ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Nonconvex problems in Control Theory and Calculus of Variations

CANDIDATE
Graziano Crasta

SUPERVISOR
Prof. Alberto Bressan

Thesis submitted for the degree of "Doctor Philosophiæ"

Academic Year 1993/94

SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

TRIESTE

Nonconvex problems
in Control Theory
and Calculus of Variations

CANDIDATE

Graziano Crasta

SUPERVISOR

Prof. Alberto Bressan

Thesis submitted for the degree of "Doctor Philosophiæ"

Academic Year 1993/94

Il presente lavoro costituisce la tesi presentata dal dott. Graziano Crasta, sotto la direzione del Prof. Alberto Bressan, al fine di ottenere l'attestato di ricerca post–universitaria "Doctor Philosophiæ" presso la S.I.S.S.A., Classe di Matematica, Settore di Analisi Funzionale ed Applicazioni. Ai sensi del Decreto del Ministro della Pubblica Istruzione 24.4.1987, n. 419, tale diploma è equipollente al titolo di "Dottore di Ricerca in Matematica".

Trieste, anno accademico 1993/94.

In ottemperanza a quanto previsto dall'art. 1 del Decreto Legislativo Luogotenenziale 31.8.1945, n. 660, le prescritte copie della presente pubblicazione sono state depositate presso la Procura della Repubblica di Trieste ed il Commissariato del Governo della Regione Friuli Venezia Giulia.

Acknowledgements

I wish to thank Prof. Alberto Bressan for his support and his interest in my work during my permanence at SISSA.

I am indebted to Prof. Arrigo Cellina for suggesting some of the problems treated in this thesis and for many stimulating conversations.

I also express my thanks to Dr. Benedetto Piccoli and Dr. Annalisa Malusa for the fruitful collaboration.

Finally I thank all my friends at SISSA for the nice time spent together, and I gratefully acknowledge the assistance of the SISSA staff.

Index

Chapter 4
An existence result for non-coercive non-convex problems

Chapter 5
Existence and regularity results for non–coercive variational problems

Introduction

It is well known that the classical approach to existence problems in differential inclusions, control theory and calculus of variations relies on suitable convexity assumptions.

A standard way to investigate non–convex problems is the so called *relaxation*, which consists of associating to the original problem a convex one, which we know to be solvable. The main difficulty of this technique is to construct a solution of the non–convex problem starting from solutions of the relaxed one.

In this thesis we use the *likelihood functional*, introduced in [10], in order to prove a selection theorem for non–convex differential inclusions, while we develop the technique of *bang–bang variations* for problems in control theory. Finally, we use the Lyapunov's theorem on the range of non–atomic vector measures in order to deal with non–convex problems in calculus of variations.

In Chapter 1 we consider a continuous multifunction $F: [0,T]\times \mathbb{R}^n \to 2^{\mathbb{R}^n}$ with compact, not necessarily convex values. If F is Lipschitz continuous, it was shown in [11] that there exists a measurable selection f of F such that, for every (t_0, x_0), the Cauchy problem

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0$$

has a unique Caratheodory solution $x(\cdot, t_0, x_0)$, depending continuously on (t_0, x_0).

We shall prove that the above selection f can be chosen so that $f(t,x)$ belongs to the extreme points ext $F(t,x)$ of $F(t,x)$, for all (t,x). Moreover, if $\varepsilon_0 > 0$ and a Lipschitz continuous selection f_0 of $\overline{\text{co}}\, F$ are given, then one can construct f with the following additional property. Denoting by $y(\cdot, t_0, x_0)$

the unique solution of

$$\dot{y}(t) = f_0(t, y(t)), \qquad y(t_0) = x_0,$$

for every $(t_0, x_0) \in [0,T] \times I\!R^n$ one has

$$\big|y(t, t_0, x_0) - x(t, t_0, x_0)\big| \leq \varepsilon_0 \qquad \forall t \in [0,T].$$

More generally, the result remains valid if F satisfies the following Lipschitz Selection Property:

(LSP) *For every (t,x), every $y \in \overline{\mathrm{co}}\, F(t,x)$ and $\varepsilon > 0$, there exists a Lipschitz selection ϕ of $\overline{\mathrm{co}}\, F$, defined on a neighborhood of (t,x), with $|\phi(t,x) - y| < \varepsilon$.*

We remark that, by [39, 44], every Lipschitz multifunction with compact values satisfies (LSP). Another interesting class, for which (LSP) holds, consists of those continuous multifunctions F whose values are compact and have convex closure with nonempty interior. Indeed, for any given t, x, y, ε, choosing $y' \in \mathrm{int}\, \overline{\mathrm{co}}\, F(t,x)$ with $|y' - y| < \varepsilon$, the constant function $\phi \equiv y'$ is a local selection from $\overline{\mathrm{co}}\, F$ satisfying the requirements.

The proof of the above theorem starts with the construction of a sequence f_n of directionally continuous selections from $\overline{\mathrm{co}}\, F$, which are piecewise Lipschitz continuous in the (t,x)-space. For every $u: [0,T] \to I\!R^n$ in a class of Lipschitz continuous functions, we then show that the composed maps $t \mapsto f_n(t, u(t))$ form a Cauchy sequence in $L^1\big([0,T];\ I\!R^n\big)$, converging pointwise almost everywhere to a map of the form $f(\cdot, u(\cdot))$, taking values within the extreme points of F. This convergence is obtained through an argument which is considerably different from previous works. Indeed, it relies on a careful use of the likelihood functional introduced in [10], interpreted here as a measure of "oscillatory non-convergence" of a set of derivatives. We recall that topological properties of the set of solutions of nonconvex differential inclusions have been studied in [9, 12] with the technique of directionally continuous selections and in [32, 33, 53] using the method of Baire category.

Among various corollaries, this result yields an extension, valid for the wider class of multifunctions with the property (LSP), of the following results, proved in [15], [11] and [32], respectively.

(i) Existence of selections from the solution set of a differential inclusion, depending continuously on the initial data.

(ii) Existence of selections from a multifunction, which generate a continuous flow.

(iii) Contractibility of the solution sets of $\dot{x} \in F(t,x)$ and $\dot{x} \in \text{ext}\, F(t,x)$.

The proofs of (i) and (ii) are straightforward, and the proof of (iii) is obtained by explicitly constructing a null homotopy of the solution sets.

Another application concerns bang–bang feedback controls. More precisely, we consider a control system

$$\dot{x} = g(t,x,u), \qquad u \in U, \tag{1}$$

with $U \subset I\!R^m$ compact, and the corresponding "relaxed system"

$$\dot{x} = g^{\#}(t,x,u^{\#}), \qquad u^{\#} \in U^{\#} \doteq U^{n+1} \times E_{n+1}, \tag{2}$$

where E_{n+1} is the standard simplex in $I\!R^{n+1}$ and

$$g^{\#}(t,x,u^{\#}) = g^{\#}\big(t,x,(u_0,\dots,u_n,(\theta_0,\dots,\theta_n))\big) \doteq \sum_{i=0}^{n} \theta_i g(t,x,u_i).$$

We prove that, given a chattering feedback control $u^{\#}(t,x) \in U^{\#}$ for (2) and $\varepsilon_0 > 0$, then there exists a feedback control $u(t,x)$ for (1), such that the solutions of the Cauchy problems with the same initial data associated to (1) and (2) with $u = u(t,x)$ and $u^{\#} = u^{\#}(t,x)$ respectively have a distance less than ε_0 in the norm of uniform convergence. Moreover the control $u(t,x)$ is bang–bang, that is it satisfies

$$g(t,x,u(t,x)) \in \text{ext}\,\{g(t,x,\omega) \,\big|\, \omega \in U\}, \quad \text{for every } t,x.$$

In Chapters 2 and 3 we consider control systems of the form

$$x' = F(x) + uG(x), \tag{3}$$

with $x \in \mathbb{R}^2$, F and G smooth vector fields defined in a open subset Ω of $\mathbb{R}^2$, and where u is a scalar control with values in $[-1,1]$.

We recall here some definitions. A control is a measurable function $u:[a,b] \to [-1,1]$, and a trajectory for u is an absolutely continuous curve $\gamma:[a,b] \to \mathbb{R}^2$ satisfying $\gamma'(t) = F(\gamma(t)) + u(t)G(\gamma(t))$ for almost all $t \in [a,b]$. We denote by $T(\gamma) \doteq b - a$ the time along γ. A bang–bang trajectory is a trajectory corresponding to a control u such that $|u(t)| = 1$ for almost all t.

The minimum time problem for control systems of the form (3) on a two dimensional manifold was studied by Sussmann in [49]–[51]. In particular, [49] contains a detailed analysis of the structure of time–optimal trajectories, and conditions which ensure that every time–optimal trajectory is bang–bang. We shall use the basic ideas of this "geometric" approach in order to study more general minimization problems.

The main technical tool used in our proofs is the concept of *bang–bang variations*. In order to illustrate this idea, consider a trajectory γ of the system (3). If γ is not bang–bang in a neighborhood of a time τ, we construct a new trajectory $\hat{\gamma}$ such that $T(\gamma) = T(\hat{\gamma})$, $\hat{\gamma}$ coincides with γ outside a neighborhood of τ and it is bang–bang in this neighborhood. Repeating this construction for various time τ, we obtain an equal time bang–bang trajectory, arbitrarily close to γ in the uniform norm. Moreover, as we shall see, this new trajectory can be constructed with additional properties which depend on the particular problem we are interested in.

The technique of bang–bang variations replaces the classical approach based on the Lyapunov's Convexity Theorem, used in previous works (see for example [16] and [47] for problems in calculus of variations and control theory respectively).

In Chapter 2 we prove a bang–bang theorem for the problem

$$\min \left\{ \int_0^1 h(\gamma(t), u(t))dt \,\middle|\, \gamma' = F(\gamma) + uG(\gamma),\ \gamma(0) = P, \gamma(1) = Q \right\} ,$$

that is, we give conditions in order to guarantee that every optimal trajectory is bang–bang. In this case, given an optimal trajectory, we construct a bang–bang variation which achieves a lower cost.

As an application, we give conditions, depending only on the vector fields F, G, for the existence of an optimal solution to the non–convex problem

$$\min\left\{\int_0^1 h(u(t))dt \,\middle|\, \gamma' = F(\gamma) + uG(\gamma),\ \gamma(0) = P, \gamma(1) = Q\right\}$$

where u takes only the values ± 1.

In Chapter 3 we consider control problems of the form

$$x'' + a(x, x') = ug(x, x'), \tag{4}$$

with $x \in I\!R$, a and g smooth functions defined on a open subset Ω of $I\!R^2$, and where u is a scalar control taking values in $[-1, 1]$.

Given a solution $x: [a, b] \to I\!R$ of (4), we prove, under mild assumptions on g, the existence of two bang–bang solutions y, z, with a finite number of switchings, satisfying

$$\begin{aligned} x(a) = y(a) = z(a), \quad x(b) = y(b) = z(b), \\ x'(a) = y'(a) = z'(a), \quad x'(b) = y'(b) = z'(b), \end{aligned} \tag{5}$$

and, for every $t \in [a, b]$,

$$y(t) \le x(t) \le z(t). \tag{6}$$

We remark that every forced semilinear second order differential equation, corresponding to (4) with g constant, satisfies the required assumptions. For example, the forced nonlinear pendulum, the forced nonlinear Duffing oscillator, and the forced Van der Pol equation belong to this class of problems.

The problem of finding a bang–bang solution satisfying (5) and (6) in the case of linear control systems $L(x) \in [\phi_1(t), \phi_2(t)]$, where L is a linear operator of order n, was studied in the case of piecewise analytic data by Andreini and Bacciotti in [4]. The techniques used in this paper are based on Lyapunov type theorems (see [23] for examples and applications). Recently Cërf and Mariconda in [19] studied the case $\phi_1, \phi_2 \in L^1$. Their approach is based on a new Lyapunov type theorem applied to the integral representation of the solutions.

As we remarked before, in order to study problem (4), a completely different approach has to be used due to the nonlinearity of the operator. Since no integral representation formula is available, we rely instead on certain geometric properties of the trajectories. Defining $x_1 = x$ and $x_2 = x'$, the control problem (4) is equivalent to a first order control system of the form (3), with

$$F(x_1, x_2) = \begin{pmatrix} x_2 \\ -a(x_1, x_2) \end{pmatrix}, \qquad G(x_1, x_2) = \begin{pmatrix} 0 \\ g(x_1, x_2) \end{pmatrix}.$$

Given a trajectory γ of the planar system, using the method of bang–bang variations, we can construct a bang–bang trajectory $\hat{\gamma}$ such that the corresponding solution of (4) satisfies (5) and (6).

As an application we obtain a closure result for the reachable set of control systems with obstacle. More precisely, if $c: [a, b] \to I\!R$ is a continuous function and (x_0, v_0) is a given initial condition, we can consider the families of solutions of (4) in $[a, b]$:

$$\mathcal{S} \doteq \{x : [a, b] \to I\!R \text{ solution of (4) } : (x(a), x'(a)) = (x_0, v_0)\},$$

$$\mathcal{T} \doteq \{x : [a, b] \to I\!R \text{ bang–bang solution of (4) } : (x(a), x'(a)) = (x_0, v_0)\}.$$

We can now define the constrained reachable sets

$$\mathcal{X} \doteq \{(x(b), x'(b)) : x(t) \leq c(t) \, \forall t \in [a, b], \, x \in \mathcal{S}\},$$

$$\mathcal{Y} \doteq \{(x(b), x'(b)) : x(t) \leq c(t) \, \forall t \in [a, b], \, x \in \mathcal{T}\}.$$

By our previous result, it follows that $\mathcal{X}$ and $\mathcal{Y}$ coincide. In particular, the reachable set $\mathcal{Y}$ associated to bang–bang constrained solutions is closed.

Another application is an existence result for Bolza problems with nonlinear dynamics. More precisely, we consider a functional of the form

$$J(x) \doteq \int_a^b [\alpha(t, x) + \beta(t, x) \cdot x' + \gamma(x') \cdot x''] \, dt,$$

with α, $\beta \in C^1([a,b] \times \mathbb{R}, \mathbb{R})$ and $\gamma \in C(\mathbb{R}, \mathbb{R})$, satisfying

$$\frac{\partial \alpha}{\partial x}(t,x) - \frac{\partial \beta}{\partial t}(t,x) \neq 0,$$

for every $(t,x) \in [a,b] \times \mathbb{R}$. We prove that, if $x: [a,b] \to \mathbb{R}$ is an optimal solution to the problem

$$\min \left\{ J(x) \,\middle|\, x \text{ solution of (4)}, \ x(a) = x_0, \ x'(a) = v_0, \ x(b) = x_1, \ x'(b) = v_1 \right\},$$

then x is bang–bang with a finite number of switchings.

In Chapters 4 and 5 we consider minimization problems of the form

$$(7) \qquad \min \left\{ \int_0^T L(t,u,u')dt \,\middle|\, u \in W^{1,1}(I, \mathbb{R}^m), \, u(0) = a, \, u(T) = b \right\},$$

with $I \doteq [0,T]$, for Lagrangeans possibly non–coercive and non–convex in the u'–argument. It is well known that, if L is a continuous function, such that $\xi \mapsto L(t,x,\xi)$ is convex and superlinear, then the variational problem (7) has a solution (see for instance [23]).

In recent years, the possibility of relaxing the convexity or the superlinearity assumption was investigated by many authors.

Some existence results for non–convex coercive problems were obtained in the case $L(t,x,\xi) = g(t,x) + f(t,\xi)$ (see for instance [16], [41], [47] and the references therein). In particular, in [16] it was proved that the convexity assumption on $f(t,\cdot)$ can be replaced by the condition of concavity of $g(t,\cdot)$.

More recently, some techniques were developed in order to treat convex but non–coercive problems. In this case, even if the functionals considered are lower semicontinuous in the weak topology of $W^{1,1}(I, \mathbb{R}^m)$, the direct method of the Calculus of Variations can not be applied, due to the lack of compactness of the minimizing sequences.

In [26], it was studied the problem (7) with L continuous, bounded from below and convex with respect to ξ, the superlinearity being replaced by a weaker condition which permits to construct a relatively compact minimizing sequence, obtained by considering the minima of suitable coercive approximating problems. The main step in the proof of the existence result in [26] was to show that every minimum point of the approximating problems solves a generalized DuBois–Reymond condition, which implies that the minimizing sequence is bounded in the space $W^{1,\infty}(I,\mathbb{R}^m)$.

A similar approach was used in [18] for the autonomous problem with Lagrangean $L(t,x,\xi) = g(x) + f(\xi)$, where g is a nonnegative continuous function, and $f \in C^1(\mathbb{R}^m,\mathbb{R})$ is a strictly convex function bounded from below, such that

$$\lim_{|\xi|\to+\infty} [f(\xi) - \langle \nabla f(\xi)\,,\,\xi\rangle] = -\infty. \tag{8}$$

In that paper, it was proved that, for every rectifiable curve C in $\mathbb{R}^m$ joining a to b there exists a unique solution to the problem (7) restricted to the class of all absolutely continuous parameterizations $u: I \to \mathbb{R}^m$ of C. Thus, every element u_n of a minimizing sequence can be replaced by the minimum corresponding to the curve parameterized by u_n. It can be shown, still using a DuBois–Reymond condition satisfied by those minima, and by (8), that this new sequence is bounded in $W^{1,\infty}(I,\mathbb{R}^m)$, so that there exists a minimum point for (7) in this space.

In Chapter 4 both the superlinearity and the convexity assumptions are dropped for Lagrangeans of the form $L(t,x,\xi) = \langle a(t)\,,\,x\rangle + f(\xi)$. We assume here that f is a lower semicontinuous function whose convexification f^{**} satisfies (8) restricted to the set where ∇f^{**} is defined. The existence of a minimum is proved by a technique relying only on a Lyapunov type theorem due to Olech (see [43]). More precisely, using a fixed point theorem for upper semicontinuous multifunctions, we prove that

$$\operatorname{Dom} f^* \doteq \{p \in \mathbb{R}^m \mid f^*(p) < +\infty\}$$

is an open subset of $\mathbb{R}^m$. Here f^* is the polar function of f. This fact, together with an existence theorem of Olech (see [43]) gives the desired result.

In Chapter 5 we consider non–autonomous problems of the form

$$(9)\quad \min\left\{\int_0^T [g(t,u)+f(t,u')]\,dt \,\Big|\, u\in W^{1,1}(I,\mathbb{R}^m),\ u(0)=a,\ u(T)=b\right\}$$

with neither coercivity nor convexity assumptions. More precisely, we introduce the class $\mathcal{E}$ of all functions $\psi : I\times\mathbb{R}^m \to \mathbb{R}$, bounded from below, such that $\psi(\cdot,\xi)$ is Lipschitz continuous for every fixed $\xi\in\mathbb{R}^m$, $\psi(t,\cdot)$ is lower semicontinuous and satisfies

$$\lim_{n\to+\infty}[\psi^{**}(t^n,\xi^n)-\langle\nabla\psi^{**}(t^n,\xi^n)\,,\,\xi^n\rangle]=-\infty$$

for every sequence $\{t^n\}\in I$ and for every choice of points ξ^n of differentiability of $\psi^{**}(t^n,\cdot)$ such that $\lim_n|\xi^n|=+\infty$. We show that, if $f\in\mathcal{E}$ and there exist two constants A and B, $B>0$ such that $f(t,\xi)\geq -A+B|\xi|$ for every $(t,\xi)\in I\times\mathbb{R}^m$, and $g(t,x)$ is a continuous function, Lipschitz continuous with respect to t, concave with respect to x, satisfying $g(t,x)\geq -\alpha-\beta|x|$ for every $(t,x)\in I\times\mathbb{R}^m$, and for suitable constants α and $0\leq\beta\leq B/T$, then the problem (9) has a solution in the space $W^{1,\infty}(I,\mathbb{R}^m)$. This result is the analogue for a class of non–coercive functionals of the one in [16], but it is not a generalization of that result, due to the additional requirement of the Lipschitz continuity of the Lagrangean with respect to the variable t. On the other hand, this extra regularity assumption allows us to obtain the necessary conditions that, used at an intermediate step, also yield a regularity result for the optimal solution, interesting by itself.

As a first step we prove an existence result for (9), requiring that f be convex with respect to ξ and dropping the concavity assumption on g. This can be done following [26] and making suitable changes, due to the the fact that the Lagrangean is not bounded from below. The second step, linking the convex to the non–convex case, is based on a result concerning the closure of the convex hull of the epigraph of functions whose convexification is strictly convex at infinity (i.e., the graph of the convexification contains no rays). This result is an extension of the classical theorem valid for superlinear functions (see [35]). We remark that the notion of strict convexity at infinity was also used in [27] in order to study non-coercive problems of the type (7) with the additional state constraint $\|u\|_{L^\infty} < R$. We shall prove that every function in the class $\mathcal{E}$ is strictly convex at infinity for every fixed t. Hence, by using the previous results and the Lyapunov theorem on the range of non-atomic measures, the existence result for the non-convex problems follows. The regularity of the solution of (9) is a consequence of the regularity of the solution to the relaxed problem.

The results stated in Chapter 1 are obtained in collaboration with Prof. A. Bressan and they are published in [13]. The results stated in Chapter 2 and 3 are achieved in collaboration with Dr. B. Piccoli of SISSA and they are published in [30] and [31] respectively, while the results stated in Chapter 5 are obtained in collaboration with Dr. A. Malusa of the University of Napoli and they are published in [29]. Finally, Chapter 4 contains results published in [28].

Notation

We shall denote by $B(x,R)$ or $B_R(x)$ (resp. $\bar{B}(x,R)$ or $\bar{B}_R(x)$) the open (resp. closed) ball of $\mathbb{R}^n$ centered at x and with radius R. $\overline{B}(D;\ R)$ is the closed neighborhood of radius R around the set D.

The Lebesgue measure of a set $A \subset \mathbb{R}^n$ is $m(A)$. We say that a point t is a Lebesgue point for the set $A \subset \mathbb{R}$ if $\lim_{\varepsilon \to 0^+} m(A \cap [t-\varepsilon, t+\varepsilon])/2\varepsilon = 1$. It is well known (see for instance [48]) that if A is Lebesgue–measurable, then almost every point of A is a Lebesgue point.

As customary, $\bar{A}$, $\operatorname{int} A$, $\overline{\operatorname{co}}\, A$ and $\operatorname{ext} A$ denote here the closure, the interior, the closed convex hull and the extreme points of A respectively, while $A \setminus B$ indicates a set–theoretic difference. The characteristic function of a set A is written as χ_A.

We shall denote by $\langle x\,,\,y\rangle$ the standard scalar product of two vectors $x,\ y \in \mathbb{R}^m$. For every $1 \le p \le +\infty$, we shall denote by $L^p(I, \mathbb{R}^m)$ and $W^{1,p}(I, \mathbb{R}^m)$, respectively, the usual Lebesgue and Sobolev spaces of functions from a interval $I \subset \mathbb{R}$ into $\mathbb{R}^m$. We shall use the symbol $\|\cdot\|_{L^p}$ to denote the norm in $L^p(I, \mathbb{R}^m)$.

We shall denote by $2^{\mathbb{R}^m}$ the family of all subset of $\mathbb{R}^m$, and $2^{\mathbb{R}^m}\backslash\emptyset$ will mean $2^{\mathbb{R}^m}\backslash\{\emptyset\}$.

Given a nonempty set Ω, a map $F\colon\Omega \to 2^{\mathbb{R}^m}$ will be called a multifunction. We let $F(\Omega) \doteq \bigcup_{x\in\Omega} F(x)$. A function $f\colon\Omega \to \mathbb{R}^m$ with $f(x) \in F(x)$ for every $x \in \Omega$ will be called a selection of F. For a detailed treatment of this subject, see for example [5] or [34].

Chapter 1
Extremal selections of multifunctions generating a continuous flow

1.1. Introduction

Let $F : [0,T] \times I\!R^n \to 2^{I\!R^n}$ be a continuous multifunction with compact, not necessarily convex values. If F is Lipschitz continuous, it was shown in [11] that there exists a measurable selection f of F such that, for every x_0, the Cauchy problem

$$\dot{x}(t) = f(t, x(t)), \qquad x(0) = x_0$$

has a unique Caratheodory solution, depending continuously on x_0.

In this chapter, we prove that the above selection f can be chosen so that $f(t,x) \in \operatorname{ext} F(t,x)$ for all t, x. More generally, the result remains valid if F satisfies the following Lipschitz Selection Property:

(LSP) *For every t, x, every $y \in \overline{\mathrm{co}}\, F(t,x)$ and $\varepsilon > 0$, there exists a Lipschitz selection ϕ of $\overline{\mathrm{co}}\, F$, defined on a neighborhood of (t,x), with $|\phi(t,x) - y| < \varepsilon$.*

We remark that, by [39, 44], every Lipschitz multifunction with compact values satisfies (LSP). Another interesting class, for which (LSP) holds, consists of those continuous multifunctions F whose values are compact and have convex closure with nonempty interior. Indeed, for any given t, x, y, ε, choosing $y' \in \operatorname{int} \overline{\mathrm{co}}\, F(t,x)$ with $|y' - y| < \varepsilon$, the constant function $\phi \equiv y'$ is a local selection from $\overline{\mathrm{co}}\, F$ satisfying the requirements.

In the following, $\Omega \subseteq I\!R^n$ is an open set, while $\mathcal{AC}$ the Sobolev space of all absolutely continuous functions $u : [0,T] \to I\!R^n$, with norm $\|u\|_{\mathcal{AC}} = \int_0^T \big(|u(t)| + |\dot{u}(t)|\big)\, dt$.

Theorem 1.1.1 *Let $F: [0,T] \times \Omega \to 2^{\mathbb{R}^n} \setminus \emptyset$ be a bounded continuous multifunction with compact values, satisfying (LSP). Assume that $F(t,x) \subseteq \overline{B}(0,M)$ for all t, x and let D be a compact set such that $\overline{B}(D;\ MT) \subset \Omega$. Then there exists a measurable function f, with*

$$f(t,x) \in \operatorname{ext} F(t,x), \qquad \forall t, x, \tag{1.1.1}$$

such that, for every $(t_0, x_0) \in [0,T] \times D$, the Cauchy problem

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0\,, \tag{1.1.2}$$

has a unique Caratheodory solution $x(\cdot) = x(\cdot, t_0, x_0)$ on $[0,T]$, depending continuously on t_0, x_0 in the norm of $\mathcal{AC}$.

Moreover, if $\varepsilon_0 > 0$ and a Lipschitz continuous selection f_0 of $\overline{\text{co}}\, F$ are given, then one can construct f with the following additional property. Denoting by $y(\cdot, t_0, x_0)$ the unique solution of

$$\dot{y}(t) = f_0(t, y(t)), \qquad y(t_0) = x_0, \tag{1.1.3}$$

for every $(t_0, x_0) \in [0,T] \times D$ one has

$$\big|y(t, t_0, x_0) - x(t, t_0, x_0)\big| \le \varepsilon_0 \qquad \forall t \in [0,T]. \tag{1.1.4}$$

The proof of the above theorem, given in section 1.3, starts with the construction of a sequence f_n of directionally continuous selections from $\overline{\text{co}}\, F$, which are piecewise Lipschitz continuous in the (t,x)-space. We then show that, for every $u: [0,T] \to \mathbb{R}^n$ in a class of Lipschitz continuous functions, the composed maps $t \mapsto f_n(t, u(t))$ form a Cauchy sequence in $L^1([0,T];\ \mathbb{R}^n)$, converging pointwise almost everywhere to a map of the form $f(\cdot, u(\cdot))$, taking values within the extreme points of F. This convergence is obtained through an argument which is considerably different from previous works. Indeed, it relies on a careful use of the likelihood functional introduced in [10], interpreted here as a measure of "oscillatory non-convergence" of a set of derivatives.

Among various corollaries, Theorem 1.1.1 yields an extension, valid for the wider class of multifunctions with the property (LSP), of the following results, proved in [15], [11] and [32], respectively.

(i) Existence of selections from the solution set of a differential inclusion, depending continuously on the initial data.

(ii) Existence of selections from a multifunction, which generate a continuous flow.

(iii) Contractibility of the solution sets of $\dot{x} \in F(t, x)$ and $\dot{x} \in \text{ext}\, F(t, x)$.

