REGULAR SYNTHESIS AND SUFFICIENCY CONDITIONS FOR OPTIMALITY

Benedetto Piccoli
SISSA-ISAS
via Beirut 2-4, 34014 Trieste, Italy

Héctor J. Sussmann*
Department of Mathematics, Rutgers University
110 Frelinghuysen Rd, Piscataway, New Jersey, NJ 08854

Abstract. The aim of this paper is to propose a new definition of "regular synthesis" more general than those suggested by other authors (e.g. Boltianskii [2], Brunovsky [5], [6]) and to give a detailed proof of the corresponding sufficiency theorem, a slightly weaker version of which had been stated in an earlier article (see [27]), with only a rough outline of the proof. We illustrate the strength of our result showing that the synthesis for the famous Fuller problem [12] satisfies our hypotheses.

Keywords. Optimal control, regular synthesis, sufficient conditions.

Ref. SISSA 68/96/M

(*) Work partially supported by NSF grant DMS95-00798. Part of this work was done while the author was a visitor of the Centre Emile Borel at the IHP in Paris for whose ospitality he is deeply grateful.

§1. Introduction

The aim of this paper is to propose a new definition of "regular synthesis" more general than those suggested by other authors (e.g. Boltianskii [2], Brunovsky [5], [6]) and to give a detailed proof of the corresponding sufficiency theorem, a slightly weaker version of which had been stated in an earlier article (see [27]), with only a rough outline of the proof. We illustrate the strength of our result showing that the synthesis for the famous Fuller problem [12] satisfies our hypotheses, but fails to obey the requirements of other sufficiency theorems such as those of [2] and [6], 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. Zelikin-Borisov [28].)

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 sufficiently smooth way, then all the members of Γ are optimal". Naturally, a good definition will be one that is 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 problem. In particular, it would be desirable for the theory to apply to well known examples where discontinuous optimal controls occur, and even controls with infinitely many switchings, as in Fuller's problem.

There are essentially two different ways of defining a synthesis. Following the approach of [2], [5], [7], [10], [16], a synthesis is a collection of trajectories generated by a piecewise regular feedback control. The set of interest (that is the set of initial points of optimal trajectories) is partitioned into a locally finite collection of submanifolds of various dimensions. More precisely, the set is stratified and for each stratum a regular feedback control is specified. An alternative strategy was followed in [14], [27], where a "synthesis" is just a collection of trajectories not necessarily arising from a feedback control. With this definition the feedback control is not sufficient for describing a synthesis, indeed a discontinuous feedback control may generate too many trajectories, some of which could fail to be optimal, as shown in Example 1 of Section 5. This same example illustrates another way in which our definition of synthesis is more general than that of [2] and [6], because we allow the cost function associated to the synthesis to be discontinuous, whereas [2] and [6] require it to be continuous

Another point that should be further investigated is the relation between the geometrical approaches of [5], [6], [16], [17], [18], [20], [21], [25], [27], and the theory of viscosity solutions [1], [11]. 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, cfr. [1], where the main result is for Lagrangian that are bounded away from zero.

§2. A sufficiency theorem

We start with some definitions that will enable us to state and prove a mild generalization of the theorem whose proof was outlined in [27]. The assumptions made here will be minimal, so as to achieve maximum generality. In subsequent sections we will impose more restrictive conditions and apply our result to other definitions of synthesis given in the literature only for more regular control systems.

We consider a control system:

$$\dot{x} = f(x,u)\,, \qquad x \in \Omega\,, \qquad u \in U\,, \tag{2.1}$$

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

$$\text{minimize} \quad \int_a^b L(\gamma(t),\eta(t))dt + \tau(\gamma(b)) \qquad \text{subject to (2.1)}\,,\ \gamma(a)=x\,,\ \text{and } \gamma(b)\in\mathcal{T}\,. \tag{2.2}$$

(A more precise formulation of the problems $\mathcal{P}(f, L, \mathcal{T}, \tau, x)$, 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) Ω *is an open subset of* $\mathbb{R}^n$,
(A2) U *is a set,*
(A3) $f : \Omega \times U \to \mathbb{R}^n$ *and* $L : \Omega \times U \to \mathbb{R}$ *are maps,*
(A4) $f(\cdot, u)$ *and* $L(\cdot, u)$ *are maps of class* $\mathcal{C}^1$ *for every fixed* $u \in U$,
(A5) $\mathcal{T}$ *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 exists 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

$$\tilde{T}_x S = T_x^c S \cap (-T_x^c S) = \{v : v \in T_x^c S \quad \text{and} \quad -v \in T_x^c S\}\,. \tag{2.3}$$

As in Sussmann [27], we use $\tilde{f}$ to denote the $\mathbb{R}^{n+1}$ valued map $(x, u) \to (f(x, u), L(x, u))$.

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 $\mathbb{R}$. A *trajectory* γ for a control η is an absolutely continuous map $\gamma : \mathrm{Dom}(\eta) \to \Omega$ that satisfies $\dot{\gamma}(t) = f(\gamma(t), \eta(t))$ 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 $\eta : [a, b] \to U$ is *admissible* if the map $\Omega \times [a, b] \ni (x, t) \to \tilde{f}_\eta(x, t) = \tilde{f}(x, \eta(t)) \in \mathbb{R}^{n+1}$ satisfies the following $\mathcal{C}^1$ Carathéodory conditions:

(C.i) $t \mapsto \tilde{f}_\eta(x, t)$ *is measurable for each* $x \in \Omega$,

(C.ii) *for every compact* $K \subseteq \Omega$ *there exists an integrable function* φ_K *such that for every* $x \in K$ *and* $t \in [a, b]$*:*

$$\|\tilde{f}_\eta(x, t)\| + \|D_x \tilde{f}_\eta(x, t)\| \leq \varphi_K(t)\,. \tag{2.4}$$

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

$$J(\gamma, \eta) = \int_T^0 L(\gamma(t), \eta(t))\, dt + \tau(\gamma^+)\,.$$

If $\Gamma \subseteq \mathrm{Adm}^0(\tilde{f}, \mathcal{T})$ is an arbitrary set of admissible pairs, we define a function $V_\Gamma : \Omega \mapsto \mathbb{R} \cup \{+\infty\}$, called the *value function of* Γ, by letting

$$V_\Gamma(x) = \inf\{J(\gamma, \eta) : (\gamma, \eta) \in \Gamma,\ \gamma^- = x\} \tag{2.5}$$

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

We can now formulate $\mathcal{P}(f, L, \mathcal{T}, \tau, x)$ precisely: $\mathcal{P}(f, L, \mathcal{T}, \tau, x)$ is the problem of finding a $(\gamma, \eta) \in \mathrm{Adm}_x^0(\tilde{f}, \mathcal{T})$ such that $J(\gamma, \eta) = V_{\tilde{f},\mathcal{T}}(x)$. Any such pair is said to be a *solution* of $\mathcal{P}(f, L, \mathcal{T}, \tau, x)$. A pair $(\gamma, \eta) \in \mathrm{Adm}^0(\tilde{f}, \mathcal{T})$ is *optimal* if it is a solution of $\mathcal{P}(f, L, \mathcal{T}, \tau, \gamma^-)$.

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:

$$\mathcal{H}(\lambda, \lambda^0, x, u) = \langle \lambda, f(x, u) \rangle + \lambda^0 L(x, u)\,, \tag{2.6}$$

$$H(\lambda, \lambda^0, x) = \inf\{\mathcal{H}(\lambda, \lambda^0, x, u) : u \in U\}\,. \tag{2.7}$$

The functions $\mathcal{H} : \mathbb{R}_n \times \mathbb{R} \times \Omega \times U \to \mathbb{R}$ and $H : \mathbb{R}_n \times \mathbb{R} \times \Omega \to \mathbb{R} \cup \{-\infty\}$ are known as the *Hamiltonian* and the *minimized Hamiltonian* of $\tilde{f}$.

We say that the pair $(\gamma, \eta) \in \mathrm{Adm}^0(\tilde{f}, \mathcal{T})$ is *extremal* if there exist an absolutely continuous map $\lambda : \mathrm{Dom}(\gamma) \mapsto \mathbb{R}_n$ and a constant $\lambda^0 \geq 0$, that satisfy:

(EX1) $(\lambda^0, \lambda(t)) \neq (0, 0)$ *for some (and hence every)* t.

(EX2) $\dot{\lambda} = -\frac{\partial \mathcal{H}}{\partial x}(\lambda(t), \lambda^0, \gamma(t), \eta(t))$ *for a.e.* $t \in \mathrm{Dom}(\eta)$.

(EX3) $H(\gamma(t), \lambda(t), \lambda^0) = \mathcal{H}(\gamma(t), \lambda(t), \lambda^0, \eta(t)) = 0$ *for a.e.* $t \in \mathrm{Dom}(\eta)$,

(EX4) $\langle \lambda(0), v \rangle = \lambda^0 \langle \nabla\tau(\gamma(0)), v \rangle$ *for all* $v \in \tilde{T}_{\gamma(0)}\mathcal{T}$.

Conditions (EX1), (EX2), (EX3), (EX4) are, respectively, the *nontriviality condition*, the *adjoint equation*, the *Hamiltonian minimization condition*, and the *transversality condition*. A pair (λ, λ^0) such that (EX2) holds is called an *adjoint vector for* (γ, η). An adjoint vector for which (EX3) 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 U of 0, and a continuous map $F : U \cap C \to S$, such that $F(x) = s + x + o(||x||)$ as $x \to 0$ via values in C.)

A *presynthesis* for our family of problems is a subset Γ of $\mathrm{Adm}^0(\tilde{f}, \mathcal{T})$ such that,

(PS) *Whenever* $(\gamma_1, \eta_1) \in \Gamma$, $(\gamma_2, \eta_2) \in \Gamma$, *and* $\gamma_1^- = \gamma_2^-$, *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 $S = \mathrm{Dom}(\Gamma)$. Clearly, giving a presynthesis on S amounts to choosing, for each $x \in S$, a $(\gamma, \eta) \in \mathrm{Adm}^0(\tilde{f}, \mathcal{T})$ such that $\gamma^- = x$.

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

Given a presynthesis Γ and an $x \in \mathrm{Dom}(\Gamma)$ we will always use (γ_x, η_x) to denote the unique pair $(\gamma, \eta) \in \Gamma$ such that $\gamma^- = x$, and will write $T_x = \min \mathrm{Dom}(\gamma_x)$, so $T_x \leq 0$ and $\mathrm{Dom}(\gamma_x) = \mathrm{Dom}(\eta_x) = [T_x, 0]$.

A presynthesis is *memoryless* if whenever $x \in S$ and $t \in \mathrm{Dom}(\eta_x)$ it follows that $y = \gamma_x(t)$ belongs to S and η_y is the restriction of η_x to the interval $[t, 0]$. A *synthesis* is a memoryless presynthesis. If each pair of a presynthesis Γ is optimal (resp. extremal) then we say that Γ is *optimal* (resp. *extremal*). In particular, a presynthesis Γ is a total optimal presynthesis if and only if $V_\Gamma \equiv V_{\tilde{f}}$.

Our goal is to prove that, under suitable regularity assumptions on Γ and V_Γ, a total extremal presynthesis is optimal. We first describe these regularity assumptions in detail.

We say that 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

$$\liminf_{y \to x} \limsup_{z \to y} V(z) \leq V(x)\,. \tag{2.8}$$

Equivalently, V is w.u.s.c. at x if there exist sequences x_j, ε_j and δ_j such that $x_j \in \Omega$, $x_j \to x$, $\varepsilon_j > 0$, $\delta_j > 0$, $\delta_j \to 0$, having the property that $V(y) \leq V(x) + \delta_j$ whenever $\|y - x_j\| < \varepsilon_j$.

Given a locally Lipschitz vector field X on Ω, we say that V has the *NDJ ("no downward jumps") property along* X if the following holds:

(NDJ) *For every γ integral curve of X, if* $\mathrm{Dom}(\gamma) = [a, b]$ *then for every* $t \in]a, b]$,

$$\limsup_{h \downarrow 0} V(\gamma(t)) - V(\gamma(t-h)) \geq 0$$

.

Finally we say that V satisfies the *weak continuity conditions* for the control problem (2.1), (2.2) 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 $u \in U$.

We now define the "weak differentiability conditions" for Γ. In this definition it is understood that $\tilde{f}(y, \eta_x(t)) = 0$ for every $t \notin \mathrm{Dom}(\eta_x)$. A presynthesis $\Gamma = \{(\gamma_x, \eta_x) : x \in S\}$ is (f, L)–*differentiable* at a point $\bar{x} \in \Omega$ if the following condition is satisfied:

(DC) *there exist a neighborhood N of $\bar{x}$ (in Ω) such that $N \subseteq \mathrm{Dom}(\Gamma)$, and an interval $[T, 0]$ such that $\mathrm{Dom}(\eta_x) \subseteq [T, 0]$ for all $x \in N$, and an $\bar{\varepsilon} > 0$, with the property that*

(DC1) *there exist, for $\varepsilon \in]0, \bar{\varepsilon}]$, integrable functions $\psi_\varepsilon : [T, 0] \to \mathbb{R}$ such that*

$$\lim_{\varepsilon \to 0} \int_T^0 \psi_\varepsilon(t)\, dt = 0\,, \tag{2.9}$$

and the inequality

$$\left\| \tilde{f}(y, \eta_x(t)) - \tilde{f}(\gamma_x(t), \eta_x(t)) - \frac{\partial \tilde{f}}{\partial x}\Big(\gamma_{\bar{x}}(t), \eta_{\bar{x}}(t)\Big) \cdot (y - \gamma_{\bar{x}}(t)) \right\| \leq \psi_\varepsilon(t).(||y - \gamma_{\bar{x}}(t)|| + ||x - \bar{x}||) \tag{2.10}$$

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

(DC2) *the map $v \to \rho_v$, where ρ_v is the integrable $\mathbb{R}^{n+1}$ valued function on $[T, 0]$ given by*

$$\rho_v(t) = \tilde{f}(\gamma_{\bar{x}}(t), \eta_{\bar{x}+v}(t)) - \tilde{f}(\gamma_{\bar{x}}(t), \eta_{\bar{x}}(t))$$

is weak*-differentiable at $v = 0$, regarded as a map into the space of $\mathbb{R}^{n+1}$–valued Borel measures. That is: for every continuous function $\alpha : [T, 0] \to \mathbb{R}$, the map $\mathbb{R}^n \ni v \to \int_T^0 \rho_v(t)\alpha(t) \in \mathbb{R}^{n+1} dt$ is differentiable at $v = 0$.

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.

We can now state the main theorem (see [27]):

Theorem 2.1. *let Ω, U, f, L be such that Assumptions (A1) – (A6) hold. Let Γ be a total extremal presynthesis. Assume that*

(2.1.a) the associated value function V_Γ satisfies the weak continuity conditions,

(2.1.b) $V_\Gamma(x) \le \tau(x)$ for every $x \in \mathcal{T}$,

(2.1.c) Γ is (f,L)–differentiable at all points in the complement of a thin set A.

Then Γ is optimal.

PROOF. We let $S = \mathrm{Dom}(\Gamma)$, so S is exactly the set of all $x \in \Omega$ that can be steered to the target $\mathcal{T}$ in finite time by an admissible control.

We first establish the relation between the differential of the value function V_Γ and the adjoint covector fields along the extremal trajectories of the presynthesis. This result is important in its own right, so we state it as a separate theorem:

Theorem 2.2. *Let Ω, U, f, L be such that Assumptions (A1) – (A6) hold. 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*

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

(2.2.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 $\lambda^0 D_y V_\Gamma(\bar{x}) = \lambda(T_{\bar{x}})$. In particular, $\lambda^0 \neq 0$, and (λ, λ^0) is unique up to multiplication by a positive constant.

PROOF OF THEOREM 2.2. We use Theorem A.5 in Appendix A with $P = \mathbb{R}^n$, letting $f_x(y,t) = f(y, \eta_x(t))$, $L_x(y,t) = L(y, \eta_x(t))$, for $y \in \Omega$, $x \in N$, $t \in [T_x, 0]$, and defining $f_x(y,t) = 0$, $L_x(y,t) = 0$, for $T \le t < T_x$. Let $\hat{V}_\Gamma(x) = \int_{T_x}^0 L(\gamma_x(t), \eta_x(t))\, dt$, so $\hat{V}_\Gamma$ is the "Lagrangian part" of our cost functional.

The map $p \mapsto \bar{x}_p$ of Theorem A.5 is the identity map, which is certainly differentiable at $\bar{x}$. So Theorem A.5 tells us that the "endpoint map" $x \mapsto \mathcal{E}(x) \stackrel{\mathrm{def}}{=} \gamma_x(0)$ and the function $\hat{V}_\Gamma$ are differentiable at $\bar{x}$, and the identity

$$\lambda(0).D_y\mathcal{E}(\bar{x}).v + \lambda^0 D_y\hat{V}_\Gamma(\bar{x}).v - \lambda(T_{\bar{x}}) \cdot v$$

$$= \lim_{\varepsilon \downarrow 0}\ \varepsilon^{-1} \int_T^0 \Big(H(\lambda(t), \lambda^0, \gamma_{\bar{x}}(t), \eta_{\bar{x}+\varepsilon v}(t)) - H(\lambda(t), \lambda^0, \gamma_{\bar{x}}(t), \eta_{\bar{x}}(t)) \Big)\, dt. \quad (2.11)$$

holds for every vector v if (λ, λ^0) is an adjoint vector for $(\gamma_{\bar{x}}, \eta_{\bar{x}})$. (Here the function λ is extended to $[T, 0]$ by letting it equal zero on $[T, T_{\bar{x}}[$. The Hamiltonian $H(\lambda(t), \lambda^0, y, \eta_x(t))$ is also defined for all $t \in [T, 0]$, by letting $H((\lambda(t), \lambda^0, y, \eta_x(t)) = \lambda(t).f(y, \eta_x(t)) + \lambda^0 L(y, \eta_x(t))$, and recalling that this expression vanishes for $t \in [T, T_x[$, because $f(y, \eta_x(t))$ and $L(y, \eta_x(t))$ have been defined to be zero for $t \in [T, T_x[$.)

Now, if (λ, λ^0) is minimizing, then the right-hand side of the inequality (2.11) is nonnegative, because $H(\lambda(t), \lambda^0, \gamma_{\bar{x}}(t), \eta_{\bar{x}}(t)) = 0$ and $H(\lambda(t), \lambda^0, \gamma_{\bar{x}}(t), \eta_{\bar{x}+\varepsilon v}(t)) \ge 0$ for almost all $t \in [T, 0]$. So

$$\lambda(0).D_y\mathcal{E}(\bar{x}).v + \lambda^0 D_y\hat{V}_\Gamma(\bar{x}).v - \lambda(T_{\bar{x}}) \cdot v \ge 0 \qquad \text{for all } v \in \mathbb{R}^n. \quad (2.12)$$

Given v, the curve $r \mapsto \gamma_{\bar{x}+rv}(0)$ is defined for r in a neighborhood of 0 in $\mathbb{R}$, and is differentiable at 0 with derivative $D_y\mathcal{E}(\bar{x}).v$. On the other hand, this curve is contained in $\mathcal{T}$, so the vector $w = D_y\mathcal{E}(\bar{x}).v$ belongs to $\tilde{T}_{\gamma_{\bar{x}}(0)}\mathcal{T}$. The transversality condition (EX4) then says that $\langle \lambda(0), w \rangle = \lambda^0 \langle \nabla\tau(\gamma(0)), w \rangle$. Therefore

$$\lambda^0 \Big(\nabla\tau(\gamma(0)).D_y\mathcal{E}(\bar{x}).v + D_y\hat{V}_\Gamma(\bar{x}).v \Big) - \lambda(T_{\bar{x}}) \cdot v \ge 0 \qquad \text{for all } v \in \mathbb{R}^n. \quad (2.13)$$

Since the left-hand side of this inequality is linear in v, it has to be equal to zero. So

$$\lambda^0\Big(\nabla\tau(\gamma(0)).D_y\mathcal{E}(\bar{x}).v + D_y\hat{V}_\Gamma(\bar{x}).v\Big) = \lambda(T_{\bar{x}})\cdot v \qquad \text{for all } v\in\mathbb{R}^n\,. \tag{2.14}$$

Clearly, the left-hand side of this equation is $\lambda^0 D_y V_\Gamma(\bar{x}).v$. Since the equality holds for all v, we have $\lambda^0 D_y V_\Gamma(\bar{x}) = \lambda(T_{\bar{x}})$.

