Control and Cybernetics

vol. 31 (2002) No. 3

Finite-dimensional representations of the value functions of some optimal control problems

by

Ştefan Mirică

University of Bucharest, Faculty of Mathematics,
Academiei 14, 70109 Bucharest, Romania
e-mail: mirica@math.math.unibuc.ro

Abstract: In this survey we analyze the possibility of obtaining information on regularity and irregularity properties of the value functions of some optimal control problems from their more precise description as **marginal functions of finite-dimensional type**, in terms of certain "generalized characteristic flows" which, in turn, may be constructed using either necessary optimality conditions **(PMP-Pontryagin's Minimum Principle)**, whenever applicable, or suitable extensions of Cauchy's Method of Characteristics for the associated Hamilton-Jacobi-Bellman equation.

This type of representation, which may be justified either by the application of **PMP** "combined" with existence theorems or by the application of a suitable verification theorem of Dynamic Programming type, not only facilitates numerical computation of the value function but also may allow identification of its discontinuity points, non-differentiability points, propagation of singularities, etc.; this approach is illustrated with three significant examples from classical Calculus of Variations.

Keywords: optimal control, value function, marginal function, regularity, generalized characteristic flow, generalized Hamiltonian flow, verification theorem

1. Introduction

We consider a *Bolza autonomous optimal control problem for differential inclusions*, $\mathcal{B}_A = (Y_0, Y_1, g(.), g_0(.,.), F(.), \Omega_a)$ which consists in *minimizing each of the cost functionals*

$$C(y; x(.)) := g(x(t_1)) + \int_0^{t_1} g_0(x(t), x'(t))\, dt, \ y \in Y_0 \subset R^n \tag{1.1}$$

over the corresponding set of *admissible trajectories* $\Omega_a(y)$, $y \in Y_0$, defined as satisfy constraints of the form:

$$x'(t) \in F(x(t)) \ a.e.(0,t_1),\ x(0) = y,\ x(t) \in Y_0$$
$$\forall\, t \in [0,t_1),\ x(t_1) \in Y_1 \subset \partial Y_0. \tag{1.2}$$

As it is apparent from this succinct formulation, *the terminal time* $t_1 > 0$ *is free*, depends on *the admissible trajectory* $x(.) \in \Omega_a(y)$ (hence also on the *initial point* $y \in Y_0$) and it may be interpreted as the first moment at which the last two conditions in (1.2) (that define *the terminating rule of the process*) are verified; in what follows we assume that $Y_0 \cap Y_1 = \emptyset$ to avoid possible ambiguities; one may note also that problem $\mathcal{B}_A$ denotes in fact *the family of optimization problems* $\mathcal{B}_A(y),\ y \in Y_0$, defined by (1.1)–(1.2).

As it is well-known (e.g. Boltiansky, 1968, Cesari, 1983, etc.), **the value function of the problem** $\mathcal{B}_A$, defined by:

$$W(y) := \begin{cases} g(y) & \text{if } y \in Y_1 \\ \inf\limits_{x(.)\in\Omega_a(y)} C(y;x(.)) & \text{if } y \in Y_0 \end{cases}, \tag{1.3}$$

is the main tool of the so called *Dynamic Programming Method* and its natural extension on the "terminal set" Y_1 is essential for the *monotonicity property* (M) in Remark 2.9 below and therefore for the derivation of the so called "verification theorems".

On the other hand, the value function it is a rather abstract *marginal function of infinite-dimensional type* and an impressive number of studies are aiming at the identification of classes of problems whose value functions have different types of regularity properties (lipschitzianity, continuity, semi-continuity, etc.); however, for many significant problems (even for certain restricted classes of problems) one may use the associated Hamiltonian:

$$H(x,p) := \inf_{v\in F(x)} \mathcal{H}(x,p,v),\ \mathcal{H}(x,p,v) := \langle p,v\rangle + g_0(x,v) \tag{1.4}$$

and suitable extensions of the Cauchy's Methods of Characteristics (e.g. Mirică, 1987, 1998, Subbotin, 1995, etc.) or even necessary optimality conditions, when applicable, to construct a "generalized characteristic flow", $C^*(.;.) = (X(.;.),P(.;.),V(.;.))$ (which is more precisely defined in Sections 2, 3, 4 of the paper), whose first component defines admissible, "possibly optimal", trajectories, $X(.;a),\ a \in A$, the last component, $V(.;.)$, gives the corresponding values of the cost functional, and the "classical" differential property

$$DV(t;a).(\bar{t},\bar{a}) = \langle P(t;a), DX(t;a).(\bar{t},\bar{a})\rangle \tag{1.5}$$

is satisfied in some generalized sense (stratified, contingent, etc.).

Using either a "combination" of necessary conditions (**PMP** – Pontryagin's Minimum Principle) with a corresponding existence theorem or suitable "veri-

Lupulescu and Mirică, 2000, Mirică, 1992b, 1995, Sussmann, 1990, etc.) one may prove that, under certain conditions, the value function is given by:

$$W(y) = \begin{cases} g(y) & \text{if } y \in Y_1 \\ \inf\limits_{X(t;a)=y} V(t;a) & \text{if } y \in X(B_0) \subset Y_0 \end{cases} \tag{1.6}$$

and in many significant cases turns out to be *a marginal function of finite-dimensional type* since $B_0 \subset (-\infty, 0) \times R^k$; this type of representation not only facilitates its numerical computation but also the identification of points of discontinuity, non-differentiability, propagation of singularities, evaluation of its generalized derivatives, etc.

By identifying the natural characteristic flows we are able to prove results of the following type:

1) in the case of the classical Brachistochrone problem (Bliss, 1925, Cesari, 1983, etc.), the value function is of class C^1 on the set $Y_0 = (0,\infty)^2$ of initial points but only locally-radially Lipschitz at the terminal point $y_1 = (0,0) \in Cl(Y_0)$; moreover, in a certain "non-singular case", the value function is locally-Lipschitz and contingent differentiable also at the terminal point y_1 (Mirică, 1996);

2) in the case of the classical Euler-Plateau problem of minimal surfaces of revolution (Bliss, 1925, Cesari, 1983, etc.), the value function is $C^1 - stratified$ and locally-Lipschitz on the whole domain $Y = Y_0 \cup Y_1 := [0,\infty)^2$ and the points of non-differentiability are more precisely identified (Mirică, 2000);

3) the value functions of some classical problems (e.g., Cesari, 1983) are neither lower nor upper semicontinuous (Lupulescu and Mirică, 2000).

The paper is organized as follows: in Section 2 we consider the "parametrized" optimal control problems for which the **PMP** is proved in its "standard" form and show that the value function may be represented as in (1.6) by all the (normal and abnormal) *extremals* $X^*(.;a) = (X(.;a), P(.;a))$, $a \in A$, provided an existence theorem may be applied; in Section 3 we consider the case of "stratified problems" and show that the same type of representation of the value function is possible using certain "stratified Hamiltonian and characteristic flows" and Dynamic Programming arguments; in Section 4 we use "contingent generalized derivatives" to treat in the same way more general problems and in the last section we illustrate these results on several significant examples from calculus of variations (considered as particular examples of optimal control problems).

2. Generalized characteristic flows generated by PMP

In this section we consider the particular case of *parametrized optimal control* *imization of each of the functionals*

$$C(y;u(.)) := g(x(t_1)) + \int_0^{t_1} f_0(x(t),u(t))dt,\ y \in Y_0 \subset R^n \tag{2.1}$$

over the corresponding set $\mathcal{U}_a(y),\ y \in Y_0$ of *admissible controls* that are mappings $u(.)$ in a prescribled class, $\mathcal{U}_a$, (usually measurable, if U is a topological space) for which the corresponding (AC) solutions, $x(.)$, of the differential system:

$$x'(t) = f(x(t),u(t)),\ u(t) \in U \text{ a.e. } ([0,t_1]),\ x(0) = y, \tag{2.2}$$

belong to given class Ω_a, of *admissible trajectories*, and satisfy the *state constraints* and the *terminal constraints* in (1.2).

