Nonlinear Analysis, Theory, Methods & Applications, Vol. 11, No. 6, pp. 665–683, 1987.
Printed in Great Britain.

0362-546X/87 $3.00 + .00
Pergamon Journals Ltd.

DIFFERENTIAL AUTOMATA AND THEIR DISCRETE SIMULATORS

Lucio Tavernini

Division of Mathematics, Computer Science and Systems Design, The University of Texas at San Antonio, San Antonio, Texas 78285, U.S.A.

(*Received* 4 *February* 1985; *received for publication* 1986)

Key words and phrases: Initial value problems with hysteresis, numerical methods, order of convergence, continuous-time system simulations.

1. INTRODUCTION

The digital computer simulation of deterministic lumped-parameter continuous-time systems leads, in simple cases, to a numerical initial value problem for ordinary differential equations. More generally, in the presence of transport delays or of "memory integrals," one faces a numerical initial data problem for functional differential equations of retarded type ([1]). While matters may be complicated further by state-dependent delays, still, arguments for the convergence of numerical approximations usually rely on some sort of continuity of the vector field with respect to the state of the system: One looks for a topology which is compatible with the time delays (see [5]).

The situation may be quite different if elements with multi-valued hysteretic input–output behavior are introduced into the mathematical model of a physical system: it may not even be entirely clear how to represent hysteresis in a computationally convenient manner. For example, consider the "system" shown in Fig. 1. How does one translate from the visual

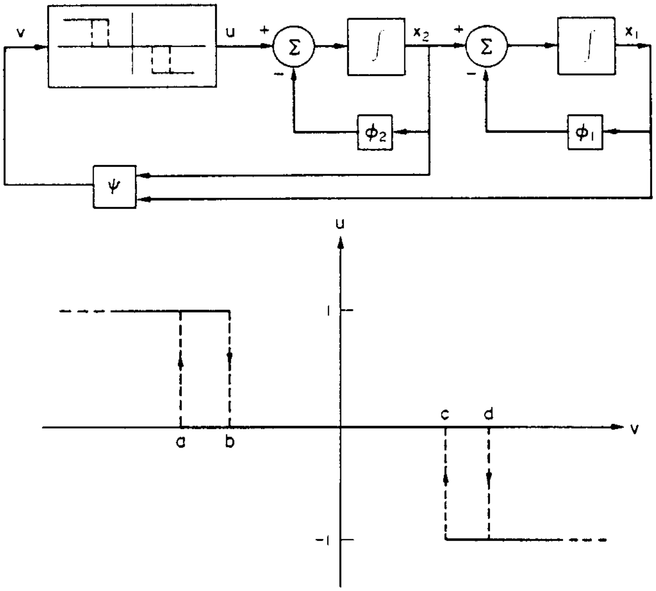

Fig. 1.

colloquialism of an engineering block diagram to a mathematical object? We are interested in a formulation which lends itself quite easily to numerical analysis.

The "system" in Fig. 1 "defines" the "differential equation" in $\mathbb{R}^2$ given by

$$\begin{aligned}\dot{x}_1 &= x_2 - \phi_1(x_1),\\ \dot{x}_2 &= u - \phi_2(x_2).\end{aligned} \tag{1}$$

Actually, the multi-valued nature of the relationship between u and the "state" (x_1, x_2) seems to suggest that (1) be split into three separate cases, depending upon whether u has one of the values -1, 0, or 1. As a visual tool, consider the directed graph shown in Fig. 2. By thinking of the set $S = \mathbb{R}^2 \times \{1, 2, 3\}$ (the disjoint union of three copies of $\mathbb{R}^2$) as state space, we can associate with each vertex of this graph, by way of (1), a unique.ODE, one for each of the three copies of $\mathbb{R}^2$. Under suitable hypotheses, one might expect that the label of each of the outgoing edges would define a time beyond which the "solution trajectory" of (1) would no longer coincide with a solution of the ODE associated with the initial vertex of that edge, but would rather coincide with a solution of the ODE associated with the terminal vertex. Still speaking rather loosely, under what conditions can one construct a "flow line" on S such that when projected on $\mathbb{R}^2$ it represents a "solution" of (1)?

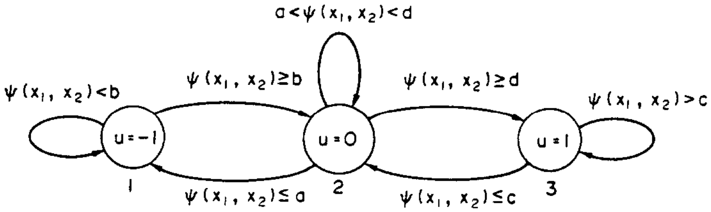

Fig. 2.

Below we present the notion of a differential automaton as a vehicle for the formulation of a class of initial value problems over hysteretic vector fields. The point here is to look at a class of problems which, while sufficiently large to be of practical interest, is sufficiently small to permit numerical analysis relatively easily. To this end, we intentionally avoid a constrained differential equation formulation (as in [4]). It must be made plain that we do not propose a mathematical formalism for understanding how hysteretic behavior comes about in a physical device. (For this see [2].) We develop an existence–uniqueness and well-posedness theory in Section 2 and a theory of numerical approximations in Section 3. In Section 4 we give an example.

