Rend. Sem. Mat. Univ. Pol. Torino - Vol. 63, 4 (2005)
Control Theory and Stabil., I

M. Guerra[*]

DISCONTINUOUS HAMILTONIAN FLOWS FOR NONLINEAR CONTROL SYSTEMS

Abstract. It is known that, under appropriate commutativity assumptions, smooth control systems that are affine with respect to controls can be extended into classes of generalized controls that contain impulses. A version of Pontryagin's maximum principle that applies in such cases as been obtained but it operates on a reduced system. We show how to construct generalized Hamiltonian trajectories for the original system in a way consistent with that version of Pontryagin's maximum principle. By lifting both the continuous and the discontinuous components of candidate optimal trajectories into the cotangent bundle this construction allows for an attractive geometric description of extremal trajectories.

1. Introduction and statement of the problem

Consider a nonlinear control system affine with respect to inputs

$$\dot{x}(t) = Y(x(t)) + \sum_{i=1}^{k} X_i(x(t))\, u_i(t)\,, \tag{1}$$

where the vector fields $Y(x)$, $X_i(x)$, $i = 1, 2, ..., k$ are smooth. The state space is $\mathbb{R}^n$ and the set of admissible controls is $L^k_{\infty,loc}(\mathbb{R})$ (the space of measurable locally essentially bounded functions with domain $\mathbb{R}$ and range $\mathbb{R}^k$). For a control system of this class, the trajectory corresponding to a given control $u \in L^k_{\infty,loc}(\mathbb{R})$ and a given initial condition, $x(0) = \overline{x}$, is uniquely defined in some open interval containing the point $t = 0$. Let $x_{u,\overline{x}}$ denote this trajectory, defined in the maximal interval. For fixed $\overline{x} \in \mathbb{R}^n$, consider the set of points which are accessible from $\overline{x}$ through trajectories of system (1) in a given time $T > 0$

$$\mathcal{A}(\overline{x}, T) = \left\{ x_{u,\overline{x}}(T)\,, u \in L^k_\infty[0, T] \right\}.$$

We wish to characterize the set $\partial\mathcal{A}(\overline{x}, T)$ (the boundary of $\mathcal{A}(\overline{x}, T)$, with $\overline{x}$ and T fixed). Pontryagin's maximum principle is a powerful tool to characterize the set $\mathcal{A}(\overline{x}, T) \cap \partial\mathcal{A}(\overline{x}, T)$: this must be contained in the set of endpoints of extremal trajectories with initial state $x(0) = \overline{x}$. However, it has some important shortcomings:

1. For an affine system (1), the maximum condition does not yield immediately a unique control in the form of a feedback function of state and adjoint vector;
2. The maximum principles applies to trajectories whose end point lies in the set $\partial\mathcal{A}(\overline{x}, T)$. Hence it can't be used to search for points in $\partial\mathcal{A}(\overline{x}, T) \setminus \mathcal{A}(\overline{x}, T)$. However, $\mathcal{A}(\overline{x}, T)$ is not, in general, closed.

---

[*] This article was partially supported by POCTI/MAT/41683/01 of FCT

The first of these difficulties can be partially overcome by the technique of so-called *Dirac's constraints* [2]. Andrei Sarychev [4] proved that, under suitable commutativity assumptions, system (1) can be extended by continuity into a space of generalized controls that includes impulsive controls. Typically, at least some points in $\partial\mathcal{A}(\overline{x},T)\setminus\mathcal{A}(\overline{x},T)$ become accessible when we consider such generalized controls. The fact that the extension of the system is obtained by continuity means that if we know a generalized control that steers $\overline{x}$ to $\overline{\overline{x}}\in\partial\mathcal{A}(\overline{x},T)\setminus\mathcal{A}(\overline{x},T)$ then we "know" which L_∞^k-controls steer $\overline{x}$ to points close to $\overline{\overline{x}}$. A. Sarychev also provided a version of the maximum principle that applies to generalized controls [4][1]. These results are an important improvement with respect to the second difficulty indicated above.

In this paper we combine the technique of Dirac's constraints with Sarychev's generalized controls to obtain some useful geometric properties of Sarychev extremals. We prove that every classical extremal that satisfies $x_{u,\overline{x}}(T)\in\partial\mathcal{A}(\overline{x},T)$ is also an extremal in the sense of Sarychev, but a classical extremal such that $x_{u,\overline{x}}(T)\notin\partial\mathcal{A}(\overline{x},T)$ may fail to be an extremal in the sense of Sarychev. We give a characterization of the classical extremals that are also Sarychev extremals. Our results also allow to compute the Sarychev extremals and/or to characterize their qualitative structure using computations that are simpler then those involved in the direct application of Sarychev's version of the maximum principle. In the end of the paper we provide an example in which these tools are used to describe the qualitative structure of the extremals in an optimal control problem.

This paper contains a partial generalization of results published in [3] concerning the case of linear-quadratic optimal control problems of arbitrary order of singularity. The generalization is only partial because here we deal essentially with the nonlinear analogous to L-Q problems whose order of singularity equals one. Research is being carried in order to obtain similar results for other types of singularity.

The results presented in this paper can be readily extended to systems where the fields Y, X_i, are time variant and/or systems whose state space is a connected smooth manifold of finite dimension. However, for the sake of simplicity, we only present the autonomous case with state space $\mathbb{R}^n$.

2. Notation, basic definitions and assumptions

In this paper, $\phi : L_{1,loc}^k(\mathbb{R}) \mapsto L_{1,loc}^k(\mathbb{R})$, denotes the "primitivation" operator, i.e., $\phi u(t)=\int_0^t u(\tau)\,d\tau$, $\forall t\in\mathbb{R}$, $\forall u\in L_{1,loc}^k(\mathbb{R})$.

A smooth vector field, F, is identified with the operator $F : C_\infty \mapsto C_\infty$ defined by $(Fg)(x)=Dg(x)\cdot F(x)$. A (local) diffeomorphism, $P:\mathbb{R}^n\mapsto\mathbb{R}^n$ generates an operator, AdP, acting in the space of smooth vector fields and defined as $(AdPF)\,g=F\left(g\circ P^{-1}\right)\circ P$, for all $g\in C_\infty$. In any system of local coordinates, the vector field $AdPF$ can be represented as $(AdPF)(x)=(DP(x))^{-1}F(P(x))$. We also use the Lie bracket, defined in the usual way: $[F,G]\,h=F(Gh)-G(Fh)$. In local coordinates this is $[F,G](x)=DG(x)\cdot F(x)-DF(x)\cdot G(x)$.

For brevity we will indicate the sum $\sum_{i=1}^{k} X_i u_i$ as Xu. Thus system (1) is indicated by $\dot{x} = Y + Xu$, being always understood that u is k-dimensional.

Let F_t denote a time-variant vector field in $\mathbb{R}^n$. $\Phi_t^{F_\tau\, d\tau}$ denotes the (local) flow generated by the time-variant vector field F_t. i.e., the map $(t, \overline{x}) \mapsto \Phi_t^{F_\tau\, d\tau}\overline{x}$ solves the differential equation

$$\frac{d}{dt}\Phi_t^{F_\tau\, d\tau}\overline{x} = F_t\left(\Phi_t^{F_\tau\, d\tau}\overline{x}\right), \qquad \Phi_0^{F_\tau\, d\tau}\overline{x} = \overline{x}. \tag{2}$$

This should not be confused with the (local) flow $\Phi_t^{F_\theta\, d\tau}$, which is the unique solution of the autonomous differential equation

$$\frac{d}{dt}\Phi_t^{F_\theta\, d\tau}\overline{x} = F_\theta\left(\Phi_t^{F_\theta\, d\tau}\overline{x}\right), \qquad \Phi_0^{F_\theta\, d\tau}\overline{x} = \overline{x}, \tag{3}$$

where θ acts as a fixed parameter, independent of the "time" variable. If these flows are well defined at least locally, then, for each sufficiently small $t \in \mathbb{R}$, the map $\overline{x} \mapsto \Phi_t^{F_\theta\, d\tau}\overline{x}$ is a local diffeomorphism.

A (possibly time-variant) vector field, F_t, generates an Hamiltonian function with domain in the cotangent bundle, $T^*\mathbb{R}^n \sim \mathbb{R}^{2n}$, defined by $h_{F_t}(x, \xi) = \xi F_t(x)$. An Hamiltonian function, h_{F_t}, defines an Hamiltonian vector field, $\overrightarrow{h}_{F_t}$, represented in local coordinates by $\overrightarrow{h}_{F_t}(x, \xi) = \left(\frac{\partial h_{F_t}}{\partial \xi}(x, \xi), -\frac{\partial h_{F_t}}{\partial x}(x, \xi)\right)$. We use these definitions also in the case when F depends on a control, i.e., a control system $\dot{x} = F_t(x, u)$ defines an Hamiltonian function, $h_{F_t(\cdot,u)}(x, \xi) = \xi F_t(x, u)$, that depends jointly on the state, adjoint vector, time and control. The Hamiltonian system generated by a control system is, from this point of view, a new control system,

$$\dot{x}(t) = \frac{\partial h_{F_t(\cdot,u(t))}}{\partial \xi}(x(t), \xi(t)), \qquad \dot{\xi}(t) = -\frac{\partial h_{F_t(\cdot,u(t))}}{\partial x}(x(t), \xi(t)), \tag{4}$$

whose trajectories lie in the cotangent bundle, $T^*\mathbb{R}^n \sim \mathbb{R}^{2n}$. In this perspective, the maximum principle states that the trajectories whose end-points lie in $\partial\mathcal{A}(\overline{x}, T)$ are to be found among the projections into the state space of the trajectories of the Hamiltonian system (4) that satisfy the maximum condition

$$h_{F_t(\cdot,u(t))}(x(t), \xi(t)) = \max_{v} h_{F_t(\cdot,v)}(x(t), \xi(t)), \qquad a.e.\, t \in [0, T]. \tag{5}$$

We will show how this point of view can be useful to deal with generalized controls and generalized trajectories.