In particular, λ^0 cannot vanish, because $\lambda^0 = 0$ would imply $\lambda(T_{\bar{x}}) = 0$, contradicting (EX1). So we may, after multiplication by a positive constant, assume that $\lambda^0 = 1$, and then

$$D_y V_\Gamma(\bar{x}) = \lambda(T_{\bar{x}}). \tag{2.15}$$

This proves, in particular, the uniqueness of $(\lambda(\cdot),\lambda^0)$ up to multiplication by a positive constant, since the Cauchy problem for the adjoint equation has unique solutions. ∎

CONTINUATION OF THE PROOF OF THEOREM 2.1. The minimization condition (EX4) of the Pontryagin Maximum Principle implies that, for a given $x\in S\setminus A$, $u\in U$, the inequality

$$\langle\lambda(t), f(\gamma_x(t),u)\rangle + L(\gamma_x(t),u)\geq 0$$

holds for almost all $t\in[T_x,0]$, where (λ,λ^0) is a minimizing adjoint vector for (γ_x,η_x), normalized so that $\lambda^0=1$. Since $t\mapsto\langle\lambda(t),f(\gamma_x(t),u)\rangle + L(\gamma_x(t),u)$ is continuous on $[T_x,0]$, the inequality is valid at time T_x. In view of Theorem 2.2, $\lambda(T_x) = dV_\Gamma(x)$, so

$$\langle dV_\Gamma(x), f(x,u)\rangle + L(x,u)\geq 0\,. \tag{2.16}$$

Now consider an arbitrary $(\gamma,\eta)\in \mathrm{Adm}^0(\tilde{f},\mathcal{T})$, with domain $[T,0]$. We want to prove that $J(\gamma,\eta)\geq V_\Gamma(x)$, if $x=\gamma^-$. It follows from Lemma 4.1 of [26] that there exists a sequence of admissible pairs (γ_j,η_j), with domain $[T,0]$, such that $\gamma_j(0)=\gamma(0)$, η_j is piecewise constant, $\gamma_j\to\gamma$ uniformly on $[T,0]$, and $J(\gamma_j,\eta_j)\to J(\gamma,\eta)$. Since the function V_Γ is lower semicontinuous, it follows that, if $V_\Gamma(\gamma_j(T))\leq J(\gamma_j,\eta_j)$, then $V_\Gamma(x)\leq J(\gamma,\eta)$. Hence it suffices to assume that the control η is piecewise constant, and we make this assumption from now on.

Let $\eta:[T,0]\to U$ be a piecewise constant control, and let $t_0,\ldots,t_m$, $u_1,\ldots,u_m$ be such that $T=t_0<t_1<\ldots<t_{m-1}<t_m=0$ and $\eta(t)=u_i$ if $t_{i-1}<t<t_i$ for $i=1,\ldots,m$. For each $y\in\Omega$, let $t\to\zeta^y(t)$ be the trajectory of η that satisfies the terminal condition $\zeta^y(0)=y$. Let W be an open subset of Ω such that $\zeta^y(T)$ is defined for all $y\in W$ and $\gamma(0)\in W$. We will show that

(#) *for almost every $y\in W$ the set $B^y=\{t:\zeta^y(t)\in A\}$ is finite or countable.*

Since A is thin, we can write $A=A_1\cup A_2$, where $A_1=\cup_j M_j$ and $\{M_j\}_{j\in J}$ is a finite or countable family of connected submanifolds of Ω of codimension $d_j>0$, while $\mathcal{H}^{n-1}(A_2)=0$. After replacing each M_j by a finite or countable family of open submanifolds of M_j, we may assume that the M_j are embedded.

For $i=1,\ldots,m$, let $\Phi_i: W\times[t_{i-1},t_i]\to\Omega$ be the map $(y,t)\to\zeta^y(t)$. Then Φ_i is a $\mathcal{C}^1$ submersion, because for each fixed t the map $y\to\Phi_i(y,t)$ is a $\mathcal{C}^1$ diffeomorphism. Therefore each set $\tilde{M}_{j,i}=\Phi_i^{-1}(M_j)$ is an embedded submanifold of $W\times[t_{i-1},t_i]$ of codimension d_j, and $\mathcal{H}^n(\Phi_i^{-1}(A_2))=0$.

Let $\pi_i: W\times[t_{i-1},t_i]\to W$ be the projection $(y,t)\to y$. Let $S_{j,i}$ be the set of points $s\in\tilde{M}_{j,i}$ such that the restriction $\pi_i\lceil\tilde{M}_{j,i}$ of π to $\tilde{M}_{j,i}$ is not regular at s. (Recall that, if S_1, S_2 are $\mathcal{C}^1$ manifolds, then a $\mathcal{C}^1$ map $F:S_1\to S_2$ is *regular* at $s\in S_1$ if the differential $d\pi(s)$ maps the tangent space T_sS_1 surjectively onto $T_{F(s)}S_2$.)

Then $\mathcal{H}^n(\pi_i(\Phi_i^{-1}(A_2))) = 0$ (see for example [8]), and $\mathcal{H}^n(\pi_i(S_{j,i})) = 0$ for each i, j, by Sard's Theorem. So the "bad" set

$$\mathcal{B} = \Big(\bigcup_{i=1}^{m} \pi_i(\Phi_i^{-1}(A_2))\Big) \cup \Big(\bigcup_{j,i} \pi_i(S_{j,i})\Big)$$

has measure zero.

Now let $y \in W \backslash \mathcal{B}$. Then $\zeta^y(t) \notin A_2$ if $T \leq t \leq 0$. (Otherwise, if $\zeta^y(t) \in A_2$, $t_{i-1} \leq t \leq t_i$, it would follow that $(y,t) \in \Phi_i^{-1}(A_2)$, so $y \in \pi_i(\Phi_i^{-1}(A_2))$, and then $y \in \mathcal{B}$.) Also, we claim that the set $\{t : \zeta^y(t) \in A_1\}$ is at most countable. To see this, it suffices to show that, for each j, i, the set $\{t : t_{i-1} \leq t \leq t_i\,,\, \zeta^y(t) \in M_j\}$ is finite. Fix i, j, and suppose $\zeta^y(t) \in M_j$. The manifold $\tilde{M}_{j,i}$ has codimension d_j in $W \times [t_{i-1}, t_i]$. Since $d_j > 0$, the dimension $\nu_{j,i}$ of $\tilde{M}_{j,i}$ is $\leq n$. Moreover, $d\pi(y,t)$ maps $T_{y,t}\tilde{M}_{j,i}$ surjectively onto $\mathbb{R}^n$ (the tangent space to Ω at y), because $y \notin \mathcal{B}$. So $\nu_{j,i} = n$, and $d\pi(y,t)$ is injective on $T_{y,t}\tilde{M}_{j,i}$. Since $\left\langle d\pi(y,t), \frac{\partial}{\partial t} \right\rangle$ is obviously equal to 0, it follows that $\frac{\partial}{\partial t} \notin T_{y,t}\tilde{M}_{j,i}$. Since $\tilde{M}_{j,i}$ is embedded, $(y,\tau) \notin \tilde{M}_{j,i}$ if $0 < |\tau - t| \leq \varepsilon$, provided that ε is sufficiently small. Therefore t is an isolated point of $\{\tau : t_{i-1} \leq \tau \leq t_i\,,\, \zeta^y(\tau) \in M_j\}$. So $\{t : t_{i-1} \leq t \leq t_i\,,\, \zeta^y(t) \in M_j\}$ is finite.

We have therefore proved that, if $y \in W \backslash \mathcal{B}$, then $\zeta^y(t) \notin A_2$ for all $t \in [T, 0]$, and the set $\{t : \zeta^y(t) \in A_1\}$ is at most countable. Since $A = A_1 \cup A_2$, and $\mathcal{B}$ has measure zero, $(\#)$ follows.

From the weak upper semicontinuity of V_Γ, there exist x_i, $\varepsilon_i > 0$, $\delta_i > 0$ such that $x_i \to \gamma(0)$, $\delta_i \to 0$ and $V_\Gamma(y) \leq V_\Gamma(\gamma(0)) + \delta_i$ whenever $\|y - x_i\| \leq \varepsilon_i$. Choose the ε_i so that $\|y - x_i\| \leq \varepsilon_i$ implies $y \in W$. Then, by $(\#)$, we can choose for each i a y_i such that $\|y_i - x_i\| \leq \varepsilon_i$ and the bad set B^{y_i} is at most finite or countable. Since Γ is total and $V_\Gamma(\zeta^{y_i}(0)) < +\infty$, it is clear that, for every i, $\zeta^{y_i}(t) \in S$ for every $t \in [T, 0]$. Hence (2.16) implies that the function:

$$\varphi_i(t) = V_\Gamma(\zeta^{y_i}(t)) + \int_T^t L(\zeta^{y_i}(s), \eta(s))ds$$

is differentiable at every t, except for a finite or countable number of values (which may include the switching times of η), with nonnegative derivative. Thus, φ_i verifies assumption (B.1.a) of Lemma B.1 (Appendix B). Moreover, the lower semicontinuity of V_Γ ensures assumption (B.1.b) of Lemma B.1 for φ_i, and the (NDJ) condition ensures assumption (B.1.c). Hence we can apply Lemma B.1 to φ_i, and conclude that $\varphi_i(T) \leq \varphi_i(0)$. So

$$V_\Gamma(\zeta^{y_i}(T)) \leq \int_T^0 L(\zeta^{y_i}(t), \eta(t))dt + V_\Gamma(y_i) \leq \int_T^0 L(\zeta^{y_i}(t), \eta(t))dt + V_\Gamma(\gamma(0)) + \delta_i.$$

Since $V_\Gamma(\gamma(0)) \leq \tau(\gamma(0))$ and $\zeta^{y_i} \to \gamma$ uniformly (and in particular $\zeta^{y_i}(T) \to \gamma(T) = x$), we conclude, using the lower semicontinuity of V_Γ, that $V_\Gamma(x) \leq J(\gamma, \eta)$.

So the desired inequality holds when η is piecewise constant. As explained earlier, this implies that Γ is optimal.

Remark 1. Notice that the weak upper semicontinuity of V_Γ everywhere is equivalent to its weak upper semicontinuity at the points of the target. Indeed, let $x \in \Omega$. If $x \notin S$, then $V_\Gamma(x) = +\infty$, and (2.8) is clearly true. Next, suppose $x \in S$, and let (γ_x, η_x) be its associated admissible pair. Assume that V_Γ is weakly upper semicontinuous at $\gamma_x(0)$. Let y_j, ε_j, δ_j be such that $y_j \to \gamma_x(0)$, $\varepsilon_j > 0$, $\delta_j > 0$, $\delta_j \to 0$, and $V_\Gamma(y) \leq V_\Gamma(\gamma_x(0)) + \delta_j$ whenever $\|y - y_j\| \leq \varepsilon_j$. Let ζ^{y_j} be the trajectory corresponding to the control η_x and such that $\zeta^{y_j}(0) = y_j$.

Since $y_j \to \gamma_x(0)$, we may assume, by picking j large enough, that ζ^{y_j} is defined on $[T_x, 0]$ for all j, and let $x_j = \zeta^{y_j}(T_x)$. Let $\tilde{\varepsilon}_j$ be such that $0 < \tilde{\varepsilon}_j < \frac{1}{j}$ and, if $\|z - x_j\| \leq \tilde{\varepsilon}_j$, then the trajectory θ^z of η_x with initial condition $\theta^z(T_x) = z$ is defined on $[T_x, 0]$ and satisfies $\|\theta^z(0) - y_j\| < \varepsilon_j$. Then, using the same reasoning as in the proof of Theorem 2.2, we can show that

$$V_\Gamma(z) \leq \int_{T_x}^{0} L(\theta^z(t), \eta_x(t))\, dt + V_\Gamma(\gamma_x(0)) + \delta_j \tag{2.17}$$

whenever $\|z - x_j\| \leq \tilde{\varepsilon}_j$. (First approximate η_x by piecewise constant controls, η^k, and let $\theta^{z,k}$ be the corresponding trajectories with terminal condition $\theta^{z,k}(0) = \theta^z(0)$. Then for each k use Statement (#) of the proof of Theorem 2.2 to approximate the terminal point $\theta^z(0)$ by points $w^{z,k,i}$ with the property that, if $\tilde{\theta}^{z,k,i}$ is the trajectory for η^k for which $\tilde{\theta}^{z,k,i}(0) = w^{z,k,i}$, it follows that $V_\Gamma(\tilde{\theta}^{z,k,i}(T_x)) \leq \int_{T_x}^{0} L(\theta^{z,k,i}(t), \eta^k(t))\, dt + V_\Gamma(w^{z,k,i})$. Since $\|\theta^z(0) - y_j\| < \varepsilon_j$, the $w^{z,k,i}$ satisfy $\|w^{z,k,i} - y_j\| < \varepsilon_j$ if i is large enough, so $V_\Gamma(\tilde{\theta}^{z,k,i}(T_x)) \leq V_\Gamma(\gamma_x(0)) + \delta_j$ when i is large enough. If we let $i \to \infty$ and then let $k \to \infty$, and use the lower semicontinuity of V_Γ, we find that (2.17) holds.)

Let $\tilde{\delta}_j = \sup\left\{ \| \int_{T_x}^{0} L(\theta^z(t), \eta_x(t))\, dt - \int_{T_x}^{0} L(\gamma_x(t), \eta_x(t))\, dt \| : \|z - x_j\| \leq \tilde{\varepsilon}_j \right\}$. Then $\tilde{\delta}_j \to 0$, because $x_j \to x$ and $\tilde{\varepsilon}_j \to 0$. On the other hand, $V_\Gamma(z) \leq V_\Gamma(x) + \delta_j + \tilde{\delta}_j$ whenever $\|z - x_j\| \leq \tilde{\varepsilon}_j$. So V_Γ is weakly upper semicontinuous at x.

§3. The Fuller phenomenon

The aim of this section is to prove that the famous control system considered by Fuller in [12]—and extensively studied by other authors, e.g. Zelikin-Borisov [28]—satisfies the hypotheses of Theorem 2.1.

Consider the system:

$$\begin{cases} \dot{x}_1 = x_2 & , \\ \dot{x}_2 = u & , \\ |u| \leq 1 & , \end{cases} \tag{3.1}$$

and the minimization problem:

$$\min \int_0^{+\infty} x_1^2(t) dt. \tag{3.2}$$

We now describe the optimal synthesis for this problem, determined by Fuller. First define the "switching locus" $\zeta = \{(x_1, x_2) : |x_1| = \xi x_2^2, x_1 x_2 \leq 0\}$, where ξ is the unique positive root of the polynomial $x^4 + x^2/12 - 1/18$, and write $\zeta = \zeta_+ \cup \zeta_-$, where $\zeta_\pm = \{(x_1, x_2) \in \zeta : \pm x_1 \geq 0\}$. Then the piecewise smooth curve ζ divides the plane into two regions A^- and A^+ consisting, respectively, of the points that lie above and below ζ. For every $x \in \mathbb{R}^2$ there exist a unique optimal trajectory $\hat{\gamma}_x$ and corresponding control $\hat{\eta}_x$, related by $\hat{\eta}_x(t) = +1$ when $\hat{\gamma}_x(t) \in A^+$, and $\hat{\eta}_x(t) = -1$ when $\hat{\gamma}_x(t) \in A^-$. The switchings occur along ζ. The curve $\hat{\gamma}_x$ reaches the origin in a finite time τ_x, and then stays there. Therefore $\hat{\eta}_x(t) = 0$ when $\hat{\gamma}_x(t) = 0$, which happens iff $t \geq \tau_x$ (see Fig. 1.)

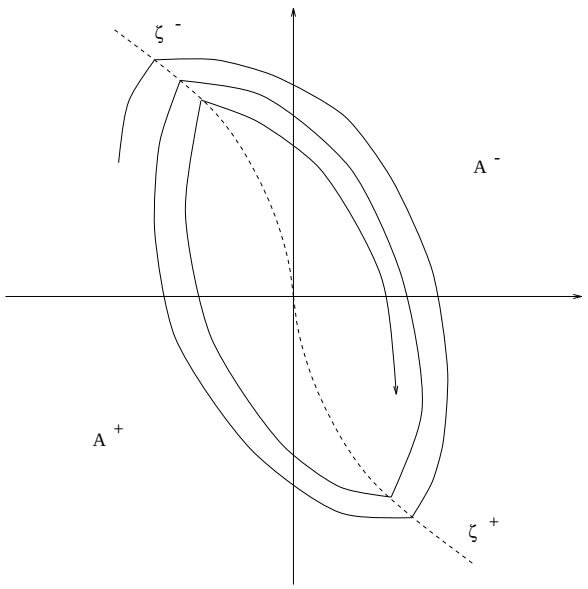

Figure 1

Therefore the admissible pairs $(\hat{\gamma}_x, \hat{\eta}_x)$ actually solve the problem of reaching the origin in finite time, subject to the dynamical constraints (3.1), and minimizing the cost (3.2). However, the sufficiency theorems of [2] and [6] do not apply, since the controls $\hat{\eta}_x$ have an infinite number of switchings. For this reason, the optimality of the synthesis obtained by Fuller is usually established by other means.

For example, one can prove optimality using the *symmetry properties* of the problem. Precisely, there exists a one parameter family $\Delta = \{\Delta_\rho\}_{\rho>0}$ of maps from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that, if (γ, η) is an admissible pair having cost c, and we let $\gamma_\rho(t) = \Delta_\rho(\gamma(\frac{t}{\rho}))$, $\eta_\rho(t) = \eta(\frac{t}{\rho})$, then (γ_ρ, η_ρ) is an admissible pair having cost $\rho^5 c$. It then follows that the Δ_ρ map optimal trajectories to optimal trajectories. The proof of optimality using the dilation symmetry is given, for example, in [28], and will not be repeated here. It will turn out, however, that the dilations Δ_ρ will play a role in our arguments, so we recall their definition and one simple property:

(1) By definition, $\Delta_\rho(x_1, x_2) = (\rho^2 x_1, \rho x_2)$.
(2) It follows from (1), in particular, that $\Delta_\rho(x_1, x_2) - x = (\rho - 1)(2x_1, x_2) + (\rho - 1)^2 (x_1, 0)$ for all ρ, x_1, x_2. This identity will be useful later.

In order to apply our Theorem 2.1, we first make a time translation and define $\gamma_x(t) = \hat{\gamma}_x(t + \tau_x)$, $\eta_x(t) = \hat{\eta}_x(t + \tau_x)$, for $-\tau_x \leq t \leq 0$. Then $\Gamma = \{(\gamma_x, \eta_x)\}_{x \in \mathbb{R}^2}$ is a memoryless family of admissible pairs such that γ_x is defined on $[-\tau_x, 0]$, $\gamma_x(-\tau_x) = x$, and $\gamma_x(0) = 0$. We want to prove that Γ is an optimal synthesis. For this purpose, we will verify the hypotheses of our Theorem 2.1, and in particular the (f, L)–differentiability of Γ at every point $x \in \mathbb{R}^2 \backslash \zeta$. Since it is very easy to check Property (DC1), we will only verify (DC2).

Fix a point $\bar{x} \in \mathbb{R}^2 \backslash \{(0,0)\}$. Then there exist s_0, s_1 such that $s_0 < s_1 < 0$, $s_1 \leq \theta s_0$,

$$\eta_{\bar{x}} = \pm \sum_{k=0}^{+\infty} (-1)^k \chi_{I_k}$$

where $I_0 = [s_0, s_1]$ and, for $k = 1, 2, \ldots$,

$$I_k = [\theta^{k-1} s_1, \theta^k s_1], \qquad \theta = \sqrt{\frac{\frac{1}{2} - \xi}{\frac{1}{2} + \xi}},$$

χ_S is the indicator function of a set S, and the sign $\pm$ depends on $\bar{x}$ as follows: it is $+$ if $\bar{x} \in A^+ \cup \zeta^+$, and $-$ if $\bar{x} \in A^- \cup \zeta^-$.