2. DIFFERENTIAL AUTOMATA

A *differential automaton* A is a triple (S, f, ν), where:

S is the *state space* of A, $S = \mathbb{R}^m \times Q$, $Q = \{1, \ldots, n\}$ is a nonempty finite set, the *discrete state space* of A, and $\mathbb{R}^m$ is the *continuous state space* of A;

f is a finite family $f(\cdot, q): \mathbb{R}^m \to \mathbb{R}^m$, $q \in Q$, of vector fields, the *continuous dynamics* of A;

ν is the *discrete state transition function* of A, $\nu: S \to Q$.

The continuous dynamics. We require that each $f(\cdot, q)$, $q \in Q$, be a globally Lipschitz C^∞ map. Thus, each vector field $f(\cdot, q)$ defines a global flow $\mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$: $(t, x) \mapsto E_q(t)x$, where $\{E_q(t) | t \in \mathbb{R}\}$ is the group of C^∞ diffeomorphisms of $\mathbb{R}^m$ with infinitesimal generator $f(\cdot, q)$.

The discrete state transition function. Denote by ν_q, $q \in Q$, the function $\nu_q : \mathbb{R}^m \to Q : x \mapsto \nu(x, q)$. Define $I(q) = \nu_q(\mathbb{R}^m) \setminus \{q\}$. We require that for each $q \in Q$ and each $p \in I(q)$ there exist a function $g_{q,p} \in C^\infty(\mathbb{R}^m, \mathbb{R})$ with 0 in the image of $g_{q,p}$ a regular value such that

$$M_{q,p} \equiv \nu_q^{-1}(p) = \{x \in \mathbb{R}^m | g_{q,p}(x) \geq 0\}.$$

Thus, $\nu_q^{-1}(p)$ is a C^∞ m-submanifold of $\mathbb{R}^m$ with boundary

$$\partial M_{q,p} = \partial \nu_q^{-1}(p) = \{x \in \mathbb{R}^m | g_{q,p}(x) = 0\},$$

which is a C^∞ $(m-1)$-submanifold of $\mathbb{R}^m$. We require that each $M_{q,p}$ be connected. The manifolds $\partial M_{q,p}$ are called the *switching manifolds* of the automaton A.

Define $\alpha_q = \min\{\text{dist}(M_{q,p}, M_{q,p'}) | p, p' \in I(q), p \neq p'\}$. We require that the inequality

$$\alpha(A) \equiv \min_{q \in Q} \alpha_q > 0 \tag{DA1}$$

be satisfied. Define $\beta_{q,p} = \min\{\text{dist}(\partial M_{q,p}, \partial M_{p,p'}) | p' \in I(p)\}$. We require that inequality

$$\beta(A) \equiv \min_{q \in Q} \min_{p \in I(q)} \beta_{q,p} > 0 \tag{DA2}$$

be satisfied. Define $M_q = \cup_{p \in I(q)} M_{q,p}$ and define the domain of capture of state q by

$$\delta(q) \equiv \mathbb{R}^m \setminus M_q = \{x \in \mathbb{R}^m | \nu(x, q) = q\}.$$

Thanks to (DA1), $\delta(q)$ is an open subset of $\mathbb{R}^m$ with frontier given by $\text{fr}\ \delta(q) = \text{fr}\ M_q = \cup_{p \in I(q)} \partial M_{q,p}$.

We require that the inclusions

$$\partial M_{q,p} \subset \delta(p), \quad p \in I(q), q \in Q, \tag{DA3}$$

be satisfied.

2.1. *The initial value problem for differential automata*

Define $S_0 = \cup_{q \in Q} \delta(q) \times \{q\}$, the *admissible state set* of A. For an admissible initial state $s_0 = (x_0, q_0) \in S_0$, the initial value problem (A, s_0) is stated as follows. Find a pair of maps u and σ, $u \in AC([0, 1], \mathbb{R}^m)$, (absolutely continuous), and $\sigma \in RC([0, 1[, Q)$, (right-continuous), such that the following hold:

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) = f(u(t), \sigma(t)), \quad \text{a.e.} \quad t \in]0, 1[,$$

$$\sigma(t) = \nu(u(t), \sigma(t)), \qquad t \in [0, 1[,$$

with initial condition $(u(0), \sigma(0)) = (x_0, q_0)$. The points where σ is not continuous shall be called the *switching points* of the solution (u, σ).

PROPOSITION 2.2. The initial value problem (A, s_0) has a unique solution. The solution has finitely many switching points.

Proof. Let $s_0 = (x_0, q_0)$. If $E_{q_0}(t)x_0 \in \delta(q_0)$ for all t in the interval $[0, 1[$, then the solution is the pair of maps $(E_{q_0}(\cdot)x_0, t \mapsto q_0)$. Otherwise, the set

$$\Lambda = \{t \in [0, 1[\ | E_{q_0}(t)x_0 \in \mathrm{fr}\,\delta(q_0)\}$$

is nonempty. Let $t_1 = \inf \Lambda$ and let $\{t_i'\}$ be a sequence in Λ converging to t_1. Since we have that

$$\Lambda = \cup_{p \in I(q_0)}\{t \in [0, 1[\,| g_{q_0,p}(E_{q_0}(t)x_0) = 0\},$$

thanks to (DA1), we have, for some $p_0 \in I(q_0)$, that $g_{q_0,p_0}(E_{q_0}(t_i')x_0) \to 0$ as $i \to \infty$. By continuity, the infimum is actually attained. Moreover, we have $t_1 > 0$, since $\delta(q_0)$ is open in $\mathbb{R}^m$. Thus, we have a maximal interval $[0, t_1[$ such that $E_{q_0}(t)x_0$ is in $\delta(q_0)$ for all t in the interval $[0, t_1[$. On this interval define $u(t) = E_{q_0}(t)x_0$, $\sigma(t) = q_0$, and define $q_1 = \nu(u(t_1), q_0)$. Clearly $q_0 \neq q_1$. Thanks to (DA3) we have that $u(t_1) \in \delta(q_1)$.