Through all our paper we will assume that the following assumptions hold:

A1: The fields $Y, X_1, X_2, ..., X_k$ are complete, i.e., $\Phi_t^{Y\, d\tau}x$, $\Phi_t^{X_i\, d\tau}x$, $i = 1, 2, ..., k$ are uniquely defined for all $t \in \mathbb{R}$, $x \in \mathbb{R}^n$.

A2: The fields X_i, $i = 1, 2, ..., k$ commute, i.e., $\left[X_i, X_j\right] \equiv 0$, $\forall i, j \in \{1, 2, ..., k\}$.

3. Dirac's constraints

In the case of an affine control system (1), the maximum condition (5) reduces to

$$\xi(t)\, X_i(x(t)) = 0, \qquad \forall t \in [0,T],\ i = 1,2,...,k. \tag{6}$$

These are what Dirac [2] called the *primary constraints* of the system. These constraints imply

$$\frac{d^j}{dt^j}\xi(t)\, X_i(x(t)) = 0, \qquad \forall t \in [0,T],\ i = 1,2,...,k,\ j \in \mathbb{N}. \tag{7}$$

These are called *Dirac's secondary constraints of* j^{th} *order*.

The trajectories which satisfy the Pontryagin maximum principle are projections into the state space of trajectories of the Hamiltonian system (4) that satisfy *all* Dirac's constraints. Since we will need to distinguish these trajectories from "generalized trajectories" that satisfy a special version of the maximum principle, we introduce the following definition:

DEFINITION 1. *A trajectory of system* (1) *is a classical extremal if it is the projection into the state space of some trajectory of the Hamiltonian system* (4) *which satisfies all Dirac's constraints* (6, 7).

In favorable cases, the set of primary and secondary constraints can be reduced to a set of equations of the type $\Psi_j(x,\xi) = 0,\ j = 1,2,...,m,\ u = \omega(x,\xi)$. Hence the maximum principle can be used in much the same way as in the cases where it gives immediately the control as a unique feedback function of the state and adjoint vector. The main difference is that in system (1), the Hamiltonian flow is restricted to the Dirac set:

$$\mathcal{D} = \left\{(x,\xi) \in \mathbb{R}^{2n} : \Psi_j(x,\xi) = 0,\ j = 1,2,...,m\right\}.$$

The first and second-order secondary constraints are specially important for our results. It can be easily checked that, due to assumption **A2**, they are, respectively

$$\xi(t)\,[Y, X_i]\,(x(t)) = 0, \qquad \forall t \in [0,T],\ i = 1,2,...,k; \tag{8}$$

$$\xi(t)\,[Y,[Y,X_i]]\,(x(t)) + \sum_{l=1}^{k} \xi(t)\,[X_l,[Y,X_i]]\,(x(t))\, u_l(t) = 0, \qquad a.e.\, t \in [0,T],\ i = 1,2,...,k. \tag{9}$$

4. Generalized controls

Here we provide a short outline of Sarychev's construction of the spaces of generalized controls and generalized trajectories. For details, see [4]. The following Theorem is fundamental

THEOREM 1. *Let* $Y_t(x)$ *denote a time-variant vector field, smooth with respect to* x*, absolutely continuous with respect to* t*, and let* $X_t(x) = \left(X_{1,t}(x), X_{2,t}(x), ..., X_{k,t}(x)\right)$ *denote an array of* k *time-variant complete vector fields, smooth with respect to* x*, absolutely continuous with respect to* t*. Then, for each* $u = (u_1, u_2, ..., u_k) \in L^k_{1,loc}$*, and every sufficiently small* t*, we have*

$$\Phi_t^{Y_\tau + X_\tau u(\tau)\ d\tau} =$$
$$= \Phi_1^{X_t \phi u(t)\ d\tau}\, \Phi_t^{Y_\tau + \int_0^1 Ad\Phi_{\theta_1}^{X_\tau \phi u(\tau)\ d\theta_2} \left([X_\tau \phi u(\tau), Y_\tau + \theta_1 X_\tau u(\tau)] - \dot{X}_\tau \phi u(\tau)\right)\ d\theta_1\ d\tau}.$$

Now, fix an arbitrary $T > 0$ and consider the space $L^k_\infty[0,T]$ provided with the norm $\|u\|_1 = \|\phi u\|_{L^k_1[0,T]} + |\phi u(T)|$. This norm is weaker then the usual L_1-norm. For each $\alpha \in\ ]0,+\infty[$, let U_α denote the topological completion of the set $\left\{u \in L^k_\infty[0,T] : \|\phi u\|_{L^k_\infty[0,T]} \le \alpha\right\}$, with respect to the norm $\|\cdot\|_1$. Under assumptions **A1** and **A2**, Theorem 1 has the following Corollary (see Theorem 4.1 in [4]).

COROLLARY 1. *For each* $u = (u_1, u_2, ..., u_k) \in L^k_1[0,T]$*, and every* $t \in [0,T]$*, the corresponding flow of system* (1) *can be represented as*

$$\Phi_t^{Y + Xu(\tau)\ d\tau} = \Phi_1^{X\phi u(t)\ d\tau}\, \Phi_t^{Y - \int_0^1 Ad\Phi_{\theta_1}^{X\phi u(\tau)\ d\theta_2}[Y, X\phi u(\tau)]\ d\theta_1\ d\tau}. \tag{10}$$

For every $\overline{x} \in \mathbb{R}^n$ *and every* $\alpha \in\ ]0,+\infty[$*, the transformation* $u \mapsto \Phi_{(\cdot)}^{Y+Xu(\tau)\ d\tau}\overline{x}$ *is a uniformly continuous map from* U_α *into* $L^n_1[0,T]$*.*

For brevity, let G denote the controlled vector field $G(x,v) = -\int_0^1 Ad\Phi_{\theta_1}^{Xv\ d\theta_2}[Y, Xv](x)\ d\theta_1$, $v \in \mathbb{R}^k$. Corollary 1 shows that there exists one unique extension of the map $u \mapsto \Phi_{(\cdot)}^{Y+Xu(\tau)\ d\tau}\overline{x}$ into the space $U = \bigcup_{\alpha \in]0,+\infty[} U_\alpha$ that is continuous in U_α, for every sufficiently large $\alpha > 0$. Thus, we have the following Definition:

DEFINITION 2. *Any element of* U *is called a generalized control. The generalized trajectory corresponding to a given generalized control,* $u \in U$*, with a given initial condition,* $x(0) = \overline{x}$*, is the function*

$$x_{u,\overline{x}}(t) = \Phi_1^{X\phi u(t)\ d\tau}\, \Phi_t^{Y + G(\cdot,\phi u(\tau))\ d\tau}\overline{x}.$$

It can be shown that U is a subspace of the Sobolev space $W_{-1,\infty}$. Notice that, for $u \in U \backslash L^k_1[0,T]$, $x_{u,\overline{x}}(t)$ is defined only as an L_1-function (i.e., for almost every $t \in [0,T]$) but the value of $\Phi_T^{Y+Xu(\tau)\ d\tau}\overline{x}$ is uniquely defined because the norm $\|\cdot\|_1$ distinguishes different values of $\phi u(T)$. Let $\widetilde{\mathcal{A}}(\overline{x},T) = \left\{\Phi_T^{Y+Xu(\tau)\ d\tau}\overline{x} : u \in U\right\}$, denote the set of points which are accessible from the point $\overline{x}$ in time T, through generalized trajectories. Since every ordinary trajectory is also a generalized trajectory, it is clear that $\mathcal{A}(\overline{x},T) \subset \widetilde{\mathcal{A}}(\overline{x},T)$. The continuity of the map $u \mapsto \Phi_{(\cdot)}^{Y+Xu(\tau)\ d\tau}\overline{x}$ implies that $\widetilde{\mathcal{A}}(\overline{x},T) \subset \overline{\mathcal{A}(\overline{x},T)}$.

5. Maximum principle for generalized controls

Sarychev proved a version of the maximum principle that applies to the time-optimal generalized trajectories of an affine time-variant control system [4],[1]. Below we present the corresponding version for the accessibility problem. The proof in [4] can easily be adapted to our case.