We recall first that **PMP** is usually proved under the following:

HYPOTHESIS 2.1 The data of the problem $P\mathcal{B}_A$ have the following properties:

(i) The set U (of *control parameters*) is non-empty (usually a subset of some Euclidean space but also a Hausdorff topological space and even an "unstructured" set), $D = Int(D) \subset R^n$ and the mappings $\widehat{f}(.,u) := (f(.,u), f_0(.,u))$: $D \times U \to R^n \times R,\ u \in U$ are of class C^1;

(ii) The class $\mathcal{U}_a$ of *admissible controls* is a prescribed set of mappings $u(.)$: $[0,t_1] \to U$ such that the mappings

$$\begin{aligned}\widehat{f}_u(t,x) &:= \widehat{f}(x,u(t)) = (f(x,u(t)), f_0(x,u(t)),\ (t,x) \in D_u \\ &:= [0,t_1] \times D\end{aligned} \tag{2.3}$$

are at least of Carathéodory - C^1 type in the sense that the mappings $\widehat{f}_u(.,.)$, $D_2\widehat{f}_u(.,.)$ are of Carathéodory type and moreover, the (unique) solutions, $x(.) = x_u(.;y),\ y \in Y_0$, of the problems in (2.2) as well as the "extended trajectories" $\widehat{x}(.)$ defined by:

$$\widehat{x}(.) := (x(.), x_0(.)),\ x_0(t) := \int_0^t f_0(x(s),u(s))ds,\ t \in [0,t_1] \tag{2.4}$$

belong to the prescribed class, Ω_a, of *admissible trajectories*;

(iii) The set of *initial states*, $Y_0 \subset D$ is **open** (i.e. $Y_0 = Int(Y_0)$), the set of *terminal states*, $Y_1 \subset \partial Y_0 := Cl(Y_0) \setminus Y_0$, is a differentiable manifold and the *terminal cost function* $g(.) : Y_1 \to R$ is differentiable. ■

We note that the class $\mathcal{U}_a$ of admissible controls *ranges* from the "smallest" one, $\mathcal{U}_{pc}$ of *piecewise continuous* admissible controls" (that "generates" the class Ω_{pc} of "piecewise-C^1" admissible trajectories) to the largest one, $\mathcal{U}_1$, that generates ("through" $\widehat{f}(.,.)$), the largest class, Ω_1, of absolutely continuous the classes of admissible control and trajectories is required by the so called "Lavrentiev phenomenon" (e.g. Cesari, 1983) and may be essential in the Dynamic Programming approach; from this point of view, an important case is that of the $\mathcal{U}_r$ of *regulated admissible controls* that generates the class Ω_r of *regular admissible trajectories* for which the derivatives, $x'(.)$, have a countable number of discontinuities, all of the first kind.

As it is well known, Pontryagin's Minimum Principle (**PMP**) is formulated in terms of the associated "pseudo-Hamiltonians" ("Pontryagin's functions", etc.):

$$\mathcal{H}^{p_0}(x,p,u) = \mathcal{H}(x,p,p_0,u) := \langle p, f(x,u)\rangle + p_0 f_0(x,u),\ p \in R^n \tag{2.5}$$

where $p_0 \in \{1,0\}$, and of the ("true") **Hamiltonians**:

$$\begin{aligned} &H^{p_0}(x,p) := \inf_{u\in U} \mathcal{H}^{p_0}(x,p,u),\ \widehat{U}^{p_0}(x,p)\\ &:= \{u \in U;\ \mathcal{H}^{p_0}(x,p,u) = H^{p_0}(x,p)\},\\ &Z^{p_0} = dom(\widehat{U}^{p_0}(.,.)) := \{(x,p) \in D \times R^n; \widehat{U}^{p_0}(x,p) \neq \emptyset\},\\ &p_0 \in \{1,0\}. \end{aligned} \tag{2.6}$$

The equivalent formulation that we will present is expressed in terms of certain "canonical Hamiltonian orientor field" and $(\Omega_a, \mathcal{U}_a)$-solutions of the corresponding "canonical Hamiltonian inclusion" that are defined as follows:

DEFINITION 2.2 *A mapping $X^*(.) = (X(.), P(.))$ is said to be an $(\Omega_a, \mathcal{U}_a)$ - solution of the canonical (Pontryagin's) Hamiltonian inclusion:*

$$\begin{aligned} &(x',p') \in d_P^{\#} H^{p_0}(x,p)\\ &:= \left\{ \left(\frac{\partial \mathcal{H}^{p_0}}{\partial p}(x,p,u), -\frac{\partial \mathcal{H}^{p_0}}{\partial x}(x,p,u) \right);\ u \in \widehat{U}^{p_0}(x,p) \right\} \end{aligned} \tag{2.7}$$

if there exists an admissible control, $\widetilde{u}(.)$, in the class $\mathcal{U}_a$, such that:

$$\begin{aligned} &(X'(t), P'(t))\\ &= \left(\frac{\partial \mathcal{H}^{p_0}}{\partial p}(X(t),P(t),\widetilde{u}(t)), -\frac{\partial \mathcal{H}^{p_0}}{\partial x}(X(t),P(t),\widetilde{u}(t)) \right)\ a.e. \end{aligned} \tag{2.8}$$

$$\widetilde{u}(t) \in \widehat{U}^{p_0}(X(t),P(t))\ a.e. \tag{2.9}$$

and, moreover, the first component, $X(.)$, as well as the real function: $X_0(.)$ defined as in (2.4), are of the type Ω_a. ■

The term "Hamiltonian inclusion" for (2.7) has been used for several reasons from which we mention the fact that, as it is easy to see, *if $H^{p_0}(.,.)$ is differentiable at* $(x,p) \in Int(Z^{p_0})$ then one has:

$$d_P^{\#} H^{p_0}(x,p) = \{d^{\#} H^{p_0}(x,p)\} := \left\{ \left(\frac{\partial H^{p_0}}{}(x,p), -\frac{\partial H^{p_0}}{}(x,p) \right) \right\}. \tag{2.10}$$

so (2.7) becomes a **classical Hamiltonian system** at such points.

We recall now the statement of Pontryagin's Minimum Principle (**PMP**) for whose proof, extensions and generalizations we refer to the abundant literature on this subject.

THEOREM 2.3 (PMP). *If Hypothesis 2.1 is satisfied and* $(\widetilde{x}(.), \widetilde{u}(.)) : [0, \widetilde{t_1}] \to Y \times U$ *is an optimal pair with respect to the initial point* $y \in Y_0$ *then there exist* $p(.) \in AC$ *and* $p_0 \in \{1, 0\}$ *such that the following properties hold:*

I (canonical Hamiltonian inclusion). *The pair* $(\widetilde{x}(.), p(.))$ *is an* $(\Omega_a, \mathcal{U}_a)$*-solution of the canonical Hamiltonian inclusion in (2.7) in the sense of Def. 2.2;*

II (minimum condition). *Besides the relations (2.8), (2.9), the following "minimum condition" is also satisfied:*

$$H^{p_0}(\widetilde{x}(t), p(t)) = \mathcal{H}^{p_0}(\widetilde{x}(t), p(t), \widetilde{u}(t)) = 0 \ a.e.\ ([0, \widetilde{t_1}]); \tag{2.11}$$

III (transversality condition). *At the terminal point* $\widetilde{t_1} > 0$ *the following condition is satisfied:*

$$\langle p(\widetilde{t_1}), v \rangle = p_0 . Dg(\widetilde{x}(\widetilde{t_1})).v \quad \forall\ v \in T_{\widetilde{x}(\widetilde{t_1})} Y_1 \tag{2.12}$$

where $T_\xi Y_1$ *denotes the tangent space to the manifold* Y_1 *at the point* $\xi \in Y_1$*;*

IV (non-triviality condition).

$$(p(t), p_0) \neq (0, 0) \in R^n \times R \quad \forall\ t \in [0, \widetilde{t_1}]. \blacksquare \tag{2.13}$$

We refer to Boltiansky (1968), Cesari (1983), Mirică (1992a), etc., for the (very difficult) proofs of the usual statement in which the canonical inclusion in (2.7) is replaced by the "adjoint equation"

$$p'(t) = -\frac{\partial \mathcal{H}^{p_0}}{\partial x}(\widetilde{x}(t), p(t), \widetilde{u}(t)) \quad \text{a.e. } [0, \widetilde{t_1}]$$

since $(\widetilde{x}(.), \widetilde{u}(.))$ satisfies (2.2) as an admissible pair.

REMARK 2.4 According to the usual terminology, an admissible pair $(\widetilde{x}(.), \widetilde{u}(.))$ that has properties I-IV from Theorem 2.3, is said to be **an extremal pair with multipliers** $(p(.), p_0)$ which is **normal** (in the sense of Mathematical Programming) if $p_0 = 1$ and **abnormal** if $p_0 = 0$; in fact, the (possibly optimal) "extremal pairs" $(\widetilde{x}(.), \widetilde{u}(.))$ may be "recovered" from the (*normal* and, respectively *abnormal* **extremals** $X^*(.) = (X(.), P(.))$ defined as $(\Omega_a, \mathcal{U}_a)$ *-solutions of the canonical Hamiltonian inclusions in* (2.7) *in the sense of Def.* 2.2 with *terminal values* in the folowing **terminal transversality sets**:

$$\begin{aligned} Z_\tau^1 &:= \{(\xi, q) \in Y_1 \times R^n;\ \langle q, v \rangle = Dg(\xi).v\ \forall\ v \in T_\xi Y_1\} \\ Z_\tau^0 &:= \{(\xi, q) \in Y_1 \times (R^n \setminus \{0\});\ \langle q, v \rangle = 0\ \forall\ v \in T_\xi Y_1\}; \end{aligned} \tag{2.14}$$

moreover, since in many cases the Hamiltonians $H^{p_0}(.,.)$, $p_0 \in \{1, 0\}$ are "first be automatically satisfied if the terminal sets in (2.14) are "diminished" upon addition of the conditions:

$$\begin{aligned} &H_1^{p_0}(\xi, q) = 0, \\ &H_1^{p_0}(\xi, q) := \lim_{Z^{p_0} \ni (x,p) \to (\xi,q)} H^{p_0}(x, p), \ (\xi, q) \in Y_1 \times R^n. \end{aligned} \tag{2.15}$$

DEFINITION 2.5 The mapping $X_1^*(.;.) = (X^1(.;.), P^1(.;.)) : B^1 \to Z^1$ is said to be a **normal Hamiltonian flow** for the problem $P\mathcal{B}_A$ in (2.1)–(2.2) if it has the following properties:

(i) for each $z = (\xi, q) \in Z_\tau^1$ there exists a (possibly empty) set $\Lambda^1(z)$ and an extended real function $t^-(.) : A^1 \to [-\infty, 0]$ such that:

$$\begin{aligned} A^1 &:= \{a = (z, \lambda);\ z = (\xi, q) \in Z_\tau^1,\ \lambda \in \Lambda^1(z)\} \\ B^1 &:= \{(t, a);\ a = (z, \lambda) \in A^1,\ t \in I(a) := (t^-(a), 0]\}; \end{aligned} \tag{2.16}$$

(ii) for each $a = (z, \lambda) \in A^1$ the mapping $X_1^*(.;a) = (X^1(.;a), P^1(.;a))$ is a "maximal to the left" (i.e non-continuable) $(\Omega_a, \mathcal{U}_a)$ *-solution of the "normal canonical Hamiltonian inclusion"*

$$(x', p') \in d_P^{\#} H^1(x, p) := \left\{ \left(f(x, u), -\frac{\partial \mathcal{H}^1}{\partial x}(x, p, u) \right);\ u \in \widehat{U}^1(x, p) \right\} \tag{2.17}$$

that satisfies the "terminal conditions":

$$X_1^*(0; a) = (X^1(0; a), P^1(0; a)) = z = (\xi, q) \ if \ a = (z, \lambda) \in A^1, \tag{2.18}$$

the "minimum condition"

$$H^1(X^1(t; a), P^1(t; a)) = 0 \text{ a.e. } (I(a))\ \forall\ a \in A^1 \tag{2.19}$$

and its first component satisfies the "state constraints":

$$X^1(t; a) \in Y_0\ \forall\ t \in I_0(a) := (t^-(a), 0),\ a \in A^1. \tag{2.20}$$

Further, the mapping $C_1^*(.;.) := (X_1^*(.;.), V^1(.;.)) : B^1 \to Z^1 \times R$ is said to be a **normal Characteristic flow** for the problem $P\mathcal{B}_A$ in (2.1)-(2.2) if $X_1^*(.;.) = (X^1(.;.), P^1(.;.))$ is a normal Hamiltonian flow in the sense above and at each point $(t, a) \in B^1,\ a = (z, \lambda) \in A^1$ the last component is given by:

$$V^1(t; a) := g(\xi) + \int_0^t \langle P^1(s; a), (X^1)'(s; a)\rangle\, ds \text{ if } z = (\xi, q) \in Z_\tau^1 \tag{2.21}$$

where $(X^1)'(.; a)$ denotes the derivative of the mapping $X^1(.; a)$. ■

REMARK 2.6 The terms "Hamiltonian" and, respectively, "Characteristic flow" may be justified by several arguments from which we mention only the fact differentiable, the (generalized) Hamiltonian inclusion in (2.17) turns out to be a smooth Hamiltonian system and the components of $C_1^*(.;.)$ satisfy the basic relation in (1.5) (e.g. Mirică, 1998) which may be very important in the Dynamic Programming approach.

On the other hand, from Definitions 2.5, 2.2 it follows that for each $(t,a) \in B_0^1 := \{(t,a) \in B^1,\ t \in I_0(a) := (t^-(a), 0)\}$, there exists an $\mathcal{U}_a$ -mapping $u^1(.;a)$ such that the mapping $(X^1(.;a), u^1(.;a))$ verifies (2.8) and defines the *admissible pair*:

$$x_{t,a}(s) := X^1(t+s;a),\ u_{t,a}(s) := u^1(t+s;a),\ s \in [0,-t] \tag{2.22}$$

with respect to the initial point $y = X^1(t;a) \in Y_0$; moreover, $(x_{t,a}(.), u_{t,a}(.))$ is a *normal extremal pair* in the sense of Remark 2.4 with the "adjoint variable" $p_{t,a}(s) := P^1(t+s;a),\ s \in [0,-t]$ and, since from condition (2.19) it follows that $\langle P^1(s;a), (X^1)'(s;a)\rangle = -f_0(X^1(s;a), u^1(s;a)$ a.e. $(I(a))$, the last component of the characteristic flow characterizes the *value of the cost functional* in (2.1) as follows:

$$V^1(t;a) = \mathcal{C}(y; u_{t,a}(.))\ if\ y = X^1(t;a) \in X^1(B_0^1) \subseteq Y_0. \tag{2.23}$$

Since the "normal trajectories" may "overlap" at some points, the optimal ones are identified from the following additional optimization problem:

$$\begin{aligned} &\widetilde{W}_0^1(y) := \inf_{X^1(t;a)=y} V^1(t;a) \\ &if\ y \in X^1(B_0^1) \subseteq Y_0, \widetilde{Y}_0^1 := dom(\widehat{B}_0^1(.)) \\ &\widehat{B}_0^1(y) := \{(t,a) \in B_0^1;\ X^1(t;a) = y,\ V^1(t;a) = \widetilde{W}_0^1(y)\} \end{aligned} \tag{2.24}$$

which define the **proper normal value function** $\widetilde{W}_0^1(.)$ (which may also be called "the value function of the normal extremals").

However, as simple examples show, some optimal trajectories of a problem may be "abnormal extremals" in the sense of Remark 2.4 which may be "organized" as "abnormal Hamiltonian and Characteristic flows" defined in the same way as the normal ones in Definition 2.5.

DEFINITION 2.7 *The mapping* $X_0^*(.;.) = (X^0(.;.), P^0(.;.)) : B^0 \to Z^0$ *is said to be an* **abnormal Hamiltonian flow** *for the problem* $\mathcal{PB}_A$ *in (2.1)-(2.2) if it has the following properties:*

(i) for each $z = (\xi, q) \in Z_\tau^0$ *there exists a (possibly empty) set* $\Lambda^0(z)$ *and an extended real function* $t^-(.) : A^0 \to [-\infty, 0]$ *such that:*

$$A^0 := \{a = (z,\lambda);\ z = (\xi,q) \in Z_\tau^0,\ \lambda \in \Lambda^0(z)\} \tag{2.25}$$

(ii) for each $a = (z, \lambda) \in A^0$ the mapping $X_0^*(.;a) = (X^0(.;a), P^0(.;a))$ is a "maximal to the left" (i.e non-continuable) $(\Omega_a, \mathcal{U}_a)$ -solution of the "abnormal canonical Hamiltonian inclusion"

$$(x', p') \in d_P^{\#} H^0(x,p) := \left\{ \left(f(x,u), -\frac{\partial \mathcal{H}^0}{\partial x}(x,p,u) \right);\ u \in \widehat{U}^0(x,p) \right\} \quad (2.26)$$

that satisfies the "terminal conditions":

$$X_0^*(0;a) = (X^0(0;a), P^0(0;a)) = z = (\xi, q) \ if\ a = (z, \lambda) \in A^0, \quad (2.27)$$

the "minimum condition"

$$H^0(X^0(t;a), P^0(t;a)) = 0 \ \ a.e.\ (I(a))\ \forall\ a \in A^0 \quad (2.28)$$

and its first component satisfies the "state constraints":

$$X^0(t;a) \in Y_0\ \forall\ t \in I_0(a) := (t^-(a), 0),\ a \in A^0. \quad (2.29)$$

Further, the mapping $C_0^*(.;.) := (X_0^*(.;.), V^0(.;.)) : B^0 \to Z^0 \times R$ is said to be an **abnormal Characteristic flow** *for the problem $P\mathcal{B}_A$ in (2.1)-(2.2) if $X_0^*(.;.) = (X^0(.;.), P^0(.;.))$ is an abnormal Hamiltonian flow in the sense above and at each point $(t,a) \in B_0^0,\ a = (z,\lambda) \in A^0$ the last component is given by:*

$$V^0(t;a) := g(\xi) + \int_t^0 f_0(X^0(s;a), u^0(s;a))ds \ if\ z = (\xi, q) \in Z_\tau^0, \quad (2.30)$$

where $u^0(.;a),\ a \in A^0$ are the admissible controls satisfying (2.8) for $X^0(.;a)$. ■

As in the case of normal extremals, the minimizing abnormal ones may be identified by the **abnormal value function**:

$$\begin{aligned} &\widetilde{W}_0^0(y) := \inf_{X^0(t;a)=y} V^0(t;a) \text{ if} \\ &y \in X^0(B_0^0) \subseteq Y_0, \widetilde{Y}_0^0 := dom(\widehat{B}_0^0(.)) \\ &\widehat{B}_0^0(y) := \{(t,a) \in B_0^0;\ X^0(t;a) = y,\ V^0(t;a) = \widetilde{W}_0^0(y)\}, \end{aligned} \quad (2.31)$$

which may also be called "the value function of the abnormal extremals".