Proceeding by induction, suppose that u and σ have been extended to $[0, t_k[$ by generating sequences $0 = t_0 < t_1 < \cdots < t_k < 1$ and $q_0, \ldots, q_k$, $q_i \neq q_{i+1}$, such that

$$u(t) = E_{q_i}(t - t_i)u(t_i), \qquad t \in [t_i, t_{i+1}],$$

$$q_i = \sigma(t) = \nu(u(t), q_i), \quad t \in [t_i, t_{i+1}[,$$

for $i = 0, \ldots, k-1$, with $q_k = \nu(u(t_k), q_{k-1}) \in \delta(q_k)$. We argue as done above to find a maximal interval $[t_k, t_{k+1}[$ such that $E_{q_k}(t - t_k)u(t_k) \in \delta(q_k)$ for t in this interval, and we define $u(t) = E_{q_k}(t - t_k)u(t_k)$, $\sigma(t) = q_k$ there. If $t_{k+1} = 1$, we are finished. Otherwise, we define $q_{k+1} = \nu(u(t_{k+1}), q_k)$. Clearly, $q_{k+1} \neq q_k$ and, thanks to (DA3), we have that $u(t_{k+1}) \in \delta(q_{k+1})$.

Suppose that this process creates an infinite sequence $t_1 < t_2 < \ldots$ of switching points in $[0, 1]$. Let t_* be the accumulation point of this sequence. Because there are finitely many switching manifolds $\partial M_{q,p}$, there is some $q \in Q$, some $p \in I(q)$, and some subsequence $\{t_i'\}$ of $\{t_i\}$ such that $u(t_i') \in \partial M_{q,p}$ for all i. The function u can be extended locally beyond each t_i' because, thanks to (DA3), we have that $\partial M_{q,p} \subset \delta(p)$, which is open. It follows that there must be another subsequence $\{t_i'\}$, of $\{t_i\}$ which separates the points of $\{t_i'\}$, i.e. $t_1' < t_1'' < t_2' < \ldots < t_*$, such that $u(t_i'') \notin \partial m_{q,p}$. Again, by the finiteness of the number of switching manifolds and by passing from $\{t_i'\}$ and $\{t_i''\}$ to subsequences, if necessary, we may suppose that $u(t_i'') \in \partial M_{p,p'}$ for some $p' \in I(p)$. If the limit $a = \lim_{i \to \infty} u(t_i)$ exists, then, because $\{u(t_i')\}$ and $\{u(t_i'')\}$ both converge to the point a, it follows that the distance between $\partial M_{q,p}$ and $\partial M_{p,p'}$ must be zero, contradicting (DA2). Thus, $\{t_i\}$ cannot have an accumulation point. To show that the above limit exists, for $|\cdot|$ a norm on $\mathbb{R}^m$, $\gamma > 0$ a Lipschitz constant for $f(\cdot, q)$, a Gronwall-type argument gives, for $t \geq 0$, $q \in Q$, the bound

$$|E_q(t)x| \leq |x|\, e^{\gamma t} + c\gamma^{-1}(e^{\gamma t} - 1).$$

for any $x \in \mathbb{R}^m$, where $c = \max_{q \in Q}|f(0, q)|$. Thus, for $t \in [0, t_*]$ we have the bound

$$|u(t)| = |E_{q_k}(t - t_k) \cdots E_{q_0}(t_1)x_0| \leq |x_0|\, e^{\gamma} + c\gamma^{-1}(e^{\gamma} - 1),$$

since for each $t \in [0, t_*[$ there is a k such that $u(t) = E_{q_k}(t - t_k) \ldots E_{q_0}(t_1)x_0$. In other words, $u(t) \in B$, where B is a ball in $\mathbb{R}^m$. On B, let λ be a bound for the first derivatives of each flow, i.e.

$$|E_q(t)x - E_q(t')y| \leq \lambda(|t - t'| + |x - y|),$$

for $t, t' \in [0, 1]$ and $x, y \in B$. This gives the bound

$$|u(t_{i+1}) - u(t_i)| = |E_{q_i}(t_{i+1} - t_i)u(t_i) - E_{q_i}(0)u(t_i)| \leq \lambda(t_{i+1} - t_i);$$

$$|u(t_{i+k}) - u(t_i)| \leq |u(t_{i+k}) - u(t_{i+k-1})| + \ldots + |u(t_{i+1}) - u(t_i)| \leq \lambda(t_{i-k} - t_i),$$

showing that $\{u(t_i)\}$ is a Cauchy sequence in $\mathbb{R}^m$.