These consequences, together with an application to bang-bang feedback controls, are described in section 1.4. Topological properties of the set of solutions of nonconvex differential inclusions have been studied in [9, 12] with the technique of directionally continuous selections and in [32, 33, 53] using the method of Baire category.

1.2. Preliminaries

In the following, $\mathcal{K}_n$ denotes the family of all nonempty compact convex subsets of $I\!R^n$, endowed with Hausdorff metric. A key technical tool used in our proofs will be the function $h\colon I\!R^n \times \mathcal{K}_n \to I\!R \cup \{-\infty\}$, defined by

$$h(y, K) \doteq \sup \left\{ \left(\int_0^1 |w(\xi) - y|^2 \, d\xi \right)^{\frac{1}{2}} ; \quad w\colon [0,1] \to K, \quad \int_0^1 w(\xi)\, d\xi = y \right\}$$

with the understanding that $h(y, K) = -\infty$ if $y \notin K$. Observe that $h^2(y, K)$ can be interpreted as the maximum variance among all random variables supported inside K, whose mean value is y. The following results were proved in [10]:

Lemma 1.2.1 *The map $(y, K) \mapsto h(y, K)$ is upper semicontinuous in both variables; for each fixed $K \in \mathcal{K}_n$ the function $y \mapsto h(y, K)$ is strictly concave down on K. Moreover, one has*

$$(1.2.1) \qquad h(y, K) = 0 \qquad \textit{if and only if} \qquad y \in \text{ext}\, K,$$

$$h^2(y, K) \le r^2(K) - \left|y - c(K)\right|^2, \tag{1.2.2}$$

where $c(K)$ and $r(K)$ denote the Chebyschev center and the Chebyschev radius of K, respectively.

Remark 1.2.2 By the above lemma, the function h has all the qualitative properties of the Choquet function d_F considered, for example, in [33, Proposition 2.6]. It could thus be used within any argument based on Baire category. Moreover, the likelihood functional

$$L(u) \doteq \left(\int_0^T h^2\big(\dot u(t), F(t, u(t))\big)\, dt \right)^{1/2}$$

provides an upper bound to the distance

$$\|\dot v - \dot u\|_{L^2}$$

between derivatives, for solutions of $\dot v \in F(t, v)$ which remain close to u uniformly on $[0, T]$. This additional quantitative property of the function h will be a crucial ingredient in our proof.

For the basic theory of multifunctions and differential inclusions we refer to [5]. As in [9], given a map $g: [0, T] \times \Omega \to I\!R^n$, we say that g is directionally continuous along the directions of the cone $\Gamma^N = \{(s, y);\ |y| \le Ns\}$ if

$$g(t, x) = \lim_{k \to \infty} g(t_k, x_k)$$

for every (t, x) and every sequence (t_k, x_k) in the domain of g such that $t_k \to t$ and $|x_k - x| \le N(t_k - t)$ for every k. Equivalently, g is Γ^N-continuous iff it is continuous with respect to the topology generated by the family of all half-open cones of the form

$$\{(s, y)\ ;\quad \hat t \le s < \hat t + \varepsilon,\ |y - \hat x| \le N(s - t)\} \tag{1.2.3}$$

with $(\hat t, \hat x) \in I\!R \times I\!R^n$, $\varepsilon > 0$. A set of the form (1.2.3) will be called an N-cone.

Under the assumptions on Ω, D made in Theorem 1.1.1, consider the set of Lipschitzean functions

$$(1.2.4) \qquad Y \doteq \{u: [0,T] \to \overline{B}(D, MT) \ : |u(t) - u(s)| \le M|t-s| \quad \forall t, s\}.$$

The Picard operator of a map $g: [0,T] \times \Omega \to I\!R^n$ is defined as

$$\mathcal{P}^g(u)(t) \doteq \int_0^t g(s, u(s))\, ds, \qquad u \in Y.$$

The distance between two Picard operators will be measured by

$$\|\mathcal{P}^f - \mathcal{P}^g\| \doteq \sup\left\{\left|\int_0^t [f(s,u(s)) - g(s,u(s))]\, ds\right| ; \quad t \in [0,T], \ \ u \in Y\right\}.$$

The next Lemma will be useful in order to prove the uniqueness of solutions of the Cauchy problems (1.1.2).

Lemma 1.2.3 *Let f be a measurable map from $[0,T] \times \Omega$ into $\overline{B}(0, M)$, with $\mathcal{P}^f$ continuous on Y. Let D be compact, with $\overline{B}(D,\ MT) \subset \Omega$, and assume that the Cauchy problem*

$$(1.2.5) \qquad \dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0, \qquad t \in [0,T],$$

has a unique solution, for each $(t_0, x_0) \in [0,T] \times D$.

Then, for every $\varepsilon > 0$, there exists $\delta > 0$ with the following property. If $g : [0,T] \times \Omega \to \overline{B}(0, M)$ satisfies $\|\mathcal{P}^g - \mathcal{P}^f\| \le \delta$, then for every $(t_0, x_0) \in [0,T] \times D$, any solution of the Cauchy problem

$$(1.2.6) \qquad \dot{y}(t) = g(t, y(t)), \qquad y(t_0) = x_0, \qquad t \in [0,T],$$

has distance $< \varepsilon$ from the corresponding solution of (1.2.5). In particular, the solution set of (1.2.6) has diameter $\le 2\varepsilon$ in $C^0([0,T]; I\!R^n)$.

Proof. If the conclusion fails, then there exist sequences of times t_ν, t'_ν, maps g_ν with $\|\mathcal{P}^{g_\nu} - \mathcal{P}^f\| \to 0$, and couples of solutions $x_\nu, y_\nu: [0,T] \to \overline{B}(D; MT)$ of

$$(1.2.7) \quad \dot{x}_\nu(t) = f(t, x_\nu(t)), \qquad \dot{y}_\nu(t) = g_\nu(t, y_\nu(t)), \qquad t \in [0,T],$$

with

$$x_\nu(t_\nu) = y_\nu(t_\nu) \in D, \qquad \left|x_\nu(t'_\nu) - y_\nu(t'_\nu)\right| \geq \varepsilon \qquad \forall \nu. \tag{1.2.8}$$

By taking subsequences, we can assume that $t_\nu \to t_0$, $t'_\nu \to \tau$, $x_\nu(t_0) \to x_0$, while $x_\nu \to x$ and $y_\nu \to y$ uniformly on $[0,T]$. From (1.2.7) it follows

$$\begin{aligned} \left|y(t) - x_0 - \int_{t_0}^t f(s,y(s))\, ds\right| &\leq \left|y(t) - y_\nu(t)\right| + \left|x_0 - y_\nu(t_0)\right| + \\ &+ \left|\int_{t_0}^t \left[f(s,y(s)) - f(s,y_\nu(s))\right]\, ds\right| + \\ &+ \left|\int_{t_0}^t \left[f(s,y_\nu(s)) - g_\nu(s,y_\nu(s))\right]\, ds\right|. \end{aligned} \tag{1.2.9}$$

As $\nu \to \infty$, the right hand side of (1.2.9) tends to zero, showing that $y(\cdot)$ is a solution of (1.2.5). By the continuity of $\mathcal{P}^f$, $x(\cdot)$ is also a solution of (1.2.5), distinct from $y(\cdot)$ because

$$|x(\tau) - y(\tau)| = \lim_{\nu\to\infty} |x_\nu(\tau) - y_\nu(\tau)| = \lim_{\nu\to\infty} \left|x_\nu(t'_\nu) - y_\nu(t'_\nu)\right| \geq \varepsilon.$$

This contradicts the uniqueness assumption, proving the lemma. □

1.3. Proof of the main theorem

Observing that ext $F(t,x) = \text{ext}\,\overline{\text{co}}\,F(t,x)$ for every compact set $F(t,x)$, it is clearly not restrictive to prove Theorem 1.1.1 under the additional assumption that all values of F are convex. Moreover, the bounds on F and D imply that no solution of the Cauchy problem

$$\dot{x}(t) \in F(t,x(t)), \qquad x(t_0) = x_0, \qquad t \in [0,T],$$

with $x_0 \in D$, can escape from the set $\overline{B}(D,\ MT)$. Therefore, it suffices to construct the selection f on the compact set $\Omega^\dagger \doteq [0,T] \times \overline{B}(D,\ MT)$. Finally, since every convex valued multifunction satisfying (LSP) admits a globally defined Lipschitz selection, it suffices to prove the second part of the theorem, with f_0 and $\varepsilon_0 > 0$ assigned.

We shall define a sequence of directionally continuous selections of F, converging a.e. to a selection from $\operatorname{ext} F$. The basic step of our constructive procedure will be provided by the next lemma.

Lemma 1.3.1 *Fix any $\varepsilon > 0$. Let S be a compact subset of $[0,T] \times \Omega$ and let $\phi : S \to \mathbb{R}^n$ be a continuous selection of F such that*

$$h\big(\phi(t,x), F(t,x)\big) < \eta \qquad \forall (t,x) \in S. \tag{1.3.1}$$

Then there exists a piecewise Lipschitz selection $g : S \to \mathbb{R}^n$ of F with the following properties:

(i) There exists a finite covering $\{\Gamma_i\}_{i=1\dots,\nu}$, consisting of Γ^{M+1}–cones, such that, if we define the pairwise disjoint sets $\Delta^i \doteq \Gamma_i \setminus \bigcup_{\ell<i} \Gamma_\ell$, then on each Δ^i the following holds:

(a) there exist Lipschitzean selections $\psi_j^i : \overline{\Delta^i} \to \mathbb{R}^n$, $j = 0, \dots, n$, such that

$$g\big|_{\Delta^i} = \sum_{j=0}^{n} \psi_j^i \, \chi_{A_j^i}, \tag{1.3.2}$$

where each A_j^i is a finite union of strips of the form $\big([t',t'')\times \mathbb{R}^n\big) \cap \Delta^i$.

(b) For every $j = 0, \dots, n$ there exists an affine map $\varphi_j^i(\cdot) \doteq \langle a_j^i, \cdot\rangle + b_j^i$ such that, for every $(t,x) \in \overline{\Delta^i}$ and $z \in F(t,x)$,

$$\varphi_j^i\big(\psi_j^i(t,x)\big) \le \varepsilon, \quad \varphi_j^i(z) \ge h(z, F(t,x)). \tag{1.3.3}$$

(ii) For every $u \in Y$ and every interval $[\tau, \tau']$ such that $(s, u(s)) \in S$ for $\tau \le s < \tau'$, the following estimates hold:

$$\left| \int_\tau^{\tau'} \big[\phi(s,u(s)) - g(s,u(s))\big] \, ds \right| \le \varepsilon, \tag{1.3.4}$$

$$\int_{\tau}^{\tau'} \left|\phi(s,u(s)) - g(s,u(s))\right| \, ds \ \leq\ \varepsilon + \eta(\tau' - \tau). \tag{1.3.5}$$

Remark 1.3.2 Thinking of $h(y,K)$ as a measure for the distance of y from the extreme points of K, the above lemma can be interpreted as follows. Given any selection ϕ of F, one can find a Γ^{M+1}-continuous selection g whose values lie close to the extreme points of F and whose Picard operator $\mathcal{P}^g$, by (1.3.4), is close to $\mathcal{P}^\phi$. Moreover, if the values of ϕ are near the extreme points of F, i.e. if η in (1.3.1) is small, then g can be chosen close to ϕ. The estimate (1.3.5) will be a direct consequence of the definition of h and of Hölder's inequality.

Remark 1.3.3 Since h is only upper semicontinuous, the two assumptions $y_\nu \to y$ and $h(y_\nu, K) \to 0$ do not necessarily imply $h(y,K) = 0$. As a consequence, the a.e. limit of a convergent sequence of approximately extremal selections f_ν of F need not take values inside ext F. To overcome this difficulty, the estimates in (1.3.3) provide upper bounds for h in terms of the affine maps φ_j^i. Since each φ_j^i is continuous, limits of the form $\varphi_j^i(y_\nu) \to \varphi_j^i(y)$ will be straightforward.

Proof of Lemma 1.3.1. For every $(t,x) \in S$ there exist values $y_j(t,x) \in F(t,x)$ and coefficients $\theta_j(t,x) \geq 0$, with

$$\phi(t,x) = \sum_{j=0}^{n} \theta_j(t,x) y_j(t,x), \qquad \sum_{j=0}^{n} \theta_j(t,x) = 1,$$

$$h\big(y_j(t,x), F(t,x)\big) < \varepsilon/2.$$

By the concavity and the upper semicontinuity of h, for every $j = 0,\ldots,n$ there exists an affine function $\varphi_j^{(t,x)}(\cdot) \doteq \langle a_j^{(t,x)}, \cdot\rangle + b_j^{(t,x)}$ such that

$$\varphi_j^{(t,x)}(y_j(t,x)) < h\big(y_j(t,x), F(t,x)\big) + \frac{\varepsilon}{2} < \varepsilon,$$

$$\varphi_j^{(t,x)}(z) > h\big(z, F(t,x)\big), \qquad \forall z \in F(t,x).$$

By (LSP) and the continuity of each $\varphi_j^{(t,x)}$, there exists a neighborhood $\mathcal{U}$ of (t,x) together with Lipschitzean selections $\psi_j^{(t,x)}: \mathcal{U} \to \mathbb{R}^n$, such that, for every j and every $(s,y) \in \mathcal{U}$,

$$\left|\psi_j^{(t,x)}(s,y) - y_j(t,x)\right| < \frac{\varepsilon}{4T}, \tag{1.3.6}$$

$$\varphi_j^{(t,x)}\big(\psi_j^{(t,x)}(s,y)\big) < \varepsilon. \tag{1.3.7}$$

Using again the upper semicontinuity of h, we can find a neighborhood $\mathcal{U}'$ of (t,x) such that

$$\varphi_j^{(t,x)}(z) \geq h\big(z, F(s,y)\big) \qquad \forall z \in F(s,y),\ (s,y) \in \mathcal{U}',\ j = 0, \ldots, n. \tag{1.3.8}$$

Choose a neighborhood $\Gamma_{t,x}$ of (t,x), contained in $\mathcal{U} \cap \mathcal{U}'$, such that, for every point (s,y) in the closure $\overline{\Gamma}_{t,x}$, one has

$$|\phi(s,y) - \phi(t,x)| < \frac{\varepsilon}{4T}\,. \tag{1.3.9}$$

It is not restrictive to assume that $\Gamma_{t,x}$ is a $(M+1)$-cone, i.e. it has the form (1.2.3) with $N = M+1$. By the compactness of S we can extract a finite subcovering $\{\Gamma^i;\ \ 1 \leq i \leq \nu\}$, with $\Gamma_i \doteq \Gamma_{t_i,x_i}$. Define $\Delta^i \doteq \Gamma_i \setminus \bigcup_{j<i} \Gamma_j$ and set $\theta_j^i \doteq \theta_j(t_i,x_i)$, $y_j^i \doteq y_j(t_i,x_i)$, $\psi_j^i \doteq \psi_j^{(t_i,x_i)}$, $\varphi_j^i = \varphi_j{}^{(t_i,x_i)}$. Choose an integer N such that

$$N > \frac{8M\nu^2 T}{\varepsilon} \tag{1.3.10}$$

and divide $[0,T]$ into N equal subintervals $J_1, \ldots, J_N$, with

$$J_k = \big[t_{k-1},\ t_k\big), \qquad\qquad t_k = \frac{kT}{N}. \tag{1.3.11}$$

For each i,k such that $\big(J_k \times \mathbb{R}^n\big) \cap \Delta^i \neq \emptyset$, we then split J_k into $n+1$ subintervals $J_{k,0}^i, \ldots, J_{k,n}^i$ with lengths proportional to $\theta_0^i, \ldots, \theta_n^i$, by setting

$$J_{k,j}^i = \big[t_{k,j-1},\ t_{k,j}\big), \qquad t_{k,j} = \frac{T}{N}\cdot\Big(k + \sum_{\ell=0}^{j} \theta_\ell^i\Big), \qquad t_{k,-1} = \frac{Tk}{N}.$$

For any point $(t,x)\in\overline{\Delta^i}$ we now set

$$\begin{cases} g^i(t,x)\doteq\psi_j^i(t,x) \\ \bar g^i(t,x)=y_j^i \end{cases} \qquad \text{if}\ \ t\in\bigcup_{k=1}^{N}J_{k,j}^i. \tag{1.3.12}$$

The piecewise Lipschitz selection g and a piecewise constant approximation $\bar g$ of g can now be defined as

$$g=\sum_{i=1}^{\nu}g^i\chi_{\Delta^i},\qquad \bar g=\sum_{i=1}^{\nu}\bar g^i\chi_{\Delta^i}. \tag{1.3.13}$$

By construction, recalling (1.3.7) and (1.3.8), the conditions (a), (b) in (i) clearly hold.

It remains to show that the estimates in (ii) hold as well. Let $\tau,\ \tau'\in[0,T]$ and $u\in Y$ be such that $(t,u(t))\in S$ for every $t\in[\tau,\tau']$, and define

$$E^i=\{t\in I\,;\quad (t,u(t))\in\Delta^i\},\qquad i=1,\dots,\nu.$$

From our previous definition $\Delta^i\doteq\Gamma_i\setminus\bigcup_{j<i}\Gamma_j$, where each Γ_j is a $(M+1)$-cone, it follows that every E^i is the union of at most i disjoint intervals. We can thus write

$$E^i=\Big(\bigcup_{J_k\subset E^i}J_k\Big)\cup\hat E^i,$$

with J_k given by (1.3.11) and

$$m(\hat E^i)\le\frac{2iT}{N}\le\frac{2\nu T}{N}. \tag{1.3.14}$$

Since

$$\phi(t_i,x_i)=\sum_{j=0}^{n}\theta_j^iy_j^i, \tag{1.3.15}$$

the definition of $\bar g$ at (1.3.12), (1.3.13) implies

$$\int_{J_k}\big[\phi(t_i,x_i)-\bar g(s,u(s))\big]\,ds=m(J_k)\cdot\left[\phi(t_i,x_i)-\sum_{j=0}^{n}\theta_j^iy_j^i\right]=0.$$

Therefore, from (1.3.9) and (1.3.6) it follows

$$\begin{aligned}\left|\int_{J_k}\left[\phi(s,u(s))-g(s,u(s))\right]\,ds\right| &\le \left|\int_{J_k}\left[\phi(s,u(s))-\phi(t_i,x_i)\right]\,ds\right| + \\ &+\left|\int_{J_k}\left[\phi(t_i,x_i)-\bar g(s,u(s))\right]\,ds\right| + \left|\int_{J_k}\left[\bar g(s,u(s))-g(s,u(s))\right]\,ds\right| \le \\ &\le m(J_k)\cdot\left[\frac{\varepsilon}{4T}+0+\frac{\varepsilon}{4T}\right] = m(J_k)\cdot\frac{\varepsilon}{2T}.\end{aligned}$$

The choice of N at (1.3.10) and the bound (1.3.14) thus imply

$$\begin{aligned}\left|\int_\tau^{\tau'}\left[\phi(s,u(s))-g(s,u(s))\right]\,ds\right| &\le 2M\cdot m\Big(\bigcup_{i=1}^{\nu}\hat E^i\Big)+(\tau'-\tau)\frac{\varepsilon}{2T}\le \\ &\le 2M\nu\cdot\frac{2\nu T}{N}+\frac{\varepsilon}{2}\le\varepsilon,\end{aligned}$$

proving (1.3.4).

We next consider (1.3.5). For a fixed $i\in\{1,\dots,\nu\}$, let E^i be as before and define

$$\xi_{-1}=0,\qquad \xi_j=\sum_{\ell=0}^{j}\theta^i_\ell,\qquad\qquad w^i(\xi)=\sum_{j=0}^{n}y^i_j\chi_{[\xi_{j-1},\,\xi_j]}.$$

Recalling (1.3.15), the definition of h and Hölder's inequality together imply

$$\begin{aligned}h\big(\phi(t_i,x_i),\ F(t_i,x_i)\big) &\ge \left(\int_0^1\left|\phi(t_i,x_i)-w^i(\xi)\right|^2\,d\xi\right)^{\frac12}\ge \\ &\ge\int_0^1\left|\phi(t_i,x_i)-w^i(\xi)\right|\,d\xi=\sum_{j=0}^{n}\theta^i_j\left|\phi(t_i,x_i)-y^i_j\right|.\end{aligned}$$

Using this inequality we obtain

$$\begin{aligned}\int_{J_k}\left|\phi(t_i,x_i)-\bar g(s,u(s))\right|\,ds &= m(J_k)\cdot\sum_{j=0}^{n}\theta^i_j\left|\phi(t_i,x_i)-y^i_j\right|\le \\ &\le m(J_k)\cdot h\big(\phi(t_i,x_i),F(t_i,x_i)\big)\le\eta\cdot m(J_k),\end{aligned}$$

and therefore, by (1.3.9) and (1.3.6),

$$\begin{aligned}
\int_{J_k} \big|\phi(s,u(s)) - g(s,u(s))\big|\, ds \le \\
&\le \int_{J_k} \big|\phi(s,u(s)) - \phi(t_i,x_i)\big|\, ds + \int_{J_k} \big|\bar g(s,u(s)) - g(s,u(s))\big|\, ds \\
&\quad + \int_{J_k} \big|\phi(t_i,x_i) - \bar g(s,u(s))\big| \le \\
&\le m(J_k)\cdot\Big[\frac{\varepsilon}{4T} + \frac{\varepsilon}{4T} + \eta\Big] = m(J_k)\cdot\Big(\frac{\varepsilon}{2T} + \eta\Big).
\end{aligned}$$

Using again (1.3.14) and (1.3.10), we conclude

$$\int_\tau^{\tau'} \big|\phi(s,u(s)) - g(s,u(s))\big|\, ds \le (\tau'-\tau)\Big(\frac{\varepsilon}{2T} + \eta\Big) + 2M\nu\cdot\frac{2\nu T}{N} \le \varepsilon + (\tau'-\tau)\eta.$$

□

Using Lemma 1.3.1, given any continuous selection $\tilde f$ of F on $\Omega^\dagger$, and any sequence $(\varepsilon_k)_{k\ge1}$ of strictly positive numbers, we can generate a sequence $(f_k)_{k\ge1}$ of selections from F as follows.

To construct f_1, we apply the lemma with $S = \Omega^\dagger$, $\phi = f_0$, $\varepsilon = \varepsilon_1$. This yields a partition $\{A_1^i;\ \ i = 1,\dots,\nu_1\}$ of $\Omega^\dagger$ and a piecewise Lipschitz selection f_1 of F of the form

$$f_1 = \sum_{i=1}^{\nu_1} f_1^i \chi_{A_1^i}.$$

In general, at the beginning of the k-th step we are given a partition of $\Omega^\dagger$, say $\{A_k^i;\ \ i = 1,\dots,\nu_k\}$, and a selection

$$f_k = \sum_{i=1}^{\nu_k} f_k^i \chi_{A_k^i},$$

where each f_k^i is Lipschitz continuous and satisfies

$$h\big(f_k(t,x), F(t,x)\big) \le \varepsilon_k \qquad \forall (t,x) \in \overline{A_k^i}.$$

We then apply Lemma 1.3.1 separately to each A_k^i, choosing $S = \overline{A_k^i}$, $\varepsilon = \varepsilon_k$, $\phi = f_k^i$. This yields a partition $\{A_{k+1}^i;\ \ i = 1, \ldots, \nu_{k+1}\}$ of $\Omega^\dagger$ and functions of the form

$$f_{k+1} = \sum_{i=1}^{\nu_{k+1}} f_{k+1}^i \chi_{A_{k+1}^i}, \qquad\qquad \varphi_{k+1}^i(\cdot) = \langle a_{k+1}^i, \cdot \rangle + b_{k+1}^i,$$

where each $f_{k+1}^i : \overline{A_{k+1}^i} \to I\!R^n$ is a Lipschitz continuous selection from F, satisfying the following estimates:

$$\varphi_{k+1}^i(z) > h\big(z, F(t,x)\big) \qquad \forall (t,x) \in A_{k+1}^i, \tag{1.3.16}$$

$$\varphi_{k+1}^i\big(f_{k+1}^i(t,x)\big) \leq \varepsilon_{k+1} \qquad \forall (t,x) \in A_{k+1}^i, \tag{1.3.17}$$

$$\left| \int_\tau^{\tau'} \big[f_{k+1}(s, u(s)) - f_k(s, u(s))\big]\ ds \right| \leq \varepsilon_{k+1}, \tag{1.3.18}$$

$$\int_\tau^{\tau'} \big|f_{k+1}(s, u(s)) - f_k(s, u(s))\big|\ ds \leq \varepsilon_{k+1} + \varepsilon_k(\tau' - \tau), \tag{1.3.19}$$

for every $u \in Y$ and every τ, τ', as long as the values $(s, u(s))$ remain inside a single set A_k^i, for $s \in [\tau, \tau')$.

Observe that, according to Lemma 1.3.1, each A_k^i is closed-open in the finer topology generated by all $(M+1)$–cones. Therefore, each f_k is Γ^{M+1}–continuous. By Theorem 2 in [8], the substitution operator $\mathcal{S}^{f_k} : u(\cdot) \mapsto f_k(\cdot, u(\cdot))$ is continuous from the set Y defined at (1.2.4) into $L^1\big([0,T];\ I\!R^n\big)$. The Picard map $\mathcal{P}^{f_k}$ is thus continuous as well.

Furthermore, there exists an integer N_k with the following property. Given any $u \in Y$, there exists a finite partition of $[0,T]$ with nodes $0 = \tau_0 < \tau_1 < \cdots < \tau_{n(u)} = T$, with $n(u) \leq N_k$, such that, as t ranges in any $[\tau_{\ell-1},\ \tau_\ell)$, the point $(t, u(t))$ remains inside one single set A_k^i. Otherwise stated, the number of times in which the curve $t \mapsto (t, u(t))$ crosses a boundary between two distinct sets A_k^i, A_k^j is smaller that N_k, for every $u \in Y$.

The construction of the A_k^i in terms of $(M+1)$–cones implies that all these crossings are transversal. Since the restriction of f_k to each A_k^i is Lipschitz continuous, it is clear that every Cauchy problem

$$\dot{x}(t) = f_k(t, x(t)), \qquad\qquad x(t_0) = x_0\,,$$

has a unique solution, depending continuously on the initial data $(t_0, x_0) \in [0,T] \times D$.

From (1.3.18), (1.3.19) and the property of N_k it follows

$$\begin{aligned} &\left| \int_0^t \left[f_{k+1}(s,u(s)) - f_k(s,u(s)) \right] \, ds \right| \leq \\ (1.3.20) \qquad &\leq \sum_{\ell=1}^{L} \left| \int_{\tau_{\ell-1}}^{\tau_\ell} \left[f_{k+1}(s,u(s)) - f_k(s,u(s)) \right] \, ds \right| \leq N_k \varepsilon_{k+1}, \end{aligned}$$

where $0 = \tau_0 < \tau_1 < \cdots < \tau_L = t$ are the times at which the map $s \to (s, u(s))$ crosses a boundary between two distinct sets A_k^i, A_k^j. Since (1.3.20) holds for every $t \in [0,T]$, we conclude

$$(1.3.21) \qquad\qquad \left\| \mathcal{P}^{f_{k+1}} - \mathcal{P}^{f_k} \right\| \leq N_k \varepsilon_{k+1}.$$

Similarly, for every $u \in Y$ one has

$$\begin{aligned} &\left\| f_{k+1}(\cdot, u(\cdot)) - f_k(\cdot, u(\cdot)) \right\|_{L^1([0,T];\ \mathbb{R}^n)} \leq \\ (1.3.22) \qquad &\leq \sum_{\ell=1}^{n(u)} \int_{\tau_{\ell-1}}^{\tau_\ell} \left| f_{k+1}(s,u(s)) - f_k(s,u(s)) \right| \, ds \leq \\ &\leq \sum_{\ell=1}^{n(u)} \left[\varepsilon_{k+1} + \varepsilon_k (\tau_\ell - \tau_{\ell-1}) \right] \ \leq\ N_k \varepsilon_{k+1} + \varepsilon_k T. \end{aligned}$$

Now consider the functions $\varphi_k : \mathbb{R}^n \times \Omega^\dagger \to \mathbb{R}$, with

$$(1.3.23) \qquad \varphi_k(y,t,x) \doteq \langle a_k^i, y \rangle + b_k^i\,, \qquad\qquad \text{if} \quad (t,x) \in A_k^i.$$

From (1.3.16), (1.3.17) it follows

$$(1.3.24) \qquad \varphi_k(y,t,x) \geq h(y, F(t,x))\,, \qquad\qquad \forall (t,x) \in \Omega^\dagger, \;\; y \in F(t,x),$$

$$\varphi_k\big(f_k(t,x),t,x\big) \le \varepsilon_k\,, \qquad \forall (t,x)\in\Omega^\dagger. \tag{1.3.25}$$

For every $u \in Y$, (1.3.18) and the linearity of φ_k w.r.t. y imply

$$\begin{aligned}
&\left|\int_0^T \big[\varphi_k\big(f_{k+1}(s,u(s)),s,u(s)\big) - \varphi_k\big(f_k(s,u(s)),s,u(s)\big)\big]\,ds\right| \le \\
&\le \sum_{\ell=1}^{n(u)} M_k\cdot\left|\int_{\tau_{\ell-1}}^{\tau_\ell} \big[f_{k+1}(s,u(s)) - f_k(s,u(s))\big]\,ds\right| \le \\
&\le N_k\cdot M_k\cdot\varepsilon_{k+1},
\end{aligned}\tag{1.3.26}$$

where $M_k \doteq \max\big\{|a_k^1|,\dots,|a_k^{\nu_k}|\big\}$. Moreover, for every $\ell \ge k$, from (1.3.19) it follows

$$\begin{aligned}
&\int_0^T \Big|\varphi_k\big(f_{\ell+1}(s,u(s)),s,u(s)\big) - \varphi_k\big(f_\ell(s,u(s)),s,u(s)\big)\Big|\,ds \le \\
&\le M_k\cdot\int_0^T \big|f_{\ell+1}(s,u(s)) - f_\ell(s,u(s))\big|\,ds \le \\
&\le M_k\cdot\big(N_\ell\varepsilon_{\ell+1} + \varepsilon_\ell T\big).
\end{aligned}\tag{1.3.27}$$

Observe that all of the above estimates hold regardless of the choice of the ε_k. We now introduce an inductive procedure for choosing the constants ε_k, which will yield the convergence of the sequence f_k to a function f with the desired properties.