Let $[T, 0]$ be an interval containing $\mathrm{Dom}(\eta_x)$ for x near $\bar{x}$. Fix a continuous function $\alpha : [T, 0] \to \mathbb{R}$. To prove (DC2) we must study the dependence on x of the integral

$$J^\alpha(x) = \int_T^0 \alpha(t).\tilde{f}(\gamma_{\bar{x}}(t), \eta_x(t))\, dt\,, \tag{3.3}$$

and show that J^α is differentiable at $x = \bar{x}$ for every $\bar{x} \in \mathbb{R}^2 \setminus \zeta$. We will do this by showing the existence, and continuity with respect to x, of the directional derivatives $D_{v_1(x)} J^\alpha$, $D_{v_2(x)} J^\alpha$, where $v_1(x)$, $v_2(x)$ are two linearly independent directions depending smoothly on x. We choose $v_1(x)$ to be the direction of the curve $\rho \to \Delta_\rho(x)$ at $\rho = 1$, and take $v_2(x) = f(x, \eta_x(-\tau_x))$. We will just determine the directional derivatives at $\bar{x}$, and the result will make it obvious that they depend continuously on $\bar{x}$, so differentiability will follow.

Let us assume, for simplicity, that $\bar{x} \in A^+$. (The case when $\bar{x} \in A^-$ is similar.) We first differentiate J^α in the direction of $v_1(\bar{x})$. If $\rho > 0$, we have (letting $s_k = \theta^{k-1} s_1$ for $k = 2, 3, \ldots$):

$$\tau_{\Delta_\rho(\bar{x})} = \rho \tau_{\bar{x}}\,, \quad \eta_{\Delta_\rho(\bar{x})} = \sum_{k=0}^{+\infty} (-1)^k \chi_{I_k^\rho}\,, \quad I_k^\rho = [\rho s_k, \rho s_{k+1}]$$

and

$$\tilde{f}(\gamma_{\bar{x}}(t), \eta_{\Delta_\rho(\bar{x})}(t)) - \tilde{f}(\gamma_{\bar{x}}(t), \eta_{\bar{x}}(t)) = \begin{pmatrix} 0 \\ \eta_{\Delta_\rho(\bar{x})}(t) - \eta_{\bar{x}}(t) \\ 0 \end{pmatrix} + R_\rho(t), \tag{3.4}$$

where $R_\rho(t)$ is a term arising because of the possibility that $\tau_{\bar{x}} \neq \tau_{\Delta_\rho(\bar{x})}$.

Then, if $\rho \geq 1$ and $\rho - 1$ is sufficiently small,

$$\begin{aligned}
\int_T^0 (\eta_{\Delta_\rho(\bar{x})}(t) - \eta_{\bar{x}}(t))\alpha(t)dt &= \int_{\rho s_0}^0 \sum_{k=0}^{+\infty} (-1)^k (\chi_{I_k^\rho}(t) - \chi_{I_k}(t))\alpha(t)dt \\
&= \sum_{k=0}^{+\infty} (-1)^k \left(\int_{\rho s_k}^{\rho s_{k+1}} \alpha(t)dt - \int_{s_k}^{s_{k+1}} \alpha(t)dt \right) \\
&= \sum_{k=0}^{+\infty} (-1)^k \left(\int_{\rho s_k}^{s_k} \alpha(t)dt - \int_{\rho s_k}^{s_{k+1}} \alpha(t)dt \right) \\
&= \int_{\rho s_0}^{s_0} \alpha(t)dt + 2 \sum_{k=1}^{+\infty} (-1)^k \left(\int_{\rho s_k}^{s_k} \alpha(t)dt \right).
\end{aligned}$$

The numbers $(\rho - 1)^{-1} \left\| \int_{\rho s_k}^{s_k} \alpha(t)dt \right\|$ are bounded by $\theta^{k-1} |s_1| . \|\alpha\|_{L^\infty}$, which is the general term of a convergent series, since $\theta < 1$. So we can divide by $\rho - 1$ and let $\rho \to 1$ taking the limit of each term separately. We then conclude that

$$\lim_{\rho \downarrow 1} \frac{1}{\rho - 1} \int_T^0 (\eta_{x_\rho}(t) - \eta_{\bar{x}}(t))\alpha(t)dt = -s_0 \alpha(s_0) - 2 s_1 \sum_{k=1}^{+\infty} (-1)^k \theta^{k-1} \alpha(\theta^{k-1} s_1)\,. \tag{3.5}$$

We now find the limit of $(\rho-1)^{-1}\int_T^0 \alpha(t).R_\rho(t)\,dt$ as $\rho\downarrow 1$. Then $-\tau_{\Delta_\rho(\bar{x})} = -\rho\tau_{\bar{x}} < -\tau_{\bar{x}}$, so

$$\int_T^0 \alpha(t).R_\rho(t)\,dt = \int_{-\rho\tau_{\bar{x}}}^{-\tau_{\bar{x}}} \alpha(t).\begin{pmatrix} \gamma^2_{\Delta_\rho(\bar{x})}(t) \\ 0 \\ \left(\gamma^1_{\Delta_\rho(\bar{x})}(t)\right)^2 \end{pmatrix} dt\,,$$

where $\gamma^i_{\Delta_\rho(\bar{x})}(t)$, for $i=1,2$, is the i-th component of $\gamma_{\Delta_\rho(\bar{x})}(t)$. Clearly, then, if $\bar{x}=(\bar{x}_1,\bar{x}_2)$, we have

$$\lim_{\rho\downarrow 1}(\rho-1)^{-1}\int_T^0 \alpha(t).R_\rho(t)\,dt = \tau_{\bar{x}}\alpha(\tau_{\bar{x}}).\begin{pmatrix} \bar{x}_2 \\ 0 \\ \bar{x}_1^2 \end{pmatrix}. \tag{3.6}$$

Similar computations shows that (3.5) and (3.6) hold as well for the limits as $\rho\uparrow 1$. So the directional derivative of J^α at $x=\bar{x}$ in the direction of $v_1(\bar{x})$ exists and is given by the sums of the right-hand sides of (3.5) and (3.6), which depend continuously on $\bar{x}$, since s_0, s_1 and $\tau_{\bar{x}}$ (which is none other than $-s_0$) depend continuously on $\bar{x}$ as long as $\bar{x}\in A^+$.

A much simpler argument verifies the existence of the derivative of J^α in the direction of $v_2(\bar{x})$, i.e. of $f(\bar{x},1)$, whose value turns out to be the vector with components $\alpha(\tau_{\bar{x}})\bar{x}_2$, $\alpha(\tau_{\bar{x}})$, and $\alpha(\tau_{\bar{x}})\bar{x}_1^2$. Once again, this expression depends continuously on $\bar{x}$, so the proof of (DC2) is complete.

The other assumptions of Theorem 2.1 are easily checked. So the theorem applies, and the optimality of the synthesis described above follows.

§4. The Boltianskii and Brunovsky definitions of synthesis

We start by reviewing the definition of synthesis proposed by Brunovsky [6], and then we explain how it differs from the more restrictive definition given by Boltianskii [2].

We consider the system (2.1), (2.2), with the restriction that f is $\mathcal{C}^1$ as a function of (x,u).

Given a $k\in\{1,2,\ldots\}\cup\{+\infty,\omega\}$, and a manifold M of class $\mathcal{C}^k$, a *$\mathcal{C}^k$ stratification in M* of a subset S of M is a partition $\mathcal{P}$ of S into connected embedded $\mathcal{C}^k$ submanifolds of M such that $\mathcal{P}$ is locally finite in M (i.e. every compact subset of M intersects finitely many members of $\mathcal{P}$) and the following "frontier axiom" holds:

(FA) *if $P_1, P_2\in\mathcal{P}$, and $P_1\cap clos_M(P_2)\neq\emptyset$, then $P_1\subset clos_M(P_2)$ and $dim(P_1)<dim(P_2)$.*

A *$\mathcal{C}^k$-stratified subset of M* is a pair $(S,\mathcal{P})$ such that $S\subseteq M$ and $\mathcal{P}$ is a C^k stratification of S in M.

Let $S\subseteq\Omega$ be an open subset such that the origin belongs to S. Brunovsky's definition (somewhat modified by us to correct what we thought were some minor imprecisions in the original formulation), says that a *regular synthesis in S for the control problem (2.1), (2.2) with target the origin*, is a 6–tuple $\Gamma=(\mathcal{P},\mathcal{P}_1,\mathcal{P}_2,\Pi,\Sigma,v)$ such that

1) $\mathcal{P}$ is the union of $\{\{0\}\}$ and a locally finite (in S) partition $\mathcal{P}\backslash\{\{0\}\}$ of $S\setminus\{0\}$ into connected embedded $\mathcal{C}^1$ submanifolds (called "cells"),

2) $\mathcal{P}\backslash\{\{0\}\}$ is the disjoint union of $\mathcal{P}_1$ (the set of "Type I cells") and $\mathcal{P}_2$ (the set of "Type II cells"),

3) $v: S\to U$ (the "closed loop control"), $\Pi:\mathcal{P}_1\to\mathcal{P}$, and $\Sigma:\mathcal{P}_2\to\mathcal{P}_1$ are maps such that the following properties are satisfied:

3.A If $\mathcal{P}'=\{P\in\mathcal{P}:\dim P<n,\ P\neq\{0\}\}$, and $S'=\bigcup\{P\in\mathcal{P}'\}$, then (S',P') is a stratified subset of S of dimension $<n$.

3.B The function v is $\mathcal{C}^1$ on each cell.

3.C If $P_1 \in \mathcal{P}_1$ and $x \in P_1$, then $f(x, v(x)) \in T_x P_1$ (the tangent space to P_1 at x) and there exists a time t_x such that the maximally defined solution ξ_x of the initial value problem

$$\dot{\xi} = f(\xi, v(\xi)) \,, \;\; \xi(0) = x \,, \;\; \xi \in P_1$$

is such that $[0, t_x\,[\subseteq Dom(\xi_x)$ and $\lim_{t \uparrow t_x} \xi_x(t)$ exists (in Ω) and belongs to $\Pi(P_1)$.

3.D If $P_2 \in \mathcal{P}_2$, then the control v is continuous on $P_2 \cup \Sigma(P_2)$, and for each $x \in P_2$ there exists a unique curve $\xi_x : [0, t_x[\to \Omega$ such that $\xi_x\,]0, t_x[$ is a maximally defined integral curve of the vector field $f(\cdot, v(\cdot))$ on $\Sigma(P_2)$, and $\xi_x(0) = x$.

3.E On every cell P, $x \to t_x$ is a continuously differentiable function, and $(t, x) \to \xi_x(t)$, $(t, x) \to u_x(t) \stackrel{\text{def}}{=} v(\xi_x(t))$ are continuously differentiable functions of t, x on the set

$$E(P) \stackrel{\text{def}}{=} \{(t, x) : x \in P \,, \; t \in [0, t_x]\}$$

in the sense that they can be prolonged to $\mathcal{C}^1$ functions on some relatively open subset of $P \times [0, \infty[$ containing $E(P)$.

3.F For every $x \in S \backslash \{0\}$, if we let $\tilde{\xi}_x$ denote the curve obtained in an obvious way by piecing together the trajectories on every single cell, and write $\tilde{\eta}_x(t) = v(\tilde{\xi}_x(t))$, then the admissible pair $(\tilde{\xi}_x, \tilde{\eta}_x)$ ends at the origin after passing through a finite number of cells, and is extremal.

3.G The value function V_Γ is continuous.

Boltianskii's definition, as given in [2], imposes the following additional requirements:

4.A $dim(\Pi(P)) = \dim(P) - 1$ whenever $P \in \mathcal{P}_1$.

4.B The only zero dimensional cell is $\{0\}$.

4.C Every trajectory ξ_x must enter a new cell "at a nonzero angle". (That is, more precisely: whenever $P \in \mathcal{P}_1$ and $x \in P$, the tangent vector to ξ_x at time t_x is not tangent to $\Pi(P)$ at $\xi_x(t_x)$.)

Remark 2. Another definition —closely related to Brunovsky's— was proposed in [16], and used to prove an existence theorem for stabilizing synthesis. The definition of Sussmann [27] —which is the one used in this paper— is more general in a number of ways, e.g. since it is not based on the existence of any feedback control generating the trajectories.

The examples of the following section show optimal syntheses that do not match the definitions above. Let Γ be the synthesis, in Brunovsky's sense, of Example 1. We have that V_Γ is discontinuous, hence Γ does not satisfy the point 3.*G* of the Brunovsky definition.

The same example show another important feature of Brunovsky's definition. It shows the necessity of specifying, for each cell $P_2 \in \mathcal{P}_2$, a cell $\Sigma(P_2) \in \mathcal{P}_1$ entered by every trajectory ξ_x (of 3.D) for $x \in P_2$. Indeed, if we do not define the map Σ for Example 1 then some not optimal trajectories would be admissible solutions of the corresponding discontinuous feedback v. The hypothesis of the uniqueness of the ξ_x entering $\Sigma(P_2)$ can be relaxed after adding a specification of ξ_x on an initial interval of time. This was done in [16], [25], where for every $x \in P_2 \in \mathcal{P}_2$ we have

an assigned trajectory ξ_x, defined for an interval of time of length ε_x, that enters $\Sigma(P_2)$. In any case it is important to specify, for every $P_2 \in \mathcal{P}_2$, the cell the trajectory ξ_x, $x \in P_2$, is supposed to enter. If we allow all the trajectories of the discontinuous feedback v then we obtain some not optimal trajectories. For example, starting from the Type II cell $R_3 = \{(-1,0)\}$ we can use the trajectory corresponding to the control $+1$ for time 2π, reaching again the same point, and then go to the origin using the control -1. The curve γ_1 defined in this way is an admissible trajectory of v, indeed at the point $(-1,0)$ the two controls ± 1 determine the same velocity vector. However, γ_1 is clearly not time optimal. We can also define the trajectory γ_2 of v, with domain $[0,+\infty[$, corresponding to the constant control $+1$ and starting from $(-1,0)$. This trajectory neither reaches the origin in finite time nor approaches it as time goes to infinity. Notice that the trajectories γ_1 and γ_2 satisfy the equation $\dot{\gamma}(t) = f(\gamma(t), v(\gamma(t)))$ for every t in their domain. Hence they satisfy the equation corresponding to the discontinuous feedback v in any reasonable *weak* sense. For a classical definition of solutions to a discontinuous ODE we refer to [9].

The synthesis of Fuller example does not match the definitions above despite its evident *regularity*. Indeed since every trajectory has an infinite number of switchings the synthesis does not satisfy assumption 3.F.

More dramatic is the situation for the definition given by Boltianskii. The synthesis of Example 1 does not match the requirements 4.A and 4.B: the cells R_7 and R_8 are one dimensional Type I cells and $\Pi(R_8) = R_7$, the cell $R_3 = \{(-1,0)\}$ is a zero dimensional cell different from the origin. The requirements 4.A and 4.B are not satisfied even by very simple synthesis Γ for which V_Γ is continuous. In [3], [4], [14], [15], [19], [20], [23], [24], [25], it was studied the time optimal synthesis and the time optimal trajectories properties for control affine single–input two dimensional systems. There are some singular trajectories called turnpike. In general these turnpikes are one dimensional cells of the time optimal synthesis with target the origin and they are not necessarily ending at the origin but enter other one dimensional or zero dimensional cells. For some explicit examples see [15].

The synthesis of Example 2 does not match the assumption 4.A as it is easily shown, since some trajectories leave a three dimensional cell entering a one dimensional cell.

Finally, in [6] it was shown an example not satisfying hypothesis 4.C.

§5. Examples

Example 1.

We study the time-optimal synthesis for the planar system:

$$\dot{q} = F(q) + u\,G(q)\,, \qquad |u| \le 1\,, \tag{5.1}$$

where

$$q \doteq \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2\,, \qquad F(q) \doteq \begin{pmatrix} 1 - \frac{y}{2} \\ \\ \frac{x+1}{2} \end{pmatrix}\,, \qquad G(q) \doteq \begin{pmatrix} -\frac{y}{2} \\ \\ \frac{x+1}{2} \end{pmatrix}\,, \tag{5.2}$$

and the target is the origin of $\mathbb{R}^2$.

The trajectories of (5.1) corresponding to the constant control $u \equiv -1$ are straight horizontal lines going from left to right, while those corresponding to $u \equiv +1$ are circles centered at the point (-1,1), running counterclockwise. The optimal synthesis, to be determined below, is shown in Fig.2.

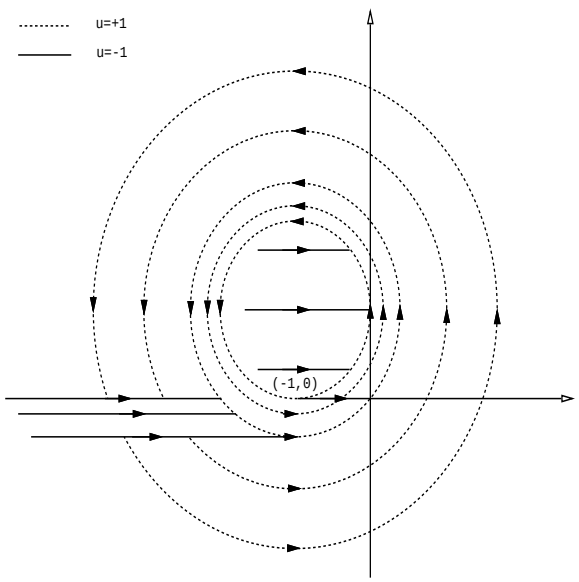

Figure 2

An arc $\gamma : [a,b] \to \mathbb{R}^2$ is *simple* if γ is an injective map. It is clear that an optimal trajectory is simple.

Since $G(q) = 0$ only for $q = (-1,0)$, it is clear that a trajectory $\gamma : [a,b] \to \mathbb{R}^2$ uniquely determines the corresponding control $u(\cdot)$, unless the set $\{t : \gamma(t) = (-1,0)\}$ is of positive measure. In particular, for simple trajectories γ —the only ones that will be of interest to us— the control is uniquely determined by γ.

We use $\mathbb{R}^k_*$ to denote the space of real row vectors with k entries, while $\mathbb{R}^k$ denotes the space of column vectors. If Ω is an open subset of $\mathbb{R}^k$ and $X : \Omega \to \mathbb{R}^k$ is a vector field of class C^1, then ∇X will denote the Jacobian matrix of X, i.e. the $k \times k$-matrix-valued function on Ω whose rows are the gradients of the components of X. If X, Y are vector fields of class C^1, then $[X,Y]$ will denote the Lie bracket of X and Y, so $[X,Y] = (\nabla Y).X - (\nabla X).Y$.

If X, Y are $\mathbb{R}^2$-valued functions on a subset of $\mathbb{R}^2$, then (X,Y) will denote the 2×2 matrix with columns X, Y. In particular, we can define the functions, see [23]:

$$\Delta_A = det\,(F,G) \qquad \Delta_B = det\,(G,[F,G]) \tag{5.3}$$

where "*det*" stands for determinant. from the results of [23], we know that every extremal trajectory of (5.1) is a finite concatenation of "bang" and "singular" arcs. A bang arc is a trajectory for a constant control with values in the set $\{-1,1\}$ of endpoints of the control interval. Singular arcs lie in the set of zeros of the function Δ_B. In our case,

$$\Delta_A = \frac{x+1}{2} \quad \text{and} \quad \Delta_B = -\frac{y}{4} \tag{5.4}$$

so the only singular arcs are those entirely contained in the x axis, and these correspond to the control -1, so they actually are bang-bang. Therefore *all the extremals of our problem are bang-bang.*

If $\Delta_A(q) \neq 0$ then the vectors $F(q)$ and $G(q)$ are linearly independent, and we can define the numbers $f(q)$ and $g(q)$ as the coefficients of the linear combination:

$$[F,G](q) = f(q)F(q) + g(q)G(q). \tag{5.5}$$

Let Ω_A, Ω_B be, respectively, the sets $\{q : \Delta_A(q) \neq 0\}$, $\{q : \Delta_B(q) \neq 0\}$, and define $\Omega_{AB} = \Omega_A \cap \Omega_B$. As was proved in [23], if γ is an extremal trajectory and $\gamma(t) \in \Omega_{AB}$ for every t, then γ is bang-bang with at most one switching, and this switching can only be from the control value -1 to the value $+1$ if $f > 0$, and from $+1$ to -1 if $f < 0$. In our case, it follows from (5.4) that

$$\Omega_A = \{(x, y) : x \neq -1\}, \quad \Omega_B = \{(x, y) : y \neq 0\}, \tag{5.6}$$

and

$$[F, G] = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} = \frac{y}{2(x+1)} F + \frac{2-y}{2(x+1)} G, \tag{5.7}$$

so Ω_{AB} is the union of four open quadrants $\Omega_{AB,i}$, $i = 1, 2, 3, 4$, defined by

$$\begin{aligned} \Omega_{AB,1} &= \{(x, y) : x > -1\,,\ y > 0\}, \\ \Omega_{AB,2} &= \{(x, y) : x < -1\,,\ y > 0\}, \\ \Omega_{AB,3} &= \{(x, y) : (x, -y) \in \Omega_{AB,2}\}, \\ \Omega_{AB,4} &= \{(x, y) : (x, -y) \in \Omega_{AB,1}\}. \end{aligned} \tag{5.8}$$

Clearly, $f > 0$ on $\Omega_{AB,1} \cup \Omega_{AB,3}$ and $f < 0$ on $\Omega_{AB,2} \cup \Omega_{AB,4}$. So the control switchings are from -1 to 1 on $\Omega_{AB,1} \cup \Omega_{AB,3}$ and from 1 to -1 on $\Omega_{AB,2} \cup \Omega_{AB,4}$.

