y}) + u$, $f \in \mathcal{C}^3(\mathbb{R}^2)$, $f(0,0) = 0$, $|u| \leq 1$, then the problem of reaching a point in the configuration space (y_0, v_0) from the origin, minimizing $\int_0^\tau L(y(t), \dot{y}(t))\,dt$, (under the hypotheses of Lemma 7) is equivalent to the minimum time problem for the system: $\dot{x}_1 = x_2/L(x)$, $\dot{x}_2 = f(x)/L(x) + 1/L(x)u$. By setting: $\alpha : \mathbb{R}^2 \to]0, 1/\delta[$, $\alpha(x) := 1/L(x)$, $\beta : \mathbb{R}^2 \to \mathbb{R}$, $\beta(x) := f(x)/L(x)$, we have: $F(x) = (x_2\alpha(x), \beta(x))$, $G(x) = (0, \alpha(x))$. From these it follows: $\Delta_A(x) = x_2\alpha^2$, $\Delta_B(x) = \alpha^2(\alpha + x_2\partial_2\alpha)$.
The equations defining turnpikes are: $\Delta_A \neq 0$, $\Delta_B = 0$, that with our expressions become the differential condition $\alpha + x_2\partial_2\alpha = 0$ that in terms of L is:

$$L(x) - x_2\partial_2 L(x) = 0 \tag{26}$$

REMARK 4. Since $L > 0$ it follows that the turnpikes never intersect the x_1-axis. Since (26) depends on $L(x)$ and not on the control system, all the properties of the turnpikes depend only on the Lagrangian.

Now we consider some particular cases of Lagrangians.

L=$L(y)$ In this case the Lagrangian depends only on the position y and not on the velocity $\dot{y}$ (i.e. $L = L(x_1)$). (26) is never satisfied so there are no turnpikes.

L=$L(\dot{y})$ In this case the Lagrangian depends only on velocity and the turnpikes are horizontal lines.

L=$V(y) + \frac{1}{2}\dot{y}^2$ In this case we want to minimize an energy with a kinetic part $\frac{1}{2}\dot{y}^2$ and a positive potential depending only on the position and satisfying $V(y) > 0$. The equation for turnpikes is $(x_2)^2 = 2V(x_1)$.

References

[1] AGRACHEV A., *Methods of Control Theory*, in "Mathematic Geometry", Proc. ICM 94, Dekker, New York 1998.

[2] AGRACHEV A., BONNARD B., CHYBA M., KUPKA I., *Sub Riemannian Sphere in Martinet Flat Case*, ESAIM: COCV **2** (1997), 337–448.

[3] AGRACHEV A., EL ALAOUI C., GAUTHIER J. P., *Sub Riemannian Metrics on* $\mathbb{R}^3$, to appear in Geometric Control and Non-Holonomic Problems in Mechanics, Conference Proceeding Series, Canad. Math Soc.

[4] AGRACHEV A., GAMKRELIDZE R., *Simplectic Methods for Optimization and Control*, in "Geometry of Feedback and Optimal Control", edited by B. Jakubczyk and W. Respondek, Dekker, New York 1998.

[5] BARDI M., CAPUZZO-DOLCETTA I., *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equation*, Birkauser Boston, Inc., Boston MA 1997.

[6] BIANCHINI R. M., STEFANI G., *Graded Approximations and Controllability along a Trajectory*, SIAM J. Control and Optimization **28** (1990), 903–924.

[7] BOLTIANSKII, *Sufficient Conditions for Optimality and the Justification of the Dynamic Programming Principle*, SIAM J. Control and Optimization **4** (1966), 326–361.

[8] BRESSAN A., PICCOLI B., *A Baire Category Approach to the Bang-Bang Property*, Journal of Differential Equations **116** (2) (1995), 318–337.

[9] BRESSAN A., PICCOLI B., *Structural Stability for Time-Optimal Planar Syntheses*, Dynamics of Continuous, Discrete and Impulsive Systems **3** (1997), 335-371.

[10] BRESSAN A., PICCOLI B., *A Generic Classification of Time Optimal Planar Stabilizing Feedbacks*, SIAM J. Control and Optimization **36** (1) (1998), 12–32.

[11] BRUNOVSKI P., *Every Normal Linear System Has a Regular Time-Optimal Synthesis*, Math. Slovaca **28** (1978), 81–100.

[12] BRUNOVSKY P., *Existence of Regular Syntheses for General Problems*, J. Diff. Equations **38** (1980), 317–343.

[13] FLEMING W. H., SONER M., *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York 1993.

[14] FULLER A. T., *Relay Control Systems Optimized for Various Performance Criteria*, Automatic and remote control, Proc. first world congress IFAC Moscow **1**, Butterworths 1961, 510–519.

[15] JURDJEVIC V., *Geometric Control Theory*, Cambridge University Press 1997.

[16] MALISOFF M., *On The Bellmann Equation for Control Problem with Exit Time and Unbounded Cost Functionals*, to appear on Proc. of 39th CDC of IEEE.

[17] MARCHAL C., *Chattering arcs and chattering controls*, J. Optim Theory Appl. **11** (1973), 441–468.

[18] PICCOLI B., *Classifications of Generic Singularities for the Planar Time-Optimal Synthesis*, SIAM J. Control and Optimization **34** (6), December 1996, 1914–1946.

[19] PICCOLI B., *Regular Time-Optimal Syntheses for Smooth Planar Systems*, Rend. Sem Mat. Univ. Padova **95** (1996), 59–79.

[20] PICCOLI B., SUSSMANN H. J., *Regular Synthesis and Sufficiency Conditions for Optimality*, to appear on SIAM J. Control and Optimization.

[21] PONTRYAGIN L. S., BOLTIANSKI V., GAMKRELIDZE R., MICHTCHENKO E., *The Mathematical Theory of Optimal Processes*, John Wiley and Sons, Inc, 1961.

[22] SUSSMANN H. J., *Subanalytic sets and feedback control*, J. Diff. Equations **31** (1) (1979), 31–52.

[23] SUSSMANN H. J., *Geometry and Optimal Control* in “Mathematical Control Theory”, Springer-Verlag, New York 1998, 140–198.

[24] SUSSMANN H. J., *The Structure of Time-Optimal Trajectories for Single-Input Systems in the Plane: the C^∞ Nonsingular Case*, SIAM J. Control Optim. **25** (2) (1987), 433–465.

[25] SUSSMANN H. J., *The Structure of Time-Optimal Trajectories for Single-Input Systems in the Plane: the General Real-Analytic Case*, SIAM J. Control Optim. **25** (4) (1987), 868–904.

[26] SUSSMANN H. J., *A Weak Regularity Theorem for Real Analytic Optimal Control Problems*, Revista Mathematica Iberoamericana **2** (1986), 307–317.

[27] SUSSMANN H. J., *Regular Synthesis for Time-Optimal Control of Single-Input Real-Analytic Systems in the Plane*, SIAM J. Control Optim. **25** (5) (1987), 1145–1162.

[28] ZELKIN M. I., BORISOV V. F., *Theory of Chattering Control, with Applications to Astronautics, Robotics, Economics and Engineering*, Birkhäuser, Boston 1994.

AMS Subject Classification: ???.

Benedetto PICCOLI
DIIMA, Università di Salerno
Via Ponte Don Melillo
84084 Fisciano (SA), ITALY
and
SISSA–ISAS
Via Beirut 2-4
34014 Trieste, ITALY

Ugo BOSCAIN
SISSA–ISAS
Via Beirut 2-4
34014 Trieste, ITALY

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

P. Brandi – A. Salvadori

ON MEASURE DIFFERENTIAL INCLUSIONS IN OPTIMAL CONTROL THEORY

1. Introduction

Differential inclusions are a fundamental tool in optimal control theory. In fact an optimal control problem

$$\min_{(x,u)\in\Omega} J[x,u]$$

can be reduced (via a deparameterization process) to a problem of Calculus of Variation whose solutions can be deduced by suitable closure theorems for differential inclusions.

More precisely, if the cost functional is of the type

$$J[x,u] = \int_I f_0(t, x(t), u(t))\, d\lambda \tag{1}$$

and Ω is a class of admissible pairs subjected to differential and state constraints

$$(t, x(t)) \in A \qquad x'(t) = f(t, x(t), u(t)), \qquad u(t) \in U(t, x(t)) \qquad t \in I \tag{2}$$

the corresponding differential inclusion is

$$(t, x(t)) \in A \qquad x'(t) \in \tilde{Q}(t, x(t)) \qquad t \in I \tag{3}$$

where multifunction $\tilde{Q}$ is related to the epigraph of the integrand i.e.

$$\tilde{Q}(t,x) = \{(z,v) : z \geq f_0(t,x,u),\ u = f(t,x,v),\ v \in U(t,x)\}\,.$$

We refer to Cesari's book [8] where the theory is developed in Sobolev spaces widely.

The extension of this theory to BV setting, motivated by the applications to variational models for plasticity [2, 3, 6, 13], allowed the authors to prove new existence results of discontinuous optimal solutions [4, 5, 9, 10, 11, 12].

This generalized formulation involved differential inclusions of the type

$$(t, x(t)) \in A \qquad x'(t) \in \tilde{Q}(t, x(t)) \qquad \text{a.e. in } I \tag{3*}$$

where u' represents the "essential gradient" of the BV function x, i.e. the density of the absolutely continuous part of the distributional derivative with respect to Lebesgue measure; moreover the Lagrangian functional (1) is replaced by the Serrin-type relaxed functional

$$J[x,u] = \inf_{(x_k,u_k)\to(x,u)} \liminf_{k\to\infty} I[x_k, u_k]\,. \tag{1*}$$

A further extension of this theory was given in [4] where we discussed the existence of L^1 solutions for the abstract evolution equation

$$(3^{**}) \qquad (t, u(t)) \in A \qquad v(t) \in Q(t, u(t)) \qquad \text{a.e. in } I$$

where u and v are two surfaces not necessarely connected. This generalization allowed us to deal with a more general class of optimization problems in BV setting, also including differential elements of higher order or non linear operators (see [4] for the details).

Note that the cost functional J takes into account of the whole distributional gradient of the BV function u, while the constraints control only the "essential" derivative.

To avoid this inconsistency a new class of inclusions involving the measure distributional derivative should be taken into consideration. This is the aim of the research we developed in the present note.

At our knowledge, the first differential inclusion involving the distributional derivative of a BV function was taken into consideration by M. Monteiro Marques [18, 19] who discussed the existence of right continuous and BV solutions for the inclusion

$$(4) \qquad u(t) \in C(t) \qquad -\frac{du}{|du|}(t) \in N_{C(t)}(u(t)) \qquad |du|\text{–a.e. in } I$$

where $C(t)$ is a closed convex set and $N_{C(t)}(a)$ is the normal cone at $C(t)$ in the point $a \in C(t)$.

These inclusions model the so called sweepping process introduced by J.J. Moreau to deal with some mechanical problems.

In [21, 22] J.J. Moreau generalized this formulation to describe general rigid body mechanics with Coulomb friction and introduced the so called measure differential inclusions

$$(4^{*}) \qquad \frac{d\mu}{d\lambda}(t) \in K(t) \qquad \lambda_\mu\text{–a.e. in } I$$

where $\lambda_\mu = \lambda + |\mu|$, with λ is the Lebesgue measure and μ is a Borel measure, and where $K(t)$ is a cone.

Both the inclusions (4) and (4*) are not suitable for our purpose since they can not be applied to multifunction $\tilde{Q}(t, u) = \text{epi } F(t, u, \cdot)$ whose values are not cones, in general.

Recently S.E. Stewart [23] extended this theory to the case of a closed convex set $K(t)$, not necessarely a cone. Inspired by Stewart's research we consider here the following measure differential inclusion

$$(4^{**}) \qquad \begin{aligned} &\frac{d\mu_a}{d\lambda}(t) \in Q(t, u(t)) \qquad &&\lambda\text{–a.e. in } I \\ &\frac{d\mu_s}{d|\mu_s|}(t) \in [Q(t, u(t))]_\infty \qquad &&\mu_s\text{–a.e. in } I \end{aligned}$$

where $\mu = \mu_a + \mu_s$ is the Lebesgue decomposition of the Borel measure μ and $[Q(t, a)]_\infty$ is the asymptotic cone of the non empty, closed, convex set $Q(t, a)$.

Note that measure μ and BV function u are not necessarely correlated, analogously to inclusion (3**). In particular, if μ coincides with the distributional derivative of u, i.e. $\frac{d\mu_a}{d\lambda} = u'$, the first inclusion is exactly (3*), while the second one involves the singular part of the measure derivative.

In other words formulation (4^{**}) is the generalization of (3^*) in the spirit of (3^{**}).

The closure theorem we prove here for inclusion (4^{**}) represents a natural extension of that given in [9, 10, 4, 5] for evolution equations of types (3^*) and (3^{**}). In particular we adopt the same assumption on multifunction Q, which fits very well for the applications to $\tilde{Q}$ and hence to optimal control problems.

Moreover, we wish to remark that our results improve those given by Stewart under stronger assumptions on multifunction Q (see Section 6).

2. Preliminaries

We list here the main notations and some preliminary results.

2.1. On asymptotic cone

DEFINITION 1. *The* asymptotic cone *of a convex set $C \subset \mathbb{R}^n$ is given by*

$$[C]_\infty = \{\lim_{k\to\infty} a_k x_k : a_k \searrow 0,\ x_k \in C,\ k \in \mathbb{N}\}\,.$$

A discussion of the properties of the asymptotic cone can be found in [16] and [23]. We recall here only the results that will be useful in the following.

P_1. *If C is non empty, closed and convex, then $[C]_\infty$ is a closed convex cone.*

P_2. *If C is a closed convex cone, then $C = [C]_\infty$.*

P_3. *If C is non empty, closed and convex, then $[C]_\infty$ is the largest cone K such that $x + K \subset C$, with $x \in C$.*

Let $(C_j)_{j\in J}$ be a family of nonempty closed convex values. Then the following results hold.

P_4. $\operatorname{cl}\operatorname{co}\bigcup_{j\in J}[C_j]_\infty \subset \left[\operatorname{cl}\operatorname{co}\bigcup_{j\in J} C_j\right]_\infty$

P_5. *if* $\bigcap_{j\in J} C_j \neq \phi$, *then* $\left[\bigcap_{j\in J} C_j\right]_\infty = \bigcap_{j\in J}[C_j]_\infty$.

2.2. On property (Q)

Let E be a given subset of a Banach space and let $Q : E \to \mathbb{R}^m$ be a given multifunction. Fixed a point $t_0 \in E$, and a number $h > 0$, we denote by $B_h = B(t_0, h) = \{t \in E : |t - t_0| \le h\}$.

DEFINITION 2. *Multifunction Q is said to satisfy Kuratowski* property (K) *at a point $t_0 \in E$, provided*

$$Q(t_0) = \bigcap_{h>0} \operatorname{cl} \bigcup_{t\in B_h} Q(t)\,. \tag{K}$$

The graph of multifunction Q is the set graph $Q := \{(t, v) : v \in Q(t),\ t \in E\}$.

It is well known that (see e.g. [8])

P_6. graph Q *is closed in* $E \times \mathbb{R}^m \iff Q$ *satisfies condition* (K) *at every point.*

Cesari [8] introduced the following strengthening of Kuratowski condition which is suitable for the differential inclusions involved in optimal control problems in BV setting.

DEFINITION 3. *Multifunction Q is said to satisfy Cesari's* property (Q) *at a point $t_0 \in E$, provided*

$$Q(t_0) = \bigcap_{h>0} \text{cl co} \bigcup_{t \in B_h} Q(t)\,. \tag{Q}$$

Note that if (Q) holds, then the set $Q(t_0)$ *is necessarily closed and convex.*

We will denote by $\mathcal{C}(\mathbb{R}^m)$ the class of non empty, closed, convex subsets of $\mathbb{R}^m$.

Property (Q) is an intermediate condition between Kuratowski condition (K) and upper semicontinuity [8] which is suitable for the applications to optimal control theory. In fact the multifunction defined by

$$\widetilde{Q}(x, u) = \text{epi}\, F(x, u, \cdot)$$

satisfies the following results (see [8]).

P_7. $\widetilde{Q}$ *has closed and convex values iff $F(x, u, \cdot)$ is lower semicontinuous and convex.*

P_8. $\widetilde{Q}$ *satisfies property* (Q) *iff F is seminormal.*

We wish to recall that seminormality is a classical Tonelli's assumption in problems of calculus of variations (see e.g. [8] for more details).

Given a multifunction $Q : E \to \mathcal{C}(\mathbb{R}^m)$, we denote by $Q_\infty : E \to \mathcal{C}(\mathbb{R}^m)$ the multifunction defined by

$$Q_\infty(t) = [Q(t)]_\infty \qquad t \in E\,.$$

PROPOSITION 1. *If Q satisfies property* (Q) *at a point t_0, then also multifunction Q_∞ does.*

Proof. Since

$$\phi \neq Q(t_0) = \bigcap_{h>0} \text{cl co} \bigcup_{t \in B_h} Q(t)$$

from P_4 and P_5 we deduce that

$$Q_\infty(t_0) = \bigcap_{h>0} \left[\text{cl co} \bigcup_{t \in B_h} Q(t) \right]_\infty \subset \bigcap_{h>0} \text{cl co} \bigcup_{t \in B_h} Q_\infty(t)\,.$$

The converse inclusion is trivial and the assertion follows.

□

3. On measure differential inclusions, weak and strong formulations

Let $Q : I \to \mathbb{R}^n$, with $I \subset \mathbb{R}$ closed interval, be a given multifunction with nonempty closed convex values and let μ be a Borel measure on I, of bounded variation.

In [23] Stewart considered the two formulations of measure differential inclusions.

Strong formulation.

$$\text{(S)} \qquad \begin{cases} \frac{d\mu_a}{d\lambda}(t) \in Q(t) & \lambda\text{–a.e. in } I \\ \frac{d\mu_s}{d|\mu_s|}(t) \in Q_\infty(t) & \mu_s\text{–a.e. in } I \end{cases}$$

where $\mu = \mu_a + \mu_s$ be the Lebesgue decomposition of measure μ.

Weak formulation.

$$\text{(W)} \qquad \frac{\int_I \phi\, d\mu}{\int_I \phi\, d\lambda} \in \operatorname{cl} \operatorname{co} \bigcup_{t \in I \cap \operatorname{Supp} \phi} Q(t)$$

for every $\phi \in \mathcal{C}_0$, where $\mathcal{C}_0$ denotes the set of all continuous functions $\phi : \mathbb{R} \to \mathbb{R}_0^+$, with compact support, such that $\int_I \phi\, d\lambda \neq 0$.

Stewart proved that the two formulations are equivalent, under suitable assumptions on Q (see Theorem 2), by means of a transfinite induction process.

We provide here a direct proof of the equivalence, under weaker assumption.

Moreover, for our convenience, we introduce also the following local version of weak formulation.

Local-weak formulation.

Let $t_0 \in I$ be fixed. There exists $\overline{h} = \overline{h}(t_0) > 0$ such that for every $0 < h < \overline{h}$,

$$\text{(LW)} \qquad \frac{\int_{B_h} \phi\, d\mu}{\int_{B_h} \phi\, d\lambda} \in \operatorname{cl} \operatorname{co} \bigcup_{t \in B_h} Q(t)$$

for every $\phi \in \mathcal{C}_0$ such that $\operatorname{Supp} \phi \subset B_h$.

Of course, *if* μ *satisfies* (W), *then* (LW) *holds for every* $t_0 \in I$.

Rather surprising also the convers hold, as we shall show in the following (Theorem 3).

In other words, also this last formulation proves to be equivalent to the previous ones.

THEOREM 1. *Every solution of* (S) *is also a solution of* (W).

Proof. Let $\phi \in \mathcal{C}_0$ be given. Note that $\int_I \phi\, d\mu = \int_I \phi\, d\mu_a + \int_I \phi\, d\mu_s$ moreover

$$(5) \qquad \int_I \phi\, d\mu_a = \int_I \frac{d\mu_a}{d\lambda} \phi\, d\lambda = \int_{I \cap \operatorname{Supp} \phi} \frac{d\mu_a}{d\lambda}\, d\lambda_\phi$$

$$(6) \qquad \int_I \phi\, d\mu_s = \int_I \frac{d\mu_s}{d|\mu_s|} \phi\, d|\mu_s| = \int_{I \cap \operatorname{Supp} \phi} \frac{d\mu_s}{d|\mu_s|}\, d\mu_{s,\phi}$$

where λ_ϕ and $\mu_{s,\phi}$ are the Borel measures defined respectively by

$$\lambda_\phi(E) = \int_E \phi\, d\lambda \qquad \mu_{s,\phi}(E) = \int_E \phi\, d|\mu_s| \qquad E \subset I\,.$$

From (5), in force of the assumption and taking Theorem 1.3 in [1] into account, we get

$$\phi_a := \frac{\int_I \phi\, d\mu_a}{\int_I \phi\, d\lambda} = \frac{\int_{I\cap \mathrm{Supp}\,\phi} \frac{d\mu_a}{d\lambda}\, d\lambda_\phi}{\lambda_\phi(I \cap \mathrm{Supp}\,\phi)} \in \mathrm{cl\,co} \bigcup_{t\in I\cap \mathrm{Supp}\,\phi} Q(t)\,. \tag{7}$$

In the case $\int_I \phi\, d|\mu_s| = 0$, then $\int_I \phi\, d\mu_s = 0$ and the assertion is an immediate consequence of (7).

Let us put

$$Q_\phi := \mathrm{cl\,co} \bigcup_{t\in I\cap \mathrm{Supp}\,\phi} Q(t)\,. \tag{7'}$$

Let us assume now that $\int_I \phi\, d|\mu_s| \neq 0$. Then from (6), in force of the assumption we get, as before

$$\begin{aligned}\frac{\int_I \phi\, d\mu_s}{\int_I \phi\, d|\mu_s|} &= \frac{\int_{I\cap \mathrm{Supp}\,\phi} \frac{d\mu_s}{d|\mu_s|}\, d\mu_{s,\phi}}{\mu_{s,\phi}(I \cap \mathrm{Supp}\,\phi)} \\ &\in \mathrm{cl\,co} \bigcup_{t\in I\cap \mathrm{Supp}\,\phi} Q_\infty(t) \subset \left[\mathrm{cl\,co} \bigcup_{t\in I\cap \mathrm{Supp}\,\phi} Q(t)\right]_\infty = [Q_\phi]_\infty\end{aligned}$$

and since the right-hand side is a cone, we deduce

$$\phi_s := \frac{\int_I \phi\, d\mu_s}{\int_I \phi\, d\lambda} = \frac{\int_I \phi\, d\mu_s}{\int_I \phi\, d|\mu_s|} \cdot \frac{\int_I \phi\, d|\mu_s|}{\int_I \phi\, d\lambda} \in [Q_\phi]_\infty\,. \tag{8}$$

From (7) and (8) we have that

$$\frac{\int_I \phi\, d\mu}{\int_I \phi\, d\lambda} = \phi_a + \phi_s \text{ with } \phi_a \in Q_\phi \quad \phi_s \in [Q_\phi]_\infty$$

and, by virtue of P_3, we conclude that

$$\frac{\int_I \phi\, d\mu}{\int_I \phi\, d\lambda} \in Q_\phi = \mathrm{cl\,co} \bigcup_{t\in \mathrm{Supp}\,\phi} Q(t)$$

which proves the assertion.

□

THEOREM 2. *Let μ be a solution of* (LW) *in $t_0 \in I$.*

(*a*) *If Q has properties* (Q) *at t_0 and the derivative $\frac{d\mu_a}{d\lambda}(t_0)$ exists, then*

$$\frac{d\mu_a}{d\lambda}(t_0) \in Q(t_0)\,.$$

(*b*) *If Q_∞ has properties* (Q) *at t_0 and the derivative $\frac{d\mu_s}{d|\mu_s|}(t_0)$ exists, then*

$$\frac{d\mu_s}{d|\mu_s|}(t_0) \in Q_\infty(t_0)\,.$$

Proof. Let S_μ denote the set where measure μ_s is concentrated, i.e. $S_\mu = \{t \in I : \mu_s\{t\} \neq 0\}$. Since μ_s is of bounded variation, then S_μ is denumerable; let us put

$$S_\mu = \{s_n,\ n \in \mathbb{N}\}\,.$$

Let us fix a point $t_0 \in I^0$. The case where t_0 is an end-point for I is analogous.

The proof will proceed into steps.

Step 1. Let us prove first that for every $B_h = B(t_0, h) \subset I$ with $0 < h < \overline{h}(t_0)$ and such that $\partial B_h \cap S_\mu = \phi$, we have

$$\mu(B_h - S_\mu) \over 2h$$

$$\tag{9} \frac{\mu(B_h - S_\mu)}{2h} = \frac{\mu_a(B_h)}{2h} \in \mathrm{cl\ co} \bigcup_{t \in B_h} Q(t)\,.$$

Let $\overline{n} \in \mathbb{N}$ be fixed. For every $1 \le i \le \overline{n}$, we consider a constant $0 < r_i = r_i(\overline{n}) \le \frac{1}{\overline{n}2^i}$ such that $B(s_i, r_i) \cap B(s_j, r_j) = \phi$, $i \neq j$, $1 \le i, j \le \overline{n}$.

Moreover, we put $I_{\overline{n}} = \displaystyle\bigcup_{i=1}^{\overline{n}} B^0(s_i, r_i)$.

Fixed a constant $0 < \eta < \min\{h,\ r_i,\ 1 \le i \le \overline{n}\}$, we denote by $I_{\overline{n},\eta} = \displaystyle\bigcup_{i=1}^{\overline{n}} B^0(s_i, r_i - \eta)$ and consider the function

$$\phi_{\overline{n},\eta}(t) = \begin{cases} 0 & t \in I - B_h \cup I_{\overline{n},\eta} \\ 1 & t \in B_{h-\eta} - I_{\overline{n}} \\ \text{linear} & \text{otherwise} \end{cases}$$

Of course $\phi_{\overline{n},\eta} \in \mathcal{C}_0$ thus, by virtue of the assumption, we have

$$\tag{10} R_{\overline{n},\eta} := \frac{\int_I \phi_{\overline{n},\eta}\, d\mu}{\int_I \phi_{\overline{n},\eta}\, d\lambda} \in \mathrm{cl\ co} \bigcup_{t \in B_h} Q(t)\,.$$

Note that, put $C_{\overline{n},\eta} = B_h - \left[I_{\overline{n},\eta} \cup (B_{h-\eta} - I_{\overline{n}})\right]$, we have

$$\tag{11} R_{\overline{n},\eta} = \frac{\int_{B_h - I_{\overline{n},\eta}} \phi_{\overline{n},\eta}\, d\mu}{\int_{B_h - I_{\overline{n},\eta}} \phi_{\overline{n},\eta}\, d\lambda} = \frac{\mu\left(B_{h-\eta} - I_{\overline{n}}\right) + \int_{C_{\overline{n},\eta}} \phi_{\overline{n},\eta}\, d\mu}{\lambda\left(B_{h-\eta} - I_{\overline{n}}\right) + \int_{C_{\overline{n},\eta}} \phi_{\overline{n},\eta}\, d\lambda}\,.$$

If we let $\eta \to 0$, we get

$$B_{h-\eta} - I_{\overline{n}} \nearrow B_h^0 - I_{\overline{n}} \qquad I_{\overline{n},\eta} \nearrow I_{\overline{n}}$$

and hence

$$C_{\overline{n},\eta} \searrow \partial B_h = \{t_0 - h, t_0 + h\}\,.$$

As a consequence, we have (see e.g. [14])

$$\tag{12} \begin{aligned} &\lim_{\eta \to 0} \mu(B_{h-\eta} - I_{\overline{n}}) = \mu(B_h - I_{\overline{n}}) \\ &\lim_{\eta \to 0} \lambda(B_{h-\eta} - I_{\overline{n}}) = \lambda(B_h - I_{\overline{n}}) \\ &\lim_{\eta \to 0} |\mu|(C_{\overline{n},\eta}) = \lim_{\eta \to 0} \lambda(C_{\overline{n},\eta}) = 0 \end{aligned}$$

and hence

$$(12') \qquad \lim_{\eta\to 0}\int_{C_{\overline{n},\eta}} \phi_{\overline{n},\eta}\, d\mu = \lim_{\eta\to 0}\int_{C_{\overline{n},\eta}} \phi_{\overline{n},\eta}\, d\lambda = 0\,.$$

From (11), (12) and (12′), we obtain

$$(13) \qquad \lim_{\eta\to 0} R_{\overline{n},\eta} = \frac{\mu(B_h - I_{\overline{n}})}{\lambda(B_h - I_{\overline{n}})} = \frac{\mu_a(B_h - I_{\overline{n}}) + \mu_s(B_h - I_{\overline{n}})}{\lambda(B_h - I_{\overline{n}})}\,.$$

Note that since

$$\lambda(I_{\overline{n}}) = \sum_{i=1}^{\overline{n}} 2r_i \le \frac{2}{\overline{n}}\sum_{i=1}^{\overline{n}}\frac{1}{2^i} < \frac{2}{\overline{n}}$$

we have

$$(14) \qquad \lim_{\overline{n}\to+\infty}\lambda(I_{\overline{n}}) = \lim_{\overline{n}\to+\infty}\mu_a(I_{\overline{n}}) = 0\,.$$

Moreover

$$|\mu_s(B_h - I_{\overline{n}})| \le |\mu_s|(B_h - I_{\overline{n}}) \le |\mu_s|(S_\mu - I_{\overline{n}}) = \sum_{n>\overline{n}} |\mu_s|(\{s_n\})$$

and, recalling that μ has bounded variation

$$(14') \qquad \lim_{\overline{n}\to+\infty}|\mu_s(B_h - I_{\overline{n}})| \le \lim_{\overline{n}\to+\infty}\sum_{n>\overline{n}}|\mu_s|(\{s_n\}) = 0\,.$$

Finally, from (13), (14) and (14′) we conclude that

$$\lim_{\overline{n}\to+\infty}\lim_{\eta\to 0} R_{\overline{n},\eta} = \frac{\mu_a(B_h)}{2h}$$

that, by virtue of (10), proves (9).