Given f_0 and ε_0, by Lemma 1.2.3 there exists $\delta_0 > 0$ such that, if $g\colon\Omega^\dagger \to \overline{B}(0,M)$ and $\|\mathcal{P}^g - \mathcal{P}^{f_0}\| \le \delta_0$, then, for each $(t_0,x_0)\in[0,T]\times D$, every solution of (1.2.6) remains ε_0-close to the unique solution of (1.1.3). We then choose $\varepsilon_1 = \delta_0/2$.

By induction on k, assume that the functions $f_1,\dots,f_k$ have been constructed, together with the linear functions $\varphi_\ell^i(\cdot) \doteq \langle a_\ell^i,\cdot\rangle + b_\ell^i$ and the integers N_ℓ, $\ell = 1,\dots,k$. Let the values $\delta_0,\delta_1,\dots,\delta_k > 0$ be inductively chosen, satisfying

$$\delta_\ell \le \frac{\delta_{\ell-1}}{2}\,, \qquad \ell = 1,\dots,k, \tag{1.3.28}$$

and such that $\|\mathcal{P}^g - \mathcal{P}^{f_\ell}\| \leq \delta_\ell$ implies that for every $(t_0, x_0) \in [0,T] \times D$ the solution set of (1.2.6) has diameter $\leq 2^{-\ell}$, for $\ell = 1, \ldots, k$. This is possible again because of Lemma 1.2.3. For $k \geq 1$ we then choose

$$(1.3.29)\quad \varepsilon_{k+1} \doteq \min\left\{ \frac{\delta_k}{2N_k}, \frac{2^{-k}}{N_k}, \frac{2^{-k}}{N_k \cdot \max\left\{|a_\ell^i|;\ 1 \leq \ell \leq k,\ 1 \leq i \leq \nu_\ell\right\}} \right\}.$$

Using (1.3.28), (1.3.29) in (1.3.21), with $N_0 \doteq 1$, we now obtain

$$(1.3.30)\qquad \sum_{k=p}^{\infty} \|\mathcal{P}^{f_{k+1}} - \mathcal{P}^{f_k}\| \leq \sum_{k=p}^{\infty} N_k \cdot \frac{\delta_k}{2N_k} \leq \sum_{k=p}^{\infty} \frac{2^{p-k}\delta_p}{2} \leq \delta_p$$

for every $p \geq 0$. From (1.3.22) and (1.3.29) we further obtain

$$(1.3.31)\qquad \begin{aligned} \sum_{k=1}^{\infty} \left\|f_{k+1}(\cdot, u(\cdot)) - f_k(\cdot, u(\cdot))\right\|_{L^1} &\leq \sum_{k=1}^{\infty} \left(N_k \cdot \frac{2^{-k}}{N_k} + \frac{2^{1-k}T}{N_k}\right) \leq \\ &\leq \sum_{k=1}^{\infty} \left(2^{-k} + 2^{1-k}T\right) \leq 1 + 2T. \end{aligned}$$

Define

$$(1.3.32)\qquad f(t,x) \doteq \lim_{k\to\infty} f_k(t,x),$$

for all $(t,x) \in \Omega^\dagger$ at which the sequence f_k converges. By (1.3.31), for every $u \in Y$ the sequence $f_k(\cdot, u(\cdot))$ converges in $L^1([0,T]; I\!R^n)$ and a.e. on $[0,T]$. In particular, considering the constant functions $u \equiv x \in \overline{B}(D, MT)$, by Fubini's theorem we conclude that f is defined a.e. on $\Omega^\dagger$. Moreover, the substitution operators $\mathcal{S}^{f_k}: u(\cdot) \mapsto f_k(\cdot, u(\cdot))$ converge to the operator $\mathcal{S}^f: u(\cdot) \mapsto f(\cdot, u(\cdot))$ uniformly on Y. Since each $\mathcal{S}^{f_k}$ is continuous, $\mathcal{S}^f$ is also continuous. Clearly, the Picard map $\mathcal{P}^f$ is continuous as well. By (1.3.30) we have

$$\|\mathcal{P}^f - \mathcal{P}^{f_k}\| \leq \sum_{k=p}^{\infty} \|\mathcal{P}^{f_{k+1}} - \mathcal{P}^{f_k}\| \leq \delta_p, \qquad \forall p \geq 1.$$

Recalling the property of δ_p, this implies that, for every p, the solution set of (1.2.6) has diameter $\leq 2^{-p}$. Since p is arbitrary, for every $(t_0, x_0) \in [0,T] \times D$

the Cauchy problem can have at most one solution. On the other hand, the existence of such a solution is guaranteed by Schauder's theorem. The continuous dependence of this solution on the initial data t_0, x_0, in the norm of $\mathcal{AC}$, is now an immediate consequence of uniqueness and of the continuity of the operators $\mathcal{S}^f$, $\mathcal{P}^f$. Furthermore, for $p = 0$, (1.3.30) yields $\|\mathcal{P}^f - \mathcal{P}^{f_0}\| \leq \delta_0$. The choice of δ_0 thus implies (1.1.4).

It now remains to prove (1.1.1). Since every set $F(t,x)$ is closed, it is clear that $f(t,x) \in F(t,x)$. For every $u \in Y$ and $k \geq 1$, by (1.3.24)–(1.3.27) the choices of ε_k at (1.3.29) yield

$$\begin{aligned}
&\int_0^T h\big(f(s,u(s)),\ F(s,u(s))\big)\ ds \leq \int_0^T \varphi_k\big(f(s,u(s)),s,u(s)\big)\ ds \leq \\
&\quad\leq \int_0^T \varphi_k\big(f_k(s,u(s)),s,u(s)\big)\ ds + \\
&\quad + \left| \int_0^T \big[\varphi_k\big(f_{k+1}(s,u(s)),s,u(s)\big) - \varphi_k\big(f_k(s,u(s)),s,u(s)\big)\big]\ ds \right| + \\
&\quad + \sum_{\ell=k+1}^{\infty} \int_0^T \big|\varphi_k\big(f_{\ell+1}((s,u(s)),s,u(s)\big) - \varphi_k\big(f_\ell(s,u(s)),s,u(s)\big)\big|\ ds \leq \\
&\quad \leq 2^{1-k}T + 2^{-k} + \sum_{\ell=k+1}^{\infty} \big(2^{-\ell} + 2^{1-\ell}T\big).
\end{aligned}$$

Observing that the right hand side of the inequality above approaches zero as $k \to \infty$, we conclude that

$$\int_0^T h\big(f(t,u(t)),\ F(t,u(t))\big)\ dt = 0.$$

By (1.2.1), given any $u \in Y$, this implies $f(t,u(t)) \in \text{ext}\, F(t,u(t))$ for almost every $t \in [0,T]$. By possibly redefining f on a set of measure zero, this yields (1.1.1). $\square$

1.4. Applications

Throughout this section we make the following assumptions.

(H) $F:[0,T]\times\Omega\to\overline{B}(0,M)$ is a bounded continuous multifunction with compact values satisfying (LSP), while D is a compact set such that $\overline{B}(D,\ MT)\subset\Omega$.

An immediate consequence of Theorem 1.1.1 is

Corollary 1.4.1 *Let the hypotheses (H) hold. Then there exists a continuous map $(t_0,x_0)\mapsto x(\cdot,t_0,x_0)$ from $[0,T]\times D$ into $\mathcal{AC}$, such that*

$$\begin{cases} \dot{x}(t,t_0,x_0)\in \mathrm{ext}\,F\big(t,x(t,t_0,x_0)\big) & \forall t\in[0,T],\\ x(t_0,t_0,x_0)=x_0 & \forall t_0,x_0. \end{cases}$$

Another consequence of Theorem 1.1.1 is the contractibility of the sets of solutions of certain differential inclusions. We recall here that a metric space X is contractible if there exist a point $\tilde{u}\in X$ and a continuous mapping $\Phi: X\times[0,1]\to X$ such that

$$\Phi(v,0)=\tilde{u},\qquad \Phi(v,1)=v,\qquad \forall v\in X.$$

The map Φ is then called a *null homotopy* of X.

Corollary 1.4.2 *Let the assumptions (H) hold. Then, for any $\bar{x}\in D$, the sets $\mathcal{M}$, $\mathcal{M}^{ext}$ of solutions of*

$$x(0)=\bar{x},\qquad \dot{x}(t)\in F(t,x(t)),\qquad t\in[0,T],$$

$$x(0)=\bar{x},\qquad \dot{x}\in \mathrm{ext}\,F(t,x(t)),\qquad t\in[0,T],$$

are both contractible in $\mathcal{AC}$.

Proof. Let f be a selection from ext F with the properties stated in Theorem 1.1.1. As usual, we denote by $x(\cdot,t_0,x_0)$ the unique solution of the Cauchy problem (1.1.2). Define the null homotopy $\Phi:\mathcal{M}\times[0,1]\to\mathcal{M}$ by setting

$$\Phi(v,\lambda)(t)\doteq\begin{cases} v(t), & \text{if } t\in[0,\lambda T],\\ x(t,\lambda T,v(\lambda T)), & \text{if } t\in[\lambda T,T]. \end{cases}$$

By Theorem 1.1.1, Φ is continuous. Moreover, setting $\tilde{u}(\cdot) \doteq u(\cdot, 0, \bar{x})$, we obtain

$$\Phi(v,0) = \tilde{u}, \qquad \Phi(v,1) = v, \qquad \Phi(v,\lambda) \in \mathcal{M} \qquad \forall v \in \mathcal{M},$$

proving that $\mathcal{M}$ is contractible. We now observe that, if $v \in \mathcal{M}^{ext}$, then $\Phi(v,\lambda) \in \mathcal{M}^{ext}$ for every λ. Therefore, $\mathcal{M}^{ext}$ is contractible as well. □

Our last application is concerned with feedback controls. Let $\Omega \subseteq I\!R^n$ be open, $U \subset I\!R^m$ compact, and let $g : [0,T] \times \Omega \times U \to I\!R^n$ be a continuous function. By a well known theorem of Filippov [36], the solutions of the control system

$$\dot{x} = g(t,x,u), \qquad u \in U, \tag{1.4.1}$$

correspond to the trajectories of the differential inclusion

$$\dot{x} \in F(t,x) \doteq \{g(t,x,\omega) \quad \omega \in U\}. \tag{1.4.2}$$

In connection with (1.4.1), one can consider the "relaxed" system

$$\dot{x} = g^{\#}(t,x,u^{\#}), \qquad u^{\#} \in U^{\#}, \tag{1.4.3}$$

whose trajectories are precisely those of the differential inclusion

$$\dot{x} \in F^{\#}(t,x) \doteq \overline{\text{co}}\, F(t,x).$$

The control system (1.4.3) is obtained defining the compact set

$$U^{\#} \doteq U \times \cdots \times U \times E_{n+1} = U^{n+1} \times E_{n+1},$$

where

$$E_{n+1} \doteq \left\{ \theta = (\theta_0, \ldots, \theta_n) \; ; \quad \sum_{i=0}^{n} \theta_i = 1, \quad \theta_i \geq 0 \quad \forall i \right\}$$

is the standard simplex in $I\!R^{n+1}$, and setting

$$g^{\#}(t,x,u^{\#}) = g^{\#}(t,x,(u_0,\ldots,u_n,(\theta_0,\ldots,\theta_n))) \doteq \sum_{i=0}^{n} \theta_i g(t,x,u_i).$$

Generalized controls of the form $u^{\#} = (u_0, \ldots, u_n, \theta)$ taking values in the set $U^{n+1} \times E_{n+1}$ are called *chattering controls.*

Corollary 1.4.3 *Consider the control system (1.4.1), with* $g\colon [0,T]\times\Omega\times U \to \overline{B}(0,M)$ *Lipschitz continuous. Let* D *be a compact set with* $\overline{B}(D;\ MT) \subset \Omega$. *Let* $u^{\#}(t,x) \in U^{\#}$ *be a chattering feedback control such that the mapping*

$$(t,x) \mapsto g^{\#}(t,x,u^{\#}(t,x)) \doteq f_0(t,x)$$

is Lipschitz continuous.

Then, for every $\varepsilon_0 > 0$ *there exists a measurable feedback control* $\bar{u} = \bar{u}(t,x)$ *with the following properties:*

(a) For every (t,x)*, one has* $g(t,x,\bar{u}(t,x)) \in \operatorname{ext} F(t,x)$*, with* F *as in (1.4.2).*

(b) For every $(t_0,x_0) \in [0,T] \times D$*, the Cauchy problem*

$$\dot{x}(t) = g\big(t, x(t), \bar{u}(t, x(t))\big), \qquad x(t_0) = x_0\,,$$

has a unique solution $x(\cdot, t_0, x_0)$.

(c) If $y(\cdot, t_0, x_0)$ *denotes the (unique) solution of the Cauchy problem*

$$\dot{y} = f_0(t, y(t)), \qquad\qquad y(t_0) = x_0,$$

then for every (t_0, x_0) *one has*

$$\big|x(t,t_0,x_0) - y(t,t_0,x_0)\big| < \varepsilon_0, \qquad\qquad \forall t \in [0,T].$$

Proof. The Lipschitz continuity of g implies that the multifunction F in (1.4.2) is Lipschitz continuous in the Hausdorff metric, hence it satisfies (LSP). We can thus apply Theorem 1.1.1, and obtain a suitable selection f of $\operatorname{ext} F$, in connection with f_0, ε_0. For every (t,x), the set

$$W(t,x) \doteq \big\{\omega \in U\ ;\quad g(t,x,\omega) = f(t,x)\big\} \subset I\!R^m$$

is a compact nonempty subset of U. Let $\bar{u}(t,x) \in W(t,x)$ be the lexicographic selection. Then the feedback control $\bar{u}$ is measurable, and it is trivial to check that $\bar{u}$ satisfies all required properties. □

Chapter 2
Bang–bang property for Bolza problems in two dimensions

2.1. Introduction

The minimum time problem for control systems on a two dimensional manifold with control appearing linearly was studied by Sussmann in [49–51]. In particular, [49] contains a detailed analysis of the structure of time optimal trajectories, and conditions which ensure that every time optimal trajectory is bang–bang. In the same paper it is shown that the bang–bang property is not generic even for the minimum time problem.

In this chapter we consider the more general minimization problem

$$\min\left\{\int_0^1 h(x,u)\,dt \,\middle|\, \dot{x} = F(x) + u\,G(x),\ x(0) = x_0,\ x(1) = x_1\right\}, \tag{2.1.1}$$

with $u \in [-1,1]$, $x \in I\!R^2$, where F, G are $\mathcal{C}^2$ vector fields, $h(\cdot,+1), h(\cdot,-1) \in \mathcal{C}^2(I\!R^2)$, and we give conditions to guarantee that every optimal trajectory is bang-bang.

The proof of our result is based on the idea of "bang-bang" variations. More precisely, we use these variations as substitutes for those based on Lyapunov's Convexity Theorem as is done in problems of the calculus of variations ([16]) or of linear control theory ([47]).

We consider an optimal trajectory γ of our system. If γ is not bang-bang in a neighborhood of a point x we construct a new trajectory which is near the previous one, is bang-bang in a neighborhood of x and achieves a lower cost. In order to ensure that this new trajectory satisfies the initial condition we use the fact that, under suitable assumptions, all time optimal trajectory are bang-bang.

Problem (2.1.1) has a particular dynamic with respect to the usual problems of calculus of variations but on the other hand the assumptions on h are not too strict.

The tools used in this chapter are geometric, so it is easy to show that the statements hold also for a smooth two dimensional manifold.

As a corollary we prove an existence result for nonconvex optimization problems.

2.2. Preliminaries

A curve in $\mathbb{R}^2$ is a continuous map $\gamma : I \to \mathbb{R}^2$. We denote by $\mathrm{Dom}(\gamma)$ its domain and by $\mathrm{Range}(\gamma)$ the set $\{\gamma(t) : t \in \mathrm{Dom}(\gamma)\}$. We use the symbol $|$ to denote the restriction, e.g. $\gamma|[a_0, b_0]$.

A vector field X on $\mathbb{R}^2$ is an $\mathbb{R}^2$–valued function. Every vector field X can be written in the form

$$X = \alpha\partial_1 + \beta\partial_2 ,$$

where ∂_1, ∂_2 are the constant vector fields with components $(1,0)$, $(0,1)$, respectively, and $\alpha, \beta : \mathbb{R}^2 \to \mathbb{R}$. We denote by $J(X)$ the Jacobian matrix of X:

$$J(X) \doteq \begin{pmatrix} \partial_1\alpha & \partial_2\alpha \\ \partial_1\beta & \partial_2\beta \end{pmatrix} .$$

The Lie–bracket of two vector fields X, Y is the vector field defined by

$$[X, Y] = J(Y)\cdot X - J(X)\cdot Y.$$

A control u is a measurable function $u : [a, b] \to [-1, 1]$ and a trajectory for u is an absolutely continuous curve γ such that

$$\gamma'(t) = F(\gamma(t)) + u(t)G(\gamma(t)) ,$$

for almost all $t \in [a, b]$. We denote by $\mathrm{Traj}(\Sigma)$ the set of all trajectories of Σ. A *bang-bang* trajectory is a trajectory corresponding to a control u such that $|u(t)| = 1$ for almost all $t \in [a, b]$.

Given a trajectory $\gamma : [a,b] \to I\!R^2$, of a control u, and a cost function h we denote by $T(\gamma)$ the time along γ, i.e. $b-a$, and by $\Gamma_h(\gamma)$ the cost of γ:

$$\Gamma_h(\gamma) \doteq \int_a^b h\big(\gamma(t), u(t)\big)\, dt.$$

Moreover we denote by $\mathrm{In}(\gamma)$ the initial point of γ, i.e. $\gamma(a)$, and by $\mathrm{Term}(\gamma)$ its terminal point. We say that γ is time optimal if $T(\gamma) \le T(\hat{\gamma})$ for every trajectory $\hat{\gamma}$ that steers $\mathrm{In}(\gamma)$ to $\mathrm{Term}(\gamma)$. We say that (γ, u) is *optimal* (or simply that γ is optimal) if $\Gamma_h(\gamma) \le \Gamma_h(\gamma')$ for every admissible pair (γ', u') that steers $\mathrm{In}(\gamma)$ to $\mathrm{Term}(\gamma)$.

If $\gamma_1 : [a,b] \to I\!R^2$, $\gamma_2 : [b,c] \to I\!R^2$ are trajectories such that $\gamma_1(b) = \gamma_2(b)$ then $\gamma_2 * \gamma_1$ is the trajectory

$$(\gamma_2 * \gamma_1)(t) \doteq \begin{cases} \gamma_1(t), & t \in [a,b], \\ \gamma_2(t), & t \in [b,c]. \end{cases}$$

In the following we shall use the notation $X \doteq F - G$, $Y \doteq F + G$. A trajectory $\gamma \in \mathrm{Traj}(\Sigma)$ is a Y–trajectory if it is an integral curve for the vector field Y, and a X–trajectory is defined similarly. Moreover a $Y * X$–trajectory is a concatenation of a X–trajectory and a Y–trajectory and similarly is defined a $X * Y$–trajectory.

An *admissible pair* is a pair (γ, u) such that $u : [a,b] \to [-1,1]$ is a control and γ is a trajectory for u. If (γ, u) is an admissible pair and γ is time optimal then (γ, u) satisfies the Pontryagin Maximum Principle (briefly PMP). See [23] and [37] for the general theory and [49] for the PMP stated for control system of the same type of Σ.

Given Σ we can define the functions

$$(2.2.1) \qquad \Delta_A \doteq \det(F, G)\,, \quad \Delta_B \doteq \det(G, [F,G])\,, \quad \Delta_C \doteq \det(F, [F,G])\,,$$

where "det" stands for the determinant. Writing F, G, $[F,G]$ as column vectors we can form a 2×2 matrix having two of them as columns. The functions Δ_A and Δ_B have been introduced in [49].

If F, G are independent at each point of an open connected set $\Omega^{\dagger} \subset I\!R^2$, then we can define the 1-differential form ω in the following way:

$$\langle \omega(x), F(x) \rangle = 1, \qquad \langle \omega(x), G(x) \rangle = 0. \tag{2.2.2}$$

Consider now two trajectories $\gamma_1, \gamma_2 \in \mathrm{Traj}(\Sigma)$ such that $\gamma_2^{-1} * \gamma_1$ is a simple closed curve oriented counterclockwise, whose interior is contained in $\Omega^{\dagger}$ (here γ_2^{-1} is γ_2 run backwards). In [49] it was proved the following

Theorem 2.2.1 *Consider $\gamma_1, \gamma_2 \in \mathrm{Traj}(\Sigma)$ as above. We have*

$$T(\gamma_1) - T(\gamma_2) = \int_{\gamma_2^{-1} * \gamma_1} \omega = \int_{\mathcal{R}} d\omega ,$$

where $\mathcal{R}$ is the region enclosed by $\gamma_2^{-1} * \gamma_1$.

If F, G are independent, there exist $f, h \in \mathcal{C}(I\!R^2, I\!R)$ such that

$$[F, G](x) = f(x)F(x) + h(x)G(x),$$

and from (2.2.1) we obtain $\Delta_B = \det(G, fF + hG)$, which gives $f = -\Delta_B/\Delta_A$. One can check that $d\omega = -(f/\Delta_A)\, dx_1 \wedge dx_2$.

We say that $x \in I\!R^2$ is an ordinary point if $\Delta_A(x) \cdot \Delta_B(x) \neq 0$. For ordinary points we have the following

Theorem 2.2.2 *Let $\Omega \subset I\!R^2$ be an open set such that each $x \in \Omega$ is an ordinary point. Then all optimal trajectories γ for the restriction of Σ to Ω are bang-bang with at most one switching. Moreover if $f > 0$ throughout Ω then γ is a X, Y or $Y * X$–trajectory, if $f < 0$ throughout Ω then γ is a X, Y or $X * Y$–trajectory.*

For the proof see [49] Theorem 3.9 p. 443.

Let us define the linear interpolation $l(x, u) \doteq m(x)u + q(x)$ between $h(x, +1)$ and $h(x, -1)$, where

$$m(x) \doteq \frac{h(x, +1) - h(x, -1)}{2}, \quad q(x) \doteq \frac{h(x, +1) + h(x, -1)}{2}. \tag{2.2.3}$$

We make the following assumptions on (2.1.1):

(H_1) $\quad \Delta_A(x)\,\Delta_B(x) \neq 0\,, \quad$ for every $x \in I\!R^2$,

(H_2) $\quad h(x,u) \geq l(x,u)\,, \quad$ for every $x \in I\!R^2$ and $u \in [-1,1]$.

In [49] it was proved that under the assumption (H_1) all time optimal trajectories are bang-bang.

From (H_1) it follows that F and G are independent at each point, and then we can define the 1-differential form ω_l in the following way:

$$\langle \omega_l(x), F(x)\rangle = q(x)\,, \quad \langle \omega_l(x), G(x)\rangle = m(x), \quad \text{for every } x \in I\!R^2\,.$$

Following [49] we have, by Stokes' theorem:

$$\Gamma_l(\gamma_1) - \Gamma_l(\gamma_2) = \int_{\gamma_2^{-1} * \gamma_1} \omega_l = \int_{\mathcal{R}} d\omega_l\,,$$

where $\mathcal{R}$, γ_1 and g_2 are as in Theorem 2.2.1.

With straightforward calculations (see [49]) we obtain

$$T(\gamma_1) - T(\gamma_2) = \int_{\mathcal{R}} \varphi\; dx \wedge dy\,, \quad \Gamma_l(\gamma_1) - \Gamma_l(\gamma_2) = \int_{\mathcal{R}} \varphi_l\; dx \wedge dy\,, \tag{2.2.4}$$

where

$$\varphi \doteq \frac{\Delta_B}{\Delta_A^2}\,, \quad \varphi_l \doteq \frac{\nabla m \cdot F - \nabla q \cdot G}{\Delta_A} + \frac{q\;\Delta_B - m\;\Delta_c}{\Delta_A^2}\,. \tag{2.2.5}$$

Given $f_1, f_2 \colon I\!R^2 \to I\!R$ we define the vector field

$$[f_1|f_2](x) \doteq f_1(x)\;\nabla f_2(x) - f_2(x)\;\nabla f_1(x). \tag{2.2.6}$$

Let us define the cone

$$\mathcal{K}(x) \doteq \{v \in I\!R^2 \mid S_{AB}\; v \cdot X(x) < 0 \text{ or } S_{AB}\; v \cdot Y(x) > 0\}\,, \tag{2.2.7}$$

where $S_{AB} \doteq \text{sign}\,(\Delta_A\;\Delta_B)$ (we remark that, from (H_1), S_{AB} is independent of x). The last assumption on (2.1.1) is

(H_3) $\quad D(x) \doteq [\varphi|\varphi_l](x) \in \mathcal{K}(x)\,, \quad$ for every $x \in I\!R^2$.

2.3. The main result

In this section we will prove the main result; the key role is played by the following lemma.

Lemma 2.3.1 *Let (γ, u) be an admissible pair of*

$$\dot{x} = F(x) + u\, G(x)\,, \tag{2.3.1}$$

and let $E \doteq \{t \in [0,1] \,:\, |u(t)| < 1\}$. Suppose that there exists a Lebesgue point τ for E of differentiability for γ. If (H_1), (H_2) hold and $D(\overline{x}) \in \mathcal{K}(\overline{x})$, $\overline{x} \doteq \gamma(\tau)$, then there exist $\varepsilon > 0$, $\sigma > 0$ and an admissible pair $(\hat{\gamma}, \hat{u})$ such that:

$$\hat{\gamma}(t) = \gamma(t)\,, \quad \text{for every } t \notin [\tau - \varepsilon, \tau + \sigma]\,,$$
$$|\hat{u}(t)| = 1\,, \quad \text{for every } t \in [\tau - \varepsilon, \tau + \sigma]\,,$$
$$\int_0^1 h(\gamma(t), u(t))\, dt \;>\; \int_0^1 h(\hat{\gamma}(t), \hat{u}(t))\, dt\,.$$

Proof. As a first step we will construct a pair of admissible trajectories $\gamma^{\pm}$, which are bang–bang in a neighborhood of τ, and correspond to controls $u^{\pm}$ which change sign two times in this neighborhood. This is possible thanks to (H_1); using (H_2) and (H_3) we will show that one of these trajectories achieves a better performance with respect to γ. We start constructing two parametric families of bang–bang trajectories $\gamma_\mu^{\pm}$ which start from $\gamma(\tau - \varepsilon)$ and reach x at times $\tau^{\pm}$ respectively, having two switchings at times $\sigma_1^{\pm}$, μ.