To construct an optimal synthesis, we will actually construct a collection $\mathcal{F}$ of simple trajectories such that every optimal trajectory ending at the target is in $\mathcal{F}$, and for every initial point q we will find a $\gamma_q \in \mathcal{F}$ such that if $\delta \in \mathcal{F}$ steers Q to 0 and $\delta \neq \gamma_q$ then δ is not optimal. Since the existence of optimal trajectories from any initial point follows by elementary arguments, we will be able to conclude that γ_q is optimal.

Suppose that $\gamma : [a, b] \to \mathbb{R}^2$ is time-optimal, $\gamma(b) = (0, 0)$, $a < b$, and γ arises from an open-loop control $t \to u(t)$. Then $\gamma$is simple, and is a concatenation of finitely many maximal bang arcs. Since γ is an extremal, we can fix a nontrivial minimizing adjoint vector λ along γ and a constant $\lambda_0 \geq 0$ for which (EX4) holds. Define the *switching function*

$$\varphi(t) = \lambda(t) \cdot G(\gamma(t)).$$

The minimization condition implies that

$$\varphi(t) \neq 0 \implies u(t) = -sgn(\varphi(t)). \tag{5.9}$$

Let $\delta : [c, d] \to \mathbb{R}^2$ be a maximal $u = -1$ piece of γ, so $c < d$ and δ is of the form:

$$\delta(t) = (x_0 + t, y_0). \tag{5.10}$$

For a.e. $t \in [c, d]$, we have $\dot{\lambda} = -\lambda \cdot \nabla(F - G) = 0$. So $\lambda = (\lambda_1, \lambda_2)$ is constant and:

$$\varphi(t) = -\lambda_1 \frac{y_0}{2} + \lambda_2 \frac{x_0 + t + 1}{2} \tag{5.11}$$

is a linear function of t. Since a linear function cannot have more than one zero unless it vanishes identically, we conclude that $a < c < d < b$ can only be true if $y_0 = 0$. If $d = b$, then of course $y_0 = 0$, since $\gamma(b) = (0, 0)$. So one of the following two possibilities occurs:

(I) δ is contained in the x axis,
(II) $a = c$.

This shows that all the maximal $u = -1$ pieces of γ are entirely contained in the x axis, with the only possible exception that γ may contain an initial horizontal segment $[a, d]$ for which $y \neq 0$.

Let us now analyze the adjoint equation for a maximal $u = +1$ piece $\delta : [c, d] \to \mathbb{R}^2$ of γ, such that $c < d$. Clearly, δ is of the form:

$$x(t) = -1 + r\cos(t + \theta)\,, \qquad y(t) = 1 + r\sin(t + \theta) \qquad c \leq t \leq d\,, \tag{5.12}$$

where $r > 0$ is a constant and $\theta \in [0, 2\pi]$. (The possibility that $r = 0$ is obviously excluded, because γ is time-optimal and $c < d$.) The components λ_1, λ_2 of λ satisfy $\dot\lambda_1 = -\lambda_2$ and $\dot\lambda_2 = -\lambda_1$, so there exist constants $A > 0$, $\theta_1 \in [0, 2\pi]$, such that

$$\lambda_1(t) = A\cos(t + \theta_1)\,, \qquad \lambda_2(t) = A\sin(t + \theta_1) \qquad c \leq t \leq d\,. \tag{5.13}$$

Therefore, up to multiplication by a positive constant,

$$\varphi(t) = -\cos(t + \theta_1) + r\sin(\theta_1 - \theta)\,. \tag{5.14}$$

Now suppose that one of the two endpoints c, d of the domain of δ is a switching time and the corresponding point in $\mathbb{R}^2$ lies in the x axis. That is, we assume that $h \in \{c, d\}$ is such that $\varphi(h) = 0$ and $y(h) = 0$. Then

$$r\sin(h + \theta) = -1\,, \tag{5.15}$$

so $\varphi(t) = -\cos(t + \theta_1) + r\Big(\sin(h + \theta_1)\cos(h + \theta) - \cos(h + \theta_1)\sin(h + \theta)\Big)$, and then

$$\varphi(t) = -\cos(t + \theta_1) + \cos(h + \theta_1) + r\sin(h + \theta_1)\cos(h + \theta)\,. \tag{5.16}$$

In particular,

$$0 = \varphi(h) = r\sin(h + \theta_1)\cos(h + \theta)\,, \tag{5.17}$$

so $\varphi(t) = -\cos(t + \theta_1) + \cos(h + \theta)$. If $r \neq 1$ then (5.15) implies $0 \geq \sin(h + \theta) \neq -1$, so $\cos(h + \theta) \neq 0$. But then (5.17) implies $\sin(h + \theta_1) = 0$, so $\cos(h + \theta_1) = \pm 1$. The possibility that $\cos(h + \theta_1) = 1$ is excluded, for in that case φ would be > 0 throughout (c, d), contradicting the minimization condition. So $\varphi(t) = -1 - \cos(t + \theta_1)$. This shows that $\varphi(t) < 0$ for $t \in [c, d] - \{h\}$, unless $d - c = 2\pi$. But if $d - c = 2\pi$ then γ would not be simple. So $d - c < 2\pi$, and we have shown that the other endpoint h' of $[c, d]$ cannot be a switching point, and therefore $h' \in \{a, b\}$.

It then follows that one of the following possibilities must occur:

(i) $d = b$;
(ii) $a = c < d < b$ and $\gamma(d)$ lies in the x axis;
(iii) $a < c < d < b$ and $\gamma(d) = (-1, 0)$.

Indeed, if $d < b$ but $\gamma(d)$ is not in the x axis, then we could let δ' be the maximal $u = -1$ piece of γ that starts at time d, and our analysis of the $u = -1$ pieces of γ, applied to δ', would show that δ' cannot in fact occur, since δ' is not an initial segment of γ and is not contained in the x axis. So, if $d < b$ then $\gamma(d)$ lies in the x axis, and d is a switching time of γ. Our previous argument shows that $c = a$, unless $r = 1$. But if $r = 1$, and $\gamma(d)$ is in the x axis, then $\gamma(d) = (-1, 0)$. So $d < b$ implies that either (ii) or (iii) must hold.

We now combine the results about the two types of pieces and provide a complete description of the optimal trajectories ending at the origin. Let us use L, L_0, C to denote, respectively, the trajectory types "$u = -1$ and not contained in the x axis," "$u = -1$ and contained in the x axis,"

and "$u = 1$." (The symbols L and C stand for "line" and "circle.") Let us use $T_1T_2 \ldots T_m$ to denote the trajectory type "T_1 followed by T_2 ... followed by T_m".

With this notation, we have established that a nontrivial optimal trajectory γ ending at the origin is of one of the following seven types: L_0, C, L_0C, LC, CL_0, LCL_0, CL_0C. (Concatenations of six or more types are ruled out because any such concatenation would have to contain two different C pieces for which (iii) above holds, so γ would go at least twice through the point $(-1,0)$, and would not be simple. excluded for the same reason. The type L_0CL_0 is excluded because the last switching would have to happen at $(-1,0)$, so C would have to be the circle with center $(-1,1)$ and radius 1, which intersects the x axis only at $(-1,0)$. So the other switching would also happen at $(-1,0)$, and γ would not be simple. The exclusion of L_0CL_0 implies in turn that the types CL_0CL_0, LCL_0CL_0 and $L_0CL_0CL_0$ are ruled out as well. The types L_0CL_0C and LCL_0C are excluded because the switching from C to L_0 would have to happen at $(-1,0)$, so the L_0 piece following the switching has to be a segment going from $(-1,0)$ to a point $(\alpha,0)$ and α must satisfy $-1 < \alpha < 0$, for if α was ≥ 0 then the L_0 piece would already reach the origin before switching to the last C. But if $-1 < \alpha < 0$ then the C arc starting at $(\alpha,0)$ does not go through the origin. Finally, having excluded L_0CL_0C we are also forced to exclude CL_0CL_0C.) The LCL_0 trajectories are further restricted by the condition that the switching from C to L_0 must happen at the point $(-1,0)$. Let us call those LCL_0 trajectories that satisfy this extra condition "good."

It turns out that the optimal synthesis involves all seven types listed above, but not all trajectories of one of these types are optimal, so we need a finer analysis. We let $\mathcal{F}$ be the set of all simple trajectories that end at the origin and are either of one of the six types L_0, C, L_0C, LC, CL_0, CL_0C, or of type LCL_0 and good.

We now observe that $\dot{x} = 1 - y$ along a $u = 1$ arc, so $\dot{x} > 1$ as long as $y < 0$, $u = -1$. Therefore, if $q_1 = (x_1, y_1)$ and $q_2 = (x_2, y_2)$ satisfy $y_1 = y_2 \leq 0$, $x_1 < x_2$, $x_1 + x_2 = -2$, so that q_1 and q_2 can be joined both by a C arc and an L (or L_0) arc, then the C arc is contained in the lower half-plane, and is faster than the L or L_0 arc. In particular, if we let $L_0(\beta)$ denote, for $\beta > 0$, the L_0 arc going from $(-\beta, 0)$ to $(0,0)$, we see that $L_0(\beta)$ *is not optimal if* $\beta > 1$, because the piece of L_0 going from $(-\beta, 0)$ to $(-2+\beta, 0)$ can be replaced by a faster type C arc.

Next we show that $L_0(1)$ *is optimal.* It suffices to notice that there is no trajectory from $(-1,0)$ to $(0,0)$ of any of the types C, LC, LCL_0, and exactly one trajectory of each of the types L_0, L_0C, CL_0, and CL_0C, but the last three are not simple, so $L_0(1)$ is left as the only possible candidate.

We now study the optimal trajectories that are of one of the types C, L_0C, LC, and CL_0C. Let $\gamma : [a,b] \to \mathbb{R}^2$ be such a trajectory, and take $b = 0$, so $\gamma(0) = (0,0)$.

If γ is a C arc, then $a > -2\pi$, for otherwise γ would contain a full loop. For $0 < a < 2\pi$, we let $C(a)$ be the C arc $\gamma : [-a, 0] \to \mathbb{R}^2$ such that $\gamma(0) = (0,0)$. It will be shown below that the arcs $C(a)$ are all optimal.

Next suppose that γ is LC. Since the C piece satisfies (5.12), we must have $r = \sqrt{2}$ and $\theta = -\frac{\pi}{4}$. Let c be the time when the switching from L to C occurs. On $[c, 0]$, the switching function φ is given, up to multiplication by a positive constant, by:

$$\varphi(t) = \sqrt{2}\sin\left(\theta_1 - \frac{\pi}{4}\right) - \cos(t + \theta_1) = \sin\theta_1 - \cos\theta_1 - \cos(t + \theta_1).$$

So

$$\varphi(t) = (1 + \sin t)\sin\theta_1 - (1 + \cos t).\cos\theta_1$$

In addition, c cannot be $-\pi$, because $c = -\pi$ would correspond to $\gamma(c) = (-2, 2) \in Q_{AB,2}$, and we know that a switching from $u = -1$ to $u = 1$ is not permitted at points in $Q_{AB,2}$.

If $\sin\theta_1 = 0$, then $\cos\theta_1 = \pm 1$, so $\cos\theta_1 = 1$, because $\varphi(0) \leq 0$. So $\varphi(t) = 1 + \cos t$, and $\varphi(c) = 0$ would imply $c = -\pi$. Since this possibility has been ruled out, we conclude that $\sin\theta_1 \neq 0$. We

can then define

$$\psi(t) = \frac{1+\sin t}{1+\cos t}\,, \quad \nu = \frac{\cos\theta_1}{\sin\theta_1}\,.$$

Then ψ is well defined and smooth except at values of t such that $t = \pi + 2k\pi$, k an integer. In our case, ψ is defined on the whole interval $[c,0]$, except at $t=-\pi$ (if $-\pi \in [c,0]$). Since $c \neq -\pi$, $\psi(c)$ is defined, and the switching condition $\varphi(c)=0$ says that $\psi(c)=\nu$. If $\sin\theta_1 > 0$, then $\varphi(t) < 0$ iff $\psi(t) < \nu$ and $\varphi(t) = 0$ iff $\psi(t) = \nu$. Then $\psi(0) \leq \nu$ and, since $\lim_{t\to-\pi}\psi(t) = +\infty$, we see that c must satisfy $-\pi < c < 0$. If $\sin\theta_1 < 0$, then $\varphi(t) < 0$ iff $\psi(t) > \nu$. So $\frac{1}{2} = \psi(0) \geq \nu$. A straightforward computation shows that $\frac{d\psi}{dt}(t) > 0$ for $-2\pi \leq t < -\pi$. Since $\psi(-2\pi) = \psi(0) \geq \nu$, we conclude that $\psi(t) > \nu$ for $-2\pi \leq t < -\pi$. So, once again, the switching cannot occur at a point in $[-2\pi, -\pi)$, and then $c \in (-\pi, 0)$.

So we have shown in all possible cases that $-\pi < c < 0$. If $-\pi < c < -\frac{\pi}{2}$, then $\gamma(c) \in Q_{AB,2}$, so a switching from L to C is not permitted. If $c = -\frac{\pi}{2}$, then $\gamma(c)$ lies in the x axis, so the previous piece is L_0 rather than L. So $-\frac{\pi}{2} < c < 0$. If $\frac{\pi}{4} < c$, then $\gamma(c) \in Q_{AB,4}$, and the switching is not permitted either. So the only possibility left to us is $-\frac{\pi}{2} < c \leq -\frac{\pi}{4}$.

For $0 < \sigma < 2\pi$, $s > 0$, let $LC(s,\sigma)$ denote the trajectory that consists of a $u=-1$ arc during time s, followed by a $u=1$ arc during time σ, and ending at the origin. We have shown that $LC(s,\sigma)$ *is not optimal unless* $\frac{\pi}{4} \leq \sigma < \frac{\pi}{2}$.

We now show that *when* $\frac{\pi}{4} \leq \sigma < \frac{\pi}{2}$ *then* $LC(s,\sigma)$ *is optimal.* To see this, let q be the starting point of $LC(s,\sigma)$, for $s > 0$, $\frac{\pi}{4} \leq \sigma < \frac{\pi}{2}$. It is clear that $LC(s,\sigma)$ is the only LC trajectory from q to $(0,0)$. Also, it is easy to see that there are no trajectories from q to $(0,0)$ of the types L_0, C, L_0C, or good LCL_0. There is exactly one CL_0 trajectory, but it is easy to see that the time $\hat{\tau}$ along this trajectory is larger than the time τ along $LC(s,\sigma)$. (Actually, $\tau < |x| < \hat{\tau} - \frac{3\pi}{2}$.) Finally, there is exactly one CL_0C trajectory γ, which can also be ruled out. (If $\bar{x} = -1 - \sqrt{2-(y+1)^2}$, so $\bar{q} = (\bar{x}, y)$ is the point where $LC(s,\sigma)$switches from L to C, then both $LC(s,\sigma)$ and γ go through $\bar{q}$ and coincide from that point on. So it suffices to compare the parts of $LC(s,\sigma)$ and γ up to $\bar{q}$. The time along $LC(s,\sigma)$ is $\bar{x} - x$, and that along γ is $> \sqrt{(1+x)^2+(1-y)^2-1} - 1 + \frac{3\pi}{2}$, so it suffices to notice that $\sqrt{(1+x)^2+(1-y)^2-1} + \frac{3\pi}{2} + x - \sqrt{2-(y+1)^2} > 0$.) for $s > 0$, we let $L_0C(s)$ denote the trajectory that consists of a $u=-1$ arc during time s followed by a $u=1$ arc during time $\frac{\pi}{2}$, and ending at the origin. Then is is easily shown that $L_0C(s)$ *is optimal for every* $s > 0$. (Indeed, if $q = (x,0)$ is the starting point of $L_0C(s)$, for $s > 0$, then $L_0C(s)$ is the only L_0C trajectory from q to $(0,0)$, and there are no simple trajectories from q to $(0,0)$ of the types C, LC, CL_0, LCL_0 or CL_0C. There is an L_0 arc, namely, $L_0(|x|)$, but we know that this arc is not optimal, because $|x| > 1$.)

Now define $R_1 = \{(0,0)\}$, $R_2 = (-1,0)\times\{0\}$, $R_3 = \{(-1,0)\}$.

For $q = (0,0)$, let Γ_q be the trajectory γ with domain $\{0\}$ such that $\gamma(0) = (0,0)$. For $q = (x,0) \in R_2 \cup R_3$, let $\Gamma_q = L_0(|x|)$. Then the Γ_q, $q \in R_1 \cup R_3 \cup R_3$ are optimal.

For $r > 0$, let $D(r)$ be the open disc with center $(-1,1)$ and radius r, and let $\partial D(r)$ denote the boundary of $D(r)$. Let $R_4 = D(1)$, $R_5 = \partial D(1)\backslash\{(-1,0)\}$. Then a point $q \in \mathbb{R}^2$ can be joined to $(0,0)$ by a good LCL_0 trajectory if and only if $q \in R_4$, and in that case the good trajectory is unique, and we will use $LCL_0(q)$ to denote it. Then *the arcs* $LCL_0(q)$*, for* $q \in R_4$*, are optimal.* To see this, observe that if $q \in R_4$ then the only trajectories in $\mathcal{F}$ other than $LCL_0(q)$ that go from q to the origin are either LCL_0 but not good, or LC with a switching from L to C at a time c that violates the requirement that $-\frac{\pi}{2} < c$. So all the alternative candidates are excluded, and $LCL_0(q)$ must be the optimal trajectory.

A similar argument shows that if $q \in R_5$ then the unique CL_0 trajectory going from q to $(0,0)$ —henceforth labeled $CL_0(q)$— is optimal.

Next, let $R_6 = D(\sqrt{2})\backslash(R_2 \cup r_3 \cup R_4 \cup R_5)$. If $q \in R_6$, then the obvious candidate for optimality

is the arc Γ_q of type CL_0 obtained by following the C arc through q until R_2 is reached, and then continuing along R_2 towards the origin. We now prove that *the arcs* Γ_q, $q \in R_6$, *are optimal.* To see this, observe that if $q = (x, y) \in R_6$ and $y \neq 0$ then there are no simple arcs from q to $(0, 0)$ of the types L_0, C, L_0C, CL_0C. There is an LC arc, but (i) if $y < 0$ then this arc switches from L to C in $Q_{AB,4}$, so it is not optimal; (ii) if $y > 0$, then the arc is $LC(s, \sigma)$ for some $\sigma \in \pi, 2\pi)$, so it is not optimal either. If $y < 0$ or $y > 2$ or $x > -1$, then there is no good LCL_0 arc. If $0 < y < 2$ and $x < -1$ then there is a good LCL_0 arc, but it switches from L to C at a point in $Q_{AB,2}$, so this arc is not optimal. Finally, if $y < 0$ then Γ_q is the only CL_0 arc, and if $y > 0$ then there are two such arcs, but the other one is not optimal, because it contains $L_0(\beta)$ for some $\beta > 1$. This proves the optimality of the arcs Γ_q for $q \in R_6$, $y \neq 0$. A similar analysis works for $y = 0$. (Alternatively, if $q = (x, 0)$ then Γ_q is a cofinal piece of $\Gamma_{q'}$ for some $q' = (x', y') \in R_6$ such that $y' \neq 0$, so Γ_q is optimal.)

Next, we let $R_7 = \partial D(\sqrt{2}) \backslash \{(0, 0)\}$. For $q \in R_7$ there is a unique trajectory Γ_q of type C going from q to $(0, 0)$. The optimality of Γ_q for $q \in R_7$ is proved by considerations similar to those used for R_6.