Step 2. Let us prove now part (a). We recall that

$$(15) \qquad \frac{d\mu_a}{d\lambda}(t_0) = \lim_{h\to 0}\frac{\mu_a(B_h)}{2h}$$

By virtue of step 1, for every fixed $\overline{h} > 0$ such that $B_{\overline{h}} \subset I$, we have

$$\frac{\mu_a(B_h)}{2h} \in \text{cl co}\bigcup_{t\in B_h} Q(t) \subset \text{cl co}\bigcup_{t\in B_{\overline{h}}} Q(t) \qquad \lambda\text{–a.e.}\quad 0 < h < \overline{h}$$

and hence, by letting $h \to 0$, and taking (15) into account, we get

$$\frac{d\mu_a}{d\lambda}(t_0) \in \text{cl co}\bigcup_{t\in B_{\overline{h}}} Q(t)\,.$$

By virtue of the arbitrariness of $\overline{h} > 0$ and in force of assumption (Q), we conclude that

$$\frac{d\mu_a}{d\lambda}(t_0) \in \bigcap_{\overline{h}>0}\text{cl co}\bigcup_{t\in B_{\overline{h}}} Q(t) = Q(t_0)\,.$$

Step 3. For the proof of part (b) let us note that

$$\frac{d\mu_s}{d|\mu_s|}(t_0) = \frac{\mu_s(\{t_0\})}{|\mu_s|(\{t_0\})} \tag{16}$$

since $\mu_s(\{t_0\}) = \int_{\{t_0\}} d\mu_s = \int_{\{t_0\}} \frac{d\mu_s}{d|\mu_s|}\, d|\mu_s| = \frac{d\mu_s}{d|\mu_s|}(t_0)\, |\mu_s|(\{t_0\})$.
Let $h > 0$ be fixed in such a way that $B_h = B(t_0, h) \subset I$. For every $0 < \eta < h$ we consider the continuous function defined by

$$\phi_\eta(t) = \begin{cases} 1 & t \in B_{\frac{\eta}{2}} \\ 0 & t \in I - B_\eta \\ \text{linear} & \text{otherwise.} \end{cases}$$

Note that (see e.g. [14])

$$\mu_s(\{t_0\}) = \lim_{\eta \to 0} \mu(B_{\frac{\eta}{2}}) = \lim_{\eta \to 0} \mu(B_\eta). \tag{17}$$

Moreover we have

$$\begin{aligned} \mu(B_{\frac{\eta}{2}}) &= \int_{B_\eta} \phi_\eta\, d\mu = \int_I \phi_\eta\, d\mu - \int_{B_\eta - B_{\frac{\eta}{2}}} \phi_\eta\, d\mu \\ &= \frac{\int_I \phi_\eta\, d\mu}{\int_I \phi_\eta\, d\lambda} \cdot \int_I \phi_\eta\, d\lambda - \int_{B_\eta - B_{\frac{\eta}{2}}} \phi_\eta\, d\mu\,. \end{aligned} \tag{18}$$

By assumption we know that

$$\frac{\int_I \phi_\eta\, d\mu}{\int_I \phi_\eta\, d\lambda} \in \operatorname{cl\,co} \bigcup_{t \in B_\eta} Q(t) \subset \operatorname{cl\,co} \bigcup_{t \in B_h} Q(t)$$

let us put

$$Q_h := \operatorname{cl\,co} \bigcup_{t \in B_h} Q(t)\,.$$

Since $\lim\limits_{\eta \to 0} \int_I \phi_\eta\, d\mu = 0$, by virtue of P_4 we get

$$\lim_{\eta \to 0} \frac{\int_I \phi_\eta\, d\mu}{\int_I \phi_\eta\, d\lambda} \cdot \int_I \phi_\eta\, d\lambda \in [Q_h]_\infty\,. \tag{19}$$

Furthermore, by virtue of (17) we have

$$\left| \int_{B_\eta - B_{\frac{\eta}{2}}} \phi_\eta\, d\mu \right| \le |\mu|(B_\eta) - |\mu|(B_{\frac{\eta}{2}})\ \overset{\eta \to 0}{longrightarrow}\ 0 \tag{20}$$

thus, from (18) and taking (17), (19) and (20) into account, we obtain

$$\mu(\{t_0\}) \in [Q_h]_\infty \qquad \text{for every } h > 0 \text{such that } B_h = B(t_0, h) \subset I\,.$$

Finally, recalling P_5 we deduce that

$$\mu(\{t_0\}) \in \bigcap_{h > 0} [Q_h]_\infty = \left[\bigcap_{h > 0} Q_h \right]_\infty = Q_\infty(t_0)$$

and taking (16) into account, since $Q_\infty(t_0)$ is a cone, the assertion follows.

□

DEFINITION 4. *Let μ be a given measure. We will say that a* property P holds (λ, μ_s)–*a.e. if property P is satisfied for every point t with the exception perhaps of a set N with $\lambda(N) + \mu_s(N) = 0$.*

From Theorem 2 the following result can be deduced.

THEOREM 3. *Assume that*

(i) *Q has properties* (Q) *λ–a.e.*

(ii) *Q_∞ has properties* (Q) *μ_s–a.e.*

Then every measure μ which is a solution of (LW) *(λ, μ_s)–a.e. is also a solution of* (S).

As we will observe in Section 6, the present equivalence result [among the three formulations (S), (W), (LW)] improves the equivalence between strong and weak formulation proved by Stewart, by means of a transfinite process in [23].

It is easy to see that Theorem 3 admits the following generalization.

THEOREM 4. *Let $Q_h : I \to \mathcal{C}(\mathbb{R}^m)$, $h \geq 0$ be a net of multifunctions and let μ be a Borel measure. Assume that*

(i) *$Q_0(t_0) = \bigcap_{h>0} Q_h(t_0)$ λ–a.e.;*

(ii) *$[Q_0]_\infty(t_0) = \bigcap_{h>0} [Q_h]_\infty(t_0)$ μ_s–a.e.;*

(iii) *for (λ, μ_s)–a.e. t_0 there exists $\overline{h} = \overline{h}(t_0) > 0$ such that for every $0 < h < \overline{h}$*

$$\frac{\int_{B_h} \phi \, d\mu}{\int_{B_h} \phi \, d\lambda} \in Q_h(t_0)$$

for every $\phi \in \mathcal{C}_0$ such that $\operatorname{Supp} \phi \subset B_h$.

Then μ is a solution of (S).

Proof. Let $t_0 \in I$ be fixed in such a way that all the assumptions hold.
Following the proof of step 1 in Theorem 3, from assumption (iii) we deduce that

$$\frac{\mu_a(B_h)}{2h} \in Q_h(t_0)$$

and hence from assumption (i) (as in step 2) we get

$$\frac{d\mu_a}{d\lambda}(t_0) \in \bigcap_{h>0} Q_h(t_0) = Q_0(t_0)\,.$$

Finally, analogously to the proof of step 3, from asumptions (iii) and (ii) we obtain

$$\mu(\{t_0\}) \in \bigcap_{h>0} [Q_h]_\infty(t_0) = Q_\infty(t_0)$$

and since $Q_\infty(t_0)$ is a cone, we get

$$\frac{d\mu_s}{d|\mu_s|}(t_0) = \frac{\mu(\{t_0\})}{|\mu|(\{t_0\})} \in Q_\infty(t_0)\,.$$

□

4. The main closure theorem

Let $I \subset \mathbb{R}$ be a closed interval and let $Q_k : I \to \mathcal{C}(\mathbb{R}^m)$, $k \geq 0$, be a sequence of multifunctions.

We introduce first the following definition.

DEFINITION 5. *We will say that* $(Q_k)_{k\geq 0}$ satisfies condition (QK) at a point $t_0 \in E$ *provided*

(QK) $$Q_0(t_0) = \bigcap_{h>0} \bigcap_{n\in N} \mathrm{cl} \bigcup_{k\geq n} \mathrm{cl\,co} \bigcup_{t\in B_h} Q_k(t)\,.$$

We are able now to state and prove our main closure result.

THEOREM 5. *Let* $Q_k : I \to \mathcal{C}(\mathbb{R}^m)$*,* $k \geq 0$ *be a sequence of multifunctions and let* $(\mu_k)_{k\geq 0}$ *be a sequence of Borel measures such that*

(i) $(Q_k)_{k\geq 0}$ *satisfies* (QK) *condition* $(\lambda, \mu_{0,s})$*–a.e.;*

(ii) μ_k w^**–converges to* μ_0*;*

(iii) $$\begin{cases} \frac{d\mu_{k,a}}{d\lambda}(t) \in Q_k(t) & \lambda\text{–a.e.} \\ \frac{d\mu_{k,s}}{d|\mu_{k,s}|}(t) \in [Q_k]_\infty(t) & \mu_{k,s}\text{–a.e.} \end{cases}$$

Then the following inclusion holds

$$\begin{cases} \frac{d\mu_{0,a}}{d\lambda}(t) \in Q_0(t) & \lambda\text{–a.e.} \\ \frac{d\mu_{0,s}}{d|\mu_{0,s}|}(t) \in [Q_0]_\infty(t) & \mu_{0,s}\text{–a.e.} \end{cases}$$

Proof. We prove this result as an application of Theorem 4 to the net

$$Q_h(t) = \bigcap_{n\in N} \mathrm{cl} \bigcup_{k\geq n} \mathrm{cl\,co} \bigcap_{\tau\in B(t,h)} Q_k(\tau)\,.$$

By virtue of P_5 assumption (i) assures that both assumptions (i) and (ii) in Theorem 4 hold.

Now, let $t_0 \in I$ be fixed in such a way that assumption (iii) holds and let $\phi \in \mathcal{C}_0$ be given with $\mathrm{Supp}\,\phi \subset B_h \cap I$.

From Theorem 1 we deduce

(21) $$\frac{\int_{\mathrm{Supp}\,\phi} \phi\, d\mu_k}{\int_{\mathrm{Supp}\,\phi} \phi\, d\lambda} \in \mathrm{cl\,co} \bigcup_{t\in \mathrm{Supp}\,\phi} Q_k(t) \qquad k \in \mathbb{N}$$

and from assumption (ii) we get

(22) $$\frac{\int_{\mathrm{Supp}\,\phi} \phi\, d\mu_0}{\int_{\mathrm{Supp}\,\phi} \phi\, d\lambda} = \lim_{k\to+\infty} \frac{\int_{\mathrm{Supp}\,\phi} \phi\, d\mu_k}{\int_{\mathrm{Supp}\,\phi} \phi\, d\lambda} \in Q_h(t_0)$$

which gives assumption (iii) in Theorem 4.

□

5. Further closure theorems for measure differential inclusions

We present here some applications of the main result to remarkable classes of measure differential inclusions.

According to standard notations, we denote by L^1 the space of summable functions $u : I \to \mathbb{R}^m$ and by BV the space of the functions $u \in L^1$ which are of bounded variation in the sense of Cesari [7], i.e. $V(u) < +\infty$.

Let $u_k : I \to \mathbb{R}^m$, $k \geq 0$, be a given sequence in L^1 and let $Q : I \times A \subset \mathbb{R}^{n+1} \to \mathcal{C}(\mathbb{R}^m)$ be a given multifunction.

DEFINITION 6. *We say that the sequence* $(u_k)_{k\geq 0}$ *satisfies the property of* local equi-oscillation *at a point* $t_0 \in I$ *provided*

(LEO)
$$\lim_{h\to 0} \limsup_{k\to\infty} \sup_{t\in B_h} |u_k(t) - u_0(t_0)| = 0\,.$$

It is easy to see that the following result holds.

PROPOSITION 2. *If* u_k *converges uniformly to a continuous function* u_0*, then condition* (LEO) *holds everywhere in* I.

In [10] an other sufficient condition for property (LEO) can be found (see the proof of Theorem 1).

PROPOSITION 3. *If* $(u_k)_{k\geq 0}$ *is a sequence of* BV *functions such that*

(i) u_k *converges to* u_0 λ*–a.e. in I;*

(ii) $\sup_{k\in\mathbb{N}} V(u_k) < +\infty$.

Then a subsequence $(u_{s_k})_{k\geq 0}$ *satisfies condition* (LEO) λ*–a.e. in I.*

Let us prove now a sufficient condition for property (QK).

THEOREM 6. *Assume that the following conditions are satisfied at a point* $t_0 \in I$

(i) Q *satisfies property* (Q)*;*

(ii) $(u_k)_{k\geq 0}$ *satisfies condition* (LEO).

Then the sequence of multifunctions $Q_k : I \to \mathcal{C}(\mathbb{R}^m)$*,* $k \geq 0$*, defined by*

$$Q_k(t) = Q(t, u_k(t)) \qquad k \geq 0$$

satisfies property (QK) *at* t_0.

Proof. By virtue of assumption (ii), fixed $\varepsilon > 0$ a number $0 < h_\varepsilon < \varepsilon$ exists such that for every $0 < h < h_\varepsilon$ an integer k_h exists with the property that for every $k \geq k_h$

$$t \in B_h(t_0) \Longrightarrow |u_0(t_0) - u_k(t)| < \varepsilon\,.$$

Then for every $k \geq k_h$

$$\operatorname{cl}\operatorname{co} \bigcup_{t\in B_h} Q(t, u_k(t)) \subset \operatorname{cl}\operatorname{co} \bigcup_{|t-t_0|\leq\varepsilon,\, |x-u_0(t_0)|\leq\varepsilon} Q(t, x) = Q_\varepsilon\,.$$

Fixed $n \geq k_h$

$$\mathrm{cl} \bigcup_{k \geq n} \mathrm{cl\,co} \bigcup_{t \in B_h} Q(t, u_k(t)) \subset Q_\varepsilon$$

and hence

$$\bigcap_{n \in \mathbb{N}} \mathrm{cl} \bigcup_{k \geq n} \mathrm{cl\,co} \bigcup_{t \in B_h} Q(t, u_k(t)) \subset Q_\varepsilon\ .$$

Finally, by virtue of assumption (i), we have

$$\bigcap_{\varepsilon > 0} \bigcap_{n \in \mathbb{N}} \mathrm{cl} \bigcup_{k \geq n} \mathrm{cl\,co} \bigcup_{t \in B_h} Q(t, u_k(t)) \subset \bigcap_{\varepsilon > 0} Q_\varepsilon = Q(t_0, u_0(t_0))$$

which proves the assertion.

□

In force of this result, the following closure Theorem 5 can be deduced as an application of the main theorem.

THEOREM 7. *Let $Q : I \times A \subset \mathbb{R}^{n+1} \to \mathcal{C}(\mathbb{R}^m)$ be a multifunction, let $(\mu_k)_{k \geq 0}$ be a sequence of Borel measures of bounded variations and let $u_k : I \to A$, $k \geq 0$ be a sequence of BV functions which satisfy the conditions*

(i) *Q has properties* (Q) *at every point (t, x) with the exception of a set of points whose t-coordinate lie on a set of $(\lambda, \mu_{0,s})$–null measure;*

(ii) $$\begin{cases} \frac{d\mu_{k,a}}{d\lambda}(t) \in Q(t, u_k(t)) & \lambda\text{–a.e.} \\ \frac{d\mu_{k,s}}{d|\mu_{k,s}|}(t) \in Q_\infty(t, u_k(t)) & \mu_{k,s}\text{–a.e.} \end{cases}$$

(iii) *μ_k w^*–converges to μ_0;*

(iv) $\sup_{k \in \mathbb{N}} V(u_k) < +\infty$*;*

(v) *u_k converges to u_0 pointwise λ–a.e. and satisfies condition* (LEO) *at $\mu_{0,s}$–a.e.*

Then the following inclusion holds

$$\begin{cases} \frac{d\mu_{0,a}}{d\lambda}(t) \in Q(t, u_0(t)) & \lambda\text{–a.e.} \\ \frac{d\mu_{0,s}}{d|\mu_{0,s}|}(t) \in Q_\infty(t, u_0(t)) & \mu_{0,s}\text{–a.e.} \end{cases}$$

REMARK 1. We recall that the distributional derivative of a BV function u is a Borel measure of bounded variation [17] that we will denote by μ_u.

Moreover u admits an "essential derivative" u' (i.e. computed by usual incremental quotients disregarding the values taken by u on a suitable Lebesgue null set) which coincides with $\frac{d\mu_{u,a}}{d\lambda}$ [25].

Note that Theorem 7 is an extension and a generalization of the main closure theorem in [10] (Theorem 1) given for a differential inclusion of the type

$$u'(t) \in Q(t, u(t)) \qquad \lambda\text{–a.e. in } I\ .$$

To this purpose, we recall that if $(u_k)_{k \geq 0}$, is a sequence of equi–BV functions, then a subsequence of distributional derivatives w^*–converges.

The following closure theorem can be considered as a particular case of Theorem 7.

THEOREM 8. *Let $Q : I \times E \to \mathcal{C}(\mathbb{R}^m)$, with E subset of a Banach space, be a multifunction, let $(\mu_k)_{k\geq 0}$ be a sequence of Borel measures of bounded variations and let $(a_k)_{k\geq 0}$ be a sequence in E. Assume that the following conditions are satisfied*

(i) *Q has properties* (Q) *at every point (t, x) with the exception of a set of points whose t-coordinate lie on a set of $(\lambda, \mu_{0,s})$–null measure;*

(ii) $\begin{cases} \dfrac{d\mu_{k,a}}{d\lambda}(t) \in Q(t, a_k) & \lambda\text{–a.e.} \\ \dfrac{d\mu_{k,s}}{d|\mu_{k,s}|}(t) \in Q_\infty(t, a_k) & \mu_{k,s}\text{–a.e.} \end{cases}$

(iii) *μ_k w^*–converges to μ_0;*

(iv) *$(a_k)_k$ converges to a_0.*

Then the following inclusion holds

$$\begin{cases} \dfrac{d\mu_{0,a}}{d\lambda}(t) \in Q(t, a_0) & \lambda\text{–a.e.} \\ \dfrac{d\mu_{0,s}}{d|\mu_{0,s}|}(t) \in Q_\infty(t, a_0) & \mu_{0,s}\text{–a.e.} \end{cases}$$

As we will prove in Section 6, this last result is an extension of closure Theorem 3 in [10].

As an application of Theorem 7 also the following result can be proved.

THEOREM 9. *Let $Q : I \times \mathbb{R}^n \times \mathbb{R}^p \to \mathcal{C}(\mathbb{R}^m)$, be a multifunction, let $f : I \times \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$ be a function and let $(u_k, v_k) : I \to \mathbb{R}^n \times \mathbb{R}^q$, $k \geq 0$, be a sequence of functions.*
Assume that

(i) *Q satisfies property* (Q) *at every point (t, x, y) with the exception of a set of points whose t-coordinate lie on a set of $(\lambda, \mu_{v_0,s})$–null measure;*

(ii) *f is a Carathéodory function and*
$|f(t, u, v)| \leq \psi_1(t) + \psi_2(t)\,|u| + \psi_3(t)\,|v|$ *with* $\psi_i \in L^1$ $i = 1, 2, 3$*;*

(iii) $\begin{cases} v_k'(t) \in Q(t, u_k(t)) - f(t, u_k(t), v_k(t)) & \lambda\text{–a.e.} \\ \dfrac{d\mu_{v_k,s}}{d|\mu_{v_k,s}|}(t) \in Q_\infty(t, u_k(t)) & \mu_{v_k,s}\text{–a.e.} \end{cases}$

(iv) $\sup_{k\in\mathbb{N}} V(v_k) < +\infty$ *and $(v_k)_k$ converges to v_0 λ–a.e.;*

(v) *$(u_k)_k$ converges uniformly to a continuous function u_0.*

Then the following inclusion holds

$$\begin{cases} v_0'(t) \in Q(t, u_0(t)) - f(t, u_0(t), v_0(t)) & \lambda\text{–a.e.} \\ \dfrac{d\mu_{v_0,s}}{d|\mu_{v_0,s}|}(t) \in Q_\infty(t, u_0) & \mu_{v_0,s}\text{–a.e.} \end{cases}$$

Proof. If we consider the sequence of Borel measures defined by

$$\nu_k([a, b]) = \int_a^b [v_k'(t) + f(t, u_k(t), v_k(t))]\, d\lambda \qquad [a, b] \subset I \qquad k \geq 0$$

it is easy to see that

$$d\nu_{k,s} = d\mu_{v_k,s} \qquad \frac{d\nu_{k,a}}{d\lambda}(t) = v_k'(t) + f(t, u_k(t), v_k(t)) \qquad \lambda\text{–a.e.}$$

It is easy to verify that assumptions assure that $(\nu_k)_{k\geq 0}$ is a sequence of BV measure which w^*–converges and the result is an immediate application of Theorem 7.

□

REMARK 2. Differential incusions of this type are adopted as a model for rigid body dynamics (see [20] for details). As we will observe in Section 6 the previous result improves the analogous theorem proved in [23] (Theorem 4).

6. On comparison with Stewart's assumptions

This section is dedicated to a discussion on the comparison between our assumptions and that adopted by Stewart in [23].

Let $Q : E \to \mathcal{C}(\mathbb{R}^n)$ be a given multifunction where E is a subset of a Banach space.

The main hypotheses adopted by Stewart in [23] on muntifunction Q are the closure of the graph (i.e. property (K)) and the following condition:

$$\text{for every } t_0 \in E \text{there exist } \sigma_0 > 0 \text{and } R_0 > 0 \text{such that} \quad \sup_{t \in B_\sigma} \inf_{x \in Q(t)} \|x\| \le R_0 \,. \tag{23}$$

We will prove here that these assumptions are stricly stronger than property (Q). As a consequence, the results of the present paper improve that given in [23].

PROPOSITION 4. *Let Q be a multifunction with closed graph and let $t_0 \in E$ be fixed. Assume that*
for a given $t_0 \in E$ there exist $\sigma_0 > 0$ and $R_0 > 0$ such that

$$\sup_{t \in B_\sigma} \inf_{x \in Q(t)} \|x\| \le R_0$$

then multifunction Q satisfies property (Q) *at t_0.*

Proof. By virtue of Lemma 5.1 in [23], fixed a number $\varepsilon > 0$, there exists $\delta = \delta(t_0, \varepsilon) > 0$ such that

$$t \in B_\delta \Longrightarrow Q(t) \subset Q(t_0) + \varepsilon\, B(0,1) + (Q_\infty(t_0))_\varepsilon$$

where $(Q_\infty(t_0))_\varepsilon$ denotes the ε–enlargement of the set $Q_\infty(t_0)$.
Since the right-hand side is closed and convex

$$\text{cl co} \bigcup_{t \in B_\delta} Q(t) \subset Q(t_0) + \varepsilon\, B(0,1) + (Q_\infty(t_0))_\varepsilon$$

then

$$Q^*(t_0) := \bigcap_{\delta > 0} \text{cl co} \bigcup_{t \in B_\delta} Q(t) \subset Q(t_0) + \varepsilon\, B(0,1) + (Q_\infty(t_0))_\varepsilon \,.$$

Now, fixed an integer $n \in \mathbb{N}$ and $0 < \varepsilon < \frac{1}{n}$, we get

$$Q^*(t_0) \subset Q(t_0) + \varepsilon\, B(0,1) + (Q_\infty(t_0))_\varepsilon \subset Q(t_0) + \varepsilon\, B(0,1) + (Q_\infty(t_0))_{\frac{1}{n}}$$

and letting $\varepsilon \to 0$, we obtain

$$Q^*(t_0) \subset Q(t_0) + (Q_\infty(t_0))_{\frac{1}{n}} \,. \tag{24}$$

Recalling that (see P_3)

$$Q(t_0) + Q_\infty(t_0) \subset Q(t_0)$$

we have

$$Q(t_0) + (Q_\infty(t_0))_{\frac{1}{n}} \subset (Q(t_0))_{\frac{1}{n}} + (Q_\infty(t_0))_{\frac{1}{n}} \subset (Q(t_0) + Q_\infty(t_0))_{\frac{1}{n}} \subset (Q(t_0))_{\frac{2}{n}}$$

and from (24) letting $n \to +\infty$ we get

$$Q^*(t_0) \subset Q(t_0)$$

which proves the assertion.

$\square$

This result proves that even if Kuratowski condition (K) is weaker than Cesari's property (Q) (see Section 2), together with hypothesis (23) it becomes a stronger assumption. The following example will show that assumption (23) and (K) are strictly stronger than property (Q).
Finally, we recall that in BV setting property (Q) can not be replaced by condition (K), as it occurs in Sobolev's setting (see [10], Remark 1).

EXAMPLE 1. Let us consider the function $F : \mathbb{R}_0^+ \times \mathbb{R} \to \mathbb{R}$ defined by

$$F(t, v) = \begin{cases} \frac{1}{t}\sin^2\frac{1}{t} + |v| & t \neq 0 \\ |v| & t = 0 \end{cases}$$

and the multifunction

$$\tilde{Q}(t, \cdot) = \mathrm{epi}\, F(t, \cdot)\,.$$

Of course assumption (ii) in Proposition 2 does not holds for $\tilde{Q}$ at the point $t_0 = 0$.
Moreover, in force of the Corollary to Theorem 3 in A.W.J. Stoddart [24], it can be easily proved that F is seminormal. Thus $\tilde{Q}$ satisfies condition (Q) at every point $t \in \mathbb{R}_0^+$ (see P_8).

References

[1] BAIOCCHI C., *Ulteriori osservazioni sull'integrale di Bochner*, Ann. Scuola Norm. Sup. Pisa **18** (1964), 283–301.

[2] BRANDI P., SALVADORI A., *Non-smooth solutions in plastic deformation*, Atti Sem. Mat. Fis. Univ. Modena **41** (1993), 483–490.

[3] BRANDI P., SALVADORI A., *A variational approach to problems of plastic deformation*, Developments in Partial Differential Equations and Applications to Mathematical Physics, G. Buttazzo, G. P. Galdi and L. Zanghirati Ed., Plenum Press (l992), 219–226.

[4] BRANDI P., SALVADORI A., *On the lower semicontinuity in BV setting*, J. Convex Analysis **1** (1994), 151–172.

[5] BRANDI P., SALVADORI A., *Closure theorems in BV setting*, Progress in Partial Differential Equations: the Metz survay 4, Chipot and Shafrir Ed., **345** (1996), 42–52, Longman Ed.

[6] BRANDI P., SALVADORI A., YANG W. H., *A nonlinear variational approach to plasticity and beyond*, Proceedings of HMM 99 - Second International Symposium on Hysteresis Modeling and Micromagnetics, Perugia 7-9.6.1999.

[7] CESARI L., *Sulle funzioni a variazione limitata*, Ann. Scuola Nor. Sup. Pisa **5** (1936), 299–313.

[8] CESARI L., *Optimization Theory and Applications*, Springer Verlag 1983.

[9] CESARI L., BRANDI P., SALVADORI A., *Discontinuous solutions in problems of optimization*, Ann. Scuola Norm. Sup. Pisa **15** (1988), 219–237.

[10] CESARI L., BRANDI P., SALVADORI A., *Existence theorems concerning simple integrals of the calculus of variations for discontinuous solutions*, Arch. Rat. Mech. Anal. **98** (1987), 307–328.

[11] CESARI L., BRANDI P., SALVADORI A., *Existence theorems for multiple integrals of the calculus of variations for discontinuous solutions*, Ann. Mat. Pura Appl. **152** (1988), 95–121.

[12] CESARI L., BRANDI P., SALVADORI A., *Seminormality conditions in calculus of variations for BV solutions*, Ann. Mat. Pura Appl. **162** (1992), 299–317.

[13] CESARI L., YANG W. H., *Serrin's integrals and second order problems of plasticity*, Proc. Royal Soc. Edimburg **117** (1991), 193–201.

[14] DINCULEANU N., *Vector Measures*, Pergamon Press 1967.

[15] HALMOS P. R., *Measure Theory*, Springer Verlag 1974.

[16] HIRIART-URRUTY B. J., LEMARÉCHAL C., *Convex Analysis and Minimization Algorithms*, Springer Verlag 1993.

[17] KRICKEBERG K., *Distributionen, Funktionen beschränkter Variation und Lebesguescher Inhalt nichtparametrischer Flächen*, Ann. Mat. Pura Appl. **44** (1957), 105–133.

[18] MONTEIRO MARQUES M., *Perturbation semi-continues superiorment de problems d'evolution dans les spaces de Hilbert*, Seminaire d'Analyse Convexe, Montpellier 1984, exposé n. 2.

[19] MONTEIRO MARQUES M., *Rafle par un convexe semi-continues inferieurement d'interieur non vide en dimension finie*, Seminaire d'Analyse Convexe, Montpellier 1984, exposé n. 6.

[20] MONTEIRO MARQUES M., *Differential Inclusions in Nonsmooth Mechanical Problems: Shocks and Dry Friction*, Birkäuser Verlag 1993.

[21] MOREAU J. J., *Une formulation du contact a frottement sec; application au calcul numerique*, C. R. Acad. Sci. Paris **302** (1986), 799–801.

[22] MOREAU J. J., *Unilateral contact and dry friction in finite freedom dynamics*, Nonsmooth mechanics and Applications, J. J. Moreau and P. D. Panagiotopoulos Ed., Springer Verlag 1988.

[23] STEWART D. E., *Asymptotic cones and measure differential inclusions*, preprint.

[24] STODDART A. W. J., *Semicontinuity of integrals*, Trans. Amer. Math. Soc. **122** (1966), 120–135.

[25] ZIEMER W., *Weakly Differentiable Functions*, Springer-Verlag 1989.

AMS Subject Classification: ???.

Primo BRANDI, Anna SALVADORI
Department of Mathematics
University of Perugia
Via L. Vanvitelli 1
06123 Perugia, Italy
e-mail: `mateas@unipg.it`

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

A. Bressan

SINGULARITIES OF STABILIZING FEEDBACKS

1. Introduction

This paper is concerned with the stabilization problem for a control system of the form

$$\dot{x} = f(x, u), \qquad u \in K\,, \tag{1}$$

assuming that the set of control values $K \subset \mathbb{R}^m$ is compact and that the map $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ is smooth. It is well known [6] that, even if every initial state $\bar{x} \in \mathbb{R}^n$ can be steered to the origin by an open-loop control $u = u^{\bar{x}}(t)$, there may not exist a continuous feedback control $u = U(x)$ which locally stabilizes the system (1). One is thus forced to look for a stabilizing feedback within a class of discontinuous functions. However, this leads to a theoretical difficulty, because, when the function U is discontinuous, the differential equation

$$\dot{x} = f\,(x, U(x)) \tag{2}$$

may not have any Carathéodory solution. To cope with this problem, two approaches are possible.

I) On one hand, one may choose to work with completely arbitrary feedback controls U. In this case, to make sense of the evolution equation (2), one must introduce a suitable definition of "generalized solution" for discontinuous O.D.E. For such solutions, a general existence theorem should be available.

II) On the other hand, one may try to solve the stabilization problem within a particular class of feedback controls U whose discontinuities are sufficiently tame. In this case, it will suffice to consider solutions of (2) in the usual Carathéodory sense.

The first approach is more in the spirit of [7], while the second was taken in [1]. In the present note we will briefly survey various definitions of generalized solutions found in the literature [2, 11, 12, 13, 14], discussing their possible application to problems of feedback stabilization. In the last sections, we will consider particular classes of discontinuous vector fields which always admit Carathéodory solutions [3, 5, 16], and outline some research directions related to the second approach.

In the following, $\overline{\Omega}$ and $\partial\Omega$ denote the closure and the boundary of a set Ω, while B_ε is the open ball centered at the origin with radius ε. To fix the ideas, two model problems will be considered.

Asymptotic Stabilization (AS). Construct a feedback $u = U(x)$, defined on $\mathbb{R}^n \setminus \{0\}$, such that every trajectory of (2) either tends to the origin as $t \to \infty$ or else reaches the origin in finite time.

Suboptimal Controllability (SOC). Consider the minimum time function

$$T(\bar{x}) \doteq \min\{t : \text{there exists a trajectory of (1) with } x(0) = \bar{x},\ x(t) = 0\}\ . \tag{3}$$

Call $R(\tau) \doteq \{x : T(x) \leq \tau\}$ the set of points that can be steered to the origin within time τ. For a given $\varepsilon > 0$, we want to construct a feedback $u = U(x)$, defined on a neighborhood V of $R(\tau)$, with the following property. For every $\bar{x} \in V$, every trajectory of (2) starting at $\bar{x}$ reaches a point inside B_ε within time $T(\bar{x}) + \varepsilon$.

Notice that we are not concerned here with time optimal feedbacks, but only with suboptimal ones. Indeed, already for systems on $\mathbb{R}^2$, an accurate description of all generic singularities of a time optimal feedback involves the classification of a large number of singular points [4, 15]. In higher dimensions, an ever growing number of different singularities can arise, and time optimal feedbacks may exhibit pathological behaviors. A complete classification thus appears to be an enormous task, if at all possible. By working with suboptimal feedbacks, we expect that such bad behaviors can be avoided. One can thus hope to construct suboptimal feedback controls having a much smaller set of singularities.

2. Nonexistence of continous stabilizing feedbacks

The papers [6, 19, 20] provided the first examples of control systems which can be asymptotically stabilized at the origin, but where no continuous feedback control $u = U(x)$ has the property that all trajectories of (2) asymptotically tend to the origin as $t \to \infty$. One such case is the following.

EXAMPLE 1. Consider the control system on $\mathbb{R}^3$

$$(\dot{x}_1, \dot{x}_2, \dot{x}_3) = (u_1, u_2, x_1u_2 - x_2u_1) . \tag{4}$$

As control set K one can take here the closed unit ball in $\mathbb{R}^2$. Using Lie-algebraic techniques, it is easy to show that this system is globally controllable to the origin. However, no smooth feedback $u = U(x)$ can achieve this stabilization.

Indeed, the existence of such a feedback would imply the existence of a compact neighborhood V of the origin which is positively invariant for the flow of the smooth vector field $g(x) \doteq f(x, U(x))$. Calling $T_V(x)$ the contingent cone [2, 8] to the set V at the point x, we thus have $g(x) \in T_V(x)$ at each boundary point $x \in \partial V$. Since g cannot vanish outside the origin, by a topological degree argument, there must be a point x^* where the field g is parallel to the x_3-axis: $g(x^*) = (0, 0, y)$ for some $y > 0$. But this is clearly impossible by the definition (4) of the vector field.
Using a mollification procedure, from a continuous stabilizing feedback one could easily construct a smooth one. Therefore, the above argument also rules out the existence of continuous stabilizing feedbacks.

We describe below a simple case where the problem of suboptimal controllability to zero cannot be solved by any continuous feedback.