Fix $0 < \varepsilon < \min(\tau, 1 - \tau)$. If ε is sufficiently small, there exist $\sigma_1^{\pm} \geq \tau - \varepsilon$ and $\tau^{\pm} \geq \sigma_1^{\pm}$, depending on ε, two one–parameter families of pairs $(\gamma_\mu^{\pm}, u_\mu^{\pm})$, $\mu \in [\tau^{\pm}, 1]$ ($\mu - \tau^{\pm}$ small), $\sigma_2^{\pm} = \sigma_2^{\pm}(\varepsilon, \mu) \in [\mu, 1]$, $\sigma_3^{\pm} = \sigma_3^{\pm}(\varepsilon, \mu) \in [\tau, 1]$ such that

$$\gamma_\mu^{\pm}(\tau - \varepsilon) = \gamma(\tau - \varepsilon)\,, \quad \gamma_\mu^{\pm}(\tau^{\pm}) = \gamma(\tau)\,, \quad \gamma_\mu^{\pm}(\sigma_2^{\pm}) = \gamma(\sigma_3^{\pm})\,,$$

$$u_\mu^{\pm}(s) = \begin{cases} \pm 1, & \text{if } s \in [\tau - \varepsilon, \sigma_1^{\pm}) \cup (\mu, \sigma_2^{\pm}], \\ \mp 1, & \text{if } s \in [\sigma_1^{\pm}, \mu]. \end{cases}$$

Let $A^{\pm}$, $B_{\mu}^{\pm}$ be the regions enclosed respectively by $(\gamma_{\mu}^{\pm}|[\tau-\varepsilon,\tau^{\pm}])^{-1} * \gamma|[\tau-\varepsilon,\tau]$ and $(\gamma_{\mu}^{\pm}|[\tau^{\pm},\sigma_2^{\pm}])^{-1} * \gamma|[\tau,\sigma_3^{\pm}]$. From (2.2.4) it follows that

$$\begin{aligned}\psi^{\pm}(\mu) &\doteq \pm\,\mathrm{sign}\,(\Delta_A\cdot\Delta_B)[T(\gamma)-T(\gamma_{\mu}^{\pm})] = \\ &= \mathrm{sign}\,(\Delta_B)\cdot\left[\int_{A^{\pm}}\varphi\;dx-\int_{B_{\mu}^{\pm}}\varphi\;dx\right].\end{aligned}\tag{2.3.2}$$

In fact if $\Delta_A > 0$ we have $\det(X,Y) > 0$, and the regions A^+ and B^- are enclosed by curves oriented counterclockwise, while A^- and B^+ are enclosed by curves oriented clockwise; the opposite happens if $\Delta_A < 0$.

Moreover $\psi^{\pm}(\tau^{\pm}) > 0$, $\psi^{\pm}$ is decreasing, and for ε sufficiently small there exists $\mu^{\pm}\in[\tau^{\pm},1]$ such that $\psi^{\pm}(\mu^{\pm})=0$. Therefore $\sigma_2^{\pm}=\sigma_3^{\pm}$, and then we can define the admissible pairs $(\gamma^{\pm},u^{\pm})$ as follows:

$$\gamma^{\pm}(t)\doteq\begin{cases}\gamma_{\mu^{\pm}}^{\pm}, & t\in I^{\pm}\doteq[\tau-\varepsilon,\sigma_3^{\pm}(\mu^{\pm})],\\ \gamma(t), & t\in[0,1]\backslash I^{\pm},\end{cases}$$
$$u^{\pm}(t)\doteq\begin{cases}u_{\mu^{\pm}}^{\pm}, & t\in I^{\pm},\\ u(t), & t\in[0,1]\backslash I^{\pm}.\end{cases}$$

Let us define $B^{\pm}\doteq B_{\mu^{\pm}}^{\pm}$. Using the expansion $\varphi(x)=\varphi(\overline{x})+\nabla\varphi(\overline{x})\cdot(x-\overline{x})+o(|x-\overline{x}|)$ we obtain

$$\begin{aligned}0&=\int_{A^{\pm}}\varphi\;dx-\int_{B^{\pm}}\varphi\;dx=\\ &=\left[m(A^{\pm})-m(B^{\pm})\right]\cdot\varphi(\overline{x})+\nabla\varphi(\overline{x})\cdot d^{\pm}+o(\varepsilon^3),\end{aligned}\tag{2.3.3}$$

where $d^{\pm}=d^{\pm}(\varepsilon)\doteq m(A^{\pm})b(A^{\pm})-m(B^{\pm})b(B^{\pm})$ and $b(\mathcal{R})$ denotes the baricenter of $\mathcal{R}$ relative to $\overline{x}$. Expanding φ_l we obtain in the same way

$$\begin{aligned}\delta^{\pm}&\doteq\Gamma_l(\gamma)-\Gamma_l(\gamma^{\pm})=\\ &=\pm\mathrm{sign}\;\Delta_A\cdot\left\{\left[m(A^{\pm})-m(B^{\pm})\right]\varphi_l(\overline{x})+\nabla\varphi_l(\overline{x})\cdot d^{\pm}\right\}+o(\varepsilon^3).\end{aligned}\tag{2.3.4}$$

By (2.3.3) and multiplying (2.3.4) by $\varphi(\overline{x})$ one has

$$\varphi(\overline{x})\delta^{\pm}=\pm\mathrm{sign}\;(\Delta_A)\;[\varphi|\varphi_l](\overline{x})\cdot d^{\pm}+o(\varepsilon^3)\,,$$

where $[\cdot|\cdot]$ is defined in (2.2.6). Let us define the cone

$$\mathcal{C}(v,w) \doteq \{\lambda v + \mu w \mid \lambda \geq 0,\ \mu \geq 0\}.$$

For ε small enough we have that $d^{\pm} \in \mathcal{C}(-X(\overline{x}), -Y(\overline{x}))$. Moreover it is easy to verify that sign $[\det(d^+, d^-)] = $ sign (Δ_A), and then, by (2.2.7),

$$\left\{v \in I\!R^2 \mid S_{AB}\, v \cdot d^{\pm} \leq 0\right\} \cap \mathcal{K}(\overline{x}) = \emptyset.$$

Since sign $\Delta_B = $ sign φ, it follows that either $\delta^+ > 0$ or $\delta^- > 0$. Suppose, for instance, that $\delta^+ > 0$. From (H_2) it follows that

$$\int_{\gamma\cap I^+} h(x,u) \geq \int_{\gamma\cap I^+} l(x,u) > \int_{\gamma^+\cap I^+} l(x,u) = \int_{\gamma^+\cap I^+} h(x,u).$$

Since the same inequality holds if $\delta^- > 0$, with γ^+ and I^+ replaced respectively by γ^- and I^-, the lemma is proved. □

The following theorem is a direct consequence of the previous lemma.

Theorem 2.3.2 *If (H_1), (H_2), (H_3) hold, then all optimal trajectories of (2.1.1) are bang–bang.*

Proof. Let (γ, u) be an optimal pair for (2.1.1), and suppose that $m(E) > 0$, where E is as in Lemma 2.3.1. Almost every $\tau \in E$ is a Lebesgue point for E and a point of differentiability for γ. If τ is such a point, all the assumptions of Lemma 2.3.1 are satisfied, and then γ is not optimal. □

2.4. Examples and Applications

In this section we give two examples of minimization problems without the bang–bang property and one application of the main result.

Example 2.4.1 Consider the following problem in $\Omega \doteq \{(x_1, x_2) \mid x_2 < 0\}$:

$$\text{(2.4.1)}\quad \min\left\{\int_0^1 u\; dt \,\Big|\, \dot{x}_1 = x_2,\ \dot{x}_2 = -x_2^2 - x_2 + u,\ x(0) = x_0,\ x(1) = x_1\right\},$$

with $x_0 = (0, -1/2)$, $x_1 = (-1/2, -1/2)$, i.e. problem (2.1.1) with

$$F(x) = \begin{pmatrix} x_2 \\ -x_2^2 - x_2 \end{pmatrix}, \quad G(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad h(x,u) = u\,.$$

The unique optimal pair (γ, u) is

$$\gamma(t) = \left(-\frac{t}{2}, -\frac{1}{2}\right), \quad u(t) = -\frac{1}{4}, \qquad t \in [0,1]\,,$$

with a cost $\Gamma(\gamma) = -1/4$. By a straightforward calculation we have that $\Delta_A = x_2$, $\Delta_B = 1$, $\Delta_C = x_2^2$, $S_{AB} = -1$, $m(x) = 1$, $q(x) = 0$.

Conditions (H_1) and (H_2) are satisfied in Ω. We now check that (H_3) does not hold. Recalling (2.2.5) it is easy to verify that

$$\varphi = \frac{1}{x_2^2}, \quad \varphi_l = -1, \quad D = \begin{pmatrix} 0 \\ -\frac{2}{x_2^3} \end{pmatrix},$$

$$X = \begin{pmatrix} x_2 \\ -x_2^2 - x_2 - 1 \end{pmatrix}, \quad Y = \begin{pmatrix} x_2 \\ -x_2^2 - x_2 + 1 \end{pmatrix},$$

$$S_{AB}\, D \cdot X = -2\,\frac{x_2^2 + x_2 + 1}{x_2^3}, \qquad S_{AB}\, D \cdot Y = -2\,\frac{x_2^2 + x_2 - 1}{x_2^3}\,.$$

Since $S_{AB}\ D(\gamma(t)) \cdot X(\gamma(t)) > 0$ and $S_{AB}\ D(\gamma(t)) \cdot Y(\gamma(t)) < 0$ for every $t \in [0,1]$, we have that $D(\gamma(t)) \notin \mathcal{K}(\gamma(t))$ for every $t \in [0,1]$.

Let us rewrite problem (2.4.1) in Mayer form, introducing the new variable x_3 such that $\dot{x_3} = u$, and consider the two–dimensional system:

$$\dot{x}_2 = -x_2^2 - x_2 + u, \quad \dot{x}_3 = u\,. \tag{2.4.2}$$

Since this system is independent of x_1, we have that every optimal trajectory of (2.4.1) gives a time–optimal trajectory of (2.4.2). The time–optimal trajectory associated to γ in the (x_2, x_3)–plane is $\tilde{\gamma}(t) = (-1/2, -t/4)$. Using the same terminology of [49], $\tilde{\gamma}$ is a turnpike; in the same paper it was shown that, under generic conditions, these are the only cases of not bang–bang time–optimal trajectories.

Example 2.4.2 Consider now the problem in $\Omega \doteq \{(x_1, x_2) \mid x_2 < 0\}$:

$$\min\left\{\int_0^1 x_2^2 + x_2 \mid \dot{x}_1 = x_2,\ \dot{x}_2 = x_2/2 + u\, x_2,\ x(0) = x_0,\ x(1) = x_1\right\},$$

with $x_0 = (0, -1/2)$, $x_1 = (-1/2, -1/2)$, i.e. problem (2.1.1) with

$$F(x) = \begin{pmatrix} x_2 \\ \frac{1}{2}x_2 \end{pmatrix}, \quad G(x) = \begin{pmatrix} 0 \\ x_2 \end{pmatrix}, \quad h(x,u) = x_2^2 + x_2\,.$$

In this case $\Delta_A = x_2^2$, $\Delta_B = x_2^2$, $\Delta_C = \frac{1}{2}x_2^2$, $S_{AB} = 1$, $m(x) = 0$, $q(x) = x_2^2 + x_2$, $\varphi = \frac{1}{x_2^2}$, $\varphi_l = -1$,

$$D = \begin{pmatrix} 0 \\ -\frac{2}{x_2^3} \end{pmatrix}, \quad X = \begin{pmatrix} x_2 \\ -\frac{1}{2}x_2 \end{pmatrix}, \quad Y = \begin{pmatrix} x_2 \\ \frac{3}{2}x_2 \end{pmatrix},$$

$$S_{AB}\, D \cdot X = 1/x_2^2, \quad S_{AB}\, D \cdot Y = -3/x_2^2\,.$$

As in Example 2.4.1 we have that $D(\gamma(t)) \notin \mathcal{K}(\gamma(t))$ for every $t \in [0,1]$.

Passing to Mayer form as in the previous example, we obtain the system:

$$\dot{x}_2 = \frac{1}{2}x_2 + u\, x_2\,, \quad \dot{x}_3 = x_2^2 + x_2\,.$$

We have $\Delta_B = -x_2^2(2x_2 + 1) = 0$ on the time–optimal trajectory $\tilde{\gamma}(t) = (-1/2, -t/4)$, i.e. $\tilde{\gamma}$ is a turnpike.

Application. We show an example of nonconvex minimization problem, and give conditions depending only on the dynamic for the existence of an optimal solution. Let us consider the problem

$$\text{(2.4.3)} \qquad \min\left\{\int_0^1 h(u)\, dt \mid \dot{x} = F(x) + u\, G(x),\ x(0) = x_0,\ x(1) = x_1\right\},$$

with $h\colon \{-1, 1\} \to \mathbb{R}$, $x \in \Omega$, and $u \in [-1,1]$. Let us define $m \doteq [h(1) - h(-1)]/2$, $q \doteq [h(1) + h(-1)]/2$, and let us consider the convexified problem (that is $u \in [-1,1]$) with cost $\tilde{h}(u) = m\, u + q$. It is not restrictive to suppose $q = 0$, otherwise we can add a constant to h; in this case we obtain $\varphi_l = -m\, \Delta_C/\Delta_A$.

If $m = 0$, that is if $h(1) = h(-1)$, then every admissible trajectory is optimal. It is easy to show, following the proof of Lemma 2.3.1, that if there exists an admissible trajectory, then there exists an admissible bang–bang trajectory.

If $m \neq 0$ we have that $D = m\tilde{D}/\Delta_A^4$, with

$$\tilde{D} \doteq -(\nabla\Delta_A)\Delta_B\Delta_C + \Delta_A(\nabla\Delta_B)\Delta_C - \Delta_A\Delta_B(\nabla\Delta_C).$$

The following proposition is now a direct application of Theorem 2.3.2.

Proposition 2.4.3 *Let us consider the problem (2.4.3). Suppose that $\Delta_A \neq 0$, $\Delta_B \neq 0$, $\pm\tilde{D}(x) \in \mathcal{K}(x)$ for every $x \in \Omega$. If there exists an admissible pair for the convexified problem, then (2.4.3) has a solution.*

Chapter 3
Special bang–bang solutions for nonlinear control systems

3.1. Introduction

In this chapter we are concerned with the control systems

$$x'' + a(x,x') = u\; g(x,x')\,, \tag{3.1.1}$$

where a, g are continuously differentiable functions on $\mathbb{R}^2$ and u is scalar control with $|u| \leq 1$. For a given solution x of (3.1.1) defined on an interval $[a,b]$, we prove, under mild assumptions on g, the existence of two bang–bang solutions y, z, with a finite number of switchings, satisfying

$$\begin{aligned} x(a) = y(a) = z(a), \quad x(b) = y(b) = z(b),\\ x'(a) = y'(a) = z'(a), \quad x'(b) = y'(b) = z'(b), \end{aligned} \tag{3.1.2}$$

and, for every $t \in [a,b]$,

$$y(t) \leq x(t) \leq z(t). \tag{3.1.3}$$

We remark that every forced semilinear second order differential equation, corresponding to (3.1.1) with g constant, satisfies the required assumptions. For example, the forced nonlinear pendulum, the forced nonlinear Duffing oscillator, and the forced Van der Pol equation belong to this class of equations.

The problem of finding a bang–bang solution satisfying (3.1.2) and (3.1.3) in the case of linear control systems $L(x) \in [\phi_1(t), \phi_2(t)]$, where L is a linear operator of order n, was studied in the case of piecewise analytic data by Andreini and Bacciotti in [4]. The techniques used in this paper are based on Lyapunov type theorems (see [23] for examples and applications). Recently Cërf and Mariconda in [19] studied the case $\phi_1, \phi_2 \in L^1$. Their approach is based on a new Lyapunov type theorem applied to the integral representation of the solutions.

In order to study problem (3.1.1), a completely different approach has to be used due to the nonlinearity of the operator. In particular we have no integral representation formula and we have to argue about the geometric properties of the trajectories. The control problem (3.1.1) can be restated as a first order control system in $\mathbb{R}^2$ with control appearing linearly. The properties of time optimal trajectories for such kind of systems were deeply studied in [49], [50], and the synthesis problem in [46], [51], with geometric methods.

Given a trajectory γ of the planar system, we construct a bang–bang trajectory $\hat{\gamma}$ such that the corresponding solution of (3.1.1) satisfies (3.1.2) and (3.1.3). We begin constructing a local bang–bang variation of γ, with the requested properties, on intervals of uniformly positive length, and then we concatenate these variations in order to obtain $\hat{\gamma}$. Closure results for the reachable sets of control systems with obstacle follow directly from our theorems.

The idea of bang–bang variations was used in Chapter 2 to prove existence results for nonconvex optimization problems. Our result can be used to avoid the classical convexity assumptions for the problem of minimizing some integral functionals with nonlinear dynamics. Existence results without convexity assumptions were recently obtained in [16], [41] and in [47] for optimal control problems.

3.2. Preliminaries

We consider the control problem

$$x'' + a(x, x') = u\, g(x, x'), \qquad u \in [-1, 1], \tag{3.2.1}$$

where a, g are continuously differentiable real function on $\mathbb{R}^2$, and the prime denotes differentiation with respect to the time variable. Defining $x_1 = x$ and

$x_2 = x'$ the control problem (3.2.1) is equivalent to the first order control system in $\mathbb{R}^2$

$$\begin{cases} x_1' = x_2\,, \\ x_2' = -a(x_1,x_2) + u\; g(x_1,x_2)\,. \end{cases} \tag{3.2.2}$$

In this system the control appears linearly and we can also write it as

$$x' \;=\; F(x) + u\; G(x), \qquad u \in [-1,1], \tag{3.2.3}$$

where

$$F = \begin{pmatrix} x_2 \\ -a(x_1,x_2) \end{pmatrix}, \qquad G = \begin{pmatrix} 0 \\ g(x_1,x_2) \end{pmatrix}. \tag{3.2.4}$$

In the following we shall denote by Σ the system (3.2.3). We shall use the notations given in Section 2.2.

Given two points $P, Q \in \mathbb{R}^2$, let us define

$$\Gamma(P,Q) \doteq \{\gamma \in \mathrm{Traj}(\Sigma) \;:\; \mathrm{In}(\gamma) = P,\ \mathrm{Term}(\gamma) = Q\}\,,$$

$$\Gamma_\Omega(P,Q) \doteq \{\gamma \in \Gamma(P,Q) \;:\; \mathrm{Range}(\gamma) \subset \Omega\}\,.$$

If $P = (x_1,x_2) \in \mathbb{R}^2$, let us define $\pi_j(P) \doteq x_j$, $j = 1,2$. For $P \in \mathbb{R}^2$, let us define γ_P^+ (resp. γ_P^-) as the maximal Y-trajectory (resp. X-trajectory) such that $0 \in \mathrm{Dom}(\gamma)$ and $\gamma_P^+(0) = P$ (resp. $\gamma_P^-(0) = P$).

3.3. The main result

In this section we shall prove the following

Theorem 3.3.1 *Consider the control problem*

$$x'' + a(x,x') = u\; g(x,x'), \tag{P}$$

with $u \in [-1,1]$, $a, g \in C^1(\mathbb{R}^2,\mathbb{R})$. Assume that the following hypotheses hold:

(i) $\mathcal{M} \doteq \{x \in \mathbb{R}^2 \;:\; g(x) = 0\}$ is a regular manifold, and $\nabla g(x) \neq 0$ for every $x \in \mathcal{M}$;

(ii) *for every* $R > 0$, *the set* $\mathcal{Z}_R \doteq \{x \in \mathcal{M} \ : \ F(x) \cdot \nabla g(x) = 0\} \cap B(0, R)$, *where* $F(x_1, x_2) \doteq (x_2, -a(x_1, x_2))$, *has a finite number of connected components.*

If x *is a solution of* (P) *defined on* $[a, b]$, *then there exist two bang–bang solutions* y, z, *defined on* $[a, b]$ *such that*

$$x(a) = y(a) = z(a), \quad x(b) = y(b) = z(b),$$
$$x'(a) = y'(a) = z'(a), \quad x'(b) = y'(b) = z'(b),$$

and, for every $t \in [a, b]$,

$$y(t) \leq x(t) \leq z(t).$$

Moreover, y *and* z *have a finite number of switchings.*

Example 3.3.2 Consider the nonlinear control system

$$x'' + a(x, x') = u,$$

that is system (P) with $g(x, x') = 1$. Since $\mathcal{M} = \emptyset$, the assumptions of the theorem are trivially satisfied. Notice that most of the second–order differential equations coming from mathematical physics and applied sciences, such as the forced nonlinear pendulum, the forced nonlinear Duffing oscillator, the forced Van der Pol equation, and so on, belong to this class.

We can associate to problem (P) a first order two–dimensional control system, as in Section 3.2. For this system, that we shall call Σ in the following, by a simple computation we have that

$$\Delta_{\mathcal{A}}(x_1, x_2) = x_2 \cdot g(x_1, x_2), \quad \Delta_{\mathcal{B}}(x_1, x_2) = g^2(x_1, x_2), \quad d\omega = \frac{1}{x_2^2}\, dx_1 \wedge dx_2.$$

Let us define the two regions of the plane

$$\Omega^+ \doteq \left\{(x_1, x_2) \in I\!R^2 \,\middle|\, x_2 > 0\right\}, \qquad \Omega^- \doteq \left\{(x_1, x_2) \in I\!R^2 \,\middle|\, x_2 < 0\right\}.$$

Lemma 3.3.3 *Suppose that Ω is an open connected subset of Ω^+, and that P and Q are two points of Ω with $\pi_1(P) < \pi_1(Q)$. Moreover, suppose that the following hypothesis hold:*

(H) Δ_A and Δ_B are both different from zero in Ω. Let $t^\pm > 0$ be the first time of intersection of $\gamma_P^\pm$ with Ω^c, or $+\infty$ if there is no intersection. Define $s^\pm < 0$ (possibly $-\infty$) in the same way for $\gamma_Q^\pm$ with reversed time. Assume that $\gamma_P^+|[0,t^+]$ intersects $\gamma_Q^-|[s^-,0]$ in a point $R \in \Omega$, and that $\gamma_P^-|[0,t^-]$ intersects $\gamma_Q^+|[s^+,0]$ in a point $S \in \Omega$.

If $\gamma \in \Gamma_\Omega(P,Q)$, then there exists a bang-bang trajectory $\hat{\gamma} \in \Gamma_\Omega(P,Q)$ such that

$$\mathrm{Dom}(\gamma) = \mathrm{Dom}(\hat{\gamma}), \qquad \pi_1(\gamma(t)) \leq \pi_1(\hat{\gamma}(t)) \quad \text{for every } t \in \mathrm{Dom}(\gamma).$$

Proof. Since $x_1' = x_2 > 0$, all the trajectories lying in Ω go from left to right in the (x_1,x_2)–plane. Since $g^2 = \Delta_B \neq 0$ in Ω, we can assume, for instance, $g > 0$ in Ω (the case $g < 0$ can be treated in a similar way). If $R = \gamma_P^+(t_1) = \gamma_Q^-(s_1)$ and $S = \gamma_P^-(t_2) = \gamma_Q^+(s_2)$, let us define the trajectories

$$\gamma_1(t) \doteq \begin{cases} \gamma_P^+(t), & \text{if } t \in [0,t_1], \\ \gamma_Q^-(t+s_1-t_1), & \text{if } t \in [t_1, t_1 - s_1]; \end{cases}$$

$$\gamma_2(t) \doteq \begin{cases} \gamma_P^-(t), & \text{if } t \in [0,t_2], \\ \gamma_Q^+(t+s_2-t_2), & \text{if } t \in [t_2, t_2 - s_2]; \end{cases}$$

i.e. γ_1 and γ_2 are respectively the $X*Y$–trajectory and the $Y*X$–trajectory steering P to Q. For every $\eta \in \Gamma_\Omega(P,Q)$, we can apply Theorem 2.2.1 to the pairs (γ_1,η) and (γ_2,η) obtaining

$$T(\gamma_1) \leq T(\eta) \leq T(\gamma_2), \qquad \eta \in \Gamma_\Omega(P,Q). \tag{3.3.1}$$

Indeed the coefficient of the differential form $d\omega$ is positive in Ω. For each $\mu \in [0,t_1]$, we construct the trajectory $\gamma_\mu \in \Gamma_\Omega(P,Q)$ in the following way. Starting from P, we follow γ_P^+ up to time μ. Then we follow the X–trajectory through $\gamma_P^+(\mu)$, until it intersects γ_Q^+ at time σ_μ, and finally we follow γ_Q^+

up to the point Q. By Theorem 2.2.1 the map $\mu \mapsto T(\gamma_\mu)$ is continuous and strictly decreasing. Hence, from (3.3.1) there exists $\hat{\mu} \in [0, t_1]$ such that $T(\gamma) = T(\hat{\gamma})$, where $\hat{\gamma} \doteq \gamma_{\hat{\mu}}$. Without loss of generality we can assume that $\mathrm{Dom}(\gamma) = \mathrm{Dom}(\hat{\gamma}) = [0, T]$. Define $\xi(t) \doteq \pi_1(\gamma(t))$, $\hat{\xi}(t) \doteq \pi_1(\hat{\gamma}(t))$, for every $t \in [0, T]$. It remains to prove that $\hat{\xi}(t) \geq \xi(t)$, for every $t \in [0, T]$. Being $\xi'(t)$ and $\hat{\xi}'(t)$ strictly positive for a.e. t, we can construct the two inverse functions $t(x_1)$, $\hat{t}(x_1)$, for $x_1 \in [\pi_1(P), \pi_1(Q)]$. Define $\phi(x_1) \doteq \pi_2(\gamma(t(x_1)))$, $\hat{\phi}(x_1) \doteq \pi_2(\hat{\gamma}(\hat{t}(x_1)))$; ξ and $\hat{\xi}$ are solutions of the Cauchy problems

$$\xi' = \phi(\xi), \qquad \xi(0) = \pi_1(P),$$

$$\hat{\xi}' = \hat{\phi}(\hat{\xi}), \qquad \hat{\xi}(0) = \pi_1(P),$$

with the same initial data. Notice that $\mathrm{Range}(\gamma)$ intersects $\mathrm{Range}(\hat{\gamma})$ in at least one point $A \doteq \hat{\gamma}(t_A)$ with $\hat{\mu} \leq t_A \leq \sigma_\mu$: indeed the range of every $\eta \in \Gamma_\Omega(P, Q)$ is contained in the region enclosed by γ_1 and γ_2. Hence $\phi(x_1) \leq \hat{\phi}(x_1)$ for every $x_1 \in [\pi_1(P), \pi_1(A)]$, and then $\hat{\xi}(t) \geq \xi(t)$ for every $t \in [0, t_A]$. In the same way, from $\xi(T) = \hat{\xi}(T)$, it follows that the same inequality holds for $t \in [t_A, T]$. □

Lemma 3.3.4 *Suppose that Ω is an open connected subset of Ω^+, and that P and Q are two points of $\overline{\Omega}$ with $\pi_1(P) < \pi_1(Q)$. Moreover, suppose that the hypothesis (H) hold. Then the conclusions of Lemma 3.3.3 hold.*

Proof. The case $P, Q \in \Omega$ has been considered in Lemma 3.3.3. Suppose now that $P \in \Omega$, $Q \in \partial\Omega$. It is not restrictive to assume $\mathrm{Dom}(\gamma) = [0, T]$. By the assumptions made on R and S, it is clear that $\gamma(t) \in \Omega$, for every $t \in [0, T)$. Indeed X and Y do not vanish and point outward Ω. Consider a sequence $(T_n)_n$, $0 < T_n < T$, converging to T for $n \to +\infty$, and define $\gamma_n \doteq \gamma | [0, T_n]$, $Q_n \doteq \mathrm{Term}(\gamma_n)$. Applying Lemma 3.3.3 to γ_n, we construct a bang-bang trajectory $\hat{\gamma}_n$ such that

$$\pi_1(\gamma_n(t)) \leq \pi_1(\hat{\gamma}_n(t)), \qquad \text{for every } t \in [0, T_n]. \tag{3.3.2}$$

Extend $\hat{\gamma}_n$ to $[0,T]$ defining $\hat{\gamma}_n(t) = \gamma(t)$ for $t \in [T_n, T]$. It is clear from the proof of Lemma 3.3.3 that we can extract a subsequence $(\hat{\gamma}_{n'})_{n'}$ of $(\hat{\gamma}_n)_n$ converging uniformly to a bang–bang trajectory $\hat{\gamma} \in \Gamma(P,Q)$, with $T(\hat{\gamma}) = T(\gamma)$. Passing to the limit in (3.3.2), we are done. Similarly we can treat the other cases. □

Remark 3.3.5 Lemma 3.3.3 and 3.3.4 hold if in (H) we replace Ω^+ by Ω^-.