We let $R_8 = (-\infty, -2) \times \{0\}$, $R_9 = \{(x, y) : 1 - \sqrt{2} < y < 0\ ,\ x < -1\ ,\ (x + 1)^2 + (1 - y)^2 > 2\}$, $R_{10} = (-\infty, -1) \times \{1 - \sqrt{2}\}$. Then $R_9 \cup R_{10}$ is exactly the set of starting points q of the trajectories $LC(s, \sigma)$, for $s > 0$, $\frac{\pi}{4} \leq \sigma < \frac{\pi}{2}$, and we know that for each $q \in R_9 \cup R_{10}$ the $LC(s, \sigma)$ trajectory is unique and optimal. We define Γ_q to be this trajectory. A similar fact is true for R_8, with $L_0C(s)$ rather than $LC(s, \sigma)$, and Γ_q is then defined to be the optimal trajectory.

Finally, we let $R_{11} = \mathbb{R}^2 \backslash (R_1 \cup \ldots \cup R_{10})$. For $q \in R_{11}$, we let Γ_q be the unique simple CL_0C trajectory from q to $(0, 0)$. We show that Γ_q *is optimal for all* $q \in R_{11}$. Let $\gamma \in \mathcal{F}$ be an extremal trajectory starting at q and verifying $\gamma(0) = 0$. There is no such trajectory of type L_0,C or L_0C. The case LC is impossible, otherwise γ would have a switching at time $c \in [-5\pi/4, -\pi/2[$. If γ is of type CL_0 then it contains a not optimal trajectory $L_0(\beta)$, $\beta > 1$.

Finally, if γ is of type LCL_0 good then $\gamma(t) = p \in R_6$ for some t, but Γ_p is not the restriction of γ to $[t, 0]$. Hence, if γ is optimal then it is of type CL_0C.

Since we have considered all the extremal trajectories this concludes the construction.

This ends the analysis and proves that the synthesis of Fig.1 is the optimal one.

Let define the feedback control $u(x, y)$ to be equal to 0 on R_1, to -1 on the set $R_2 \cup R_3 \cup R_4 \cup R_8 \cup R_9 \cup R_{10}$, and to $+1$ on $R_5 \cup R_6 \cup R_7 \cup R_{11}$. Then the $+1$ trajectory starting from the point $(-1, 0)$ is an admissible trajectory for this feedback control but it is not optimal.

Example 2.
The following example provides a synthesis that is optimal but does not match the restriction 4.A of Boltianskii's definition. Consider the system:

$$\begin{cases} \dot{x} = u \frac{\sqrt{x^2+y^2}}{\sqrt{1+x^2+y^2}} \\ \dot{y} = v \frac{\sqrt{x^2+y^2}}{\sqrt{1+x^2+y^2}} \\ \dot{z} = w \frac{1}{\sqrt{1+x^2+y^2}} \end{cases} \qquad u^2 + v^2 + w^2 \leq 1 \ \cdot$$

and the problem of stabilization to the origin in minimum time. Consider the following regions:

$$R_0 = \{(0, 0, 0)\}, \qquad R_1 = \{(x, y, z) : x = y = 0,\ z > 0\}, \qquad R_2 = \{(x, y, z) : x = y = 0,\ z < 0\},$$

$$R_3 = \{(x, y, z) : x^2 + y^2 > 0,\ z = 0\}, \qquad R_4 = \{(x, y, z) : x^2 + y^2 > 0,\ z > 0\},$$

$$R_5 = \{(x, y, z) : x^2 + y^2 > 0,\ z < 0\}.$$

We assign a feedback control $(u_i(x,y,z), v_i(x,y,z), w_i(x,y,z))$ for every region R_i in the following way:

$$u_0 = v_0 = w_0 = 0, \qquad u_1 = v_1 = 0,\ w_1 = -1 \qquad u_2 = v_2 = 0,\ w_2 = 1,$$

$$u_3 = u_4 = u_5 = -\frac{x}{\sqrt{x^2+y^2}},\ v_3 = v_4 = v_5 = -\frac{y}{\sqrt{x^2+y^2}},\ w_3 = w_4 = w_5 = 0.$$

Hence starting from a point $(x,y,z) \in R_3 \cup R_4 \cup R_5$ we first reach the z axis using control (u_3, v_3, w_3), and then, if we are not yet at the origin, we reach the origin running on the z axis. In this way we have constructed a synthesis Γ. It is possible to prove that this synthesis Γ is optimal applying Theorem 2.1. However, in this case it is immediate to verify that the associated value function

$$V_\Gamma = |z| + \sqrt{x^2+y^2}$$

satisfies the Hamilton–Jacobi equation on the set $\{(x,y,z) : x^2+y^2 > 0, z \neq 0\}$. Indeed:

$$\inf_{u^2+v^2+z^2 \le 1} \left(\frac{\partial V}{\partial x}\, u\, \frac{\sqrt{x^2+y^2}}{\sqrt{1+x^2+y^2}} + \frac{\partial V}{\partial y}\, v\, \frac{\sqrt{x^2+y^2}}{\sqrt{1+x^2+y^2}} + \frac{\partial V}{\partial z}\, w\, \frac{1}{\sqrt{1+x^2+y^2}} \right) =$$

$$= -\left\| \left(\frac{x}{\sqrt{1+x^2+y^2}}, \frac{y}{\sqrt{1+x^2+y^2}}, \frac{1}{\sqrt{1+x^2+y^2}} \right) \right\| = -1.$$

We can deduce that equation (2.16) holds at every point of $R_4 \cup R_5$. The other assumptions of Theorem 2.1, necessary to conclude the proof, are easily checked. Notice that every trajectory starting from a point of the three dimensional cell R_4, respectively R_5, reaches the one dimensional cell R_1, respectively R_2. Every trajectory starting from the two dimensional cell R_3 reaches the zero dimensional cell R_0. Hence, the synthesis Γ does not match the restriction 4.A of section 4.

Appendix A . Differentiation of trajectories with respect to a parameter .

The purpose of this appendix is to prove a general theorem on differentiation with respect to a parameter p, taking values in a normed space P, of a family of trajectories $x_p : [a,b] \mapsto \mathbb{R}^n$ of vector fields f_p. We will state and prove the theorem under minimal hypotheses, much weaker than what is actually needed for the main results of this paper, because the theorem has other applications, such as the theory of envelopes, where the more general statement is useful.

If $\dot{x}_p(t) = f_p(x_p(t), t)$, then formal differentiation with respect to p at $p = p_0$ in some direction v yields the variational equation

$$\dot{y}_{p_0,v}(t) = \frac{\partial f_{p_0}}{\partial x}\Big(x^{p_0}(t), t\Big).y_{p_0,v}(t) + \lim_{\varepsilon \downarrow 0} \frac{w_{p_0+\varepsilon v}(t) - w_{p_0}(t)}{\varepsilon}, \tag{A.1}$$

where $y_{p_0,v}(t)$ is the directional derivative of $x_p(t)$ at p_0 in the v-direction, and $w_p(t) = f_p(x_{p_0}(t), t)$. The technical problem that will concern us here is to make (A.1) rigorous even when the limit in the right-hand side only exists in a weak sense.

The variational equation that arises when the limit of (A.1) exists in a weak sense is a linear equation $\dot{y}(t) = A(t)y(t) + w(t)$ where the "input" w is the formal derivative of a function W which is not necessarily absolutely continuous or even continuous. So our first task will be to study the solutions of linear systems with "generalized inputs" w.

Let $-\infty < a < b < \infty$, and let n be a positive integer. If $A \in L^1([a,b];\mathbb{R}^{n\times n})$, and $\bar{Y} \in \mathbb{R}^n$, then it is well known that the initial value problem

$$\dot{Y}(t) = A(t)\cdot Y(t) + w(t)\,, \quad Y(a) = \bar{Y} \tag{A.2}$$

has a unique solution for every initial condition and every map $w \in L^1([a,b],\mathbb{R}^n)$. Because (A.2) can also be written formally as

$$dY(t) = A(t)\cdot Y(t)\,dt + w(t)\,dt\,, \quad Y(a) = \bar{Y} \tag{A.3}$$

we will use $Y_{A,w\,dt,\bar{Y}}$ to denote the unique solution of (A.2).

As explained before, it will be important for us to be able to solve (A.2) for "inputs" w more general than integrable functions. The appropriate class of inputs turns out to be that of *formal derivatives of bounded measurable functions*. Precisely, we let $BM([a,b],\mathbb{R}^n)$ denote the space of all bounded Lebesgue measurable functions $W : [a,b] \to \mathbb{R}^n$, and use $BM_0([a,b],\mathbb{R}^n)$ to denote the set of those $W \in BM([a,b],\mathbb{R}^n)$ such that $W(a) = 0$. (In $BM([a,b],\mathbb{R}^n)$ we do *not* identify two functions that are equal almost everywhere.) We endow $BM([a,b],\mathbb{R}^n)$ with the norm

$$||W||_{BM} = \sup\{||W(t)|| : t \in [a,b]\}\,. \tag{A.4}$$

Then $BM([a,b],\mathbb{R}^n)$ is a Banach space, and $BM_0([a,b],\mathbb{R}^n)$ is a closed subspace of $BM([a,b],\mathbb{R}^n)$. The space $C^0([a,b],\mathbb{R}^n)$ of continuous maps $W : [a,b] \mapsto \mathbb{R}^n$ is of course a closed subspace of $BM([a,b],\mathbb{R}^n)$, but $L^\infty([a,b],\mathbb{R}^n)$ is *not* a subspace of $BM([a,b],\mathbb{R}^n)$, because in $L^\infty([a,b],\mathbb{R}^n)$ we identify functions that are equal a.e., but in $BM([a,b],\mathbb{R}^n)$ we do not.

If $W \in BM_0([a,b],\mathbb{R}^n)$ is absolutely continuous, and $w = \dot{W}$, so that, formally, $dW = w\,dt$, then we can rewrite (A.2) formally as

$$dY(t) = A(t)\cdot Y(t)\,dt + dW(t) \qquad Y(a) = \bar{Y}\,. \tag{A.5}$$

A solution of (A.5) is then a continuous map $t \to Y(t)$ such that

$$Y(t) = \bar{Y} + \int_a^t A(s)\cdot Y(s)\,ds + W(t) \qquad \text{for} \quad \text{all} \quad t \in [a,b]\,. \tag{A.6}$$

Now (A.6) makes sense for arbitrary $Y \in BM([a,b],\mathbb{R}^n)$, $W \in BM_0([a,b],\mathbb{R}^n)$, even if W is not absolutely continuous. So we turn it into a definition: for a $W \in BM_0([a,b],\mathbb{R}^n)$, a *solution* of (A.5) is a map $Y \in BM([a,b],R^n)$ such that (A.6) holds. It then follows that

(EU) *for every $W \in BM_0([a,b],\mathbb{R}^n)$ and every $\bar{Y} \in \mathbb{R}^n$ there exists a unique solution Y of (A.5).*

Indeed, uniqueness is trivial, since the difference $Y^1 - Y^2$ of two solutions is a solution of the homogeneous problem (i.e. of (A.2) with $w \equiv 0$ and $\bar{Y} = 0$), so $Y^1 - Y^2 \equiv 0$. To prove existence, we observe that (A.6) is equivalent to

$$Y(t) - W(t) = \bar{Y} + \int_a^t A(s)\cdot(Y(s) - W(s))\,ds + \int_a^t A(s)\cdot W(s)\,ds \qquad \text{for} \quad \text{all} \quad t \in [a,b]\,, \tag{A.7}$$

i.e. to $dZ = AZ\,dt + AW\,dt$, $Z(a) = \bar{Y}$, with $Z = Y - W$. Since $AW \in L^1$, because $A \in L^1$ and W is bounded and measurable, the equation $\dot{Z}(t) = A(t)\cdot Z(t) + A(t)W(t)$ has a unique solution Z^* for which $Z^*(a) = \bar{Y}$. If we let $Y(t) = Z^*(t) + W(t)$, and recall that $W(a) = 0$, we see that Y is the desired solution of (A.5).

We will use $Y_{A,dW,\bar{Y}}$ to denote the solution of (A.5), whose existence and uniqueness has just been established. In addition to proving (EU), the preceding argument also gives an explicit formula for $Y_{A,dW,\bar{Y}}$, namely:

$$Y_{A,dW,\bar{Y}} = W + Y_{A,AW\,dt,\bar{Y}}\,. \tag{A.8}$$

Clearly, the function $Y_{A,AW\,dt,\bar{Y}}$ is absolutely continuous, so the regularity properties of $Y_{A,dW,\bar{Y}}$ —modulo absolutely continuous functions— are exactly the same as those of W. (For example: Y is continuous, right-continuous, or left-continuous at a point $t \in [a,b]$ if and only if W is; Y is of bounded variation iff W is; Y is absolutely continuous iff W is.)

Gronwall's inequality, applied to (A.7), tells us that

$$||Y_{A,dW,\bar{Y}}(t) - W(t)|| \le e^{||A||_{L^1}}\Big(||\bar{Y}|| + ||A||_{L^1}||W||_{BM}\Big)\,. \tag{A.9}$$

Therefore

$$||Y_{A,dW,\bar{Y}}(t)||_{BM} \le ||W||_{BM} + e^{||A||_{L^1}}\Big(||\bar{Y}|| + ||A||_{L^1}||W||_{BM}\Big)\,. \tag{A.10}$$

On the other hand, (A.6) also implies that

$$||W||_{BM} \le ||\bar{Y}|| + ||Y_{A,dW,\bar{Y}}||_{BM}\Big(1 + ||A||_{L^1}\Big)\,. \tag{A.11}$$

Theorem A.1. *Let $\{W_j\}$, $\{\bar{Y}_j\}$, $\{A_j\}$ be sequences in $BM_0([a,b],\mathbb{R}^n)$, $\mathbb{R}^n$, $L^1([a,b],\mathbb{R}^{n\times n})$, respectively. Assume that $\{A_j\}$ is bounded in $L^1([a,b],\mathbb{R}^{n\times n})$. Let $Y_j = Y_{A_j,dW_j,\bar{Y}_j}$. Then*

(A.1.1) *$\{Y_j\}$ is bounded in $BM([a,b],\mathbb{R}^n)$ if and only if $\{W_j\}$ is bounded in $BM([a,b],\mathbb{R}^n)$ and $\{\bar{Y}_j\}$ is bounded in $\mathbb{R}^n$.*

(A.1.2) *if $\{A_j\}$ converges in $L^1([a,b],\mathbb{R}^n)$ to a limit A, $\bar{Y} \in \mathbb{R}^n$, $\bar{Y}_j \to \bar{Y}$, $\{W_j\}$ is bounded in $BM([a,b],\mathbb{R}^n)$, $W \in BM_0([a,b],\mathbb{R}^n)$, and $Y = Y_{A,dW,\bar{Y}}$, then $Y_j \to Y$ almost everywhere iff $W_j \to W$ almost everywhere. In that case, $Y_j - W_j$ converges to $Y - W$ uniformly on $[a,b]$. In particular, for every $t \in [a,b]$, $Y_j(t) \to Y(t)$ iff $W_j(t) \to W(t)$.*

PROOF. Statement (A.1.1) follows from the bounds (A.10) and (A.11). To prove (A.1.2), we first observe that under the hypotheses of (A.1.2) we can apply (A.1.1) and conclude that $\{Y_j\}$ is bounded in $BM([a,b],\mathbb{R}^n)$.

Let $\tilde{Y}_j = Y_j - W_j = Y_{A_j,A_jW_j\,dt,\bar{Y}_j}$, $\tilde{Y} = Y - W = Y_{A,AW\,dt,\bar{Y}_j}$. Then

$$\tilde{Y}_j(t) = \bar{Y}_j + \int_a^t A_j(s)Y_j(s)\,ds = \bar{Y}_j + \int_a^t A_j(s)\tilde{Y}_j(s)\,ds + \int_a^t A_j(s)W_j(s)\,ds\,, \tag{A.12}$$

so

$$\tilde{Y}_j(t) - \tilde{Y}(t) = \int_a^t A(s)\Big(\tilde{Y}_j(s) - \tilde{Y}(s)\Big)\,ds + R_j(t)\,, \tag{A.13}$$

where

$$R_j(t) = \bar{Y}_j - \bar{Y} + \int_a^t \Big(A_j(s) - A(s)\Big)\tilde{Y}_j(s)\,ds - \int_a^t A(s)\Big(W_j(s) - W(s)\Big)\,ds\,, \tag{A.14}$$

and also

$$\tilde{Y}_j(t) - \tilde{Y}(t) = \bar{Y}_j - \bar{Y} + \int_a^t \Big(A_j(s) - A(s)\Big)Y_j(s)\,ds + \int_a^t A(s)\Big(Y_j(s) - Y(s)\Big)\,ds\,. \tag{A.15}$$

If $W_j - W \to 0$ a.e., then, since the sequences $\{\tilde{Y}_j\}$, $\{Y_j\}$ and $\{W_j\}$ are uniformly bounded, and $A_j \to A$ in L^1, it is clear that the first two terms of the sum defining $R_j(t)$ go to zero uniformly as $j \to \infty$. Also,

$$\left\| \int_a^t A(s)\Big(W_j(s) - W(s)\Big)\, ds \right\| \leq \int_a^b ||A(s)|| \cdot ||W_j(s) - W(s)||\, ds\,, \tag{A.16}$$

which goes to zero because $s \to ||A(s)||$ is integrable and the sequence $\{W_j - W\}$ is uniformly bounded and goes to zero almost everywhere. So, if we let $K_j = \sup\{||R_j(t)|| : a \leq t \leq b\}$, we see that $K_j \to 0$ as $j \to \infty$. Then Gronwall's inequality, applied to (A.13), gives $||\tilde{Y}_j(t) - \tilde{Y}(t)|| \leq K_j e^{||A||_{L^1}}$. So $\tilde{Y}_j - \tilde{Y} \to 0$ uniformly.

If $Y_j - Y \to 0$ a.e., then the first two terms of the right-hand side of (A.15) go to zero uniformly, and the third one is bounded by $\int_a^b ||A(s)|| \cdot ||Y_j(s) - Y(s)||\, ds$, which also goes to zero. So $\tilde{Y}_j - \tilde{Y} \to 0$ uniformly.

We have thus shown that $\tilde{Y}_j - \tilde{Y} \to 0$ uniformly if either $W_j \to W$ a.e. or $Y_j \to Y$ a.e. This clearly implies all the conclusions of (A.1.2). ∎

A very important subspace of $BM([a,b], \mathbb{R}^n)$ is $BV([a,b], \mathbb{R}^n)$, the space of all functions $W : [a,b] \mapsto \mathbb{R}^n$ such that (a) W is of bounded variation, (b) W is right-continuous at every point of $]a,b]$, and (c) $W(a) = 0$. There is a canonical correspondence between $BV([a,b], \mathbb{R}^n)$ and the dual space $C^0([a,b], \mathbb{R}_n)^*$ of the Banach space $C^0([a,b], \mathbb{R}_n)$ of continuous functions $[a,b] \mapsto \mathbb{R}_n$. (We write R_n rather than R^n because we want to think of the members of $C^0([a,b], \mathbb{R}_n)$ as *row-vector-valued functions*, cf. below.)

The members of $C^0([a,b], \mathbb{R}_n)^*$ are the *$\mathbb{R}^n$-valued Borel measures on $[a,b]$*. When $n = 1$, they are the *finite signed Borel measures on* $[a,b]$. The identification map from $BV([a,b], \mathbb{R}^n)$ to $C^0([a,b], \mathbb{R}_n)^*$ is the one that assigns to a function $W \in BV([a,b], \mathbb{R}^n)$ the unique Borel measure μ_W such that $\mu_W([a,t]) = W(t)$ for $a < t \leq b$.

From now on we identify a function $W \in BV([a,b], \mathbb{R}^n)$ with its corresponding measure μ_W, and write $\int \alpha \cdot dW$ for $\int \alpha \cdot d\mu_W$. Clearly, $\int \alpha \cdot dW$ is defined, more generally, for an arbitrary bounded Borel measurable $\mathbb{R}_n$-valued map α on $[a,b]$.

Naturally, the integral $\int \alpha\, dW$ of a scalar bounded Borel measurable function $\alpha : [a,b] \mapsto \mathbb{R}$ is also well defined, and is a vector in $\mathbb{R}^n$.