2.3. *Well-posed initial valued problems*

If (u, σ) denotes the solution of the initial value problem (A, x, q_0), then σ defines a sequence $q_0, \ldots, q_k$ of discrete states of A, the "discrete trajectory of A started in state (x, q_0)". If (v, μ) denotes the solution of (A, x', q_0), where x' is near x, then we should have (u, σ) "near" (v, μ) in the sense that μ should define the same "discrete trajectory" $q_0, \ldots, q_k$ with possibly different switching points. Corresponding switching points of the two solutions should be close, however, in the sense that if σ defines the transition from q_{i-1} to q_i at time t_i, and if μ defines the same transition at time t_i', $|t_i - t_i'|$ should be "small" whenever $|x - x'|$ is "small" (in some norm). Below we define a topology to make this precise. Although the set of points x with the above property may not be the whole of $\delta(q_0)$, we show that it is generic. In fact, we show that it is an open dense subset of $\delta(q_0)$. To this end, below we define transversals of a differential automaton. The transversals are simply the solutions (u, σ) of the initial value problem which are transverse to the switching manifolds in the sense that each local flow line of $f(\cdot, q)$ given by $E_q(\cdot)u(t)$ in an open neighborhood of a switching point t is transverse to the switching manifold intersected by u at that switching point.

Topologize $Q \simeq \{1, \ldots, n\} \subset \mathbb{R}$ with the subspace topology. Let $PRC([0, b[, Q)$ denote the set of functions on $[0, b[$ to Q which are piecewise continuous and also right-continuous. Then, each $\sigma \in PRC([0, b[, Q)$ is completely characterized by the sequence $\{(t_i, \sigma(t_i))\}_{i=0}^{k}$, where $t_0 = 0$, and where $t_1 < \ldots < t_k$ are the points where σ is not continuous. The integer $k \geq 0$ depends, of course, on σ. Define $\pi_1, \pi_2 : PRC([0, b[, Q) \to l_\infty(\mathbb{R})$ by

$$\pi_1 \sigma = \{t_0, \ldots, t_k, 0, \ldots\},$$

$$\pi_2 \sigma = \{\sigma(t_0), \ldots, \sigma(t_k), 0, \ldots\},$$

and define the metric ρ_b on $PRC([0, b[, Q)$ by

$$\rho_b(\sigma, \mu) = \|\pi_1 \sigma - \pi_1 \mu\|_{l_\infty} + \|\pi_2 \sigma - \pi_2 \mu\|_{l_\infty}.$$

Note. The ε-ball of $\sigma \in PRC([0, b[, Q)$ consists of all functions $\mu \in PRC([0, b[, Q)$ which can be obtained from σ by perturbing by an amount less than ε the points $t_1, \ldots, t_k$ where σ jumps in value.

For $q \in Q$, define $\mathfrak{E}_1(\cdot, q) : \delta(q) \to C([0, 1], \mathbb{R}^m)$ and $\mathfrak{E}_2(\cdot, q) : \delta(q) \to PRC([0, 1[, Q)$ by $\mathfrak{E}_1(x, Q) = u$ and $\mathfrak{E}_2(\cdot, q) = \sigma$, where (u, σ) is the solution of the initial value problem

(A, x, q). Further, define $\mathfrak{E}(x, q) = (u, \sigma)$. Thus, $\mathfrak{E}(\cdot, q): \delta(q) \to F$, where F is the metric space $F = C([0, 1], \mathbb{R}^m) \times PRC(([0, 1[, Q)$ with the topology given by

$$\operatorname{dist}((u_1, \sigma_1), (u_2, \sigma_2)) = \|u_1 - u_2\|_{C([0,1],\mathbb{R}^m)} + \rho_1(\sigma_1, \sigma_2).$$

$\mathfrak{E}(x, q)$ is called a *transversal* of A if at every switching point $t_1 < \ldots < t_k$ of (u, σ) the C^∞ function given by

$$t \mapsto g_{\sigma(t_{i-1}),\sigma(t_i)}(E_{\sigma(t_{i-1})}(t - t_{i-1})u(t_{i-1}))$$

changes sign at t_i and its first derivative does not vanish there. Here, $i = 1, \ldots, k$ and $t_0 = 0$.

Thanks to the chain rule, the above condition is equivalent to the requirement

$$f(E_{\sigma(t_{i-1})}(t_i - t_{i-1})u(t_{i-1}), \sigma(t_{i-1})) \notin \ker Dg_{\sigma(t_{i-1}),\sigma(t_i)}(u(t_i)),$$

where $Dg_{q,p}(x): \mathbb{R}^m \to \mathbb{R}$ denotes the derivative of $g_{q,p}$ at the point x.

An admissible state $(x, q) \in S_0$ of A is called a *regular state* of A if $\mathfrak{E}(x, q)$ is a transversal of A. We define, for $q \in Q$,

$$\delta^0(q) = \{x \in \delta(q) | \mathfrak{E}(x, q) \text{ is a transversal of } A\}.$$

PROPOSITION 2.4. For each $q \in Q$, the set $\delta^0(q)$ is open in $\mathbb{R}^m$. The restriction $\mathfrak{E}(\cdot, q)|\delta^0(q)$ is continuous. The restriction $\pi_1\mathfrak{E}_2(\cdot, q)|\delta^0(q)$ is a C^∞ map.

Proof. Let (x_0, q_0) be given, with $x_0 \in \delta^0(q_0)$, $q_0 \in Q$. Let $(u, \sigma) = \mathfrak{E}(x_0, q_0)$, $(v, \mu) = \mathfrak{E}(x_0', q_0)$, with $|x_0 - x_0'| < \varepsilon$, where $\varepsilon > 0$ is yet to be determined. (Here $|\cdot|$ denotes any convenient norm on $\mathbb{R}^m$.)

Define $t_0 = t_0' = 0$ and let $t_1 < \ldots < t_k$ and $t_1' < \ldots . < t_l'$ denote the switching points of (u, σ) and of (v, μ), respectively. Define $q_i = \sigma(t_i)$, $i = 1, \ldots, k$; $q_i' = \mu(t_i')$, $i = 1, \ldots, l$. Further, let $t_{k+1} = 1$.