Lemma 3.3.6 *Suppose $\pi_2(P) = 0$ and $|g(P)| > |a(P)|$. Then there exists $\tau > 0$ with the following property. For every trajectory γ starting from P, with $T(\gamma) = \tau$ and lying in Ω^+ for positive times, there exists a bang–bang trajectory $\hat{\gamma}$ satisfying:*

$$\begin{aligned} &\mathrm{In}(\gamma) = \mathrm{In}(\hat{\gamma}), && \mathrm{Term}(\gamma) = \mathrm{Term}(\hat{\gamma}), \\ &\mathrm{Dom}(\gamma) = \mathrm{Dom}(\hat{\gamma}), && \pi_1(\gamma(t)) \le \pi_1(\hat{\gamma}(t)) \quad \text{for every } t \in \mathrm{Dom}(\gamma). \end{aligned} \tag{3.3.3}$$

Proof. Let $R(t)$ be the reachable set in time t from P, that is the set of points Q for which there exists a trajectory steering P to Q in time less than or equal to t. Since $|g(P)| > |a(P)|$, there exists a neighborhood of P such that X and Y are non vanishing vectors which point to opposite sides of the x_1–axis, and $\Delta_B \neq 0$. Therefore there exists $\tau > 0$ small enough such that $R(\tau) \cap \Delta_B^{-1}(0) = \emptyset$, $\gamma_P^{\pm}((0,\tau]) \cap \Delta_A^{-1}(0) = \emptyset$ and X, Y do not vanish in $R(\tau)$. The system is locally controllable at P (see [52]), and we can cover a neighborhood of P with bang–bang trajectories η having at most one switching, satisfying the PMP and with $T(\eta) \le \tau$. It is well known that these trajectories form an optimal synthesis covering $R(\tau)$ (see for instance [37]). These are the only time optimal trajectories starting from P. Consider a trajectory γ as in the statement and let $Q \doteq \gamma(\tau)$. Assume that $Y(P)$ points into Ω^+, the other case being similar. Consider the one parameter family γ_μ of bang–bang trajectories defined as follows. Let τ' be the first time of intersection of γ_Q^+ with $\partial R(\tau)$ for negative times, and define $\mu' = \min\{\tau, -\tau'\}$. Let $Q_\mu \doteq \gamma_Q^+(-\mu)$, $\mu \in [0, \mu']$, and define η_μ as the time optimal trajectory steering P to Q_μ. We define γ_μ as the trajectory obtained following η_μ up to the point Q_μ, and then $\gamma_{Q_\mu}^+$

up to the point Q. By the local controllability at the point P, the map $\mu \mapsto T(\gamma_\mu)$ is continuous; moreover $T(\gamma_0) \leq T(\gamma) \leq T(\gamma_{\mu'})$. Hence there exists $\hat{\mu} \in [0,\tau]$ such that $T(\gamma) = T(\hat{\gamma})$, with $\hat{\gamma} \doteq \gamma_{\hat{\mu}}$. Notice that $\hat{\gamma}$ has at most two switchings. If $\hat{\gamma}$ has no switchings or only one switching, then it is time optimal and $\gamma \equiv \hat{\gamma}$, so we have done. Suppose now that $\hat{\gamma}$ has two switchings at times σ_1 and σ_2, $\sigma_1 < \sigma_2$. It is not restrictive to assume that $\mathrm{Dom}(\gamma) = [0,\tau]$. Let us define the time σ_3 either as the last positive intersection time (if any) of $\hat{\gamma}$ with the x_1-axis, or set $\sigma_3 = \sigma_2$ if there are no such intersections. By a comparison argument as in the proof of Lemma 3.3.3, we have that the inequality in (3.3.3) holds for $t \in [0,\sigma_1] \cup [\sigma_3,\tau]$. Let σ_4 be the first time at which $\pi_1(\hat{\gamma}(t)) = \pi_1(\hat{\gamma}(\sigma_3))$. By construction the inequality in (3.3.3) holds for $t \in [\sigma_4,\sigma_3]$: indeed

$$\pi_1(\hat{\gamma}(t)) \geq \pi_1(\hat{\gamma}(\sigma_4)), \qquad \text{for every } t \in [\sigma_4,\sigma_3],$$

$t \mapsto \pi_1(\gamma(t))$ is increasing, and $\pi_1(\gamma(\sigma_3)) \leq \pi_1(\hat{\gamma}(\sigma_3))$. Notice that $\hat{\gamma}((0,\sigma_4))$ lies in Ω^+. If $\sigma_4 \leq \sigma_1$ we are done. If $\sigma_4 > \sigma_1$, suppose by contradiction that there exist $s \in (\sigma_1,\sigma_4)$ and $\varepsilon > 0$ such that the inequality in (3.3.3) holds for $t \in [0,s]$ and does not hold for $t \in (s,s+\varepsilon]$. But in this case we obtain $\pi_2(\gamma(s)) \geq \pi_2(\hat{\gamma}(s))$, hence by a comparison argument $\pi_1(\gamma(\sigma_4)) > \pi_1(\hat{\gamma}(\sigma_4))$, reaching a contradiction. □

Lemma 3.3.7 *The conclusions of Lemma 3.3.6 hold if we replace Ω^+ by Ω^-.*

Proof. We choose $\tau > 0$ and, given γ as in the statement, we construct $\hat{\gamma}$ as in the proof of Lemma 3.3.6. We assume again that $Y(P)$ points into Ω^+, and that $\hat{\gamma}$ has two switchings at times σ_1 and σ_2. Let σ_3 be the last time for which $\pi_1(\hat{\gamma}(t)) \geq \pi_1(P)$. With a comparison argument as in the proof of Lemma 3.3.3, we obtain that the inequality in (3.3.3) holds for $t \in [\sigma_2,\tau]$. Since the map $t \mapsto \pi_1(\gamma(t))$ is monotone decreasing, we have that $\pi_1(\gamma(t)) \leq \pi_1(P)$ for every $t \in [0,\tau]$. Hence the inequality in (3.3.3) holds for $t \in [0,\sigma_3]$. If $\sigma_3 \geq \sigma_2$ we are done. Otherwise, if $\sigma_3 < \sigma_2$, suppose by contradiction that there exist $s \in (\sigma_3, \sigma_2)$ and $\varepsilon > 0$ such that the inequality in (3.3.3) holds for $t \in [s, \tau]$ and does not hold for $t \in [s-\varepsilon, s)$. But in this case $\pi_2(\gamma(s)) \leq \pi_2(\hat{\gamma}(s))$, hence $\pi_1(\gamma(\sigma_3)) > \pi_1(\hat{\gamma}(\sigma_3))$, reaching a contradiction. $\square$

Lemma 3.3.8 *Suppose $\pi_2(P) = 0$ and $|g(P)| > |a(P)|$. Then there exists $\tau > 0$ with the following property: for every trajectory γ starting from P, with $T(\gamma) = \tau$, there exists a bang–bang trajectory $\hat{\gamma}$ satisfying (3.3.3).*

Proof. Choose $\tau > 0$ as in the proof of Lemma 3.3.6, and let γ be a trajectory as in the statement. It is not restrictive to assume $\mathrm{Dom}(\gamma) = [0, \tau]$. If γ does not intersect the x_1–axis for positive times, we can apply Lemma 3.3.6 or Lemma 3.3.7. Otherwise let $t_1 \leq \tau$ be the last intersection time of γ with the x_1–axis for positive times. Clearly we can apply Lemma 3.3.6 or 3.3.7 to $\gamma|[t_1, \tau]$. Hence from now on we shall consider $\gamma|[0, t_1]$. Assume that $Y(P)$ points into Ω^+, define $Q \doteq \gamma(t_1)$ and construct $\hat{\gamma}$ as in the proof of Lemma 3.3.6. We can assume again that $\hat{\gamma}$ is a bang–bang trajectory with two switchings at times σ_1 and σ_2. Let σ_3 be the first positive time of intersection of $\hat{\gamma}$ with the x_1–axis. Notice that $\hat{\gamma}([0, \sigma_3]) \subset \overline{\Omega}^+$, $\hat{\gamma}([\sigma_3, t_1]) \subset \overline{\Omega}^-$. Suppose, by contradiction, that there exist $s \in [0, \sigma_1]$, $\varepsilon > 0$ such that

$$\pi_1(\gamma(t)) \leq \pi_1(\hat{\gamma}(t)), \qquad \text{for every } t \in [0, s],$$
$$\pi_1(\gamma(t)) > \pi_1(\hat{\gamma}(t)), \qquad \text{for every } t \in (s, s+\varepsilon].$$

By these conditions and by comparing the two trajectories we have that $\pi_2(\gamma(s)) > \pi_2(\hat{\gamma}(s))$. Consider the trajectory η which follows γ up to time s and then the X–trajectory up to the intersection point R with γ_P^+. By a comparison argument we have that, to steer P to R, η takes a time strictly smaller than γ_P^+, contradicting the optimality of γ_P^+.

Suppose now that there exist s, ε as above, but now with $s \in [\sigma_1, \sigma_3]$. We have again $\pi_2(\gamma(s)) > \pi_2(\hat{\gamma}(s))$. Let $P_1 \doteq \hat{\gamma}(\sigma_1)$, and consider the trajectory $\gamma_{P_1}^-$. There exists a time $s_1 < s$ in which γ intersects $\gamma_{P_1}^-$ and $\gamma(s_1) \in \Omega^+$. Let η be the trajectory obtained following γ up to time s_1 and then the time optimal trajectory that steers $\gamma(s_1)$ to Q. If $\gamma(s_1) \in \mathrm{Range}(\hat{\gamma})$, then $T(\eta) \geq T(\hat{\gamma})$, and the equality holds if and only if $\eta = \hat{\gamma}$. Indeed, in this case, $\hat{\gamma}$ is time optimal to steer P to $\gamma(s_1)$ and to steer $\gamma(s_1)$ to Q. Otherwise, $\eta(t) = P_1$ for some $t > \sigma_1$, and again $T(\eta) \geq T(\hat{\gamma})$, the equality holding if and only if $\eta = \hat{\gamma}$. On the other hand $T(\eta) \leq T(\gamma) = T(\hat{\gamma})$ by construction. Hence $T(\eta) = T(\hat{\gamma})$, and then $\eta = \gamma = \hat{\gamma}$, which contradicts the definition of s. In the same way, but with reversed time, we can argue for $t \in [\sigma_3, t_1]$. □

In Lemmas 3.3.6, 3.3.7 and 3.3.8 we have considered trajectories starting from points of the x_1-axis, while in Lemmas 3.3.9, 3.3.10 and 3.3.11 we shall consider trajectories arriving to points of this kind.

Lemma 3.3.9 *Suppose $\pi_2(P) = 0$ and $|g(P)| > |a(P)|$. Then there exists $\tau > 0$ with the following property. If γ is a trajectory steering a point $Q \in \Omega^+$ to P in time τ and lying in Ω^+ for $t < \tau$, then there exists a bang–bang trajectory $\hat{\gamma}$ satisfying (3.3.3).*

Proof. Let $R'(t)$ be the reachable set for reversed time from P, that is the set of points S for which there exists a trajectory steering S to P in time less than or equal to t. As in the proof of Lemma 3.3.6, we can choose $\tau > 0$ such that $R'(\tau) \cap \Delta_B^{-1}(0) = \emptyset$, and $\gamma_P^{\pm}([-\tau, 0)) \cap \Delta_A^{-1}(0) = \emptyset$. Moreover we can construct an optimal synthesis for the problem of reaching P in minimum time that covers $R'(\tau)$, which is formed of bang–bang trajectories with at most one switching. Consider a trajectory γ satisfying the required assumptions and assume that $Y(P)$ points into Ω^+. We construct a one–parameter family γ_μ of bang–bang trajectories defined as follows. Let τ' be the first time of intersection of γ_Q^+ with $\partial R'(\tau)$ for positive times, and define $\mu' = \min\{\tau, \tau'\}$. Given $\mu \in [0, \mu']$, let $Q_\mu \doteq \gamma_Q^+(\mu)$, and define η_μ as the time optimal trajectory which steers Q_μ to P. Define γ_μ as the trajectory obtained following γ_Q^+ up to the point Q_μ and then η_μ up to the point P. By the local controllability at the point P, the map $\mu \mapsto T(\gamma_\mu)$ is continuous and $T(\gamma_0) \leq T(\gamma) \leq T(\gamma_{\mu'})$, hence there exists $\hat{\mu}$ such that $T(\gamma) = T(\hat{\gamma})$, with $\hat{\gamma} \doteq \gamma_{\hat{\mu}}$. We can assume $\mathrm{Dom}(\gamma) = \mathrm{Dom}(\hat{\gamma}) = [0, \tau]$. We have that $\pi_1(\gamma(t)) \leq \pi_1(P)$ for every $t \in [0, \tau]$, so that we are in a case similar to the one of Lemma 3.3.7. Reasoning in the same way we obtain the conclusion. □

With the same arguments we obtain the following lemmas.

Lemma 3.3.10 *Suppose $\pi_2(P) = 0$ and $|g(P)| > |a(P)|$. Then there exists $\tau > 0$ with the following property. If γ is a trajectory steering a point $Q \in \Omega^-$ to P in time τ and lying in Ω^- for $t < \tau$, then there exists a bang–bang trajectory $\hat{\gamma}$ satisfying (3.3.3).*

Lemma 3.3.11 *Suppose $\pi_2(P) = 0$ and $|g(P)| > |a(P)|$. Then there exists $\tau > 0$ with the following property. For every trajectory γ steering a point Q to P in time τ, there exists a bang–bang trajectory $\hat{\gamma}$ satisfying (3.3.3).*

In Lemmas 3.3.12 and 3.3.14 we shall deal with the singular case in which one of the vector fields X, Y vanishes at the initial or terminal point of the trajectory.

Lemma 3.3.12 *Suppose $\pi_2(P) = 0$, $|g(P)| = |a(P)|$. Then there exists $\tau > 0$ with the following property. For every trajectory γ starting from P, with $T(\gamma) = \tau$, there exists a bang–bang trajectory $\hat{\gamma}$ satisfying (3.3.3).*

Proof. If $g(P) = a(P) = 0$, then $F(P) = G(P) = 0$, and $\gamma(t) = P$ for every $t \in [0, \tau]$. The trajectory is independent of the control, and we can define $\hat{\gamma}$ as the Y–trajectory through P. Suppose now $|g(P)| = |a(P)| \neq 0$. By continuity, there exists a neighborhood $\mathcal{N}_1$ of P such that g does not vanish on $\mathcal{N}_1$.

Let us rewrite the control system in polar coordinates (ρ, θ) centered at P. We obtain:

$$\begin{cases} \rho' = \sin\theta \cdot (\rho\cos\theta - \hat{a}(\rho,\theta) + u \cdot \hat{g}(\rho,\theta)), \\ \theta' = \rho^{-1}\cos\theta \cdot (-\hat{a}(\rho,\theta) + u \cdot \hat{g}(\rho,\theta)) - \sin^2\theta, \end{cases}$$

where $\hat{a}(\rho,\theta) \doteq a(\rho\cos\theta + \pi_1(P), \rho\sin\theta)$, $\hat{g}(\rho,\theta) \doteq g(\rho\cos\theta + \pi_1(P), \rho\sin\theta)$. Along the X–trajectories we have that

$$\theta' = -\rho^{-1}\cos\theta \cdot (\hat{a}(\rho,\theta) + \hat{g}(\rho,\theta)) - \sin^2\theta.$$

Since $a(P)+g(P)=0$, and $a,g\in\mathcal{C}^1(\mathbb{R}^2,\mathbb{R})$, we have that

$$|\hat{a}(\rho,\theta)+\hat{g}(\rho,\theta)|\leq C\cdot\rho+\rho\cdot r(\rho,\theta),$$

with $r(\rho,\theta)\to 0$ as $\rho\to 0^+$, uniformly with respect to θ. Hence there exist a neighborhood $\mathcal{N}_2$ of P and a positive constant M such that $|\theta'|<M$ along every X–trajectory contained in $\mathcal{N}_2$. Let us now define $r(\alpha)$ as the half line through P with positive slope α, that is $r(\alpha)\doteq\{(x_1,x_2)\in\mathbb{R}^2\ :\ x_2=\alpha\cdot(x_1-\pi_1(P)),\ x_1>\pi_1(P)\}$. Since the first component of $Y(P)$ vanishes, for every $\sigma>0$ small enough there exists $\beta(\sigma)>0$ such that for every $t\in(0,\sigma)$ the Y–trajectory γ_P^+ through P lies in the cone with sides $r(\beta(\sigma))$ and $\{x_1=\pi_1(P),\ x_2>0\}$. Clearly $\beta(\sigma)$ can be chosen in such a way that $\beta(\sigma)\to+\infty$ as $\sigma\to 0^+$.

Let $r=r(1)$, and let σ be small enough such that $\beta\doteq\beta(\sigma)>\sqrt{3}=\tan(\pi/3)$. If S is a reachable point from P lying on r, then by Theorem 2.2.2 the time optimal trajectory η_S steering P to S is a $X*Y$–trajectory. Choose $\tau'\in(0,\sigma)$ such that the reachable set $R(\tau')$ from P in time τ' is contained in $\mathcal{N}_1\cap\mathcal{N}_2$. Since $|\theta'|<M$ in $R(\tau')$ along the X–trajectories, we have that, for every $S\in r\cap R(\tau')$

$$T(\eta_S)>(\arctan\beta-\arctan\alpha)/M>\tau''\doteq\frac{\pi}{12\,M}.$$

Let us define τ as the minimum between τ' and τ''. It is clear that $R(\tau)$ does not intersect r: indeed every trajectory takes a time strictly greater than τ to steer P to a point of r.

Consider now a trajectory γ as in the statement. Define $Q\doteq\gamma(\tau)$ and notice that Q lies in the region above r. Moreover the time optimal trajectory which steers P to Q is a $X*Y$–trajectory. Consider the one parameter family of bang–bang trajectories γ_μ constructed in the following way. Given $\mu>0$, let $Q_\mu\doteq\gamma_Q^+(-\mu)$, and let μ' be the first positive μ such that $Q_{\mu'}\in r$. Let I be the set of all $\mu\in[0,\mu']$ such that the trajectory $\gamma_{Q_\mu}^-$ intersects γ_P^+ for negative times in a point R_μ. We have that either $I\equiv[0,\mu']$, or $I=[0,\mu'')$, with $\mu''\leq\mu'$.

For $\mu \in I$, let us define γ_μ as the trajectory starting from P that follows γ_P^+ up to the point R_μ, then $\gamma_{R_\mu}^-$ up to the point Q_μ, and finally $\gamma_{Q_\mu}^+$ up to the point Q. By construction the map $\mu \mapsto T(\gamma_\mu)$ is continuous on I. Since γ_0 is time optimal, we have that $T(\gamma_0) \leq T(\gamma)$. If $I = [0, \mu']$, then $T(\gamma_{\mu'}) \geq \tau = T(\gamma)$: indeed $\gamma_{\mu'}$ intersects r in $Q_{\mu'}$. If $I = [0, \mu'')$, then necessarily $\gamma_{Q_{\mu''}}^-(t) \to P$ as $t \to -\infty$, hence $T(\gamma_\mu) \to +\infty$ as $\mu \to \mu''$. In both cases there exists $\hat{\mu} \in I$ such that $T(\gamma) = T(\hat{\gamma})$, with $\hat{\gamma} \doteq \gamma_{\hat{\mu}}$. By construction $\hat{\gamma}$ is bang–bang with at most two switchings, and (3.3.3) can be proved with the same comparison argument of Lemma 3.3.3. □

Remark 3.3.13 If we fix a compact set $K \subset I\!R^2$, then the time σ in the proof of Lemma 3.3.12 can be chosen uniformly positive in $P \in K$.

In the same way we can prove the following

Lemma 3.3.14 *Suppose $\pi_2(P) = 0$, $|g(P)| = |a(P)|$. Then there exists $\tau > 0$ with the following property: for every trajectory γ with terminal point P, and with $T(\gamma) = \tau$, there exists a bang–bang trajectory $\hat{\gamma}$ satisfying (3.3.3).*

We can now prove the main result of the chapter.

Proof of Theorem 3.3.1. We rewrite the problem (P) as a control system in $I\!R^2$

$$(3.3.4) \qquad \begin{cases} x_1' = x_2\,, \\ x_2' = -a(x_1, x_2) + u\, g(x_1, x_2)\,. \end{cases}$$

Let $\gamma(t) \doteq (x(t), x'(t))$ be the trajectory of this system corresponding to x. We shall construct a bang–bang trajectory $\hat{\gamma}$ by making a concatenation of local bang–bang variations of γ.

Let $R > 0$ be such that $\text{Range}(\gamma) \subset B(0, R)$. Consider a connected component $\mathcal{A}$ of $\mathcal{Z}_R$, and assume that $\gamma(\bar{t}) \in \mathcal{A}$ for some $\bar{t} \in [a, b]$. If $F(\gamma(\bar{t})) = 0$, then every trajectory through $\gamma(\bar{t})$ is a constant trajectory, hence there is nothing to prove. Otherwise let $\mathcal{A}'$ be the connected component of $\mathcal{A} \setminus \{x \in I\!R^2 \,:\, F(x) = 0\}$ which contains $\gamma(\bar{t})$. Since the vector field F is tangent to

$\mathcal{A}$ in every point of $\mathcal{A}$, there exists a trajectory $\eta : I \to I\!R^2$ of (3.3.4), with I unbounded if $\mathcal{A}' \neq \mathcal{A}$, such that Range$(\eta) = \mathcal{A}'$, and $\eta(\bar{t}) = \gamma(\bar{t})$. Moreover, the control has no effect in $\mathcal{A}$, hence $\gamma \equiv \eta$ on $I \cap [a,b]$.

Consider the set $S_1 \doteq \{t \in [a,b] \ : \ g(\gamma(t)) = 0\}$. From (i) and (ii) we have that S_1 is a finite union of isolated points and closed intervals, say

$$S_1 = \{t_1, \ldots, t_n\} \cup [r_1, s_1] \cup \cdots \cup [r_m, s_m].$$

For every t_i, $i = 1, \ldots, n$, there exists $\tau_i > 0$ such that Lemma 3.3.4 can be applied to $\gamma|[t_i - \tau_i, t_i]$, and to $\gamma|[t_i, t_i + \tau_i]$. In this way we construct a bang–bang trajectory $\hat{\gamma}_i$ defined on $I_i \doteq [t_i - \tau_i, t_i + \tau_i]$. Define $\hat{\gamma}(t) \doteq \hat{\gamma}_i(t)$ for $t \in I_i$. For every r_j, s_j, $j = 1, \ldots, m$, there exists $\rho_j > 0$ such that we can apply Lemma 3.3.4 to $\gamma|[r_j - \rho_j, r_j]$ and to $\gamma|[s_j, s_j + \rho_j]$. Notice that every trajectory η with Range$(\eta) \subset S_1$ is bang–bang because it corresponds, for example, to the constant control $u \equiv +1$. Indeed, on S_1, the control has no effect. Then we can construct a bang–bang trajectory $\hat{\gamma}_j$ on $J_j \doteq [r_j - \rho_j, s_j + \rho_j]$, $j = 1, \ldots, m$, by applying Lemma 3.3.4 and choosing control $u \equiv 1$ for $t \in [r_j, s_j]$. Define $\hat{\gamma}(t) \doteq \hat{\gamma}_j(t)$ for $t \in J_j$.

Consider now the compact set

$$S_2 \doteq \left\{ t \in [a,b] \backslash \big(\bigcup_i \operatorname{int} I_i \cup \bigcup_j \operatorname{int} J_j \big) \ : \ \pi_2(\gamma(t)) = 0, \ |g(\gamma(t))| = |a(\gamma(t))| \right\},$$

and let $\kappa \doteq \min\{dist(\gamma(t), \mathcal{M}) \ : \ t \in S_2\}$. Since $S_1 \cap S_2 = \emptyset$, we have that $\kappa > 0$.

Let us define the compact sets

$$D^- \doteq \{x_1 \in I\!R \ : \ (x_1, 0) \in \gamma(S_2), \ X(\gamma((x_1, 0))) = 0\},$$

$$D^+ \doteq \{x_1 \in I\!R \ : \ (x_1, 0) \in \gamma(S_2), \ Y(\gamma((x_1, 0))) = 0\},$$

$$\tilde{D}^\pm \doteq D^\pm \times \bar{B}(0, \kappa/2) \subset I\!R^3.$$

Let us consider the set of singular points D^-, the case D^+ being similar. Let us define the function $h(x_1, x_2) \doteq a(x_1, x_2) + g(x_1, x_2)$. Since h is of class C^1, there exists a continuous function $r\colon \tilde{D}^- \to I\!R$ such that

$$h(x_1 + \delta_1, \delta_2) = h'(x_1, 0) \cdot \delta + |\delta|\, r(x_1; \delta), \quad \text{for every } (x_1, \delta) \in \tilde{D}^-,$$

with $\delta = (\delta_1, \delta_2)$. Since r and h' are continuous on the compact set $\tilde{D}^-$, there exists $C > 0$ such that

$$|h(x_1 + \delta_1, \delta_2)| \le C\,|\delta|, \quad \text{for every } (x_1, \delta) \in \tilde{D}^-.$$

Passing in polar coordinates centered at a point $P = (x_1, 0)$, $x_1 \in D^-$, as in the proof of Lemma 3.3.12, we have that there exists M independent of $x_1 \in D^-$ such that $|\theta'| < M$ along every X–trajectory in $B(P, \kappa/2)$. This estimate, and Remark 3.3.13, imply that we can choose the time τ of Lemma 3.3.12 independent of the point $P \in D^- \times \{0\}$.

Hence, given $t_1, t_2 \in S_2$, belonging to different connected components of S_2, we have that $|t_1 - t_2| \ge \tau$. This implies that ∂S_2 is a finite set.

If $[c, d]$ is a nontrivial connected component of S_2, and u is the control corresponding to γ, then necessarily either $u \equiv +1$ or $u \equiv -1$ on $[c, d]$. Reasoning as for S_1, applying Lemmas 3.3.12 and 3.3.14, we can define $\hat{\gamma}$ on a finite union of closed intervals K_k, $k = 1, \ldots, p$, of positive length σ_k, covering a neighborhood of S_2.

Consider now the compact set

$$S_3 \doteq \left\{ t \in [a, b] \setminus \left(\bigcup_i \operatorname{int} I_i \cup \bigcup_j \operatorname{int} J_j \cup \bigcup_k \operatorname{int} K_k \right) \;:\; \pi_2(\gamma(t)) = 0 \right\}.$$

Since $|g(\gamma(t))| \ne |a(\gamma(t))|$, and $g(\gamma(t)) \ne 0$ for every $t \in S_3$, we have that

$$dist(\gamma(S_3), \mathcal{M}) > 0, \quad dist(\gamma(S_3), \{x \in I\!R^2 \;:\; |g(x)| = |a(x)|\}) > 0.$$

Hence there exists $\varepsilon > 0$ such that for every $t \in S_3$ we can apply, to $\gamma|[t - \varepsilon, t + \varepsilon]$, Lemma 3.3.4 if $|g(\gamma(t))| < |a(\gamma(t))|$, and Lemmas 3.3.8 and 3.3.11 if the opposite inequality holds. Indeed, if $|g(P)| < |a(P)|$, $P \doteq \gamma(t)$, the vector fields $X(P)$ and $Y(P)$ point to the same side with respect to the x_1-axis. We can define $\hat{\gamma}$ on a finite union of intervals L_l, $l = 1, \ldots, q$, of positive length, covering a neighborhood of S_3.