On $BV([a,b], R^n)$ there are at least three important topologies , namely, (a) the one induced by the norm $||\ldots||_{BM}$, (b) the strong topology of $BV([a,b], \mathbb{R}^n)$ as the dual of the Banach space $C^0([a,b], \mathbb{R}_n)$, (c) the weak* topology, also corresponding to the duality with $C^0([a,b], \mathbb{R}_n)$. The second one is induced by the *total variation norm*:

$$||W||_{TV} = \sup\left\{ \left| \int \alpha \cdot dW \right| : \alpha \in C^0([a,b], \mathbb{R}_n)\,,\ ||\alpha||_{sup} \leq 1 \right\}. \tag{A.17}$$

It is clear that $||W||_{BM} \leq ||W||_{TV}$ for all W in $BV([a,b], \mathbb{R}^n)$. Therefore, *a subset of $BV([a,b], \mathbb{R}^n)$ which is bounded in TV norm is also bounded in BM norm.*

We will need the following

Lemma A.1. *If a sequence $\{W_j\}$ in $BV([a,b], \mathbb{R}^n)$ weak*-converges to a $W \in BV([a,b], \mathbb{R}^n)$, then (a) $\{W_j\}$ is bounded in TV norm, (b) $W_j(a) \to W(a)$ and $W_j(b) \to W(b)$, (c) every subsequence of $\{W_j\}$ has a subsequence that converges to W at all but at most countably many points of $[a,b]$, (d) $\int_a^b \beta(s)||W_j(t) - W(t)||\, dt \to 0$ for every function $\beta \in L^1([a,b], \mathbb{R})$.*

PROOF. Statement (a) follows from the uniform boundedness theorem, and Statement (b) is trivial, since $W_j(a) = 0 = W(a)$, $W_j(b) = \int 1\, dW_j$, and $W(b) = \int 1\, dW$. Statement (d) clearly follows from (c), using the dominated convergence theorem together with the fact that boundedness in TV norm implies boundedness in BM norm.

So all we need to do is prove (c). If Z^i is the i-th component of Z for $Z \in BM([a,b], \mathbb{R}^n)$, $i = 1, \ldots, n$, then $\{W_j\}$ weak*-converges to W iff $\{W_j^i\}$ weak*-converges to W^i for each i. From this it clearly follows that it suffices to prove (c) for $n = 1$.

If $n = 1$, then every finite signed Borel measure μ on $[a,b]$ admits a unique decomposition $\mu = \mu^+ - \mu^-$, where μ^+ and μ^- are nonnegative Borel measures that are mutually singular (that is: there exist two μ-measurable sets E and F, such that $[a,b] = E \cup F$, $E \cap F = \emptyset$, $\mu^+(E) = \mu^-(F) = 0$). This gives, using the identification of $BV([a,b], R)$ with $C^0([a,b], R)^*$, a decomposition $Z = Z^+ - Z^-$ of every $Z \in BV([a,b], \mathbb{R})$ into a difference of two members Z^+, Z^- of $\mathcal{M}(a,b)$, the space of all monotonically nondecreasing functions $W : [a,b] \mapsto \mathbb{R}$ such that $W(a) = 0$ and W is right-continuous at every $t \in]a,b]$. It is well known that $||Z||_{TV} = Z^+(b) + Z^-(b)$.

Now let $\{W_{j(k)}\}_{k=1}^{\infty}$ be an arbitrary subsequence of $\{W_j\}$, and write $Z_k = W_{j(k)}$. Letting $Z_k = Z_k^+ - Z_k^-$, we have $||Z_k^+||_{TV} = Z_k^+(b) \leq ||Z_k||_{TV}$, so the sequence $\{Z_k^+\}$ is bounded in $BV([a,b], \mathbb{R})$. Similarly, $\{Z_k^-\}$ is bounded in $BV([a,b], \mathbb{R})$. Since the closed balls of $BV([a,b], R)$ are weak*-compact and metrizable, there exists a subsequence $\{Z_{k(\ell)}\}$ of $\{Z_k\}$ such that $Z_{k(\ell)}^+$ and $Z_{k(\ell)}^-$ weak*-converge to limits U, V. Then of course $Z_{k(\ell)} = Z_{k(\ell)}^+ - Z_{k(\ell)}^-$ weak*-converges to $U - V$, and also to W, so $U - V = W$. (But this does *not* mean that $U = W^+$ and $V = W^-$.)

Since the $Z_{k(\ell)}^+$ are in $\mathcal{M}(a,b)$, it is well known (and very easy to prove directly) that $\{Z_{k(\ell)}^+(t)\}$ converges to $U(t)$ at every point of continuity of U. Since U is monotonically nondecreasing, $Z_{k(\ell)}^+(t) \to U(t)$ at all but countably many t's in $[a,b]$. A similar conclusion holds for $Z_{k(\ell)}^-$ and V. So $Z_{k(\ell)}(t) \to W(t)$ for all but countably many t's, concluding our proof. ∎

As a corollary of Lemma A.1, we get the integration by parts formula:

Lemma A.2. *If $\lambda : [a,b] \mapsto \mathbb{R}_n$ is absolutely continuous, and $W \in BV([a,b], \mathbb{R}^n)$, then*

$$\lambda(b) \cdot W(b) = \int \lambda \cdot dW + \int_a^b \dot{\lambda}(t) \cdot W(t)\, dt\,. \tag{A.18}$$

PROOF. The result is obviously true, from the standard product rule for derivatives, if W is also absolutely continuous. In the general case, we use the fact that an arbitrary $W \in BV([a,b], \mathbb{R}^n)$ is weak*-limit of a sequence $\{W_j\}$ of absolutely continuous functions. Then (i) $\lambda(b) \cdot W_j(b) \to \lambda(b) \cdot W(b)$, because $W_j(b) \to W(b)$, (ii) $\int \lambda \cdot dW_j \to \int \lambda \cdot dW$, because $\{W_j\}$ weak*-converges to W, and λ is continuous, and (iii) $\int_a^b \dot{\lambda}(t) \cdot W_j(t)\, dt \to \int_a^b \dot{\lambda}(t) \cdot W(t)\, dt$, because Part (d) of Lemma A.1 implies that $\int_a^b ||\dot{\lambda}(t)||.||W_j(t) - W(t)||\, dt \to 0$, since $\dot{\lambda} \in L^1$. ∎

An important consequence of Lemma A.2 is the following

Lemma A.3. *Let $W \in BV([a,b], \mathbb{R}^n)$, $A \in L^1([a,b], R^{n \times n})$, $\bar{Y} \in \mathbb{R}^n$, and write $Y = Y_{A, dW, \bar{Y}}$. Suppose $\lambda : [a,b] \to \mathbb{R}_n$ is absolutely continuous and satisfies $\dot{\lambda}(t) = -\lambda(t)A(t)$ for almost all t. Then*

$$\lambda(b) \cdot Y(b) - \lambda(a) \cdot Y(a) = \int \lambda \cdot dW\,. \tag{A.19}$$

PROOF. Write $Y(t) = V(t) + W(t)$, where $V(t) = Y(a) + \int_a^t A(s).Y(s)\, ds$. Then

$$\lambda(b) \cdot Y(b) - \lambda(a) \cdot Y(a) = \lambda(b) \cdot V(b) - \lambda(a) \cdot V(a) + \lambda(b) \cdot W(b)\,. \tag{A.20}$$

The function $t \to \lambda(t) \cdot V(t)$ is absolutely continuous and its derivative is $\dot\lambda(t) \cdot V(t) + \lambda(t) \cdot \dot V(t)$, which equals $-\lambda(t)A(t)Y(t) + \lambda(t)A(t)Y(t) - \dot\lambda(t) \cdot W(t)$, i.e. $-\dot\lambda(t) \cdot W(t)$. So

$$\lambda(b) \cdot Y(b) - \lambda(a) \cdot Y(a) = \lambda(b) \cdot W(b) - \int_a^b \dot\lambda(t) \cdot W(t)\, dt\,. \tag{A.21}$$

Equation (A.19) follows from this and the integration by parts formula (A.18). ∎

It is clear that, if $A \in L^1([a,b], \mathbb{R}^{n\times n})$, $\bar Y \in \mathbb{R}^n$, and $W \in BM_0([a,b], \mathbb{R}^n)$, then $Y_{A,dW,\bar Y} - W$ is absolutely continuous, so $W \in BV([a,b], \mathbb{R}^n)$ iff $Y_{A,dW,\bar Y}$ is of bounded variation, which happens iff $Y_{A,dW,\bar Y} - \bar Y \in BV([a,b], \mathbb{R}^n)$.

Theorem A.2. *Let $\{W_j\}$, $\{\bar Y_j\}$, $\{A_j\}$ be sequences in $BV([a,b], \mathbb{R}^n)$, $\mathbb{R}^n$, $L^1([a,b], \mathbb{R}^{n\times n})$, respectively. Assume that $\{A_j\}$ is bounded in $L^1([a,b], \mathbb{R}^{n\times n})$ and $\{\bar Y_j\}$ is bounded in $\mathbb{R}^n$. Let $Y_j = Y_{A_j, dW_j, \bar Y_j}$. Then*

(A.2.1) *$\{Y_j - \bar Y_j\}$ is bounded in $BV([a,b], \mathbb{R}^n)$ if and only if $\{W_j\}$ is bounded in $BV([a,b], \mathbb{R}^n)$.*

(A.2.2) *If $\{A_j\}$ converges in $L^1([a,b], \mathbb{R}^n)$ to a limit A, $\bar Y_j \to \bar Y$, $W \in BV([a,b], \mathbb{R}^n)$, and we write $Y = Y_{A,dW,\bar Y}$, then $\{W_j\}$ weak*-converges to W iff $\{Y_j - \bar Y_j\}$ weak*-converges to $Y - \bar Y$. In that case, $Y_j - W_j - \bar Y_j$ converges to $Y - W - \bar Y$ strongly in $BV([a,b], \mathbb{R}^n)$. In particular, (a) $W_j(b) \to W(b)$, (b) $Y_j(b) \to Y(b)$, (c) for every $t \in [a,b]$, $Y_j(t) \to Y(t)$ iff $W_j(t) \to W(t)$.*

PROOF. We know from theorem A.1 that $\{Y_j - \bar Y_j\}$ is bounded in BM norm iff $\{W_j\}$ is. Since boundedness in BV norm implies boundedness in BM norm, we see that boundedness of one of the two sequences in BV norm implies that there is a constant C such that $||Y_j(t)|| \le C$ for all j, t, and then the integral $I_j = \int_a^b ||A_j(t)Y_j(t)||\, dt$ is bounded by a fixed constant K. But $I_j = ||Y_j - \bar Y_j - W_j||_{BV}$. So boundedness in BV of one of the sequences $\{W_j\}$, $\{Y_j\}$ implies boundedness in BV of the other one.

If either W_j weak*-converges to W, or $Y_j - \bar Y_j$ weak*-converges to $Y - \bar Y$, then the uniform boundedness theorem implies that either $\{W_j\}$ or $\{Y_j - \bar Y_j\}$ is bounded in BV norm, so both sequences are bounded in BV norm —and hence also in BM— norm by a constant C. Writing $Z_j = Y_j - \bar Y_j - W_j$, $Z = Y - \bar Y - W$, we have then $Z_j - Z = R_j^1 + R_j^2$, where

$$R_j^1(t) = \int_a^t \Big(A_j(s) - A(s)\Big) Y_j(s)\, ds\,, \qquad R_j^2(t) = \int_a^t A(s)\Big(Y_j(s) - Y(s)\Big)\, ds\,. \tag{A.22}$$

Then $||R_j^1||_{BV} \le C||A_j - A||_{L^1}$, so $||R_j^1||_{BV} \to 0$. Also,

$$||R_j^2||_{BV} \le ||A||_{L^1}.||\bar Y_j - \bar Y|| + \int_a^b ||A(t)||.||Y_j(t) - \bar Y_j - (Y(t) - \bar Y)||\, dt\,. \tag{A.23}$$

If $\{Y_j - \bar Y_j\}$ weak*-converges to $Y - \bar Y$, then Lemma A.1 implies that $||R_j^2||_{BV} \to 0$, so $Z_j \to Z$ in $BV([a,b], \mathbb{R}^n)$.

If $\{W_j\}$ weak*-converges to W, then we rewrite R_j^2 as

$$R_j^2(t) = \int_a^t A(s)\Big(Z_j(s) - Z(s)\Big)\, ds + R_j^3(t) + R_j^4(t)\,, \tag{A.24}$$

where $R_j^3(t) = \int_a^t A(s)\Big(\bar Y_j - \bar Y\Big)\, ds$ and $R_j^4(t) = \int_a^t A(s)\Big(W_j(s) - W(s)\Big)\, ds$. We then have $||R_j^3||_{BV} \le ||A||_{L^1}.||\bar Y_j - \bar Y||$ and $||R_j^4||_{BV} \le \int_a^b ||A(s)||.||W_j(s) - W(s)||\, ds$. Lemma A.1 then implies

that $||R_j^4||_{BV} \to 0$. So $Z_j(t) - Z(t) = \int_a^t A(s).(Z_j(s) - Z(s))\, ds + R_j(t)$, where $R_j = R_j^1 + R_j^3 + R_j^4$, so $||R_j||_{BV} \to 0$. Then Gronwall's inequality implies that $||Z_j - Z||_{sup} \leq e^{||A||_{L^1}} ||R_j||_{BV}$. If $Q_j(t) = \int_a^t A(s).(Z_j(s) - Z(s))\, ds$, then $||Q_j||_{BV} \leq ||A||_{L^1} ||Z_j - Z||_{sup}$, so $||Q_j||_{BV} \to 0$. Since $Z_j - Z = Q_j + R_j$, we see that $||Z_j - Z||_{BV} \to 0$.

So we have shown that if one of the two sequences $\{W_j - W\}$, $\{Y_j - \bar{Y}_j - (Y - \bar{Y})\}$ weak*-converges to 0, then $Y_j - \bar{Y}_j - W_j \to Y - \bar{Y} - W$ strongly in $BV([a,b], \mathbb{R}^n)$. This implies all the conclusions of (A.2.2). ∎

Recall that, if P, Q are normed spaces, and $F : \Omega \to Q$ is a map defined on a neighborhood Ω of a point p of P, then F is said to be *Fréchet differentiable* at p if there exists a bounded linear map L —called the *Fréchet differential* of F at p— such that $F(p') = F(p) + L(p' - p) + o(||p' - p||)$ as $p' \to p$. If X is a Banach space and $Q = X^*$, then we call F *weak* Fréchet differentiable* at p if for every $x \in X$ the map $\Omega \ni p' \mapsto \langle F(p'), x\rangle \in \mathbb{R}$ is Fréchet differentiable at p. (We use $\langle x^*, x\rangle$, for $x \in X$, $x^* \in X^*$, as an alternative notation for $x^*(x)$.)

Suppose P is a normed space, X is a Banach space, $p \in P$, $F : \Omega \to X^*$ is a map defined on a neighborhood Ω of p, and F is weak* Fréchet differentiable at p. Then for each $x \in X$ there is a bounded linear functional θ_x on P such that $\langle F(p+v), x\rangle = \langle F(p), x\rangle + \theta_x(v) + o(||v||)$ as $v \to 0$ in P. Clearly, θ_x is unique, and is given by the formula

$$\theta_x(v) = \lim_{h \to 0} h^{-1} \langle F(p + hv) - F(p), x\rangle\,. \tag{A.25}$$

For any given $v \in P$, use θ^v to denote the map $x \to \theta_x(v)$. Then θ^v is a limit of linear maps, so it is linear. If $\{h_j\}$ is any sequence going to 0, then the sequence of linear functionals ν_j on X given by $\nu_j(x) = h_j^{-1} \langle F(p + h_j v) - F(p), x\rangle$ converges pointwise to θ^v. Since each ν_j is a bounded linear functional on X, it follows for the uniform boundedness theorem that θ^v is bounded. Since $\theta^v(x) = \theta_x(v)$, which is linear with respect to v, we can conclude that $v \mapsto \theta_v$ is a linear map from P to X^*. We use $D^w F(p)$ to denote this map, and $D^w F(p).v$ to denote its value for a particular v, so $D^w F(p).v = \theta^v$, and

$$\langle D^w F(p).v, x\rangle = \lim_{h \to 0} h^{-1} \langle F(p + hv) - F(p), x\rangle\,. \tag{A.26}$$

Lemma A.4. *If P is a normed space, X is a Banach space, $p \in P$, $F : \Omega \to X^*$ is a map defined on a neighborhood Ω of p, and F is weak* Fréchet differentiable at p. Then:*

(A.4.i) *there is a $C > 0$ such that*

$$||F(p + v) - F(p)|| \leq C||v|| \qquad \textit{for all sufficiently small} \qquad v \in P\,, \tag{A.27}$$

(A.4.ii) *the linear map $D^w F(p) : P \mapsto X^*$ defined by (A.26) is bounded.*

PROOF. Suppose (A.4.i) is not true. Then there exists a sequence $\{v_j\}$ in P, converging to 0 and such that $||F(p + v_j) - F(p)|| > j||v_j||$. Define, for each j, a member σ_j of X^*, by

$$\sigma_j = \frac{F(p + v_j) - F(p) - \theta^{v_j}}{||v_j||}\,. \tag{A.28}$$

Then $\sigma_j(x) \to 0$ for each x, so the sequence $\{\sigma_j\}$ is pointwise bounded and hence uniformly bounded.

On the other hand, for each $x \in X$, the numbers $\frac{\theta^{v_j}(x)}{||v_j||}$ are bounded, because $\theta^{v_j}(x) = \theta_x(v_j)$, and θ_x is a bounded linear functional on P, so $|\theta^{v_j}(x)| \leq ||\theta_x||.||v_j||$. So, using the uniform boundedness theorem once again, we conclude that the sequence $\left\{\frac{\theta^{v_j}}{||v_j||}\right\}$ is bounded in X^*. Since $\{\sigma_j\}$ is bounded, we conclude that the sequence $\left\{\frac{F(p+v_j)-F(p)}{||v_j||}\right\}$ is bounded, contradicting the fact that $||F(p+v_j)-F(p)|| > j||v_j||$. So (A.27) holds for some $C > 0$, and we have proved (A.4.i).

It follows from (A.4.i) that $|\langle F(p+v)-F(p), x\rangle| \leq C||x||.||v||$ for all $x \in X$, if $v \in P$ is small enough. Applying this with hv in the role of v, and letting $h \downarrow 0$, we find that

$$|\langle D^w F(p).v, x\rangle| \leq C||x||.||v|| \qquad \text{for all} \qquad x \in X\,,\ \ v \in P\,. \tag{A.29}$$

Therefore $||D^w F(p).v|| \leq C||v||$, so $D^w F(p)$ is a bounded linear operator from P to X^*. ∎

We are now ready to study the differentiability properties with respect to a parameter p, belonging to a normed space P, of a family of trajectories $x_p : [a,b] \to \mathbb{R}^n$ of time-varying vector fields f_p. We will prove, under minimal technical hypotheses, that if the initial condition map $p \mapsto x_p(a)$ is differentiable at $p = 0$, and the map $p \mapsto f_p$ satisfies a "weak differentiability condition" at the trajectory x_0, then the map $p \mapsto x_p$, is weak*- differentiable at $p = 0$, as a map into $BV([a,b], \mathbb{R}^n)$, and in particular the endpoint map $p \mapsto x_p(b)$ is differentiable at $p = 0$.

A *curve in* $\mathbb{R}^n$ is a continuous map $\xi : I \mapsto \mathbb{R}^n$, whose domain $I = \mathrm{Dom}(\xi)$ is a nonempty interval. A curve ξ is an *arc* if $\mathrm{Dom}(\xi)$ is compact. The *graph* $G(\xi)$ of a curve ξ is the set $G(\xi) = \{(\xi(t), t) : t \in \mathrm{Dom}(\xi)\}$. If $\xi : [a,b] \mapsto \mathbb{R}^n$ is an arc in $\mathbb{R}^n$, and $\varepsilon > 0$, then the *ε-tube about* ξ is the set $\mathcal{T}(\xi, \varepsilon) = \{(x,t) : x \in \mathbb{R}^n, a \leq t \leq b, ||x - \xi(t)|| \leq \varepsilon\}$.