For $i = 0, \ldots, k$, define

$$\Phi_i: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R} \times \mathbb{R}^m: (t, x) \mapsto (t, E_{q_i}(t - t_i)x).$$

Because each flow $(t, x) \mapsto E_q(t)x$ is necessarily C^∞, [3, theorem 5.2, pp. 144–145], each Φ_i is also a C^∞ map. It follows that each of the maps

$$(t, x) \mapsto E_{q_i}(t - t_i) \ldots E_{q_0}(t)x$$

is also C^∞, since each is a composition of C^∞ maps:

$$(t, E_{q_i}(t - t_i) \ldots E_{q_0}(t)x) = \Phi_i \circ \ldots \circ \Phi_0(t, x).$$

For $i = 0, \ldots, k - 1$, define

$$\phi_i: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}: (t, x) \mapsto g_{q_i,q_{i+1}}(E_{q_i}(t - t_i) \ldots E_{q_0}(t)x).$$

Each ϕ_i is a C^∞ function. Moreover, because (u, σ) is a transversal of A, we have that $D_1\phi_i(t_{i+1}, x_0) > 0$. Thus, thanks to the implicit function theorem, for ε sufficiently small, there are uniquely determined C^∞ functions $\theta_1, \ldots, \theta_k: B_\varepsilon \to \mathbb{R}$, where B_ε is an open ball in $\mathbb{R}^m$

centered at x_0, such that $\phi_i(\theta_{i+1}(x), x) = 0$ for all $x \in B_\varepsilon$. Moreover, if ε is small enough, $\phi_i(\cdot, x)$ changes sign at $\theta_{i+1}(x)$, since $D_1\phi_i(\theta_{i+1}(x), x) > 0$ also.

For $b_0 > t_1$, but sufficiently close to t_1, the function $\phi_0(\cdot, x_0)$ changes sign at t_1 and is nonvanishing elsewhere in $[0, b_0]$. Therefore, the point $\theta_1(x_0')$ is the least $t > 0$ such that $\phi_0(t, x_0') = 0$.

It follows from the minimality of $\theta_1(x_0')$ that

$$E_{q_0}(t)x_0' \notin M_{q_0,q_1} \quad \text{for} \quad t \in [0, \theta_1(x_0')[.$$

To show that we can choose ε so small that

$$E_{q_0,p}(t)x_0' \notin M_{q_0,p} \quad \text{for} \quad p \in I(q_0)\backslash\{q_1\}, \quad t \in [0, \theta_1(x_0')],$$

consider the inequality, for $p \in I(q_0)\backslash\{q_1\}$,

$$q_{q_0,p}(E_{q_0}(t)x_0) < 0,$$

which holds for $t \in [0, t_1]$, thanks to (DA1). By continuity, it also holds for $t \in [0, c_1]$, where $c_1 = \max\{t_1, \theta_1(x_0')\}$, provided that ε is sufficiently small. This means that for ε small enough we have

$$g_{q_0,p}(E_{q_0}(t)x_0') < 0$$

for $t \in [0, c_1]$, $p \in I(q_0)\backslash\{q_1\}$. We conclude that $t_1' = \theta_1(x_0')$ and that $q_1' = q_1$.

Let K be a compact subset of $\mathbb{R}^m$ such that $u(t)$ and $v(t)$ are both in K for $t \in [0, 1]$. Let $\lambda \geqslant 1$ be a bound for the derivatives of the flows $E_q(\cdot)(\cdot)$ on $[0, 1] \times K$, i.e.

$$|E_q(t)x - E_q(t')y| \leqslant \lambda(|t - t'| + |x - y|),$$

for $t, t' \in [0, 1]$, $x, y \in K$, $q \in Q$.

For $0 \leqslant t \leqslant \min\{t_1, t_1'\}$, we have the bound

$$|u(t) - v(t)| = |E_{q_0}(t)x_0 - E_{q_0}(t)x_0'| \leqslant \lambda|x_0 - x_0'|.$$

For $t_1 \leqslant t_1'$ and $t_1 \leqslant t \leqslant t_1'$, we have the bound

$$\begin{aligned} |u(t) - v(t)| &= |E_{q_1}(t - t_1)u(t_1) - E_{q_0}(t - t_1)v(t_1)| \\ &\leqslant |E_{q_1}(t - t_1)u(t_1) - E_{q_1}(0)u(t_1)| + |u(t_1) - v(t_1)| \\ &\quad + |E_{q_0}(0)v(t_1) - E_{q_0}(t - t_1(v(t_1))| \\ &\leqslant 2\lambda|t_1 - t_1'| + \lambda|x_0 - x_0'|. \end{aligned}$$

For $t_1 \geqslant t_1'$ and $t_1' \leqslant t \leqslant t_1$, we obtain the same bound. In other words, we have the inequality

$$|u(t) - v(t)| \leqslant 2\lambda(|t_1 - t_1'| + |x_0 - x_0'|), \quad t \in [0, c_1].$$

Proceeding by induction, suppose that we have, for some $j < \min\{k, l\}$, that $q_i = q_i'$ and $t_i' = \theta_i(x_0')$, for $i = 1, \ldots, j$. Further suppose that

$$|u(t) - v(t)| \leqslant 2\lambda^j(\Sigma_{1 \leqslant i \leqslant j}|t_i - t_i'| + |x_0 - x_0'|)$$

for $t \in [0, c_j]$, where $c_j = \max\{t_j, t_j'\}$. Note: This means that we have

$$\rho_{c_j}(\sigma|[0, c_j[, \mu|[0, c_j[) \leqslant \max_{1 \leqslant i \leqslant j} |\theta_i(x_0) - \theta_i(x_0')|.$$

For $a_j = \min\{t_j, t_j'\}$ and $b_j > t_{j+1}$, but sufficiently close to t_{j+1}, and for ε sufficiently small, the function $\phi_j(\cdot, x_0)$ changes sign at t_{j+1} and is nonvanishing elsewhere in $[a_j, b_j]$. Therefore, the point $\theta_{j+1}(x_0')$ is the least $t > a_j$ such that $\phi_j(t, x_0') = 0$, provided that $\varepsilon > 0$ is small enough.