THEOREM 2. *Let $\widehat{u} \in U$, such that $x_{\widehat{u},\overline{x}}(T) \in \partial\widetilde{\mathcal{A}}(\overline{x}, T)$. Then, there exists an adjoint vector $\overline{\xi} \in \mathbb{R}^n$ such that the trajectory of the Hamiltonian system $\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{Y+G(\cdot,\phi\widehat{u})}(x, \xi)$, $(x(0), \xi(0)) = \left(\overline{x}, \overline{\xi}\right)$ satisfies*

1. $h_{Y+G(\cdot,\phi\widehat{u}(t))}(x(t), \xi(t)) = \max_{\upsilon\in\mathbb{R}^k} h_{Y+G(\cdot,\upsilon)}(x(t), \xi(t))$, *a.e.* $t \in [0, T]$;
2. $\xi(T)\, X(x(T)) = 0$.

Sketch of the proof. Let $\widehat{u} \in U$, such that $x_{\widehat{u},\overline{x}}(T) \in \partial\widetilde{\mathcal{A}}(\overline{x}, T)$. From Corollary 1, we have $x_{\widehat{u},\overline{x}}(T) = \Phi_1^{X\phi\widehat{u}(T)}\Phi_T^{Y+G(\cdot,\phi\widehat{u}(\tau))\,d\tau}\overline{x}$. Since $x \mapsto \Phi_1^{X\phi\widehat{u}(T)}x$ is a diffeomorphism, it follows that $\Phi_T^{Y+G(\cdot,\phi\widehat{u}(\tau))\,d\tau}\overline{x}$ must lie in the boundary of the accessible set of the system

$$\dot{x} = Y(x) + G(x, \upsilon), \qquad x(0) = \overline{x}, \qquad \upsilon \in L_\infty^k[0, T]. \tag{11}$$

Let K denote the Pontryagin cone for system (11) at the point $\Phi_T^{Y+G(\cdot,\phi\widehat{u}(\tau))\,d\tau}\overline{x}$, and let S denote the tangent space to the integral manifold of X at the point $\Phi_T^{Y+G(\cdot,\phi\widehat{u}(\tau))\,d\tau}\overline{x}$. Now, $\widehat{u} \in U$, implies that $\phi\widehat{u}$ is defined only as a function of class $L_\infty^k[0, T]$, i.e., only "almost every where". This means that $\phi\widehat{u}(T)$ (the value of $\phi\widehat{u}$ at the particular point $t = T$) can be chosen independently of $\phi\widehat{u}$ as an element of $L_\infty^k[0, T]$. It follows that $K + S$ is an approximation of $\widetilde{\mathcal{A}}(\overline{x}, T)$ in the neighborhood of $x_{\widehat{u},\overline{x}}(T)$. The proof follows by a procedure analogous to the classical proof of the Pontryagin maximum principle, by using $K + S$ instead of the Pontryagin cone of system (1), which may fail to exist at the point $x_{\widehat{u},\overline{x}}(T)$. ☐

This version of the maximum principle gives a second definition of extremal trajectory.

DEFINITION 3. *A generalized trajectory, $x \in L_1^n[0, T]$ is a Sarychev extremal with initial condition $x(0) = \overline{x}$ if it can be represented in the form*

$$x(t) = \Phi_1^{X\phi u(t)\,d\tau}\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\overline{x},$$

for some $u \in U$ which satisfies the conditions (1) *and* (2) *of Theorem 2.*

We will show that Definition 1 includes curves that are not Sarychev extremals. Hence, Theorem 2 is a necessary condition that is stronger than the classical maximum principle.

6. Generalized Hamiltonian flows

Theorem 2 states that we should look for extremals of the reduced system, $\dot{x} = Y(x) + G(x, v)$ and then recover the extremals for the original system (1) by using equality (10). Now we show that the extension of system (1) into the class of generalized controls induces naturally an extension of the Hamiltonian flow (understood as the flow of the Hamiltonian system (4)) into the same class of generalized controls. This gives useful insights into the structure of Sarychev extremals. It also gives a method to compute the Sarychev extremals without having to compute explicitly the field G.

Any smooth map, $\varphi : \mathbb{R}^m \mapsto \mathbb{R}^n$, generates a map, $T\varphi$, that assigns to each vector $V \in T_x\mathbb{R}^n$, a new vector, $T\varphi V \in T_{\varphi(x)}\mathbb{R}^n$, defined by $(T\varphi V)\, f = V\,(f \circ \varphi)$, $\forall f \in C_\infty$. If $\varphi : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a (local) diffeomorphism, then it also defines a map, φ_*, which assigns to each adjoint vector $\xi \in T_x^*\mathbb{R}^n$, a new adjoint vector $\varphi_*\xi \in T_{\varphi(x)}^*\mathbb{R}^n$, defined by $(\varphi_*\xi)\, V = \xi T\left(\varphi^{-1}\right) V$, $\forall V \in T_{\varphi(x)}\mathbb{R}^n$. $\varphi_*\xi$ is called the *pushforward* of ξ by the map φ.

LEMMA 1. *Let $F_t\,(x)$ denote a time-variant vector field, smooth with respect to x, absolutely continuous with respect to t. Then,*

$$\Phi_t^{\overrightarrow{h}_{F_\tau}\,d\tau} = \left(\Phi_t^{F_\tau\,d\tau}\right)_*.$$

Proof. Consider a fixed $\overline{x} \in \mathbb{R}^n$, $\overline{\xi} \in T_{\overline{x}}\mathbb{R}^n$. In local coordinates, we have $\left(\Phi_t^{F_\tau\,d\tau}\right)_* \overline{\xi} = \overline{\xi} D_x \Phi_t^{F_\tau\,d\tau}\overline{x}$. Use the Theorem of differentiability of the solution of an ODE with respect to initial conditions to prove that the pair $\left(\Phi_t^{F_\tau\,d\tau}\overline{x}, \overline{\xi} D_x \Phi_t^{F_\tau\,d\tau}\overline{x}\right)$ is the unique solution of the Hamiltonian system $\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{F_t}\,(x, \xi)$, with initial point $(x\,(0)\,, \xi\,(0)) = \left(\overline{x}, \overline{\xi}\right)$. □

Using Corollary 1, we obtain a representation of the Hamiltonian flow in the form of a composition of flows that depends only on the primitives of the controls.

LEMMA 2. *For any $u \in L_1^k\,[0, T]$, we have*

$$\Phi_t^{\overrightarrow{h}_{Y+Xu(\tau)}\,d\tau} = \Phi_1^{\overrightarrow{h}_{X\phi u(t)}\,d\tau}\,\Phi_t^{\overrightarrow{h}_{Y+G(\cdot,\phi u(\tau))}\,d\tau}.$$

For any fixed $\left(\overline{x}, \overline{\xi}\right) \in \mathbb{R}^{2n}$ and any $\alpha \in\,]0, +\infty[$, the map $u \mapsto \Phi_t^{\overrightarrow{h}_{Y+Xu(\tau)}\,d\tau}\left(\overline{x}, \overline{\xi}\right)$ is uniformly continuous from U_α into $L_1^{2n}\,[0, T]$.

Proof. Lemma 1 states that $\Phi_t^{\overrightarrow{h}_{Y+Xu(\tau)}\,d\tau} = \left(\Phi_t^{Y+Xu(\tau)\,d\tau}\right)_*$. By Corollary 1, this is $\Phi_t^{\overrightarrow{h}_{Y+Xu(\tau)}\,d\tau} = \left(\Phi_1^{X\phi u(t)\,d\tau}\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\right)_*$. Since $(\varphi \circ \psi)_* = \varphi_* \circ \psi_*$, this reduces to $\Phi_t^{\overrightarrow{h}_{Y+Xu(\tau)}\,d\tau} = \left(\Phi_1^{X\phi u(t)\,d\tau}\right)_* \left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\right)_*$, and the equality follows by using again Lemma 1. The continuity of the map follows immediately from Corollary 1. □

It follows from Lemma 2 that the map $u \mapsto \Phi_{(\cdot)}^{\overrightarrow{h}_{Y+Xu(\tau)}\,d\tau}\overline{x}$ has one unique extension into the space U. Thus, the generalized Hamiltonian flow corresponding to a given generalized control, $u \in U$, is defined in the same way as the generalized flow introduced in Definition 2.

We will now show that Lemma 2 provides a link between Theorem 2 and Dirac's constraints. We start with the following Proposition which shows that the primary constraints can be represented using the flow of the reduced Hamiltonian system, $\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{Y+G(\cdot,\phi\widehat{u})}\left(x, \xi\right)$.