EXAMPLE 2. Consider the system

$$(\dot{x}_1, \dot{x}_2) = (u, -x_1^2), \qquad u \in [-1, 1] . \tag{5}$$

The set of points that can be steered to the origin within time $\tau = 1$ is found to be

$$R(1) = \left\{ (x_1, x_2) : x_1 \in [-1, 1], \quad \frac{1}{3}|x_1^3| \leq x_2 \leq \frac{1}{4}\left(\frac{1}{3} + |x_1| + x_1^2 - |x_1^3|\right) \right\} . \tag{6}$$

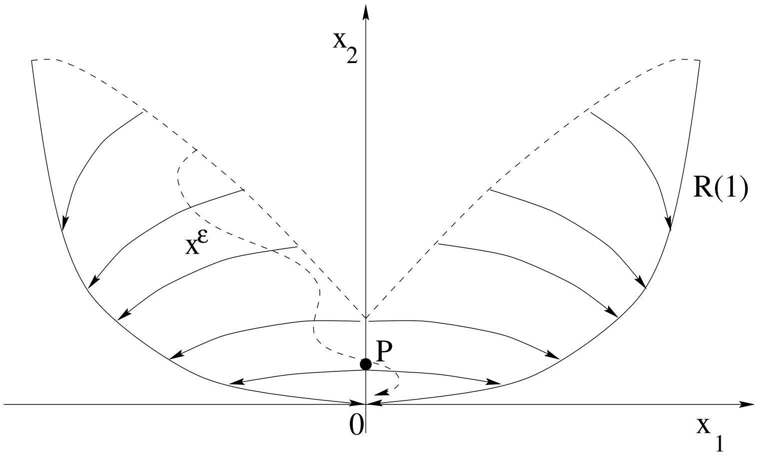

figure 1

Moreover, all time-optimal controls are bang-bang with at most one switching, as shown in fig. 1.

Assume that for every $\varepsilon > 0$ there exists a continuous feedback U_ε such that all trajectories of

$$\dot{x} = (\dot{x}_1, \dot{x}_2) = \left(U_\varepsilon(x), -x_1^2\right)$$

starting at some point $\bar{x} \in R(1)$ reach the ball B_ε within time $T(\bar{x})+\varepsilon$. To derive a contradiction, fix the point $P = (0, 1/24)$. By continuity, for each ε sufficiently small, there will be at least one trajectory $x^\varepsilon(\cdot)$ starting from a point on the upper boundary

$$\partial^+ R(1) \doteq \left\{(x_1, x_2) : x_1 \in [-1, 1], \quad x_2 = \frac{1}{4}\left(\frac{1}{3} + |x_1| + x_1^2 - |x_1^3|\right)\right\} \tag{7}$$

and passing through P before reaching a point in B_ε. By compactness, as $\varepsilon \to 0$ we can take a subsequence of trajectories $x^\varepsilon(\cdot)$ converging to function $x^*(\cdot)$ on $[0, 1]$. By construction, $x^*(\cdot)$ is then a time optimal trajectory starting from a point on the upper boundary $\partial^+ R(1)$ and reaching the origin in minimum time, passing through the point P at some intermediate time $s \in]0, 1[$. But this is a contradiction because no such trajectory exists.

3. Generalized solutions of a discontinuous O.D.E.

Let g be a bounded, possibly discontinuous vector field on $\mathbb{R}^n$. In connection with the O.D.E.

$$\dot{x} = g(x)\,, \tag{8}$$

various concepts of "generalized" solutions can be found in the literature. We discuss here the two main approaches.

(A) Starting from g, by some regularization procedure, one constructs an upper semicontinuous multifunction G with compact convex values. Every absolutely continuous function which satisfies a.e. the differential inclusion

$$\dot{x} \in G(x) \tag{9}$$

can then be regarded as generalized solutions of (8).

In the case of *Krasovskii solutions*, one takes the multifunction

(10) $$G(x) \doteq \bigcap_{\varepsilon>0} \overline{co}\,\{g(y) : |y-x| < \varepsilon\} .$$

Here $\overline{co}\,A$ denotes the closed convex hull of the set A. The *Filippov solutions* are defined similarly, except that one now excludes sets of measure zero from the domain of g. More precisely, calling $\mathcal{N}$ the family of sets $A \subset \mathbb{R}^n$ of measure zero, one defines

(11) $$G(x) \doteq \bigcap_{\varepsilon>0} \bigcap_{A\in\mathcal{N}} \overline{co}\,\{g(y) : |y-x| < \varepsilon,\ y \notin A\} .$$

Concerning solutions of the multivalued Cauchy problem

(12) $$x(0) = \bar{x}, \qquad \dot{x}(t) \in G\,(x(t)) \qquad t \in [0, T]\,,$$

one has the following existence result [2].

THEOREM 1. *Let g be a bounded vector field on $\mathbb{R}^n$. Then the multifunction G defined by either (10) or (11) is upper semicontinuous with compact convex values. For every initial data $\bar{x}$, the family $\mathcal{F}^{\bar{x}}$ of Carathéodory solutions of (12) is a nonempty, compact, connected, acyclic subset of $\mathcal{C}\left([0, T]\,;\,\mathbb{R}^n\right)$. The map $\bar{x} \mapsto \mathcal{F}^{\bar{x}}$ is upper semicontinuous. If g is continuous, then $G(x) = \{g(x)\}$ for all x, hence the solutions of (8) and (9) coincide.*

It may appear that the nice properties of Krasovskii or Filippov solutions stated in Theorem 1 make them a very attractive candidate toward a theory of discontinuous feedback control. However, quite the contrary is true. Indeed, by Theorem 1 the solution sets for the multivalued Cauchy problem (12) have the same topological properties as the solution sets for the standard Cauchy problem

(13) $$x(0) = \bar{x}, \qquad \dot{x}(t) = g\,(x(t)) \qquad t \in [0, T]$$

with continuous right hand side. As a result, the same topological obstructions found in Examples 1 and 2 will again be encountered in connection with Krasovskii or Filippov solutions. Namely [10, 17], for the system (4) one can show that for every discontinuous feedback $u = U(x)$ there will be some Filippov solution of the corresponding discontinuous O.D.E. (2) which does not approach the origin as $t \to \infty$. Similarly, for the system (5), when $\varepsilon > 0$ is small enough there exists no feedback $u = U(x)$ such that every Filippov solution of (2) starting from some point $\bar{x} \in R(1)$ reaches the ball B_ε within time $T(\bar{x}) + \varepsilon$.

The above considerations show the necessity of a new definition of "generalized solution" for a discontinuous O.D.E. which will allow the solution set to be possibly disconnected. The next paragraph describes a step in this direction.

(B) Following a second approach, one defines an algorithm which constructs a family of ε-approximate solutions x_ε. Letting the approximation parameter $\varepsilon \to 0$, every uniform limit $x(\cdot) = \lim_{\varepsilon\to 0} x_\varepsilon(\cdot)$ is defined to be a generalized solution of (8).

Of course, there is a wide variety of techniques [8, 13, 14] for constructing approximate solutions to the Cauchy problem (13). We describe here two particularly significant procedures.

Polygonal Approximations. By a general *polygonal ε-approximate* solution of (13) we mean any function $x : [0, T] \mapsto \mathbb{R}^n$ constructed by the following procedure. Consider a partition of the interval $[0, T]$, say $0 = t_0 < t_1 < \cdots < t_m = T$, whose mesh size satisfies

$$\max_i (t_i - t_{i-1}) < \varepsilon .$$

For $i = 0, \dots, m-1$, choose arbitrary outer and inner perturbations $e_i, e'_i \in \mathbb{R}^n$, with the only requirement that $|e_i| < \varepsilon$, $|e'_i| < \varepsilon$. By induction on i, determine the values x_i such that

$$|x_0 - \bar{x}| < \varepsilon, \qquad x_{i+1} = x_i + (t_{i+1} - t_i)\big(e_i + g(x_i + e'_i)\big) \tag{14}$$

Finally, define $x(\cdot)$ as the continuous, piecewise affine function such that $x(t_i) = x_i$ for all $i = 0, \dots, m$.

Forward Euler Approximations. By a *forward Euler ε-approximate* solution of (13) we mean any polygonal approximation constructed without taking any inner perturbation, i.e. with $e'_i \equiv 0$ for all i.

In the following, the trajectories of the differential inclusion (12), with G given by (10) or (11) will be called respectively *Krasovskii* or *Filippov solutions* of (13). By a *forward Euler solution* we mean a limit of forward Euler ε-approximate solutions, as $\varepsilon \to 0$. Some relations between these different concepts of solutions are illustrated below.

THEOREM 2. *The set of Krasovskii solutions of (13) coincides with the set of all limits of polygonal ε-approximate solutions, as $\varepsilon \to 0$.*

For a proof, see [2, 9].

EXAMPLE 3. On the real line, consider the vector field (fig. 2)

$$g(x) = \begin{cases} 1 & \text{if } x \geq 0\,, \\ -1 & \text{if } x < 0\,. \end{cases}$$

The corresponding multifunction G, according to both (10) and (11) is

$$G(x) = \begin{cases} \{1\} & \text{if } x > 0\,, \\ [-1, 1] & \text{if } x = 0\,, \\ \{-1\} & \text{if } x < 0\,. \end{cases}$$

The set of Krasovskii (or Filippov) solutions to (13) with initial data $\bar{x} = 0$ thus consists of all functions of the form

$$x(t) = \begin{cases} 0 & \text{if } t \leq \tau\,, \\ t - \tau & \text{if } t > \tau\,, \end{cases}$$

together with all functions of the form

$$x(t) = \begin{cases} 0 & \text{if } t \leq \tau\,, \\ \tau - t & \text{if } t > \tau\,, \end{cases}$$

for any $\tau \geq 0$. On the other hand, there are only two forward Euler solutions:

$$x_1(t) = t, \qquad x_2(t) = -t\,.$$

In particular, this set of limit solutions is not connected.

EXAMPLE 4. On $\mathbb{R}^2$ consider the vector field (fig. 3)

$$g(x_1, x_2) \doteq \begin{cases} (0, -1) & \text{if } x_2 > 0\,, \\ (0, 1) & \text{if } x_2 < 0\,, \\ (1, 0) & \text{if } x_2 = 0\,. \end{cases}$$

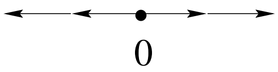

figure 2

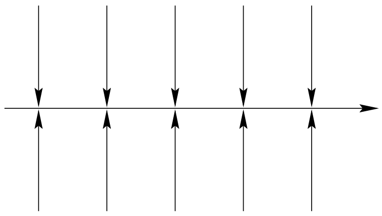

figure 3

The corresponding Krasovskii multivalued regularization (10) is

$$G_K(x_1, x_2) = \begin{cases} \{(0, -1)\} & \text{if } x_2 > 0\,, \\ \{(0, 1)\} & \text{if } x_2 < 0\,, \\ \overline{co}\,\{(0, -1), (0, 1), (1, 0)\} & \text{if } x_2 = 0\,. \end{cases}$$

Given the initial condition $\bar{x} = (0, 0)$, the corresponding Krasovskii solutions are all the functions of the form $t \mapsto (x_1(t), 0)$, with $\dot{x}_1(t) \in [0, 1]$ almost everywhere. These coincide with the limits of forward Euler approximations. On the other hand, since the line $\{x_2 = 0\}$ is a null set, the Filippov multivalued regularization (11) is

$$G_F(x_1, x_2) = \begin{cases} \{(0, -1)\} & \text{if } x_2 > 0\,, \\ \{(0, 1)\} & \text{if } x_2 < 0\,, \\ \overline{co}\,\{(0, -1), (0, 1)\} & \text{if } x_2 = 0\,. \end{cases}$$

Therefore, the only Filippov solution starting from the origin is the function $x(t) \equiv (0, 0)$ for all $t \geq 0$.

4. Patchy vector fields

For a general discontinuous vector field g, the Cauchy problem for the O.D.E.

$$\dot{x} = g(x) \tag{15}$$

may not have any Carathéodory solution. Or else, the solution set may exhibit very wild behavior. It is our purpose to introduce a particular class of discontinuous maps g whose corresponding trajectories are quite well behaved. This is particularly interesting, because it appears that various stabilization problems can be solved by discontinuous feedback controls within this class.

DEFINITION 1. *By a* patch *we mean a pair (Ω, g) where $\Omega \subset \mathbb{R}^n$ is an open domain with smooth boundary and g is a smooth vector field defined on a neighborhood of $\overline{\Omega}$ which points strictly inward at each boundary point $x \in \partial\Omega$.*

Calling $\mathbf{n}(x)$ the outer normal at the boundary point x, we thus require

$$\langle g(x), \mathbf{n}(x) \rangle < 0 \qquad \text{for all } x \in \partial\Omega\,. \tag{16}$$

DEFINITION 2. *We say that* $g : \Omega \mapsto \mathbb{R}^n$ *is a* patchy vector field *on the open domain* Ω *if there exists a family of patches* $\{(\Omega_\alpha, g_\alpha) : \alpha \in \mathcal{A}\}$ *such that*
- $\mathcal{A}$ is a totally ordered set of indices,
- the open sets Ω_α form a locally finite covering of Ω,
- the vector field g can be written in the form

$$g(x) = g_\alpha(x) \text{ if } x \in \Omega_\alpha \setminus \bigcup_{\beta > \alpha} \Omega_\beta \,. \tag{17}$$

By defining

$$\alpha^*(x) \doteq \max\{\alpha \in \mathcal{A} : x \in \Omega_\alpha\} \,, \tag{18}$$

we can write (17) in the equivalent form

$$g(x) = g_{\alpha^*(x)}(x) \text{ for all } x \in \Omega \,. \tag{19}$$

We shall occasionally adopt the longer notation $\big(\Omega, g, (\Omega_\alpha, g_\alpha)_{\alpha\in\mathcal{A}}\big)$ to indicate a patchy vector field, specifying both the domain and the single patches. Of course, the patches $(\Omega_\alpha, g_\alpha)$ are not uniquely determined by the vector field g. Indeed, whenever $\alpha < \beta$, by (17) the values of g_α on the set $\Omega_\beta \setminus \Omega_\alpha$ are irrelevant. This is further illustrated by the following lemma.

LEMMA 1. *Assume that the open sets Ω_α form a locally finite covering of Ω and that, for each $\alpha \in \mathcal{A}$, the vector field g_α satisfies the condition (16) at every point $x \in \partial\Omega_\alpha \setminus \cup_{\beta>\alpha}\Omega_\beta$. Then g is again a patchy vector field.*

Proof. To prove the lemma, it suffices to construct vector fields $\tilde{g}_\alpha$ which satisfy the inward pointing property (16) at every point $x \in \partial\Omega_\alpha$ and such that $\tilde{g}_\alpha = g_\alpha$ on $\Omega_\alpha \setminus \cup_{\beta>\alpha}\Omega_\beta$. To accomplish this, for each α we first consider a smooth vector field v_α such that $v_\alpha(x) = -\mathbf{n}(x)$ on $\partial\Omega_\alpha$. The map $\tilde{g}_\alpha$ is then defined as the interpolation

$$\tilde{g}_\alpha(x) \doteq \varphi(x) g_\alpha(x) + (1 - \varphi(x))\, v_\alpha(x) \,,$$

where φ is a smooth scalar function such that

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in \Omega_\alpha \setminus \cup_{\beta>\alpha}\Omega_\beta \,, \\ 0 & \text{if } x \in \partial\Omega_\alpha \text{ and } \langle g(x), \mathbf{n}(x)\rangle \geq 0 \,. \end{cases}$$

□

The main properties of trajectories of a patchy vector field (fig. 4) are collected below.

THEOREM 3. *Let* $\big(\Omega, g, (\Omega_\alpha, g_\alpha)_{\alpha\in\mathcal{A}}\big)$ *be a patchy vector field.*

(i) *If $t \to x(t)$ is a Carathéodory solution of (15) on an open interval J, then $t \to \dot{x}(t)$ is piecewise smooth and has a finite set of jumps on any compact subinterval $J' \subset J$. The function $t \mapsto \alpha^*(x(t))$ defined by (18) is piecewise constant, left continuous and non-decreasing. Moreover there holds*

$$\dot{x}((t-)) = g\,(x(t)) \text{ for all } t \in J \,. \tag{20}$$

(ii) *For each $\bar{x} \in \Omega$, the Cauchy problem for (15) with initial condition $x(0) = \bar{x}$ has at least one local forward Carathéodory solution and at most one backward Carathéodory solution.*

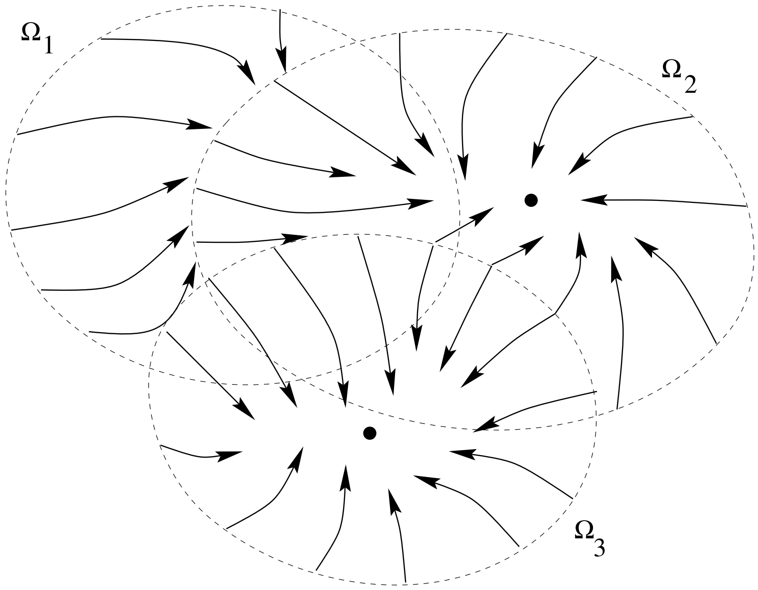

figure 4

(iii) *The set of Carathéodory solutions of (15) is closed. More precisely, assume that* $x_\nu : [a_\nu, b_\nu] \mapsto \Omega$ *is a sequence of solutions and, as* $\nu \to \infty$*, there holds*

$$a_\nu \to a, \qquad b_\nu \to b, \qquad x_\nu(t) \to \hat{x}(t) \textit{ for all } t \in]a, b[\,.$$

Then $\hat{x}(\cdot)$ *is itself a Carathéodory solution of (15).*

(iv) *The set of a Carathéodory solutions of the Cauchy problem (13) coincides with the set of forward Euler solutions.*

Proof. We sketch the main arguments in the proof. For details see [1].

To prove (i), observe that on any compact interval $[a, b]$ a solution $x(\cdot)$ can intersect only finitely many domains Ω_α, say those with indices $\alpha_1 < \alpha_2 < \cdots < \alpha_m$. It is now convenient to argue by backward induction. Since Ω_{α_m} is positively invariant for the flow of g_{α_m}, the set of times $\{t \in [a, b] : x(t) \in \Omega_{\alpha_m}\}$ must be a (possibly empty) interval of the form $]t_m, b]$. Similarly, the set $\{t \in [a, b] : x(t) \in \Omega_{\alpha_{m-1}}\}$ is an interval of the form $]t_{m-1}, t_m]$. After m inductive steps we conclude that

$$\dot{x}(t) = g_{\alpha_j}(x(t)) \qquad t \in]t_j, t_{j+1}[$$

for some times t_j with $a = t_1 \leq t_2 \leq \cdots \leq t_{m+1} = b$. All statements in (i) now follow from this fact. In particular, (20) holds because each set Ω_α is open and positively invariant for the flow of the corresponding vector field g_α.

Concerning (ii), to prove the local existence of a forward Carathéodory solution, consider the index

$$\bar{\alpha} \doteq \max\left\{\alpha \in \mathcal{A} : \bar{x} \in \overline{\Omega}_\alpha\right\}.$$

Because of the transversality condition (16), the solution of the Cauchy problem

$$\dot{x} = g_{\bar{\alpha}}(x), \qquad x(0) = \bar{x}$$

remains inside $\Omega_{\bar\alpha}$ for all $t \geq 0$. Hence it provides also a solution of (15) on some positive interval $[0, \delta]$.

To show the backward uniqueness property, let $x_1(\cdot)$, $x_2(\cdot)$ be any two Carathéodory solutions to (15) with $x_1(0) = x_2(0) = \bar x$. For $i = 1, 2$, call

$$\alpha_i^*(t) \doteq \max\{\alpha \in \mathcal{A} : x_i(t) \in \Omega_\alpha\}\ .$$

By (i), the maps $t \mapsto \alpha_i^*(t)$ are piecewise constant and left continuous. Hence there exists $\delta > 0$ such that

$$\alpha_1^*(t) = \alpha_2^*(t) = \bar\alpha \doteq \max\{\alpha \in \mathcal{A} : \bar x \in \Omega_\alpha\} \text{ for all } t \in]-\delta, 0]\ .$$

The uniqueness of backward solutions is now clear, because on $]-\delta, 0]$ both x_1 and x_2 are solutions of the same Cauchy problem with smooth coefficients

$$\dot x = g_{\bar\alpha}(x), \qquad x(0) = \bar x\ .$$

Concerning (iii), to prove that $\hat x(\cdot)$ is itself a Carathéodory solution, we observe that on any compact subinterval $J \subset]a, b[$ the functions u_ν are uniformly continuous and intersect a finite number of domains Ω_α, say with indices $\alpha_1 < \alpha_2 < \cdots < \alpha_m$. For each ν, the function

$$\alpha_\nu^*(t) \doteq \max\{\alpha \in \mathcal{A} : x_\nu(t) \in \Omega_\alpha\}$$

is non-decreasing and left continuous, hence it can be written in the form

$$\alpha_\nu^*(t) = \alpha_j \text{ if } t \in]t_j^\nu, t_{j+1}^\nu]\ .$$

By taking a subsequence we can assume that, as $\nu \to \infty$, $t_j^\nu \to \hat t_j$ for all j. By a standard convergence result for smooth O.D.E's, the function $\hat x$ provides a solution to $\dot x = g_{\alpha_j}(x)$ on each open subinterval $I_j \doteq]\hat t_j, \hat t_{j+1}[$. Since the domains Ω_β are open, there holds

$$\hat x(t) \notin \Omega_\beta \text{ for all } \beta > \alpha_j, \quad t \in I_j\ .$$

On the other hand, since g_{α_j} is inward pointing, a limit of trajectories $\dot x_\nu = g_{\alpha_j}(x_\nu)$ taking values within Ω_{α_j} must remain in the interior of Ω_{α_j}. Hence $\alpha^*\big(\hat x(t)\big) = \alpha_j$ for all $t \in I_j$, achieving the proof of (iii).

Regarding (iv), let $x_\varepsilon : [0, T] \mapsto \Omega$ be a sequence of forward Euler ε-approximate solutions of (13), converging to $\hat x(\cdot)$ as $\varepsilon \to 0$. To show that $\hat x$ is a Carathéodory solution, we first observe that, for $\varepsilon > 0$ sufficiently small, the maps $t \mapsto \alpha^*(x_\varepsilon(t))$ are non-decreasing. More precisely, there exist finitely many indices $\alpha_1 < \cdots < \alpha_m$ and times $0 = t_0^\varepsilon \leq t_1^\varepsilon \leq \cdots \leq t_m^\varepsilon = T$ such that

$$\alpha^*(x_\varepsilon(t)) = \alpha_j \qquad t \in]t_{j-1}^\varepsilon, t_j^\varepsilon]\ .$$

By taking a subsequence, we can assume $t_j^\varepsilon \to \hat t_j$ for all j, as $\varepsilon \to 0$. On each open interval $]\hat t_{j-1}, \hat t_j[$ the trajectory $\hat x$ is thus a uniform limit of polygonal approximate solutions of the smooth O.D.E.

$$\dot x = g_{\alpha_j}(x)\ . \tag{21}$$

By standard O.D.E. theory, $\hat x$ is itself a solution of (21). As in the proof of part (iii), we conclude observing that $\alpha^*\big(\hat x(t)\big) = \alpha_j$ for all $t \in]\hat t_{j-1}, \hat t_j]$.

To prove the converse, let $x : [0, T] \mapsto \Omega$ be a Carathéodory solution of (13). By (i), there exist indices $\alpha_1 < \cdots < \alpha_m$ and times $0 = t_0 < t_1 < \cdots < t_m = T$ such that $\dot x(t) = g_{\alpha_j}(x(t))$

for $t \in]t_{j-1}, t_j[$. For each $n \geq 1$, consider the polygonal map $x_n(\cdot)$ which is piecewise affine on the subintervals $[t_{j,k}, t_{j,k+1}]$, $j = 1, \ldots, m$, $k = 1, \ldots, n$ and takes values $x_n(t_{j,k}) = x_{j,k}$. The times $t_{j,k}$ and the values $x_{j,k}$ are here defined as

$$t_{j,k} \doteq t_{j-1} + \frac{k}{n}(t_j - t_{j-1}), \qquad x_{j,k} \doteq x(t_{j,k} + 2^{-n}).$$

As $n \to \infty$, it is now clear that $x_n \to x$ uniformly on $[0, T]$. On the other hand, for a fixed $\varepsilon > 0$ one can show that the polygonals $x_n(\cdot)$ are forward Euler ε-approximate solutions, for all $n \geq N_\varepsilon$ sufficiently large. This concludes the proof of part (iv).

□

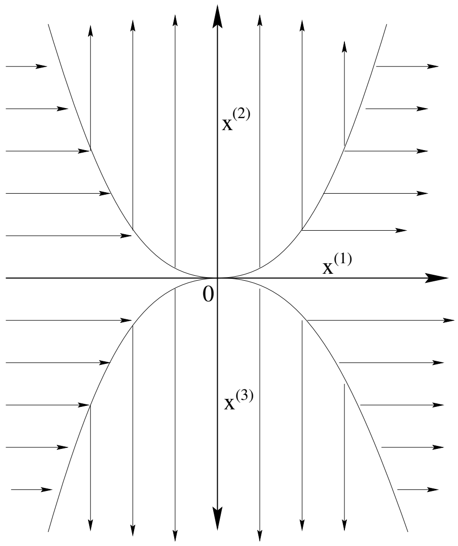

figure 5

EXAMPLE 5. Consider the patchy vector field on the plane (fig. 5) defined by (17), by taking

$$\Omega_1 \doteq \mathbb{R}^2, \qquad \Omega_2 \doteq \{x_2 > x_1^2\}, \qquad \Omega_3 \doteq \{x_2 < -x_1^2\},$$
$$g_1(x_1, x_2) \equiv (1, 0), \quad g_2(x_1, x_2) \equiv (0, 1), \quad g_3(x_1, x_2) \equiv (0, -1).$$

Then the Cauchy problem starting from the origin at time $t = 0$ has exactly three forward Carathéodory solutions, namely

$$x^{(1)}(t) = (t, 0), \qquad x^{(2)}(t) = (0, t), \qquad x^{(3)}(t) = (0, -t) \qquad t \geq 0.$$

The only backward Carathéodory solution is

$$x^{(1)}(t) = (t, 0) \qquad t \leq 0.$$

On the other hand there exist infinitely many Filippov solutions. In particular, for every $\tau < 0 < \tau'$, the function

$$x(t) = \begin{cases} (t - \tau, 0) & \text{if } t < \tau\,, \\ (0, 0) & \text{if } t \in [\tau, \tau']\,, \\ (t - \tau', 0) & \text{if } t > \tau' \end{cases}$$

provides a Filippov solution, and hence a Krasovskii solution as well.

5. Directionally continuous vector fields

Following [3], we say that a vector field g on $\mathbb{R}^n$ is *directionally continuous* if, at every point x where $g(x) \neq 0$ there holds

$$\lim_{n\to\infty} g(x_n) = g(x) \tag{22}$$

for every sequence $x_n \to x$ such that

$$\left| \frac{x_n - x}{|x_n - x|} - \frac{g(x)}{|g(x)|} \right| < \delta \text{ for all } n \geq 1\,. \tag{23}$$

Here $\delta = \delta(x) > 0$ is a function uniformly positive on compact sets. In other words (fig. 6), one requires $g(x_n) \to g(x)$ only for the sequences converging to x contained inside a cone with vertex at x and opening δ around an axis having the direction of $g(x)$.

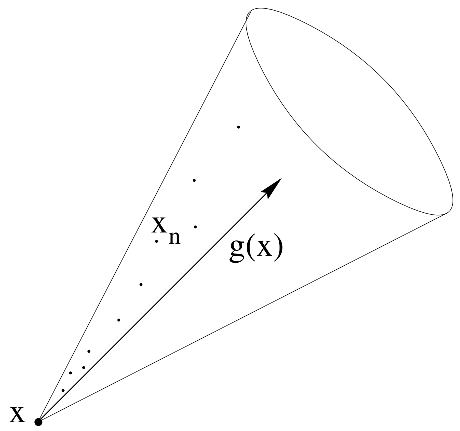

figure 6

For these vector fields, the local existence of Carathéodory trajectories is known [16]. It seems natural to ask whether the stabilization problems (AS) or (SOC) can be solved in terms of feedback controls generating a directionally continuous vector field. The following lemma reduces the problem to the construction of a patchy vector field.

LEMMA 2. *Let* $\big(\Omega, g, (\Omega_\alpha, g_\alpha)_{\alpha\in\mathcal{A}}\big)$ *be a patchy vector field. Then the map* $\tilde g$ *defined by*

$$\tilde g(x) = g_\alpha(x) \text{ if } x \in \overline{\Omega}_\alpha \setminus \bigcup_{\beta>\alpha} \overline{\Omega}_\beta \tag{24}$$

is directionally continuous. Every Carathéodory solution of

$$\dot x = \tilde g(x) \tag{25}$$

is also a solution of $\dot x = g(x)$. *The set of solutions of (25) may not be closed.*

Since directionally continuous vector fields form a much broader class of maps than patchy vector fields, solving a stabilization problem in terms of patchy fields thus provides a much better result. To see that the solution set of (25) may not be closed, consider

EXAMPLE 6. Consider the patchy vector field on $\mathbb{R}^2$ defined as follows.

$$\Omega_1 \doteq \mathbb{R}^2, \quad \Omega_2 \doteq \{x_2 < 0\}, \qquad g_1(x_1, x_2) = (1, 0), \quad g_2(x_1, x_2) = (0, -1)\,.$$

$$g(x_1, x_2) = \begin{cases} (1,0) & \text{if } x_2 \geq 0\,, \\ (0,-1) & \text{if } x_2 < 0\,. \end{cases} \tag{26}$$

The corresponding directionally continuous field is (fig. 7)

$$\tilde{g}(x_1, x_2) = \begin{cases} (1,0) & \text{if } x_2 > 0\,, \\ (0,-1) & \text{if } x_2 \leq 0\,. \end{cases} \tag{27}$$

The functions $t \mapsto x_\varepsilon(t) = (t, \varepsilon)$ are trajectories of both (26) and (27). However, as $\varepsilon \to 0$, the limit function $t \mapsto x(t) = (t, 0)$ is a trajectory of (26) but not of (27).

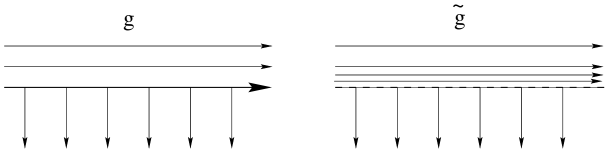

figure 7

6. Stabilizing feedback controls

In this section we discuss the applicability of the previous theory of discontinuous O.D.E's toward the construction of a stabilizing feedback. We first recall a basic definition [7, 18].

DEFINITION 3. *The system (1) is said to be globally* asymptotically controllable *to the origin if the following holds.*

1 - Attractivity. *For each $\bar{x} \in \mathbb{R}^n$ there exists some admissible control $u = u^{\bar{x}}(t)$ such that the corresponding solution of*

$$\dot{x}(t) = f\left(x(t), u^{\bar{x}}(t)\right), \qquad x(0) = \bar{x} \tag{28}$$

either tends to the origin as $t \to \infty$ or reaches the origin in finite time.

2 - Lyapunov stability. *For each $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. For every $\bar{x} \in \mathbb{R}^n$ with $|\bar{x}| < \delta$ there is an admissible control $u^{\bar{x}}$ as in* **1.** *steering the system from $\bar{x}$ to the origin, such that the corresponding trajectory of (28) satisfies $|x(t)| < \varepsilon$ for all $t \geq 0$.*

The next definition singles out a particular class of piecewise constant feedback controls, generating a "patchy" dynamics.