Therefore, **the PMP-value function** (actually, "the value function of all the extremals") is naturally defined by:

$$\widetilde{W}(y) := \begin{cases} g(y) & \text{if } y \in Y_1 \\ \widetilde{W}_0(y) := \min\{\widetilde{W}_0^1(y), \widetilde{W}_0^0(y)\} & \text{if } y \in \widetilde{Y}_0 := \widetilde{Y}_0^1 \cup \widetilde{Y}_0^0 \end{cases} \quad (2.32)$$

and may obviously be written in the form of (1.6) if one "concatenates" the normal and the abnormal characteristic flows:

$$C^*(t;a) := \begin{cases} C_1^*(t;a) & \text{if } (t,a) \in A^1 \\ C_0^*(t;a) & \text{if } (t,a) \in A^0. \end{cases}$$

THEOREM 2.8 (PMP solution). *Let the data of the problem $\mathcal{PB}_A$ in (2.1)–(2.2) satisfy Hypothesis 2.1 and also the following ones:*

(i) the data of the problem $\mathcal{PB}_A$ satisfy the hypotheses of **one** *of the theorems stating the existence of an optimal control for each initial point $y \in \widetilde{Y}_0$ (e.g. Cesari, 1983);*

(ii) the mappings $X^1(.;a)$, $a \in A^1$, $X^0(.;a)$, $a \in A^0$ in Definitions 2.5, 2.7 are **all** *the normal and, respectively, abnormal extremals of the problem.*

Then the function $\widetilde{W}(.)$ defined in (2.32) coincides with the restriction to $\widetilde{Y} := \widetilde{Y}_0 \cup Y_1$ of value function in (1.3) of the problem $\mathcal{PB}_A$ and, moreover, the pairs $(\widetilde{x}_{t,a}(.), \widetilde{u}_{t,a}(.))$ in (2.22) that correspond to the minimizing points $(t,a) \in \widehat{B}_0(y)$, $y \in \widetilde{Y}_0$ of the problem in (2.32) are the only optimal pairs. ■

REMARK 2.9 We note that the very restrictive hypotheses of Theorem 2.8 are severely limiting the class of problems for which its conclusion is valid; in the first place, the hypotheses of the available "existence theorems" are not only restrictive but also very difficult to verify (e.g. Cesari, 1983); on the other hand, from a slightly different point of view, the description of **all the extremals** as solutions of the differential inclusions in (2.7) may be rather difficult; finally, Hypothesis 2.1 itself (under which the **PMP** in its classical form is proved) is very restrictive, eliminating the problems with *active state space constraints* (for which $Y_0 \neq Int(Y_0)$), with "non-smooth data" $f(.,u)$, $f_0(.,u)$, $g(.)$, Y_1 or the more general (non-parametrized) problems defined by differential inclusions.

In the case Hypotheses (i), (ii) in Theorem 2.8 are not satisfied one may still obtain the same conclusion using suitable *Dynamic Programming arguments* that are based on the following rather obvious statement (e.g. Cesari, 1983, Proposition 4.5.i): *the function $\widetilde{W}(.)$ in* (2.32) *coincides with the value function in* (1.3) *of the restriction $\mathcal{PB}_A|\widetilde{Y}_0$* **iff** *it has the following:*

Monotonicity property (M): *for any $y \in \widetilde{Y}_0$ and any admissible pair $(u(.), x(.))$ that satisfies:*

$$x(t) \in \widetilde{Y}_0 \ \forall\, t \in [0, t_1) \tag{2.33}$$

the real function

$$\omega_x(t) := \widetilde{W}(x(t)) + \int_0^t f_0(x(s), u(s))ds, \ t \in [0, t_1] \tag{2.34}$$

is increasing (i.e. $\omega_x(s_1) \leq \omega_x(s_2) \ \forall\, 0 \leq s_1 \leq s_2 \leq t_1$). ■

In turn, the monotonicity property (M) is implied by hypotheses of the so called **verification theorems** containing different types of regularity properties (i.e. Lipschitzianity, continuity, semi-continuity) of the function $\widetilde{W}(.)$ accompanied by suitable *basic differential inequalities* of the form:

satisfied by the restriction $\widetilde{W}_0(.) := \widetilde{W}(.)|\widetilde{Y}_0$, where $\widetilde{D}\widetilde{W}_0(x;.)$ denotes a suitable generalized directional derivative, usually, either a stratified or a contingent one (e.g. Lupulescu and Mirică, 2000, Mirică, 1992b, 1995, Sussmann, 1990, etc.).

The experience shows that in many significant examples the "normal Hamiltonian inclusion" in (2.17) turns out to be a "piecewise smooth" Hamiltonian system and therefore the components $X^1(.;.)$, $P^1(.;.)$, $V^1(.;.)$ of a "normal characterisitic flow" in Definition 2.5 satisfy relations of the form in (1.5) in a piecewise manner (e.g. Mirică, 1998); in turn, these relations usually imply the fact that the "proper normal value function" $\widetilde{W}_0^1(.)$ in (2.24) satisfies differential inequalities of the form from (2.35), which are fundamental in any **DP** verification theorem; on the other hand, the "abnormal extremals" $X^0(.;a)$, $a \in A^0$ in (2.27)–(2.29) seem to have a very "singular" nature, the minimizing ones in (2.32) "filling-up" in a certain sense the domain covered by the normal extremals.

THEOREM 2.10 (Dynamic Programming partial solution). *Let the data of the problem $P\mathcal{B}_A$ in (2.1)–(2.2) satisfy Hypothesis 2.1 and also the following ones:*

(i) the mappings $X^1(.;a)$, $a \in A^1$, $X^0(.;a)$, $a \in A^0$ in Definitions 2.5, 2.7 are specific ("chosen") families of the normal and, respectively, abnormal extremals of the problem.

(ii) the function $\widetilde{W}(.)$ in (2.32) satisfies the hypotheses of **one of the existing DP verification theorems** *(e.g. Lupulescu and Mirică, 2000, Mirică, 1992b, 1995, Sussmann, 1990, etc.).*

Then the function $\widetilde{W}(.)$ defined in (2.32) coincides with the value function in (1.3) of the restriction $P\mathcal{B}_A|\widetilde{Y}_0$ and moreover, the pairs $(\widetilde{x}_{t,a}(.), \widetilde{u}_{t,a}(.))$ in (2.22) that correspond to the minimizing points $(t,a) \in \widehat{B}_0(y)$, $y \in \widetilde{Y}_0$ of the problem in (2.32) are optimal pairs for the problem $P\mathcal{B}_A|\widetilde{Y}_0$. ■

REMARK 2.11 We note that besides avoiding the very restrictive hypotheses (i) and (ii) of Theorem 2.8, the Dynamic Programming (**DP**) argument in Theorem 2.10 may have the following possible advantages:

– one may choose remarkable families of extremals (possibly, only the "normal" ones) for which one may check the hypotheses of a suitable verification theorem;

– one may extend the procedure to much more general problems for which Hypothesis 2.1 is not verified.

One may note here that although the hypotheses of a verification theorem are not always easy to check, in the absence of Hypotheses of Theorem 2.8, this is the only possibility left to prove the optimality of the minimizing trajectories of the additional "parametrized optimization problem" in (2.32).

3. Stratified optimal control problems

In this section we use the DP arguments in Theorem 2.10 to obtain representations of the form in (1.6) for the particular case of *stratified optimal control problems* in (1.1)–(1.2) for which not only the data but also the Hamiltonian in (1.4) and the value function in (2.32) are "stratified" in the following very weak sense:

DEFINITION 3.1 *A non-empty subset $X \subseteq R^n$ is said to be (weakly) C^1-stratified by $\mathcal{S}_X$ if $\mathcal{S}_X$ is a countable partition of X into C^1-submanifolds of R^n (called "strata"); in this case, the tangent space (with respect to the stratification $\mathcal{S}_X$) at $x \in X$ is defined by: $T_xX := T_xS$ if $x \in S \in \mathcal{S}_X$; next, the mapping $f(.) : X \subseteq R^n \to R^k$ is said to be differentiably stratified if there exists a stratification $\mathcal{S}_f$ of X such that for each $S \in \mathcal{S}_f$ the restriction $f_S(.) := f(.)|S$ is differentiable (in the classical sense); in this case, the derivative of $f(.)$ with respect to the stratification $\mathcal{S}_f$ is defined by: $Df(x) := Df_S(x) \in L(T_xS; R^k)$ if $x \in S \in \mathcal{S}_f$.* ■

As illustrative examples we consider the functions $f_1(x) := |x|$, $f_2(x) := \sqrt[3]{x}$, $x \in R$ which are, both, analytically-stratified by $\mathcal{S}_R = \{(-\infty, 0), \{0\}, (0, \infty)\}$.

We note first that if the *set of initial states* Y_0 in (1.2) is stratified in the sense above then from the state constraints in (1.2), $x(t) \in Y_0\ \forall\ t \in [0, t_1)$ it follows that an admissible trajectory should satisfy also the condition $x'(t) \in T_{x(t)}Y_0$ a.e. $(0, t_1)$ (e.g. Mirică, 1995) hence the differential inclusion in (1.2) may be replaced by the following one:

$$x'(t) \in F_T(x(t)), \quad \text{a.e. } (0, t_1), \; F_T(x) := F(x) \cap T_xY_0, \; x \in Y_0 \tag{3.1}$$

which "produces" the (restricted) *Geometric Hamiltonians*:

$$\begin{aligned} &H_T^{p_0}(x,p) := \inf_{v \in F_T(x)} \mathcal{H}^{p_0}(x,p,v), \\ &\widehat{F}_T^{p_0}(x,p) := \{v \in F_T(x);\ \mathcal{H}^{p_0}(x,p,v) = H_T^{p_0}(x,p)\}, \\ &\mathcal{H}^{p_0}(x,p,v) := \langle p, v\rangle + p_0 g_0(x,v), p_0 \in \{1, 0\} \end{aligned} \tag{3.2}$$

which at the boundary points $x \in \partial Y_0$ may be considerably larger than the "original" ones in (2.6) (e.g. see Example 3.2 below).