PROPOSITION 1. *Let $u \in U$. The Dirac's constraints are satisfied at every point of the trajectory $t \mapsto \Phi_t^{\overrightarrow{h}_{Y+Xu}\,d\tau}\left(\overline{x}, \overline{\xi}\right)$ if and only if they are satisfied at every point of the trajectory $t \mapsto \Phi_t^{\overrightarrow{h}_{Y+G(\cdot,\phi u)}\,d\tau}\left(\overline{x}, \overline{\xi}\right)$.*

Proof. Using Lemma 1, we have

$$\xi\left(t\right) X_i\left(x\left(t\right)\right) = \left(\left(\Phi_t^{Y+Xu(\tau)\,d\tau}\right)_* \overline{\xi}\right) X_i \left(\Phi_t^{Y+Xu(\tau)\,d\tau}\overline{x}\right).$$

Using Lemma 2, this reduces to

$$\begin{aligned}
&\xi\left(t\right) X_i\left(x\left(t\right)\right) = \\
&= \left(\left(\Phi_1^{X\phi u(t)\,d\tau}\right)_* \left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\right)_* \overline{\xi}\right) X_i \left(\Phi_1^{X\phi u(t)\,d\tau}\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\overline{x}\right) = \\
&= \left(\left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\right)_* \overline{\xi}\right)\left(Ad\Phi_1^{X\phi u(t)\,d\tau}X_i\right)\left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\overline{x}\right).
\end{aligned}$$

Since assumption **A2** implies $Ad\Phi_1^{X\phi u(t)\,d\tau}X_i = X_i$, this proves the Proposition. □

Proposition 1 has the following Corollary:

COROLLARY 2. *Consider fixed $\widehat{u} \in U$, $\left(\overline{x}, \overline{\xi}\right) \in \mathbb{R}^{2n}$. Dirac's primary constraints are satisfied at every point of the trajectory of the Hamiltonian system*

$$\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{Y+X\widehat{u}}\left(x, \xi\right), \qquad \left(x\left(0\right), \xi\left(0\right)\right) = \left(\overline{x}, \overline{\xi}\right) \tag{12}$$

if and only if all the following conditions are satisfied:

1. $\overline{\xi}X_i\left(\overline{x}\right) = 0, \qquad i = 1, 2, ..., k;$
2. $\xi\left(t\right)\left[Y + G\left(\cdot, \phi\widehat{u}\left(t\right)\right), X_i\right]\left(x\left(t\right)\right) = 0$, $i = 1, 2, ..., k$ *hold at almost every* $t \in [0, T]$, *along the trajectory of the reduced Hamiltonian system*

   $$\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{Y+G(\cdot,\phi\widehat{u})}\left(x, \xi\right), \qquad \left(x\left(0\right), \xi\left(0\right)\right) = \left(\overline{x}, \overline{\xi}\right). \tag{13}$$

Proof. Proposition 1 states that a trajectory of system (12) satisfies Dirac's primary constraints if and only if the corresponding trajectory of system (13) satisfies them. Since

$$\begin{aligned}&\tfrac{d}{dt}\left(\left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\right)_*\overline{\xi}\right) X_i\left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\overline{x}\right) =\\ &=\left(\left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\right)_*\overline{\xi}\right)\left[Y+G\left(\cdot,\phi u\left(t\right)\right),X_i\right]\left(\Phi_t^{Y+G(\cdot,\phi u(\tau))\,d\tau}\overline{x}\right),\end{aligned}$$

we proved the Corollary. □

The next Lemma, together with Corollary 2, shows that classical extremals correspond to critical points of the reduced Hamiltonian function, $h_{Y+G(\cdot,v)}$. These may happen to be minima, saddle points or just local maxima instead of global maxima as required in Theorem 2 (see example in next Section).

LEMMA 3. *$\frac{\partial}{\partial v_i}G\left(x,v\right)=\left[X_i,Y+G\left(\cdot,v\right)\right]\left(x\right)$, for every $\left(x,v\right)\in\mathbb{R}^{n+k}$, $i=1,2,...,k$.*

Proof. Assumption **A2** implies that

$$\Phi_{\theta_1}^{X\cdot(v+\delta v)\ d\theta_2}=\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\Phi_{\theta_1}^{X\cdot\delta v\ d\theta_2}.$$

It follows that,

$$\begin{aligned}G\left(x,v+\delta v\right)&=-\int_0^1 Ad\left(\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\Phi_{\theta_1}^{X\cdot\delta v\ d\theta_2}\right)\left[Y,X\cdot\left(v+\delta v\right)\right]\left(x\right)d\theta_1=\\ &=-\int_0^1 Ad\Phi_{\theta_1}^{X\cdot\delta v\ d\theta_2}\left(Ad\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\left[Y,X\cdot\left(v+\delta v\right)\right]\right)\left(x\right)d\theta_1=\\ &=-\int_0^1\left(Id-\theta_1 D\left(X\cdot\delta v\right)\left(x\right)+o\left(\delta v\right)\right)\cdot\\ &\qquad\left(Ad\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\left[Y,X\cdot\left(v+\delta v\right)\right]\right)\left(x+\theta_1 X\left(x\right)\cdot\delta v+o\left(\delta v\right)\right)\ d\theta_1=\\ &=-\int_0^1 Ad\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\left[Y,X\cdot\left(v+\delta v\right)\right]\left(x\right)d\theta_1-\\ &\qquad-\int_0^1\theta_1 D\left(Ad\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\left[Y,X\cdot v\right]\right)\cdot X\left(x\right)\cdot\delta v d\theta_1+\\ &\qquad+\int_0^1\theta_1 D\left(X\cdot\delta v\right)\left(x\right)\cdot Ad\Phi_{\theta_1}^{X\cdot v\ d\theta_2}\left[Y,X\cdot v\right]\left(x\right)\ d\theta_1+o\left(\delta v\right)=\end{aligned}$$

$$
\begin{aligned}
&= G(x,v) - \int_0^1 Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2}\left[Y, X\cdot\delta v\right](x)\; d\theta_1 + \\
&\quad + \int_0^1 \theta_1 \left[Ad\Phi_{\theta_1}^{X\cdot v d\theta_2}\left[Y, X\cdot v\right], X\cdot\delta v\right](x) d\theta_1 + o(\delta v) = \\
&= G(x,v) - \int_0^1 \left[Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2} Y, Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2}\left(X\cdot\delta v\right)\right](x)\; d\theta_1 + \\
&\quad + \int_0^1 \theta_1 \left[Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2}\left[Y, X\cdot v\right], X\cdot\delta v\right](x)\, d\theta_1 + o(\delta v) = \\
&= G(x,v) - \int_0^1 \left[Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2} Y, X\cdot\delta v\right](x)\; d\theta_1 + \\
&\quad + \int_0^1 \theta_1 \left[Ad\Phi_{\theta_1}^{X\cdot v d\theta_2}\left[Y, X\cdot v\right], X\cdot\delta v\right](x)\; d\theta_1 + o(\delta v) = \\
&= G(x,v) + \left[X\cdot\delta v, \int_0^1 Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2} Y\; d\theta_1\right](x) - \\
&\quad - \left[X\cdot\delta v, \int_0^1 \theta_1 Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2}\left[Y, X\cdot v\right]\; d\theta_1\right](x) + o(\delta v) = \\
&= G(x,v) + \left[X\cdot\delta v, Y + \int_0^1 (1-\theta_1)\, Ad\Phi_{\theta_1}^{X\cdot v\; d\theta_2}\left[X\cdot v, Y\right] d\theta_1\right](x) - \\
&\quad - \left[X\cdot\delta v, \int_0^1 \theta_1 Ad\Phi_{\theta_1}^{X\cdot v d\theta_2}\left[Y, X\cdot v\right]\; d\theta_1\right](x) + o(\delta v) = \\
&= G(x,v) + \left[X\cdot\delta v, Y + G(\cdot, v)\right](x) + o(\delta v).
\end{aligned}
$$

□

To explore further the relationship between Dirac's constraints and Sarychev extremals, we will use the following Proposition that states an important invariance property of the integral manifolds of the distribution $\overrightarrow{h}_{X_i}$, $i = 1, 2, ..., k$.

PROPOSITION 2. *For every* $(x, \xi) \in \mathbb{R}^{2n}$, $v, w \in \mathbb{R}^k$, $j \in \mathbb{N}$, $i_1, i_2, ..., i_j \in \{1, 2, ..., k\}$, *we have:*

$$
\begin{aligned}
&h_{Y+G(\cdot,v+w)}(x,\xi) = h_{Y+G(\cdot,v)} \circ \Phi_1^{\overrightarrow{h}_{Xw}\, d\tau}(x,\xi); \\
&h_{\left[X_{i_j},\dots\left[X_{i_2},\left[X_{i_1}, Y+G(\cdot,v+w)\right]\right]\right]}(x,\xi) = \\
&\qquad\qquad = h_{\left[X_{i_j},\dots\left[X_{i_2},\left[X_{i_1}, Y+G(\cdot,v)\right]\right]\right]} \circ \Phi_1^{\overrightarrow{h}_{Xw}\, d\tau}(x,\xi).
\end{aligned}
$$

Proof. Fix arbitrary $(x, \xi) \in \mathbb{R}^{2n}$, $v, w \in \mathbb{R}^k$, and consider the function

$$
f(\varepsilon) = h_{\left[X_{i_j},\dots\left[X_{i_1}, Y+G(\cdot,v+\varepsilon w)\right]\right]} \circ \Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\, d\tau}(x,\xi).
$$

Then,

$$
\begin{aligned}
f'(\varepsilon) = {} & \frac{\partial h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]}}{\partial\varepsilon}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)+\\
&+\frac{\partial h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]}}{\partial(x,\xi)}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)\cdot\frac{\partial\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}}{\partial\varepsilon}(x,\xi)=\\
={} & h_{\left[X_{i_j},\dots\left[X_{i_1},\frac{\partial G}{\partial\varepsilon}(\cdot,v+\varepsilon w)\right]\right]}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)-\\
&-\frac{\partial h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]}}{\partial(x,\xi)}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)\cdot\\
&\qquad\cdot\frac{\partial\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}}{\partial(x,\xi)}(x,\xi)\cdot\overrightarrow{h}_{Xw}(x,\xi)=\\
={} & h_{\left[X_{i_j},\dots\left[X_{i_1},\frac{\partial G}{\partial\varepsilon}(\cdot,v+\varepsilon w)\right]\right]}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)-\\
&-\left(\frac{\partial h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]}}{\partial(x,\xi)}\cdot Ad\,\Phi_1^{\overrightarrow{h}_{X(\varepsilon-1)w}\,d\tau}\overrightarrow{h}_{Xw}\right)\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right).
\end{aligned}
$$