DEFINITION 4. *Let $\left(\Omega, g, (\Omega_\alpha, g_\alpha)_{\alpha\in\mathcal{A}}\right)$ be a patchy vector field. Assume that there exist control values $k_\alpha \in K$ such that, for each $\alpha \in \mathcal{A}$*

$$g_\alpha(x) \doteq f(x, k_\alpha) \text{ for all } x \in \Omega_\alpha \setminus \bigcup_{\beta>\alpha} \Omega_\beta\,. \tag{29}$$

Then the piecewise constant map

$$U(x) \doteq k_\alpha \text{ if } x \in \Omega_\alpha \setminus \bigcup_{\beta>\alpha} \Omega_\beta \tag{30}$$

is called a patchy feedback *control on* Ω.

The main results concerning stabilization by discontinuous feedback controls can be stated as follows. For the proofs, see [7] and [1] respectively.

THEOREM 4. *If the system (1) is asymptotically controllable, then there exists a feedback control* $U : \mathbb{R}^n \setminus \{0\} \mapsto K$ *such that every uniform limit of sampling solutions either tends asymptotically to the origin, or reaches the origin in finite time.*

THEOREM 5. *If the system (1) is asymptotically controllable, then there exists a patchy feedback control U such that every Carathéodory solution of (2) either tends asymptotically to the origin, or reaches the origin in finite time.*

Proof. In view of part (iv) of Theorem 3, the result stated in Theorem 4 can be obtained as a consequence of Theorem 5. The main part of the proof of Theorem 5 consists in showing that, given two closed balls $B' \subset B$ centered at the origin, there exists a patchy feedback that steers every point $\bar{x} \in B$ inside B' within finite time. The basic steps of this construction are sketched below. Further details can be found in [1].

1. By assumption, for each point $\bar{x} \in B$, there exists an open loop control $t \mapsto u^{\bar{x}}(t)$ that steers the system from $\bar{x}$ into a point x' in the interior of B' at some time $\tau > 0$. By a density and continuity argument, we can replace $u^{\bar{x}}$ with a piecewise constant open loop control $\bar{u}$ (fig. 8), say

$$\bar{u}(t) = k_\alpha \in K \text{ if } t \in]t_\alpha, t_{\alpha+1}],$$

for some finite partition $0 = t_0 < t_1 < \cdots < t_m = \tau$. Moreover, it is not restrictive to assume that the corresponding trajectory $t \mapsto \gamma(t) \doteq x(t;\ \bar{x}, \bar{u})$ has no self-intersections.

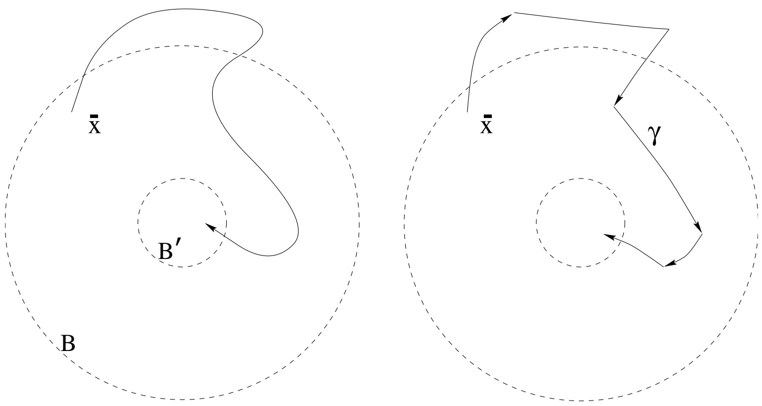

figure 8

2. We can now define a piecewise constant feedback control $u = U(x)$, taking the constant values $k_{\alpha_1}, \ldots, k_{\alpha_m}$ on a narrow tube Γ around γ, so that all trajectories starting inside Γ eventually reach the interior of B' (fig. 9).

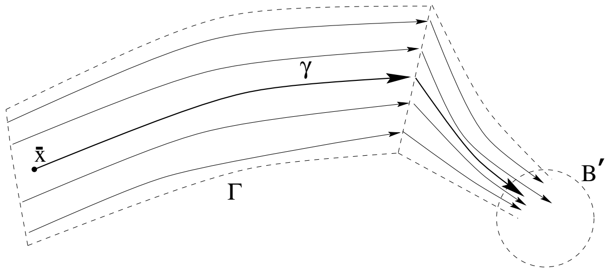

figure 9

3. By slightly bending the outer surface of each section of the tube Γ, we can arrange so that the vector fields $g_\alpha(x) \doteq f(x, k_\alpha)$ point strictly inward along the portion $\partial\Omega_\alpha \setminus \Omega_{\alpha+1}$. Recalling Lemma 1, we thus obtain a patchy vector field (fig. 10) defined on a small neighborhood of the tube Γ, which steers all points of a neighborhood of $\bar{x}$ into the interior of B'.

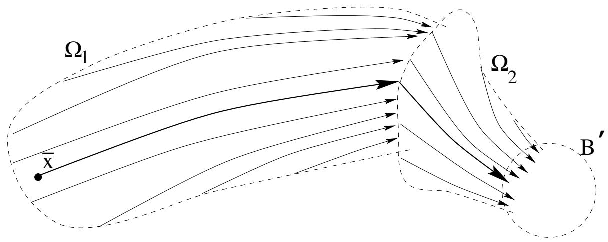

figure 10

4. The above construction can be repeated for every point $\bar{x}$ in the compact set B. We now select finitely many points $x_1, \ldots, x_N$ and patchy vector fields, $\Big(\Omega_i, g_i, (\Omega_{i,\alpha}, g_{i,\alpha})_{\alpha\in\mathcal{A}_i}\Big)$ with the properties that the domains Ω_i cover B, and that all trajectories of each field g_i eventually reach the interior of B'. We now define the patchy feedback obtained by the superposition of the g_i, in lexicographic order:

$$g(x) = g_{i,\alpha}(x) \text{ if } x \in \Omega_{i,\alpha} \setminus \bigcup_{(j,\beta)\succ(i,\alpha)} \Omega_{j,\beta}\,.$$

This achieves a patchy feedback control (fig. 11) defined on a neighborhood of $B \setminus B'$ which steers each point of B into the interior of B'.

5. For every integer ν, call B^ν be the closed ball centered at the origin with radius $2^{-\nu}$. By the previous steps, for every ν there exists a patchy feedback control U_ν steering each point in B_ν inside $B_{\nu+1}$, say

$$U_\nu(x) = k_{\nu,\alpha} \text{ if } x \in \Omega_{\nu,\alpha} \setminus \bigcup_{\beta>\alpha} \Omega_{\nu,\beta}\,. \tag{31}$$

The property of Lyapunov stability guarantees that the family of all open sets $\{\Omega_{\nu,\alpha} : \nu \in \mathbb{Z}, \alpha = 1, \ldots, N_\nu\}$ forms a locally finite covering of $\mathbb{R}^n \setminus \{0\}$. We now define the patchy feedback control

$$U_\nu(x) = k_{\nu,\alpha} \text{ if } x \in \Omega_{\nu,\alpha} \setminus \bigcup_{(\mu,\beta)\succ(\nu,\alpha)} \Omega_{\mu,\beta}\,, \tag{32}$$

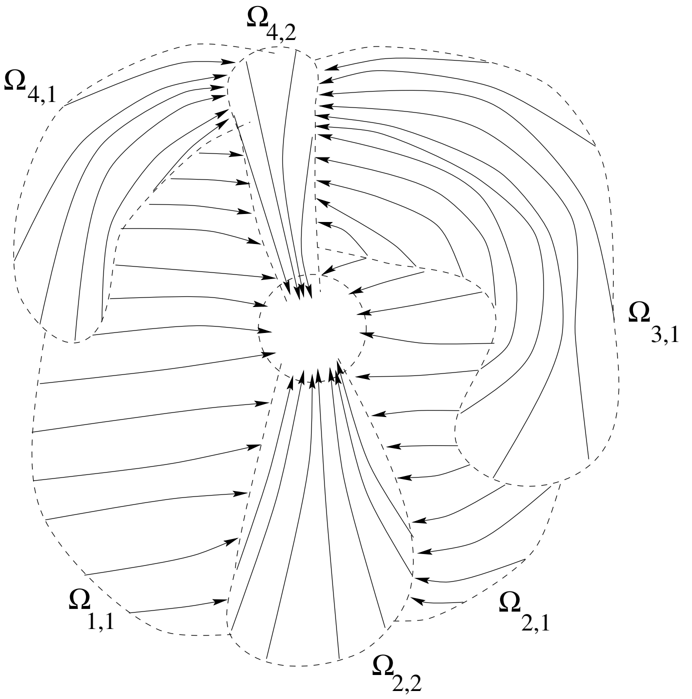

figure 11

where the set of indices (ν, α) is again ordered lexicographically. By construction, the patchy feedback (32) steers each point $x \in B^\nu$ into the interior of the smaller ball $B^{\nu+1}$ within finite time. Hence, every trajectory either tends to the origin as $t \to \infty$ or reaches the origin in finite time.

□

7. Some open problems

By Theorem 5, the asymptotic stabilization problem can be solved within the class of patchy feedback controls. We conjecture that the same is true for the problem of suboptimal controllability to zero.

Conjecture 1. Consider the smooth control system (1). For a fixed $\tau > 0$, call $R(\tau)$ the set of points that can be steered to the origin within time τ. Then, for every $\varepsilon > 0$, there exists a patchy feedback $u = U(x)$, defined on a neighborhood V of $R(\tau)$, with the following property. For every $\bar{x} \in V$, every trajectory of (2) starting at $\bar{x}$ reaches a point inside B_ε within time $T(\bar{x}) + \varepsilon$.

Although the family of patchy vector fields forms a very particular subclass of all discontinuous maps, the dynamics generated by such fields may still be very complicated and structurally unstable. In this connection, one should observe that the boundaries of the sets Ω_α may be taken in generic position. More precisely, one can slightly modify these boundaries so that the following property holds. If $x \in \partial\Omega_{\alpha_1} \cap \cdots \cap \partial\Omega_{\alpha_m}$, then the unit normals $\mathbf{n}_{\alpha_1}, \ldots, \mathbf{n}_{\alpha_m}$ are linearly independent. However, since no assumption is placed on the behavior of a vector field g_α at boundary points of a different domain Ω_β with $\beta \neq \alpha$, even the local behavior of the set of trajectories may be quite difficult to classify. More detailed results may be achieved for the special case of planar systems with control entering linearly:

$$\dot{x} = \sum_{i=1}^{m} f_i(x)\, u_i, \qquad u = (u_1, \ldots, u_m) \in K\,, \tag{33}$$

where $K \subset \mathbb{R}^m$ is a compact convex set. In this case, it is natural to conjecture the existence of stabilizing feedbacks whose dynamics has a very limited set of singular points. More precisely, consider the following four types of singularities illustrated in fig. 12. By a *cut* we mean a smooth curve γ along which the field g has a jump, pointing outward from both sides. At points at the of a cut, the field g is always tangent to γ. We call the endpoint an *incoming edge* or an *outgoing edge* depending on the orientation of g. A point where three distinct cuts join is called a *triple point*. Notice that the Cauchy problem with initial data along a cut, or an incoming edge of a cut, has two forward local solutions. Starting from a triple point there are three forward solutions.

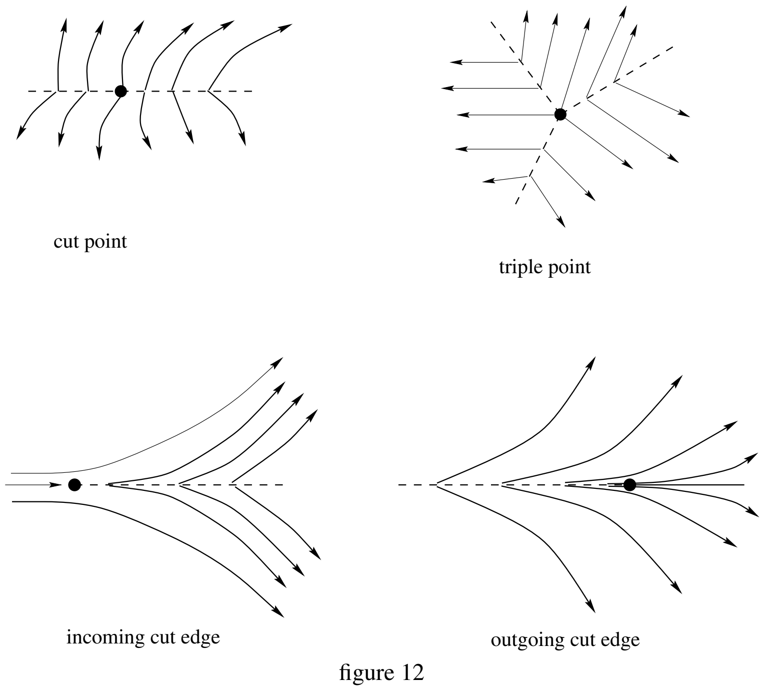

figure 12

Conjecture 2. Let the planar control system (33) be asymptotically controllable, with smooth coefficients. Then both the asymptotic stabilization problem (AS) and the suboptimal zero controllability problem (SOC) admit a solution in terms of a feedback $u = U(x) = (U_1(x), \ldots, U_n(x)) \in K$, such that the corresponding vector field

$$g(x) \doteq \sum_{i=1}^{m} f_i(x) U_i(x)$$

has singularities only of the four types described in fig. 12.

References

[1] ANCONA F., BRESSAN A., *Patchy vector fields and asymptotic stabilization*, preprint SISSA, Trieste 1998.

[2] AUBIN J. P., CELLINA A., *Differential Inclusions*, Springer Verlag, Berlin 1984.

[3] BRESSAN A., COLOMBO G., *Existence and continuous dependence for discontinuous O.D.E's*, Boll. Un. Mat. Ital. **4-B** (1990), 295–311.

[4] BRESSAN A., PICCOLI B., *A generic classification of optimal planar stabilizing feedbacks*, S.I.A.M. J. Control Optim. **36** (1998), 12–32.

[5] BRESSAN A., SHEN W., *On discontinuous differential equations*, in "Differential Inclusions and Optimal Control", J. Andres, L. Gorniewicz and P. Nistri Eds., Lecture Notes in Nonlinear Analysis, Vol. 2, 1998.

[6] BROCKETT R. W., *Asymptotic stability and feedback stabilization*, in "Differential Geometric Control Theory", R. W. Brockett, R. S. Millman and H. J. Sussmann Eds., Birkhauser, Boston 1983, 181–191.

[7] CLARKE F. H., LEDYAEV YU. S., SONTAG E. D., SUBBOTIN A. I., *Asymptotic Controllability Implies Feedback Stabilization*, IEEE Trans. Autom. Control **42** (1997), 1394–1407.

[8] CLARKE F. H., LEDYAEV YU. S., STERN R. J., WOLENSKI P. R., *Qualitative Properties of Trajectories of Control Systems: A survey*, J. Dynamic. Control Systems **1** (1995), 1–47.

[9] COLOMBO G., *On Extremal Solutions of Differential Inclusions*, Bull. Polish. Acad. Sci. **40** (1992), 97–109.

[10] CORON J. M., ROSIER L., *A Relation between Continuous Time-Varying and Discontinuous Feedback Stabilization*, J. Math. Systems, Estimation, and Control **4** (1994), 67–84.

[11] FILIPPOV A. F., *Differential equations with discontinuous right-hand side*, Amer. Math. Soc. Transl. **42** (1964), 199–231.

[12] FILIPPOV A. F., *Differential Equations with Discontinuous Right-Hand Sides*, Kluwer Academic Publisher 1988.

[13] HAJEK O., *Discontinuous differential equations I - II*, J. Differential Equations **32** (1979), 149–170 and 171–185.

[14] HERMES H., *Discontinuous Vector Fields and Feedback Control*, in "Differential Equations and Dynamical Systems", J. K. Hale and J. P. La Salle Eds., Academic Press, New York 1967, 155–165.

[15] PICCOLI B., *Classification of generic singularities for the planar time optimal synthesis*, S.I.A.M. J. Control Optim. **95** (1996), 59–79.

[16] PUCCI A., *Traiettorie di campi vettori discontinui*, Rend. Ist. Mat. Univ. Trieste **8** (1976), 84–93.

[17] RYAN E. P., *On Brockett's Condition for Smooth Stabilizability and its Necessity in a context of Nonsmooth Feedback*, SIAM J. Control Optim. **32** (1994), 1597–1604.

[18] SONTAG E. D., *Mathematical Control Theory, Deterministic Finite Dimensional Systems*, Springer-Verlag, New York 1990.

[19] SUSSMANN H. J., *Subanalytic sets and feedback control*, J. Differential Equations **31** (1979), 31–52.

[20] SONTAG E. D., SUSSMANN H. J., *Remarks on continuous feedback*, in "Proc. IEEE Conf. Decision and Control", Aulbuquerque, IEEE Publications, Piscataway 1980, 916–921.

AMS Subject Classification: ???.

Alberto BRESSAN
S.I.S.S.A.
Via Beirut 4
Trieste 34014 Italy

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

F. Bucci

THE NON-STANDARD LQR PROBLEM FOR BOUNDARY CONTROL SYSTEMS[†]

Abstract.

An overview of recent results concerning the non-standard, finite horizon Linear Quadratic Regulator problem for a class of boundary control systems is provided.

1. Introduction

In the present paper we give an account of recent results concerning the regulator problem with non-coercive, quadratic cost functionals over a finite time interval, for a class of abstract linear systems in a Hilbert space X, of the form

$$\begin{cases} x'(t) = Ax(t) + Bu(t), & 0 \le \tau < t < T \\ x(\tau) = x_0 \in X\,. \end{cases} \tag{1}$$

Here, A (free dynamics operator) is at least the generator of a strongly continuous semigroup on X, and B (input operator) is a linear operator subject to a suitable regularity assumption. The control function u is L^2 in time, with values in a Hilbert space U. Through the abstract assumptions on the operators A and B, a class of partial differential equations, with boundary/point control, is identified. We shall mostly focus our attention on systems which satisfy condition $(H2) = (8)$, see §1.2 below. It is known ([13]) that this condition amounts to a trace regularity property which is fulfilled by the solutions to a variety of hyperbolic (hyperbolic-like) partial differential equations.

With system (1), we associate the following cost functional

$$J_{\tau,T}(x_0; u) = \int_\tau^T F(x(t), u(t))\,\mathrm{d}t + \langle P_T x(T), x(T)\rangle\,, \tag{2}$$

where F is a continuous quadratic form on $X \times U$,

$$F(x, u) = \langle Qx, x\rangle + \langle Su, x\rangle + \langle x, Su\rangle + \langle Ru, u\rangle\,, \tag{3}$$

and $x(t) = x(t; \tau, x_0, u)$ is the solution to system (1) due to $u(\cdot) \in L^2(\tau, T; U)$. It is asked to provide conditions under which, for each $x_0 \in X$, a constant $c_{\tau,T}(x_0)$ exists such that

$$\inf_{u \in L^2(\tau,T;U)} J_{\tau,T}(x_0; u) \ge c_{\tau,T}(x_0)\,. \tag{4}$$

[†]This research was supported by the Italian Ministero dell'Università e della Ricerca Scientifica e Tecnologica within the program of GNAFA–CNR.

The special, important case, where

$$S = 0, \quad Q,\ P_T \geq 0, \quad R \geq \gamma > 0\,, \tag{5}$$

is now referred to as the classical (or, *standard*) LQR problem. We can say that this problem is now pretty well understood even for boundary control systems: the corresponding Riccati operator yields the synthesis of the optimal control (see [13]).

Functionals which do not display property (5) arise in different fields of systems/control theory. To name a few, the study of dissipative systems ([25]), where typical cases are

$$F(x,u) = |u|^2 - |x|^2\,, \qquad F(x,u) = \langle x, u\rangle\,;$$

the analysis of second variations of nonlinear control problems; H_∞ theory. It is worth recalling that the theory of infinite horizon linear quadratic control developed in [17], including the case of *singular* functionals, with $R = 0$, has more recently provided new insight in the study of the standard LQR problem for special classes of boundary control systems, see [21, 12].

In conclusion, the characterization of property (4), in a more general framework than the one defined by (5), is the object of the *non-standard* Linear Quadratic Regulator (LQR) problem.

Most results of the theory of the non-standard LQR problem for finite dimensional systems have been extended to the boundary control setting. We shall see that in particular, necessary conditions or sufficient conditions in order that (4) is satisfied can be provided, in term of non-negativity of suitable functionals. Unlike the infinite time horizon case, a gap still remains between necessary (non-negativity) conditions and sufficient (non-negativity) conditions, even when system (1) is exactly controllable. We shall examine this issue more in detail in §3.

The infinite dimensional problem reveals however new distinctive features. It is well known that in the finite dimensional case, the condition $R \geq 0$ has long been recognized as necessary in order that (4) is fulfilled; this applies even to time-dependent systems, see [7]. This property extends to infinite dimensional systems, when $P_T = 0$ (see [14, 6]). In contrast, in [6] an example is provided where, in spite of the fact that R is negative definite, the cost functional is coercive in $L^2(0,T;U)$, so that (4) is obviously satisfied. Crucially in that example $P_T \neq 0$, while the dynamics is given by a first order hyperbolic equation in one dimension, with control acted on the boundary.

Finally, we note that over an infinite horizon, the non-negativity condition which is necessary (and sufficient, under controllability of system (1)) for boundedness from below of the cost, is in fact equivalent to a suitable frequency domain inequality, (15) in §2, whose validity can be easily checked. In contrast, when T is finite, there is a lack of a frequency domain interpretation of the conditions provided.

The plan of the paper is the following. In §1.1 we provide a brief outline of the literature concerning the non-standard, finite horizon LQR problem for infinite dimensional systems. In §1.2 we introduce the abstract assumptions which characterize the class of dynamics of interest. In §2 we derive necessary conditions in order that (4) is satisfied, whereas §3 contains the statement of sufficient conditions. Most results of §2 and §3 are extracted from [6].

1.1. Literature

In this section we would like to provide a broad outline of contributions to the non-standard, *finite* horizon LQR problem for infinite dimensional systems. For a review of the richest literature on the same problem over an *infinite* horizon, we refer to [20]. We just recall that most recent extensions to the boundary control setting are given in [11], [14, Ch. 9], [18, 22, 23, 24].

Application to stability of holomorphic semigroup systems with boundary input is obtained, e.g., in [4].

The LQR problem with non-coercive functionals over a finite time interval has been the object of research starting around the 1970s. The most noticeable contribution to the study of this problem has been given, in our opinion, in [19]. For a comprehensive account of the theory developed in a finite dimensional context, and an extensive list of references, we refer to the monography [7].

The first paper which deals with the non-standard LQR problem over a finite time interval in infinite dimensions is, to our knowledge, [27]. The author considers dynamics of the form (1), which model distributed systems, with distributed control. Partial results are provided in order to characterize (4), without constraints on the form (3), except for $S = 0$. Moreover, the issue of the existence (and uniqueness) of an optimal control is considered, under the additional assumption that R is coercive.

A paper which deals with minimization of possible non-convex and non-coercive functionals, in a context which is more general than ours, is [1]. Necessary conditions or sufficient conditions for the existence of minimizers are stated therein, which involve a suitable 'recession functional' associated with the original functional.

In [9], the analysis is again restricted to cost functionals for which $R = I$, $S = 0$ (Q, P_T are allowed indefinite). Since R is coercive, the issue of the existence of solutions to the Riccati equation associated with the control problem is investigated. A new feature of the non-standard problem is pointed out, that the existence of an optimal control is not equivalent to the existence of a solution to the Riccati equation on $[0, T]$.

The study of the LQR problem with general cost functionals, still in the case of distributed systems with distributed control, has been carried out in [5]. Extensions of most finite dimensional results of [19] are provided. The application of the Bellman optimality principle to the infimization problem leads to introduce a crucial integral operator inequality, the so called 'Dissipation Inequality',

$$\langle P(a)x(a), x(a)\rangle \le \langle P(b)x(b), x(b)\rangle + \int_a^b F(x(s), u(s))\,\mathrm{d}s\,, \quad \tau \le a < b \le T\,,$$

whose solvability is equivalent to (4). Moreover, in [5] the regularity properties of the value function

$$V(\tau; x_0) = \inf_{u\in L^2(\tau,T;U)} J_{\tau,T}(x_0; u) \tag{6}$$

of the infimization problem are investigated, and new results are provided in this direction. In particular, it is showed that – unlike the standard case – the function $\tau \to V(\tau; x_0)$ is in general only upper semicontinuous on $[0, T]$, and that lack of continuity in the interior of $[0, T]$ may occurr, for instance, in the case of delay systems.

We remark that in all the aforementioned papers [27, 9, 5], among necessary conditions for finiteness of (6), a basic non-negativity condition is provided, namely (13) below, which in turn implies $R \ge 0$. On the other hand, sufficient conditions are so far given in a form which requires coercivity of the operator R.

Finally, more recently, extensions to the boundary control setting have been provided for a class of holomorphic semigroup systems ([14, Ch. 9], [26]), and for a class of 'hyperbolic-like' dynamics ([6]), respectively. We note that in [14] and [26] a greater emphasis is still placed on the *non-singular* case, since R is assumed coercive.

1.2. Notations, basic assumptions and abstract classes of dynamics

As explained in the introduction, we consider systems of the form (1) in abstract spaces of infinite dimension. A familiarity with the representation of controlled infinite dimensional systems is assumed, compatible with, e.g., [2].

Most notation used in the paper is standard. We just point out that inner products in any Hilbert space are denoted by $\langle\cdot,\cdot\rangle$; norms and operator norms are denoted by the symbols $|\cdot|$ and $\|\cdot\|$, respectively. The linear space of linear, bounded operators from X to Y is denoted by $\mathcal{L}(X,Y)$ ($\mathcal{L}(X)$, if $X=Y$).

Throughout the paper we shall make the following standing assumptions on the state equation (1) and the cost functional (2):

(i) *$A : D(A) \subset X \to X$ is the generator of a strongly continuous (s.c.) semigroup e^{At} on X, $t>0$;*

(ii) *$B \in \mathcal{L}(U, (D(A^*))')$; equivalently,*

$$A^{-\gamma}B \in \mathcal{L}(U,X) \text{ for some constant } \gamma \in [0,1]\,. \tag{7}$$

(iii) *Q, $P_T \in \mathcal{L}(X)$, $S \in \mathcal{L}(U,X)$, $R \in \mathcal{L}(U)$; Q, P_T, R are selfadjoint.*

REMARK 1. Assumptions (i)-(ii) identify dynamics which model distributed systems with distributed/boundary/point control. More specifically, the case of distributed control leads to a bounded input operator B, namely $\gamma=0$ in (ii), whereas $\gamma>0$ refers to the more challenging case of boundary/point control.

In order to characterize two main classes of partial differential equations problems of interest, roughly the 'parabolic' class and the 'hyperbolic' class, we follow [13] and introduce two distinct abstract conditions:

$(H1)$ *the s.c. semigroup e^{At} is analytic on X, $t>0$, and the constant γ appearing in (7) is strictly <1;*

$(H2)$ *there exists a positive constant k_T such that*

$$\int_0^T |B^* e^{A^*t}x|^2\,\mathrm{d}t \le k_T|x|^2 \qquad \forall x \in D(A^*)\,. \tag{8}$$

It is well known that under either $(H1)$ or $(H2)$, the (input-solution) operator

$$L_\tau : u \to (L_\tau u)(t) := \int_\tau^t e^{A(t-s)}Bu(s)\,rmds\,, \tag{9}$$

is continuous from $L^2(\tau,T;U)$ to $L^2(\tau,T;X)$. Consequently, system (1) admits a unique *mild* solution on (τ,T) given by

$$x(t) = e^{A(t-\tau)}x_0 + (L_\tau u)(t)\,, \tag{10}$$

which is (at least) L^2 in time. For a detailed analysis of examples of partial differential equations with boundary/point control which fall into either class, we refer to [13].

Let us recall that $(H2)$ is in fact equivalent to ([8])

$$L_\tau \textit{ continuous } : L^2(\tau,T;U) \to C(\tau,T;X)\,, \tag{11}$$

and that the following estimate holds true, for a positive constant $C_{\tau,T}$ and for any $u(\cdot)$ in $L^2(\tau, T; U)$:

$$|(L_\tau u)(t)| \le C_{\tau,T}\, |u|_{L^2(\tau,T;U)} \quad \forall t \in [\tau, T] \tag{12}$$

Therefore, for any initial datum $x_0 \in X$, the unique mild solution $x(\cdot; \tau, x_0, u)$ to equation (1), given by (10), is continuous on $[\tau, T]$, in particular at $t = T$. Thus, the term $\langle x(T), P_T\, x(T)\rangle$ makes sense for every control $u(\cdot) \in L^2(\tau, T; U)$.

REMARK 2. We note that $(H2)$, hence (11), follows as well from $(H1)$, when $\gamma \in [0, 1/2[$. Instead, when $(H1)$ holds with $\gamma \in [1/2, 1[$, counterexamples can be given to continuity of solutions at $t = T$, see [15, p. 202]. In that case, unless smoothing properties of P_T are required, the class of admissible controls need to be restricted. Comprehensive surveys of the theory of the standard LQR problem for systems subject to $(H1)$ are provided in [13] and [3]. Partial results for the corresponding non-standard regulator can be found in [14, Ch. 9].

In the present paper we shall mainly consider systems of the form (1) which satisfy assumption $(H2)$. This model covers many partial differential equations with boundary/point control, including, e.g., second order hyperbolic equations, Euler–Bernoulli and Kirchoff equations, the Schrödinger equation (see [13]).

2. Necessary conditions

In this section we are concerned with necessary conditions in order that (4) is satisfied, with special regard to the role of condition $R \ge 0$.

We begin with the statement of two basic necessary conditions, in the case of distributed systems with *distributed* control. For the sake of completeness, an outline of the proof is given; we refer to [5] for details. Condition (13) below is often referred to as the *non-negativity condition.*

THEOREM 1. *Assume that $B \in \mathcal{L}(U, X)$ (equivalently, $(H2)$ holds, with $\gamma = 0$). If there exist a $0 \le \tau < T$ and an $x_0 \in X$ such that (4) is satisfied, then*

$$J_{\tau,T}(0; u) \ge 0 \qquad \forall u \in L^2(\tau, T; U)\,, \tag{13}$$

which in turn implies

$$R \ge 0\,. \tag{14}$$

Sketch of the proof. For simplicity of exposition we assume that (4) is satisfied, with $\tau = 0$. In order to show that this implies (13), one first derives a representation of the cost $J_{0,T}(x_0; u)$ as a quadratic functional on $L^2(0, T; U)$, when x_0 is fixed, namely

$$J_{0,T}(x_0; u) = \langle \mathcal{M}x_0, x_0\rangle_X + 2\, Re\, \langle \mathcal{N}x_0, u\rangle_{L^2(0,T;U)} + \langle \mathcal{R}u, u\rangle_{L^2(0,T;U)}\,,$$

with $\mathcal{M}$, $\mathcal{N}$ and $\mathcal{R}$ suitable bounded operators. Readily $\langle \mathcal{R}u, u\rangle = J_{0,T}(0; u)$, and condition (13) follows from general results pertaining to infimization of quadratic functionals (see [5]).

Next, we use the actual expression of the operator $\mathcal{R}$ and the regularity of the input-solution operator L_0 defined by (9). Boundedness of the input operator B has here a crucial role. Proceeding by contradiction, (14) follows as a consequence of (13).

□

REMARK 3. A counterpart of Theorem 1 can be stated in infinite horizon, namely when $T = +\infty$ in (2) (set $P_T = 0$). In this case, if the semigroup e^{At} is not exponentially stable, the cost is not necessarily finite for an arbitrary control $u(\cdot) \in L^2(0, \infty; U)$. Consequently, the class of admissible controls need to be restricted. However, under stabilizability of the system (1), a non-negativity condition and (14) follow as well from (4). Even more, as remarked in the introduction, the non-negativity condition has a frequency domain counterpart ([17]), which in the stable case reads as

$$\begin{aligned}\Pi(i\omega) :=& B^*(-i\omega I - A^*)^{-1} Q(i\omega I - A)^{-1} B + S^*(i\omega I - A)^{-1} B \\ &+ B^*(-i\omega I - A^*)^{-1} S + R \geq 0 \quad \forall \omega \in \mathbb{R}\,.\end{aligned} \tag{15}$$

Theorem 1 can be extended to boundary control systems only in part.