From the minimality of $\theta_{j+1}(x_0')$ we have that

$$E_{q_j}(t - t_j')v(t_j') \notin M_{q_j, q_{j+1}}, t \in [t_j', \theta_{j-1}(x_0')[.$$

To show that for sufficiently small $\varepsilon > 0$ we also have

$$E_{q_j}(t - t_j')v(t_j') \notin M_{q_j, p}, p \in I(q_j)\backslash\{q_{j+1}\}, t \in [t_j', \theta_{j+1}(x_0')],$$

consider the inequality, for $p \in I(q_j)\backslash\{q_{\mathrm{j}+1}\}$,

$$g_{q_j, p}(E_{q_j}(t - t_j)u(t_j)) < 0,$$

which holds for $t \in [t_j, t_{j+1}]$. By continuity, it also holds for $t \in [a_j, c_{j-1}]$, where $c_j = \max\{t_{j+1}, \theta_{j+1}(x_0')\}$, provided that ε is sufficiently small. This means that for ε small enough we have the inequality

$$g_{q_j, p}(E_{q_j}(t - t_j')v(t_j')) < 0, \qquad t \in [a_j, c_{j+1}].$$

We conclude that $t_{j+1}' = \theta_{j+1}(x_0')$ and that $q_{j+1}' = q_{j+1}$.

For $c_j \leqslant t \leqslant \min\{t_{j+1}, t_{j+1}'\}$ we have the bound

$$|u(t) - v(t)| = |E_{q_j}(t - c_j)u(c_j) - E_{q_j}(t - c_j)v(c_j)| \leqslant \lambda |u(c_j) - v(c_j)|.$$

For $t_{j+1} \leqslant t_{j+1}'$ and $t_{j+1} \leqslant t \leqslant t_{j+1}'$ we have the bound

$$\begin{aligned} |u(t) - v(t)| &= |E_{q_{j+1}}(t - t_{j+1})u(t_{j+1}) - E_{q_j}(t - t_{j+1})v(t_{j+1})| \\ &\leqslant |E_{q_{j+1}}(t - t_{j+1})u(t_{j+1}) - E_{q_{j-1}}(0)u(t_{j+1})| \\ &\quad + |u(t_{j+1}) - v(t_{j+1})| \\ &\quad + |E_{q_j}(0)v(t_{j+1}) - E_{q_j}(t - t_{j+1})v(t_{j+1})| \\ &\leqslant 2\lambda |t_{j+1} - t_{j+1}'| + \lambda |u(c_j) - v(c_j)|. \end{aligned}$$

For the case where $t_{j+1} \geqslant t_{j-1}'$ and $t_{j+1}' \leqslant t \leqslant t_{j+1}$ we obtain the same bound. Thus, we have obtained the bound

$$|u(t) - v(t)| \leqslant 2\lambda^{j+1}(\Sigma_{1 \leqslant i \leqslant j+1} |t_i - t_i'| + |x_0 - x_0'|)$$

for $t \in [0, c_{j+1}]$.

From the above, it follows that we have $k = l$ and

$$\pi_2 \sigma = \pi_2 \mu,$$

$$\|\pi_1 \sigma - \pi_1 \mu\|_{l_\infty} \leqslant \max_{1 \leqslant j \leqslant k} |\theta_j(x_0) - \theta_j(x_0')|,$$

$$\|u - v\|_{C([0,1], \mathbb{R}^m)} \leqslant 2\lambda^{k+1}\left(\sum_{1 \leqslant i \leqslant k} |e_i(x_0) - \theta_i(x_0')| + |x_0 - x_0'|\right).$$

Thus, the map $\mathfrak{E}(\cdot, q_0)$ is continuous on $\delta^0(q_0)$. Moreover, we have, for $x \in B_\varepsilon$, and for integer $p \geqslant 0$,

$$D^p(\pi_1[\mathfrak{E}_2(\cdot, q)])(x): \mathbb{R}^{mp} \to l_\infty(\mathbb{R})$$
$$: y \mapsto \{0, D^p\theta_1(x)y, \ldots, D^p\theta_k(x)y, 0, \ldots\}.$$

Moreover, we have also shown that $\delta^0(q_0)$ contains a neighborhood of each of its points, since (v, μ) is a transversal of A for all sufficiently small $\varepsilon > 0$.