Using Lemma 3 and assumption **A2**, this reduces to

$$
\begin{aligned}
f'(\varepsilon) = {} & h_{\left[X_{i_j},\dots\left[X_{i_1},\left[Xw,Y+G(\cdot,v+\varepsilon w)\right]\right]\right]}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)-\\
&-\left(\frac{\partial h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]}}{\partial(x,\xi)}\cdot\overrightarrow{h}_{Xw}\right)\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right).
\end{aligned}
$$

Taking into account the fact that $h_{[F_1,F_2]}=\frac{\partial h_{F_2}}{\partial(x,\xi)}\cdot\overrightarrow{h}_{F_1}$, this is

$$
\begin{aligned}
f'(\varepsilon) = {} & h_{\left[X_{i_j},\dots\left[X_{i_1},\left[Xw,Y+G(\cdot,v+\varepsilon w)\right]\right]\right]}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right)-\\
&-h_{\left[Xw,\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]\right]}\left(\Phi_1^{\overrightarrow{h}_{X(1-\varepsilon)w}\,d\tau}(x,\xi)\right).
\end{aligned}
$$

Assumption **A2** implies

$$
\begin{aligned}
&\left[X_{i_j},\dots\left[X_{i_1},\left[Xw,Y+G(\cdot,v+\varepsilon w)\right]\right]\right]=\\
&\qquad\qquad\left[Xw,\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+\varepsilon w)\right]\right]\right],
\end{aligned}
$$

and hence f is constant. Since $f(0)=h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v)\right]\right]}\left(\Phi_1^{\overrightarrow{h}_{Xw}\,d\tau}(x,\xi)\right)$, and $f(1)=h_{\left[X_{i_j},\dots\left[X_{i_1},Y+G(\cdot,v+w)\right]\right]}(x,\xi)$, this proves the Proposition. □

Now consider a generalized control, $\widehat{u} \in U$, and a point, $\left(\overline{x}, \overline{\xi}\right) \in \mathbb{R}^{2n}$. Let $\left(x^1(t), \xi^1(t)\right)$ denote the trajectory of the reduced Hamiltonian system,

$$\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{Y+G(\cdot,\phi\widehat{u})}(x, \xi), \qquad (x(0), \xi(0)) = \left(\overline{x}, \overline{\xi}\right),$$

and let $\left(x^0(t), \xi^0(t)\right)$ denote the generalized trajectory of the original Hamiltonian system,

$$\left(\dot{x}, \dot{\xi}\right) = \overrightarrow{h}_{Y+X\widehat{u}}(x, \xi), \qquad (x(0), \xi(0)) = \left(\overline{x}, \overline{\xi}\right).$$

Suppose that $\widehat{u}$ satisfies Sarychev's version of the maximum principle (Theorem 2). Then Lemma 3 states that

$$h_{[X_i, Y+G(\cdot,\phi\widehat{u}(t))]}\left(x^1(t), \xi^1(t)\right) = 0,$$

at almost every $t \in [0, T]$. On the other hand, Proposition 2 and Lemma 2 imply that

$$\begin{aligned} h_{[X_i, Y+G(\cdot,\phi\widehat{u}(t))]}\left(x^1(t), \xi^1(t)\right) &= h_{[X_i, Y+G(\cdot,0)]} \circ \Phi_1^{\overrightarrow{h}_{X\phi\widehat{u}(t)}}\left(x^1(t), \xi^1(t)\right) = \\ &= h_{[X_i, Y]}\left(x^0(t), \xi^0(t)\right) = \\ &= \xi^0(t)\,[X_i, Y]\left(x^0(t)\right). \end{aligned}$$

This shows that Sarychev extremals must satisfy the first order secondary constraint but only at "almost every $t \in [0, T]$", the exception of a null but nonempty set being possible. This differs from the primary constraints, which must be satisfied at *every point* of the time interval.

At almost every $t \in [0, T]$, a Sarychev extremal must satisfy

$$\left.\frac{\partial^2 h_{Y+G(\cdot,v)}}{\partial v^2}\right|_{v=\phi\widehat{u}(t)}\left(x^1(t), \xi^1(t)\right) \leq 0.$$

By Lemma 3, we have

$$\frac{\partial^2 h_{Y+G(\cdot,v)}}{\partial v_i \partial v_j} = h_{[X_i[X_j, Y+G(\cdot,v)]]}.$$

Hence Proposition 2 implies that the matrix with entries

$$L_{ij} = h_{[X_i[X_j,Y]]}\left(x^0(t), \xi^0(t)\right) = \xi^0(t)\left[X_i, \left[X_j, Y\right]\right]\left(x^0(t)\right)$$

must be semidefinite negative at almost every $t \in [0, T]$. This is a version for generalized controls of the well known generalized Legendre-Clebsch condition. If the generalized Hamiltonian trajectory satisfies the strong Legendre-Clebsch condition (i.e., the matrix with entries $L_{ij} = \xi^0(t)\left[X_i, \left[X_j, Y\right]\right]\left(x^0(t)\right)$ is strictly negative at almost every $t \in [0, T]$), then it corresponds to local maxima of the reduced Hamiltonian function, $h_{Y+G(\cdot,v)}$, but it may fail to correspond to global maxima (see example below).

Hence the strong Legendre-Clebsch condition is not sufficient in order for a classical extremal to be a Sarychev extremal.

Now, let $\Phi^{\overrightarrow{h}_X}(x,\xi))$ denote the integral manifold of $\overrightarrow{h}_X$ through the point (x,ξ). By the same arguments as above, the maximum condition, $h_{Y+G(\cdot,\phi\widehat{u}(t))}\left(x^1(t),\xi^1(t)\right) = \max\limits_{v\in\mathbb{R}^k} h_{Y+G(\cdot,v)}\left(x^1(t),\xi^1(t)\right)$, reduces to $h_Y\left(x^0(t),\xi^0(t)\right) = \max\left\{h_Y(x,\xi) : (x,\xi)\in\Phi^{\overrightarrow{h}_X}\left(x^0(t),\xi^0(t)\right)\right\}$. Hence, an absolutely continuous arc, $(x(t),\xi(t))$, of a generalized Hamiltonian trajectory, satisfies the maximum condition in Theorem 2 if and only if the Dirac's constraints hold along the trajectory and

$$h_Y(x(t),\xi(t)) = \max\left\{h_Y(x,\xi) : (x,\xi)\in\Phi^{\overrightarrow{h}_X}(x(t),\xi(t))\right\} \tag{14}$$

at almost every t. Notice that, if the existence of the right handside of (14) is granted, we only have to check that

$$h_Y(x(t),\xi(t)) = \max\left\{h_Y(x,\xi) : (x,\xi)\in\mathcal{D}\cap\Phi^{\overrightarrow{h}_X}(x(t),\xi(t))\right\}$$

for almost every t.

The results above allow for a geometric description of Sarychev's extremals. In order to see this, consider an affine system (1) such that every secondary constraint of order higher than 2 is redundant (i.e., the set of Dirac's constraints reduces to (6), (8) and (9)). Define the set

$$\mathcal{W} = \left\{(x,\xi)\in\mathbb{R}^{2n} : \xi X_i(x) = 0,\ i = 1,2,...,k\right\}.$$

Then the Dirac set is

$$\mathcal{D} = \{(x,\xi)\in\mathcal{W} : \xi[Y,X_i](x) = 0,\ i = 1,2,...,k\}.$$

Suppose that, for each pair $(x,\xi)\in\mathcal{W}$, the set $\left\{v : v = \arg\max h_{Y+G(\cdot,v)}(x,\xi)\right\}$ is nonempty and finite. We will describe the generalized Hamiltonian trajectories whose projection into the state space coincide with the Sarychev extremals with initial condition $x(0) = \overline{x}$ ($\overline{x}$ fixed but arbitrary).