THEOREM 2 ([6]). *Assume $(H2)$. Then the following statements hold true:*

(i) *if there exists an $x_0 \in X$ such that (4) is satisfied, then (13) holds;*

(ii) *if $P_T = 0$, then (13) implies (14); hence (14) is a necessary condition in order that (4) is satisfied.*

(iii) *if $P_T \neq 0$, then (14) is not necessary in order that (13) is satisfied.*

Sketch of the proof. Item (i) can be shown by using essentially the same arguments as in the proof of Theorem 1, which still apply to the present case, due to assumption $(H2)$. Similarly, when $P_T = 0$, (ii) follows as well.

The following example ([6, Ex. 4.4]) illustrates the third item. Let us consider, for a fixed $T \in (0, 1)$ and $\epsilon > 0$, the cost functional

$$J_{0,T}(x_0(\cdot); u) = \int_0^T \left\{ \int_T^1 |x(t, \xi)|^2 \, \mathrm{d}\xi - \epsilon |u(t)|^2 \right\} \mathrm{d}t + \int_0^T |x(T, \xi)|^2 \, \mathrm{d}\xi \,,$$

where $x(t, \xi)$ solves the boundary value problem

$$\begin{cases} x_t(t, \xi) = -x_\xi(t, \xi) & \\ x(0, \xi) = x_0(\xi) & 0 < \xi < 1 \\ x(t, 0) = u(t) & 0 < t < T\,. \end{cases} \tag{16}$$

Note that here $R = -\epsilon I$, $P_T = I$.

The solution to (16), corresponding to $x_0 \equiv 0$, is given by

$$x(t, \xi) = \begin{cases} 0 & t < \xi \\ u(t - \xi) & t > \xi\,, \end{cases} \tag{17}$$

so that

$$\begin{aligned} J_{0,T}(x_0 \equiv 0; u) &= -\epsilon \int_0^T |u(t)|^2 \, \mathrm{d}t + \int_0^T |u(T - \xi)|^2 \, \mathrm{d}\xi \\ &= (1 - \epsilon) \int_0^T |u(t)|^2 \, \mathrm{d}t\,. \end{aligned}$$

Therefore, if $0 < \epsilon < 1$, $J_{0,T}(0; u)$ is not only positive but even coercive in L^2, which implies (4). Nevertheless, $R < 0$.

□

A better result can be provided in the case of holomorphic semigroup systems, by using the smoothing properties of the operator L_τ. Somehow the 'analytic case' parallels the case when the input operator is bounded. See [14, Ch. 9, Theorem 3.1] for the proof.

THEOREM 3. *Assume that* $(H1)$ *holds, with* $\gamma < 1/2$. *Then (14) is a necessary condition in order that (4) is satisfied, even when* $P_T \neq 0$.

3. Sufficient conditions

In this section we provide sufficient conditions in order that (4) is satisfied. Let us go back to the non-negativity conditions of Theorem 1. It can be easily shown that neither (14), nor (13), are, by themselves, sufficient to guarantee that the cost functional is bounded from below.

EXAMPLE 1. Let $X = U = \mathbb{R}$, and set $\tau = 0$, $A = -1$, $B = 0$ in (1); moreover, let $F(x, u) = xu$. Note that here $R = 0$. For any x_0, the solution to (1) is given by $x(t) = x_0 e^{-t}$, so that

$$J_{0,T}(x_0; u) = x_0 \int_0^T e^{-t} u(t)\,\mathrm{d}t$$

for any admissible control u. Therefore $J_{0,T}(0; u) \equiv 0$, and (13) holds true, whereas it is readily verified that when $x_0 \neq 0$, $\inf_u J_{0,T}(x_0; u) = -\infty$ (if $x_0 > 0$ take, for instance, the sequence $u_k(t) = -k$ on $[0, T]$).

If $T = +\infty$, the same example shows that Theorem 1 cannot be reversed without further assumptions. However it turns out that, over an infinite horizon, the necessary non-negativity condition (13) is also sufficient in order that (4) is satisfied, if system (1) is completely controllable. This property is well known in the finite dimensional case, since the early work [10].

Recently, the aforementioned result has been extended to boundary control systems, under the following assumptions:

(i') *$A : D(A) \subset X \to X$ is the generator of a s.c. group e^{At} on X, $t \in \mathbb{R}$;*

$(H2')$ *there exists a $T > 0$ and a constant $k_T > 0$ such that*

$$\int_0^T |B^* e^{A^* t} x|^2\,\mathrm{d}t \leq k_T |x|^2 \qquad \forall x \in D(A^*)\,; \tag{18}$$

$(H3')$ *system (1) is completely controllable, namely for each pair $x_0, x_1 \in X$ there is a T and an admissibile control $v(\cdot)$ such that $x(T; 0, x_0, v) = x_1$.*

For simplicity of exposition, we state the theorem below under the additional condition that e^{At} is exponentially stable.

THEOREM 4 ([22]). *Assume* (i')–$(H2')$–$(H3')$. *If*

$$J_\infty(0; u) \geq 0 \qquad \forall u \in L^2(0, \infty; U)\,,$$

then for each $x_0 \in X$ there exists a constant $c_\infty(x_0)$ such that

$$\inf_{L^2(0,\infty;U)} J_\infty(x_0; u) \geq c_\infty(x_0) \qquad \forall u \in L^2(0, \infty; U)\,.$$

We now return to the finite time interval $[0, T]$ and introduce the assumption that the system is exactly controllable on a certain interval $[0, r]$ (see, e.g., [3]):

$(H3)$ *there is an* $r > 0$ *such that, for each pair* $x_0, x_1 \in X$*, there exists an admissibile control* $v(\cdot) \in L^2(0, r; U)$ *yielding* $x(r; 0, x_0, v) = x_1$*. Equivalently,*

$$\exists\, r,\ c_r > 0 : \quad \int_0^r |B^* e^{A^* t} x|^2 \, \mathrm{d}t \ge c_r |x|^2 \qquad \forall x \in D(A^*) \,. \tag{19}$$

On the basis of Theorem 4, one would be tempted to formulate the following claim.

CLAIM 5. Assume (i')–$(H2)$–$(H3)$. If

$$J_{0,T}(0; u) \ge 0 \qquad \forall u \in L^2(0, T; U) \,,$$

then (4) is satisfied for $0 \le \tau \le T$.

It turns out that this claim is false, as it can been shown by means of examples: see [7, 6].

A correct counterpart of Theorem 4 over a finite time interval has been given in [6].

THEOREM 6 ([6]). *Assume* (i')–$(H2)$–$(H3)$*. If*

$$J_{0,T+r}(0; u) \ge 0 \qquad \forall u \in L^2(0, T + r; U) \,, \tag{20}$$

then (4) is satisfied for $0 \le \tau \le T$.

We point out that in fact a proof of Theorem 6 can be provided which does not make *explicit* use of assumption $(i)'$, see Theorem 7 below. Let us recall however that, when the input operator B is bounded, controllability of the pair (A, B) on $[0, r]$, namely assumption $(H3)$ above, implies that the semigroup e^{At} is *right* invertible, [16]. Therefore, the actual need of some kind of 'group property' in Theorem 6 is an issue which is left for further investigation.

THEOREM 7. *Assume* $(H2)$–$(H3)$*. If (20) holds, then (4) is satisfied for* $0 \le \tau \le T$.

Proof. Let $x_1 \in X$ be given. By $(H3)$ there exists a control $v(\cdot) \in L^2(0, r; U)$ steering the solution of (1) from $x_0 = 0$ to x_1 in time r, namely $x(r; 0, 0, v) = x_1$. Obviously, v depends on x_1: more precisely, it can be shown that, as a consequence of assumptions $(H2)$ and $(H3)$, a constant K exists such that

$$|v|_{L^2(0,r;U)} \le K |x_1| \,,$$

see [6]. For arbitrary $u \in L^2(r, T + r; U)$, set now

$$u_v(t) = \begin{cases} v(t) & 0 \le t \le r \\ u(t) & r < t \le T + r \,. \end{cases}$$

Readily $u_v(\cdot) \in L^2(0, T + r; U)$, and $J_{0,T+r}(0; u_v) \ge 0$ due to (20). On the other hand,

$$J_{0,T+r}(0; u_v) = \int_0^r F(x(s; 0, 0, v), v(s)) \, \mathrm{d}s + J_{r,T+r}(x_1; u) \,,$$

where the first summand is a constant which depends only on x_1. A straightforward computation shows that the second summand equals $J_{0,T}(x_1; u_r)$, with $u_r(t) = u(t+r)$ an arbitrary admissible control on $[0, T]$. In conclusion,

$$J_{0,T}(x_1; u_r) \geq -\int_0^r F(x(s; 0, 0, v), v(s))\, \mathrm{d}s =: c(x_1)\,,$$

and (4) holds for $\tau = 0$. The case $\tau > 0$ can be treated by using similar arguments.

□

REMARK 4. In conclusion, we have provided the sufficiency counterpart of item (i) of Theorem 2, under the additional condition that system (1) is exactly controllable in time $r > 0$. Apparently, in order that (4) is satisfied, the non-negativity condition need to be required on a larger interval than $[0, T]$, precisely on an interval of lenght $T + r$. This produces a gap between necessary conditions and sufficient conditions, which was already pointed out in finite dimensions ([7]).

Finally, we stress that the exact controllability assumption cannot be weakened to null controllability, as pointed out in [6, Ex. 4.5].

References

[1] BAIOCCHI C., BUTTAZZO G., GASTALDI F., TOMARELLI F., *General existence theorems for unilateral problems in continuum mechanics*, Arch. Ration. Mech. Anal. **100** (2) (1988), 149–189.

[2] BENSOUSSAN A., DA PRATO G., DELFOUR M. C., MITTER S. K., *Representation and Control of Infinite Dimensional Systems*, Vol. I, Birkhäuser, Boston 1992.

[3] BENSOUSSAN A., DA PRATO G., DELFOUR M. C., MITTER S. K., *Representation and Control of Infinite Dimensional Systems*, Vol. II, Birkhäuser, Boston 1993.

[4] BUCCI F., *Frequency domain stability of nonlinear feedback systems with unbounded input operator*, Dynamics of Continuous, Discrete and Impulsive Systems (to appear).

[5] BUCCI F., PANDOLFI L., *The value function of the singular quadratic regulator problem with distributed control action*, SIAM J. Control Optim. **36** (1) (1998), 115–136.

[6] BUCCI F., PANDOLFI L., *The regulator problem with indefinite quadratic cost for boundary control systems: the finite horizon case*, Politecnico di Torino, Dipartimento di Matematica, Preprint N. 22, 1998.

[7] CLEMENTS D. J., ANDERSON B. D. O., *Singular optimal control: the linear-quadratic problem*, Lect. Notes in Control Inf. Sci. No.5, Springer Verlag, Berlin 1978.

[8] FLANDOLI F., LASIECKA I., TRIGGIANI R., *Algebraic Riccati equations with non-smoothing observation arising in hyperbolic and Euler-Bernoulli boundary control problems*, Ann. Mat. Pura Appl. **153** (1988), 307–382.

[9] JACOB B., *Linear quadratic optimal control of time-varying systems with indefinite costs on Hilbert spaces: the finite horizon problem*, J. Math. Syst. Estim. Control **5** (1995), 1–28.

[10] KALMAN R. E., *Lyapunov functions for the problem of Lur'e in automatic control*, Proc. Nat. Acad. Sci. USA **49** (1963), 201–205.

[11] VAN KEULEN B., *Equivalent conditions for the solvability of the nonstandard LQ-Problem for Pritchard-Salamon systems*, SIAM J. Contr. Optim. **33** (1995), 1326–1356.

[12] LASIECKA I., PANDOLFI L., TRIGGIANI R., *A singular control approach to highly damped second-order abstract equations and applications*, in: Control of Partial Differential Equations, (E. Casas Ed.), M. Dekker, New York 1995.

[13] LASIECKA I., TRIGGIANI R., *Differential and Algebraic Riccati Equations with Applications to Boundary/Point Control Problems: Continuous Theory and Approximation Theory*, Lect. Notes in Control Inf. Sci. No.164, Springer Verlag, Berlin 1991.

[14] LI X., YONG J., *Optimal control theory for infinite dimensional systems*, Birkhäuser, Basel 1995.

[15] LIONS J. L., *Optimal control of systems governed by Partial Differential Equations*, Springer Verlag, Berlin 1971.

[16] LOUIS J-CL., WEXLER D., *On exact controllability in Hilbert spaces*, J. Differ. Equations **49** (1983), 258–269.

[17] LOUIS J-CL., WEXLER D., *The Hilbert space regulator problem and operator Riccati equation under stabilizability*, Ann. Soc. Sci. Bruxelles, Ser.I, **105** (1991), 137–165.

[18] MCMILLAN C., *Equivalent conditions for the solvability of non-standard LQ-problems with applications to partial differential equations with continuous input-output solution map*, J. Math. Syst. Estim. Control **7** (3) (1997), 379–382 (short printed version); full electronic version 27 pages (retrieval code: 71413).

[19] MOLINARI B. P., *Nonnegativity of a quadratic functional*, SIAM J. Control **13** (1975), 92–806.

[20] PANDOLFI L., *The Kalman–Yakubovich–Popov Theorem: an overview and new results for hyperbolic control systems*, Nonlinear Anal. **30** (1997), 735–745.

[21] PANDOLFI L., *The standard regulator problem for systems with input delays: an approach through singular control theory*, Appl. Math. Optim. **31** (1995), 119–136.

[22] PANDOLFI L., *The Kalman–Yakubovich–Popov Theorem for stabilizable hyperbolic boundary control systems*, Integral Equations Operator Theory, to appear.

[23] PANDOLFI L., *Dissipativity and the Lur'e problem for parabolic boundary control systems*, SIAM J. Control Optim. **36** (1998), 2061–2081.

[24] WEISS M., WEISS G., *Optimal control of stable weakly regular linear systems*, Math. Control Signal Systems **10**, 287–330.

[25] WILLEMS J. C., *Dissipative dynamical systems, Part I: General theory, Part II: Linear systems with quadratic supply rates*, Arch. Ration. Mech. Anal. **45** (1972), 321–351 and 352–392.

[26] WU H., LI X., *Linear quadratic problem with unbounded control in Hilbert spaces*, Chin. Sci. Bull. **43** (20) (1998), 1712–1717.

[27] YOU Y., *Optimal control for linear system with quadratic indefinite criterion on Hilbert spaces*, Chin. Ann. Math., Ser. B 4, (1983), 21–32.

AMS Subject Classification: 49N10, 49J20.

Francesca BUCCI
Università degli Studi di Firenze
Dipartimento di Matematica Applicata "G. Sansone"
Via S. Marta 3, I–50139 Firenze, Italy
e-mail: fbucci@dma.unifi.it

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

F. Ceragioli

EXTERNAL STABILIZATION OF DISCONTINUOS SYSTEMS AND NONSMOOTH CONTROL LYAPUNOV-LIKE FUNCTIONS

Abstract.
The main result of this note is an external stabilizability theorem for discontinuous systems affine in the control (with solutions intended in the Filippov's sense). In order to get it we first prove a sufficient condition for external stability which makes use of nonsmooth Lyapunov-like functions.

1. Introduction

In this note we deal with discontinuous time-dependent systems affine in the control:

$$\dot{x} = f(t,x) + G(t,x)u = f(t,x) + \sum_{i=1}^{m} u_i g_i(t,x) \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $f \in L^\infty_{loc}(\mathbb{R}^{n+1};\mathbb{R}^n)$, for all $i \in \{1,\dots,m\}$, $g_i \in C(\mathbb{R}^{n+1};\mathbb{R}^n)$ and G is the matrix whose columns are $g_1,\dots,g_m$.

Admissible inputs are $u \in L^\infty_{loc}(\mathbb{R};\mathbb{R}^m)$.

Solutions of system (1) (as well as solutions of all the systems considered in the following) are intended in the Filippov's sense. In other words, for each admissible input $u(t)$, (1) is replaced by the differential inclusion

$$\dot{x} \in K(f+Gu)(t,x) = \bigcap_{\delta>0} \bigcap_{\mu(N)=0} \overline{\mathrm{co}}\,\{(f+Gu)(t, B(x,\delta)\backslash N)\}\,,$$

where $B(x,\delta)$ is the ball of center x and radius δ, $\overline{\mathrm{co}}$ denotes the convex closure and μ is the usual Lebesgue measure in $\mathbb{R}^n$.

For the general theory of Filippov's solutions we refer to [6]. We denote by $S_{t_0,x_0,u}$ the set of solutions $\varphi(\cdot)$ of system (1) with the initial condition $\varphi(t_0) = x_0$ and the function $u:\mathbb{R}\to\mathbb{R}^m$ as input.

We are interested in the external behaviour of system (1), in particular in its uniform bounded input bounded state (UBIBS) stability.

Roughly speaking a system is said to be UBIBS stable if its trajectories are bounded whenever the input is bounded. More precisely we have the following definition.

DEFINITION 1. *System (1) is said to be* UBIBS stable *if for each $R>0$ there exists $S>0$ such that for each $(t_0,x_0)\in\mathbb{R}^{n+1}$, $t_0\ge 0$, and each input $u\in L^\infty_{loc}(\mathbb{R};\mathbb{R}^m)$ one has*

$$\|x_0\| < R, \quad \|u\|_\infty < R \Rightarrow \forall\varphi(\cdot)\in S_{t_0,x_0,u} \quad \|\varphi(t)\| < S \quad \forall t \ge t_0\,.$$

We associate to system (1) the unforced system

$$\dot{x} = f(t, x) \tag{2}$$

obtained from (1) by setting $u = 0$. We denote $S_{t_0,x_0,0} = S_{t_0,x_0}$.

DEFINITION 2. *System (2) is said to be* uniformly Lagrange stable *if for each $R > 0$ there exists $S > 0$ such that for each $(t_0, x_0) \in \mathbb{R}^{n+1}$, $t_0 \geq 0$, one has*

$$\|x_0\| < R \Rightarrow \forall \varphi(\cdot) \in S_{t_0,x_0} \quad \|\varphi(t)\| < S \quad \forall t \geq t_0 .$$

If system (1) is UBIBS stable, then system (2) is uniformly Lagrange stable, but the converse is not true in general. In Section 3 we prove that, if not only system (2) is uniformly Lagrange stable, but some additional conditions on f and G are satisfied, then there exists an externally stabilizing feedback law for system (1), in the sense of the following definition.

DEFINITION 3. *System (1) is said to be* UBIBS stabilizable *if there exists a function $k \in L^{\infty}_{loc}(\mathbb{R}^{n+1}; \mathbb{R}^m)$ such that the closed loop system*

$$\dot{x} = f(t, x) + G(t, x)k(t, x) + G(t, x)v \tag{3}$$

(with v as input) is UBIBS stable.

The same problem has been previously treated in [1, 2, 4, 9]. We give our result (Theorem 2) and discuss the differences with the results obtained in the mentioned papers in Section 3.

In order to achieve Theorem 2 we need a preliminary theorem (Theorem 1 in Section 2). It is a different version of Theorem 1 in [13] and Theorem 6.2 in [4]. It provides a sufficient condition for UBIBS stability of system (1) by means of a nonsmooth control Lyapunov-like function. Finally the proof of the main result is given in Section 4.

2. UBIBS Stability

In this section we give a sufficient condition for UBIBS stability of system (1) by means of a nonsmooth control Lyapunov-like function. (See [11, 12] for control Lyapunov functions).

The following Theorem 1 (and also its proof) is analogous to Theorem 1 in [13] and Theorem 6.2 in [4]. It differs from both for the fact that it involves a control Lyapunov-like function which is not of class C^1, but just locally Lipschitz continuous and regular in the sense of Clarke (see [5], page 39).

DEFINITION 4. *We say that a function $V : \mathbb{R}^{n+1} \to \mathbb{R}$ is* regular *at $(t, x) \in \mathbb{R}^{n+1}$ if*

(i) *for all $v \in \mathbb{R}^n$ there exists the usual right directional derivative $V'_+((t, x), (1, v))$,*

(ii) *for all $v \in \mathbb{R}^n$, $V'_+((t, x), (1, v)) = \limsup_{(s,y)\to(t,x)\; h\downarrow 0} \frac{V(s+h, y+hv)-V(s,y)}{h}$.*

The fact that the control Lyapunov-like function for system (1) is regular allows us to characterize it by means of its set-valued derivative with respect to the system instead of by means of Dini derivatives.

Let us recall the definition of set-valued derivative of a function with respect to a system introduced in [10] and then used (with some modifications) in [3]. Let us denote by $\partial V(t, x)$ Clarke generalized gradient of V at (t, x) (see [5], page 27).

DEFINITION 5. *Let $t > 0$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ be fixed, $V : \mathbb{R}^{n+1} \to \mathbb{R}$. We call* set-valued derivative of V with respect to system (1) *the set*

$$\dot{\overline{V}}^{(1)}(t, x, u) = \{a \in \mathbb{R} : \exists v \in K(f(t, x) + G(t, x)u) \text{ such that } \forall p \in \partial V(t, x) \ p \cdot (1, v) = a\}.$$

Analagously, if $t > 0$, $x \in \mathbb{R}^n$, $u \in L^\infty_{loc}(\mathbb{R}, \mathbb{R}^m)$ are fixed, we set

$$\dot{\overline{V}}^{(1)}_{u(\cdot)}(t, x) = \{a \in \mathbb{R} : \exists v \in K(f(t, x) + G(t, x)u(t)) \text{ such that } \forall p \in \partial V(t, x) \ p \cdot (1, v) = a\}$$

and, if $t > 0$ and $x \in \mathbb{R}^n$ are fixed, we define

$$\dot{\overline{V}}^{(2)}(t, x) = \{a \in \mathbb{R} : \exists v \in Kf(t, x) \text{ such that } \forall p \in \partial V(t, x) \ p \cdot (1, v) = a\}.$$

Let us remark that $\dot{\overline{V}}^{(1)}(t, x, u)$ is a closed and bounded interval, possibly empty and

$$\max \dot{\overline{V}}^{(1)}(t, x, u) \leq \max_{v \in K(f(t,x)+G(t,x)u)} \overline{D^+} V((t, x), (1, v)),$$

where $\overline{D^+} V((t, x), (1, v))$ is the Dini derivative of V at (t, x) in the direction of $(1, v)$.

LEMMA 1. *Let $\varphi(\cdot)$ be a solution of the differential inclusion (1) corresponding to the input $u(\cdot)$ and let $V : \mathbb{R}^{n+1} \to \mathbb{R}$ be a locally Lipschitz continuous and regular function. Then $\frac{d}{dt} V(t, \varphi(t))$ exists almost everywhere and $\frac{d}{dt} V(t, \varphi(t)) \in \dot{\overline{V}}^{(1)}_{u(\cdot)}(t, \varphi(t))$ almost everywhere.*

We omit the proof of the previous lemma since it is completely analogous to the proofs of Theorem 2.2 in [10] (which involves a slightly different kind of set-valued derivative with respect to the system) and of Lemma 1 in [3] (which is given for autonomous differential inclusions and V not depending on time).

We can now state the main theorem of this section.

THEOREM 1. *Let $V : \mathbb{R}^{n+1} \to \mathbb{R}$ be such that there exists $L > 0$ such that*

(V0) *there exist two continuous, strictly increasing, positive functions $a, b : \mathbb{R} \to \mathbb{R}$ such that $\lim_{r \to +\infty} a(r) = +\infty$ and for all $t > 0$ and for all x*

$$\|x\| > L \Rightarrow a(\|x\|) \leq V(t, x) \leq b(\|x\|)$$

(V1) *V is locally Lipschitz continuous and regular in $\mathbb{R}^+ \times \{x \in \mathbb{R}^n : \|x\| > L\}$.*

If

(fG) *for all $R > 0$ there exists $\rho > L$ such that for all $x \in \mathbb{R}^n$ and for all $u \in \mathbb{R}^m$ the following holds:*

$$\|x\| > \rho, \ \|u\| < R \Rightarrow \max \dot{\overline{V}}^{(1)}(t, x, u) \leq 0 \text{ for a.e. } t \geq 0$$

then system (1) is UBIBS stable.

Proof. We prove the statement by contradiction, by assuming that there exists $\overline{R}$ such that for all $S > 0$ there exist $\overline{x_0}$ and $\overline{u} : [0, +\infty) \to \mathbb{R}^m$ such that $\|\overline{x_0}\| < \overline{R}$, $\|\overline{u}\|_\infty < \overline{R}$ and there exist $\overline{\varphi}(\cdot) \in S_{t_0, x_0, \overline{u}}$, and $\overline{t} > 0$ such that $\|\overline{\varphi}(\overline{t})\| \geq S$.

Let us choose $\overline{\rho}$ corresponding to $\overline{R}$ as in (fG). Without loss of generality we can suppose that $\overline{\rho} > \overline{R}$. Because of (V0), there exists $S_M > 0$ such that if $\|x\| > S_M$, then $V(t,x) > M = b(\overline{\rho}) \geq \max\{V(t,x),\ \|x\| = \overline{\rho},\ t \geq 0\}$ for all t.

Let us consider $S > \overline{\rho}$, S_M. There exist $t_1, t_2 > 0$ such that $\overline{t} \in [t_1, t_2]$, $\|\overline{\varphi}(t_1)\| = \overline{\rho}$, $\|\overline{\varphi}(t)\| \geq \overline{\rho}$ in $[t_1, t_2]$, $\|\overline{\varphi}(t_2)\| \geq S$. Then

$$V(t_2, \overline{\varphi}(t_2)) > M \geq V(t_1, \overline{\varphi}(t_1))\,. \tag{4}$$

On the other hand, by Lemma 1, $\frac{d}{dt}V(t,\overline{\varphi}(t)) \in \dot{\overline{V}}^{(1)}_{\overline{u}(\cdot)}(t,\overline{\varphi}(t))$ a.e. It is clear that $\dot{\overline{V}}^{(1)}_{\overline{u}(\cdot)}(t,\overline{\varphi}(t)) \subseteq \dot{\overline{V}}^{(1)}(t,\overline{\varphi}(t),\overline{u}(t))$. Since $|\overline{u}(t)| < \overline{R}$ a.e. and $\|\varphi(t)\| > \overline{\rho}$ for all $t \in [t_1, t_2]$, by virtue of (fG) we have $\frac{d}{dt}V(t,\overline{\varphi}(t)) \leq 0$ for a.e. $t \in [t_1, t_2]$. By [7] (page 207) we get that $V \circ \overline{\varphi}$ is decreasing in $[t_1, t_2]$, then

$$V(t_2, \overline{\varphi}(t_2)) \leq V(t_1, \overline{\varphi}(t_1))$$

that contradicts (4).

□

REMARK 1. In order to get a sufficient condition for system (2) to be uniformly Lagrange stable, one can state Theorem 1 in the case $u = 0$. In this case the control Lyapunov-like function simply becomes a Lyapunov-like function.

REMARK 2. For sake of simplicity we have given the definition of UBIBS stability and stated Theorem 1 for systems affine in the control. Let us remark that exactly analogous definition, theorem and proof hold for more general systems of the form

$$\dot{x} = f(t,x,u)$$

where $f : \mathbb{R}^{m+n+1} \to \mathbb{R}^n$ is locally bounded and measurable with respect to the variables t and x and continuous with respect to u.

REMARK 3. If system (1) is autonomous it is possible to state a theorem analogous to Theorem 1 for a control Lyapunov-like function V not depending on time.

3. The Main Result

The main result of this note is the following Theorem 2. It essentially recalls Theorem 6.2 in [4] and Theorem 5 in [9], with the difference that the control Lyapunov-like function involved is not smooth.

We don't give a unique condition for system (1) to be externally stabilizable, but some alternative conditions which, combined together, give the external stabilizability of the system. Before stating the theorem we list these conditions. Note that the variable x is not yet quantified. Since its role depend on different situations, it is convenient to specify it later.

(f1) $\max \dot{\overline{V}}^{(2)}(t,x) \leq 0$;

(f2) for all $z \in Kf(t,x)$ there exists $\overline{p} \in \partial V(t,x)$ such that $\overline{p} \cdot (1,z) \leq 0$;

(f3) for all $z \in Kf(t,x)$ and for all $p \in \partial V(t,x)$, $p \cdot (1,z) \leq 0$;

(G1) for each $i \in \{1,\ldots,m\}$ there exists $c^i_{t,x} \in \mathbb{R}$ such that for all $p \in \partial V(t,x)$, $p \cdot (1, g_i(t,x)) = c^i_{t,x}$;

(G2) for each $i \in \{1, \ldots, m\}$ only one of the following mutually exclusive conditions holds:

- for all $p \in \partial V(t, x)$ $p \cdot (1, g_i(t, x)) > 0$,
- for all $p \in \partial V(t, x)$ $p \cdot (1, g_i(t, x)) < 0$,
- for all $p \in \partial V(t, x)$ $p \cdot (1, g_i(t, x)) = 0$;

(G3) there exists $\bar{\imath} \in \{1, \ldots, m\}$ such that for each $i \in \{1, \ldots, m\} \setminus \{\bar{\imath}\}$ only one of the following mutually exclusive conditions holds:

- for all $p \in \partial V(t, x)$ $p \cdot (1, g_i(t, x)) > 0$,
- for all $p \in \partial V(t, x)$ $p \cdot (1, g_i(t, x)) < 0$,
- for all $p \in \partial V(t, x)$ $p \cdot (1, g_i(t, x)) = 0$;

Let us remark that (f3) $\Rightarrow$ (f2) $\Rightarrow$ (f1) and (G1) $\Rightarrow$ (G2) $\Rightarrow$ (G3).

THEOREM 2. *Let $V : \mathbb{R}^{n+1} \to \mathbb{R}$ be such that there exists $L > 0$ such that* (V0) *and* (V1) *hold.*

If for all $x \in \mathbb{R}^n$ with $\|x\| > L$ one of the following couples of conditions holds for a.e. $t \geq 0$:

(i) (f1) *and* (G1), (ii) (f2) *and* (G2), (iii) (f3) *and* (G3),

then system (1) is UBIBS stabilizable.

Let us make some remarks.
If for all $x \in \mathbb{R}^n$ with $\|x\| > L$ assumption (f1) (or (f2) or (f3)) holds for a.e. $t \geq 0$, then, by Theorem 1 in Section 2, system (2) is uniformly Lagrange stable. Actually in [4] the authors introduce the concept of robust uniform Lagrange stability and prove that it is equivalent to the existence of a locally Lipschitz continuous Lyapunov-like function. Then assumption (f1) (or (f2) or (f3)) implies more than uniform Lagrange stability of system (2). In [9], the author has also proved that, under mild additional assumptions on f, robust Lagrange stability implies the existence of a C^∞ Lyapunov-like function, but the proof of this result is not actually constructive. Then we could still have to deal with nonsmooth Lyapunov-like functions even if we know that there exist smooth ones.

Moreover Theorem 2 can be restated for autonomous systems with the function V not depending on time. In this case the feedback law is autonomous and it is possible to deal with a situation in which the results in [9] don't help.

Finally let us remark that if f is locally Lipschitz continuous, then, by [14] (page 105), the Lagrange stability of system (2) implies the existence of a time-dependent Lyapunov-like function of class C^∞. In this case, in order to get UBIBS stabilizability of system (1), the regularity assumption on G can be weakened to $G \in L^\infty_{loc}(\mathbb{R}^{n+1}; \mathbb{R}^m)$ (as in [2]).

4. Proof of Theorem 2

We first state and prove a lemma.

LEMMA 2. *Let $V : \mathbb{R}^{n+1} \to \mathbb{R}$ be such that there exists $L > 0$ such that* (V0) *and* (V1) *hold. If $(\bar{t}, \bar{x})$, with $\|\bar{x}\| > L$, is such that, for all $p \in \partial V(\bar{t}, \bar{x})$ $p \cdot (1, g_i(\bar{t}, \bar{x})) > 0$, then there exists $\delta_{\bar{x}} > 0$ such that, for all $x \in B(\bar{x}, \delta_{\bar{x}})$, for all $p \in \partial V(\bar{t}, x)$, $p \cdot (1, g_i(\bar{t}, x)) > 0$.*

Analagously if $(\bar{t}, \bar{x})$, *with* $\|\bar{x}\| > L$, *is such that for all* $p \in \partial V(\bar{t}, \bar{x})$, $p \cdot (1, g_i(\bar{t}, \bar{x})) < 0$, *then there exists* $\delta_{\bar{x}} > 0$ *such that, for all* $x \in B(\bar{x}, \delta_{\bar{x}})$, *for all* $p \in \partial V(\bar{t}, x)$, $p \cdot (1, g_i(\bar{t}, x)) < 0$.