Finally let us define

$$S_4 \doteq [a,b] \setminus \left(\bigcup_i \operatorname{int} I_i \cup \bigcup_j \operatorname{int} J_j \cup \bigcup_k \operatorname{int} K_k \cup \bigcup_l \operatorname{int} L_l \right) \equiv \bigcup_{\nu=1}^{N} [a_\nu, b_\nu].$$

There exists $\delta > 0$ such that for every $t \in [a_\nu, b_\nu]$, $\nu = 1, \ldots, N$, the hypotheses of Lemma 3.3.3 hold for $\gamma|([t-\delta, t+\delta] \cap [a_\nu, b_\nu])$. We can cover S_4 with a finite number of such intervals, and we can construct $\hat{\gamma}$ on S_4 by applying Lemma 3.3.3 on each interval. It is clear that $\hat{\gamma}$ is well defined on $[a,b]$ because, on the boundary of every subinterval on which it was defined, it coincides with γ. Now $z(t) \doteq \pi_1(\hat{\gamma}(t))$, $t \in [a,b]$, is a bang–bang solution of (P) and by construction it verifies all the requirements. If we consider the system (P) with reversed time, we obtain in the same manner the bang–bang trajectory y with the stated properties. □

3.4. Applications

Consider the control problem (P), with $u \in [-1,1]$. Let $c : [a,b] \to I\!R$ be an arbitrary function. For every initial condition (x_0, v_0) we define the families of solutions of (P) in $[a,b]$:

$$\mathcal{S} \doteq \{x : [a,b] \to I\!R \text{ solution of } (P) : (x(a), x'(a)) = (x_0, v_0)\},$$

$$\mathcal{T} \doteq \{x : [a,b] \to I\!R \text{ bang–bang solution of } (P) : (x(a), x'(a)) = (x_0, v_0)\}.$$

We can now define the constrained reachable sets

$$\mathcal{X} \doteq \{(x(b), x'(b)) : x(t) \leq c(t)\, \forall t \in [a,b],\, x \in \mathcal{S}\},$$

$$\mathcal{Y} \doteq \{(x(b), x'(b)) : x(t) \leq c(t)\, \forall t \in [a,b],\, x \in \mathcal{T}\}.$$

If the assumptions of Theorem 3.3.1 are satisfied, we immediately obtain the following

Theorem 3.4.1 *The sets $\mathcal{X}$, $\mathcal{Y}$ coincide; in particular, the reachable set associated to bang-bang constrained solutions $\mathcal{Y}$ is closed.*

As another application, we prove now an existence result for Bolza problems with nonlinear dynamics.

Theorem 3.4.2 *Let*

$$J(x) \doteq \int_a^b [\alpha(t,x) + \beta(t,x) \cdot x' + \gamma(x') \cdot x''] \, dt,$$

with α, $\beta \in C^1([a,b] \times \mathbb{R}, \mathbb{R})$ and $\gamma \in C(\mathbb{R}, \mathbb{R})$. Let us consider the following minimization problem

$$\min \left\{ J(x) \,\middle|\, x \in \mathcal{F},\ x(a) = x_0,\ x'(a) = v_0,\ x(b) = x_1,\ x'(b) = v_1 \right\},$$

where $\mathcal{F}$ is the family of solutions of (3.2.1). Assume that the hypotheses of Theorem 3.3.1 are satisfied, and that

$$\frac{\partial \alpha}{\partial x}(t,x) - \frac{\partial \beta}{\partial t}(t,x) > 0 \qquad (\text{resp. } < 0), \tag{3.4.1}$$

for every $(t,x) \in [a,b] \times \mathbb{R}$. If $x : [a,b] \to \mathbb{R}$ is an optimal solution, then x is bang-bang with a finite number of switchings.

Proof. Let

$$B(t,x) \doteq \int_0^x \beta(t,z) \, dz, \qquad C(x') \doteq \int_0^{x'} \gamma(\xi) \, d\xi.$$

We have that

$$\begin{aligned}
&\int_a^b [\alpha(t,x) + \beta(t,x) \cdot x' + \gamma(x') \cdot x''] dt = \\
&\quad = \int_a^b \left[\alpha(t,x) - \frac{\partial B}{\partial t}(t,x) + \frac{d}{dt}(B(t,x) + C(x')) \right] dt = \\
&\quad = C(v_1) - C(v_0) + B(b,x_1) - B(a,x_0) + \int_a^b \left[\alpha(t,x) - \frac{\partial B}{\partial t}(t,x) \right] dt.
\end{aligned}$$

Clearly

$$\frac{\partial B}{\partial t}(t,x) = \int_0^x \frac{\partial \beta}{\partial t}(t,y)\, dy,$$

and then $\frac{\partial^2 B}{\partial x \partial t} = \frac{\partial \beta}{\partial t}$. Hence, by (3.4.1),

$$\frac{\partial}{\partial x}\left[\alpha(t,x) - \frac{\partial B}{\partial t}(t,x)\right] = \frac{\partial \alpha}{\partial x}(t,x) - \frac{\partial \beta}{\partial t}(t,x) > 0, \qquad (\text{resp. } < 0),$$

that is the map $x \mapsto \delta(t,x) \doteq \alpha(t,x) - \frac{\partial B}{\partial t}(t,x)$ is strictly increasing (resp. decreasing) for every fixed $t \in [a,b]$. By Theorem 3.3.1, there exists a bang–bang solution $y : [a,b] \to I\!R$ satisfying the boundary conditions and such that $y(t) \le x(t)$ (resp. $x(t) \le y(t)$) for every $t \in [a,b]$. By the strictly monotonicity of $x \mapsto \delta(t,x)$, we have that

$$\int_a^b \delta(t, y(t))\, dt \le \int_a^b \delta(t, x(t))\, dt,$$

the equality holding if and only if $x(t) = y(t)$ for a.e. $t \in [a,b]$, and then

$$J(y) = C(v_1) - C(v_0) + B(b, x_1) - B(a, x_0) + \int_a^b \delta(t, y(y))\, dt \le J(x),$$

with $J(y) = J(x)$ if and only if $x(t) = y(t)$ for a.e. $t \in [a,b]$. By the optimality of x, it follows that $x = y$ a.e., reaching the conclusion. □

Chapter 4
An existence result for non-coercive non-convex problems

4.1. Introduction

In this chapter we consider the minimization problem

$$\min \int_0^T [\langle a(t), u(t)\rangle + f(u'(t))]\, dt, \tag{4.1.1}$$

in the class $\mathcal{AC}$ of the absolutely continuous functions from $[0,T]$ into $\mathbb{R}^m$ satisfying the boundary conditions

$$u(0) = u_0, \qquad u(T) = u_1. \tag{4.1.2}$$

Here $a: [0,T] \to \mathbb{R}^m$ is a continuous function.

It is well known (see, for example, [23]) that this problem has a solution if f satisfies the so called growth condition of Tonelli:

$$f(\xi) \geq \psi(|\xi|), \quad \text{for every } \xi \in \mathbb{R}^m, \tag{4.1.3}$$

with $\psi: [0, +\infty) \to \mathbb{R}$ superlinear, that is

$$\lim_{s \to +\infty} \frac{\psi(s)}{s} = +\infty. \tag{4.1.4}$$

Recently, in [18], it was introduced a new class $\mathcal{G}$ of convex functions, satisfying a growth condition strictly weaker than superlinearity, and for which the problem

$$\min \left\{ \int_0^T [g(u(t)) + f(u'(t))]\ dt \colon u \in \mathcal{AC},\ u(0) = u_0,\ u(T) = u_1 \right\},$$

has a solution for every non-negative, continuous function g. This class $\mathcal{G}$ consists of all strictly convex functions $f \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R})$ satisfying the "energy condition"

$$\lim_{|x| \to +\infty} [f(x) - \langle x \,,\, \nabla f(x) \rangle] = -\infty \,. \tag{4.1.5}$$

The technique used in this chapter is completely different from the one used in [18], and it is based on an existence result due to Olech ([43]), which turns to be a corollary of a Lyapunov–type theorem on the range of nonatomic, finite, vector-valued measures. We consider the class $\mathcal{F}$ of all lower semicontinuous functions $f: \mathbb{R}^m \to \mathbb{R}$, such that the convexification f^{**} satisfies the growth condition:

$$\lim_{n \to +\infty} [f^{**}(x_n) - \langle x_n \,,\, \nabla f^{**}(x_n) \rangle] = -\infty \,,$$

for every sequence $(x_n) \subset \mathbb{R}^m$ of points of differentiability of f^{**} such that $\lim_n |x_n| = +\infty$. We prove that, for every $f \in \mathcal{F}$, the minimization problem (4.1.1)-(4.1.2) has a solution.

For related results on non-coercive problems, we mention the papers [3], [6], [7], [26]. In connection with problems without convexity assumptions, we cite [16] and [41] in the case of calculus of variations, and the paper [47] for control theory.

4.2. Preliminaries

We recall that, by Carathéodory's theorem, we have the following characterization of the convex hull of a set $A \subset \mathbb{R}^m$:

$$\operatorname{co} A = \left\{ x \in \mathbb{R}^m \;\middle|\; x = \sum_{i=1}^{m+1} \lambda_i x_i ,\ \tilde{\lambda} \in E_{m+1},\ x_i \in A\ \forall i \right\} , \tag{4.2.1}$$

where $\tilde{\lambda} \doteq (\lambda_1, \dots \lambda_{m+1})$, and E_{m+1} denotes the standard simplex:

$$E_{m+1} \doteq \left\{ (\lambda_1, \dots, \lambda_{m+1}) \in \mathbb{R}^{m+1} \;\middle|\; \lambda_i \geq 0\ \forall i,\ \sum_{i=1}^{m+1} \lambda_i = 1 \right\} .$$

A multifunction $F: \Omega \to 2^{\mathbb{R}^m} \setminus \emptyset$ is upper semicontinuous (u.s.c.) if, for every $x_0 \in \Omega$ and every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$F(x) \subset F(x_0) + B_\varepsilon, \quad \text{for every } x \in B_\delta(x_0) \cap \Omega.$$

We shall say that F is monotone if, for every $x, y \in \Omega$,

$$\langle p_x - p_y \,,\, x - y \rangle \geq 0, \quad \text{for every } p_x \in F(x),\ p_y \in F(y).$$

Given a function $\psi : \mathbb{R}^m \to \mathbb{R}$, we will denote by ψ^{**} its convexification, that is the largest convex function satisfying $\psi^{**} \leq \psi$.

Let $\psi: \mathbb{R}^m \to \mathbb{R}$ be a convex function. For every $x \in \mathbb{R}^m$ we define the subdifferential of ψ at x as the set

$$\partial\psi(x) \doteq \left\{ p \in \mathbb{R}^m \,\middle|\, \psi(y) \geq \psi(x) + \langle p \,,\, y - x \rangle,\ \forall y \in \mathbb{R}^m \right\}.$$

We collect here some properties of the subdifferential; for a proof see, for example, [24] or [35]. We denote by $\mathcal{D}(\psi)$ the set of points of differentiability of ψ.

Proposition 4.2.1 *Let $\psi: \mathbb{R}^m \to \mathbb{R}$ be a convex function. Then:*

(i) $\partial\psi(x)$ is a nonempty, convex, compact subset of $\mathbb{R}^m$ for every $x \in \mathbb{R}^m$;

(ii) $\partial\psi$ is a monotone u.s.c. multifunction;

(iii) ψ is Lipschitz continuous, it is differentiable at almost every point $x \in \mathbb{R}^m$, and $\partial\psi(x) = \{\nabla\psi(x)\}$ for every $x \in \mathcal{D}(\psi)$;

(iv) for every $x \in \mathbb{R}^m$, we have that

$$\partial\psi(x) = \operatorname{co}\left\{ \lim_{k \to +\infty} \nabla\psi(x_k) \,\middle|\, x_k \in \mathcal{D}(\psi)\ \forall k \in \mathbb{N},\ \lim_{k \to +\infty} x_k = x \right\}.$$

Remark 4.2.2 By property (iv) and the characterization (4.2.1) of the convex hull, we have that, if $p \in \partial\psi(x)$, there exist $m+1$ sequences $(x_k^j)_k \subset \mathcal{D}(\psi)$, $j = 1, \ldots, m+1$, and $\tilde{\lambda} \in E_{m+1}$, such that:

$$p = \sum_{j=1}^{m+1} \lambda_j p_j, \qquad p_j = \lim_{k \to +\infty} \nabla\psi(x_k^j), \quad \lim_{k \to +\infty} x_k^j = x, \quad j = 1, \ldots, m+1.$$

This implies that, for every $\varepsilon > 0$ and every $\delta > 0$, there exist $y_1, \dots, y_{m+1} \in \mathcal{D}(\psi) \cap B_\delta(x)$, such that

$$|\nabla\psi(y_j) - p_j| < \varepsilon, \quad \text{for every } j = 1, \dots, m+1.$$

Now we will prove some technical lemmas that will be used in the subsequent section.

Lemma 4.2.3 *Let $D \doteq \overline{B}_R(0) \subset I\!R^m$, $F\colon D \to 2^{I\!R^m}\setminus\emptyset$ u.s.c., with compact convex values. Assume that, for every $x \in \partial D$,*

$$\langle p\,,\, x\rangle \geq 0, \quad \text{for every } p \in F(x).$$

Then there exists $y \in D$ such that $0 \in F(y)$.

Proof. Let us define the multifunction $G \doteq I - F$. Clearly, G is u.s.c. with compact convex values.

Moreover, $\lambda x \notin G(x)$ for every $x \in \partial D$ and every $\lambda > 1$. Indeed, suppose by contradiction that there exist $x \in \partial D$, $p \in F(x)$ and $\lambda > 1$ such that $\lambda x = x - p$. We have that

$$0 \leq \langle p\,,\, x\rangle = -(\lambda - 1)R^2 < 0,$$

reaching a contradiction.

Hence G has a fixed point (see [34], Theorem 11.6), that is there exists $y \in D$ such that $y \in G(y)$. But this condition is equivalent to $0 \in F(y)$. □

Lemma 4.2.4 *Let D and F be as in the previous lemma. If $0 \notin F(D)$, then there exist $y \in \partial D$ and $q \in F(y)$ such that $\langle q\,,\, y\rangle \leq 0$.*

Proof. Suppose that, for every $x \in \partial D$, $\langle p\,,\, x\rangle > 0$, for every $p \in F(x)$. Then Lemma 4.2.3 implies that $0 \in F(D)$, giving a contradiction. □

4.3. The main result

We start this section by stating an existence result due to Olech ([43]).

Theorem 4.3.1 *Assume that $f\colon I\!R^m \to I\!R$ is a lower semicontinuous function. If the set*

$$H \doteq \left\{ p \in I\!R^m \;\Big|\; \sup_{x \in I\!R^m} [\langle p\,,\,x\rangle - f(x)] < +\infty \right\} \tag{4.3.1}$$

is open, then the minimization problem (4.1.1)-(4.1.2) has a solution.

This result is a consequence of a generalization of the Lyapunov theorem (see, for example, [23] for a treatment of this subject).

It is easy to verify that, if f satisfies the superlinearity condition (4.1.3)-(4.1.4), then $H \equiv I\!R^m$. Indeed, for every $p \in I\!R^m$ there exists $R > 0$ such that

$$\frac{f(x)}{|x|} \geq \frac{\psi(|x|)}{|x|} \geq |p|, \quad \text{for every } |x| \geq R. \tag{4.3.2}$$

If we set $M \doteq \max\{|p| \cdot |x| - f(x)\colon |x| \leq R\}$, by (4.3.2) we have that:

$$\langle p\,,\,x\rangle - f(x) \leq |p| \cdot |x| - f(x) \leq \max\{0, M\},$$

for every $x \in I\!R^m$, and then $p \in H$.

We will show that the family of functions for which the set H is open includes the class $\mathcal{F}$ of all lower semicontinuous functions $f : I\!R^m \to I\!R$ satisfying

$$\lim_{|x| \to +\infty} E_f(x) = -\infty\,, \tag{4.3.3}$$

for every selection $E_f\colon I\!R^m \to I\!R$ of the multifunction

$$\mathcal{E}_f(x) \doteq \{f^{**}(x) - \langle p\,,\,x\rangle \,|\, p \in \partial f^{**}(x)\}. \tag{4.3.4}$$

Clearly, $f \in \mathcal{F}$ if and only if $f^{**} \in \mathcal{F}$. Moreover, by property (iii) in Proposition 4.2.1, every selection E_f coincides for almost every $x \in I\!R^m$ with $f^{**}(x) - \langle \nabla f^{**}(x)\,,\,x\rangle$, and then it is measurable.

Before proving the main theorem, we give a characterization of the class $\mathcal{F}$ that involves only the points of differentiability of the functions. In particular, this implies that the property (4.3.3) is independent of the selection.

Proposition 4.3.2 *Let $f\colon \mathbb{R}^m \to \mathbb{R}$ be a convex function. Then $f \in \mathcal{F}$ if and only if the following property holds:*

(P) $\lim_{n\to+\infty}[f(x_n) - \langle \nabla f(x_n), x_n\rangle] = -\infty$, for every sequence $(x_n) \subset \mathcal{D}(f)$ such that $|x_n| \to +\infty$.

Proof. Clearly, if $f \in \mathcal{F}$ then f satisfies (P). Suppose now that (P) holds, and prove that $f \in \mathcal{F}$.

Let E be a selection of $\mathcal{E}_f$. Let (x^n) be a sequence of points of $\mathbb{R}^m$ such that $|x^n| \to +\infty$ for $n \to +\infty$, and let $p^n \in \partial f(x^n)$ be such that $E(x^n) = f(x^n) - \langle p^n , x^n\rangle$, for every $n \in \mathbb{N}$.

Fix $n \in \mathbb{N}$. By Remark 4.2.2, there exist $p_j^n \in \partial f(x^n)$, $y_j^n \in \mathcal{D}(f) \cap B_1(x^n)$, $j \in J \doteq \{1, \ldots, m+1\}$, and $\tilde{\lambda}^n \in E_{m+1}$ such that

$$(4.3.5) \qquad p^n = \sum_{j=1}^{m+1} \lambda_j^n p_j^n, \qquad \left|\nabla f(y_j^n) - p_j^n\right| < \frac{1}{|x^n|+1}, \quad \text{for every } j \in J.$$

The last inequality implies that

$$(4.3.6) \qquad \left|\langle \nabla f(y_j^n) - p_j^n , y_j^n\rangle\right| < \frac{|y_j^n|}{|x^n|+1} < 1, \quad \text{for every } j \in J.$$

By the convexity of f we have that

$$(4.3.7) \qquad f(x^n) - f(y_j^n) \le \langle p_j^n , x^n - y_j^n\rangle, \quad \text{for every } j \in J.$$

Using (4.3.6) and (4.3.7), for every $j \in J$ we obtain

$$(4.3.8) \qquad \begin{aligned} f(x^n) - \langle p_j^n , x^n\rangle =& [f(x^n) - f(y_j^n)] + f(y_j^n) - \langle p_j^n , x^n - y_j^n\rangle + \\ &+ \langle \nabla f(y_j^n) - p_j^n , y_j^n\rangle - \langle \nabla f(y_j^n), y_j^n\rangle \le \\ \le& E(y_j^n) + \langle \nabla f(y_j^n) - p_j^n , y_j^n\rangle \le E(y_j^n) + 1. \end{aligned}$$

Recalling (4.3.5), and using the estimate (4.3.8), one obtains

$$
\tag{4.3.9}
\begin{aligned}
E(x^n) = f(x^n) - \langle p^n , x^n \rangle &= \sum_{j=1}^{m+1} \lambda_j^n [f(x^n) - \langle p_j^n , x^n \rangle] \leq \\
&\leq \sum_{j=1}^{m+1} \lambda_j^n E(y_j^n) + 1 \leq \mu^n + 1,
\end{aligned}
$$

where $\mu^n \doteq \max\{E(y_j^n) : j \in J\}$.

For every fixed $j \in J$, we have that $y_j^n \in B_1(x^n)$ for every $n \in I\!N$, hence $|y_j^n| \to +\infty$, as $n \to +\infty$. Moreover, $y_j^n \in \mathcal{D}(f)$ for every n and j, hence property (P) implies that $\mu^n \to -\infty$ for $n \to +\infty$. By the estimate (4.3.9) we have that

$$
\tag{4.3.10}
\lim_{n \to +\infty} E(x^n) = -\infty \,.
$$

Since (4.3.10) is true for every diverging sequence $(x^n) \subset I\!R^m$, one has

$$
\tag{4.3.11}
\lim_{|x| \to +\infty} E(x) = -\infty \,.
$$

But now we have that (4.3.11) holds for every selection E of $\mathcal{E}_f$, hence $f \in \mathcal{F}$.

□

Now we will prove two lemmas concerning the structure of the set H. These results, together with Theorem 4.3.1, will imply an existence result for the problem (4.1.1)-(4.1.2) with $f \in \mathcal{F}$.

Lemma 4.3.3 *If $f \in \mathcal{F}$ is convex, then the set H coincides with the set $S \doteq \partial f(I\!R^m)$. Moreover, these two sets are open.*

Proof. Let $p \in S$, and let $y \in I\!R^m$ be such that $p \in \partial f(y)$. By the convexity of f we have that, for every $x \in I\!R^m$,

$$
\langle p , x \rangle - f(x) \leq \langle p , y \rangle - f(y),
$$

and then $p \in H$.

Now let $p \notin S$. For every $R > 0$ we can apply Lemma 4.2.4 to the multifunction $F \doteq \partial f - p$, obtaining $y \in \mathbb{R}^m$, $|y| = R$, and $q \in \partial f(y)$ such that $\langle q - p\,,\, y\rangle \leq 0$.

Hence, we can construct a sequence (x_n, p_n), $n \in \mathbb{N}$, such that $\lim_n |x_n| = +\infty$ and

$$p_n \in \partial f(x_n), \quad \langle p_n - p\,,\, x_n\rangle \leq 0, \quad \text{for every } n \in \mathbb{N}.$$

Let E be a selection of the multifunction $\mathcal{E}_f$, defined in (4.3.4), such that $E(x_n) = f(x_n) - \langle p_n\,,\, x_n\rangle$, for every $n \in \mathbb{N}$. We have that

$$\langle p\,,\, x_n\rangle - f(x_n) = \langle p - p_n\,,\, x_n\rangle - E(x_n) \geq -E(x_n) \to +\infty, \quad n \to +\infty,$$

and then $p \notin H$.

It remains to prove that S is open. We consider a sequence $p_n \notin S$, $n \in \mathbb{N}$, $p_n \to p$, and we will prove that $p \notin S$. Clearly, we can assume that $p_n \neq p$ for every $n \in \mathbb{N}$, otherwise we are done. By Lemma 4.2.4 applied to the multifunction $F \doteq \partial f - p_n$, we have that, for every $n \in \mathbb{N}$, there exist $x_n \in \mathbb{R}^m$ and $q_n \in \partial f(x_n)$ such that

$$|x_n| = \frac{1}{|p - p_n|}, \qquad \langle q_n - p_n\,,\, x_n\rangle \leq 0.$$

Now, let $\tilde{E}$ be a selection of $\mathcal{E}_f$ such that $\tilde{E}(x_n) = f(x_n) - \langle q_n\,,\, x_n\rangle$, for every $n \in \mathbb{N}$. Since $\lim_n |x_n| = +\infty$, we have that the r.h.s. of the following inequality

$$\langle p\,,\, x_n\rangle - f(x_n) = \langle p - p_n\,,\, x_n\rangle + \langle p_n - q_n\,,\, x_n\rangle - \tilde{E}(x_n) \geq -1 - \tilde{E}(x_n)$$

tends to $+\infty$ as $n \to +\infty$, and then $p \notin S$. $\square$

Lemma 4.3.4 *Assume $f\colon \mathbb{R}^m \to \mathbb{R}$ is lower semicontinuous, and let H and H^{**} be respectively the sets defined in (4.3.1) for f and f^{**}. Then $H = H^{**}$; in particular, if $f \in \mathcal{F}$, then H is open.*

Proof. The inclusion $H^{**} \subset H$ is trivial: indeed for every $x \in \mathbb{R}^m$ we have

$$\langle p, x\rangle - f^{**}(x) \geq \langle p, x\rangle - f(x).$$

Now suppose that $p \notin H^{**}$. By the very definition of H^{**}, there exists a sequence $(x^n) \subset \mathbb{R}^m$ such that

$$\lim_{n\to+\infty}[\langle p, x^n\rangle - f^{**}(x^n)] = +\infty\,. \tag{4.3.12}$$

Fix $\varepsilon > 0$. Since, for every $\xi \in \mathbb{R}^m$,

$$f^{**}(\xi) = \inf\left\{\sum_{j=1}^{m+1}\lambda_j f(\xi_j) \,\Big|\, \sum_{j=1}^{m+1}\lambda_j\xi_j = \xi,\ \tilde{\lambda} \in E_{m+1}\right\},$$

then for every $n \in \mathbb{N}$ there exist $x_j^n \in \mathbb{R}^m$, $j = 1, \ldots, m+1$, and $\tilde{\lambda}^n \in E_{m+1}$ such that

$$\sum_{j=1}^{m+1}\lambda_j^n x_j^n = x^n, \quad f^{**}(x^n) > \sum_{j=1}^{m+1}\lambda_j^n f(x_j^n) - \varepsilon.$$

This implies that

$$\langle p, x^n\rangle - f^{**}(x^n) < \sum_{j=1}^{m+1}\lambda_j^n[\langle p, x_j^n\rangle - f(x_j^n)] + \varepsilon. \tag{4.3.13}$$

By (4.3.12) and (4.3.13), there exists at least one index $j \in \{1, \ldots, m+1\}$ such that, up to a subsequence,

$$\lim_{n\to+\infty}[\langle p, x_j^n\rangle - f(x_j^n)] = +\infty\,.$$

Hence $p \notin H$, and then we have the other inclusion $H \subset H^{**}$. □

The following theorem is a direct consequence of Theorem 4.3.1 and Lemma 4.3.4.

Theorem 4.3.5 *If $f \in \mathcal{F}$, then the minimization problem (4.1.1)-(4.1.2) has a solution.*

Chapter 5
Existence and regularity results for non–coercive variational problems

5.1. Introduction

It is well known that, if L is a continuous function, such that $\xi \mapsto L(t,x,\xi)$ is convex and superlinear, then the variational problem

$$(5.1.1)\quad \min\left\{ \int_0^T L(t,u,u')dt \;\middle|\; u \in W^{1,1}([0,T],\mathbb{R}^m),\, u(0)=a,\, u(T)=b \right\},$$

has a solution (see for instance [23]).

In recent years, the possibility of relaxing the convexity or the superlinearity assumption was investigated by many authors.

Some existence results for non–convex coercive problems were obtained in the case $L(t,x,\xi) = g(t,x) + f(t,\xi)$ (see for instance [16], [41], [47] and the references therein). In particular, in [16] it was proved that the convexity assumption on $f(t,\cdot)$ can be replaced by the condition of concavity of $g(t,\cdot)$.

More recently, some techniques were developed in order to treat convex but non–coercive problems. In this case, even if the functionals considered are lower semicontinuous in the weak topology of $W^{1,1}([0,T],\mathbb{R}^m)$, the direct method of the Calculus of Variations can not be applied, due to the lack of compactness of the minimizing sequences.

In [26], it was studied the problem (5.1.1) with L continuous, bounded from below and convex with respect to ξ, the superlinearity being replaced by a weaker condition which permits to construct a relatively compact minimizing sequence, obtained by considering the minima of suitable coercive approximating problems. The main step in the proof of the existence result in [26] was to show that every minimum point of the approximating problems solves a generalized DuBois–Reymond condition, which implies that the minimizing sequence is bounded in the space $W^{1,\infty}([0,T],\mathbb{R}^m)$.

A similar approach was used in [18] for the autonomous problem with Lagrangean $L(t,x,\xi) = g(x) + f(\xi)$, where g is a nonnegative continuous function, and $f \in C^1(\mathbb{R}^m,\mathbb{R})$ is a strictly convex function bounded from below, such that

$$\lim_{|\xi|\to+\infty} [f(\xi) - \langle \nabla f(\xi)\,,\,\xi\rangle] = -\infty. \tag{5.1.2}$$

In that paper, it was proved that, for every rectifiable curve C in $\mathbb{R}^m$ joining a to b there exists a unique solution to the problem (5.1.1) restricted to the class of all absolutely continuous parameterizations $u: I \to \mathbb{R}^m$ of C. Thus, every element u_n of a minimizing sequence can be replaced by the minimum corresponding to the curve parameterized by u_n. It can be shown, still using a DuBois–Reymond condition satisfied by those minima, and by (5.1.2), that this new sequence is bounded in $W^{1,\infty}([0,T],\mathbb{R}^m)$, so that there exists a minimum point for (5.1.1) in this space.