The so called "stratified problems" are characterized in the first place by the following properties of the data:

HYPOTHESIS 3.2 *The data of the problem $\mathcal{B}_A$ in (1.1)–(1.2) have the following properties: the set $Y_0 \subset R^n$ of admissible initial states is C^1-stratified and the terminal cost function $g(.) : Y_1 \to R$ as well as the geometric Hamiltonians $H_T^{p_0}(.,.) : Z^{p_0} \to R$, $p_0 \in \{1, 0\}$ in (3.2) are differentiably stratified in the sense of Definition 3.1.* ■

In this case the *terminal transversality sets* $Z_\tau^{p_0}$, $p_0 \in \{1,0\}$ are defined in the same way as in (2.14) by the "stratified derivative" $Dg(.)$ but the "canonical Hamiltonian inclusion" in (2.7) is replaced by the following ("stratified") one:

$$\begin{aligned}&(x',p') \in d_s^\# H_T^{p_0}(x,p),\ (x(0),p(0)) = z = (\xi,q) \in Z_\tau^{p_0},\ p_0 \in \{1,0\}\\ &d_s^\# H_T^{p_0}(x,p) := \{(x',p') \in T_{(x,p)}Z^{p_0};\ x' \in \widehat{F}_T^{p_0}(x,p),\\ &\langle x',\overline{p}\rangle - \langle p',\overline{x}\rangle = DH_T^{p_0}(x,p).(\overline{x},\overline{p})\ \forall\ (\overline{x},\overline{p}) \in T_{(x,p)}Z^{p_0}\}.\end{aligned} \tag{3.3}$$

We note that on open (i.e. $2n$-dimensional) strata $S \in \mathcal{S}_{H_T^{p_0}}$ the "stratified Hamiltonian orientor field" in (3.3) coincides with the classical one in (2.10), while on lower dimensional strata it either may have empty values (on "transversal strata") or may be multi-valued (on strata corresponding to the so called "singular extremals" in optimal control); one may note also that in the case both Hypotheses 2.1 and 3.2 are satisfied the canonical Hamiltonian field in (2.7) and the stratified one in (3.2) are related a follows:

$$\begin{aligned}&d_p^\# H^{p_0}(x,p) \cap T_{(x,p)}Z^{p_0} \subseteq d_s^\# H_T^{p_0}(x,p)\\ &\forall\ (x,p) \in Z^{p_0} := dom\widehat{F}_T^{p_0}(.,.)\end{aligned} \tag{3.4}$$

so the use of the "stratified Hamiltonian inclusion" in (3.3) to construct generalized Hamiltonian and Characteristic flows as in Definitions 2.5, 2.7 may "produce" non-extremal admissible trajectories which, however, could be eliminated by the minimizing processes in (2.24), (2.31), (2.32).

In this case the optimality of the minimizing trajectories of the additional problem (2.32) may be proved only in the framework of Dynamic Programming:

THEOREM 3.3 (DP partial solution). *Let the data of the problem $\mathcal{B}_A$ in (1.1)–(1.2) satisfy Hypothesis 3.2 and also the following ones:*

(i) the mappings $C_{p_0}^*(.;a) = (X^{p_0}(.;a), P^{p_0}(.;a), V^{p_0}(.;a))$, $a \in A^{p_0}$, $p_0 \in \{1,0\}$ are specific ("chosen") families of the generalized "generalized" normal and, respectively, abnormal extremals of the problem in the sense of Definitions 2.5, 2.7, respectively, generated by the stratified Hamiltonian inclusion in (3.3);

(ii) the function $\widetilde{W}(.)$ in (2.32) satisfies the hypotheses of **one of the existing DP verification theorems**.

Then the function $\widetilde{W}(.)$ defined in (2.32) coincides with the value function in (1.3) of the restriction $\mathcal{B}_A|\widetilde{Y}_0$ and moreover, the mappings $\widetilde{x}_{t,a}(.)$ in (2.22) that correspond to the minimizing points $(t,a) \in \widehat{B}_0(y)$, $y \in \widetilde{Y}_0$ of the problem in (2.32) are optimal trajectories for the problem $\mathcal{B}_A|\widetilde{Y}_0$. ■

REMARK 3.4 Experience shows that in many examples the stratified Hamiltonian orientor fields $d_s^\# H_T^{p_0}(.,.)$ in (3.3) have "smooth selections" on the singular (non-transversal) strata so (3.3) becomes a piecewise smooth Hamiltonian sysmoreover, one may choose the Hamiltonian and the Characteristic flows such that its components are stratified and verify (1.5) at each point in stratified sense and this fact may facilitate the verification of the differential inequality in (2.35) that is needed in the verification theorems.

4. General optimal control problems

In case the problem $\mathcal{B}_A$ in (1.1)-(1.2) is no more stratified in the sense of Hypothesis 3.2 one may use other concepts and results from the so called "Nonsmooth Analysis" (e.g. Aubin and Frankowska, 1990) to obtain similar results.

Natural generalizations of the concepts and results in Sections 2 and 3 may be obtained using the *contingent* and, respectively, the *quasitangent cones* to a subset $X \subset R^n$ at a point $x \in X$:

$$\begin{array}{l} K_x^{\pm}X := \{v \in R^n;\ \exists\ (s_k, v_k) \to (0_{\pm}, v):\ x + s_k v_k \in X\ \forall\ k \in N\} \\ Q_x^{\pm}X := \{v \in R^n;\ \forall\ s_k \to 0_{\pm}\ \exists\ v_k \to v:\ x + s_k v_k \in X\ \forall\ k \in N\} \\ K_x X := K_x^+ X \cap K_x^- X,\ Q_x X := Q_x^+ X \cap Q_x^- X, \end{array} \tag{4.1}$$

and the *extreme contingent derivatives* of a real-value function $g(.) : X \to R$ at a point $x \in X$ in a direction $v \in K_x^{\pm}X$:

$$\begin{array}{l} \overline{D}_K^{\pm} g(x; v) = \limsup\limits_{(s,u)\to(0_{\pm},v)} \dfrac{g(x + s.u) - g(x)}{s} \\ \underline{D}_K^{\pm} g(x; v) = \liminf\limits_{(s,u)\to(0_{\pm},v)} \dfrac{g(x + s.u) - g(x)}{s}. \end{array} \tag{4.2}$$

We note first that from the state constraints in (1.2), $x(t) \in Y_0\ \forall\ t \in [0, t_1)$ it follows that an admissible trajectory should satisfy also the condition $x'(t) \in Q_{x(t)}Y_0$ a.e. $(0, t_1)$ (e.g. Mirică, 1995) hence the differential inclusion in (1.2) may be replaced by the following one:

$$x'(t) \in F_Q(x(t)),\ \ \text{a.e. } (0, t_1),\ F_Q(x) := F(x) \cap Q_x Y_0,\ x \in Y_0 \tag{4.3}$$

which "produces" the (restricted) *Quasitangent Hamiltonians*:

$$\begin{array}{l} H_Q^{p_0}(x, p) := \inf\limits_{v \in F_Q(x)} \mathcal{H}^{p_0}(x, p, v), \widehat{F}_Q^{p_0}(x, p) := \{v \in F_Q(x); \\ \mathcal{H}^{p_0}(x, p, v) = H_Q^{p_0}(x, p)\}, \\ \mathcal{H}^{p_0}(x, p, v) := \langle p, v\rangle + p_0 g_0(x, v),\ p_0 \in \{1, 0\} \end{array} \tag{4.4}$$

which at the boundary points $x \in \partial Y_0$ may be considerably larger than the "original" ones in (2.6).

In this case the *terminal transversality sets* $Z_\tau^{p_0}$, $p_0 \in \{1, 0\}$ in (2.14) should be replaced by the following ones, that are defined by the "extreme contingent derivatives" of $g(.)$:

where $(q, p_0) \neq (0, 0)$, and the "canonical Hamiltonian inclusion" in (2.7) as well as the stratified one in (3.3) are replaced by the following ("contingent") one:

$$
\begin{aligned}
&(x', p') \in \overline{d}^{\#}_K H^{p_0}_Q(x, p),\ (x(0), p(0)) = z = (\xi, q) \in Z^{p_0}_K,\ p_0 \in \{1, 0\} \\
&\overline{d}^{\#}_K H^{p_0}_Q(x, p) := \{(x', p') \in K^-_{(x,p)} Z^{p_0};\ x' \in \widehat{F}^{p_0}_Q(x, p), \\
&\langle x', \overline{p} \rangle - \langle p', \overline{x} \rangle \leq \underline{D}^-_K H^{p_0}_Q((x, p); (\overline{x}, \overline{p}))\ \forall\, (\overline{x}, \overline{p}) \in K^-_{(x,p)} Z^{p_0}\}.
\end{aligned}
\tag{4.6}
$$

We note that at interior points $(x, p) \in Int(Z^{p_0})$ at which $H^{p_0}_Q(.,.)$ is differentiable, the "contingent Hamiltonian orientor field" in (4.6) coincides with the classical one in (2.10), while at other points either it may have empty values or may be multi-valued; one may note also that in the case Hypothesis 2.1 is also satisfied, the canonical Hamiltonian field in (2.7) and the contingent one in (4.6) are related a follows:

$$
d^{\#}_P H^{p_0}(x, p) \cap K^-_{(x,p)} Z^{p_0} \subseteq \overline{d}^{\#}_K H^{p_0}_Q(x, p)\ \forall\, (x, p) \in Z^{p_0} := dom \widehat{F}^{p_0}_Q(.,.),
$$

so the use of the "contingent Hamiltonian inclusion" in (4.6) to construct generalized Hamiltonian and Characteristic flows as in Definitions 2.5, 2.7 may "produce" non-extremal admissible trajectories which, however, could be eliminated by the minimizing processes in (2.24), (2.31), (2.32).