Proof. Let $\gamma > 0$ be such that $\|\bar{x}\| > L + \gamma$, and let $L_{\bar{x}} > 0$ be the Lipschitz constant of V in the set $\{\bar{t}\} \times B(\bar{x}, \gamma)$. For all $(\bar{t}, x) \in \{\bar{t}\} \times B(\bar{x}, \gamma)$ and for all $p \in \partial V(\bar{t}, x)$ $\|p\| \leq L_{\bar{x}}$ (see [5], page 27).

Since g_i is continuous there exist η and M such that $\|(1, g_i(\bar{t}, x))\| \leq M$ in $\{\bar{t}\} \times B(\bar{x}, \eta)$.

Let $d = \min\{p \cdot (1, g_i(\bar{t}, \bar{x})),\ p \in \partial V(\bar{t}, \bar{x})\}$. By assumption $d > 0$.

Let us consider $\epsilon < \frac{d}{2(L_{\bar{x}}+M)}$.

By the continuity of g_i, there exists δ_i such that, if $\|x - \bar{x}\| < \delta_i$, then $\|(1, g_i(\bar{t}, x)) - (1, g_i(\bar{t}, \bar{x}))\| < \epsilon$.

By the upper semi-continuity of ∂V (see [5], page 29), there exists $\delta_V > 0$ such that, if $\|x - \bar{x}\| < \delta_V$, then $\partial V(\bar{t}, x) \subseteq \partial V(\bar{t}, \bar{x}) + \epsilon B(0, 1)$, i.e. for all $p \in \partial V(\bar{t}, x)$ there exists $\bar{p} \in \partial V(\bar{t}, \bar{x})$ such that $\|p - \bar{p}\| < \epsilon$.

Let $\delta_{\bar{x}} = \min\{\gamma, \eta, \delta_i, \delta_V\}$, x be such that $\|x - \bar{x}\| < \delta_{\bar{x}}$ and $p \in \partial V(\bar{t}, x)$, $\bar{p} \in \partial V(\bar{t}, \bar{x})$ be such that $\|p - \bar{p}\| < \epsilon$.

It is easy to see that $|p \cdot (1, g_i(\bar{t}, x)) - \bar{p} \cdot (1, g_i(\bar{t}, \bar{x}))| < \frac{d}{2}$, hence $p \cdot (1, g_i(\bar{t}, x)) > \bar{p} \cdot (1, g_i(\bar{t}, \bar{x})) - \frac{d}{2} = \frac{d}{2} > 0$.

The second part of the lemma can be proved in a perfectly analogous way.

□

Proof of Theorem 2. For each $x \in \mathbb{R}^n$, let N_x be the zero-measure subset of $\mathbb{R}^+$ in which no one of the couples of conditions (i), (ii) and (iii) holds. Let $k : \mathbb{R}^{n+1} \to \mathbb{R}^m$, $k(t, x) = (k_1(t, x), \ldots, k_m(t, x))$, be defined by

$$k_i(t, x) = \begin{cases} -\|x\| & \text{if } \forall p \in \partial V(t, x)\ p \cdot (1, g_i(t, x)) > 0 \\ 0 & \text{if } \forall p \in \partial V(t, x)\ p \cdot g_i(t, x) = 0\,, \\ & \text{or (f3) and (G3) hold and } i = \bar{\imath},\ \text{ or } t \in N_x \\ \|x\| & \text{if } \forall p \in \partial V(t, x)\ p \cdot (1, g_i(t, x)) < 0\,. \end{cases}$$

It is clear that $k \in L^\infty_{loc}(\mathbb{R}^{n+1}, \mathbb{R}^m)$.

By Theorem 1 it is sufficient to prove that for all $R > 0$ there exists $\rho > L, R$ such that for all $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ the following holds:

$$\|x\| > \rho, \quad \|v\| < R \Rightarrow \max \dot{\overline{V}}^{(3)}(t, x) \leq 0 \text{ for all } t \in \mathbb{R}^+ \backslash N_x$$

where $\dot{\overline{V}}^{(3)}(t, x) = \{a \in \mathbb{R} : \exists w \in K(f(t, x) + G(t, x)k(t, x) + G(t, x)v) \text{ such that } \forall p \in \partial V(t, x)\ p \cdot (1, w) = a\}$.

Let x be fixed and $t \in \mathbb{R}^+ \backslash N_x$. Let $a \in \dot{\overline{V}}^{(3)}(t, x)$, $w \in K(f(t, x) + G(t, x)k(t, x) + G(t, x)v)$ be such that for all $p \in \partial V(t, x)$ $p \cdot w = a$.

By Theorem 1 in [8] we have that
$K(f(t, x) + G(t, x))(k(t, x) + v)(x) \subseteq Kf(t, x) + \sum_{i=1}^m g_i(t, x) K(k_i(t, x) + v_i)$, then there exists $z \in Kf(t, x)$, $z_i \in K(k_i(t, x) + v_i)$, $i \in \{1, \ldots, m\}$, such that $w = z + \sum_{i=1}^m g_i(t, x) z_i$.

Let us show that $a \leq 0$. We distinguish the three cases (i), (ii), (iii).

(i) $b = p \cdot (1, z) = a - \sum_{i=1}^m c^i_{t,x} z_i$ does not depend on p, then $b \in \dot{\overline{V}}^{(2)}(t, x)$ and, by (f1), $b \leq 0$.

Let us now show that for each $i \in \{1, \ldots, m\}$ $c^i_{t,x} z_i \leq 0$. If i is such that $c^i_{t,x} = 0$, obviously $c^i_{t,x} z_i \leq 0$. If i is such that $c^i_{t,x} > 0$ then, by Lemma 1, there exists δ_x such that $k_i(t, y) = -\|y\|$ in $\{\overline{t}\} \times B(x, \delta_x)$, then k_i is continuous at x with respect to y. This implies that $K(k_i(t, x) + v_i) = -\|x\| + v_i$, i.e. $z_i = -\|x\| + v_i$ and $c^i_{t,x} z_i \leq 0$, provided that $\|v\| > \rho \geq \max\{L, R\}$.

The case in which i is such that $c^i_{t,x} < 0$ can be treated analogously. We finally get that $a = b + \sum_{i=1}^m c^i_{t,x} z_i \leq 0$.

(ii) By (f2) there exists $\overline{p} \in \partial V(t, x)$ such that $\overline{p} \cdot (1, z) \leq 0$. $a = \overline{p} \cdot (1, z) + \sum_{i=1}^m \overline{p} \cdot (1, g_i(t, x)) z_i$. The fact that for each $i \in \{1, \ldots, m\}$ we have $\overline{p} \cdot (1, g_i(t, x)) z_i \leq 0$ can be proved as in (i) we have proved that for each $i \in \{1, \ldots, m\}$ $c^i_{t,x} z_i \leq 0$. We finally get that $a \leq 0$.

(iii) Let us remark that if (G2) is not verified, i.e. we are not in the case (ii), there exists $\overline{p} \in \partial V(t, x)$ corresponding to $\overline{\imath}$ such that $\overline{p} \cdot (1, g_{\overline{\imath}}(t, x)) = 0$. Indeed, because of the convexity of $\partial V(t, x)$, for all $v \in \mathbb{R}^n$, if there exist $p_1, p_2 \in \partial V(t, x)$ such that $p_1 \cdot v > 0$ and $p_2 \cdot v < 0$, then there also exists $p_3 \in \partial V(t, x)$ such that $p_3 \cdot v = 0$.

Let $\overline{p} \in \partial V(t, x)$ be such that $\overline{p} \cdot (1, g_{\overline{\imath}}(t, x)) = 0$. For all $p \in \partial V(t, x)$ $a = p \cdot (1, w)$. In particular we have $a = \overline{p} \cdot (1, w) = \overline{p} \cdot (1, z) + \sum_{i \neq \overline{\imath}} \overline{p} \cdot (1, g_i(t, x)) z_i + \overline{p} \cdot (1, g_{\overline{\imath}}(t, x)) z_{\overline{\imath}}$. By (f3), $\overline{p} \cdot (1, z) \leq 0$. If $i \neq \overline{\imath}$ the proof that $\overline{p} \cdot (1, g_i(t, x)) z_i \leq 0$ is the same as in (ii). If $i = \overline{\imath}$, because of the choice of $\overline{p}$, $\overline{p} \cdot (1, g_{\overline{\imath}}(t, x)) = 0$. Also in this case we can then conclude that $a \leq 0$.

□

References

[1] Bacciotti A., *External Stabilizability of Nonlinear Systems with some applications*, International Journal of Robust and Nonlinear Control **8** (1998), 1–10.

[2] Bacciotti A., Beccari G., *External Stabilizabilty by Discontinuous Feedback*, Proceedings of the second Portuguese Conference on Automatic Control 1996.

[3] Bacciotti A., Ceragioli F., *Stability and Stabilization of Discontinuous Systems and Nonsmooth Lyapunov Functions*, Preprint Dipartimento di Matematica del Politecnico di Torino.

[4] Bacciotti A., Rosier L., *Liapunov and Lagrange Stability: Inverse Theorems for Discontinuous Systems*, Mathematics of Control, Signals and Systems **11** (1998), 101–128.

[5] Clarke F. H., *Optimization and Nonsmooth Analysis*, Wiley and Sons 1983.

[6] Filippov A. F., *Differential Equations with Discontinuous Right Handsides*, Kluwer Academic Publishers 1988.

[7] McShane E. J., *Integration*, Princeton University Press 1947.

[8] Paden B., Sastry S., *A Calculus for Computing Filippov's Differential Inclusion with Application to the Variable Structure Control of Robot Manipulators*, IEEE Transaction on Circuits and Systems, Vol. Cas-34, No. 1, January 1997, 73–81.

[9] Rosier L., *Smooth Lyapunov Functions for Lagrange Stable Systems*, Preprint.

[10] Shevitz D., Paden B., *Lyapunov Stability Theory of Nonsmooth Systems*, IEEE Transaction on Automatic Control **39** (9), September 1994, 1910–1914.

[11] SONTAG E. D., *A Lyapunov-like Characterization of Asymptotic Controllability*, SIAM J. Control and Opt. **21** (1983), 462–471.

[12] SONTAG E. D., SUSSMANN H., *Nonsmooth Control Lyapunov Functions*, Proc. IEEE Conf. Decision and Control, New Orleans, Dec. 1995, IEEE Publications 1995, 2799–2805.

[13] VARAIYA P. P., LIU R., *Bounded-input Bounded-output Stability of Nonlinear Time-varying Differential Systems*, SIAM Journal Control **4** (1966), 698–704.

[14] YOSHIZAWA T., *Stability Theory by Liapunov's Second Method*, The Mathematical Society of Japan 1966.

AMS Subject Classification: ???.

Francesca CERAGIOLI
Dipartimento di Matematica "U. Dini"
Università di Firenze
viale Morgagni 67/A
50134 Firenze, ITALY
e-mail: `ceragio@udini.math.unifi.it`

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

L. Pandolfi*

ON THE SOLUTIONS OF THE DISSIPATION INEQUALITY

Abstract.
We present some recent results on the existence of solutions to the Dissipation Inequality.

1. Introduction

In this review paper we outline recent results on the properties of the **Dissipation Inequality**, shortly **(DI)**. The **(DI)** is the following inequality in the unknown operator P:

$$\textbf{(DI)} \qquad 2\Re e\,\langle Ax,\, P(x+Du)\rangle + F(x+Du,u) \geq 0\,.$$

Here A is the generator of a C_0-semigroup e^{At} on a Hilbert space X and $D \in \mathcal{L}(U,X)$ where U is a second Hilbert space; $F(x,u)$ is a continuous quadratic form on $X \times U$,

$$F(x,u) = \langle x,\, Qx\rangle + 2\Re e\,\langle Sx,\, u\rangle + \langle u,\, Ru\rangle\,.$$

Positivity of $F(x,u)$ is not assumed.

We require that $P = P^* \in \mathcal{L}(X)$.

We note that the unknown P appears linearly in the **(DI)**, which is also called **Linear Operator Inequality** for this reason.

The **(DI)** has a central role in control theory. We shortly outline the reason by noting the following special cases:

- The case $D = 0$, $S = 0$, $R = 0$. In this case, **(DI)** takes the form of a Lyapunov type inequality,
$$2\Re e\,\langle Ax,\, Px\rangle \geq -\langle x,\, Qx\rangle\,.$$
- If $Q = 0$ and $R = 0$ (but $S \neq 0$) and if $B = -AD \in \mathcal{L}(U,X)$ we get the problem
$$(1) \qquad 2\Re e\,\langle Ax,\, Px\rangle \geq 0 \qquad B^*P = -S\,.$$
This problem is known as *Lur'e Problem* and it is important for example in stability theory, network theory and operator theory.
- The case $S = 0$, $R = I$ and $Q = -I$ is encountered in scattering theory while the case $S = 0$, $Q \geq 0$ and coercive R corresponds to the **standard** regulator problem of control theory.

* Paper written with financial support of the Italian MINISTERO DELLA RICERCA SCIENTIFICA E TECNOLOGICA within the program of GNAFA–CNR.

We associate to **(DI)** the following quadratic regulator problem "with stability": we consider the control system

$$\dot{x} = A(x - Du)\,. \tag{2}$$

We call a pair $(x(\cdot), u(\cdot))$ an *evolution of system (2) with initial datum* x_0 when $x(\cdot)$ is a (mild) solution to (2) with input $u(\cdot)$ and $x(0) = x_0$.

We associate to control system (2) the quadratic cost

$$J(x_0; u) = \int_0^{+\infty} F(x(t), u(t))\,\mathrm{d}t\,. \tag{3}$$

The relevant problem is the following one: we want to characterize the condition $V(x_0) > -\infty$ for each x_0 where $V(x_0)$ **is the infimum of (3) over the class of those square integrable evolutions which have initial datum** x_0. (The term "with stability" refers to the fact that we only consider the square integrable evolutions of the system).

Of course, Eq. (2) has no meaning in general. One case in which it makes sense is the case that $B = -AD$ is a bounded operator (*distributed* control action). In this case the problem has been essentially studied in [7] but for one crucial aspect that we describe below.

More in general, large classes of boundary control systems can be put in the form (2), as shown in [6], where two main classes have been singled out, the first one which corresponds to "hyperbolic" systems and the second one which corresponds to "parabolic" systems.

We illustrate the two classes introduced in [6]:

- The class that models in particular most control problems for the heat equation: the semigroup e^{At} is holomorphic (we assume exponentially stable for simplicity) and $\mathrm{im}D = \mathrm{im}[-A^{-1}B] \subseteq \mathrm{dom}(-A)^{\tilde{\gamma}}$, $\tilde{\gamma} < 1$.
- The class that models in particular most control problems for string and membrane equations: e^{At} is a C_0-semigroup, $A^{-1}B \in \mathcal{L}(X)$ and

  $$\int_0^T \|B^* e^{A^*t} x\|^2\,\mathrm{d}t \le k_T \|x\|^2\,. \tag{4}$$

  It is sufficient to assume that the previous inequality holds for one value of T since then it holds for every T.

As we said, for simplicity of exposition, we assume exponential stability. The simplification which is obtained when the semigroup is exponentially stable is that the class of the controls is $L^2(0, +\infty; U)$, independent of x_0. However, this condition can be removed.

The crucial result in the case of *distributed control action* is as follows (see [14] for the finite dimensional theory and [7] for distributed systems with distributed control action):

THEOREM 1. *If $AD \in \mathcal{L}(U, X)$, then $V(x_0)$ is finite for every x_0 if and only if there exists a solution to* **(DI)** *and in this case $V(x_0)$ is a continuous quadratic form on X: $V(x_0) = \langle x_0, Px_0\rangle$. The operator P of the quadratic form is the* maximal *solution to* **(DI)**.

The result just quoted can be extended to both the classes of boundary control systems introduced in [6], see [9, 11]. Rather than repeating the very long proof, it is possible to use a device, introduced in [10, 8], which associates to the boundary control system an "augmented" system, with distributed control action. From this distributed system it is possible to derive many properties of the **(DI)** of the original boundary control system. This device is illustrated in sect. 2.

With the same method it is possible to extend the next result:

THEOREM 2. *If $V(x_0) > -\infty$, i.e. if* **(DI)** *is solvable, then*

$$\Pi(i\omega) = F(-i\omega(i\omega I - A)^{-1}Du + Du, u) \geq 0 \qquad \forall \omega \in \mathbb{R}\,. \tag{5}$$

The function $\Pi(i\omega)$ was introduced in [12] and it is called the **Popov function**.

As the number $i\omega$ are considered "frequencies", condition (5) is a special "frequency domain condition".

At the level of the frequency domain condition we encounter a crucial difference between the class of "parabolic" and "hyperbolic" systems:

THEOREM 3. *In the parabolic case if $V(x_0) > -\infty$, then $R \geq 0$. Instead, in the "hyperbolic" case, we can have $V(x_0) > -\infty$ even if $R = -\alpha I$, $\alpha > 0$.*

Proof. It is clear that

$$\Pi(i\omega) = F((i\omega I - A)^{-1}Bu, u)$$

($B = -AD$) and $\lim_{|\omega|\to+\infty}(i\omega I - A)^{-1}Bu = 0$ because $\mathrm{im}D = \mathrm{im}[-A^{-1}B] \subseteq \mathrm{dom}(-A)^{\gamma}$ (here we use exponential stability, but the proof can be adapted to the unstable case.) Hence, $0 \leq \lim_{|\omega|\to+\infty} \Pi(i\omega) = \langle u, Ru\rangle$ for each $u \in U$. This proves that $R \geq 0$.

Clearly an analogous proof cannot be repeated in the "hyperbolic" case; and the analogous result does not hold, as the following example shows:

the system is described by

$$x_t = -x_\theta \qquad 0 < \theta < 1, \qquad t > 0 \qquad x(t,0) = u(t)$$

(this system is exponentially stable since the free evolution is zero for $t > 1$).

The functional $F(x,u)$ is

$$F(x,u) = \|x(\cdot)\|^2_{L^2(0,1)} - \alpha|u|^2$$

so that

$$J(x_0; u) = \int_0^{+\infty} \{\|x(t,\cdot)\|^2_{L^2(0,1)} - \alpha|u(t)|^2\}\,\mathrm{d}t\,.$$

If $x(0,\theta) \equiv 0$ then

$$\hat{x}(z,\theta) = e^{-z\theta}\hat{u}(z)$$

so that

$$\langle u, \Pi(i\omega)u\rangle = [1-\alpha]|u|^2\,.$$

This is nonnegative for each $\alpha \leq 1$ in spite of the fact that $R = -\alpha I$ can be negative. Hence, in the hyperbolic boundary control case, **the condition $R \geq 0$ does not follows from the positivity of the Popov function**.

□

It is clear that the frequency domain condition may hold even if the **(DI)** is not solvable, as the following example shows:

EXAMPLE 1. The example is an example of a scalar system,

$$\dot{x} = -x + 0u \qquad y = x\,.$$

It is clear that $\Pi(i\omega) \geq 0$, is nonnegative; but $PB = C$, i.e. $P0 = 1$, is not solvable.

A problem that has been studied in a great deal of papers is the problem of finding additional conditions which imply solvability of the **(DI)** in the case that the frequency domain condition (5) holds. A special instance of this problem is the important Lur'e problem of stability theory.

This problem is a difficult problem which is not completely solved even for finite dimensional systems. Perhaps, the most complete result is in [2]: if a system is finite dimensional and $\Pi(i\omega) \geq 0$, then a sufficient condition for solvability of **(DI)** is the existence of a number ω_0 such that $\det \Pi(i\omega_0) \neq 0$.

It is easy to construct examples which show that this condition is far from sufficient.

In the context of hyperbolic systems, the following result is proved in [11].

THEOREM 4. *Let condition (4) hold and let the system be* exactly controllable. *Under these conditions, if the Popov function is nonnegative then there exists a solution to* **(DI)** *and, moreover, the* maximal *solution P of* **(DI)** *is the strong limit of the decreasing sequence $\{P_n\}$, where P_n is the maximal solution of the* **(DI)**

$$2\Re e\,\langle Ax, P(x+Du)\rangle + F(x+Du, u) + \frac{1}{n}\{\|u\|^2 + \|x\|^2\} \geq 0\,. \tag{6}$$

The last statement is important because it turns out that P_n solves a Riccati equation, while there is no equation solved by P in general.

The proof of Theorem 4 essentially reproduces the finite dimensional proof in [14]. Hence, the "hyperbolic" case is "easy" since the finite dimensional proof can be adapted. In contrast with this, the "parabolic" case requires new ideas and it is "difficult". Consistent with this, only very partial results are available in this "parabolic" case, and under quite restrictive conditions. These results are outlined in sect. 3.

Before doing this we present, in the next section, the key idea that can be used in order to pass from a boundary control system to an *"augmented"* but *distributed* control system.

2. The augmented system

A general model for the analysis of boundary control systems was proposed by Fattorini ([4]). Let X be a Hilbert space and σ a linear closed densely defined operator, $\sigma : X \to X$. A second operator τ is linear from X to a Hilbert space U.

We assume:

Assumption We have: $\operatorname{dom}\sigma \subseteq \operatorname{dom}\tau$ and τ is continuous on the Hilbert space $\operatorname{dom}\sigma$ with the graph norm.

The "boundary control system" is described by:

$$\begin{cases} \dot{x} = \sigma x \\ \tau x = u \end{cases} \qquad x(0) = x_0 \tag{7}$$

where $u(\cdot) \in L^2_{\mathrm{loc}}(0, +\infty; U)$.

We must define the "strong solutions" $x(\cdot; x_0, u)$ to system (7). Following [3] the function $x(\cdot) = x(\cdot; x_0, u)$ is a strong solution if there exists a sequence $\{x_n(\cdot)\}$ of C^1-functions such that

$x_n(t) \in \operatorname{dom}\sigma$ for each $t \geq 0$ and:

$$(8) \qquad \begin{cases} \dot{x}_n(\cdot) - \sigma x_n(\cdot) \to 0 & \text{in } L^2_{\text{loc}}(0, +\infty; X) \\ x_n(0) \to x_0 & \text{in } X \\ \tau x_n(\cdot) \to u(\cdot) & \text{in } L^2_{\text{loc}}(0, +\infty; U) \end{cases}$$

and

- $x_n(\cdot)$ **converges uniformly to** $x(\cdot)$ **on compact intervals in** $[0, +\infty)$.

In the special case that the sequence $x_n(\cdot)$ is stationary, $x_n(\cdot) = x(\cdot)$, we shall say that $x(\cdot)$ is a *classical* solution to problem (7).

Assumption 1. Let us consider the "elliptic" problem $\sigma x = u$. We assume that it is "well posed", i.e. that there exists an operator $D \in \mathcal{L}(U, X)$ such that

$$x = Du \text{ iff } \{\sigma x = 0 \text{ and } \tau x = u\}.$$

Moreover we assume that the operator A defined by

$$\operatorname{dom} A = \operatorname{dom}\sigma \cap \ker\tau \qquad Ax = \sigma x$$

generates a strongly continuous semigroup on X.

As we said already, for simplicity of exposition, we assume that the semigroup e^{At} is exponentially stable.

Now we recall the following arguments from [1]. Classical solutions to Eq. (7) solve

$$(9) \qquad \dot{x} = A(x - Du) \qquad x(0) = x_0 .$$

Let $u(\cdot)$ be an absolutely continuous control and $\xi(t) = x(t) - Du(t)$. Then, $\xi(\cdot)$ is a classical solution to

$$(10) \qquad \dot{\xi} = A\xi - D\dot{u} \qquad \xi_0 = \xi(0) = x(0) - Du(0)$$

and conversely.

As the operator A generates a C_0-semigroup, it is possible to write a "variation of constants" formula for the solution ξ. "Integration by parts" produces a variation of constants formula, which contains unbounded operators, for the function $x(\cdot)$. This is the usual starting point for the study of large classes of boundary control systems. Instead, we "augment" system (9) and we consider the system:

$$(11) \qquad \begin{cases} \dot{\xi} = A\xi - Dv \\ \dot{u} = v \end{cases}$$

Here we consider formally $v(\cdot)$ as a new "input", see [10, 8].

Moreover, we note that it is possible to stabilize the previous system with the simple feedback $v = -u$, since e^{At} is exponentially stable.

The cost that we associate to (11) is the cost

$$(12) \qquad J(x_0; u) = \int_0^{+\infty} F(\xi(t) + Du(t), u(t))\, \mathrm{d}t .$$

This cost does not depend explicitly on the new input $v(\cdot)$: it is a quadratic form of the state, which is now $\Xi = [\xi, u]$.

It is proved in [9] that the value function $\mathcal{V}(\xi_0, u_0)$ of the augmented system has the following property:

$$\mathcal{V}(\xi_0 + Du_0, u_0) = V(x_0)\,.$$

We apply the stabilizing feedback $v = -u$ and we write down the **(DI)** and the Popov function for the stabilized augmented system. The **(DI)** is

$$2\Re e\,\langle \mathcal{A}\Xi, W\Xi\rangle + \langle \Xi, \mathcal{Q}\Xi\rangle \geq 0 \qquad \forall \Xi \in \mathrm{dom}\mathcal{A}\,, \qquad W\mathcal{D} = 0\,. \tag{13}$$

where

$$\mathcal{A} = \begin{bmatrix} A & -D \\ 0 & I \end{bmatrix}, \qquad \Xi = \begin{bmatrix} \xi \\ u \end{bmatrix},$$

$$\mathcal{Q} = \begin{bmatrix} Q & S^* + QD \\ D^*Q + S & R + D^*S^* + SD + D^*QD \end{bmatrix}, \quad \mathcal{D} = \begin{bmatrix} -D \\ I \end{bmatrix}.$$

The Popov function is:

$$P(i\omega) = \frac{\Pi(i\omega)}{1+\omega^2} \tag{14}$$

It is clear that the transformations outlined above from the original to the augmented system do not affect the positivity of the Popov function and that if $\omega^s\Pi(i\omega)$ is bounded from below, then $\omega^{s+2}P(i\omega)$ is bounded from below.

In the next section we apply the previous arguments to the case that the operator A generates a holomorphic semigroup and $\mathrm{im}D \subseteq (\mathrm{dom}(-A)^\gamma)$, $\gamma < 1$.

3. "Parabolic" case: from the Frequency domain condition to the (DI)

We already said that in the parabolic case only partial results are available. In particular, available results require that the control be scalar so that S is an element of X. This we shall assume in this section. We assume moreover that the operator A has only point spectrum with simple eigenvalues z_k and the eigenvectors v_k form a complete set in X. Just for simplicity we assume that the eigenvalues are real (hence negative). Moreover, we assume that we already wrote the system in the form of a distributed (augmented and stabilized) control system. Hence we look for conditions under which there exists a solution W to (13).

We note that $\mathcal{D} \in X \times U$ and that $P(i\omega)$ is a scalar function: it is the restriction to the imaginary axis of the analytic function

$$P(z) = -\mathcal{D}(zI + \mathcal{A}^*)^{-1}\mathcal{Q}(zI - \mathcal{A})^{-1}\mathcal{D}\,.$$

The function $P(z)$ is analytic in a strip which contains the imaginary axis in its interior.

We assume that $P(i\omega) \geq 0$ and we want to give additional conditions under which (13) is solvable. In fact, we give conditions for the existence of a solution to the following more restricted problem: to find an operator W and a vector $q \in (\mathrm{dom}A)'$ such that

$$2\Re e\,\langle \mathcal{A}\Xi, W\Xi\rangle + \langle \Xi, \mathcal{Q}\Xi\rangle = |\langle\langle \Xi, q\rangle\rangle\|^2 \quad \forall \Xi \in \mathrm{dom}\mathcal{A}\,. \tag{15}$$

The symbol $\langle\langle \cdot, \cdot\rangle\rangle$ denotes the pairing of $(\mathrm{dom}A)'$ and $\mathrm{dom}A$.

The previous equation suggests a form for the solution W:

$$\langle \Xi, W\Xi\rangle = \int_0^{+\infty} \langle e^{\mathcal{A}t}\,\Xi, \mathcal{Q}e^{\mathcal{A}t}\,\Xi\rangle\,\mathrm{d}t - \int_0^{+\infty} |\langle\langle \Xi, e^{\mathcal{A}t}q\rangle\rangle\|^2\,\mathrm{d}t\,. \tag{16}$$

However, it is clear that in general the operator W so defined will not be continuous, unless q enjoys further regularity. We use known properties of the fractional powers of the generators of holomorphic semigroups and we see that W is bounded if $q \in [\mathrm{dom}(-\mathcal{A}^{\alpha})]'$ with $\alpha < 1/2$.

It is possible to prove that if a solution W to (15) exists then there exists a factorization

$$P(i\omega) = m^*(i\omega)m(i\omega)$$

and $m(i\omega)$ does not have zeros in the right half plane. This observation suggests a method for the solution of Eq. (15), which relies on the computation of a factorization of $P(i\omega)$. The factorization of functions which takes nonnegative values is a classical problem in analysis. The key result is the following one:

LEMMA 1. *If $P(i\omega) \geq 0$ and if $|\ln P(i\omega)|/(1+\omega^2)$ is integrable, then there exists a function $m(z)$ with the following properties:*

- *$m(z)$ is holomorphic and bounded in $\Re e\, z > 0$;*
- *$P(i\omega) = m(-i\omega)m(i\omega)$;*
- *let $z = x + iy$, $x > 0$. The following equality holds:*

$$\ln|m(z)| = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \ln P(i\omega)\frac{x}{x^2+(\omega-y)^2}\,d\omega \quad \forall z = x+iy, \quad x > 0\,. \tag{17}$$

See [13, p. 121], [5, p. 67].

A function which is holomorphic and bounded in the right half plane and which satisfies (17) is called an *outer function.*

The previous arguments show that an outer factor of $P(z)$ exists when $P(i\omega) \geq 0$ and when $P(i\omega)$ decays for $|\omega| \to +\infty$ of the order $1/|\omega|^{\beta}$, $\beta < 1$. Let us assume this condition (which will be strengthened below). Under this condition $P(z)$ can be factorized and, moreover,

$$\begin{aligned}
\ln|m(z)| &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \ln P(i\omega)\frac{x}{x^2+(\omega-y)^2}\,\mathrm{d}\omega \\
&\leq \frac{1}{2\pi}\int_{-\infty}^{+\infty} \ln\frac{M}{1+\omega^2}\,\frac{x}{x^2+(\omega-y)^2}\,\mathrm{d}\omega \\
&= \ln\left|\frac{1}{1+z^2}\right|.
\end{aligned}$$

This estimates implies in particular that the integrals $\int_{-\infty}^{+\infty} |m(x+iy)|^2\,\mathrm{d}y$ are uniformly bounded in $x > 0$. Paley Wiener theorem (see [5]) implies that

$$m(i\omega) = \int_0^{+\infty} e^{-i\omega t}\check{m}(t)\,\mathrm{d}t, \qquad \check{m}(\cdot) \in L^2(0,+\infty)\,.$$

The function $\check{m}(t)$ being square integrable, we can write the integral

$$\int_0^{+\infty} e^{A^*s}q\check{m}(t)\,\mathrm{d}t$$

and we can try to solve the following equation for q:

$$\int_0^{+\infty} e^{A^* s} q \check{m}(t)\,\mathrm{d}t = -s = \int_0^{+\infty} e^{\mathcal{A}^* t} Q e^{\mathcal{A} t} \mathcal{D}\,\mathrm{d}t\,. \tag{18}$$

This equation is suggested by certain necessary conditions for the solvability of (1) which are not discussed here.

We note that

$$s \in \mathrm{dom}(-\mathcal{A})^{1-\epsilon} \text{ for each } \epsilon > 0\,. \tag{19}$$

It turns out that equation (18) can always be **formally** solved, a solution being

$$q_k = \langle v_k, q\rangle = -\frac{\langle v_k, s\rangle}{m(-\bar{z}_k)}$$

since $m(z)$ does not have zeros in the right half plane.

Moreover, we can prove that the operator W defined by (16) **formally** satisfies the condition $W\mathcal{D} = 0$. Hence, this operator W will be the required solution of (15) if it is a bounded operator, i.e. if $q \in [\mathrm{dom}(-\mathcal{A}^\alpha)]'$.

