This paper was presented at the 38th IEEE Conference on Decision and Control, held in Phoenix, Arizona, December 7-10, 1999, and has appeared in the conference proceedings. The date of this version is August 12, 1999.

Regular Presynthesis and Synthesis, and Optimality of Families of Extremals

Héctor J. Sussmann [1]
Department of Mathematics
Rutgers University
Piscataway, NJ 08854-8019, USA
sussmann@hamilton.rutgers.edu
http://www.math.rutgers.edu/~sussmann

Benedetto Piccoli
SISSA–ISAS
via Beirut 2-4
34014 Trieste, Italy
piccoli@shannon.sissa.it

1. Introduction

The purpose of the theory of "regular synthesis" is to turn into a rigorous, precise theorem the vague assertion that "if a collection Γ of extremals covers the whole space in a smooth enough way, then all the members of Γ are optimal." Naturally, a good definition should at the same time be stringent enough to lead to a correct proof of the sufficiency assertion, and broad enough to cover as large as possible a class of optimal control problems. In particular, it would be desirable for the theory to apply to well-known examples such as the famous "Fuller problem" (cf. Fuller [11], Marchal [12], Zelikin-Borisov [20]), where the optimal controls have infinitely many switchings.

The main goals of this paper are (a) to propose a new definition of regular synthesis, more general than those suggested by other authors (e. g. Boltyanskii [3], Brunovský [4, 5]), as well as an even more general notion of "regular presynthesis," and (b) to give a detailed statement of the corresponding sufficiency theorem.

The paper is an abridged version of our journal article [14], where we give a complete proof of the sufficiency theorem, and compare our definition with other concepts of synthesis such as those of Boltyanskii and Brunovský, and with the simpler notion of a "family of solutions of the closed-loop equation arising from an optimal feedback law."

The main sufficiency result is Theorem 1. The key to the proof turns out to be a technical result—Theorem 2—on the relation between the adjoint vector on the Pontryagin Maximum Principle and the gradient of the cost functional. This in turn depends on results about differentiability of trajectories with respect to a parameter under weak differentiability assumptions on the dependence of the data on the parameter. These results are of interest in themselves and are presented separately in Appendix A of [14].

Our definitions of regular presynthesis and regular synthesis are presented in §2, and the main theorems are stated in §3. In §4 we point out that Theorem 1 applies to the synthesis for Fuller's problem, for which other sufficiency theorems such as those of [3] and [5] cannot be used because of the technical problem arising from the occurrence of an infinite number of switchings. (Not surprisingly, the optimality of Fuller's synthesis was originally proved without resorting to regular synthesis arguments, and using instead special symmetry properties of the problem, cf., e. g., Fuller [11], Marchal [12], Zelikin-Borisov [20].)

Other definitions of synthesis—including those of Boltyanskii and Brunovský—are briefly mentioned and compared with ours in §5, where we also discuss why the concept of "optimal synthesis" cannot just be defined by simply identifying it with that of "optimal feedback law."

As explained in §5, there are essentially two opposite starting points for defining the notion of a "regular" optimal synthesis. One approach, following Boltyanskii [3], Brunovský [4, 5], Cesari [6], Fleming and Soner [10], and Sussmann [16], is to take as the primary object a smooth or "piecewise smooth" feedback control v, and then let the synthesis be the collection $Traj(v)$ of all trajectories generated by v. If this collection is too large (for example, if the closed-loop equation determined by v does not have unique solutions), then a synthesis would be a suitably defined subclass Γ of $Traj(v)$. The other strategy, proposed in Piccoli [13] and Sussmann [18], is to take a "synthesis" to be just a collection of trajectories—or, more precisely, trajectory-control pairs—not necessarily arising from a feedback control. To carry out the first strategy, Boltyanskii and Brunovský were forced to make precise sense of the notion of piecewise smoothness, which they did by requiring that the set of interest—that is, the set of initial points of optimal trajectories—be partitioned into submanifolds of various dimensions on which a smooth feedback control is specified. It turns out, however, that with this definition the feedback control by itself is not sufficient for the synthesis to be completely determined, since a discontinuous feedback control may generate too many trajectories, some of which could fail to be optimal, as shown in Example 5.3 of [14]. So, in any case, the final outcome of the effort to implement the first strategy is a result more consistent with the second one, since one ends up with a notion of synthesis as a pair consisting of a feedback control v together with a selection Γ of a family of trajectories generated by v, suggesting that the specification of Γ is needed but that of v might not be. This leads naturally to the approach of this paper, in which v is dispensed with altogether.

An important point that ought to be further investigated but will not be touched upon here is the relation between the geometrical approaches of [4, 5], [16, 17, 18], and the theory of viscosity solutions (cf. Bardi and Capuzzo Dolcetta [1], Fleming and Soner [10]). We point out that the existing results of this theory for deterministic optimal control do not appear to cover cases such as Fuller's problem. (Cf. for example [1], where the main result is for Lagrangians that are bounded away from zero.)

[1] Research supported in part by NSF Grant DMS98-03411-00798 and AFOSR Grant 0923.

2. Regular Presynthesis and Synthesis

We start with some definitions that will enable us to state and prove a mild generalization of the theorem whose proof was outlined in [18]. Our assumptions will be minimal, so as to achieve maximum generality.