In Chapter 4 (see [28]) both the superlinearity and the convexity assumptions were dropped for Lagrangeans of the form $L(t,x,\xi) = \langle a(t)\,,\,x\rangle + f(\xi)$ where f is a lower semicontinuous function whose convexification f^{**} satisfies (5.1.2) for every diverging sequence of points of differentiability of the Lipschitz continuous function f^{**}. The existence of a minimum is proved by a technique relying only on a Lyapunov type theorem due to Olech (see [43]).

For other results concerning non–coercive problems we mention [3], [6] and [7].

In this chapter we consider non–autonomous problems of the form

$$\min_{u\in W^{1,1}([0,T],\mathbb{R}^m)} \left\{ \int_0^T [g(t,u) + f(t,u')]\,dt \;\middle|\; u(0) = a\,,\ u(T) = b \right\} \tag{5.1.3}$$

with neither coercivity nor convexity assumptions. More precisely, we introduce the class $\mathcal{E}$ of all functions $\psi : [0,T] \times \mathbb{R}^m \to \mathbb{R}$, bounded from below, such that $\psi(\cdot,\xi)$ is Lipschitz continuous for every fixed $\xi \in I\!R^m$, $\psi(t,\cdot)$ is lower semicontinuous and satisfies

$$\lim_{n\to+\infty} [\psi^{**}(t^n,\xi^n) - \langle \nabla\psi^{**}(t^n,\xi^n)\,,\,\xi^n\rangle] = -\infty$$

for every sequence $\{t^n\} \in [0,T]$ and for every choice of points ξ^n of differentiability of $\psi^{**}(t^n,\cdot)$ such that $\lim_n |\xi^n| = +\infty$. We show that, if $f \in \mathcal{E}$ and there exist two constants A and B, $B > 0$ such that $f(t,\xi) \geq -A + B|\xi|$ for every $(t,\xi) \in [0,T] \times I\!R^m$, and $g(t,x)$ is a continuous function, Lipschitz continuous with respect to t, concave with respect to x, satisfying $g(t,x) \geq -\alpha - \beta|x|$ for every $(t,x) \in [0,T] \times I\!R^m$, and for suitable constants α and $0 \leq \beta \leq B/T$, then the problem (5.1.3) has a solution in the space $W^{1,\infty}([0,T], I\!R^m)$. This result is the analogue for a class of non–coercive functionals of the one in [16], but it is not a generalization of that result, due to the additional requirement of the Lipschitz continuity of the Lagrangean with respect to the variable t. On the other hand, this extra regularity assumption allows us to obtain the necessary conditions that, used at an intermediate step, also yield a regularity result for the optimal solution, interesting by itself.

As a first step we prove an existence result for (5.1.3), requiring that f be convex with respect to ξ and dropping the concavity assumption on g. This can be done following [26] and making suitable changes, due to the the fact that the Lagrangean is not bounded from below. The second step, linking the convex to the non–convex case, is based on a result concerning the closure of the convex hull of the epigraph of functions whose convexification is strictly convex at infinity (i.e., the graph of the convexification contains no rays). This result is an extension of the classical theorem valid for superlinear functions (see [35]). We remark that the notion of strict convexity at infinity was also used in [27] in order to study non–coercive problems of the type (5.1.1) with the additional state constraint $\|u\|_{L^\infty} < R$. We shall prove that every function in the class $\mathcal{E}$ is strictly convex at infinity for every fixed t. Hence, by using the previous results and the Lyapunov theorem on the range of non–atomic measures, the existence result for the non–convex problems follows.

The regularity of the solution of (5.1.3) is a consequence of the regularity of the solution to the relaxed problem.

5.2. Preliminaries

Given a function $\psi\colon \mathbb{R}^m \to \overline{\mathbb{R}}$, we shall denote by $\operatorname{dom}(\psi)$ its effective domain, defined as the subset of $\mathbb{R}^m$ $\{\xi \mid \psi(\xi) < +\infty\}$, and by $\operatorname{epi}\psi$ its epigraph, that is the set:

$$\operatorname{epi}\psi \doteq \{(x,a) \in \mathbb{R}^m \times \mathbb{R} \mid \psi(x) \le a\}.$$

If $\psi\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is Lipschitz continuous in a neighborhood of a point ξ, we shall denote by $\partial\psi(\xi)$ the generalized gradient of ψ at ξ, defined by

$$\partial\psi(\xi) \doteq \operatorname{co}\left\{\lim_{i\to+\infty} \nabla\psi(\xi_i) \mid \xi_i \to \xi,\ \xi_i \in \mathcal{D}(\psi)\right\}, \tag{5.2.1}$$

where $\mathcal{D}(\psi)$ denotes the set of points of differentiability of ψ. We recall that a Lipschitz continuous function ψ is almost everywhere differentiable in $\operatorname{int}(\operatorname{dom}(\psi))$.

A function $\psi\colon \mathbb{R}^m \to (-\infty,+\infty]$ is convex if, for every $\xi,\eta \in \mathbb{R}^m$ and for every $\lambda \in [0,1]$, we have $\psi(\lambda\xi + (1-\lambda)\eta) \le \lambda\psi(\xi) + (1-\lambda)\psi(\eta)$. We say that ψ is concave if $-\psi$ is convex.

Given a function $\psi\colon \mathbb{R}^m \to (-\infty,+\infty]$, we shall denote by ψ^* its dual function, defined for every $p \in \mathbb{R}^m$ by

$$\psi^*(p) \doteq \sup_{\xi\in\mathbb{R}^m} \{\langle p\,,\,\xi\rangle - \psi(\xi)\}\,.$$

It is well known that the bidual function ψ^{**} coincides with the convexification of ψ, which is the largest convex function φ satisfying $\varphi \le \psi$.

If $\psi\colon \mathbb{R}^m \to (-\infty,+\infty]$ is convex, then the generalized gradient of ψ coincides in $\operatorname{int}(\operatorname{dom}(\psi))$ with the subgradient of ψ in the sense of convex analysis, defined at every point $\xi \in \operatorname{dom}(\psi)$ by

$$\partial\psi(\xi) \doteq \{p \in \mathbb{R}^m \mid \psi(\eta) \ge \psi(\xi) + \langle p\,,\,\eta - \xi\rangle, \text{ for every } \eta \in \mathbb{R}^m\} \tag{5.2.2}$$

(see [24], Proposition 2.2.7). By definition, we set $\partial\psi(\xi) \doteq \emptyset$ for every $\xi \notin \mathrm{dom}(\psi)$. We recall that, if ψ is differentiable at ξ, then $\partial\psi(\xi) = \{\nabla\psi(\xi)\}$.

In the following proposition we collect some well known properties of the subgradient (see [24] and [35]).

Proposition 5.2.1 *Let $\psi: \mathbb{R}^m \to (-\infty, +\infty]$ be a convex function. Then the following properties hold:*

(i) if ψ is bounded from above in a non–empty open set A, then ψ is locally Lipschitz continuous in A;

(ii) for every $\xi \in \mathbb{R}^m$, the set $\partial\psi(\xi)$ (possibly empty) is convex and closed in $\mathbb{R}^m$;

(iii) if $\xi \in \mathrm{int}(\mathrm{dom}(\psi))$, then $\partial\psi(\xi)$ is a non–empty compact set.

5.3. The closure result

In this section we shall prove a result concerning the closure of the convex hull of the epigraph of functions possibly without superlinear growth.

We recall the notion of strict convexity at infinity, introduced by Clarke and Loewen in [27].

Definition 5.3.1 *A convex function $\psi: \mathbb{R}^m \to \mathbb{R}$ is said to be strictly convex at infinity if its graph contains no rays, that is for every $\nu \in \mathbb{R}^m$, $\nu \neq 0$, and for every $\xi \in \mathbb{R}^m$, the function $\psi_{\nu,\xi}(s) \doteq \psi(s\nu + \xi)$ has the following property: for every $s_0 \in \mathcal{D}(\psi_{\nu,\xi})$ there exists $s_1 \in \mathcal{D}(\psi_{\nu,\xi})$, $s_1 > s_0$, such that $\psi'_{\nu,\xi}(s_1) > \psi'_{\nu,\xi}(s_0)$.*

Remark 5.3.2 It is easy to see that, if $\psi: \mathbb{R}^m \to \mathbb{R}$ is convex, then ψ is strictly convex at infinity if and only if $\partial\psi^*(p)$ is either empty or bounded for every $p \in \mathbb{R}^m$.

Definition 5.3.3 *We shall denote by $\mathcal{G}$ the family of all lower semicontinuous functions $\psi : \mathbb{R}^m \to \mathbb{R}$ such that $\psi^{**} \not\equiv -\infty$ and ψ^{**} is strictly convex at infinity.*

Remark 5.3.4 Clearly, every strictly convex function is strictly convex at infinity. Moreover, every lower semicontinuous superlinear function $\psi: \mathbb{R}^m \to \mathbb{R}$ belongs to $\mathcal{G}$. Indeed, denoting by φ the convexification ψ^{**}, for every fixed ν, $\xi \in \mathbb{R}^m$, $\nu \neq 0$, by (5.2.2) it follows that the inequality $\langle \nabla\varphi(s\nu+\xi)\,,\, s\nu\rangle \geq \varphi(s\nu+\xi) - \varphi(\xi)$ holds for every $s \in \mathcal{D}(\varphi_{\nu,\xi})$. This implies that

$$\varphi'_{\nu,\xi}(s) = \langle \nabla\varphi(s\nu+\xi)\,,\, \nu\rangle \geq \frac{\varphi(s\nu+\xi) - \varphi(\xi)}{s}, \quad \text{for every } s \in \mathcal{D}(\varphi_{\nu,\xi}),\ s > 0.$$

Since ψ is superlinear, the last term tends to $+\infty$ as s goes to $+\infty$.

Lemma 5.3.5 *For every function $\psi \in \mathcal{G}$, satisfying $\psi \geq 0$ and $\psi(0) = 0$, there exist two positive constants C, ρ such that $\psi(\xi) \geq C|\xi|$ for every $|\xi| > \rho$.*

Proof. We can certainly assume that ψ is convex, for if not, we replace ψ by ψ^{**}. We start by proving that ψ is coercive, that is $\psi(\xi) \to +\infty$ as $|\xi| \to +\infty$. Since ψ is convex, the sets $\psi^a \doteq \{\xi \in \mathbb{R}^m \mid \psi(\xi) < a\}$ are convex subsets of $\mathbb{R}^m$ for every $a \geq 0$. By contradiction, suppose that there exists $a > 0$ such that ψ^a is unbounded. Since ψ^a is convex, it contains at least one half line $\{s\nu \mid s \geq 0\}$ for some $\nu \in \mathbb{R}^m$, $\nu \neq 0$. This means that $\psi_{\nu,0}(s) < a$ for every $s \geq 0$. Since $\psi_{\nu,0}$ is an absolutely continuous function, then for every $\tau > 0$ we have

$$0 \leq \psi_{\nu,0}(\tau) - \psi_{\nu,0}(0) = \int_0^\tau \psi'_{\nu,0}(\sigma)\, d\sigma.$$

Hence, there exists $s_0 \in \mathcal{D}(\psi_{\nu,0}) \cap [0,\tau]$ such that $\psi'_{\nu,0}(s_0) \geq 0$. Since ψ is strictly convex at infinity, there exists $s_1 \in \mathcal{D}(\psi_{\nu,0})$, $s_1 > s_0$, such that $\psi'_{\nu,0}(s_1) > 0$. By the convexity of $\psi_{\nu,0}$ it follows that

$$\psi_{\nu,0}(s) \geq \psi_{\nu,0}(s_1) + (s - s_1)\psi'_{\nu,0}(s_1), \quad \text{for every } s \geq 0,$$

and this implies that $\lim_{s\to+\infty} \psi_{\nu,0}(s) = +\infty$, in contradiction with $\psi_{\nu,0} < a$.

Since ψ is coercive, there exist two positive constants ρ, δ such that

$$\psi(\eta) \geq \delta, \quad \text{for all } |\eta| = \rho.$$

If $|\xi| > \rho$, let us define $\lambda \doteq \rho/|\xi|$ and $\eta \doteq \lambda\xi$. By the convexity of ψ, and recalling that $\psi(0) = 0$, we get

$$\psi(\xi) \geq \frac{1}{\lambda}\psi(\eta) = \frac{\psi(\eta)}{\rho}|\xi| \geq \frac{\delta}{\rho}|\xi|,$$

so that we have done by choosing $C = \delta/\rho$. □

We are now in a position to prove the closure result. The proof is based on the fact that, if f belongs to the class $\mathcal{G}$, then for every support hyperplane r of f^{**}, the function $f - r$ belongs to $\mathcal{G}$. Applying the estimate of Lemma 5.3.5 to this function, we can follow the lines of the proof of Lemma IX.3.3 in [35].

Theorem 5.3.6 *For every $f \in \mathcal{G}$ the set* co epi f *is closed.*

Proof. Let $(\xi, a) \in \partial(\text{co epi f})$, where ∂S denotes the boundary of the set S, and let $r(\eta) \doteq \langle c\,,\,\eta\rangle + d$ be an affine function such that the hyperplane $H \doteq \{(\eta, r(\eta))\}$ weakly separates co epi f and the point (ξ, a). Let us define the function

$$\phi(\eta) \doteq f(\eta + \xi) - r(\eta + \xi).$$

We have $\phi^{**}(\eta) = f^{**}(\eta + \xi) - r(\eta + \xi)$, $\phi^{**} \geq 0$, $\phi^{**}(0) = 0$. Moreover, for every $\nu \in I\!R^m$, $\nu \neq 0$, for every $\eta \in I\!R^m$ and for every $s \in \mathcal{D}(f^{**}_{\nu,\xi+\eta})$ we have $(\phi^{**}_{\nu,\eta})'(s) = (f^{**}_{\nu,\xi+\eta})'(s) - \langle c\,,\,\nu\rangle$. Since f^{**} is strictly convex at infinity, then so is ϕ^{**}. By Lemma 5.3.5, there exist two positive constants C, ρ such that

$$\phi^{**}(\eta) \geq C|\eta|, \quad \text{for every } |\eta| \geq \rho. \tag{5.3.1}$$

Notice that $(\xi, a) \in$ co epi f if and only if $(0,0) \in$ co epi ϕ. Moreover, $(\xi, a) \in \partial(\text{co epi f})$ if and only if $(0,0) \in \partial(\text{co epi}\,\phi)$. Hence, to prove the proposition, it suffices to show that $(0,0) \in$ co epi ϕ.

Let $(\xi^n, a^n) \in$ co epi ϕ be such that $\lim_n(\xi^n, a^n) = (0,0)$. By the characterization (4.2.1) of the convex hull, for every n there exist $\tilde{\lambda}^n \in E_{m+2}$ and $(\xi_j^n, a_j^n) \in$ epi ϕ, $j = 1, \ldots, m+2$, such that

$$\sum_{j=1}^{m+2} \lambda_j^n(\xi_j^n, a_j^n) = (\xi^n, a^n).$$

By the very definition of epigraph it follows that

$$a^n = \sum_{j=1}^{m+2} \lambda_j^n a_j^n \geq \sum_{j=1}^{m+2} \lambda_j^n \phi(\xi_j^n). \tag{5.3.2}$$

Moreover, (5.3.2) and the fact that $\phi \geq \phi^{**}$ imply that $a^n \geq \sum_{j=1}^{m+2} \lambda_j^n \phi^{**}(\xi_j^n)$. Since $\phi^{**} \geq 0$, the inequality

$$a^n \geq \lambda_j^n \phi^{**}(\xi_j^n) \tag{5.3.3}$$

holds for every $j = 1, \ldots, m+2$. Let $J \subset \{1, \ldots, m+2\}$ be the set of all j such that $\{|\xi_j^n|\}_n$ is unbounded, and let $I \doteq \{1, \ldots, m+2\} \setminus J$. By passing to a subsequence, we can assume that there exist $\overline{\xi}_j$, $j \in I$, and $\tilde{\lambda} \in E_{m+2}$, such that

$$\begin{aligned}
&\lim_{n \to +\infty} |\xi_j^n| = +\infty, \quad j \in J, \\
&\lim_{n \to +\infty} \xi_j^n = \overline{\xi}_j, \quad j \in I, \\
&\lim_{n \to +\infty} \lambda_j^n = \lambda_j, \quad j \in \{1, \ldots, m+2\}.
\end{aligned}$$

For every $j \in J$, we have $|\xi_j^n| > \rho$ for n large enough, and then from (5.3.1) and (5.3.3) it follows that $a^n \geq C\lambda_j^n |\xi_j^n|$. Since $\lim_n a^n = 0$, we get

$$\lim_{n \to +\infty} \lambda_j^n |\xi_j^n| = 0, \quad j \in J. \tag{5.3.4}$$

From (5.3.4), and recalling that $\lim_n \xi^n = 0$, we deduce that

$$\begin{aligned}
\sum_{j \in I} \lambda_j \overline{\xi}_j &= \lim_{n \to +\infty} \sum_{j \in I} \lambda_j^n \xi_j^n = \lim_{n \to +\infty} \left(\sum_{j=1}^{m+2} \lambda_j^n \xi_j^n - \sum_{j \in J} \lambda_j^n \xi_j^n \right) = \\
&= \lim_{n \to +\infty} \left(\xi^n - \sum_{j \in J} \lambda_j^n \xi_j^n \right) = 0.
\end{aligned} \tag{5.3.5}$$

Moreover, since $\lim_n \lambda_j^n = 0$ for every $j \in J$, we obtain

$$\sum_{j \in I} \lambda_j = \lim_{n \to +\infty} \sum_{j \in I} \lambda_j^n = 1. \tag{5.3.6}$$

Since ϕ is a non–negative lower semicontinuous function, we get

$$0 \le \sum_{j\in I} \lambda_j \phi(\bar{\xi}_j) \le \liminf_{n\to+\infty} \sum_{j\in I} \lambda_j^n \phi(\xi_j^n) \le \liminf_{n\to+\infty} a^n = 0. \tag{5.3.7}$$

There is no loss of generality in assuming that $\lambda_j > 0$ for every $j \in I$, hence (5.3.7) implies that $\phi(\bar{\xi}_j) = 0$ for every $j \in I$, that is $(\bar{\xi}_j, 0) \in \operatorname{epi}\phi$ for every $j \in I$. Thus, by (5.3.5) and (5.3.6), we can conclude that $(0,0)$ belongs to $\operatorname{co}\operatorname{epi}\phi$. □

Now we state two direct consequences of Theorem 5.3.6.

Corollary 5.3.7 *If $f \in \mathcal{G}$, then*

$$f^{**}(\xi) = \min\left\{ \sum_{j=1}^{m+1} \lambda_j f(\xi_j) \;\middle|\; \sum_{j=1}^{m+1} \lambda_j \xi_j = \xi,\ \tilde{\lambda} \in E_{m+1} \right\},$$

for every $\xi \in I\!R^m$.

Proof. See [35], Lemma IX.3.3. □

We recall that a function $f: I \times I\!R^m \to I\!R$ is said to be a normal integrand (see [35]) if $f(t,\cdot)$ is lower semicontinuous for a.e. $t \in I$, and there exists a Borel function $\tilde{f}: I \times I\!R^m \to I\!R$ such that $\tilde{f}(t,\cdot) = f(t,\cdot)$ for a.e. $t \in I$.

Corollary 5.3.8 *Let $f: I \times I\!R^m \to I\!R$ be a normal integrand, and suppose that $f(t,\cdot) \in \mathcal{G}$ for every $t \in I$. Then for any measurable mapping $p: [0,T] \to I\!R^m$, there exist a measurable mapping $\tilde{\lambda}: [0,T] \to E_{m+1}$ and $m+1$ measurable mappings $q_j: [0,T] \to I\!R^m$, such that*

$$\sum_{j=1}^{m+1} \lambda_j(t) q_j(t) = p(t), \quad \sum_{j=1}^{m+1} \lambda_j(t) f(t, q_j(t)) = f^{**}(t, p(t)),$$

for almost all $t \in [0,T]$.

Proof. See [35], Proposition IX.3.1. □

5.4. Existence results for variational problems

In this section we shall show that the existence result proved by Cellina and Colombo in [16] holds even for functions of the class $\mathcal{E}$ defined below. In the following, the convexification and the gradient of a function $\psi(t,\xi)$ are understood with respect to ξ.

Definition 5.4.1 *We shall denote by $\mathcal{E}$ the family of all functions $\psi: I \times \mathbb{R}^m \to \mathbb{R}$, bounded from below, such that $\psi(\cdot,\xi)$ is Lipschitz continuous for every fixed $\xi \in \mathbb{R}^m$, $\psi(t,\cdot)$ is lower semicontinuous for every fixed $t \in I$, and*

$$(5.4.1) \qquad \lim_{R\to+\infty} \sup_{\substack{t\in I\\ |\xi|>R}} \sup\left\{\psi^{**}(t,\xi) - \langle p\,,\, \xi\rangle \;\middle|\; p \in \partial_\xi \psi^{**}(t,\xi)\right\} = -\infty.$$

The following proposition gives a characterization of the family $\mathcal{E}$. The proof is similar to the one of Proposition 4.3.2.

Proposition 5.4.2 *The condition (5.4.1) in Definition 5.4.1 is equivalent to:*

$$(5.4.2) \qquad \lim_{n\to+\infty} [\psi^{**}(t^n,\xi^n) - \langle \nabla\psi^{**}(t^n,\xi^n)\,,\, \xi^n\rangle] = -\infty$$

for every sequence $(t^n,\xi^n) \in I \times \mathbb{R}^m$ such that $\xi^n \in \mathcal{D}(\psi^{**}(t^n,\cdot))$, $\lim_n |\xi^n| = +\infty$.

Proof. We have to prove that (5.4.2) implies (5.4.1), the other implication being trivial. Let us denote by $\chi(R)$ the argument of the limit in (5.4.1), and let $\{R_n\}$ be a diverging sequence. For every fixed $n \in \mathbb{N}$, by definition of supremum, there exists $(t^n,\xi^n,p^n) \in I \times \mathbb{R}^m \times \mathbb{R}^m$, with $p^n \in \partial_\xi\psi^{**}(t^n,\xi^n)$ and $|\xi^n| > R_n$, such that

$$(5.4.3) \qquad \chi(R_n) \le \psi^{**}(t^n,\xi^n) - \langle p^n\,,\, \xi^n\rangle + 1.$$

From (5.2.1) and (4.2.1), there exist $p_j^n \in \partial_\xi\psi^{**}(t^n,\xi^n)$, $\xi_j^n \in \mathcal{D}(\psi^{**}(t^n,\cdot))$, with $|\xi_j^n - \xi^n| < 1$, $j \in J \doteq \{1,\dots,m+1\}$, and $\tilde\lambda^n \in E_{m+1}$, such that

$$p^n = \sum_{j=1}^{m+1} \lambda_j^n p_j^n, \qquad \left|\nabla\psi^{**}(t^n,\xi_j^n) - p_j^n\right| < \frac{1}{|\xi^n|+1}, \quad \text{for every } j \in J.$$

For every $j \in J$, the last inequality and the fact that $|\xi_j^n - \xi^n| < 1$ imply that

$$(5.4.4) \qquad \left|\langle \nabla\psi^{**}(t^n, \xi_j^n) - p_j^n\,,\, \xi_j^n\rangle\right| < \frac{|\xi_j^n|}{|\xi^n|+1} < 1\,.$$

By the convexity of $\psi^{**}(t^n,\cdot)$ we have

$$(5.4.5) \qquad \psi^{**}(t^n,\xi^n) - \psi^{**}(t^n,\xi_j^n) \le \langle p_j^n\,,\, \xi^n - \xi_j^n\rangle, \quad \text{for every } j \in J.$$

Using (5.4.4) and (5.4.5) we obtain

$$(5.4.6) \qquad \psi^{**}(t^n,\xi^n) - \langle p_j^n\,,\, \xi^n\rangle \le \psi^{**}(t^n,\xi_j^n) - \langle \nabla\psi^{**}(t^n,\xi_j^n)\,,\, \xi_j^n\rangle + 1\,.$$

Multiplying (5.4.6) by λ_j^n and summing over j it follows that $\psi^{**}(t^n,\xi^n) - \langle p^n\,,\, \xi^n\rangle \le \mu^n$, where $\mu^n \doteq 1 + \max_j\{\psi^{**}(t^n,\xi_j^n) - \langle \nabla\psi^{**}(t^n,\xi_j^n)\,,\, \xi_j^n\rangle\}$.

Since $\lim_n |\xi_j^n| = +\infty$ for every $j \in J$, (5.4.2) implies that $\lim_n \mu^n = -\infty$. Hence, by (5.4.3), it follows that

$$\lim_{n\to+\infty} \chi(R_n) \le \lim_{n\to+\infty} (\mu^n + 1) = -\infty\,.$$

Since χ is a monotone non–increasing function, (5.4.1) holds. □

Remark 5.4.3 The Definition 5.4.1 agrees with the one given in [18] and in Chapter 4 (see [28]), respectively in the case of convex time–independent smooth functions and non–convex time–independent functions.

Lemma 5.4.4 *If $\psi \in \mathcal{E}$, then $\psi(t,\cdot) \in \mathcal{G}$ for every $t \in I$.*

Proof. Let us fix $t \in I$, and denote by φ the convexification with respect to ξ of $\psi(t,\xi)$. By Lemma 4.3.3, the effective domain $\mathrm{dom}(\varphi^*)$ of φ^* is an open subset of $\mathbb{R}^m$. Hence, by Proposition 5.2.1(iii), $\partial\varphi^*(p)$ is either bounded, if $p \in \mathrm{dom}(\varphi^*)$, or empty, if $p \notin \mathrm{dom}(\varphi^*)$. By Remark 5.3.2, the result is thus proved. □

Lemma 5.4.5 *Let $\varphi: I \times I\!R^m \times I\!R^m \to I\!R$ be a lower semicontinuous function, Lipschitz continuous with respect to the first variable. Assume that $\varphi(t, x, \cdot)$ is convex for a.e. $t \in I$ and for every $x \in I\!R^m$, and that there exist three constants C_i, $i = 0, 1, 2$, such that*

$$|v| \leq C_0|\varphi(t, x, \xi)| + C_1|x| + C_2\,, \tag{5.4.7}$$

for every $(t, x, \xi) \in I \times I\!R^m \times I\!R^m$ and for every $v \in \partial_t\varphi(t, x, \xi)$, where $\partial_t\varphi$ denotes the generalized gradient of φ with respect to t.

Let $u \in W^{1,1}(I, I\!R^m)$, and assume that the function $t \mapsto \varphi(t, u(t), u'(t))$ belongs to $L^1(I)$. Then there exists $k_0 \in L^1(I)$ such that

$$|\varphi(s_2, u(t), u'(t)) - \varphi(s_1, u(t), u'(t))| \leq k_0(t)|s_2 - s_1|,$$

for every $t, s_1, s_2 \in I$.