We can choose for initial adjoint vector any $\overline{\xi}$ such that $\left(\overline{x},\overline{\xi}\right)\in\mathcal{W}$. By choosing different $\overline{\xi}$, we are choosing different sets $\left\{v : v = \arg\max h_{Y+G(\cdot,v)}\left(\overline{x},\overline{\xi}\right)\right\}$. For each fixed $\overline{\xi}$, the choice of a $v\in\left\{v : v = \arg\max h_{Y+G(\cdot,v)}\left(\overline{x},\overline{\xi}\right)\right\}$ amounts to the choice of a point $\left(x\left(0^+\right),\xi\left(0^+\right)\right)\in\mathcal{D}\cap\Phi^{\overrightarrow{h}_X}\left(\overline{x},\overline{\xi}\right)$, which maximizes $h_Y(x,\xi) = \xi Y(x)$ over the set $\mathcal{D}\cap\Phi^{\overrightarrow{h}_X}\left(\overline{x},\overline{\xi}\right)$. If in the $\mathcal{D}$-neighborhood of $\left(x\left(0^+\right),\xi\left(0^+\right)\right)$ there exists a feedback, $u = u(x,\xi)$ which satisfies the second-order secondary constraints (9) along the curve $\Phi_t^{\overrightarrow{h}_{Y+Xu}}\left(x\left(0^+\right),\xi\left(0^+\right)\right)$ and, for every sufficiently small $t > 0$, $h_Y\circ\Phi_t^{\overrightarrow{h}_{Y+Xu}}\left(x\left(0^+\right),\xi\left(0^+\right)\right)$ is maximal among the set

$$\mathcal{D}\cap\Phi^{\overrightarrow{h}_X}\Phi_t^{\overrightarrow{h}_{Y+Xu}}\left(x\left(0^+\right),\xi\left(0^+\right)\right),$$

then we may follow the curve $\Phi_t^{\overrightarrow{h}_{Y+Xu}}\left(x\left(0^+\right),\xi\left(0^+\right)\right)$ during at least a small interval of time. If $[0,t_1]$ is the maximum interval of time where the feedback u satisfies the stated conditions, then a new jump must occur at time $t = t_1$, this time to a point $\left(x\left(t_1^+\right),\xi\left(t_1^+\right)\right) \in \mathcal{D}\cap\Phi^{\overrightarrow{h}_X}\left(x\left(t_1\right),\xi\left(t_1\right)\right)$, which maximizes $h_Y(x,\xi)$ over the set $\mathcal{D}\cap\Phi^{\overrightarrow{h}_X}\left(x\left(t_1\right),\xi\left(t_1\right)\right)$, and so on. If at some point of the generalized trajectory, $\left(x\left(\widehat{t}\right),\xi\left(\widehat{t}\right)\right)$ there exists more than one $\upsilon = \arg\max h_{Y+G(\cdot,\upsilon)}\left(x\left(\widehat{t}\right),\xi\left(\widehat{t}\right)\right)$ (i.e., there exists more than one (x,ξ) which maximizes $h_Y(x,\xi)$ over the set $\mathcal{D}\cap\Phi^{\overrightarrow{h}_X}\left(x\left(\widehat{t}\right),\xi\left(\widehat{t}\right)\right)$) then we may have alternative jumps at time $t = \widehat{t}$, corresponding to different generalized Hamiltonian trajectories. At the final time, $t = T$, we are free to choose as end-point of the trajectory any point in $\Phi^{\overrightarrow{h}_X}\left(x\left(T^-\right),\xi\left(T^-\right)\right)$ (not necessarily in $\mathcal{D}$).

7. An Example

Consider the optimal control problem

$$\begin{aligned}
&\textstyle\int_0^T x_1x_2\,dt \to \min;\\
&\dot{x}_1 = x_1 + u, \qquad \dot{x}_2 = -x_1 + x_2 + \left((x_1+1)^2+1\right)u, \qquad (15)\\
&x_1 \in AC\,[0,T], \qquad x_2 \in AC\,[0,T], \qquad u \in L_\infty\,[0,T],
\end{aligned}$$

with fixed boundary conditions, $x_1(0) = \overline{x}_1,\ x_2(0) = \overline{x}_2,\ x_1(T) = \overline{\overline{x}}_1,\ x_2(T) = \overline{\overline{x}}_2$, and fixed $T > 0$. We can consider the augmented system,

$$\dot{x}_1 = x_1 + u, \qquad \dot{x}_2 = -x_1 + x_2 + \left((x_1+1)^2+1\right)u, \qquad \dot{x}_3 = x_1x_2.$$

This can be represented in the form $\dot{x} = Y(x) + X(x)u$, with $Y(x) = (x_1,\ -x_1+x_2,\ x_1x_2)$, $X(x) = \left(1,\ (x_1+1)^2+1,\ 0\right)$. Since the problem has one single input, assumption **A2** holds trivially. The first Lie brackets are

$$\begin{aligned}
[Y,X] &= \left(-1,\ x_1^2-1, -x_1^3-2x_1^2-2x_1-x_2\right),\\
[Y\,[Y,X]] &= \left(1,\ x_1^2,\ -4x_1^2\,(x_1+1)\right),\\
[X\,[Y,X]] &= \left(0,\ 4x_1+2,\ -2\left(2x_1^2+3x_1+2\right)\right).
\end{aligned}$$

The primary constraint, first-order secondary constraint and second-order secondary constraint reduce to $\xi_1 + \xi_2\left((x_1+1)^2+1\right) = 0$, $-\xi_1+\xi_2\left(x_1^2-1\right)-\xi_3\left(x_1^3+2x_1^2+2x_1+x_2\right) = 0$ and $\xi_1+\xi_2x_1^2-4\xi_3x_1^2\,(x_1+1)+\left(\xi_2\,(4x_1+2)-2\xi_3\left(2x_1^2+3x_1+2\right)\right)u = 0$, respectively. This three constraints give u in the feedback form

$$u = -\frac{(x1+1)\left(4x_1^4+5x_1^3+4x_1^2+2x_1+x_2\right)}{2x_1^4+5x_1^3+6x_1^2+5x_1+2-(2x_1+1)x_2}. \qquad (16)$$

Thus, any further constraint is redundant and we have

$$\mathcal{W} = \left\{(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3) : \xi_1 = -\xi_2 \left((x_1+1)^2+1\right)\right\},$$

$$\mathcal{D} = \left\{(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3) : \xi_1 = -\xi_3 \frac{(x_1^2+2x_1+2)\left(x_1^3+2x_1^2+2x_1+x_2\right)}{2x_1^2+2x_1+1}, \right.$$
$$\left. \xi_2 = \xi_3 \frac{x_1^3+2x_1^2+2x_1+x_2}{2x_1^2+2x_1+1}\right\}.$$

Since the constraints are homogeneous with respect to (ξ_1, ξ_2, ξ_3), it follows that no abnormal extremal exists for this problem. Hence we can set $\xi_3 = -1$ and analyze the generalized Hamiltonian flow in the 4-dimensional space with coordinates (x_1, x_2, ξ_1, ξ_2). Since any generalized trajectory must lie in $\mathcal{W}$ and the coordinate ξ_1 for a point in $\mathcal{W}$ is uniquely defined by its coordinates (x_1, ξ_2), we can reduce further the dimension of the phase space and analyze the generalized Hamiltonian flow in the 3-dimensional space with coordinates (x_1, x_2, ξ_2). Thus, we can use the identifications $\mathcal{W} \sim \mathbb{R}^3$, $\mathcal{D} \sim \left\{(x_1, x_2, \xi_2) : \xi_2 = -\frac{x_1^3+2x_1^2+2x_1+x_2}{2x_1^2+2x_1+1}\right\}$. On the other hand,

$$\Phi_1^{\overrightarrow{h}\, X_v}(x,\xi) = \begin{pmatrix} x_1 + v \\ x_2 + \left(x_1^2+2x_1+2\right)v + (x_1+1)\,v^2 + \frac{v^3}{3} \\ x_3 \\ \xi_1 - \xi_2 v\,(v + 2x_1 + 2) \\ \xi_2 \\ \xi_3 \end{pmatrix}.$$

Hence any jump changes only the coordinates (x_1, x_2, ξ_1), leaving the coordinates (x_3, ξ_2, ξ_3) unchanged. In particular, since the coordinate x_3 represents the running cost, in this example jumps have zero cost. This is a particular case: in general the cost of one particular jump can be any real number.

For $(x, \xi) \in \mathcal{W}$, $v \in \mathbb{R}$, $h_Y \circ \Phi_1^{\overrightarrow{h}\, X_v}(x,\xi)$ is a 4^{th} degree polynomial with leading term $-\frac{v^4}{3}$, hence it has a maximum. It follows that there exist extremals starting at every point $(x_1, x_2) \in \mathbb{R}^2$.

Different extremal trajectories with initial point $(x_1(0), x_2(0)) = (\overline{x}_1, \overline{x}_2)$ can be selected by choosing different initial adjoint states, $\overline{\xi}_2$. An initial jump will transfer the (fixed) initial point, $\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right)$, to a point $\left(x_1\left(0^+\right), x_2\left(0^+\right), \xi_2\left(0^+\right)\right)$ located in the intersection of the surface

$$\xi_2 = -\frac{x_1^3 + 2x_1^2 + 2x_1 + x_2}{2x_1^2 + 2x_1 + 1}$$

with the curve

$$\mathcal{J}_{(\overline{x}_1, \overline{x}_2, \overline{\xi}_2)} = \left\{\left(\overline{x}_1 + v,\ \overline{x}_2 + \left(\overline{x}_1^2 + 2\overline{x}_1 + 2\right)v + (\overline{x}_1 + 1)\,v^2 + \frac{v^3}{3},\ \overline{\xi}_2\right),\ v \in \mathbb{R}\right\}.$$

The intersection of these two submanifolds is given by the zeros of the polynomial

$$\begin{aligned} P\left(\upsilon\right) = {} & \tfrac{4}{3}\upsilon^3 + \left(2\overline{\xi}_2 + 4\overline{x}_1 + 3\right)\upsilon^2 + \left(4\overline{x}_1^2 + 6\overline{x}_1 + 4 + \left(4\overline{x}_1 + 2\right)\overline{\xi}_2\right)\upsilon + \\ & + \overline{x}_2 + \overline{x}_1^3 + 2\overline{x}_1^2 + 2\overline{x}_1 + \left(2\overline{x}_1^2 + 2\overline{x}_1 + 1\right)\overline{\xi}_2 . \end{aligned}$$

It follows that for each $\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right) \in \mathbb{R}^3$, the set $\mathcal{J}_{\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right)} \cap \mathcal{D}$ is nonempty and has at most three different points.