We consider a control system:

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

and a family

$$\boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{P}}(\Omega, U, f, L, \mathcal{T}, \tau) = \{\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)\}_{x\in\Omega}$$

of minimization problems, where $\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)$ is the problem:

$$\begin{cases} \text{minimize} & \tau(\gamma(b)) + \int_a^b L(\gamma(t), \eta(t))dt \\ \text{subject to} & (1)\,,\ \gamma(a) = x\,, \\ \text{and} & \gamma(b) \in \mathcal{T}\,. \end{cases} \tag{2}$$

(A more precise formulation of the problems $\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)$, carefully specifying the space of of curves to be considered, will be given below. Since both f and L have no explicit dependence on t, we can always take $b = 0$.) We assume that

- *(A1) n is a positive integer and Ω—the "state space"—is an open subset of $\mathbb{R}^n$;*
- *(A2) U—the "control set"—is a nonempty set;*
- *(A3) $f : \Omega \times U \to \mathbb{R}^n$—the "dynamics"—and $L : \Omega \times U \to \mathbb{R}$—the "Lagrangian"—are maps;*
- *(A4) $f(\cdot, u)$ and $L(\cdot, u)$ are maps of class $\mathcal{C}^1$ for every fixed $u \in U$;*
- *(A5) $\mathcal{T}$—the "target"—is a nonempty subset of Ω;*
- *(A6) $\tau : \Omega \to \mathbb{R}$ is a function of class $\mathcal{C}^1$.*

We use $T_x^c S$ to denote the *contingent tangent cone* to a set S at a point $x \in S$, defined as follows: a vector v is in $T_x^c S$ iff there exist a sequence $\{x_j\}$ of points of S and a sequence $\{h_j\}$ of strictly positive numbers such that $h_j \to 0$ and $x_j = x + h_j v + o(h_j)$ as $j \to \infty$. We write

$$\tilde{T}_x S = T_x^c S \cap (-T_x^c S) = \{v : v \in T_x^c S \quad \text{and} \quad -v \in T_x^c S\}\,.$$

As in Sussmann [18], $\tilde{f}$ will denote the map

$$\Omega \times U \ni (x,u) \to \tilde{f}(x,u) \stackrel{\text{def}}{=} (f(x,u), L(x,u)) \in \mathbb{R}^{n+1}\,.$$

If $\Omega \times \mathbb{R} \ni (x,t) \to g(x,t) \in \mathbb{R}^m$ is a time-varying vector-valued map on Ω, then $Dg(x,t)$ denotes, for each $(x,t) \in \Omega$, the differential (i. e., the Jacobian matrix) at x of the map $y \to g(y,t)$, so $Dg(x,t) \in \mathbb{R}^{m\times n}$.

If $\mu : A \to B$ is a map, we use $\mathrm{Dom}(\mu)$ to indicate the domain of μ, i.e. the set A.

A *control* is a U-valued map η such that $\mathrm{Dom}(\eta)$ is a compact subinterval $[a,b]$ of $\mathbb{R}$. If η is a control, we define maps $f_\eta : \Omega \times \mathbb{R} \to \mathbb{R}^n$, $L_\eta : \Omega \times \mathbb{R} \to \mathbb{R}$, $\tilde{f}_\eta : \Omega \times \mathbb{R} \to \mathbb{R}^{n+1}$ by letting

$$\left.\begin{array}{rcl} f_\eta(x,t) &=& f(x,\eta(t)), \\ L_\eta(x,t) &=& L(x,\eta(t)), \\ \tilde{f}_\eta(x,t) &=& \tilde{f}(x,\eta(t)), \end{array}\right\} \quad \text{if} \quad t \in \mathrm{Dom}(\eta)\,;$$

$$\left.\begin{array}{rcl} f_\eta(x,t) &=& 0, \\ L_\eta(x,t) &=& 0, \\ \tilde{f}_\eta(x,t) &=& 0, \end{array}\right\} \quad \text{if} \quad t \notin \mathrm{Dom}(\eta)\,.$$

A *trajectory* for a control η is an absolutely continuous map $\gamma : \mathrm{Dom}(\eta) \to \Omega$ such that $\dot{\gamma}(t) = f(\gamma(t), \eta(t))$ for almost every $t \in \mathrm{Dom}(\eta)$. If $\mathrm{Dom}(\eta) = [a,b]$ (so that $\mathrm{Dom}(\gamma) = [a,b]$ as well), and $x = \gamma(a)$, $y = \gamma(b)$, we say that γ (or the pair (γ, η)) *goes from x to y*, and that η *steers x to y*, and we write $\gamma^- = x$, $\gamma^+ = y$.

A control η is $\mathcal{C}^1$*-admissible* for $\tilde{f}$ if $\tilde{f}_\eta$ satisfies the following $\mathcal{C}^1$ Carathéodory conditions:

- (C.i) *the map $\mathrm{Dom}(\eta) \ni t \to \tilde{f}_\eta(x,t) \in \mathbb{R}^{n+1}$ is measurable for each $x \in \Omega$;*
- (C.ii) *for every compact subset K of Ω there exists an integrable function $\varphi_K : \mathbb{R} \to \mathbb{R}$ such that for every $x \in K$ and $t \in \mathrm{Dom}(\eta)$:*

$$\|\tilde{f}_\eta(x,t)\| + \|D\tilde{f}_\eta(x,t)\| \le \varphi_K(t)\,. \tag{3}$$