In this case the optimality of the minimizing trajectories of the additional problem (2.32) may be proved only in the framework of Dynamic Programming:

THEOREM 4.1 (DP partial solution). *Let $\mathcal{B}_A$ be the problem in (1.1)–(1.2) and let the mappings $C^*_{p_0}(.; a) = (X^{p_0}(.; a), P^{p_0}(.; a), V^{p_0}(.; a))$, $a \in A^{p_0}$, $p_0 \in \{1, 0\}$ be specific ("chosen") families of the "generalized" normal and, respectively, abnormal extremals of the problem in the sense of Definitions 2.5, 2.7, respectively, generated by the contingent Hamiltonian inclusion in (4.6).*

If the function $\widetilde{W}(.)$ in (2.32) satisfies the hypotheses of **one of the existing DP verification theorems** *then the function $\widetilde{W}(.)$ defined in (2.32) coincides with the value function in (1.3) of the restriction $\mathcal{B}_A|\widetilde{Y}_0$ and moreover, the mappings $\widetilde{x}_{t,a}(.)$ in (2.22) that correspond to the minimizing points $(t, a) \in \widehat{B}_0(y)$, $y \in \widetilde{Y}_0$ of the problem in (2.32) are optimal trajectories for the problem $\mathcal{B}_A|\widetilde{Y}_0$.* ■

REMARK 4.2 The comments in Remark 3.4 remain obviously valid for the generalized characteristic flows generated by the contingent Hamiltonian orientor fields $\overline{d}^{\#}_K H^{p_0}_Q(.,.)$ in (4.6), which may turn out to be a piecewise smooth Hamiltonian system for which the basic relations in (1.5) are verified in a piecewise manner; moreover, in certain cases one may choose the Hamiltonian and the Characteristic flows in Definitions 2.5, 2.7 such that its components verify the basic relation in (1.5) in the following generalized sense:

where $K^{-}X((t;a);.,.)$ denotes the set-valued "left" contingent derivative of the mapping $X(.;.)$ at the point $(t,a) \in B_0$; as in the case of relation (1.5) verified in classical or "stratified sense", certain results on the marginal functions of the form in (1.6) show that the inequality in (4.7) may facilitate the verification of the "basic" differential inequality of the form in (2.35) that is needed in the verification theorems.

5. Examples

In this section we illustrate the method above on several significant examples for whose value functions one may obtain representations of the form in (1.6) as marginal functions of finite-dimensional type.

EXAMPLE 5.1 **The Brachistochrone problem**, formulated and solved first by Johann Bernoulli in 1696, is considered by most authors to mark the beginning of Calculus of Variations and it is more or less completely studied in most books and monographs in the field, using the multitude of classical results in Calculus of Variations (e.g. Bliss, 1925, Cesari, 1983, etc.).

The Dynamic Programming solution we are going to describe very shortly (for details see Mirică, 1996) is not only simpler and more complete but also allows a more precise description of the value function and of its regularity properties.

We recall first that "geometrically", the problem consists in *finding a curve joining two given points, P_0, P_1, in a vertical plane, such that a material point of mass $m > 0$ falling under gravity and without friction, travels from P_0 to P_1* **in the shortest time**; analytically, fixing one of the points, $P_1 = (0,0)$, as a problem of the form in (1.1)–(1.2), the Brachistochrone problem may be formulated as follows:

Given $k := (v_0)^2/2g \geq 0$ (where $v_0 \geq 0$ is a possible "initial" velocity), *minimize each of the functionals*:

$$\mathcal{C}(y;x(.)) := \int_0^{t_1} \frac{\|x'(t)\|}{\sqrt{x_2+k}}dt, \; y = (y_1,y_2) \in Y_0 := (0,\infty) \times (-k,\infty) \quad (5.1)$$

subject to:

$$x(.) \in AC, \; x(0) = y, \; x(t) \in Y_0 \; \forall \, t \in [0,t_1), \; x(t_1) \in Y_1 := \{(0,0)\}. \quad (5.2)$$

The "normal Hamiltonian" in (1.4), which in our case is defined by:

$$H(x,p) := \inf_{v \in R^2} \mathcal{H}(x,p,v), \; \mathcal{H}(x,p,v) := \langle p,v \rangle + \frac{\|v\|}{\sqrt{x_2+k}}$$

turns out to be given by:

$$H(x,p) = \begin{cases} -\infty & \text{if } \|p\| > \frac{1}{\sqrt{x_2+k}} \end{cases}$$

$$\widehat{F}(x,p) = \begin{cases} -\emptyset & \text{if } \|p\| > \frac{1}{\sqrt{x_2+k}} \\ \{\mu.p; \mu \leq 0\} & \text{if } \|p\| = \frac{1}{\sqrt{x_2+k}} \\ \{0\} & \text{if } \|p\| < \frac{1}{\sqrt{x_2+k}} \end{cases}$$

and is obviously C^1-stratified in the sense of Definition 3.1 by $\mathcal{S}_H = \{S_1, S_2\}$,

$$S_1 := \{(x,p);\ \|p\| = 1/\sqrt{x_2+k}\},$$
$$S_2 := \{(x,p);\ \|p\| < 1/\sqrt{x_2+k}\}.$$

As it is easy to see, on the only "stratum" of interest, S_1, the "geometric Hamiltonian field" in (3.3) is given by:

$$d_s^{\#}H(x,p) = \{p_2'(-2(x_2+k)p, (0,1));\ p_2' \geq 0\} \text{ if } (x,p) \in S_1$$

and therefore upon choosing $p_2' = p_2/p_1$, $p_1 > 0$, we obtain the following smooth Hamiltonian system:

$$\begin{cases} x_1' = -2(x_2+k), & x_1(0) = 0 \\ x_2' = -2(x_2+k)\frac{p_2}{p_1}, & x_2(0) = 0, \end{cases} \begin{cases} p_1' = 0,\ p(0) = q \in Q_\tau^1 \\ p_2' = \frac{1}{p_1}((p_1)^2 + (p_2)^2) \end{cases} \tag{5.3}$$

on the "non-symplectic" (3-dimensional) differentiable manifold S_1.

Regarding the problem of choosing suitable Hamiltonian and Characteristic flows (and of the end-point transversality values in (2.14)) we have to treat separately the "singular case" in which $k = 0$ and for which the function $g_0(x,v) = \|v\|/\sqrt{x_2}$ as well as the Hamiltonian are not defined at the terminal point $y_1 = (0,0)$.

The nonsingular case $k > 0$ may be treated in the framework of Sections 2-4 choosing the following set of "terminal transversality values" (that correspond to Z_τ^1 in (2.14), (4.5)):

$$Q_\tau^1 := \left\{\frac{1}{\sqrt{k}}(\cos\theta, \sin\theta);\ \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)\right\}; \tag{5.4}$$

standard computations show that the differential system in (5.3) with the terminal values in the set Q_τ^1 in (5.4) "produces" the following smooth Hamiltonian flow:

$$X_1(t;\theta) := k.\tan\theta - \frac{k}{1+\cos(2\theta)}[2t + \cos(2(t+\theta))]$$

$$X_2(t;\theta) := -k + \frac{k}{1+\cos(2\theta)}[1 + \cos(2(t+\theta))] \tag{5.5}$$

$$P(t;\theta) := (1, \tan(t+\theta)).\frac{\cos\theta}{\sqrt{k}},\ t \in I(\theta) := \left(-\frac{\pi}{2} - \theta, 0\right],\ \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

while the V^1-function in (2.21) is given by:

$$V^1(t;\theta) = -2\sqrt{k}\frac{t}{\ldots},\quad (t,\theta) \in B_\tau := \left\{(t,\theta);\ \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),\ t \in I(\theta)\right\}$$

and therefore the "would be" value function of the problem is given by the formula of the form in (1.6):

$$\widetilde{W}_0^1(y) := \inf_{X(t;\theta)=y}\left[-2\sqrt{k}\frac{t}{\cos\theta}\right],\ y \in X(B_0) \subset Y_0. \tag{5.6}$$

A further "Calculus result" shows that the mapping $X(.;.) : B_0 \to Y_0$ in (5.5) (that defines the "cycloids", $X(.;\theta)$, $\theta \in (-\frac{\pi}{2},\frac{\pi}{2})$), is a C^1-diffeomorphism with the inverse $\widehat{B}_0(.) = (\widehat{t}(.),\widehat{\theta}(.))$, hence the "proper normal value function" in (5.6) is given by the formula: $\widetilde{W}_0^1(y) = -2\sqrt{k}\frac{\widehat{t}(y)}{\cos\widehat{\theta}(y)}$, $y \in X(B_0) = Y_0$ and moreover, its derivative is given by: $D\widetilde{W}_0^1(y) = P(\widehat{B}_0(y)) = (1, \tan(\widehat{t}(y) + \widehat{\theta}(y))).\frac{\cos\widehat{\theta}(y)}{\sqrt{k}}$, $y \in Y_0$. Finally, from the properties of the inverse $\widehat{B}_0(.)$ it follows that $D\widetilde{W}_0^1(y) \to \frac{1}{\sqrt{k}}(1,0)$ as $y \to y_1 := (0,0)$ hence the value function in (1.3), (1.6) is of class C^1 on the set Y_0 of initial states, locally-Lipschitz on the set $Y := Y_0 \cup Y_1$ and also contingent differentiable at the terminal point $y_1 = (0,0)$.