Consider $\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right) \in \mathcal{D}$. Then, the polynomial P reduces to

$$P\left(\upsilon\right) = \upsilon^3 + \tfrac{3}{4}\tfrac{6\overline{x}_1^3 + 10\overline{x}_1^2 + 6\overline{x}_1 + 3 - 2\overline{x}_2}{2\overline{x}_1^2 + 2\overline{x}_1 + 1}\upsilon^2 + \tfrac{3}{2}\tfrac{2\overline{x}_1^4 + 5\overline{x}_1^3 + 6\overline{x}_1^2 + 5\overline{x}_1 + 2 - (2\overline{x}_1 + 1)\overline{x}_2}{2\overline{x}_1^2 + 2\overline{x}_1 + 1}\upsilon .$$

$\upsilon = 0$ is the unique zero of P if and only if $(\overline{x}_1, \overline{x}_2)$ lies in the region of the plane between the curves $x_2 = \gamma_1\left(x_1\right) = \frac{-14x_1^3 - 18x_1^2 - 14x_1 + 1 + 4\left(2x_1^2 + 2x_1 + 1\right)\sqrt{(2x_1+1)^2 + 6}}{6}$, $x_2 = \gamma_2\left(x_1\right) = \frac{-14x_1^3 - 18x_1^2 - 14x_1 + 1 - 4\left(2x_1^2 + 2x_1 + 1\right)\sqrt{(2x_1+1)^2 + 6}}{6}$. If $\overline{x}_2 \geq \gamma_1\left(\overline{x}_1\right)$ (resp., $\overline{x}_2 \leq \gamma_2\left(\overline{x}_1\right)$), then $\mathcal{J}_{\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right)} \cap \mathcal{D}$ contains either two or three points. If it contains only two points, $\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right)$ and $\left(\widetilde{x}_1, \widetilde{x}_2, \widetilde{\xi}_2\right)$, then either $\overline{x}_2 = \gamma_1\left(\overline{x}_1\right)$ (resp., $\overline{x}_2 = \gamma_2\left(\overline{x}_1\right)$) or $\widetilde{x}_2 = \gamma_1\left(\widetilde{x}_1\right)$ (resp., $\widetilde{x}_2 = \gamma_2\left(\widetilde{x}_1\right)$). In the case when $\mathcal{J}_{\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right)} \cap \mathcal{D}$ contains more than one point, then the jump must proceed to the point that maximizes x h_Y, i.e., we must choose υ that maximizes

$$\begin{aligned} h_Y \circ \Phi_1^{\vec{h}\,X_\upsilon}\left(\overline{x}, \overline{\xi}\right) = {} & \frac{-1}{3}\upsilon^4 - \frac{6\overline{x}_1^3 + 10\overline{x}_1^2 + 6\overline{x}_1 + 3 - 2\overline{x}_2}{3\left(2\overline{x}_1^2 + 2\overline{x}_1 + 1\right)}\upsilon^3 - \\ & \tfrac{2\overline{x}_1^4 + 5\overline{x}_1^3 + 6\overline{x}_1^2 + 5\overline{x}_1 + 2 - (2\overline{x}_1 + 1)\overline{x}_2}{2\overline{x}_1^2 + 2\overline{x}_1 + 1}\upsilon^2 + \tfrac{\overline{x}_1^6 + 4\overline{x}_1^5 + 9\overline{x}_1^4 - 5\overline{x}_1^3 + 6\overline{x}_1^2 + 5\overline{x}_1 + 2 - \overline{x}_2^2 - 2\left(\overline{x}_1^3 + \overline{x}_1^2\right)\overline{x}_2}{2\overline{x}_1^2 + 2\overline{x}_1 + 1} . \end{aligned}$$

One may check that $\upsilon = 0$ is the maximizer of $h_Y \circ \Phi_1^{\vec{h}\,X_\upsilon}\left(\overline{x}, \overline{\xi}\right)$ if and only if the polynomial

$$Q\left(\upsilon\right) = \frac{-1}{3}\upsilon^2 - \frac{6\overline{x}_1^3 + 10\overline{x}_1^2 + 6\overline{x}_1 + 3 - 2\overline{x}_2}{3\left(2\overline{x}_1^2 + 2\overline{x}_1 + 1\right)}\upsilon - \tfrac{2\overline{x}_1^4 + 5\overline{x}_1^3 + 6\overline{x}_1^2 + 5\overline{x}_1 + 2 - (2\overline{x}_1 + 1)\overline{x}_2}{2\overline{x}_1^2 + 2\overline{x}_1 + 1}$$

has at most one zero. This happens if and only if $(\overline{x}_1, \overline{x}_2)$ lies on or between the curves

$$x_2 = \gamma_3\left(x_1\right) = \frac{-6x_1^3 - 8x_1^2 - 6x_1 + \sqrt{3}\left(2x_1^2 + 2x_1 + 1\right)\sqrt{(2x_1 + 1)^2 + 4}}{2}$$

and

$$x_2 = \gamma_4\left(x_1\right) = \frac{-6x_1^3 - 8x_1^2 - 6x_1 - \sqrt{3}\left(2x_1^2 + 2x_1 + 1\right)\sqrt{(2x_1 + 1)^2 + 4}}{2} .$$

Thus, the state space can be divided in 3 regions: $A_1 = \{(x_1, x_2) : x_2 > \gamma_3\left(x_1\right)\}$, $A_2 = \{(x_1, x_2) : x_2 < \gamma_4\left(x_1\right)\}$, $A_3 = \{(x_1, x_2) : \gamma_4\left(x_1\right) \leq x_2 \leq \gamma_3\left(x_1\right)\}$. If $\left(x_1\left(t\right), x_2\left(t\right)\right)$

is a Sarychev extremal, then the set $\{t > 0 : (x_1(t), x_2(t)) \in A_1 \cup A_2\}$ must be a set of zero measure. If $(\overline{x}_1, \overline{x}_2) \in A_1 \cup A_2$, then every Sarychev extremal must start by a jump that transfers the state into A_3.

Now, let's consider the case when $\left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right) \in \mathcal{D}$, $(\overline{x}_1, \overline{x}_2) \in int\,(A_3)$. For such initial points, a classical extremal exists and coincides with the Sarychev extremal, at least for small time intervals. Hence the extremal control can be determined through Dirac's constraints (16). By substituting this control in system (15), we obtain the differential equations

$$\begin{aligned} \dot{x}_1 &= -\frac{\left(2x_1^2+2x_1+1\right)\left(x_1^3+x_1^2+x_2\right)}{2x_1^4+5x_1^3+6x_1^2+5x_1+2-(2x_1+1)x_2}, \\ \dot{x}_2 &= -x_1 + x_2 - \frac{(x_1+1)\left(x_1^2+2x_1+2\right)\left(4x_1^4+5x_1^3+4x_1^2+2x_1+x_2\right)}{2x_1^4+5x_1^3+6x_1^2+5x_1+2-(2x_1+1)x_2}. \end{aligned} \tag{17}$$

This field has one unique critical point in A_3 which is $(x_1, x_2) = (0, 0)$. This is a saddle and its stable and unstable manifolds divide A_3 in four regions, with one branch of the unstable manifold passing through the narrow space between A_1 and A_2, and one branch of the stable manifold intersecting the boundary of A_2 (see figure). Along the curve $x_2 = \gamma_3(x_1)$, the field (17) points outward from A_1 at every point to the left of the point $x_1 = \frac{\sqrt{2}-1}{2}$ and points inward to A_1 at every point to the right of $x_1 = \frac{\sqrt{2}-1}{2}$. Hence, A_1 is attractive for every extremal trajectory lying in A_3 that passes sufficiently close to the curve $x_2 = \gamma_3(x_1)$, $x_1 > \frac{\sqrt{2}-1}{2}$ and is repulsive for every extremal trajectory lying in A_3 that passes sufficiently close to the curve $x_2 = \gamma_3(x_1)$, $x_1 < \frac{\sqrt{2}-1}{2}$. Similarly, A_2 is attractive for every extremal trajectory lying in A_3 that passes sufficiently close to the curve $x_2 = \gamma_4(x_1)$, $x_1 < \frac{-\sqrt{2}-1}{2}$ and is repulsive for every extremal trajectory lying in A_3 that passes sufficiently close to the curve $x_2 = \gamma_4(x_1)$, $x_1 > \frac{-\sqrt{2}-1}{2}$.