If η is a $\mathcal{C}^1$-admissible control for $\tilde{f}$, and γ is a trajectory corresponding to η, then we say that (γ, η) is a $\mathcal{C}^1$*-admissible pair* for $\tilde{f}$. We use $\mathrm{Adm}^{\mathcal{C}^1}(\tilde{f})$ to denote the set of all $\mathcal{C}^1$-admissible pairs for $\tilde{f}$, and write $\mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f})$, $\mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$, $\mathrm{Adm}_x^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$, to denote, respectively, the set of all $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1}(\tilde{f})$ such that $\mathrm{Dom}(\eta)$ is an interval of the form $[T, 0]$ with $T \le 0$, the set of those $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f})$ such that $\gamma(0) \in \mathcal{T}$, and the set of those $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$ such that $\gamma^- = x$. For a pair $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$, we use $J(\gamma, \eta)$ to denote the *cost* of (γ, η), given by

$$J(\gamma, \eta) = \tau(\gamma^+) + \int_{\mathrm{Dom}(\gamma)} L(\gamma(t), \eta(t))\,dt\,. \tag{4}$$

If $\Gamma \subseteq \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$ is an arbitrary set of $\mathcal{C}^1$-admissible pairs for $\tilde{f}$ ending at the target $\mathcal{T}$ at time 0, we define a function $V_\Gamma : \Omega \to \mathbb{R} \cup \{+\infty\}$, called the *cost function of* Γ, by letting

$$V_\Gamma(x) = \inf\{J(\gamma, \eta) : (\gamma, \eta) \in \Gamma,\ \gamma^- = x\} \tag{5}$$

for $x \in \Omega$. (In particular, $V_\Gamma(x) = +\infty$ iff there is no $(\gamma, \eta) \in \Gamma$ starting at x.) If $\Gamma = \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$, then the cost function V_Γ is the *value function* of our problem, and in that case we write $V_{\tilde{f},\mathcal{T}}$ rather than $V_{\mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f},\mathcal{T})}$. Naturally, $V_\Gamma \ge V_{\tilde{f},\mathcal{T}}$ pointwise for every subset Γ of $\mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$.

We can now formulate $\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)$ precisely: $\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)$ is the problem of finding a pair $(\gamma, \eta) \in \mathrm{Adm}_x^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$ that satisfies $J(\gamma, \eta) = V_{\tilde{f},\mathcal{T}}(x)$. Any such pair (γ, η) is said to be a *solution* of $\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)$. A pair $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f}, \mathcal{T})$ is *optimal* if it is a solution of $\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, \gamma^-)$. We use $\boldsymbol{\mathcal{P}}(\Omega, U, f, L, \mathcal{T}, \tau)$—or, simply, $\boldsymbol{\mathcal{P}}$, in any context where the meaning of Ω, U, f, L, $\mathcal{T}$, τ is clear—to denote the family of problems $\Big\{\mathcal{P}(\Omega, U, f, L, \mathcal{T}, \tau, x)\Big\}_{x\in\Omega}$.

Remark 1 We emphasize that in the optimal control problems studied here *the trajectories have to end at* *the target but could pass through it before.* For example, consider the dynamical law $\dot{x} = 1$ in $\mathbb{R}$, with Lagrangian $L(x) \equiv -1$ and target $\mathcal{T} = \{0, 1\}$. Then the value function V is given by $V(x) = x - 1$ for $x \le 1$ and $V(x) = +\infty$ for $x > 1$. If $x < 1$, then the optimal trajectory γ_x from x to the target is given by $\gamma_x(t) = t + 1$ for $T_x \le t \le 0$, where $T_x = x - 1$. Notice that if $x < 0$ then $t = 0$ is not the first t such that $\gamma_x(t) \in \mathcal{T}$. $\diamondsuit$

Given $\lambda \in \mathbb{R}_n$ (the space of row n-vectors), $\lambda^0 \in \mathbb{R}$, $x \in \Omega$ and $u \in U$, we define:

$$\begin{aligned} H^c(\lambda, \lambda^0, x, u) &= \langle \lambda, f(x, u) \rangle + \lambda^0 L(x, u)\,, \\ H(\lambda, \lambda^0, x) &= \inf\{H^c(\lambda, \lambda^0, x, u) : u \in U\}\,. \end{aligned}$$

The real-valued functions $H^c : \mathbb{R}_n \times \mathbb{R} \times \Omega \times U \to \mathbb{R}$ and $H : \mathbb{R}_n \times \mathbb{R} \times \Omega \to \mathbb{R} \cup \{-\infty\}$ are, respectively, the *control Hamiltonian* and the *minimized Hamiltonian* of $\tilde{f}$.

Definition 1 A pair $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1, 0}(\tilde{f}, \mathcal{T})$ is *extremal* if there exist an absolutely continuous map $\lambda : \mathrm{Dom}(\gamma) \to \mathbb{R}_n$ and a constant $\lambda^0 \ge 0$, such that:

(E1) $(\lambda^0, \lambda(t)) \ne (0, 0)$ *for some (and hence every)* t*;*

(E2) $\dot{\lambda} = -\frac{\partial H^c}{\partial x}(\lambda(t), \lambda^0, \gamma(t), \eta(t))$ *for a. e.* t*;*

(E3) $H(\lambda(t), \lambda^0, \gamma(t)) = H^c(\lambda(t), \lambda^0, \gamma(t), \eta(t)) = 0$ *for a. e.* t*;*

(E4) $\langle \lambda(0), v \rangle = \lambda^0 \langle \nabla \tau(\gamma(0)), v \rangle$ *for all* $v \in \tilde{T}_{\gamma(0)} \mathcal{T}$. $\diamondsuit$

Properties (E1), (E2), (E3), (E4) are, respectively, the *nontriviality condition*, the *adjoint equation*, the *Hamiltonian minimization condition*, and the *transversality condition*. A pair (λ, λ^0) such that (E2) holds is an *adjoint vector for* (γ, η). An adjoint vector for which (E3) holds is a *minimizing adjoint vector for* (γ, η).