Proof. For every fixed t_1, $t_2 \in I$, let us define the function

$$g(s) \doteq |\varphi(t_1 + sd, x, \xi) - \varphi(t_1, x, \xi)|\,, \quad s \in [0, 1],$$

where $d \doteq t_2 - t_1$. By (5.4.7), it follows that for a.e. $s \in [0, 1]$

$$g'(s) \leq |d||\partial_t\varphi(t_1 + sd, x, \xi)| \leq |d|\left(C_0 g(s) + C_0|\varphi(t_1, x, \xi)| + C_1|x| + C_2\right).$$

We can apply Gronwall's inequality to the non–negative absolutely continuous function g, obtaining

$$\begin{aligned} |\varphi(t_2, x, \xi) - \varphi(t_1, x, \xi)| &= g(1) \leq \\ &\leq |t_2 - t_1|e^{C_0T}\left(C_0|\varphi(t_1, x, \xi)| + C_1|x| + C_2\right). \end{aligned} \tag{5.4.8}$$

This inequality, with $t_1 = t$ and $t_2 = s_1$, implies that

$$|\varphi(s_1, x, \xi)| \leq |\varphi(t, x, \xi)| + Te^{C_0T}\left(C_0|\varphi(t, x, \xi)| + C_1|x| + C_2\right). \tag{5.4.9}$$

Again by (5.4.8), with $t_1 = s_1$, $t_2 = s_2$, and by (5.4.9), it follows that

$$|\varphi(s_2, x, \xi) - \varphi(s_1, x, \xi)| \leq |s_2 - s_1|(\tilde{C}_0|\varphi(t, x, \xi)| + \tilde{C}_1|x| + \tilde{C}_2),$$

where $\tilde{C}_i \doteq C_i e^{C_0T}(1 + TC_0e^{C_0T})$, $i = 0, 1, 2$. Finally, by hypothesis, the function

$$k_0(t) \doteq \tilde{C}_0|\varphi(t, u(t), u'(t))| + \tilde{C}_1|u(t)| + \tilde{C}_2$$

belongs to $L^1(I)$, completing the proof. □

Definition 5.4.6 *We shall say that $\theta \in C^1((0,+\infty), I\!R)$ is a Nagumo function if θ is convex, increasing and it satisfies* $\lim_{r\to+\infty} \theta(r)/r = +\infty$.

We begin the study of minimization problems, starting with an existence result for convex functionals. We collect here the basic hypotheses on the integrand.

(H_0) $f \in \mathcal{E}$, and $f(t,\cdot)$ is a convex function for every $t \in I$.

(H_1) There exist two constants A and B, with $B > 0$, such that $f(t,\xi) \geq -A + B|\xi|$ for every $(t,\xi) \in I \times I\!R^m$.

(H_2) $g\colon I \times I\!R^m \to I\!R$ is Lipschitz continuous with respect to the first variable, continuous with respect to the second, and there exist two constants α, β, with $0 \leq \beta < B/T$, such that $g(t,x) \geq -\alpha - \beta|x|$ for every $(t,x) \in I \times I\!R^m$.

(H_3) There exist three constants C_i, $i = 0,1,2$, such that the condition (5.4.7) holds with $\varphi(t,x,\xi) \doteq g(t,x) + f(t,\xi)$.

Remark 5.4.7 If $f \in \mathcal{E}$ is independent of t, then it is easily seen that Lemma 5.3.5 and Lemma 5.4.4 imply that condition (H_1) is always satisfied for suitable constants A, B, with $B > 0$.

Theorem 5.4.8 *Let f and g satisfy the hypotheses (H_0), (H_1), (H_2), (H_3). Then there exists a solution $\tilde{u}$ to the problem*

$$\min\left\{F(u) \;\middle|\; u \in W^{1,1}(I, I\!R^m),\, u(0) = a,\, u(T) = b\right\} \tag{5.4.10}$$

where

$$F(u) \doteq \int_I [f(t,u'(t)) + g(t,u(t))]\, dt\,.$$

Moreover $\tilde{u}$ belongs to $W^{1,\infty}(I, I\!R^m)$ and satisfies for a.e. $t \in I$

$$f(t,\tilde{u}'(t)) - \langle p(t)\,,\, \tilde{u}'(t)\rangle + g(t,\tilde{u}(t)) = c + \int_0^t v(\tau)\, d\tau\,, \tag{5.4.11}$$

where $(v(t),p(t)) \in (\partial_t f(t,\tilde{u}'(t)) + \partial_t g(t,\tilde{u}(t)), \partial_\xi f(t,\tilde{u}'(t)))$ for almost every $t \in I$ and c is a constant.

Proof. The proof follows the lines of the one of Theorem 3 in [26], with some changes due to the fact that in this case the lagrangian is not bounded from below. As in [26] one can prove, using the De Giorgi's semicontinuity result (see [14]) and the Dunford–Pettis criterion of weak compactness in $L^1(I, I\!R^m)$, that for every Nagumo function θ and for every $l > 0$ there exists a solution u_l to the problem

$$\min\left\{F(u) \;\middle|\; u \in AC^l_\theta(I, I\!R^m),\, u(0) = a,\, u(T) = b\right\} ,$$

where $AC^l_\theta(I, I\!R^m)$ denotes the class of all function $u \in W^{1,1}(I, I\!R^m)$ such that $\Theta(u) \le l$, with $\Theta(u) \doteq \int_I \theta(|u'(t)|)\, dt$. Let us set $V_\theta(l) \doteq F(u_l)$.

One can easily check that, if $V_\theta(l) = V_\theta(l_0)$ for every $l \ge l_0$, then u_{l_0} is a solution to the problem

$$\text{(5.4.12)} \quad \min\left\{F(u) \;\middle|\; u \in W^{1,1}(I, I\!R^m),\, \Theta(u) < +\infty,\; u(0) = a,\, u(T) = b\right\} .$$

Finally, as in [26], if we are able to prove that u_{l_0} belongs to $W^{1,\infty}(I, I\!R^m)$, then we can conclude that such a function is a solution to (5.4.10).

Thus it remains to prove that V_θ is eventually constant and that, for l large enough, u_l belongs to $W^{1,\infty}(I, I\!R^m)$ and satisfies (5.4.11). Since V_θ is lower semicontinuous, for every $l > 0$ there exists a proximal subgradient (see [25]) of V_θ at l and, since V_θ is nonincreasing, it is nonpositive. If V_θ is not eventually constant, by Proposition 6.1 in [26], there exists a diverging sequence $\{l_k\}$ such that the proximal subgradient of V_θ at l_k takes the form $-r_k$, with $r_k > 0$. Moreover, it is easy to check that, if we set $u_k \equiv u_{l_k}$, then $\Theta(u_k) = l_k$, so that

$$\text{(5.4.13)} \qquad \lim_{k\to+\infty} \|u'_k\|_{L^\infty} \ge \lim_{k\to+\infty} \theta^{-1}(l_k/T) = +\infty .$$

By definition of r_k and the fact that $\Theta(u_k) = l_k$, it follows that for every $k \in I\!N$ there exists a positive constant σ_k such that, if we define

$$G(u) \doteq F(u) + r_k\Theta(u) + \sigma_k|\Theta(u) - \Theta(u_k)|^2 ,$$

then we get that $G(u_k) \le G(u)$ for every u admissible for (5.4.12) and such that $\Theta(u)$ is sufficiently near to $\Theta(u_k)$ (see [26]). By (H_3) and Lemma 5.4.5, it follows that there exists $k_0 \in L^1(I)$ such that for every $s_1, s_2, t \in I$

$$|f(s_1, u'_k(t)) + g(s_1, u_k(t)) - f(s_2, u'_k(t)) - g(s_2, u_k(t))| \le k_0(t)|s_1 - s_2| ,$$

so that we can apply Theorem 5 of [26]. Thus we obtain that u_k satisfies

$$E_f(t,u_k'(t)) + g(t,u_k(t)) + r_k E_\theta(|u_k'(t)|) = c_k + \int_0^t v_k(\tau)\,d\tau\,, \tag{5.4.14}$$

where $E_f(t,u_k'(t)) \doteq f(t,u_k'(t)) - \langle p_k(t)\,,\,u_k'(t)\rangle$, $E_\theta(s) \doteq \theta(s) - s\theta'(s)$, c_k is a constant, and $(v_k(t),p_k(t)) \in (\partial_t f(t,u_k'(t)) + \partial_t g(t,u_k(t)), \partial_\xi f(t,u_k'(t)))$ for a.e. $t \in I$.

Moreover there exists $M_1 > 0$ such that $\|u_k\|_{L^\infty} \leq M_1$ for every $k \in \mathbb{N}$. Actually, if there exists $t_k \in I$ such that $\limsup_k |u_k(t_k)| = +\infty$, then

$$\limsup_{k\to+\infty} \int_I |u_k'(t)|\,dt \geq \limsup_{k\to+\infty} \left|\int_0^{t_k} u_k'(t)\,dt\right| = \limsup_{k\to+\infty} |u_k(t_k) - a| = +\infty\,,$$

while, if we define $u_0(t) \doteq a + \xi t$, with $\xi \doteq (b-a)/T$, then u_0 is admissible for (5.4.12), $F(u_0) < +\infty$, and

$$\begin{aligned} F(u_0) \geq F(u_k) &\geq (-A-\alpha)T + B\|u_k'\|_{L^1} - \beta\|u_k\|_{L^1} \geq \\ &\geq \tilde{A} + (B - \beta T)\|u_k'\|_{L^1}\,, \end{aligned} \tag{5.4.15}$$

so that, by (H_2), $\{u_k'\}$ must be bounded in $L^1(I,\mathbb{R}^m)$.

The boundedness of $\{u_k\}$ in $L^\infty(I,\mathbb{R}^m)$ and the continuity of g guarantee that there exists M_2 such that

$$|g(t,u_k(t)| \leq M_2\,, \tag{5.4.16}$$

for a.e. $t \in I$ and for every k. Moreover, by (H_3) we obtain

(5.4.17)
$$\begin{aligned} \left|\int_0^t v_k(s)\,ds\right| &\leq \int_I [C_0\,|f(s,u_k'(s)) + g(s,u_k(s))| + C_1|u_k(s)| + C_2]\,ds \leq \\ &\leq \int_I \left[C_0\,|\alpha + \beta|u_k(s)| + f(s,u_k'(s)) + g(s,u_k(s))| + \tilde{C}_1|u_k(s)| + \tilde{C}_2\right]\,ds\,, \end{aligned}$$

where $\tilde{C}_1 \doteq C_0\beta + C_1$ and $\tilde{C}_2 \doteq C_0|\alpha| + C_2$. Without loss of generality we can assume that f is positive, so that, thanks to (H_2), it follows that for every $k \in \mathbb{N}$

$$f(s,u_k'(s)) + g(s,u_k(s)) + \alpha + \beta|u_k(s)| \geq 0, \quad \text{a.e. } s \in I\,. \tag{5.4.18}$$

By (5.4.15), (5.4.17) and (5.4.18) there exist $M_3 > 0$ and two constants $\hat{C}_1$, $\hat{C}_2$ such that

$$(5.4.19) \quad \left| \int_0^t v_k(s)\, ds \right| \leq C_0 F(u_k) + \hat{C}_1 \|u_k\|_{L^1} + \hat{C}_2 \leq M_3 \,, \quad \text{for every } t \in I.$$

By (5.4.14), (5.4.16), and (5.4.19) we obtain

$$E_f(t, u_k'(t)) + r_k E_\theta(|u_k'(t)|) \leq c_k + M_2 + M_3 \,,$$

for every $t \in I$ and for every $k \in \mathbb{N}$.

We claim that it is not possible that there exists a subsequence of $\{c_k\}$, still denoted by $\{c_k\}$, such that $\lim_k c_k = -\infty$. Indeed, if this is the case, then for every $t \in I$ we should have

$$(5.4.20) \qquad \lim_{k \to +\infty} E_f(t, u_k'(t)) + r_k E_\theta(|u_k'(t)|) = -\infty \,.$$

Since $f \in \mathcal{E}$ and θ is superlinear, (5.4.20) implies that $\lim_k |u_k'(t)| = +\infty$ for every $t \in I$, which, by Fatou's Lemma, contradicts the boundedness of u_k' in $L^1(I, \mathbb{R}^m)$.

Thus there exists c^* such that $c_k \geq c^*$ for every k. From (5.4.14) we obtain, for every $t \in I$,

$$(5.4.21) \qquad E_f(t, u_k'(t)) + r_k E_\theta(|u_k'(t)|) \geq c^* - M_2 - M_3 \,.$$

Now let us suppose that for every k there exists $t_k \in I$ such that $\limsup_k |\xi_k| = +\infty$, where $\xi_k \doteq u_k'(t_k)$. Since f and θ belong to $\mathcal{E}$, we have

$$\liminf_{k \to +\infty} [E_f(t_k, \xi_k) + r_k E_\theta(|\xi_k|)] \leq \liminf_{k \to +\infty} \sup_{t \in I} \{E_f(t, \xi_k) + r_k E_\theta(|\xi_k|)\} = -\infty \,,$$

in contradiction with (5.4.21). This implies that $\|u_k'\|_{L^\infty}$ is bounded, which contradicts (5.4.13).

So we can conclude that V_θ is eventually constant. Hence for k sufficiently large $u_k \in W^{1,\infty}(I, \mathbb{R}^m)$ is a solution of (5.4.12). Moreover $r_k = 0$, so that u_k satisfies (5.4.11). Then the proof is complete. □

The last part of this section is devoted to the study of the non-convex case. The hypotheses (H_0) and (H_3) will be replaced respectively by:

(H_0') $f \in \mathcal{E}$.

(H_3') There exist three constants C_i, $i = 0, 1, 2$, such that the condition (5.4.7) holds with $\varphi(t, x, \xi) \doteq g(t, x) + f^{**}(t, \xi)$.

Notice that (H_3') requires the Lipschitz continuity of f^{**} with respect to t. The following two lemmas show that this conclusion follows from (H_0') and

(H_4) For every $R > 0$ there exists a constant L such that

$$|f(t,\xi) - f(s,\xi)| \leq L|t - s|, \quad \text{for every } t, s \in I, \text{ and } \xi \in \overline{B}_R .$$

Lemma 5.4.9 *Let $\psi \in \mathcal{E}$, and let us define, for every $(t,p) \in I \times I\!R^m$, the set*

$$W(t,p) \doteq \{\xi \in I\!R^m \mid p \in \partial_\xi \psi^{**}(t,\xi)\} .$$

Then for every $r > 0$ there exists $R > 0$ such that for every $(t,p) \in I \times I\!R^m$ the condition $W(t,p) \cap \overline{B}_r \neq \emptyset$ implies $W(t,p) \subset \overline{B}_R$.

Proof. Suppose, by contradiction, that there exist sequences $(t_n, p_n) \subset I \times I\!R^m$, $(\eta_n) \subset \overline{B}_r$, $(\xi_n) \subset I\!R^m$, with $\lim_n |\xi_n| = +\infty$, such that, for every $n \in I\!N$,

$$(5.4.22) \qquad p_n \in \partial_\xi \psi^{**}(t_n, \eta_n), \quad p_n \in \partial_\xi \psi^{**}(t_n, \xi_n) .$$

From (5.4.22) it follows that, for every $n \in I\!N$,

$$(5.4.23) \qquad \psi^{**}(t_n, \eta_n) - \langle p_n \,, \eta_n \rangle = \psi^{**}(t_n, \xi_n) - \langle p_n \,, \xi_n \rangle .$$

Since (η_n) is a bounded sequence, there exists a constant C such that the left hand side of (5.4.23) is bounded from below by C. Thus

$$(5.4.24) \qquad C \leq \psi^{**}(t_n, \xi_n) - \langle p_n \,, \xi_n \rangle \leq \chi(|\xi_n|), \quad \text{for every } n \in I\!N,$$

where $\chi(R)$ is the argument of the limit in (5.4.1). Since $\lim_n |\xi_n| = +\infty$, from (5.4.1) we have that $\lim_n \chi(|\xi_n|) = -\infty$, which contradicts (5.4.24). □

Remark 5.4.10 Let us fix $\xi \in I\!R^m$. Let $t \in I$, $\tilde{\lambda} \in E_{m+1}$, $\xi_j \in I\!R^m$, $j = 1, \dots, m+1$ satisfy

$$f^{**}(t,\xi) = \sum_{j=1}^{m+1} \lambda_j f(t,\xi_j), \quad \xi = \sum_{j=1}^{m+1} \lambda_j \xi_j.$$

Since for every j there exists $p_j \in \partial_\xi f^{**}(t,\xi)$ such that $\xi_j \in W(t,p_j)$, by Lemma 5.4.9 we obtain that there exists $R > 0$, depending only on $|\xi|$, such that $\xi_j \in \overline{B}_R$ for every $j = 1, \dots, m+1$.

Lemma 5.4.11 *If $f \in \mathcal{E}$ satisfies (H_4), then $f^{**}(\cdot,\xi)$ is Lipschitz continuous for every $\xi \in I\!R^m$.*

Proof. Let us fix $\xi \in I\!R^m$, and consider t, $s \in I$. By Corollary 5.3.7, there exist $\tilde{\lambda}$, $\tilde{\mu} \in E_{m+1}$, ξ_j, $\eta_j \in I\!R^m$, $j = 1, \dots, m+1$, such that

$$f^{**}(t,\xi) = \sum_{j=1}^{m+1} \lambda_j f(t,\xi_j), \quad f^{**}(s,\xi) = \sum_{j=1}^{m+1} \mu_j f(s,\eta_j),$$

and $\xi = \sum_j \lambda_j \xi_j = \sum_j \mu_j \eta_j$. Moreover, one has

$$f^{**}(t,\xi) \le \sum_{j=1}^{m+1} \mu_j f(t,\eta_j), \quad f^{**}(s,\xi) \le \sum_{j=1}^{m+1} \lambda_j f(s,\xi_j).$$

Then, by Remark 5.4.10 and (H_4), there exists $L > 0$, depending only on $|\xi|$, such that

$$f^{**}(s,\xi) - f^{**}(t,\xi) \le \sum_{j=1}^{m+1} \lambda_j [f(s,\xi_j) - f(t,\xi_j)] \le \sum_{j=1}^{m+1} \lambda_j L|t-s| = L|t-s|.$$

In the same way one obtains

$$f^{**}(t,\xi) - f^{**}(s,\xi) \le \sum_{j=1}^{m+1} \mu_j [f(t,\eta_j) - f(s,\eta_j)] \le L|t-s|,$$

completing the proof. □

We are now in a position to prove the existence result for the non–convex case.

Theorem 5.4.12 *Let g and f satisfy the basic hypotheses (H_0'), (H_1), (H_2), (H_3'), (H_4), and assume that $g(t,\cdot)$ is concave for every $t \in I$. Then the problem (5.4.10) has a solution $u \in W^{1,\infty}([0,T],\mathbb{R}^m)$.*

Proof. The proof follows the same lines of the one of Theorem 1 in [16]. It is enough to use Theorem 5.4.8 to obtain a solution $\tilde{u} \in W^{1,\infty}([0,T],\mathbb{R}^m)$ of the relaxed problem, and to replace Lemma IX.3.3 and Proposition IX.3.1 of [35] with Corollaries 5.3.7 and 5.3.8. Since $\tilde{u}' \in L^\infty([0,T],\mathbb{R}^m)$, it is easily seen, using Lemma 5.4.9, that we obtain a solution $u \in W^{1,\infty}([0,T],\mathbb{R}^m)$. □

Example 5.4.13 Consider an elastic, incompressible, isotropic, and homogeneous circular tube of inner and outer radius r_1 and r_2 respectively. In [38] it was shown that, if the tube undergo a helical shear deformation, expressed in cilindrical coordinates as $(r,\vartheta,z) \mapsto (r,\theta+\omega(r),z+\phi(r))$, then the vector function $u \doteq (\phi,\omega)$ is a solution of the problem

$$\min_{u\in W^{1,1}([r_1,r_2],\mathbb{R}^2)} \left\{ \int_{r_1}^{r_2} f(r,u'(r))dr \;\Big|\; u(r_1)=(0,0),\; u(r_2)=(\phi_0,\omega_0) \right\},$$

where $f(r,\xi) \doteq r\cdot W\left(\sqrt{\xi_1^2+r^2\xi_2^2}\right)$. The function W, possibly non–convex, is related to the stored energy density per unit volume of the material. The only physical assumption about its behaviour at infinity is $\lim_{\kappa\to+\infty} W(\kappa)=+\infty$. It is easy to see that, if W satisfies the growth condition

$$\lim_{\kappa\to+\infty}\left(W^{**}(\kappa)-\kappa\cdot\frac{dW^{**}(\kappa)}{d\kappa}\right)=-\infty,$$

then $f \in \mathcal{E}$. Moreover, if W is a Lipschitz function with at most polynomial growth, then all the other assumptions of Theorem 5.4.12 are satisfied (with $g \equiv 0$).

References

[1] Amar, M. and Cellina, C.: On passing to the limit for non convex variational problems, *Asymptotic Analysis*, **9**, 1994, p. 135–148.

[2] Amar, M. and Mariconda, C.: A non-convex variational problem with constraints, *SIAM Journal on Control and Optimization*, **38**, 1995, p. 1–10.

[3] Ambrosio, L., Ascenzi, O. and Buttazzo, G.: Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands, *J. Math. Anal. Appl.*, **142**, 1989, p. 301–316.

[4] Andreini, A. and Bacciotti, A.: Bang–bang and relay principles for time–dependent constrained sets, *Journal of Optimization Theory and Applications*, **31**, 1980, p. 207–216.

[5] Aubin, J.P. and Cellina, A.: *Differential Inclusions*, Springer–Verlag, Berlin, 1984.

[6] Botteron, B. and Dacorogna, B.: Existence and non-existence results for non-coercive variational problems and applications in ecology, *J. Differential Equations*, **85**, 1990, p. 214–235.

[7] Botteron, B. and Marcellini, P.: A general approach to the existence of minimizers of one dimensional non-coercive integrals of the calculus of variations, *Ann. Inst. H. Poincaré*, **8**, 1991, p. 197–223.

[8] Bressan, A.: Directionally continuous selections and differential inclusions, *Funkc. Ekvac*, **31**, 1988, p. 459–470.

[9] Bressan, A.: On the qualitative theory of lower semicontinuous differential inclusions, *J. Differential Equations*, **77**, 1989, p. 379–391.

[10] BRESSAN, A.: The most likely path of a differential inclusion, *J. Differential Equations*, **88**, 1990, p. 155–174.

[11] BRESSAN, A.: Selections of Lipschitz multifunctions generating a continuous flow, *Diff. & Integ. Equat.*, **4**, 1991, p. 483–490.

[12] BRESSAN, A. AND COLOMBO, G.: Boundary value problems for lower semicontinuous differential inclusions, *Funkc. Ekvac.*, **36**, 1993, p. 359–373.

[13] BRESSAN, A. AND CRASTA, G.: Extremal selections of multifunctions generating a continuous flow, *Annales Polonici Mathematici*, **60**, 1994, p. 101–117.

[14] BUTTAZZO, G.: *Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations*, Pitman Res. Notes Math. Ser. **207**, Longman, Harlow, 1989.

[15] CELLINA, A.: On the set of solutions to Lipschitzean differential inclusions, *Diff. & Integ. Equat.*, **1**, 1988, p. 495–500.

[16] CELLINA, A. AND COLOMBO, G.: On a Classical Problem of the Calculus of Variations without Convexity Assumptions, *Annales de l'Institut Henri Poincaré*, **7**, 1990, p. 97–106.

[17] CELLINA, A. AND MARICONDA, C.: The existence question in the calculus of variations: a density result, *Proceedings of the American Mathematical Society*, **120**, 1994, p. 1145–1150.

[18] CELLINA, A., TREU, G. AND ZAGATTI, S.: On the minimum problem for a class of non-coercive functionals, *Preprint SISSA*, 1994.

[19] CËRF, R. AND MARICONDA, C.: On bang–bang constrained solutions of a control system, *SIAM Journal on Control and Optimization*, to appear.

[20] CËRF, R. AND MARICONDA, C.: Oriented measures with continuous densities and the bang–bang principle, *J. Funct. Anal.*, **126**, 1994, p. 476–505.

[21] CËRF, R. AND MARICONDA, C.: Oriented measures and the bang–bang principle, *Compte Rendus de l'Académie de Sciences*, to appear.

[22] CËRF, R. AND MARICONDA, C.: Oriented measures, *Preprint Dip. di Matematica Pura e Appl., Padova*, 1993.

[23] CESARI, L.: *Optimization – theory and applications*, Springer–Verlag, New York, 1983.

[24] CLARKE, F. H.: *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983.

[25] CLARKE, F. H.: *Methods of dynamic and nonsmooth optimization*, CBMS-NSF Regional Conf. Ser. in Appl. Math., Vol. 57, SIAM, Philadelphia, 1989.

[26] CLARKE, F. H.: An indirect method in the calculus of variations, *Trans. Amer. Math. Soc.*, **336**, 1993, p. 655–673.

[27] CLARKE, F. H. AND LOEWEN, P. D.: An intermediate existence theory in the Calculus of Variations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, **16**, 1989, p. 487–526.

[28] CRASTA, G.: An existence result for non-coercive non-convex problems in the Calculus of Variations, *Nonlinear Analysis TMA*, to appear.

[29] CRASTA, G. AND MALUSA, A.: Existence results for non–coercive variational problems, *Preprint SISSA Ref. 130/94/M*, 1994.

[30] CRASTA, G. AND PICCOLI, B.: Bang–bang property for Bolza problems in two dimensions, *Journal of Optimization Theory and Applications*, **83**, 1994, p. 155–165.

[31] CRASTA, G. AND PICCOLI, B.: Special bang–bang solutions of nonlinear control systems, *Nonlinear Differential Equations and Applications*, to appear.

[32] DE BLASI, F. S. AND PIANIGIANI, G.: On the solution set of nonconvex differential inclusions, *J. Differential Equations*, to appear.

[33] De Blasi, F. S. and Pianigiani, G.: Topological properties of nonconvex differential inclusions, *Nonlinear Analysis, Theory, Methods, Appl.*, **20**, 1993, p. 871–894.

[34] Deimling, K.: *Multivalued Differential Equations*, De Gruyter, Berlin, 1992.

[35] Ekeland, I. and Temam, R.: *Convex Analysis and Variational Problems*, North–Holland, Amsterdam, 1977.

[36] Filippov, A. F.: On certain questions in the theory of optimal control, *SIAM J. Control*, **1**, 1962, p. 76–84.

[37] Fleming, W.H. and Rishel, R.W.: *Deterministic and stochastic optimal control*, Springer–Verlag, New York, 1975.

[38] Fosdick, R.L. and MacSithigh, G.: Helical shear of an elastic, circular tube with a nonconvex stored energy, *Arch. Rational Mech. Anal.*, **84**, 1983, p. 31–53.

[39] LeDonne, A. and M. V. Marchi, M. V.: Representation of Lipschitz compact convex valued mappings, *Rend. Accad. Naz. Lincei*, **68**, 1980, p. 278–280.

[40] Lee, E.B. and Markus, L.: *Foundations of Optimal Control Theory*, Wiley, New York, 1967.

[41] Marcellini, P.: Alcune osservazioni sull'esistenza del minimo di integrali nel calcolo delle variazioni, *Rendiconti di Matematica (2)*, **13**, 1980, p. 271–281.

[42] Olech, C.: Extremal solutions of a control system, *Journal of Differential Equations*, **2**, 1966, p. 74–101.

[43] Olech, C.: *The Lyapunov theorem: its extensions and applications*, in "Methods of Non-convex Analysis", A. Cellina ed., Springer–Verlag, New York, 1990.

[44] ORNELAS, A.: Parametrization of Caratheodory multifunctions, *Rend. Sem. Mat. Univ. Padova*, **83**, 1990, p. 33–44.

[45] PERROTTA, S.: On the closure of reachable sets for control systems, *Journal of Optimization Theory and Applications*, **82**, 1994, p. 555–569.

[46] PICCOLI, B.: Regular time optimal syntheses for smooth planar systems, *Rend. Sem. Mat. Univ. Padova*, to appear.

[47] RAYMOND, J.P.: Existence Theorems in Optimal Control Problems without Convexity Assumptions, *Journal of Optimization Theory and Applications*, **67**, 1990, p. 109–132.

[48] RUDIN, W.: *Real and Complex Analysis*, McGraw–Hill, New York, 1974.

[49] SUSSMANN, H.J.: The Structure of Time–Optimal Trajectories for Single–Input Systems in the Plane: the C^∞ Nonsingular Case, *SIAM Journal on Control and Optimization*, **25**, 1987, p. 433–465.

[50] SUSSMANN, H.J.: The Structure of Time–Optimal Trajectories for Single–Input Systems in the Plane: the General Real Analytic case, *SIAM Journal on Control and Optimization*, **25**, 1987, p. 868–904.

[51] SUSSMANN, H.J.: Regular Synthesis for Time–Optimal Control of Single–Input Real–Analytic Systems in the Plane, *SIAM Journal on Control and Optimization*, **25**, 1987, p. 1145–1162.

[52] SUSSMANN, H.J.: A general theorem on local controllability, *SIAM Journal on Control and Optimization*, **25**, 1987, p. 158–194.

[53] TOLSTONOGOV, A. A.: Extreme continuous selectors of multivalued maps and their applications, *Preprint*, 1992.