A *time-varying vector field* on $\mathbb{R}^n$ is an $\mathbb{R}^n$-valued map f whose domain $\mathrm{Dom}(f)$ is a —possibly empty— subset S of the product $\mathbb{R}^n \times \mathbb{R}$. We will use $TVVF(\mathbb{R}^n)$ to denote the set of all time-varying vector fields on $\mathbb{R}^n$. A *trajectory*, or *integral curve*, of an $f \in TVVF(\mathbb{R}^n)$ is a locally absolutely continuous curve ξ in $\mathbb{R}^n$ such that $G(\xi) \subseteq \mathrm{Dom}(f)$ and $\dot{\xi}(t) = f(\xi(t), t)$ for almost all $t \in \mathrm{Dom}(\xi)$. We use $\mathrm{Traj}(f)$ to denote the set of all trajectories of f, and $\mathrm{Traj}_c(f)$ to denote the set of all $\xi \in \mathrm{Traj}(f)$ such that $\mathrm{Dom}(\xi)$ is compact.

If $f \in TVVF(\mathbb{R}^n)$ and $\xi \in \mathrm{Traj}(f)$, we say that ξ is a *maximal* trajectory of f if it cannot be extended to a trajectory $\tilde{\xi} : \tilde{I} \mapsto \mathbb{R}^n$ of f defined on an interval $\tilde{I}$ such that $I \subseteq \tilde{I}$ but $I \neq \tilde{I}$. We use $\mathrm{MTraj}(f)$ to denote the set of all maximal trajectories of f. Given $\bar{x} \in \mathbb{R}^n$, $\bar{t} \in \mathbb{R}$, we use $\mathrm{Traj}(f, \bar{x}, \bar{t})$, $\mathrm{MTraj}(f, \bar{x}, \bar{t})$ to denote, respectively, the set of all $\xi \in \mathrm{Traj}(f)$ such that $\xi(\bar{t}) = \bar{x}$, and the set $\mathrm{Traj}(f, \bar{x}, \bar{t}) \cap \mathrm{MTraj}(f)$. Zorn's lemma implies that every $\xi \in \mathrm{Traj}(f, \bar{x}, \bar{t})$ can be extended to a $\tilde{\xi} \in \mathrm{MTraj}(f, \bar{x}, \bar{t})$, so $\mathrm{Traj}(f, \bar{x}, \bar{t}) \neq \emptyset$ iff $\mathrm{MTraj}(f, \bar{x}, \bar{t}) \neq \emptyset$. Clearly, $\mathrm{Traj}(f, \bar{x}, \bar{t}) \neq \emptyset$ iff $(\bar{x}, \bar{t}) \in \mathrm{Dom}(f)$. (Indeed, if $(\bar{x}, \bar{t}) \in \mathrm{Dom}(f)$, then the map ξ with domain $\{\bar{t}\}$ such that $\xi(\bar{t}) = \bar{x}$ is in $\mathrm{Traj}(f)$.)

We say that an $f \in TVVF(\mathbb{R}^n)$ has the *forward existence property* at a point $(\bar{x}, \bar{t})$ if for some $\varepsilon > 0$ there exists a $\xi \in \mathrm{Traj}(f, \bar{x}, \bar{t})$ with domain $[\bar{t}, \bar{t} + \varepsilon]$. We say that f has the *forward limit property* on a subset S of $\mathrm{Dom}(f)$ if, whenever $\xi : [a, b[\mapsto \mathbb{R}^n$ is a trajectory of f which is contained in a compact subset of S, it follows that $\lim_{t \uparrow b} \xi(t)$ exists.

Suppose we are given the following data:

(D.1) a normed space P,
(D.2) a family $\mathbf{f} = \{f_p\}_{p \in P}$ of time-varying vector fields on $\mathbb{R}^n$, depending on a parameter $p \in P$,
(D.3) a $p_0 \in P$,
(D.4) a trajectory $\xi : [a,b] \mapsto \mathbb{R}^n$ of f_{p_0}.

We then define functions $w_p : [a,b] \mapsto \mathbb{R}^n$ by

$$w_p(t) = f_p(\xi(t), t) - f_{p_0}(\xi(t), t) \quad \text{for} \quad t \in [a,b]\,, \tag{A.30}$$

so w_p is defined on $[a,b]$ whenever $G(\xi) \subseteq \mathrm{Dom}(f_p)$. If w_p is integrable, we let

$$W_p(t) = \int_a^t w_p(s)\, ds \qquad \text{for} \qquad t \in [a,b]\,. \tag{A.31}$$

We then let $\mathcal{W}$ be the map that sends p to $\mathcal{W}(p) = W_p$, with domain

$$\mathrm{Dom}(\mathcal{W}) = \{p \in P : G(\xi) \subseteq \mathrm{Dom}(f_p),\ w_p \in L^1([a,b], \mathbb{R}^n)\}\,. \tag{A.32}$$

We regard $\mathcal{W}$ as a map from $\mathrm{Dom}(\mathcal{W})$ to $BV([a,b], \mathbb{R}^n)$.

We then say that $\mathbf{f}$ is *weakly differentiable at p_0 along ξ* if there exists $\bar{\varepsilon} > 0$ such that

(WD1) *$\mathcal{T}(\xi, \bar{\varepsilon}) \subseteq \mathrm{Dom}(f_p)$ for every p such that $||p - p_0|| \le \bar{\varepsilon}$.*

(WD2) *For every p such that $||p - p_0|| \le \bar{\varepsilon}$, f_p has the forward limit property on $\mathcal{T}(\xi, \bar{\varepsilon})$ and the forward existence property at every (x,t) such that $a \le t < b$ and $||x - \xi(t)|| < \bar{\varepsilon}$.*

(WD3) *There exist $A \in L^1([a,b], \mathbb{R}^{n\times n})$, and functions $\psi_\varepsilon \in L^1([a,b], \mathbb{R})$, for $\varepsilon \in]0, \bar{\varepsilon}]$, such that*

$$||f_p(x,t) - f_p(\xi(t),t) - A(t).(x - \xi(t))|| \le \psi_\varepsilon(t)(||x - \xi(t)|| + ||p - p_0||) \tag{A.33}$$

whenever $a \le t \le b$, $||x - \xi(t)|| \le \varepsilon$, $||p - p_0|| \le \varepsilon$, and

$$\lim_{\varepsilon \downarrow 0} \int_a^b \psi_\varepsilon(t)\, dt = 0\,. \tag{A.34}$$

(WD4) *The function w_p defined by (A.30) is integrable for every p such that $||p - p_0|| \le \bar{\varepsilon}$ and, if $\mathcal{W}$ is defined as above, then $\mathcal{W}$ is weak* Fréchet differentiable at p_0.*

The second part of condition (WD4) says that for each $\alpha \in C^0([a,b], \mathbb{R}_n)$ the map $p \to \int \alpha \cdot dW_p$ is Fréchet differentiable at p_0. Lemma A.4 then implies —since $W_{p_0} = 0$— that there exists a bounded linear map $D^w\mathcal{W}(p_0) : P \mapsto BV([a,b], \mathbb{R}^n)$ having the property that

$$\int \alpha \cdot dW_p = \langle \mathcal{W}(p) - \mathcal{W}(p_0), \alpha \rangle = \int \alpha \cdot \Big(D^w\mathcal{W}(p_0).(p - p_0) \Big) + o(||p - p_0||) \qquad \text{as} \qquad p \to p_0 \tag{A.35}$$

for each $\alpha \in C^0([a,b], \mathbb{R}_n)$, and also that there is a constant C such that

$$||W_p||_{TV} \le C||p - p_0|| \qquad \text{for all } p \in P \text{ such that } \|p - p_0\| \text{ is sufficiently small.} \tag{A.36}$$

Theorem A.3. *Let P, $\mathbf{f}$, p_0, ξ be data as in (D.1,2,3,4). Assume that $\mathbf{f}$ is weakly differentiable at p_0 along ξ. Let $\bar{\varepsilon} > 0$ be such that (WD1) to (WD4) hold. Then*

(A.3.a) *For every $\varepsilon \in]0, \bar{\varepsilon}]$ there exists an $\delta \in]0, \varepsilon[$ such that, whenever $||p - p_0|| \le \delta$, $||\bar{x} - \xi(a)|| \le \delta$, and $\zeta \in \mathrm{MTraj}(f_p, \bar{x}, a)$, it follows that $[a,b] \subseteq \mathrm{Dom}(\zeta)$ and $||\zeta(t) - \xi(t)|| \le \varepsilon$ for all $t \in [a,b]$.*

(A.3.b) *If $\{\bar{x}_p\}_{||p-p_0|| \le \bar{\varepsilon}}$ is a family of points of $\mathbb{R}^n$ such that the map $p \mapsto \bar{x}_p$ is Fréchet differentiable at p_0 and $\bar{x}_{p_0} = \xi(a)$, then, if $\{\xi_p\}_{||p-p_0|| \le \bar{\varepsilon}}$ is a family of curves such that $\xi_p \in \mathrm{MTraj}(f_p, \bar{x}_p, a)$, and x_p denotes the restriction of ξ_p to $[a,b]$ —which is well defined for small enough $||p - p_0||$ by (A.3.a)— then the map $\mathcal{X}$ that sends each p to the function x_p is weak* Fréchet differentiable at p_0 as a map from P to $BV([a,b], \mathbb{R}^n)$. Moreover, $D^w\mathcal{X}(p_0)$ can be computed as follows. For $v \in P$,*

$$D^w\mathcal{X}(p_0).v = Y_{A,\, d(D^w\mathcal{W}(0).v),\, L.(p-p_0)}\,, \tag{A.37}$$

where $L : P \to \mathbb{R}^n$ is the differential at p_0 of the map $p \to \bar{x}_p$.

PROOF. We will assume, as we clearly may without loss of generality, that $p_0 = 0$. Also, we know that there is a $C > 0$ such that the bound (A.36) holds. So we can assume, by making $\bar{\varepsilon}$ smaller, if necessary, that $||W_p||_{TV} \leq C||p||$ whenever $||p|| \leq \bar{\varepsilon}$. By making C larger and $\bar{\varepsilon}$ smaller, if necessary, we may also assume that $C \geq 1$, and

$$8e^{||A||_{L^1}}||\psi_\varepsilon||_{L^1} < 1 \qquad \text{for} \qquad \varepsilon \in]0, \bar{\varepsilon}] \,. \tag{A.38}$$

Pick $\varepsilon \in]0, \bar{\varepsilon}[$, and choose $\delta > 0$ such that

$$C\delta < \varepsilon \,. \tag{A.39}$$

Notice that $\delta < \varepsilon$, since $C \geq 1$.

Let p, $\bar{x}$ be such that $||p|| \leq \delta$ and $||\bar{x} - \xi(a)|| \leq \delta$. Let $\zeta \in \mathrm{MTraj}(f_p, \bar{x}, a)$, and let $I = \mathrm{Dom}(\zeta)$. Then $a \in I$. Let c be the supremum of those t such that $a \leq t \leq b$ and $||\zeta(s) - \xi(s)|| \leq \varepsilon$ for $a \leq s \leq t$. Then $a < c \leq b$, $c \in I$, and either $c = b$ or $||\zeta(c) - \xi(c)|| = \varepsilon$. (The fact that $a < c$ follows from the local existence property, since $||\bar{x} - \xi(a)|| \leq \delta < \varepsilon < \bar{\varepsilon}$, and ζ is continuous. The fact that $c \in I$ follows from the forward limit property, since the graph of the restriction of ζ to $[a, c[$ is contained in $\mathcal{T}(\xi, \varepsilon)$. If $c < b$ and $||\zeta(c) - \xi(c)|| < \varepsilon$, then the forward existence property implies that ζ is not maximal.)

We will estimate $||\zeta(t) - \xi(t)||$ for $t \in [a, c]$, and show that —if ε is small enough, and δ is suitably chosen— $||\zeta(t) - \xi(t)|| < \varepsilon$ for all such t, including $t = c$. This will imply that $c = b$,

From now on, we work on the interval $[a, c]$. We have

$$\zeta(t) - \xi(t) = \bar{x} - \xi(a) + \int_a^t \Big(f_p(\zeta(s), s) - f_0(\xi(s), s) \Big)\, ds \,, \tag{A.40}$$

so

$$\zeta(t) - \xi(t) = \bar{x} - \xi(a) + W_p(t) + \int_a^t \Big(f_p(\zeta(s), s) - f_p(\xi(s), s) \Big)\, ds \,. \tag{A.41}$$

Let $r_p(t) = f_p(\zeta(t), t) - f_p(\xi(t), t) - A(t).(\zeta(t) - \xi(t))$. Then

$$\zeta(t) - \xi(t) = \bar{x} - \xi(a) + W_p(t) + \int_a^t A(s).(\zeta(s) - \xi(s))\, ds + \int_a^t r_p(s)\, ds \,, \tag{A.42}$$

so

$$||\zeta(t) - \xi(t)|| \leq ||\bar{x} - \xi(a)|| + ||W_p||_{sup} + \int_a^t ||A(s)||.||\zeta(s) - \xi(s)||\, ds + \int_a^t r_p(s)\, ds \,, \tag{A.43}$$

and then

$$||\zeta(t) - \xi(t)|| \leq ||\bar{x} - \xi(a)|| + ||W_p||_{sup} + 2\varepsilon||\psi_\varepsilon||_{L^1} + \int_a^t ||A(s)||.||\zeta(s) - \xi(s)||\, ds \,, \tag{A.44}$$

because $||r_p(s)|| \leq 2\varepsilon\psi_\varepsilon(s)$ for $s \in [a, c]$, since $||\zeta(s) - \xi(s)|| \leq \varepsilon$ and $||p|| \leq \varepsilon$.

Since $||\bar{x} - \xi(a)|| \leq \delta$, and $||W_p||_{sup} \leq ||W||_{TV} \leq C||p|| \leq C\delta$, Gronwall's inequality implies

$$||\zeta(t) - \xi(t)|| \leq e^{||A||_{L^1}} \Big((C + 1)\delta + 2\varepsilon||\psi_\varepsilon||_{L^1} \Big) \,. \tag{A.45}$$

In view of (A.38),we have

$$2\varepsilon e^{||A||_{L^1}}||\psi_\varepsilon||_{L^1} < \frac{\varepsilon}{4} \,. \tag{A.46}$$

Now suppose that

$$\delta \leq \kappa\varepsilon\,, \qquad \text{where} \qquad \kappa = (4C+4)^{-1}e^{-||A||_{L^1}}\,. \tag{A.47}$$

Then

$$||\zeta(t)-\xi(t)|| \leq \frac{\varepsilon}{2} \qquad \text{for} \qquad \text{all} \qquad t \in [a,c]\,. \tag{A.48}$$

This is our desired estimate, from which it follows, as explained above, that $c = b$, and then $||\zeta(t)-\xi(t)|| \leq \varepsilon$ for all $t \in [a,b]$. This completes the proof of (A.3.a).

Besides proving (A.3.a), we have also shown that, if κ is the constant given by (A.47), then the conclusion of (A.3.a) holds for every δ such that $0 < \delta \leq \kappa\varepsilon$. (Notice that if $\delta \leq \kappa\varepsilon$ then (A.39) holds.)

Now suppose we are given a map $P \ni p \mapsto \bar{x}_p \in \mathbb{R}^n$ which is defined for $||p|| \leq \bar{\delta}$ for some $\bar{\delta} > 0$. Assume that this map is Fréchet differentiable at 0, with differential L, and is such that $\bar{x}_0 = \xi(a)$. Write $z_p = Y_{A,dV_p,L.p}$, where $V_p = D^w\mathcal{W}(0).p$.

It then follows that there exist a constant $C' \geq 1$ and a $\bar{\delta}' \in]0,\bar{\delta}]$ such that (i) $C'\bar{\delta}' \leq \kappa\bar{\varepsilon}$, and (ii) $||\bar{x}_p - \xi(a)|| \leq C'||p||$ for $||p|| \leq \bar{\delta}'$.

Pick in an arbitrary fashion $\xi_p \in \mathrm{MTraj}(f_p, \bar{x}_p, a)$ for $||p|| \leq \bar{\delta}'$. Then, if $||p|| \leq \bar{\delta}'$, it follows that (i) ξ_p is defined, (ii) $\xi_p \in \mathrm{MTraj}(f_p, \bar{x}, a)$ for an $\bar{x}$ —namely, $\bar{x}_p$— such that $||\bar{x} - \xi(a)|| \leq C'\bar{\delta}' \leq \kappa\bar{\varepsilon}$. Since $C' \geq 1$, we also have $||p|| \leq \kappa\bar{\varepsilon}$. Therefore we can conclude —using $\varepsilon = \bar{\varepsilon}$ and $\delta = \kappa\bar{\varepsilon}$, that $[a,b] \subseteq \mathrm{Dom}(\xi_p)$ and $(\xi_p(t),t) \in \mathcal{T}(\xi,\bar{\varepsilon})$ for $t \in [a,b]$.

Let x_p be the restriction of ξ_p to $[a,b]$. For $||p|| \leq \bar{\delta}'$, we have

$$x_p(t)-\xi(t)-z_p(t) = \bar{x}_p-\xi(a)-L.p+\int_a^t A(s).(x_p(s)-\xi(s)-z_p(s))\,ds+W_p(t)-V_p(t)+R_p(t)\,, \tag{A.49}$$

where $R_p(t) = \int_a^t r_p(s)\,ds$. If $p \neq 0$, let

$$Y_p = ||p||^{-1}(x_p - x_0 - z_p)\,, \qquad U_p = ||p||^{-1}(W_p - V_p + R_p)\,, \qquad \bar{Y}_p = ||p||^{-1}(\bar{x}_p - x_0(a) - L.p)\,. \tag{A.50}$$

Then $Y_p = Y_{A,dU_p,\bar{Y}_p}$.

We have to prove that Y_p weak* converges to 0 as $p \to 0$. Since $\bar{Y}_p \to 0$, our conclusion will follow if we show that U_p weak* converges to 0. Since $V_p = D^w\mathcal{W}(0).p$, it is clear that $||p||^{-1}(W_p - V_p)$ weak*-converges to 0. So all we need is to prove that $||p||^{-1}R_p$ weak*-converges to 0. If $||p||$ is sufficiently small, then we can apply the previous estimates using $\delta = ||p||$, $\varepsilon = \frac{\delta}{\kappa}$. We conclude that $||r_p(s)|| \leq 2\varepsilon\psi_\varepsilon(s)$ for all $s \in [a,b]$. This implies that

$$||r_p||_{L^1} \leq \frac{2||p||}{K}||\psi_{\kappa^{-1}||p||}||_{L^1}\,. \tag{A.51}$$

This shows that $||R_p||_{TV} = ||r_p||_{L^1} = o(||p||)$ as $p \to 0$. So $||p||^{-1}R_p$ converges to 0 strongly —and *a fortiori* weakly*— in $BV([a,b],\mathbb{R}^n)$. This completes our proof. ∎

If $f \in TVVF(\mathbb{R}^n)$ and $\xi : [a,b] \mapsto \mathbb{R}^n$ is a trajectory of f, we say that f *is differentiable along* ξ if the map $x \mapsto f(x,t)$ is differentiable at $\xi(t)$ for almost every t. In that case, we write $D_\xi f(t) \stackrel{\mathrm{def}}{=} \frac{\partial f}{\partial x}(\xi(t),t)$, and refer to the map $[a,b] \ni t \mapsto D_\xi f(t) \in \mathbb{R}^{n\times n}$ as the *differential of* f *along* ξ.

Lemma A.5. *If* $\mathbf{f}$ *is weakly differentiable at* p_0 *along a trajectory* $\xi : [a,b] \to \mathbb{R}^n$ *of* f_{p_0}, *then* f_{p_0} *is differentiable along* ξ *and the map* A *of (WD3) is such that* $A(t) = D_\xi f_{p_0}(t)$ *for almost all* $t \in [a,b]$.

PROOF. We begin by showing that the functions ψ_ε for which (WD3) holds can be chosen so that $\psi_\varepsilon(t)$ is a monotonically nonincreasing function of ε for each t. To see this, start with some family $\{\psi_\varepsilon\}_{\varepsilon\in]0,\bar\varepsilon]}$ of functions for which (WD3) holds, and define $\hat\psi_\varepsilon = \psi_{2^{-k}\bar\varepsilon}$ whenever $2^{-k-1}\bar\varepsilon < \varepsilon \le 2^{-k}\bar\varepsilon$ and k is a nonnegative integer. Then (WD3) also holds if the $\hat\psi_\varepsilon$ are substituted for the ψ_ε. Next define $\bar\psi_\varepsilon(t) = \min\{\hat\psi_{\varepsilon'}(t) : \varepsilon \le \varepsilon' \le \bar\varepsilon\}$. Then the $\bar\psi_\varepsilon$ are well defined integrable functions of t, because for each ε the set of functions $\{\hat\psi_{\varepsilon'} : \varepsilon \le \varepsilon' \le \bar\varepsilon\}$ is finite. It is clear that (WD3) also holds with the $\bar\psi_\varepsilon$ instead of the $\hat\psi_\varepsilon$, and in addition $\hat\psi_\varepsilon(t)$ is nonincreasing as a function of ε.