The *maximum principle* says that, under some special conditions on the target set $\mathcal{T}$, *every optimal pair* (γ, η) *is extremal.* (This is true, for example, if $\mathcal{T}$ is a submanifold of Ω of class $\mathcal{C}^1$, or a closed convex subset of Ω, or a set which is locally equivalent to a convex set via a $\mathcal{C}^1$ diffeomorphism. More generally, it is true if the cone $\tilde{T}_{\gamma(0)} \mathcal{T}$ is convex—in which case it is of course a subspace, since $v \in \tilde{T}_{\gamma(0)} \mathcal{T} \Rightarrow -v \in \tilde{T}_{\gamma(0)} \mathcal{T}$—and is an approximating cone to $\mathcal{T}$ at $\gamma(0)$. Recall that a closed convex cone C is an *approximating cone* to a set S at a point $s \in S$ if there exist a neighborhood W of 0, and a continuous map $F : W \cap C \to S$, such that $F(v) = s + v + o(\|v\|)$ as $v \to 0$ via values in C.)

Definition 2 A *presynthesis* for $\boldsymbol{\mathcal{P}}$ is a subset Γ of $\mathrm{Adm}^{\mathcal{C}^1, 0}(\tilde{f}, \mathcal{T})$ such that,

(PS) *Whenever* $(\gamma_1, \eta_1) \in \Gamma$, $(\gamma_2, \eta_2) \in \Gamma$, $\gamma_1^- = \gamma_2^-$, *it follows that* $(\gamma_1, \eta_1) = (\gamma_2, \eta_2)$.

The set $\mathrm{Dom}(\Gamma) \stackrel{\mathrm{def}}{=} \{\gamma^- : (\gamma, \eta) \in \Gamma\}$ is called the *domain* of Γ. We say that Γ is a *presynthesis on* S if Γ is a presynthesis and $S = \mathrm{Dom}(\Gamma)$. $\diamondsuit$

Clearly, giving a presynthesis on S amounts to choosing, for each $x \in S$, a $(\gamma, \eta) \in \mathrm{Adm}^{\mathcal{C}^1, 0}(\tilde{f}, \mathcal{T})$ such that $\gamma^- = x$.

If Γ is a presynthesis such that $\mathrm{Dom}(\Gamma)$ consists of all points that can be steered in finite time to a point of $\mathcal{T}$ by means of an admissible pair, then we say that Γ is *total*.

Given a presynthesis Γ and a point $x \in \mathrm{Dom}(\Gamma)$, we will always use the expression $(\gamma_x^\Gamma, \eta_x^\Gamma)$—or, simply, (γ_x, η_x) when the meaning of Γ is clear from the context—to denote the unique pair $(\gamma, \eta) \in \Gamma$ such that $\gamma^- = x$, and will write $T_x = \min \mathrm{Dom}(\gamma_x)$, so $T_x \le 0$ and $\mathrm{Dom}(\gamma_x) = \mathrm{Dom}(\eta_x) = [T_x, 0]$.

Definition 3 A presynthesis is *memoryless* if whenever $x \in S$ and $t \in \mathrm{Dom}(\eta_x)$ it follows that $y = \gamma_x(t) \in S$ and η_y is the restriction of η_x to the interval $[t, 0]$. A *synthesis* is a memoryless presynthesis. $\diamondsuit$

Definition 4 If each pair of a presynthesis Γ is optimal (resp. extremal) then we say that Γ is *optimal* (resp. *extremal*). $\diamondsuit$

In particular, a presynthesis Γ is a total optimal presynthesis if and only if $V_\Gamma \equiv V_{\tilde{f}, \mathcal{T}}$.

As explained in the introduction, we wish to propose a definition of "regular synthesis" that will give rise to a sufficiency theorem. The theorem will say that a total extremal regular synthesis that satisfies the appropriate boundary conditions is optimal. We will actually define the more general notion of "regular presynthesis," and state the sufficiency theorem in this broader setting, so the desired optimality result will be true in particular when the presynthesis is a true synthesis.

The definition will say that a "regular presynthesis" is a presynthesis that satisfies suitable regularity assumptions. So we begin by describing these regularity assumptions in detail.

Definition 5 A function $V : \Omega \to \mathbb{R} \cup \{-\infty, +\infty\}$ is *weakly upper semicontinuous* (w.u.s.c) at a point $x \in \Omega$ if $\liminf_{y \to x} \limsup_{z \to y} V(z) \le V(x)$. $\diamondsuit$

Equivalently, V is w.u.s.c. at x if there exist sequences $\{x_j\}$, $\{\varepsilon_j\}$ and $\{\delta_j\}$ such that $x_j \in \Omega$, $\varepsilon_j > 0$, and $\delta_j > 0$ for all j, and $x_j \to x$ and $\delta_j \to 0$ as $j \to \infty$, having the property that $V(y) \le V(x) + \delta_j$ whenever $\|y - x_j\| < \varepsilon_j$.