An analysis of formula (17) shows the following result:

THEOREM 5. *The vector q belongs to* $(\mathrm{dom}(-\mathcal{A}^*)^{1/2-\epsilon})'$ *for some* $\epsilon > 0$ *if there exist numbers* $\gamma < 1$ *and* $M > 0$ *such that*

$$|\omega|^\gamma \Pi(i\omega) > M$$

for $|\omega|$ *large.*

Examples in which the condition of the theorem holds exist, see [9].

Let $\zeta_k = -z_k \in \mathbb{R}$. The key observation in the proof of the theorem is the following equality, derived from (17):

$$\begin{aligned}
\log|\zeta|^{\frac{3}{2}-\epsilon} m(\zeta_k) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} [\log|\zeta_k|^{3-2\epsilon} P(i\zeta_k s)]\frac{1}{1+s^2}\,\mathrm{d}s \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} [\log \zeta_k^{3-\gamma-2\epsilon}\frac{1}{|s|^\gamma}]\frac{1}{1+s^2}\,\mathrm{d}s \\
&\quad + \frac{1}{2\pi}\int_{-\infty}^{+\infty} [\log \zeta_k |s|^\gamma P(i\zeta_k s)]\frac{1}{1+s^2}\,\mathrm{d}s\,.
\end{aligned}$$

The first integral is bounded below if $\gamma \le 3 - 2\epsilon$ and the second one is bounded below in any case.

We recapitulate: the condition $q \in (\mathrm{dom}(-\mathcal{A}^*)^{1/2-\epsilon})'$ holds if $P(i\omega)$ decays at ∞ of order less than 3. We recall (14) and we get the result.

Acknowledgment. The author thanks the referee for the carefull reading of this paper.

References

[1] BALAKRISHNAN A. V., *On a generalization of the Kalman-Yakubovich Lemma*, Appl. Math. Optim. **31** (1995), 177–187.

[2] CHURILOV A. N., *On the solvability of matrix inequalities*, Mat. Zametki **36** (1984), 725–732.

[3] DESCH W., LASIECKA I., SCHAPPACHER W., *Feedback boundary control problems for linear semigroups,* Israel J. of Mathematics **51** (1985), 177–207.

[4] FATTORINI O., *Boundary control systems*, SIAM J. Control Optim. **6** (1968), 349–385.

[5] GARNETT J. B., *Bounded Analytic Functions*, Academic Press, New York 1981.

[6] LASIECKA I., TRIGGIANI R., *Differential and Algebraic Riccati Equations with Applications to Boundary/Point Control Problems: Continuous Theory and Approximation Theory*, Lect. Notes in Control Inf. Sci. n. 164, Springer-Verlag, Berlin 1991.

[7] LOUIS J-CL., WEXLER D., *The Hilbert space regulator problem and operator Riccati equation under stabilizability*, Ann. Soc. Sci. Bruxelles Sér I **105** (1991), 137–165.

[8] PANDOLFI L., *From singular to regular control systems*, in "Control of Partial Differential Equations", G. Da Prato and M. Tubaro Ed.s, M. Dekker, New York 1994, 153–165.

[9] PANDOLFI L., *Dissipativity and the Lur'e Problem for parabolic boundary control systems*, SIAM J. Control Optimization **36** (1998), 2061–2081.

[10] PANDOLFI L., *The standard regulator problem for systems with input delays: an approach through singular control theory*, Appl. Math. Optim. **31** (1995), 119–136.

[11] PANDOLFI L., *The Kalman–Yakubovich–Popov Theorem for stabilizable hyperbolic boundary control systems*, Integral Equations Operator Theory, to appear.

[12] POPOV V. M., *Absolute stability of nonlinear systems of automatic control*, Automat. Remote Control **22** (1961), 857–875.

[13] ROSENBLUM M., ROVNYAK J., *Topics in Hardy classes and univalent functions*, Birkäuser Verlag, Basel 1994.

[14] YAKUBOVICH V. A., *The frequency theorem in control theory*, Siberian Math. J. **14** (1973), 384–419.

AMS Subject Classification: ???.

L. PANDOLFI
Politecnico di Torino
Dipartimento di Matematica
Corso Duca degli Abruzzi, 24
10129 Torino, Italy
e-mail: `lucipan@polito.it`

Rend. Sem. Mat. Univ. Pol. Torino
Vol. 56, 4 (1998)

F. Rampazzo – C. Sartori

ON PERTURBATIONS OF MINIMUM PROBLEMS WITH UNBOUNDED CONTROLS

Abstract.
A typical optimal control problem among those considered in this work includes dynamics of the form $f(x,c) = g_0(x) + \tilde{g}_0(x)|c|^\alpha$ (here x and c represent the state and the control, respectively) and a Lagrangian of the form $l(x,c) = l_0(x) + \tilde{l}_0(x)|c|^\beta$, with $\alpha \leq \beta$, and c belonging to a closed, unbounded subset of $\mathbb{R}^m$. We perturb this problem by considering dynamics and Lagrangians $f_n(x,c) = g_n(x) + \tilde{g}_n(x)|c|^{\alpha_n}$, and $l_n(x,c) = l_{0_n}(x) + \tilde{l}_{0_n}(x)|c|^\beta$ respectively, with $\alpha_n \leq \beta$, and f_n and l_n approaching f and l. We show that the value functions of the perturbed problems converge, uniformly on compact sets, to the value function of the original problem. For this purpose we exploit some comparison results for Bellman equations with fast gradient-dependence which have been recently established in a companion paper. Of course the fast growth in the gradient of the involved Hamiltonians is connected with the presence of unbounded controls. As an easy consequence of the convergence result, an optimal control for the original problem turns out to be nearly optimal for the perturbed problems. This is true in particular, for very general perturbations of the LQ problem, including cases where the perturbed problem is *not* coercive, that is, $\alpha_n = \beta (= 2)$.

1. Introduction

Let us consider a Boltz optimal control problem,

$$(P) \qquad \begin{array}{c} \text{minimize } \int_{\bar{t}}^{T} l(t,x,c)\,dt + g(x(T)) \\ \dot{x} = f(t,x,c) \quad x(\bar{t}) = \bar{x} \\ (\bar{t},\bar{x}) \in [0,T] \times \mathbb{R}^k\,, \end{array}$$

where $c = c(t)$ is a control which takes values in $\mathbb{R}^m$. Let us also consider a sequence of *perturbations* of this problem,

$$(P_n) \qquad \begin{array}{c} \text{minimize } \int_{\bar{t}}^{T} l_n(t,x,c)dt + g_n(x(T)) \\ \dot{x} = f_n(t,x,c) \quad x(\bar{t}) = \bar{x} \end{array}$$

where the triples (f_n, l_n, g_n) converge to (f, l, g), in a sense to be made precise.

In the present note we address the following question:

Q$_1$. *Assume that for every initial data $(\bar{t},\bar{x})$ an optimal control $c_{(\bar{t},\bar{x})} : [\bar{t},T] \to \mathbb{R}^m$ is known. Are these controls nearly optimal for the problem (P_n)?*

(Here *nearly optimal* means that the value of the cost functional of (P_n) when the control $c_{(\bar{t},\bar{x})}$ is implemented differs from the optimal value by an error which approaches zero when n tends to ∞).

An analogous question can be posed when an optimal feedback control $c = c(t, x)$ of problem (P) is known:

$\mathbf{Q}_2$. *Is the feedback control $c(t, x)$ nearly optimal for the problem (P_n)?*

The practical usefulness of studying such a theoretical problem is evident: it may happen that the construction of an optimal control for problem (P) is relatively easy, while the same task for the perturbed problem (P_n) might result hopeless. In this case, one could be tempted to implement the (P) optimal control for problem (P_n) as well. And positive answers to questions like $\mathbf{Q}_1$ and $\mathbf{Q}_2$ would guarantee that these strategies would be safe. (For a general account on perturbation theory see e.g. [3]).

Since we are interested in the case when the controls c are unbounded, questions concerning the growth in c of f and l turn out to be quite relevant. The crucial hypotheses (see $\mathbf{A}_1$-$\mathbf{A}_5$ in Section 2) here assumed on the dynamics f and the Lagrangian l are as follows: there exist α, β, both greater than or equal to 1, such that if $Q \subset \mathbb{R}^k$ is a compact subset and $x, y \in Q$, then

$$|f(t,x,c) - f(t,y,c)| \leq L(1+|c|^{\alpha})|x-y| \tag{1}$$

$$|l(t,x,c)| \geq l_0|c|^{\beta} - C \tag{2}$$

for all $c \in \mathbb{R}^m$, where L depends only on Q. The same kind of hypotheses are assumed on the perturbed pairs f_n, l_n, with the same growth-exponent β for the Lagrangians l_n, while the growth-exponents α_n of the f_n are allowed to depend on n. Moreover, *weak coercivity* relations, namely $\alpha_n \leq \beta$, $\alpha \leq \beta$, are assumed. Let us observe that when $\alpha < \beta$ (*strict coercivity*) the optimal trajectories turn out to be (absolutely) continuous, while, if $\alpha = \beta$, an optimal path may contain *jumps* (in a non trivial sense which *cannot* be resumed by a distributional approach, see e.g. [7, 8]).

Answers to questions $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are given in Theorems 6, 7 below, respectively. The main theoretical tool on which these results rely consists in a so-called stability theorem (see Theorem 1) for a class of Hamilton-Jacobi-Bellman equations with fast gradient-dependence. In order to prove the stability theorem we exploit some uniqueness and regularity results for this class of equations that have been recently established in a companion paper [8] (see also [1] and [6]). Let us notice that questions like $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can be approached with more standard uniqueness results as soon as the controls c are bounded.

Similar questions were addressed in a paper by M. Bardi and F. Da Lio [1], where the authors assumed the following stronger hypothesis on f:

$$|f(x,c) - f(y,c)| \leq L|x-y| \tag{3}$$

(actually a monotonicity hypothesis, weaker than (3) is assumed in [1]; however this is irrelevant at this stage, while the main point in assuming (3) consists in the fact that it is Lipschitz in x *uniformly with respect to* c). Observe that hypothesis (3) still allows for fields growing as $|c|^{\alpha}$ in the variable c. Yet, while a field of the form $f(x,c) \doteq g_0(x) + g_1(x)|c|^{\alpha}$ agrees with hypothesis (1), it does not satisfy hypothesis (3) unless $g_1(x)$ is constant. Furthermore, in [1] the exponent α is required to be strictly less than β (strict coercivity).

The relevance of weakening hypothesis (3) (and the position $\alpha < \beta$) is perhaps better understood by means of an application to a perturbation question for the linear quadratic problem. In this case one has: $\alpha = 1, \beta = 2$, $f(x,c) = Ax + Bc$, $l(x,c) = x^*Dx + x^*Ec + c^*Fc$, $g(x) = x^*Sx$. Here the coercivity hypothesis reduces to the fact that F is positive definite. As it is well known, (see e.g. [4]) under suitable hypotheses on A, B, D, E and F, this problem admits a smooth optimal feedback, which can be actually computed by solving the corresponding Riccati equation. It is obvious that a crucial point in questions $\mathbf{Q}_1$ and $\mathbf{Q}_2$ consists in the specification of *which* class of perturbation problems (P_n) has to be considered. Of course, since in practical situations the nature of the perturbation is only partially known, the larger this class is the better. In [1] a positive answer to $\mathbf{Q}_1$ is provided when the perturbed fields are of the form

$$f_n(x) = Ax + Bx + \epsilon(n)\varphi(x, c)$$

with φ verifying (3) and $\epsilon(n)$ infinitesimal. So, for instance, a perturbed dynamics like

$$f_n = Ax + Bx + \frac{1}{n}xc$$

is *not* allowed. On the contrary, hypothesis (1) assumed in the present paper is not in contrast with this (and much more general) kind of perturbation. A further improvement is represented by the fact that the f_n's growth exponents α_n are allowed to be different from the f's growth exponent α (=1, in this case), and moreover, they can be less than or *equal* to β (which in this example is equal to 2). So, for example, perturbed dynamics like

$$f_n = Ax + Bc + \epsilon(n)(g(x)c + h(x)|c|^2)$$

may be well considered. In this case, the possibility of implementing a (P)-optimal control c in the perturbed problem (P_n) may be of particular interest. Indeed, the problems (P_n) are quite irregular, in that the lack of a sufficient degree of coercivity may give rise to optimal trajectories with *jumps* (see Remark 2).

The general approach of the present paper, which is partially inspired by [1], relies on proving the convergence of the value functions of the problems (P_n) to the value function of (P) via a PDE argument. However, the enlarged generality of the considered problems makes the exploitation of very recent results on Hamilton-Jacobi-Bellman equations with fast gradient-dependence crucial (see [8]). In particular, by allowing mixed type boundary conditions, these results cover the weak coercivity case ($\alpha = \beta$). Moreover they do not require an assumption of local Lipschitz continuity of the solution of the associated dynamic programming equation. Actually, as a consequence of the fact that we allow value functions which are not equicontinuous, the Ascoli-Arzelà argument exploited in the stability theorem of [1] does not work here. In order to overcome this difficulty we join ordinary convergence arguments originally due to G. Barles and B. Perthame [2] with the reparameterization techniques introduced in [8].

2. A convergence result

For every $\bar{t} \in [0, T]$, let $\mathcal{C}(\bar{t})$ denote the set of Borel-measurable maps which belong to $L^\beta([\bar{t}, T], \mathbb{R}^m)$. $\mathcal{C}(\bar{t})$ is called the set of controls starting at $\bar{t}$. Let us point out that the choice of the whole $\mathbb{R}^m$ as the set where the controls take values is made just for the sake of simplicity. Indeed, in view of the Appendix in [8] it is straightforward to generalize the results presented here to situations where the controls can take values in a (possibly unbounded) closed subset of $\mathbb{R}^m$. For every $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^k$ and every $c \in \mathcal{C}(\bar{t})$, by the assumptions $\mathbf{A}_1$-$\mathbf{A}_5$ listed below, there exists a unique solution of the Cauchy problem

$$(E) \qquad \begin{cases} \dot{x} = f(t, x, c) \text{ for } t \in [\bar{t}, T] \\ x(\bar{t}) = \bar{x}\,, \end{cases}$$

(where the dot means differentiation with respect to t). We will denote this solution by $x_{(\bar{t},\bar{x})}[c](\cdot)$ (or by $x[c](\cdot)$ if the initial data are meant by the context). For every $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^k$ let us consider the optimal control problem

$$(P) \qquad \underset{c \in \mathcal{C}(\bar{t})}{\text{minimize}}\ J(\bar{t}, \bar{x}, c)$$

where

$$J(\bar{t},\bar{x},c) \doteq \int_{\bar{t}}^{T} l(t,x[c](t),c(t))\,dt + g(x[c](T))\,,$$

and let us define the *value function* $V : [0,T[\times\mathbb{R}^k \to \mathbb{R}$, by setting

$$V(\bar{t},\bar{x}) \doteq \inf_{c\in\mathcal{C}(\bar{t})} J(\bar{t},\bar{x},c)\,.$$

We consider also a sequence of *perturbed* problems

$$(P_n) \qquad \underset{c\in\mathcal{C}(\bar{t})}{\text{minimize}}\ J_n(\bar{t},\bar{x},c)$$

where

$$J_n(\bar{t},\bar{x},c) \doteq \int_{\bar{t}}^{T} l_n(t,x_n[c](t),c(t))\,dt + g_n(x_n[c](T))\,,$$

where $x_n[c]$ (or $x^n_{(\bar{t},\bar{x})}[c]$ if one wishes to specify the initial data), denotes the solution – existing unique by hypotheses $\mathbf{A}_1$-$\mathbf{A}_5$ below – of

$$(E_n) \qquad \begin{cases} \dot{x} = f_n(t,x,c) \text{ for } t\in[\bar{t},T] \\ x(\bar{t}) = \bar{x} \end{cases}$$

Let us define the value function V_n of (P_n) by setting

$$V_n(\bar{t},\bar{x}) \doteq \inf_{c\in\mathcal{C}(\bar{t})} J_n(\bar{t},\bar{x},c)\,.$$

We assume that there exist numbers α,α_n,β satisfying $1\le\alpha\le\beta$, $1\le\alpha_n\le\beta$, such that the following hypotheses hold true:

$\mathbf{A}_1$ *the maps f and f_n are continuous on $[0,T]\times\mathbb{R}^k\times\mathbb{R}^m$ and, for every compact subset $Q\subset\mathbb{R}^k$, there exists a positive constant L and a modulus ρ_f verifying*

$$\begin{aligned} |f(t_1,x_1,c)-f(t_2,x_2,c)| &\le (1+|c|^{\alpha})(L|x_1-x_2|+\rho_f(|t_1-t_2|)\,, \\ |f_n(t_1,x_1,c)-f_n(t_2,x_2,c)| &\le (1+|c|^{\alpha_n})(L|x_1-x_2|+\rho_f(|t_1-t_2|) \end{aligned}$$

for all $(t_1,x_1,c),(t_2,x_2,c)\in[0,T]\times Q\times\mathbb{R}^m$ and $n\in\mathbb{N}$, (by modulus *we mean a positive, nondecreasing function, null and continuous at zero);*

$\mathbf{A}_2$ *there exist two nonnegative constants M_1 and M_2 such that*

$$\begin{aligned} |f(t,x,c)| &\le M_1(1+|c|^{\alpha})(1+|x|)+M_2(1+|c|^{\alpha}) \\ |f_n(t,x,c)| &\le M_1(1+|c|^{\alpha_n})(1+|x|)+M_2(1+|c|^{\alpha_n}) \end{aligned}$$

for every $(t,x,c)\in[0,T]\times\mathbb{R}^k\times\mathbb{R}^m$;

$\mathbf{A}_3$ *the maps l and l_n are continuous on $[0,T]\times\mathbb{R}^k\times\mathbb{R}^m$ and, for every compact subset $Q\subset\mathbb{R}^k$, there is a modulus ρ_l satisfying*

$$\begin{aligned} |l(t_1,x_1,c)-l(t_2,x_2,c)| &\le (1+|c|^{\beta})\rho_l(|(t_1,x_1)-(t_2,x_2)|) \\ |l_n(t_1,x_1,c)-l_n(t_2,x_2,c)| &\le (1+|c|^{\beta})\rho_l(|(t_1,x_1)-(t_2,x_2)|) \end{aligned}$$

for every $(t_1,x_1,c),(t_2,x_2,c)\in[0,T]\times Q\times\mathbb{R}^m$ and $n\in\mathbb{N}$;

A_4 *there exist positive constants Λ_0 and Λ_1 such that the following* coercivity conditions

$$\begin{aligned} l(t,x,c) &\geq \Lambda_0|c|^\beta - \Lambda_1 \\ l_n(t,x,c) &\geq \Lambda_0|c|^\beta - \Lambda_1\,, \end{aligned}$$

are verified for every $(t,x,c) \in [0,T] \times \mathbb{R}^k \times \mathbb{R}^m$ and every $n \in \mathbb{N}$;

A_5 *the maps g, g_n are bounded below by a constant $\overline{G}$ and, for every compact $Q \subset \mathbb{R}^k$, there is a modulus ρ_g such that*

$$\begin{aligned} |g_n(x_1) - g_n(x_2)| &\leq \rho_g(|x_1 - x_2|)\,, \\ |g(x_1) - g(x_2)| &\leq \rho_g(|x_1 - x_2|) \end{aligned}$$

for every $x_1, x_2 \in Q$.

When $\alpha = \beta$, we also assume a condition of regularity of f and l at infinity in the variable c. Precisely, we posit the existence of continuous functions f^∞ and l^∞, the *recessions functions* of f and l, respectively, verifying

$$\begin{aligned} \lim_{r\to 0} r^\beta f(t,x,r^{-1}w) &\doteq f^\infty(t,x,w) \\ \lim_{r\to 0} r^\beta l(t,x,r^{-1}w) &\doteq l^\infty(t,x,w)\,, \end{aligned}$$

on compact sets of $[0,T] \times \mathbb{R}^k \times \mathbb{R}^m$ (e.g., if $f(t,x,c) = f_0(t,x) + f_1(t,x)|c| + f_2(t,x)|c|^2$ then $f^\infty(t,x,w) = f_2(t,x)|w|^2$). When $\alpha_n = \beta$ we likewise assume the existence of the recession functions f_n^∞, l_n^∞, respectively.

Theorem 1 below is the main result of the paper and concerns the convergence of the value functions V_n to V. We point out that, unlike previous results on this subject (see [1]), the triples (f_n, l_n, g_n) are allowed to tend to (f, l, g) not uniformly with respect to x and c.

THEOREM 1. *Let us assume that for every set $[0,T] \times Q$, where Q is a compact subset of $\mathbb{R}^k$, there exists a function $\epsilon : \mathbb{N} \to [0,\infty)$ infinitesimal for $n \to \infty$ such that*

$$|f_n(t,x,c) - f(t,x,c)| \leq \epsilon(n)(1 + |c|^\beta)\,, \tag{4}$$

$$|l_n(t,x,c) - l(t,x,c)| \leq \epsilon(n)(1 + |c|^\beta) \tag{5}$$

for $(t,x,c) \in [0,T] \times Q \times \mathbb{R}^m$ and

$$|g_n(x) - g(x)| \leq \epsilon(n)$$

for every $x \in Q$. Then the value functions V_n converge uniformly, as n tends to ∞, to V on compact subsets of $[0,T] \times \mathbb{R}^k$.

This theorem will be proved in Section 4 via some arguments which rely on the fact that the considered value functions are solutions of suitable Hamilton-Jacobi-Bellman equations. Actually, due to the non standard growth properties of the data, the Hamiltonians involved in these equations do not satisfy a uniform growth assumption in the adjoint variable which is shared by most of the uniqueness results existing in literature. In a recent paper [8] we have established some uniqueness and regularity results for this kind of equations. In the next section we recall briefly the points of this investigation that turn out to be essential in the proof of Theorem 1.

3. Reparameterizations and Bellman equations

The contents of this section thoroughly relies on the results of [8]. Let us embed the unperturbed and the perturbed problems in a class of extended problems which have the advantage of involving only bounded controls. There is a reparameterization argument behind this embedding which allows one to transform a L^β constraint (implicitly imposed by the coercivity assumptions) into a L^∞ constraint.

Let us introduce the extended fields

$$\overline{f}(t,x,w_0,w) \doteq \begin{cases} f\left(t,x,\frac{w}{w_0}\right)\cdot w_0^\beta & \text{if } w_0 \neq 0 \\ f^\infty(t,x,v,w) & \text{if } w_0 = 0 \text{ and } \alpha = \beta \end{cases}$$

and

$$\overline{l}(t,x,w_0,w) \doteq \begin{cases} l\left(t,x,\frac{w}{w_0}\right)\cdot w_0^\beta & \text{if } w_0 \neq 0 \\ l^\infty(t,x,v,w) & \text{if } w_0 = 0 \text{ and } \alpha = \beta\,. \end{cases}$$

Similarly, for every n we define the extended fields $\overline{f_n}$ and $\overline{l_n}$ of f_n and l_n, respectively. Hypotheses $\mathbf{A}_1$-$\mathbf{A}_5$ imply the following properties for the maps $\overline{f_n}$, $\overline{l_n}$, $\overline{f}$, and $\overline{l}$.

PROPOSITION 1. **(i)** *The functions $\overline{f_n}$, $\overline{l_n}$, $\overline{f}$, and $\overline{l}$ are continuous on $[0,T]\times\mathbb{R}^k\times[0,+\infty[\times\mathbb{R}^m$ and for every compact $Q\subset\mathbb{R}^k$ we have*

$$(A_{e_1})\quad \begin{aligned} |\overline{f}(t_1,x_1,w_0,w)-\overline{f}(t_2,x_2,w_0,w)| &\le (w_0^\alpha+|w|^\alpha)w_0^{\beta-\alpha}(L|x_1-x_2| \\ &\quad + \rho_f(|t_1-t_2|))\,, \\ |\overline{f}_n(t_1,x_1,w_0,w)-\overline{f}_n(t_2,x_2,w_0,w)| &\le (w_0^{\alpha_n}+|w|^{\alpha_n})w_0^{\beta-\alpha_n}(L|x_1-x_2| \\ &\quad + \rho_f(|t_1-t_2|)) \end{aligned}$$

and

$$(A_{e_3})\quad \begin{aligned} |\overline{l}(t_1,x_1,w_0,w)-\overline{l}(t_2,x_2,w_0,w)| &\le (w_0^\beta+|w|^\beta)\rho_l(|(t_1,x_1)-(t_2,x_2)|)\,, \\ |\overline{l}_n(t_1,x_1,w_0,w)-\overline{l}_n(t_2,x_2,w_0,w)| &\le (w_0^\beta+|w|^\beta)\rho_l(|(t_1,x_1)-(t_2,x_2)|) \end{aligned}$$

$\forall (t_1,x_1,w_0,w),(t_2,x_2,w_0,w)\in[0,T]\times\mathbb{R}^k\times[0,+\infty[\times\mathbb{R}^m$, where α, α_n, β, L, ρ_f, and ρ_l are the same as in assumptions $\mathbf{A}_1$ and $\mathbf{A}_3$.

Moreover,

$$(A_{e_2})\quad \begin{aligned} |\overline{f}(t,x,w_0,w)| &\le (w_0^\alpha+|w|^\alpha)w_0^{\beta-\alpha}(M_1(1+|x|)+M_2)\,, \\ |\overline{f}_n(t,x,w_0,w)| &\le (w_0^{\alpha_n}+|w|^{\alpha_n})w_0^{\beta-\alpha_n}(M_1(1+|x|)+M_2) \end{aligned}$$

and

$$(A_{e_4})\quad \begin{aligned} \overline{l}(t,x,w_0,w) &\ge \Lambda_0|w|^\beta-\Lambda_1|w_0|^\beta\,, \\ \overline{l}_n(t,x,w_0,w) &\ge \Lambda_0|w|^\beta-\Lambda_1|w_0|^\beta \end{aligned}$$

$\forall (t,x,w_0,w)\in[0,T]\times\mathbb{R}^k\times[0,+\infty[\times\mathbb{R}^m$, where M_1, M_2, Λ_0 and Λ_1 are the same as in $\mathbf{A}_2$ and $\mathbf{A}_4$.

(ii) (Positive homogeneity in (w_0, w)). *The map* $\overline{f}$, $\overline{l}$, $\overline{f}_n$, *and* $\overline{l}_n$ *are positively homogeneous of degree* β *in* (w_0, w)*, that is,*

$$\overline{f}(t, x, rw_0, rw) = r^\beta \overline{f}(t, x, w_0, w), \quad \overline{l}(t, x, rw_0, rw) = r^\beta \overline{l}(t, x, w_0, w)$$
$$\overline{f}_n(t, x, rw_0, rw) = r^\beta \overline{f}_n(t, x, w_0, w), \overline{l}_n(t, x, rw_0, rw) = r^\beta \overline{l}_n(t, x, w_0, w)$$

$\forall r > 0$, $\forall (t, x, w_0, w) \in [0, T] \times \mathbb{R}^k \times]0, +\infty[\times \mathbb{R}^m$.

For every $\bar{t} \in [0, T]$ let us introduce the following sets of *space-time controls*

$$\Gamma(\bar{t}) \doteq \left\{ (w_0, w) \in \mathcal{B}([0, 1], [0, +\infty) \times \mathbb{R}^m) \text{ such that } \bar{t} + \int_0^1 w_0^\beta(s)\, ds = T \right\}$$

and

$$\Gamma^+(\bar{t}) \doteq \left\{ (w_0, w) \in \Gamma(\bar{t}) \text{ such that } w_0 > 0 \text{ a.e.} \right\}$$

where $\mathcal{B}([0, 1], [0, +\infty)\times\mathbb{R}^m)$ is the set of L^∞, Borel maps, which take values in $[0, +\infty[\times\mathbb{R}^m$. If $\alpha < \beta$ [resp.$\alpha = \beta$], for every $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^k$ and every $(w_0, w) \in \Gamma^+(\bar{t})$ [resp. $(w_0, w) \in \Gamma(\bar{t})$], let us denote by $(t, y)_{(\bar{t},\bar{x})}[w_0, w](\cdot)$ the solution of the (*extended*) Cauchy problem

$$(E_e) \qquad \begin{cases} t'(s) = w_0^\beta(s) \\ y'(s) = \overline{f}(t(s), y(s), w_0(s), w(s)) \\ (t(0), y(0)) = (\bar{t}, \bar{x}), \end{cases}$$

where the parameter s belongs to the interval $[0, 1]$ and the prime denotes differentiation with respect to s. When the initial conditions are meant by the context we shall write $(t, y)[w_0, w](\cdot)$ instead of $(t, y)_{(\bar{t},\bar{x})}[w_0, w](\cdot)$. Let us consider the following (*extended*) cost functional

$$J_e(\bar{t}, \bar{x}, w_0, w) \doteq \int_0^1 \overline{l}\,((t, y)[w_0, w], w_0, w)\,(s)\, ds + g(y[w_0, w](1))$$

and the corresponding (*extended*) value function

$$V_e : [0, T] \times \mathbb{R}^k \to \mathbb{R}$$
$$V_e(\bar{t}, \bar{x}) \doteq \inf_{(w_0, w)\in\Gamma(\bar{t})} J_e(\bar{t}, \bar{x}, w_0, w)\,.$$

Similarly, for every $n \in \mathbb{N}$, for every $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^k$ and every $(w_0, w) \in \Gamma(\bar{t})$ let us introduce the system

$$(E_{e_n}) \qquad \begin{cases} t'(s) = w_0^\beta(s) \\ y'(s) = \overline{f}_n(t(s), y(s), w_0(s), w(s)) \quad s \in [0, 1] \\ (t(0), y(0)) = (\bar{t}, \bar{x}), \end{cases}$$

and let us denote its solution by $(t, y)^n_{(\bar{t},\bar{x})}[w_0, w](\cdot)$. Let us introduce the cost functionals

$$J_{e_n}(\bar{t}, \bar{x}, w_0, w) \doteq \int_0^1 \overline{l}_n\left((t, y)^n_{(\bar{t},\bar{x})}, w_0, w\right)(s)\, ds + g_n(y_n[w_0, w](1))$$

and the corresponding value functions

$$\begin{gathered} V_{e_n} : [0,T] \times \mathbb{R}^k \to \mathbb{R} \\ V_{e_n}(\bar{t},\bar{x}) \doteq \inf_{(w_0,w)\in\Gamma(\bar{t})} J_{e_n}(\bar{t},\bar{x},w_0,w)\,. \end{gathered}$$

Next theorem establishes the coincidence of the value functions of the original problems with those of the extended problems.

THEOREM 2. *Assume* $\mathbf{A}_1$-$\mathbf{A}_5$.

(i) *For every* $(t,x) \in [0,T[\times\mathbb{R}^k$ *and for every* $n \in \mathbb{N}$ *one has* $V_e(t,x) = V(t,x)$*; and* $V_{e_n}(t,x) = V_n(t,x)$*;*

(ii) *the maps* V_e *and* V_{e_n} *are continuous on* $[0,T]\times\mathbb{R}^k$.

Thanks to this theorem – which, in particular, implies that V and V_n can be continuously extended on $[0,T]\times\mathbb{R}^k$ – the problem of the convergence of the V_n is transformed in the analogous problem for the V_{e_n}.

We now recall that each of these value functions is the unique solution of a suitable boundary value problem. This is a consequence of the comparison theorem below. To state these results, let us introduce the *extended Hamiltonians*

(6)

$$H_e(t,x,p_0,p) \doteq \sup_{(w_0,w)\in([0,+\infty[\times\mathbb{R}^m)\cap S_m^+} \{-p_0 w_0^\beta - \langle p, \overline{f}(t,x,w_0,w)\rangle - \overline{l}(t,x,w_0,w)\}$$

where $S_m^+ \doteq \{(w_0,w) \in [0,+\infty[\times\mathbb{R}^m : |(w_0,w)| = 1\}$,

$$H_{e_n}(t,x,p_0,p) \doteq \sup_{(w_0,w)\in([0,+\infty[\times\mathbb{R}^m)\cap S_m^+} \{-p_0 w_0^\beta - \langle p, \overline{f_n}(t,x,w_0,w)\rangle - \overline{l_n}(t,x,w_0,w)\}\,,$$

and the corresponding Hamilton-Jacobi-Bellman equations

$$(HJ_e) \qquad H_e(t,x,u_t,u_x) = 0\,,$$

$$(HJ_{e_n}) \qquad H_{e_n}(t,x,u_t,u_x) = 0\,.$$

For the sake of self consistency let us recall the definition of (possibly discontinuous) viscosity solution, which was introduced by H. Ishii in [5].