In the singular case $k = 0$ the only Hamiltonian flow that verifies the conditions in (2.18)-(2.20) is obtained from the differential system in (5.3) with the terminal condition $p_1(0) = q_1 = \lambda \in (0,\infty)$ which "produces" the mappings:

$$\begin{aligned} &X(t;\lambda) := ((\sin(2t) - 2t)/2\lambda^2, \sin^2 t/\lambda^2),\ t \in I(\lambda) := (-\pi, 0],\\ &P(t;\lambda) := \lambda(1, -\cot(t)),\ \lambda \in (0,\infty) \end{aligned} \tag{5.7}$$

while the function in (2.21) is given by: $V^1(t;\lambda) = -2t/\lambda$, $(t,\lambda) \in B_0 := (-\pi, 0) \times (0,\infty)$ and therefore the "would be" value function of the problem is given by the formula of the form from (1.6): $\widetilde{W}_0^1(y) := \inf_{X(t;\theta)=y}[-2t/\lambda]$, $y \in X(B_0) \subset Y_0$.

A similar "Calculus result" shows that the mapping $X(. : .) : B_0 \to Y_0$ in (5.7) is a C^1-diffeomorphism with the inverse $\widehat{B}_0(.) = (\widehat{t}(.),\widehat{\lambda}(.))$, hence the "proper normal value function" above is given by the formula: $\widetilde{W}_0^1(y) = -2\widehat{t}(y)/\widehat{\lambda}(y)$, $y \in X(B_0) = Y_0$ and moreover, its derivative is given by: $D\widetilde{W}_0^1(y) = \widehat{\lambda}(y)(1, -\cot\widehat{t}(y))$. Finally, from the properties of the inverse $\widehat{B}_0(.)$ it follows that the derivative $D\widetilde{W}_0^1(y)$ *is unbounded* as $y \to y_1 := (0,0)$, hence the value function in (1.3), (1,6) is of class C^1 on the set Y_0 of initial states, but not locally-Lipschitz at the terminal point $y_1 = (0,0)$; moreover, a further study leads to the conclusion that the value function $W(.)$ is "locally radially-Lipschitz" (hence continuous) but not contingent differentiable at the terminal point $y_1 = (0,0)$.

We note that in both cases, the *justification* of the representation in (1.6) relies on the straightforward application of the **DP** "elementary verification theorem" in which the value function $W(.)$ is continuous and the restriction $W_0(.) := W(.)|Y_0$ is differentiable (in our case, of class C^1) and satisfies (2.35)

EXAMPLE 5.2 **The Euler–Plateau problem on minimal surfaces of revolution,** as famous as the Brachistochrone, "geometrically", is formulated as follows: *find a curve joining two given points, P_0, P_1, in the same half-plane, such that the surface obtained by rotating the curve around the axis has the* **minimal area**; analytically, fixing $P_1 = (0,1)$, as a problem of the form from (1.1)–(1.2), the problem may be formulated as follows:

Minimize each of the functionals:

$$\mathcal{C}(y;x(.)) := \int_0^{t_1} x_2(t)\sqrt{(x_1'(t))^2 + (x_2'(t))^2}dt, y = (y_1, y_2) \in Y_0, \tag{5.8}$$

subject to:

$$\begin{array}{l} x(0) = y,\ x(t) \in Y_0 := [R \times [0,\infty)] \setminus Y_1 \\ \forall\, t \in [0, t_1),\ x(t_1) \in Y_1 := \{(0,1)\}. \end{array} \tag{5.9}$$

Since the set Y_0 is C^1-stratified in the sense of Def.3.1 and its tangent space is given by:

$$T_x Y_0 = \begin{cases} R^2 & \text{if } x \in Y_0^1 := (0,\infty)^2 \setminus Y_1 \\ Y_0^2 & \text{if } x \in Y_0^2 := R \times \{0\}, \end{cases}$$

the "restricted" (geometric) orientor field in (3.1) is given by: $F_T(x) \equiv T_x Y_0$ and the "normal geometric Hamiltonian" in (3.2), which in our case is defined by: $H_T(x,p) := \inf_{v \in F_T(x)} \mathcal{H}(x,p,v)$, $\mathcal{H}(x,p,v) := \langle p, v\rangle + x_2\|v\|$ turns out to be given by:

$$H_T(x,p) = \begin{cases} -\infty & \text{if } (x,p) \in Y_0 \times R^2 \setminus \cup_1^3 S_j \\ 0 & \text{if } (x,p) \in \cup_1^3 S_j, \end{cases}$$

$$\widehat{F}_T(x,p) = \begin{cases} \{\mu.p; \mu \le 0\} & \text{if } (x,p) \in S_1 \\ R \times \{0\} & \text{if } (x,p) \in S_2 \\ \{(0,0)\} & \text{if } (x,p) \in S_3 \end{cases}$$

where the "strata" S_j, $j = 1,2,3$ are defined by: $S_1 := \{(x,p); x_2 = \|p\| > 0\}$, $S_2 := R \times \{(0,0)\} \times R$, $S_3 := \{(x,p) \in Y_0^1 \times R^2; \|p\| < x_2\}$.

As it is easy to see, on the only "strata" of interest, S_1, S_2, the "geometric Hamiltonian field" in (3.3) is given by:

$$d_s^{\#} H(x,p) = \begin{cases} \{(\frac{p_2'}{x_2}p, (0, p_2'));\ p_2' \le 0\} & \text{if } (x,p) \in S_1 \\ S_2 & \text{if } (x,p) \in S_2 \end{cases}$$

and therefore by choosing $p_2' = -1$ we obtain the following smooth Hamiltonian system:

$$\begin{cases} x_1' = -p_1/x_2, & x_1(0) = 0 \\ x_2' = -p_2/x_2, & x_2(0) = 0 \\ p_1' = 0,\ p(0) = q_1,\ q = (q_1, q_2) \in Q_\tau^1 \end{cases} \tag{5.10}$$

on the "non-symplectic" (3-dimensional) differentiable manifold S_1.

For the set of "terminal transversality values" (corresponding to Z_τ^1 in (2.14), (4.5)):

$$Q_\tau^1 := \left\{(\cos\theta, \sin\theta);\ \theta \in \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right)\right\}; \tag{5.11}$$

standard computations show that the differential system in (5.10) with the terminal values in the set Q_τ^1 in (5.11) "produces" the following stratified Hamiltonian flow:

$$\begin{aligned} &X_1^{(1)}(t;\theta) := \begin{cases} 0 & \text{if } \theta \in \{-\frac{\pi}{2}, \frac{\pi}{2}\} \\ \ln \frac{X_2^{(1)}(t,\theta)-t-\sin\theta}{1+\sin\theta}.\cos\theta & \text{if } \theta \neq \pm\frac{\pi}{2} \end{cases} \\ &X_2^{(1)}(t;\theta) := \sqrt{(t-\sin\theta)^2 + (\cos\theta)^2} = \|P^{(1)}(t,\theta)\| \\ &P^{(1)}(t;\theta) := (\cos\theta, -t+\sin\theta)),\ t \in I(\theta) := (t_1(\theta), 0], \\ &t_1(\theta) := \begin{cases} -\infty & \text{if } \theta \in (-\frac{\pi}{2}, \frac{3\pi}{2}) \\ -1 & \text{if } \theta = -\frac{\pi}{2} \end{cases} \end{aligned} \tag{5.12}$$

while the function in (2.21) is given by:

$$V_{(1)}^1(t;\theta) = (1/2)[X_1^{(1)}(t;\theta)\cos\theta - \sin\theta) + (\sin\theta - t)X_2^{(1)}(t;\theta)]. \tag{5.13}$$