Definition 6 Suppose we are given a Lipschitz-continuous vector field X on Ω and a function $V : \Omega \to \mathbb{R} \cup \{-\infty, +\infty\}$. We say that V has the *NDJ* ("no downward jumps") *property along* X if

(NDJ) *if* $\gamma : [a, b] \to \Omega$ *is an integral curve of* X*, and* $a \le t < b$*, then* $\liminf_{h \downarrow 0} V(\gamma(t - h)) \le V(\gamma(t))$. $\diamondsuit$

Definition 7 We say that V satisfies the *weak continuity conditions* for the control system (1) if V is lower semicontinuous and weakly upper semicontinuous, and has the NDJ property along the vector field $x \to f(x, u)$ for every $u \in U$. $\diamondsuit$

We now define the "weak differentiability conditions" for Γ. In this definition, we let $\rho_{\tilde{f}, \Gamma, \bar{x}, T}(v)$ be the integrable $\mathbb{R}^{n+1}$-valued function on $[T, 0]$ given by

$$\rho_{\tilde{f}, \Gamma, \bar{x}, T}(v)(t) = \tilde{f}_{\eta_{\bar{x}+v}}(\gamma_{\bar{x}}(t), t) - \tilde{f}_{\eta_{\bar{x}}}(\gamma_{\bar{x}}(t), t)\,,$$

and use the symbol $\rho_{\tilde{f}, \Gamma, \bar{x}, T}$ to denote the correspondence $v \to \rho_{\tilde{f}, \Gamma, \bar{x}, T}(v)$, regarded as a map into the

$\mathrm{Bor}([T,0],\mathbb{R}^{n+1})$ of $\mathbb{R}^{n+1}$-valued Borel measures on $[T,0]$. (Recall that $\mathrm{Bor}([T,0],\mathbb{R}^{n+1})$ is the dual of the space $C^0([T,0],\mathbb{R}^{n+1})$ of continuous $\mathbb{R}^{n+1}$-valued functions on $[T,0]$.)

Definition 8 A presynthesis $\Gamma = \{(\gamma_x,\eta_x) : x \in S\}$ is said to be (f,L)*–differentiable* at a point $\bar{x} \in \Omega$ if there exist: (a) a neighborhood N of $\bar{x}$ in Ω such that $N \subseteq \mathrm{Dom}(\Gamma)$, (b) an interval $[T,0]$ such that $\mathrm{Dom}(\eta_x) \subseteq [T,0]$ for all $x \in N$, and (c) an $\bar{\varepsilon} > 0$, for which

(DC1) *there are integrable functions* $\psi_\varepsilon : [T,0] \to \mathbb{R}$, *for each* $\varepsilon \in]0,\bar{\varepsilon}]$, *with the property that* $\lim_{\varepsilon\downarrow 0}\int_T^0 \psi_\varepsilon(t)\,dt = 0$, *for which the inequality*

$$\left\|\tilde{f}_{\eta_x}(y,t) - \tilde{f}_{\eta_x}(\gamma_x(t),t) - D\tilde{f}_{\eta_{\bar{x}}}(\gamma_{\bar{x}}(t),t)\cdot(y-\gamma_{\bar{x}}(t))\right\| \le \psi_\varepsilon(t).(||y-\gamma_{\bar{x}}(t)|| + ||x-\bar{x}||)$$

holds for every $y \in \Omega$, $x \in N$, $t \in [T,0]$ *such that* $\|y-\gamma_{\bar{x}}(t)\| \le \varepsilon$ *and* $||x-\bar{x}|| \le \varepsilon$;

(DC2) *the map* $\rho_{\tilde{f},\Gamma,\bar{x},T}$ *is weak*-differentiable at* $v = 0$; *that is: if* $\alpha : [T,0] \to \mathbb{R}$ *is continuous, then the map* $\mathbb{R}^n \ni v \to \int_T^0 \alpha(t).\rho_{\tilde{f},\Gamma,\bar{x},T}(v)(t)dt \in \mathbb{R}^{n+1}$ *is differentiable at* $v = 0$. ♢

Definition 9 A subset A of Ω is *thin* if there exist A_1, A_2 such that $A = A_1 \cup A_2$, A_1 is a finite or countable union of connected $\mathcal{C}^1$ submanifolds of positive codimension, and $\mathcal{H}^{n-1}(A_2) = 0$, where $\mathcal{H}^k$ denotes k-dimensional Hausdorff measure. ♢

If $S \subseteq \Omega$, then $\overset{\circ}{S}$ will denote the interior of S in Ω.

Definition 10 *If* n, Ω, U, f, L, $\mathcal{T}$, τ *are such that Assumptions (A1) – (A6) hold,* Γ *is a presynthesis for* $\boldsymbol{\mathcal{P}}(\Omega,U,f,L,\mathcal{T},\tau)$, *and* $S = \mathrm{Dom}(\Gamma)$, *we say that* Γ *has the interior approximation property if*

(IAP) *If* $(\gamma,\eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f})$ *is such that* $\gamma(t) \in S$ *for all* $t \in \mathrm{Dom}(\gamma)$, *then there exists a sequence* $\{(\gamma_j,\eta_j)\}_{j=1}^\infty$ *in* $\mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f})$ *such that*