Given a function $F : \mathcal{Q} \to \mathbb{R}$, $\mathcal{Q} \subseteq \mathbb{R}^k$, let us consider the *upper and lower semicontinuous envelopes,* defined by

$$\begin{aligned} F^*(x) &\doteq \lim_{r\to 0^+} \sup\{F(y) : y \in \mathcal{Q},\ |x-y| \le r\}\,, \\ F_*(x) &\doteq \lim_{r\to 0^+} \inf\{F(y) : y \in \mathcal{Q},\ |x-y| \le r\}\,, \qquad x \in \overline{\mathcal{Q}}\,, \end{aligned}$$

respectively. Of course, F^* is upper semicontinuous and F_* is lower semicontinuous.

DEFINITION 1. *Let* E *be a subset of* $\mathbb{R}^s$ *and let* G *be a real map, the* Hamiltonian, *defined on* $E \times \mathbb{R} \times \mathbb{R}^s$. *An upper[resp. lower]-semicontinuous function* u *is a viscosity subsolution [resp. supersolution] of*

$$(7) \qquad G(y,u,u_y) = 0$$

at $y \in E$ if for every $\phi \in \mathcal{C}^1(\mathbb{R}^s)$ such that y is a local maximum [resp. minimum] point of $u - \phi$ on E one has

$$G_*(y, \phi(y), \phi_y(y)) \leq 0$$

[resp.

$$G^*(y, \phi(y), \phi_y(y)) \geq 0] .$$

A function u is a viscosity solution of (7) at $y \in E$ if u^* is a viscosity subsolution at y and u_* is a viscosity supersolution at y.

THEOREM 3 (COMPARISON). *Assume $\mathbf{A}_1$-$\mathbf{A}_5$. Let $u_1 : [0, T] \times \mathbb{R}^k \to \mathbb{R}$ be an upper semicontinuous, bounded below, viscosity subsolution of (HJ_e) in $]0, T[\times\mathbb{R}^k$, continuous on $(\{0\} \times \mathbb{R}^k) \cup (\{T\} \times \mathbb{R}^k)$. Let $u_2 : [0, T] \times \mathbb{R}^k \to \mathbb{R}$ be a lower semicontinuous, bounded below, viscosity supersolution of (HJ_e) in $[0, T[\times\mathbb{R}^k$. For every $x \in \mathbb{R}^k$, assume that*

$$\begin{cases} u_1(T, x) \leq u_2(T, x) \\ \text{or} \\ u_2 \text{ is a viscosity supersolution of } (HJ_e) \text{ at } (T, x) . \end{cases}$$

Then

$$u_1(t, x) \leq u_2(t, x) \quad \forall (t, x) \in [0, T] \times \mathbb{R}^k .$$

The same statement holds true for the equations (HJ_{e_n}).

As a consequence of this theorem and of a suitable dynamic programming principle for the extended problems one can prove the following:

THEOREM 4. *The value function V_e is the unique map which*

i) is continuous on $(\{0\} \times \mathbb{R}^k) \cup (\{T\} \times \mathbb{R}^k)$;

ii) is a viscosity solution of (HJ_e) in $]0, T[\times\mathbb{R}^k$;

iii) satisfies the following mixed type boundary condition:

$$(BC_{e_m}) \qquad \begin{cases} V_e(T, x) \leq g(x) \quad \forall x \in \mathbb{R}^k \text{ and} \\ \begin{cases} V_e(T, x) = g(x) \\ \text{or} \\ V_e \text{ is a viscosity supersolution of } (HJ_e) \text{ at } (T, x) . \end{cases} \end{cases}$$

Once we replace (HJ_e) by (HJ_{e_n}), the same statement holds true for the maps V_{e_n}.

Finally let us recall a regularity result which will be useful in the proof of Theorem 1.

THEOREM 5. *Assume $\mathbf{A}_1$-$\mathbf{A}_5$ and fix $R > 0$. Then there exists $R' \geq R$ and positive constants C_1, C_2 such that*

$$|V_e(t, x_1) - V_e(t, x_2)| \leq C_1 \rho_l(C_2|x_2 - x_1|) + \rho_g(C_2|x_2 - x_1|)$$

for every $(t, x_1)\,(t, x_2) \in [0, T] \times B[0; R]$, where ρ_l and ρ_g are the modulus appearing in $\mathbf{A}_3$ and the modulus of uniform continuity of g, respectively, corresponding to the compact $[0, T] \times B[0; R']$. Moreover for every $\bar{t} \in [0, T[$ one has

$$|V_e(t, x) - V_e(\bar{t}, x)| \leq \eta_{\bar{t}}(|t - \bar{t}|)$$

for every $(t,x) \in [0,T[\times B[0;R]$*, where* $\eta_{\bar t}$ *is a suitable modulus, and for every* s*,* $\bar t \to \eta_{\bar t}(s)$ *is an increasing map. The same statement holds true for the maps* V_{e_n}*, with the same* $\eta_{\bar t}$*.*

REMARK 1. We do not need, for our purposes, an explicit expression of $\eta_{\bar t}$, which, however, can be found in [8]. Also in that paper sharper regularity results are established. Finally let us point out that though an estimate like the second one in Theorem 5 is not available for $\bar t = T$ the map V_e is continuous on $\{T\} \times \mathbb{R}^k$, (see Theorem 2).

4. Proof of the convergence theorem

Proof of Theorem 1. In view of Theorem 2 it is sufficient to show that the maps V_{e_n} converge to V_e. Observe that the assumptions (4), (5) imply

$$|\overline{f}_n(t,x,w_0,w) - \overline{f}(t,x,w_0,w)| \le \epsilon(n)(w_0^\beta + |w|^\beta) \tag{8}$$

and

$$|\bar l_n(t,x,w_0,w) - \bar l(t,x,w_0,w)| \le \epsilon(n)(w_0^\beta + |w|^\beta)\,, \tag{9}$$

for every $(t,x,w_0,w) \in [0,T] \times Q \times [0,\infty[\times\mathbb{R}^m$ and every $n \in \mathbb{N}$.

Moreover, by the coercivity condition $\mathbf{A}_{e_4}$ and by the obvious local uniform boundedness of V_{e_n}, and V_e when the initial conditions are taken in a ball $B[0,R]$ it is not restrictive to consider only those space time controls such that

$$\int_0^1 (w_0(s) + |w(s)|)^\beta \, ds \le K_R \tag{10}$$

where K_R is a suitable constant depending on R. By Hölder's inequality we have also that

$$\int_0^1 (w_0(s) + |w(s)|)^{\alpha_n} w_0(s)^{\beta-\alpha_n} \, ds \le (T+1)(K_R+1)\,.$$

Hence by Gronwall's Lemma, we can assume that there exists a ball $B[0,R'] \subset \mathbb{R}^k$ containing all the trajectories issuing from $B[0,R]$.

Let us fix $\overline{T} < T$: by Theorem 5 the maps V_{e_n} are equicontinuous and equibounded on $[0,\overline{T}] \times B[0,R]$, so we can apply Ascoli-Arzela's Theorem to get a subsequence of V_{e_n}, still denoted by V_{e_n}, converging to a continuous function. Actually by taking R larger and larger, via a standard diagonal procedure we can assume that the V_{e_n} converge to a continuous function $\mathcal{V} : [0,T] \times \mathbb{R}^k \to \mathbb{R}$, uniformly on compact sets of $[0,\overline{T}] \times \mathbb{R}^k$. Now, for every $(t,x) \in [0,T] \times \mathbb{R}^k$, let us consider the *weak limits*

$$\overline{V}(t,x) \doteq \limsup_{\substack{n\to\infty \\ (s,y)\to(t,x) \\ (s,y)\in[0,T]\times\mathbb{R}^k}} V_{e_n}(s,y)$$

and

$$\underline{V}(t,x) \doteq \liminf_{\substack{n\to\infty \\ (s,y)\to(t,x) \\ (s,y)\in[0,T]\times\mathbb{R}^k}} V_{e_n}(s,y)\,.$$

Our goal is to apply a method (see [2]) based on the application of the comparison theorem (see Theorem 3) to these weak limits. Let us observe that both $\overline{V}$ and $\underline{V}$ coincide with $\mathcal{V}$ on the boundary $\{0\} \times \mathbb{R}^k$: in particular they are continuous on $\{0\} \times \mathbb{R}^k$. Since the Hamiltonians H_{e_n} converge to H_e uniformly on compact subsets of $[0, T] \times \mathbb{R}^k \times \mathbb{R} \times \mathbb{R}^k$, standard arguments imply that $\overline{V}$ is a (upper semicontinuous) viscosity subsolution of (HJ_e) in $[0, T) \times \mathbb{R}^k$, while $\underline{V}$ is a (lower semicontinuous) viscosity supersolution of (HJ_e) in $[0, T) \times \mathbb{R}^k$. Hence the convergence result is proven as soon as one shows that $\overline{V} \leq \underline{V}$ in $[0, T] \times \mathbb{R}^k$. For this purpose it is sufficient to show that $\overline{V}$ and $\underline{V}$ verify the hypotheses of Theorem 3. Actually the only hypothesis which is left to be verified is the one concerning the boundary subset $\{T\} \times \mathbb{R}^k$. We claim that

$$\lim_{\substack{n\to\infty \\ (s,y)\to(T,x) \\ (s,y)\in[0,T]\times\mathbb{R}^k}} V_{e_n}(s, y) = V_e(T, x) \tag{11}$$

which implies $\overline{V}(T, \cdot) = \underline{V}(T, \cdot) = V_e(T, \cdot)$. In particular the maps $\overline{V}(T, \cdot)$ and $\underline{V}(T, \cdot)$ turn out to be continuous, so all assumptions of Theorem 3 are verified. The remaining part of this proof is thus devoted to prove (11). Let us consider $x_1, x_2 \in B[0, R]$ and controls $(0, w_n) \in \Gamma(T)$ such that, setting $(t_n, x_n) \doteq (t, y)^n_{(T,x_1)}[0, w_n](\cdot)$, we have

$$V_{e_n}(T, x_1) \geq \int_0^1 \bar{l}_n(t_n, x_n, 0, w_n)(s)\, ds + g_n(x_n(1)) - \epsilon\,.$$

Hence, setting $(\tilde{t}_n, \tilde{x}_n) \doteq (t, y)_{(T,x_2)}[0, w_n](\cdot)$ and noticing that $\tilde{t}_n(s) = t_n(s) = T\ \forall s \in [0, 1]$, we have

$$\begin{aligned} V_e(T, x_2) - V_{e_n}(T, x_1) \quad &\leq \quad \int_0^1 \bar{l}(T, \tilde{x}_n, 0, w_n)(s)\, ds + g_n(\tilde{x}_n(1)) \\ &\qquad - \int_0^1 \bar{l}_n(T, x_n, 0, w_n)(s)\, ds - g_n(x_n(1)) + \epsilon \\ &\leq \quad \int_0^1 |w_n(s)|^\beta [\epsilon(n) + \rho_l(|\tilde{x}_n(s) - x_n(s)|)]\, ds \\ &\qquad + \rho_g(|\tilde{x}_n(1) - x_n(1)|) + \epsilon(n) + \epsilon\,, \end{aligned}$$

where $\epsilon(n)$, ρ_l and ρ_g (see $\mathbf{A}_3$ and $\mathbf{A}_5$) are determined with reference to the compact subset $Q = B[0, R']$. If $L_{R'}$ is the determination of L in (A_{e_1}) for $B[0, R']$ then

$$|\tilde{x}_n(s) - x_n(s)| \leq (|x_1 - x_2| + \epsilon(n)(T + 1)(K_R + 1))e^{L_{R'}(T+1)(K_R+1)}\,.$$

This, together with the fact that a similar inequality can be proved (in a similar way) when the roles of V_e and V_{e_n} are interchanged, implies

$$\begin{aligned} |V_e(T, x_2) - V_{e_n}(T, x_1)| \leq K_R \rho_l &\Big[(|x_1 - x_2| \\ &+ \epsilon(n)(T + 1)(K_R + 1))e^{L_{R'}(T+1)(K_R+1)}\Big] \\ &+ \rho_g \Big[(|x_1 - x_2| + \epsilon(n)(T + 1)(K_R + 1))e^{L_{R'}(T+1)(K_R+1)}\Big] \\ &+ (K_R + 1)\epsilon(n)\,. \end{aligned} \tag{12}$$

Now, for $\tau \leq T$, let us estimate the difference $V_{e_n}(\tau, x) - V_e(T, x)$, assuming that this difference is non negative. Let us set $(t_n, x_n)(\cdot) \doteq (t, y)^n_{(\tau,x)}(\tilde{w}_0, 0)(\cdot)$ with $\tilde{w}_0(s) \doteq (T - \tau)^{\frac{1}{\beta}}$ $\forall s \in [0, 1]$. Then the Dynamic Programming Principle

$$V_{e_n}(\tau, x) - V_e(T, x) \leq \int_0^1 \overline{l}_n(t_n, x_n, \tilde{w}_0, 0)(s)\, ds + V_{e_n}(T, x_n(1)) - V_e(T, x)\,.$$

If $M \doteq \max\{M_1 + M_2, 1\}$, by (A_{e_2}) we have $|x_n(1) - x| \leq M(1 + R')|T - \tau|$. Hence, if $K'_R \geq \max_{\substack{(t,x)\in[0,T]\times B[0,R'] \\ n\in\mathbb{N}}} \overline{l}_n(t, x, 1, 0)$, by the positive homogeneity of $\overline{l}_n$ and by the first part of the proof we obtain

$$V_{e_n}(\tau, x) - V_e(T, x) \leq K'_R|T - \tau| + \sigma_n(|T - \tau|) \tag{13}$$

where

$$\begin{aligned}\sigma_n(s) =& K_R\rho_l[(M(1 + R')s + \epsilon(n)(T + 1)(K_R + 1))e^{L_{R'}(T+1)(K_R+1)}] \\ &+ \rho_g[(M(1 + R')s + \epsilon(n)(T + 1)(K_R + 1))e^{L_{R'}(T+1)(K_R+1)}] + (K_R + 1)\epsilon(n)\,.\end{aligned}$$

Now let us estimate the difference $V_e(T, x) - V_{e_n}(\tau, x)$, assuming it non negative. Let us consider a sequence of controls $(w_{0_n}, w_n) \in \Gamma(\tau)$ such that, setting $(t_n, x_n) \doteq (t, y)^n_{(\tau,x)}[w_{0_n}, w_n](\cdot)$, one has

$$V_{e_n}(\tau, x) \geq \int_0^1 \overline{l}_n(t_n, x_n, w_{0_n}, w_n)(s)\, ds + g_n(x_n(1)) - \epsilon\,.$$

Then the controls $(0, w_n)$ belong to $\Gamma(T)$, and, setting $(\tilde{t}_n, \tilde{x}_n) \doteq (t, y)[0, w_n](\cdot)$, we obtain

$$\begin{aligned}V_e(T, x) - V_{e_n}(\tau, x) \leq& \int_0^1 \overline{l}(\tilde{t}_n, \tilde{x}_n, 0, w_n)(s)\, ds + g(\tilde{x}_n(1)) \\ &- \int_0^1 \overline{l}_n(t_n, x_n, w_{0_n}, w_n)(s)\, ds - g_n(x_n(1)) + \epsilon\end{aligned} \tag{14}$$

for every $n \in \mathbb{N}$. Now one has

$$\begin{aligned}|x_n(s) - \tilde{x}_n(s)| \leq& \int_0^1 |\overline{f}_n(t_n, x_n, w_{0_n}, w_n)(s) - \overline{f}(\tilde{t}_n, \tilde{x}_n, 0, w_n)(s)|\, ds \\ \leq& \int_0^1 |\overline{f}_n(t_n, x_n, w_{0_n}, w_n)(s) - \overline{f}(t_n, x_n, w_{0_n}, w_n)(s)|\, ds \\ &+ \int_0^1 |\overline{f}(t_n, x_n, w_{0_n}, w_n)(s) - \overline{f}(\tilde{t}_n, x_n, w_{0_n}, w_n)(s)|\, ds \\ &+ \int_0^1 |\overline{f}(\tilde{t}_n, x_n, w_{0_n}, w_n)(s) - \overline{f}(\tilde{t}_n, \tilde{x}_n, w_{0_n}, w_n)(s)|\, ds \\ &+ \int_0^1 |\overline{f}(\tilde{t}_n, \tilde{x}_n, w_{0_n}, w_n)(s) - \overline{f}(\tilde{t}_n, \tilde{x}_n, 0, w_n)(s)|\, ds\,.\end{aligned} \tag{15}$$

for all $s \in [0, 1]$. In view of the parameter-free character of the system (see e.g. [7] for the case $\alpha = \beta = 1$), it is easy to show that one can transform the integral bound (10) into the pointwise bound

$$|(w_0, w)(s)| \leq \tilde{K}_R\ \forall s \in [0, 1]\,,$$

where $\tilde{K}_R$ is a constant depending on R. Therefore, in view of basic continuity properties of the composition operator, there exists a modulus ρ such that the last integral in the above inequality is smaller than or equal to $\rho(|T-\tau|)$. Therefore, applying Gronwall's inequality to (15) we obtain

$$
\begin{aligned}
|x_n(s)-\tilde{x}_n(s)| \le &(T+1)(K_R+1)[\epsilon(n)+\rho_f(|T-\tau|) \\
&+\rho(|T-\tau|)]e^{L_{R'}(T+1)(K_R+1)}\,.
\end{aligned}
\tag{16}
$$

Hence (14) yields

$$
\begin{aligned}
V_e(T,x)-V_{e_n}(\tau,x) \le &\int_0^1 |\bar{l}(\tilde{t}_n,\tilde{x}_n,0,w_n)(s)-\bar{l}(\tilde{t}_n,x_n,0,w_n)(s)|\,ds \\
&+\int_0^1 |\bar{l}(\tilde{t}_n,x_n,0,w_n)(s)|\,ds-\bar{l}(t_n,x_n,0,w_n)(s)|\,ds \\
&+\int_0^1 |\bar{l}(t_n,x_n,0,w_n)(s)-\bar{l}(t_n,x_n,w_{0_n},w_n)(s)\,ds \\
&+\int_0^1 |\bar{l}(t_n,x_n,w_{0_n},w_n)(s)-\bar{l}_n(t_n,x_n,w_{0_n},w_n)(s)|\,ds \\
&+\rho_g(|x_n(1)-\tilde{x}_n(1)|)\,.
\end{aligned}
\tag{17}
$$

Again, an argument based on the continuity properties of the composition operator allows one to conclude that there exists a modulus $\tilde{\rho}$ such that

$$
\int_0^1 |\bar{l}(t_n,x_n,0,w_n)(s)-\bar{l}(t_n,x_n,w_{0_n},w_n)(s)|\,ds \le \tilde{\rho}(|T-\tau|)
$$

Therefore, plugging (15) into (16), we obtain

$$
\begin{aligned}
V_e(T,x)-V_{e_n}(\tau,x) \le &P_R(\rho_l+\rho_g)[P_R(\epsilon(n)+\rho_f(|T-\tau|)+\rho(|T-\tau|))e^{L_{R'}P_R}] \\
&+P_R[\rho_l(|T-\tau|)+\epsilon(n)]+\tilde{\rho}(|T-\tau|)\,,
\end{aligned}
\tag{18}
$$

where $P_R \doteq (T+1)(K_R+1)$.

Estimates (12), (13) and (18) imply the claim, so the theorem is proved.

□

5. Implementing optimal controls in the presence of perturbations

As an application of Theorem 1, Theorems 6 and 7 below provide, for the special case of the linear quadratic problem, an answer to the general questions $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively. Let us remark that the perturbation we consider is not the most general among those allowed by Theorem 1. However, it well illustrates the degree of improvement with respect to previous results concerning questions like $\mathbf{Q}_1$ and $\mathbf{Q}_2$ (see Introduction). Let us also remark that the linear-quadratic problem is just a model case. Indeed, it is evident that Theorem 6 below holds also if we replace the linear-quadratic problem with a problem that (satisfy hypotheses $\mathbf{A}_1$-$\mathbf{A}_5$, (4), and (5) and) admits an optimal L^β control, while Theorem 7 is still valid for any problem for which (f is Lipschitz in c uniformly for (t,x) in a compact subset of $[0,T]\times\mathbb{R}^k$ and) a Lipschitz continuous feedback control $c(x)$ does exist.

Let us be more precise by stating that by linear-quadratic problem we mean here an optimal control problem as the ones considered in the previous sections, with

$$f = f(x,c) \doteq Ax + Bc \quad l(x,c) \doteq x^t Dx + c^t Ec \quad g(x) = x^t Sx\,,$$

where D, E, S are symmetric matrices (of suitable dimensions), D and S are nonnegative definite, F is positive definite, while no assumptions are made on A, B, C. Let us observe that the fields f, l, g satisfies hypotheses $\mathbf{A}_1$-$\mathbf{A}_5$. In particular, one has $\alpha = 1$, $\beta = 2$, and Λ_0 is the smallest eigenvalue of E.

Let us consider the following perturbations of the maps f, l, g:

$$\begin{aligned} f_n &\doteq Ax + Bc + \varphi_n(t,x,c) \\ l_n &\doteq x^t Dx + c^t Ec + \theta_n(t,x,c) \\ g_n &\doteq x^t Sx + \psi_n(x)\,. \end{aligned}$$

We assume that for each compact subset $Q \subset \mathbb{R}^k$ there exist a constant λ and moduli ρ and $\overline{\rho}$ such that:

i) *For every n, the map $\varphi_n : [0,T] \times \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^k$ is continuous and verifies*

$$|\varphi_n(t_1,x_1,c) - \varphi_n(t_2,x_2,c)| \le (1+|c|^{\alpha_n})(\Lambda|x_1 - x_2| + \rho(|t_1 - t_2|)$$

for all $(t_1,x_1,c), (t_2,x_2,c) \in [0,T] \times Q \times \mathbb{R}^m$ and for a suitable $\alpha_n \in [1,2]$ (varying with n and independent of Q).

ii) *There exist constants μ_1, μ_2 such that for every $n \in \mathbb{N}$ one has*

$$|\varphi_n(t,x,c)| \le \mu_1(1+|c|^{\alpha_n})(1+|x|) + \mu_2(1+|c|^{\alpha_n})$$

for every $(t,x,c) \in [0,T] \times \mathbb{R}^k \times \mathbb{R}^m$.

iii) *For every $n \in \mathbb{N}$, $\theta_n : [0,T] \times \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^k$ is continuous and verifies*

$$|\theta_n(t_1,x_1,c) - \theta_n(t_2,x_2,c)| \le (1+|c|^2)\overline{\rho}(|(t_1,x_1) - (t_2,x_2)|)$$

for every $(t_1,x_1,c), (t_2,x_2,c) \in [0,T] \times Q \times \mathbb{R}^m$.

iv) *There exist a (possibly negative) constant λ_0, strictly larger than the opposite of the smallest eigenvalue of E, and a positive constant λ_1 such that*

$$\theta_n(t,x,c) \ge \lambda_0|c|^2 - \lambda_1$$

for every n and every $(t,x,c) \in [0,T] \times \mathbb{R}^k \times \mathbb{R}^m$.

v) *$\psi_n : \mathbb{R}^k \to \mathbb{R}$ is continuous and $\psi_n \ge 0$.*

Moreover we assume that for every compact $Q \subset \mathbb{R}^k$ there exists a function $\epsilon : \mathbb{N} \to \mathbb{N}$, infinitesimal as $n \to \infty$ such that

$$\begin{aligned} |\phi_n(x,c)| &\le \epsilon(n)(1+|c|^2)\,, \\ |\theta_n(x,c)| &\le \epsilon(n)(1+|c|^2) \end{aligned}$$

for every $(x,c) \in Q \times \mathbb{R}^m$ and

$$|\psi_n(x)| \le \epsilon(n)$$

for every $x \in Q$.

REMARK 2. Let us observe that the above assumptions imply that the hypotheses of the convergence theorem (Theorem 1) are verified. Let us also point out that we allow α_n to be *equal* to $\beta(=2)$ (see Remark 1).

THEOREM 6 (OPEN LOOP). *Fix* $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^k$. *Assume that* $\overline{c}$ *is an optimal control for the unperturbed problem that is* $J(\bar{t}, \bar{x}, \overline{c}) = V(\bar{t}, \bar{x})$. *Then* $\overline{c}$ *is nearly optimal for the perturbed problem, i.e.,*

$$\lim_{n\to\infty} |J_n(\bar{t}, \bar{x}, \overline{c}) - V_{e_n}(\bar{t}, \bar{x})| = 0 . \tag{19}$$

Proof. As in the proof of Theorem 1, when the initial condition are taken in a ball $B[0, R]$, by the coercivity condition $\mathbf{A}_4$ we can consider only controls such that

$$\int_0^T (1 + |c(s)|)^2 \, ds \le K_R ,$$

where K_R is a suitable constant depending on R. Then, by Hölder's inequality we have also,

$$\int_0^T (1 + |c(s)|)^{\alpha_n} \, ds \le (K_R + 1)(T + 1)$$

which, by Gronwall's inequality, implies that there is a ball $B[0, R']$ which contains all the trajectories issuing from $B[0, R]$. Setting $x(\cdot) \doteq x_{(\bar{t},\bar{x})}[\overline{c}](\cdot)$ and $x_n(\cdot) \doteq x^n_{(\bar{t},\bar{x})}[\overline{c}](\cdot)$, we have

$$\begin{aligned} |x_n(t) - x(t)| &\le \int_{\bar{t}}^T |f_n(s, x_n(s), \overline{c}(s)) - f(s, x(s), \overline{c}(s))| \, ds \\ &\le \int_{\bar{t}}^T |Ax_n(s) - Ax(s) + \phi_n(x_n, \overline{c})(s)| \, ds \\ &\le (T + 1)(K_R + 1)\epsilon(n) + \|A\| \int_{\bar{t}}^T (|x_n(s) - x(s)|) \, ds \end{aligned} \tag{20}$$

where $\epsilon(n)$ is relative to $B[0, R]$ and $\|A\|$ is the operator norm of the matrix A. Hence, Gronwall's Lemma implies

$$|x_n(s) - x(s)| \le (T + 1)(K_R + 1)\epsilon(n)e^{\|A\|T} , \tag{21}$$

for every $t \in [\bar{t}, T]$. Since

$$\begin{aligned} |J_n(t, x, \overline{c}) - J(t, x, \overline{c})| \le &\int_{\bar{t}}^T |x_n(s)^t Dx_n(s) - x(s)^t Dx(s)| \, ds \\ &+ \int_{\bar{t}}^T |\theta_n(x_n(s), \overline{c}(s))| \, ds \\ &+ |x_n(T)^t Sx_n(T) - x(T)^t Sx(T)| + |\psi_n(x_n(T))| \\ \le &\|D\| \int_{\bar{t}}^T |x_n(s) - x(s)|(|x_n(s|) + |x(s)|) \, ds \\ &+ \|S\||x_n(s) - x(s)|(|x_n(T)| + |x(T)|) + \epsilon(n)(1 + K_R) , \end{aligned} \tag{22}$$

in view of estimate (22) and of Theorem 1, the theorem is proven.

□

THEOREM 7. *Let $c(x)$ be a locally Lipschitz continuous optimal feedback control for the unperturbed problem. Then this control is nearly optimal for the perturbed problem, that is*

$$\lim_{n\to\infty} |J_n(\bar{t}, \bar{x}, c) - V_{e_n}(\bar{t}, \bar{x})| = 0 .$$

Proof. If we denote by $x(\cdot)$ and $x_n(\cdot)$ the solutions to (E) to (E_n), respectively, corresponding to the feedback control $c(x)$, we obtain

$$\begin{aligned}
|x_n(t) - x(t)| &\leq \int_{\bar{t}}^{T} |f_n(s, x_n(s), c(x_n(s))) - f(s, x(s), c(x(s)))|\, ds \\
&\leq \int_{\bar{t}}^{T} |Ax_n(s) + Bc(x_n(s)) - Ax(s) - Bc(x(s))|\, ds \\
&\quad + \int_{\bar{t}}^{T} |\phi_n(x_n(s), c(x_n(s))|\, ds \\
&\leq \int_{\bar{t}}^{T} (\|A\| + \|B\|\gamma)(|x_n(s) - x(s)|\, ds + (T+1)(K_R + 1)\epsilon(n)
\end{aligned}$$

where γ is the Lipschitz constant of the map $c(x)$ corresponding to the compact set $B[0, R']$. Hence one has

$$|x_n(s) - x(s)| \leq (T+1)(K_R + 1)\epsilon(n)e^{T\|A\|+\|B\|\gamma} ,$$

and from here on one can proceed as in the proof of Theorem 6.

□

REMARK 3. As we have mentioned in the Introduction, when $\alpha_n = \beta$ it may happen that a perturbed problem (P_n) does not possess a minimum in the class of absolutely continuous trajectories. Indeed, due to the fact that the growth ratio $\frac{\beta}{\alpha_n}(= 1)$ is not greater than 1, the minimizing sequences could converge to a *discontinuous trajectory*. In this case, the possibility of implementing a control that is optimal for the unperturbed system – which is now assumed sufficiently coercive, that is, satisfying $\alpha < \beta$ – turns out to be of some interest whenever one is worried to avoid a discontinuous performance of the system under consideration.

To be more concrete, let us consider the very simple (linear-quadratic) minimum-problem where $l = x^2 + \Lambda_0 c^2$ and $f = 0$. In this case $\beta = 2$ and α can be taken equal to 1. Let us perturb this problem by taking $l_n = l$ and $f_n = f + \varphi_n = \varphi_n \doteq \frac{-c^2}{n}$. Observe that these perturbations give rise to *quadratic-quadratic* problems, that is problems where $\alpha_n = \beta = 2$. Let us consider the initial data $\bar{t} = 0$ and $\bar{x} > 0$. The constant map $x(t) = \bar{x}$ is the unique trajectory of the unperturbed system, so the control $\hat{c}(t) = 0 \forall t \in [0, T]$ turns out to be optimal. In view of Theorem 6 this control is nearly optimal for the perturbed problems as well. However, as soon $\bar{x}$ is sufficiently large and Λ_0 is sufficiently small with respect to $\frac{1}{n}$, an application of the Maximum Principle to the space-time extension of the perturbed system shows that the "optimal trajectory" of the perturbed problem is the concatenation of an "initial jump" (from $\bar{x}$ to a point $x_n \in]0, \bar{x}[$) and a suitable absolutely continuous map.

References

[1] BARDI M., DA LIO F., *On the Bellman equation for some unbounded control problems*, NoDEA – Nonlinear Differential Equations and Applications **4** (1997), 491–510.

[2] BARLES G., PERTHAME B., *Discontinuous solutions of deterministic optimal stopping time problems*, RAIRO Modél. Math. Anal. Numér. **21** (1987), 557–579.

[3] BENSOUSSAN A., *Perturbations methods in optimal control*, John Wiley and Sons, New York 1988.

[4] FLEMING W. H., RISHEL R. W., *Deterministic and stochastic optimal control*, Springer, New York 1975.

[5] ISHII H., *A boundary value problem of the Dirichlet type for Hamilton-Jacobi equations*, Ann. Sc. Norm. Sup. Pisa **(IV) 16** (1989), 105–135.

[6] ISHII H., *A comparison result for Hamilton-Jacobi equations without growth condition on solutions from above*, Appl. Anal. **67** (1997), 357–372.

[7] MOTTA M., RAMPAZZO F., *Nonlinear systems with unbounded controls and state constraints: a problem of proper extension*, NoDEA – Nonlinear Differential Equations and Applications **3** (1996), 191–216.

[8] RAMPAZZO F., SARTORI C., *Hamilton-Jacobi-Bellman equations with fast gradient-dependence*, preprint (1999).

AMS Subject Classification: 49K40, 35B37, 49L25, 49N25.

F. RAMPAZZO
Dipartimento di Matematica Pura ed Applicata
Via Belzoni 7 - 35131 Padova, Italy
e-mail: `rampazzo@math.unipd.it`

C. SARTORI
Dipartimento di Metodi e Modelli Matematici
per le Scienze Applicate
Via Belzoni 7 - 35131 Padova, Italy
e-mail: `sartori@dmsa.unipd.it`