Definition 2.5. Let X be a C^∞ $(m-1)$-submanifold (without boundary) of $\mathbb{R}^m$. Let $E = \{E(t) | t \in \mathbb{R}\}$ be a one-parameter group of C^∞ diffeomorphisms of $\mathbb{R}^m$. We define the "tangent set"

$$\mathfrak{T}(E, X) = \{x \in \mathbb{R}^m | E(\cdot)x \quad \text{is tangent to} \quad X \quad \text{at some}$$
$$\text{point} \quad E(t)x \in X \quad \text{for some} \quad t \in]0, 1[\}.$$

In other words:

$$x \in \mathfrak{T}(E, X) \Leftrightarrow (y, D_1E(t_x)x) \in T_yX,$$

for some $t_x \in]0, 1[$, where T_yX is (the obvious identification of) the tangent space of X at the point $y \in X$, and where $y = E(t_x)x$.

Proposition 2.6. If E and X are as in the preceding definition, then $\mathfrak{T}(E, X)$ has empty interior.

Proof. Suppose the contrary. Then there is an open ball B in $\mathbb{R}^m$ such that for each $x \in B$ there is a real number $t_x \in]0, 1[$ such that $E(\cdot)x$ is tangent to X at $E(t_x)x$. B can be chosen so that $B \cap X = \varnothing$. Because $E(t)$ is an open map for each t, we have that $\cup_{x \in B} E(t_x)B$ is open in $\mathbb{R}^m$. Therefore, the set

$$V = X \cap (\cup_{x \in B} E(t_x)B)$$

is open in X. Thus, there is nonempty open (in X) connected $U \subset V$. Let $\phi \in C^\infty(\mathbb{R}^m, \mathbb{R}^m)$ be the infinitesimal generator of E. The restriction $\phi|U$ defines a map $U \to TU$ which is a smooth section of TU, the tangent bundle of U. This defines a flow on U. That is, given $y \in U$, there is an $\varepsilon > 0$ such that $t \mapsto E(t)y$, $t \in]-\varepsilon, \varepsilon[$, is an integral curve of $\phi|U$ passing through y with values in U. Because every point on this integral curve is also a point on the integral curve of ϕ passing through some $x \in B$, and because $B \cap U = \varnothing$, we obtain a contradiction on the uniqueness of integral curves, since ϕ cannot vanish at the point y.

Proposition 2.7. For each $q \in Q$, $\delta^0(q)$ is dense in $\delta(q)$.

Proof. For each $q_0 \in Q$ we have the inclusion

$$\delta(q_0) \backslash \delta^0(q_0) \subset \bigcup_{\substack{q \in Q \\ p \in I(q)}} \mathfrak{T}(E_q, \partial M_{q,p}),$$

showing that the set $\delta(q_0) \backslash \delta^0(q_0)$, which is closed in $\delta(q_0)$, has empty interior, thanks to proposition 2.6. We conclude that the closure of $\delta(q_0) \backslash \delta^0(q_0)$ is nowhere dense in $\delta(q_0)$.

3. NUMERICAL APPROXIMATIONS

Let $\mathfrak{P}$ denote the collection of all partitions of the interval $[0, 1]$. Let $\mathfrak{P}_0$ denote any subset of $\mathfrak{P}$ which is dense in $[0, 1]$. Let $\Delta \in \mathfrak{P}_0$. For $t \in \Delta \cap]0, 1]$ define

$$|\Delta|_t^- = \min\{|s - t| \,|\, s \in \Delta, s < t\}.$$

Further, define

$$|\Delta| = \max\{|\Delta|_t^- \,|\, t \in \Delta \cap]0, 1]\}.$$

We shall suppose that we have a collection $\mathfrak{M}$ of maps

$$\bar{E}_q(\Delta, s, \cdot) \in C([s, 1] \times \mathbb{R}^m, \mathbb{R}^m),$$

defined for $\Delta \in \mathfrak{P}_0$, $s \in \Delta$, such that given any bounded subset B of $\mathbb{R}^m$ there are continuous monotone increasing functions ζ, $\xi: \mathbb{R}_+ \to \mathbb{R}_+$ with $\zeta(0) = \xi(0) = 0$ such that the inequalities

$$\max_{\substack{s,t \in \Delta \\ s \leq t}} |E_q(t - s)x - \bar{E}_q(\Delta, s, t)x| \leq \zeta(|\Delta|), \tag{NA1}$$

$$\max_{\substack{t \in [s,1] \\ s \in \Delta}} |E_q(t - s)x - \bar{E}_q(\Delta, s, t)x| \leq \xi(|\Delta|), \tag{NA2}$$

hold for $q \in Q$, $\Delta \in \mathfrak{P}_0$, $x \in B$. There are one-step and multi-step methods of all orders ($\zeta(|\Delta|) = \xi(|\Delta|) = O(|\Delta|^\kappa)$, for any integer $\kappa \geq 1$) which satisfy the above [6, 7]. For some $c \geq 0$, $\gamma > 0$, we require that the bound

$$|\bar{E}_q(\Delta, s, t)x| \leq |x|e^{\gamma(t-s)} + c\gamma^{-1}(e^{\gamma(t-s)} - 1) \tag{NA3}$$

hold for all $q \in Q$, $\Delta \in \mathfrak{P}_0$, $s \in \Delta$, $t \in [s, 1]$, $x \in \mathbb{R}^m$. This is usually satisfied thanks to the Lipschitz continuity of the approximate increment function of the numerical method. Even though we shall only require that $\bar{E}_q(\Delta, s, t)$ be evaluated at the gridpoints to construct approximate solutions, (NA2) shall be required to prove the "right" order of convergence of the approximate switching points.