Moreover, the trajectory $X_{(1)}^*(.; -\frac{\pi}{2})$ may be continued "backwards", for $t < -1$, first in the stratum S_2, then again in the stratum S_1 to obtain the following "flow of Goldschmidt trajectories":

$$\begin{aligned} &X_{(2)}^*(t;\lambda) \\ &:= \begin{cases} ((0, t+1), 0, -t-1)) & \text{if } t \in [t_1^2(\lambda), 0],\ \lambda \in R^* \\ ((-(t+1)sign(\lambda), 0), (0,0)) & \text{if } t \in [t_2^2(\lambda), t_1^2(\lambda)] \\ ((\lambda, -t+t_2^2(\lambda)), (0, -t+t_2^2(\lambda))) & \text{if } t \in (-\infty, t_2^2(\lambda)] \end{cases} \\ &t_1^2(\lambda) := -1,\ t_2^2(\lambda) := -1-|\lambda|,\ \lambda \in R^* := R \setminus \{0\} \end{aligned} \tag{5.14}$$

for which the V-function in (2.21) is given by:

$$V_{(2)}^1(t;\lambda) = \begin{cases} \frac{1}{2}[1 - (X_2^{(2)}(t;\lambda))^2] & \text{if } t \in [t_1^2(\lambda), 0] \\ \frac{1}{2}[1 + (X_2^{(2)}(t;\lambda))^2] & \text{if } t \in (-\infty, t_2^2(\lambda)] \end{cases} \tag{5.15}$$

and therefore the corresponding value function of the "Goldschmidt trajectories" may be obtained in the explicit form:

$$\begin{aligned} &W_0^2(y) = \inf_{X^{(2)}(t;\lambda)=y} V_{(2)}^1(t;\lambda) \\ &\begin{cases} \frac{1}{2}[1-(y_2)^2] & \text{if } y = (0, y_2), y_2 \in [0, 1] \end{cases} \end{aligned} \tag{5.16}$$

On the other hand, a rather complicated but straightforward analysis of the "catenaries mapping" $X^{(1)}(.;\theta)$, $\theta \in (-\frac{\pi}{2}, \frac{3\pi}{2})$ in (5.12) shows that the corresponding value function (along the "catenaries"):

$$W_0^1(y) := \inf_{X^{(1)}(t;\theta)=y} V_{(1)}^1(t;\theta),\ y \in \widetilde{Y}_0^1 := X^{(1)}(B_0^1) \subset Y_0 \tag{5.17}$$

is locally-Lipschitz and C^1-stratified on the proper subset $\widetilde{Y}_0^1 \subset Y_0$ that is "covered" by the classical catenaries.

Finally, further analysis, also of "Calculus" type, shows that the "would be" value function given by the formula:

$$\widetilde{W}(y) := \begin{cases} 0 & \text{if } y \in Y_1 := \{(0,1)\} \\ \widetilde{W}_0(y) := \min\{W_0^1(y), W_0^2(y)\} & \text{if } y \in Y_0 \end{cases} \tag{5.18}$$

has the same regularity properties i.e. it is locally-Lipschitz and C^1-stratified in the sense of Definition 3.1 and, moreover, the standard application of the "verification theorem for stratified value function" (e.g. Lupulescu and Mirică, 2000, Mirică, 1995, 2000) proves that it coincides with the value function in (1.3) of the problem.

EXAMPLE 5.3 Let us consider the problem of *minimizing each of the functionals*

$$\mathcal{C}(y;u(.)) := -(x_2(t_1))^2 + \int_0^{t_1} (u(t))^2 dt,\ y \in Y_0 := (-\infty, 0) \times R \tag{5.19}$$

subject to:

$$\begin{cases} x_1' = 1,\ x_1(0) = y_1 < 0, & x(t) \in Y_0\ \forall\ t \in [0, t_1) \\ x_2' = u(t) \in U := R,\ x_2(0) = y_2 \in R, & x(t_1) \in Y_1 := \{0\} \times R. \end{cases} \tag{5.20}$$

The normal Hamiltonian in (1.4), which in our case is defined by: $H(x,p) := \inf_{u \in R}[p_1 + p_2 u + u^2]$, turns out to be the smooth function: $H(x,p) = p_1 - (p_2)^2/4$, $\widehat{U}(x,p) = \{-p_2/2\}$ and the "normal transversality set" in (2.14) is given by: $Z_\tau^1 = \{(\xi, q);\ \xi = (0, \xi_2) \in Y_1,\ q = (q_1, -2\xi_2),\ q_1, \xi_2 \in R\}$; therefore, adding the natural condition $H(\xi, q) = 0$ in (2.15) we obtain the smooth Hamiltonian system:

$$\begin{cases} x_1' = 1, & x_1(0) = 0 \\ x_2' = -1/2p_2, & x_2(0) = \xi_2 \in R, \end{cases} \quad \begin{cases} p_1' = 0, & p_1(0) = (\xi_2)^2 \\ p_2' = 0, & p_2(0) = -2\xi_2, \end{cases}$$

which "produces" the smooth Hamiltonian flow:

$$X(t;\xi_2) = (t, (t+1)\xi_2),\ P(t;\xi_2) = ((\xi_2)^2, -2\xi_2)),$$

while the corresponding V^1-function in (2.21) is given by: $V^1(t;\xi_2) = -(t+1)(\xi_2)^2$, $(t,\xi_2) \in B_0 := (-\infty,0) \times R$; therefore the "would be" proper value function in (2.24) is given by:

$$\widetilde{W}_0(y) = \begin{cases} -\frac{(y_2)^2}{y_1+1} & \text{if } y \in Y_0 \setminus (\{-1\} \times R) \\ 0 & \text{if } y = (-1,0) \end{cases} \tag{5.21}$$

and is, obviously, *neither lower nor upper semiccontinuos* at the "singular" point $y_0 = (-1,0)$ of its domain, $\widetilde{Y}_0 := Y_0 \setminus (\{-1\} \times R^*)$.

However, using a more sophisticated "verification theorem", for discontinuous value functions, (e.g. Lupulescu and Mirică, 2000) one may prove the fact that the function $\widetilde{W}_0(.)$ in (5.21) coincides with the value function in (1.3) of the restriction $\mathcal{B}|\widetilde{Y}_0$ of the problem in (5.19)-(5.20) in the restricted class, Ω_∞, of *Lipschitzian trajectories*; on the other hand, the optimality in the largest class, Ω_2, of trajectories $x(.)$ for which $x'(.) \in L^2$, remains an open problem whose (probably negative) answer seems very difficult to find.

One may note also that though the problem in (5.19)-(5.20) satisfies Hypotheses 2.1 under which **PMP** is proved, *none of the existence theorems* is valid (at least for the initial points $y \in Y_0,\ y_1 < -1$) hence Theorem 2.8 cannot be applied to obtain the optimality; in fact, this problem is simple enough to show "directly" that for initial points $y \in Y_0,\ y_1 < -1$ there does not exist an optimal control since one may prove that: $\inf_{u(.)\in\mathcal{U}_a(y)} \mathcal{C}(y;u(.)) = -\infty\ \forall\ y \in Y_0 \setminus \widetilde{Y}_0$.

6. Conclusions

The results and examples above allow us to conclude that the representation in the form in (1.6) of the value function and its "validation" either by the use of **PMP** or by the application of a suitable "verification theorem" is possible for significant classes of optimal control problems; moreover, the formula in (1.6) (which may lead also to different types of "Hopf-Lax formulas" for solutions of certain types of Hamilton-Jacobi equations) may solve a multitude of problems regarding not only the complete and rigorous solutions of the optimal control problems but also problems concerning numerical solutions, the identification of regularity and/or irregularity properties of the value functions, the identification of the "relevant restrictions" of the problems, etc.

References

AUBIN, J.P. and FRANKOWSKA, H. (1990) *Set-Valued Analysis*. Birkhäuser, Boston.

BARDI, M. and CAPUZZO-DOLCETTA, I. (1997) *Optimal Control and Viscosity Solutions for Hamilton-Jacobi-Bellman Equations*. Birkhäuser, Boston.

BOLTIANSKY, V.G. (1968) *Mathematical Methods of Optimal Control.* Nauka, Moskva, in Russian (English translation: Holt & Rinehart, New York, 1971).

CESARI, L. (1983) *Optimization—Theory and applications.* Springer-Verlag, New York.

LUPULESCU, V. and MIRICĂ, Ş. (2000) Verification theorems for discontinuous value functions in optimal control. *Mathematical Reports*, **2** (52), 299–326.

MIRICĂ, Ş. (1987) Generalized solutions by Cauchy's Method of Characteristics. *Rend. Sem. Mat. Univ. Padova*, **77**, 317–349.

MIRICĂ, Ş. (1992a) New proof and some generalizations of the minimum principle in optimal control. *J. Opt. Theory Appl.*, **74**, 487–508.

MIRICĂ, Ş. (1992b) *Verification theorems of Dynamic Programming type in optimal control.* In: *Optimization, Optimal Control and PDE*, V.Barbu et al., eds., Birkhäuser, Boston, 181–192.

MIRICĂ. Ş. (1995) Invariance and monotonicity with respect to autonomous differential inclusions. *Studii Cerc. Mat.*, **44**, 49–60.

MIRICĂ, Ş. (1996) A Dynamic Programming solution to Bernoulli's Brachistochrone problem. *Bull. Math. Soc. Sci. Math. Roumanie*, **39** (87), 201–222.

MIRICĂ, Ş. (1998) Hamilton–Jacobi equations on possibly non-symplectic differentiable manifolds. *Bull. Math. Soc. Sci. Math. Roumanie*, **41** (87), 23–39.

MIRICĂ, ŞT. (2000) Dynamic Programming solution to minimal surfaces of revolution. *Bull. Math. Soc. Sci. Math. Roumanie*, **43** (91), 67–98.

SUBBOTIN, A.I. (1995) *Generalized Solutions of First Order PDEs; The Dynamical Optimization Perspective.* Birkhäuser, Boston.

SUSSMANN, H. (1990) Synthesis, presynthesis, sufficient conditions for optimality and subanalytic sets. In: *Nonlinear Controllability and Optimal Control*, H. Sussmann, ed., Marcel Dekker, New York, 1–19.