- *(IAP.1)* $\gamma_j(t) \in \overset{\circ}{S}$ *for every* $t \in \mathrm{Dom}(\gamma_j)$;
- *(IAP.2)* $V_\Gamma(\gamma_j^+) \to V_\Gamma(\gamma^+)$;
- *(IAP.3)* $\gamma_j^- \to \gamma^-$;
- *(IAP.4)* $\int_{\mathrm{Dom}(\gamma_j)} L(\gamma_j(t),\eta_j(t))\,dt$ *converges to* $\int_{\mathrm{Dom}(\gamma)} L(\gamma(t),\eta(t))\,dt$ *as* $j \to \infty$. ♢

We are now ready to present our definition or "regular presynthesis":

Definition 11 Let n, Ω, U, f, L, $\mathcal{T}$, τ be such that Assumptions (A1) – (A6) hold. Let Γ be a presynthesis for $\boldsymbol{\mathcal{P}}(\Omega,U,f,L,\mathcal{T},\tau)$. We say that Γ is *regular* if

(11.a) *the associated cost function* V_Γ *satisfies the weak continuity conditions;*

(11.b) *there exists a thin subset* A *of* Ω *such that* Γ *is* (f,L)*–differentiable at every point belonging to* $\mathrm{Dom}(\Gamma)\backslash A$. ♢

3. The Sufficiency Theorem

With the above definitions, we are now in a position to state our main sufficiency theorem. (A weaker version of this result was already stated in [18], with only an outline of the proof. A detailed proof is given in [14].)

Theorem 1 *Let* n, Ω, U, f, L, $\mathcal{T}$, τ *be such that Assumptions (A1) – (A6) hold. Let* Γ *be an extremal regular presynthesis for* $\boldsymbol{\mathcal{P}}(\Omega,U,f,L,\mathcal{T},\tau)$ *and let* $S = \mathrm{Dom}(\Gamma)$. *Then:*

(1.a) *The "dynamic programming inequality"*

$$V_\Gamma(\gamma^-) \le \int_{\mathrm{Dom}(\gamma)} L(\gamma(t),\eta(t))\,dt + V_\Gamma(\gamma^+) \quad (6)$$

holds for every $(\gamma,\eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f})$ *such that* γ *is contained in* $\overset{\circ}{S}$.

(1.b) *If* V_Γ *satisfies the boundary condition*

$$V_\Gamma(x) \le \tau(x) \quad \text{for every} \quad x \in \mathcal{T}\,, \quad (7)$$

then $V_\Gamma(\gamma^-) \le J(\gamma,\eta)$ *for all* (γ,η) *in* $\mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f},\mathcal{T})$ *such that* γ *is contained in* $\overset{\circ}{S}$.

(1.c) *If* Γ *has the interior approximation property, then (6) holds for all* $(\gamma,\eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f})$ *such that* γ *is contained in* S.

(1.d) *If* Γ *has the interior approximation property and satisfies (7) then* $V_\Gamma(\gamma^-) \le J(\gamma,\eta)$ *for all* $(\gamma,\eta) \in \mathrm{Adm}^{\mathcal{C}^1,0}(\tilde{f},\mathcal{T})$ *such that* γ *is contained in* S.

(1.e) *If* Γ *is total and (7) holds,* Γ *is optimal.* ♢

The key fact in the proof of Theorem 1 is the relation between the differential of the cost function V_Γ and the adjoint fields of covectors along the extremal trajectories of the presynthesis. This result is important in its own right, so we state it as a separate theorem:

Theorem 2 *Assume that* n, Ω, U, f, L, $\mathcal{T}$, τ *are such that (A1)-(A6) hold, and let* $\Gamma = \{(\gamma_x,\eta_x) : x \in S\}$ *be a presynthesis. Let* $\bar{x} \in S$ *be such that* Γ *is* (f,L)*–differentiable at* $\bar{x}$. *Then*

(2.a) *the function* V_Γ *is differentiable at* $\bar{x}$;

(2.b) *if the pair* $(\gamma_{\bar{x}},\eta_{\bar{x}})$ *is extremal, then every minimizing adjoint vector* (λ,λ^0) *for* $(\gamma_{\bar{x}},\eta_{\bar{x}})$ *satisfies the identity* $\lambda^0(\nabla V_\Gamma)(\bar{x}) = \lambda(T_{\bar{x}})$. *In particular,* $\lambda^0 \ne 0$, *and* (λ,λ^0) *is unique up to multiplication by a positive constant.* ♢

We refer the reader to [14] for the proof of Theorems 1 and 2.

4. The Fuller Problem

The family $\hat{\boldsymbol{\mathcal{P}}} = \{\hat{\mathcal{P}}_x\}_{x\in\mathbb{R}^2}$ of minimization problems:

$$\hat{\mathcal{P}}_x \; : \qquad \left\{ \begin{array}{ll} \text{minimize} & \int_0^{+\infty} x_1^2(t)\,dt \\ \text{subject to} & (x_1(T),x_2(T)) = x\,, \end{array} \right.$$

associated to the control system

$$\dot{x}_1 = x_2\,, \qquad \dot{x}_2 = u\,, \qquad |u| \le 1\,, \qquad (8)$$

was studied by Fuller in [11], and extensively analyzed by other authors, e.g. Zelikin-Borisov [20]. The optimal synthesis for this problem was determined by Fuller, and is described in detail in [20], Ch. 2, and in our paper [14], to which we refer the reader for further details. It is shown in [14] that *all the assumptions of Theorem 1 hold.* So the theorem applies, and the optimality of the synthesis follows.