Now, if we pick the ψ_ε to be monotonically nonincreasing with respect to ε, it follows that the limit $\rho(t) = \lim_{\varepsilon\downarrow 0}\psi_\varepsilon(t)$ exists for every t. By the dominated convergence theorem, $\int_a^b \rho(t)dt = 0$. Since $\rho(t) \ge 0$ for all t, it follows that $\rho(t) = 0$ almost everywhere.

Let t be such that $\rho(t) = 0$. Let $\theta(x) = \psi_{||x-\xi(t)||}(t)$. Then $\lim_{x\to\xi(t)}\theta(x) = 0$. On the other hand, (WD3) implies that

$$||f_{p_0}(x,t) - f_{p_0}(\xi(t),t) - A(t).(x-\xi(t))|| \le \theta(x)||x-\xi(t)||\,. \tag{A.52}$$

Since $\theta(x) \to 0$ as $x \to \xi(t)$, (A.52) says precisely that $A(t)$ is the differential at $\xi(t)$ of the map $x \mapsto f_{p_0}(x,t)$, completing our proof. ∎

If $f \in TVVF(\mathbb{R}^n)$ is differentiable along a trajectory $\xi : [a,b] \to \mathbb{R}^n$ of f, and the matrix-valued function $D_\xi f$ is integrable, then the ordinary differential equation

$$\dot\lambda(t) = -\lambda(t).D_\xi f(t)\,, \tag{A.53}$$

for an $\mathbb{R}_n$-valued function λ on $[a,b]$, is the *adjoint variational equation* —or, simply, the *adjoint equation*— for f along ξ. A solution of the adjoint equation (i.e. an absolutely continuous function $\lambda : [a,b] \mapsto \mathbb{R}_n$ such that (A.53) holds for almost all t) is known as an *adjoint vector for f along ξ*.

Theorem A.4. *Let P, $\mathbf{f} = \{f_p\}_{p\in P}$, $p_0 \in P$, $\xi : [a,b] \to \mathbb{R}^n$ be as in Theorem A.3. Let $P \ni p \mapsto \bar x_p \subset \mathbb{R}^n$ be a map which is Fréchet differentiable at p_0 and such that $\bar x_{p_0} = \xi(a)$, and let L_a be its differential at p_0. Let ξ_p be, for p close enough to p_0, maximal trajectories of f_p such that $\xi_p(a) = \bar x_p$, and let x_p be the restriction of ξ_p to $[a,b]$. Then the map $p \mapsto x_p(b)$ is differentiable at p_0. Let L_b be its differential at p_0. Let $\lambda : [a,b] \to \mathbb{R}^n$ be an adjoint vector for f_{p_0} along ξ. Then*

$$\lambda(b)\cdot L_b(v) - \lambda(a)\cdot L_a(v) = \lim_{\varepsilon\downarrow 0}\varepsilon^{-1}\int_a^b \Big\langle \lambda(t), f_{p_0+\varepsilon v}(\xi(t),t) - f_{p_0}(\xi(t),t)\Big\rangle dt\,. \tag{A.54}$$

PROOF. Assume, without loss of generality, that $p_0 = 0$. According to Theorem A.3 and Lemma A.5, the map $p \mapsto x_p$ is weak* Fréchet differentiable at 0 —as a map from P to $BV([a,b],\mathbb{R}^n)$— and its differential at 0 is the map that sends $v \in P$ to $Y_{A,d(U.v),L_a(v)}$, where $U = D^w\mathcal{W}(0)$, and $\mathcal{W}$ is the map $p \mapsto W_p$, where $W_p(t) = \int_a^t \Big(f_p(\xi(s),s) - f_0(\xi(s),s)\Big)dt$, and $A = D_\xi f$. Then $p \mapsto x_p(b)$ is Fréchet differentiable at 0, and the differential L_b is given by $L_b(v) = Y_{A,d(U.v),L_a(v)}(b)$. Clearly, $L_a(v) = Y_{A,d(U.v),L_a(v)}(a)$. Since λ satisfies $\dot\lambda = -\lambda.A$, Lemma A.3 tells us that

$$\lambda(b)\cdot L_b(v) - \lambda(a)\cdot L_a(v) = \int_a^b \lambda.d(U.v)\,.$$

Since $U.v$ is the weak*-limit of $\varepsilon^{-1}W_{\varepsilon v}$ as $\varepsilon \downarrow 0$, our conclusion follows. ∎

We now turn to the differentiation of a Lagrangian cost functional.

If $f \in TVVF(\mathbb{R}^n)$, a *Lagrangian* for f in $\mathbb{R}^n$ is a real-valued function L such that

(L.1) $\mathrm{Dom}(f) \subseteq \mathrm{Dom}(L)$;

(L.2) the function $t \mapsto L(\xi(t), t)$ is locally integrable for every $\xi \in \mathrm{Traj}(f)$.

If $f \in TVVF(\mathbb{R}^n)$ has components $f^1, \ldots, f^n$, and L is a Lagrangian for f, then we can consider the time-varying vector field $[L; f]$ in $\mathbb{R}^{n+1}$ whose components are $(f^0, f^1, \ldots, f^n)$, where $f^0 \equiv L$. A trajectory of $[L; f]$ can then be regarded a pair (ξ^0, ξ), where $\xi \in \mathrm{Traj}(f)$ and ξ^0 is a real-valued locally absolutely continuous function on $\mathrm{Dom}(\xi)$ such that $\dot{\xi}^0(t) = L(\xi(t), t)$ for a.e. t. Any such function ξ^0 will be called a *running L-cost along ξ*. Our definition of a Lagrangian implies that a running L-cost along ξ exists for all $\xi \in \mathrm{Traj}(f)$, and it is clear that any two running L-costs along ξ must differ by a constant. If ξ is such that $\mathrm{Dom}(\xi)$ is of the form $[a, b]$ or $[a, b[$ or $[a, \infty[$, then there is a canonical choice of running L-cost along ξ, obtained by letting $\xi^0(t) = \int_a^t L(\xi(s), s)\, ds$. In that case, we will write ξ^L to denote the curve $t \mapsto (\xi^0(t), \xi(t))$, where ξ^0 is the canonical running L-cost.

Now suppose we are given data P, $\mathbf{f}$, p_0, ξ as in (D.1,2,3,4), as well as

(D.5) a family $\mathbf{L} = \{L_p\}_{p \in P}$ such that each L_p is a Lagrangian for f_p.

Given a trajectory $\xi : [a, b] \mapsto \mathbb{R}^n$ of f_{p_0}, we say that the pair $(\mathbf{f}, \mathbf{L})$ is *weakly differentiable at p_0 along ξ* if the family $\{[L_p; f_p]\}_{p \in P}$ is weakly differentiable along $\xi^{L_{p_0}}$.

It is easy to see that $(\mathbf{f}, \mathbf{L})$ is weakly differentiable at p_0 along ξ iff there exists $\bar{\varepsilon} > 0$ such that (WD1,2,3,4) hold, and in addition

(WD5) *There exists $h \in L^1([a, b], \mathbb{R}_n)$ such that*

$$||L_p(x, t) - L_p(\xi(t), t) - h(t).(x - \xi(t))|| \leq \psi_\varepsilon(t)(||x - \xi(t)|| + ||p||) \qquad \text{(A.55)}$$

whenever $a \leq t \leq b$, $||x - \xi(t)|| \leq \varepsilon$ and $||p|| \leq \varepsilon$.

(WD6) *If we let $w_p^0(t) = L_p(\xi(t), t) - L_{p_0}(\xi(t), t)$, $W_p^0(t) = \int_a^t w_p^0(s)\, ds$, then the map $\mathcal{W}^0$ that sends p to $\mathcal{W}^0(p) = W_p^0 \in BV([a, b], \mathbb{R})$ is weak* Fréchet differentiable at p_0.*

If $f \in TVVF(\mathbb{R}^n)$, $\xi : [a, b] \mapsto \mathbb{R}^n$ is a trajectory of f, and L is a Lagrangian for f, we say that *L is differentiable along ξ* if the map $x \mapsto L(x, t)$ is differentiable at $\xi(t)$ for almost every t. In that case, we write $\nabla_\xi L(t) \stackrel{\text{def}}{=} \nabla_x L(\xi(t), t)$, and refer to the map $[a, b] \ni t \to \nabla_\xi L(t) \in \mathbb{R}^{n \times n}$ as the *differential of L along ξ*.

Lemma A.5, applied to the curve $\xi^{L_{p_0}}$, yields

Lemma A.6. *If $(\mathbf{f}, \mathbf{L})$ is weakly differentiable at p_0 along a trajectory $\xi : [a, b] \to \mathbb{R}^n$ of f_{p_0}, then f_{p_0} and L_{p_0} are differentiable along ξ and the maps A, h of (WD3,5) are such that $A(t) = D_\xi f(t)$ and $h(t) = \nabla_\xi L_{p_0}(\xi(t), t)$ for almost all $t \in [a, b]$.* ∎

Under the hypothesis of Lemma A.6, we can apply Theorem A.4 to the family $\{[L_p; f_p]\}_{p \in P}$ and the curve $\xi^{L_{p_0}}$. Define

$$J_p(\zeta) = \int_a^b L_p(\zeta(t), t)\, dt\,, \qquad \text{(A.56)}$$

if $\zeta : [a, b] \to \mathbb{R}^n$ is a curve such that the integral exists. (We will only be using this when $\zeta \in \mathrm{Traj}(f_p)$, in which case $J_p(\zeta)$ is well defined.) An adjoint vector for $[L_{p_0}; f_{p_0}]$ along $\xi^{L_{p_0}}$ is then a pair (λ^0, λ) such that $\lambda^0 \in \mathbb{R}$, $\lambda : [a, b] \mapsto \mathbb{R}_n$ is absolutely continuous, and

$$\dot{\lambda}(t) = -\lambda(t).A(t) - \lambda^0 h(t) \qquad \text{for} \qquad \text{a.e.} \qquad t \in [a, b]\,. \qquad \text{(A.57)}$$

Theorem A.4 then gives:

Theorem A.5. *Assume that we are given data P, $\mathbf{f}$, p_0, ξ, $\mathbf{L}$ as in (D.1,2,3,4,5). Assume that $(\mathbf{f},\mathbf{L})$ is weakly differentiable at p_0 along ξ. Let $P \ni p \mapsto \bar{x}_p \subset \mathbb{R}^n$ be a map which is Fréchet differentiable at p_0 and such that $\bar{x}_{p_0} = \xi(a)$, and let L_a be its differential at p_0. Let ξ_p be, for p close enough to p_0, maximal trajectories of f_p such that $\xi_p(a) = \bar{x}_p$, and let x_p be the restriction of ξ_p to $[a,b]$. Then the maps $p \mapsto x_p(b)$ and $p \mapsto J_p(x_p)$ are differentiable at p_0. Let L_b, L_b^0 be their differentials at p_0. Then, if (λ^0,λ) is an adjoint vector for $[L_{p_0}, f_{p_0}]$ along ξ, it follows that*

$$\lambda(b)\cdot L_b(v) + \lambda^0 L_b^0(v) - \lambda(a)\cdot L_a(v) = \lim_{\varepsilon\downarrow 0} \varepsilon^{-1} \int_a^b \Big(H_{p_0+\varepsilon v}(\xi(t),\lambda(t),\lambda_0,t) - H_{p_0}(\xi(t),\lambda(t),\lambda_0,t) \Big)\, dt\,, \tag{A.58}$$

where

$$H_p(x,z,z^0,t) = \langle z, f_p(x,t)\rangle + z^0 L_p(x,t) \tag{A.59}$$

for $x \in \mathbb{R}^n$, $z \in \mathbb{R}_n$, $z^0 \in \mathbb{R}$, $t \in [a,b]$. ∎

Appendix B . A lemma .

Lemma B.1. *Let f be a real-valued function on a compact interval $[a,b]$. Assume that there exists a finite or countable subset E of $[a,b]$ with the following properties:*

(B.1.a) $\liminf_{h\downarrow 0} \frac{f(x+h)-f(x)}{h} \geq 0$ *for all* $x \in [a,b[\backslash E$,
(B.1.b) $\liminf_{h\downarrow 0} f(x+h) - f(x) \geq 0$ *for all* $x \in [a,b[$,
(B.1.c) $\limsup_{h\downarrow 0} f(x) - f(x-h) \geq 0$ *for all* $x \in]a,b]$.
Then $f(b) \geq f(a)$.

PROOF. Let $\varepsilon > 0$ be arbitrary. We will prove that $f(b) \geq f(a) - \varepsilon$. Let $\varepsilon' = \frac{\varepsilon}{1+b-a}$ and $E = \{x_k\}_{k\in N}$, $N \subset \mathbb{N}$.

For each $x \in [a,b]$, let $I(x) = \{k : x_k < x\}$. Define $g : [a,b] \mapsto \mathbb{R}$ by

$$g(x) = \sum_{k\in I(x)} 2^{-k}\varepsilon'\,.$$

Then g is left-continuous and monotonically nondecreasing.

Let S be the set of all $x \in [a,b]$ such that $f(y) \geq f(a) - (y-a)\varepsilon' - g(y)$ for all $y \in [a,x]$. It is clear that $a \in S$, and also that S is an interval. So S is of the form $[a,c[$ or $[a,c]$, with $c \in [a,b]$.

Suppose $S = [a,c[$. Then $c > a$. By (B.1.c), given $\delta > 0$ there is a sequence $\{z_j\}$ in $[a,b]$ such that $z_j < c$, $z_j \to c$, and $f(c) - f(z_j) \geq -\delta$. Then $z_j \in S$, so $f(z_j) \geq f(a) - (z_j - a)\varepsilon' - g(z_j) \geq f(a) - (c-a)\varepsilon' - g(c)$. It then follows that $f(c) \geq f(a) - (c-a)\varepsilon' - g(c) - \delta$. Since δ is arbitrary, we have $f(c) \geq f(a) - (c-a)\varepsilon' - g(c)$, so $c \in S$, which is a contradiction.

Therefore $S = [a,c]$ for some $c \in [a,b]$. We now show that $c = b$. Assume that $c < b$, and consider separately the cases $c \notin E$, $c \in E$.

If $c \notin E$, then (B.1.a) implies that there exists a $c' > c$ such that $f(x) - f(c) \geq -(x-c)\varepsilon'$ for $x \in [c,c']$. But then, since $f(c) \geq f(a) - (c-a)\varepsilon' - g(c)$, we have $f(x) \geq f(a) - (x-c)\varepsilon' - g(c) \geq f(a) - (x-a)\varepsilon' - g(x)$ for $x \in [c,c']$. Therefore $c' \in S$, contradicting the fact that $S = [a,c]$.

Now suppose $c \in E$. Then $c = x_k$ for some k, and $x_k \neq b$. By (B.1.b), there exists $h > 0$ such that $c + h \leq b$ and $f(y) - f(c) \geq -2^{-k}\varepsilon'$ for all $y \in [c, c+h]$. If $y \in]c, c+h]$, then $k \notin I(c)$, and $I(c) \cup \{k\} \subseteq I(y)$. Therefore $g(y) \geq g(c) + 2^{-k}\varepsilon'$. Since $f(c) \geq f(a) - (c-a)\varepsilon' - g(c)$, we have $f(y) \geq f(c) - 2^{-k}\varepsilon' \geq f(a) - (c-a)\varepsilon' - 2^{-k}\varepsilon' - g(c) \geq f(a) - (y-a)\varepsilon' - g(y)$. So $c + h \in S$, which is a contradiction.

We have therefore proved that $c = b$. Then

$$f(b) \geq f(a) - (b-a)\varepsilon' - g(b) \geq f(a) - (b-a)\varepsilon' - \varepsilon' = f(a) - \varepsilon\,.$$

Since ε is arbitrary, we have shown that $f(b) \geq f(a)$, as desired. ∎

REFERENCES

1. Bardi, M.,and I. Capuzzo Dolcetta, "Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equation", Birkäuser, Boston, to appear.
2. Boltianskii, V. G., "Sufficient conditions for optimality and the justification of the dynamic programming principle", *S.I.A.M. J. Control and Optimization*, **4**, (1966), pp. 326-361.
3. Bressan, A., and B. Piccoli, "Structural stability for time-optimal planar syntheses", to appear on *Dynamics of Continuous, Discrete and Impulsive Systems.*
4. Bressan, A., and B. Piccoli, "A generic classification of time optimal planar stabilizing feedbacks", to appear on *S.I.A.M. J. Control and Optimization.*
5. Brunovsky, P., "Every normal linear system has a regular time–optimal synthesis", *Math. Slovaca*, **28**, (1978), pp. 81-100.
6. Brunovsky, P., "Existence of regular syntheses for general problems", *J. Diff. Equations*, **38**, (1980), pp. 317-343.
7. Cesari, L., "Optimization – Theory and applications", Springer–Verlag, New York, 1983.
8. Federer, H., "Geometric measure theory", Springer, Berlin, 1969.
9. Filippov, A. F., "Differential equations with discontinuous right-hand side", Reidel, 1989.
10. Fleming, W.H., and R. W. Rishel, "Deterministic and stochastic optimal control", Springer–Verlag, New York, 1975.
11. Fleming, W.H., and M. Soner, "Controlled Markov processes and viscosity solutions", Springer–Verlag, New York, 1993.
12. 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.
13. Lojasiewicz Jr., S., and H. J. Sussmann, "Some examples of reachable sets and optimal cost functions that fail to be subanalytic," *S.I.A.M. J. Control and Optimization* **23**, No. 4 (1985), pp. 584-598.
14. Piccoli, B., "Regular time-optimal syntheses for smooth planar systems", *Rend. Sem. Mat. Univ. Padova* **95**, (1996), pp. 59-79.
15. Piccoli, B., "Classification of generic singularities for the planar time optimal syntheses", to appear on *S.I.A.M. J. Control and Optimization.*
16. Sussmann, H. J., "Subanalytic sets and feedback control," *J. Diff. Equations* **31**, No.1 (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., "Analytic stratifications and control theory," in *Proc. 1978 Int. Congress of Mathematicians*, Helsinki (1980), pp. 865-871.
19. Sussmann, H.. J., "Time-optimal control in the plane," in *Feedback Control of Linear and Nonlinear Systems*, D. Hinrichsen and A. Isidori Eds., Springer-Verlag Lect. Notes in Control and Information Sciences, No. 39, New York (1982), pp. 244-260.
20. Sussmann, H. J., "Subanalytic sets and optimal control in the plane," *Proc. 21st I.E.E.E. Conference on Decision and Control*, Orlando, Fla. (Dec 1982), pp. 295-299.

21. Sussmann, H. J., "Lie brackets, real analyticity and geometric control theory," in *Differential Geometric Control Theory*, R.W. Brockett, R.S. Millman and H.J. Sussmann Eds., Birkhäuser Boston Inc. (1983), pp. 1-115.

22. Sussmann, H. J.., "Lie Brackets and real analyticity in control theory," in *Mathematical Control Theory*, C. Olech Ed., Banach Center Publications, Volume 14, PWN-Polish Scientific Publishers,Warsaw, Poland, 1985, pp. 515-542.

23. Sussmann, H. J., "The structure of time-optimal trajectories for single-input systems in the plane : the C^∞ nonsingular case, *S.I.A.M. J. Control and Optimization,* **25**, No. 2 (1987), pp. 433-465.

24. Sussmann, H. J., "The structure of time-optimal trajectories for single-input systems in the plane: the general real-analytic case," *S.I.A.M. J. Control and Optimization* **25**, No. 4 (1987), pp. 868-904.

25. Sussmann, H. J., "Regular synthesis for time-optimal control of single-input real-analytic systems in the plane," *S.I.A.M. J. Control and Optimization* **25**, No. 5 (1987), pp. 1145-1162.

26. 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.

27. Sussmann, H. J., "Synthesis, presynthesis, sufficient conditions for optimality and subanaltic sets", in Nonlinear controllability and optimal control, H. J. Sussmann ed., Marcel Dekker, New York, 1990, pp. 1-19.

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