From the above arguments it follows that the structure of Sarychev extremals with initial point $(\overline{x}_1, \overline{x}_2) \in \mathbb{R}^2$ is as follows. At time $t = 0$, execute a jump to any point in $\Phi^X(\overline{x}_1, \overline{x}_2) \cap A_3$. This jump is selected by choosing the appropriated lift $(\overline{x}_1, \overline{x}_2) \mapsto \left(\overline{x}_1, \overline{x}_2, \overline{\xi}_2\right)$. If $(\overline{x}_1, \overline{x}_2) \in A_3$, a jump with zero length is possible, otherwise the jump must transfer the initial state to A_3 and hence it must have positive length. In the interval $]0, T[$, the generalized extremal must be of one of the following types:

1. The point $\left(x_1\left(0^+\right), x_2\left(0^+\right)\right)$ lies in the region to the left of the stable and the unstable manifolds of (17) and to the left of A_2, but not close to the curve $x_2 = \gamma_4(x_1)$, $x_1 < \frac{-\sqrt{2}-1}{2}$. In this case the generalized extremal is an upward moving continuous trajectory in the interval $]0, T[$, for every $T > 0$.

2. The point $\left(x_1\left(0^+\right), x_2\left(0^+\right)\right)$ lies in the region to the left of the stable and the unstable manifolds of (17), to the left of A_2 and sufficiently close to the curve $x_2 = \gamma_4(x_1)$, $x_1 < \frac{-\sqrt{2}-1}{2}$. In this case, for sufficiently large $T > 0$, the

generalized extremal has one unique discontinuity in the interval $]0,T[$. This discontinuity transfers the state $\left(x_1\left(t_1^-\right),x_2\left(t_1^-\right)\right)$ which lies in the curve $x_2=\gamma_4\left(x_1\right),\ x_1<\frac{-\sqrt{2}-1}{2}$ to a point $\left(x_1\left(t_1^+\right),x_2\left(t_1^+\right)\right)$, lying in the curve $x_2=\gamma_4\left(x_1\right),\ x_1>\frac{-\sqrt{2}-1}{2}$. Two subcases can occur:

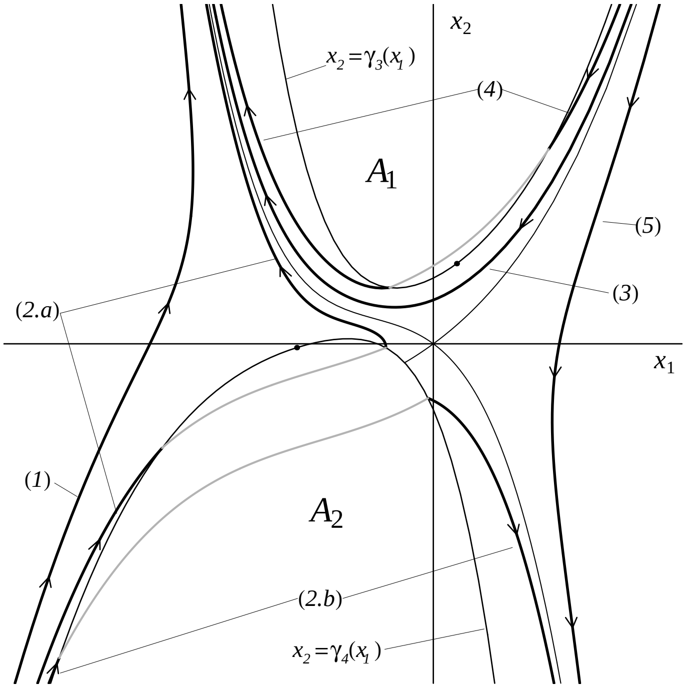
(a) If the point $\left(x_1\left(t_1^+\right),x_2\left(t_1^+\right)\right)$ lies to the left of the stable manifold of (17), then the generalized extremal in the interval $]t_1,T[$ is a trajectory moving upwards to the left;

(b) If the point $\left(x_1\left(t_1^+\right),x_2\left(t_1^+\right)\right)$ lies to the right of the stable manifold of (17), then the generalized extremal in the interval $]t_1,T[$ is a trajectory moving downwards to the right.

3. The point $\left(x_1\left(0^+\right),x_2\left(0^+\right)\right)$ lies in the region above the stable and the unstable manifolds of (17) and below A_1, but not close to the curve $x_2=\gamma_3\left(x_1\right),\ x_1>\frac{\sqrt{2}-1}{2}$. In this case the generalized extremal is a continuous trajectory in the interval $]0,T[$, for every $T>0$. It evolves in the region between the stable and unstable manifolds of (17) and A_1, from right to left.
4. The point $\left(x_1\left(0^+\right),x_2\left(0^+\right)\right)$ lies in the region above the stable and the unstable manifolds of (17), below A_1 and close to the curve $x_2=\gamma_3\left(x_1\right),\ x_1>\frac{\sqrt{2}-1}{2}$. In this case, for sufficiently large $T>0$, the generalized extremal has one unique discontinuity in the interval $]0,T[$. This discontinuity transfers the state $\left(x_1\left(t_1^-\right),x_2\left(t_1^-\right)\right)$ which lies in the curve $x_2=\gamma_3\left(x_1\right),\ x_1>\frac{\sqrt{2}-1}{2}$ to a point $\left(x_1\left(t_1^+\right),x_2\left(t_1^+\right)\right)$, lying in the curve $x_2=\gamma_3\left(x_1\right),\ x_1<\frac{\sqrt{2}-1}{2}$. In the intervals $]0,t_1[$, $]t_1,T[$ the generalized extremal is a continuous trajectory lying in the region between the stable and unstable manifolds of (17) and A_1, moving from right to left.
5. The point $\left(x_1\left(0^+\right),x_2\left(0^+\right)\right)$ lies in the region to the right of the stable and the unstable manifolds of (17). In this case the generalized extremal is a downward moving continuous trajectory in the interval $]0,T[$, for every $T>0$.
6. If the point $\left(x_1\left(0^+\right),x_2\left(0^+\right)\right)$ lies in the region below the stable and the unstable manifolds of (17) and to the right of A_2, Then the generalized extremal is a continuous curve in the interval $]0,T[$, for every $T>0$, of the same type as the arc described by an extremal of the type $(2.b)$ in the interval $]t_1,T[$.

At time $t=T$, we may choose a final jump that transfers the state $\left(x_1\left(T^-\right),x_2\left(T^-\right)\right)$ to any point in $\Phi^X\left(x_1\left(T^-\right),x_2\left(T^-\right)\right)$ (including points in $A_1\cup A_2$).

Arcs of extremals of types (1) to (5) are shown in the figure, identified by the corresponding numbers. The continuous arcs of the generalized extremals are drawn in black, while the segments of the integral manifolds Φ^X that describe jumps during the interval $]0,T[$ are drawn in grey (initial and final jumps are not represented).



It should be noticed that the denominator in feedback (16) has zeroes in $\mathbb{R}^2$. However, the only zeroes that lie in A_3 are the points

$$\left(\frac{-\sqrt{2}-1}{2}, \frac{9-7\sqrt{2}}{8}\right) \text{ and } \left(\frac{\sqrt{2}-1}{2}, \frac{9+7\sqrt{2}}{8}\right),$$

which are, respectively, the points where the curves $x_2 = \gamma_4(x_1)$, and $x_2 = \gamma_3(x_1)$ are neither repulsive nor attractive for extremal trajectories lying in A_3 (black dots in the figure). The direction of the field is well defined in both these points and it is tangent to the boundary of A_3, pointing from the attractive segment towards the repulsive segment. The extremals that pass through these points attain infinite speed when they pass over the point but lie entirely on A_3 and are continuous.

Also, notice that every point $(\overline{x}_1, \overline{x}_2) \in A_1 \cup A_2$ which is not a zero of the denominator of (16), admits a classical extremal with initial condition $(x_1(0), x_2(0)) = (\overline{x}_1, \overline{x}_2)$, but none of these classical extremals is a Sarychev extremal. However, every classical extremal arc that lies in one of the regions, between the curves

$$x_2 = \frac{(x_1+1)^2\left(2x_1^2+x_1+2\right)}{2x_1+1}, \; x_1 > -\frac{1}{2} \text{ and } x_2 = \gamma_3(x_1),$$

or between the curves

$$x_2 = \frac{(x_1+1)^2\left(2x_1^2+x_1+2\right)}{2x_1+1},\ x_1 < -\frac{1}{2} \text{ and } x_2 = \gamma_4\left(x_1\right),$$

satisfies the strong Legendre-Clebsch condition at every point.

References

[1] AGRACHEV A.A. AND SARYCHEV A.V., *On reduction of a smooth system linear in the control*, Math USSR Sbornik **58** (1) (1987), 15–30.

[2] DIRAC P.A.M., *Lectures on quantum mechanics*, Belfer Graduate School of Science, 1964.

[3] GUERRA M. *Distribution-like Hamiltonian flows and generalized optimal controls*, J. Math Sci. **120** (1) (2004), 895-918.

[4] SARYCHEV A.V. *Nonlinear Systems with Impulsive and Generalized Function Controls*, in: “Proceedings of a IIASA Workshop, Sopron 1989” (Eds. Byrnes C.I. and Kurzhansky A.) Birkhäuser, Boston 1991, 244–257.

AMS Subject Classification: 93C15, 93B50, 49K15.

Manuel GUERRA, Department of Mathematics, ISEG, Technical University of Lisbon, R. do Quelhas 6,
1200 Lisboa, PORTUGAL
e-mail: `mguerra@iseg.utl.pt`