5. Comparison with Other Definitions

Our concept of regular presynthesis can be compared with other definitions of synthesis that either have been proposed by other authors or are sufficiently natural to be worth considering as possible ways to define this notion. A detailed comparison is carried out in [14]. Here we shall limit ourselves to outlining the results.

Perhaps the simplest conceivable notion of an "optimal synthesis" is that of an "optimal feedback control." In principle, one could define a "feedback control law" on a set S to be a map $v : S \to U$. If v is such that the vector field $S \ni x \to f(x, v(x)) \in \mathbb{R}^n$ is Lipschitz-continuous, and we assume for simplicity that S is open, then v gives rise to unique maximally defined trajectories $\hat{\gamma}_x$ starting at each $x \in S$, and corresponding open-loop controls $\hat{\eta}_x$, given by $\hat{\eta}_x(t) = v(\hat{\gamma}_x(t))$. If we make the additional assumption that every trajectory $\hat{\gamma}_x$ obtained in this way reaches the target at a time $\hat{T}_x$, define $T_x = -\hat{T}_x$, and let $\gamma_x(t) = \hat{\gamma}(t + \hat{T}_x)$ and $\eta_x(t) = \hat{\eta}(t+\hat{T}_x)$ for $T_x \le t \le 0$, then we will have constructed a synthesis $\Gamma = \{(\gamma_x, \eta_x)\}_{x\in S}$—in the sense of Definition 3—with domain S.

The main drawback of such a definition is that, as is well known, for most reasonable optimal control problems there does not exist an optimal feedback that renders the map $x \to f(x, v(x))$ Lipschitz-continuous, or even continuous. So it is absolutely essential to allow "discontinuous feedback laws," and when this is done one immediately runs into the problems of

(a) the lack of a truly satisfactory notion of solution

and

(b) the lack of good theorems guaranteeing existence of solutions.

Difficulty (b) can be handled in at least two ways, namely, by

(A) incorporating into the definition requirements that imply existence and uniqueness of solutions of the closed-loop equation (CLE) $\dot{x} = f(x, v(x))$;

(B) incorporating into the definition additional requirements that imply existence—but not necessarily uniqueness—of solutions of the CLE, and adding to the specification of v a definite prescription for choosing a solution when uniqueness fails, so that the "feedback law" is no longer the closed-loop control v alone, but the *pair* (v, Γ), where $\Gamma = \{(\gamma_x, \eta_x)\}$ is a family that selects one solution of the CLE for each initial condition x.

Moreover, whether we choose (A) or (B), one has to be precise about the concept of "solution." In [14] we consider three such concepts, namely,

(1) classical solutions (i. e., absolutely continuous curves $t \to x(t)$ having the property that the equality $\dot{x}(t) = f(x(t), v(x(t)))$ holds for almost every t),

(2) Filippov solutions (cf. [8]),

(3) "CLSS solutions," that is, the "feedback solutions" defined by Clarke, Ledyaev, Subbotin and Sontag in [7].

In §5 of [14] it is shown that the concepts of CLSS solution and Filippov solution are not adequate. Example 5.4 of [14] exhibits an optimal synthesis whose trajectories are not Filippov solutions of the optimal closed-loop equation. In this example, the optimal trajectories are CLSS solutions, but the optimal closed-loop equation has many other CLSS solutions that are not optimal. In Example 5.6 of [14], we describe an optimal control problem whose optimal trajectories are not CLSS solutions of the optimal closed-loop equation.

This leaves the notion of classical solution as the only viable candidate for the concept of solution to be used in the definition of optimal synthesis. Example 5.3 of [14] shows that it may happen that the set of classical solutions of the optimal closed-loop equation contains, in addition to all the optimal trajectories, some other arcs that are not optimal.

It then follows that (A) is not an adequate way of handling difficulty (b), and we are left with (B). This was indeed the strategy followed by Boltyanskii in [3], Brunovský in [5], and Sussmann in [16]. In all these cases, an "optimal synthesis" is defined by first specifying a—not necessarily continuous—feedback control law $x \to v(x)$. The feedback v is supposed to be "piecewise smooth" in a technical sense that, among other things, guarantees the existence of trajectories for every initial condition. Uniqueness is not assumed, but the specification of v is supplemented with a prescription for selecting a trajectory when non-uniqueness occurs. So, a "synthesis" in the sense of [3], [5], and [16] is more than just an optimal feedback: it is really a pair (v, Γ) of the kind discussed in (B).

Three concepts of regular synthesis were proposed by Boltyanskii, Brunovský, and Sussmann, in [3], [5], and [16], respectively.

We show in [14], by means of examples, that the two concepts of a regular synthesis defined by Boltyanskii and Brunovský are not comparable. A Brunovský synthesis can fail to be a Boltyanskii synthesis because

a. it violates the "nonzero angle" condition (condition (Bo.8) of [14]), as in the example given by Brunovský in [5],

or

b. it violates the dimension condition (labeled (Bo.4) in [14]), as in Example 5.4 of [14],

or

c. it violates the condition called (Bo.5) in [14], as in Example 5.4 of that paper.

On the other hand, a Boltyanskii synthesis can fail to be a Brunovský synthesis because, for example, the set of cells could fail to be locally finite, as in Example 5.2 of [14]. (Most readers who are familiar with the theory of regular synthesis probably believe that Boltyanskii's definition requires the set of cells to be locally finite. A careful reading of his paper shows that this is not so. In [14] we give an example of a Boltyanskii regular synthesis whose set of cells is not locally finite.) Also, the Brunovský definition, as stated in [5], does not allow for an extra "singular set" N, which is permitted by Boltyanskii.

Brunovský imposes a rather strong uniqueness requirement (labeled (Br.3.D) in [14]). In [16] we proposed an even less restrictive formulation, involving continuous "exiting maps."

It is not hard to produce a concept of regular synthesis that contains as special cases the three definitions of [3], [5], and [16] but does not differ too much from any of them. This can be done by removing from the definitions of [3], [5], and [16] conditions that not really needed, such as the local finiteness requirements and the Boltyanskii conditions called (Bo.4), (Bo.5) and (Bo.8) in [14]. One can also eliminate the requirement that the control space U be a subset of some Euclidean space $\mathbb{R}^m$, and take U to be an arbitrary set with no extra structure, provided only that all the conditions involving differentiability of the feedback control $x \to v(x)$ are replaced by a differentiability requirement for the map $x \to \tilde{f}(x, v(x))$. One possible result is the notion that we call "BB-regular synthesis" in [14], using the letters BB to stand for "Boltyanskii and Brunovský."

It is then easy to show that if Ξ is a regular synthesis in the sense of Boltyanskii [3], Brunovský [5], or Sussmann [16], then Ξ is a BB-regular synthesis in the sense of the definition of [14]. Moreover, a BB-regular synthesis gives rise in a natural way to a regular presynthesis $\Gamma(\Xi)$ in the sense of our Definition 11, which is a synthesis if the exceptional set N is just $\{0\}$.

Our Definition 11—which was already proposed in [18]—is even more general, because:

a. We do not require that V_Γ be continuous, and only assume the weak continuity conditions.

b. We do not require that the domain of the synthesis be an open set. This is quite important because one wants the theory to apply to systems that are not locally controllable, and in those cases one usually wants to take S to be the set $\hat{S}$ of all points of Ω that can be steered to the target, and in general $\hat{S}$ is not open.

c. We do not require the existence of a partition into "cells." From our point of view, such a partition is only needed to guarantee the weak differentiability condition (11.b).

d. Even when a good partition into "cells" exists, satisfying the most restrictive conditions of both definitions, it may still happen that the crucial "finite number of steps" condition of Boltyanskii and Brunovský is violated, as in Fuller's example, but our requirements for a regular presynthesis are still met.

References

[1] Bardi, M., and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Birkäuser, Boston, 1997.

[2] Berkovitz, L. D., *Optimal Control Theory,* Springer-Verlag, New York, 1974.

[3] Boltyanskii, V. G., *Sufficient conditions for optimality and the justification of the dynamic programming principle.* SIAM J. Control Optim. **4**, 1966, pp. 326–361.

[4] Brunovský, P., *Every normal linear system has a regular time–optimal synthesis.* Math. Slovaca **28**, 1978, pp. 81–100.

[5] Brunovský, P., *Existence of regular synthesis for general problems.* J. Diff. Equations **38**, 1980, pp. 317–343.

[6] Cesari, L., *Optimization – Theory and applications.* Springer–Verlag, New York, 1983.

[7] Clarke, F. H., Yu. S. Ledyaev, A. I. Subbotin and E. D. Sontag, *Asymptotic controllability implies feedback stabilization.* IEEE Trans. Autom. Control **42**, 1997, pp. 1394–1407.

[8] Filippov, A. F., *Differential equations with discontinuous right-hand side.* Reidel, 1989.

[9] Fleming, W. H., and R. W. Rishel, *Deterministic and stochastic optimal control.* Springer–Verlag, New York, 1975.

[10] Fleming, W. H., and M. Soner, *Controlled Markov processes and viscosity solutions.* Springer–Verlag, New York, 1993.

[11] Fuller, A. T., *Relay control systems optimized for various performance criteria.* Automatic and remote control, Proc. First World Congress IFAC Moscow, vol. 1, Butterworths, 1961, pp. 510-519.

[12] Marchal, C., *Chattering arcs and chattering controls.* J. Optim Theory Appl. **11**, 1973, pp. 441–468.

[13] Piccoli, B., *Regular time-optimal syntheses for smooth planar systems.* Rend. Sem. Mat. Univ. Padova **95**, 1996, pp. 59–79.

[14] Piccoli, B., and H. J. Sussmann, *Regular synthesis and sufficient conditions for optimality.* To appear in SIAM J. Control and Optimization.

[15] Pontryagin, L.S., V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley, New York, 1962.

[16] Sussmann, H. J., *Subanalytic sets and feedback control.* J. Diff. Equations **31**, 1979, pp. 31-52.

[17] Sussmann, H. J., *Piecewise analyticity of optimal cost functions and optimal feedback.* In *Proc. Joint Automatic Control Conference*, Denver, Colo, August 1979.

[18] Sussmann, H. J., *Synthesis, presynthesis, sufficient conditions for optimality and subanalytic sets.* In *Nonlinear Controllability and Optimal Control,* H. J. Sussmann ed., Marcel Dekker, New York, 1990, pp. 1–19.

[19] Sussmann, H. J., *Recent developments in the regularity theory of optimal trajectories.* In *Linear and Nonlinear Mathematical Control Theory,* Rend. Sem. Mat. Univ. e Pol. Torino, Fascicolo speciale, 1987, pp. 149–182.

[20] Zelikin, M. I., and V. F. Borisov, *Theory of Chattering Control, with applications to astronautics, robotics, economics and engineering.* Birkhäuser, Boston, 1994.