The motivation is as follows. The point $\bar{E}_q(\Delta, s, t)x$ in $\mathbb{R}^m$ denotes the approximate solution at the gridpoint t defined by a polyalgorithmic numerical integrator applied to the initial value problem

$$u'(\tau) = f(u(\tau), q), \qquad s < \tau,$$
$$u(s) = x,$$

using the grid $\Delta \cap [s, 1]$. Because $\bar{E}_q$ may be defined by a polyalgorithm, we do not assume the "semigroup" property

$$\bar{E}_q(\Delta, s, t)x = \bar{E}_q(\Delta, t', t)\bar{E}_q(\Delta, s, t')x$$

for $0 \leq s \leq t' \leq t \leq 1$. The reason is that the approximate solution defined over $\Delta \cap [t', t]$ by the application of $\bar{E}_q(\Delta, s, \cdot)$ may be evaluated by a multistep method for the case $s < t'$, while the approximate solution defined over $\Delta \cap [t', t]$ by the application of $\bar{E}_q(\Delta, t', \cdot)$ needs a self-starting numerical method, which may yield a different approximation.

3.1. *Approximate trajectories*

Fix $\mathfrak{M}$. $\Delta \in \mathfrak{P}_0$, $q_0 \in Q$, $x_0 \in \delta(q_0)$. By the $(\mathfrak{M}, \Delta)$-*approximate trajectory* of (A, x_0, q_0) we mean the pair of maps $(\bar{u}, \bar{\sigma})$, $\bar{u}: \Delta \to \mathbb{R}^m$ and $\bar{\sigma}: \Delta \cap [0, 1[\to Q$ defined as follows.

Let t_1 denote the maximal element of Δ such that

$$g_{q0,p}(\bar{E}_{q0}(\Delta, 0, t)x_0) < 0$$

for all $p \in I(q_0)$ and all $t \in \Delta_1 \equiv \Delta \cap [0, t_1[$. (Note: Δ_1 may be a singleton, but cannot be empty.) Define

$$\bar{u}(t) = \bar{E}_{q_0}(\Delta, 0, t)x_0, \qquad t \in \Delta \cap [0, t_1],$$
$$\bar{\sigma}(t) = q_0, \qquad t \in \Delta_1.$$

If $t_1 = 1$, we are finished. Otherwise, we have that

$$g_{q_0,q_1}(\bar{u}(t_1)) \geq 0$$

for some $q_1 \in I(q_0)$. Thanks to (DA1), this q_1 is unique. Proceeding inductively, suppose that we have

$$0 = t_0 < t_1 < \cdots < t_j < 1 \text{ in } \Delta;$$
$$q_1 \in I(q_0), \ldots, q_j \in I(q_{j-1});$$
$$\bar{u}(t_1) \in \delta(q_1), \ldots, \bar{u}(t_{j-1}) \in \delta(q_{j-1});$$

$$\bar{u}(t) = \bar{E}_{q_i}(\Delta, t_i, t)u(t_i), \qquad t \in \Delta \cap [t_i, t_{i+1}],$$
$$\bar{\sigma}(t) = q_i, \qquad t \in \Delta_{i+1} \equiv \Delta \cap [t_i, t_{i+1}[,$$

for $i = 0, \ldots, j-1$; such that for $i = 0, \ldots, j-1$ we have

$$g_{q_i,p}(\bar{u}(t)) < 0, \qquad t \in \Delta_{i+1},$$

for all $p \in I(q_i)$, and such that

$$g_{q_i,q_{i+1}}(\bar{u}(t_{i+1})) \geq 0,$$

for $i = 0, \ldots, j-1$.

Let t_{j+1} denote the maximal element of Δ such that

$$g_{q_j,p}(\bar{E}_{q_j}(\Delta, t_j, t)\bar{u}(t_j)) < 0$$

for all $p \in I(q_j)$ and for all $t \in \Delta_{j+1} \equiv \Delta \cap [t_j, t_{j+1}[$. Define

$$\bar{u}(t) = \bar{E}_q(\Delta, t_j, t)\bar{u}(t_j), \qquad t \in \Delta \cap [t_j, t_{j+1}],$$
$$\bar{\sigma}(t) = q_j, \qquad t \in \Delta_{j+1}.$$

If $t_{j+1} = 1$, we are finished. Otherwise, we have that

$$g_{q_j,q_{j+1}}(\bar{u}(t_{j+1})) \geq 0$$

for some $q_{j+1} \in I(q_j)$. Thanks to (DA1), this q_{j+1} is unique.

A point $t \in \Delta \cap]0, 1[$ where $\bar{\sigma}(t - |\Delta|_t^-) \neq \bar{\sigma}(t)$ is called a *switching point* of $(\bar{u}, \bar{\sigma})$.

The motivation is as follows [8]. Given the initial state (x_0, q_0), we apply a numerical integrator to the initial value problem $u_1'(\tau) = f(u_1(\tau), q_0)$, $\tau > 0$, $u_1(0) = x_0$. We keep up the numerical integration to define $\bar{u}_1(t)$ for successive values of $t \in \Delta$ as long as we have that $q_{q,p}(\bar{u}_1(t)) < 0$ for all $p \in I(q_0)$. Let t_1 denote the least $t \in \Delta$ such that any one of these inequalities fails, say for $p = q_1$. To extend the approximate solution past the point t_1, we apply a numerical integrator to the initial value problem $u_2'(\tau) = f(u_2(\tau), q_1)$, $\tau > t_1$, $u_2(t_1) = u_1(t_1)$. We keep up the numerical integration as long as the inequality $g_{q_1,p}(\bar{u}_2(t)) < 0$ is satisfied for all $p \in I(q_1)$. By induction, we define the approximate solution $\bar{u}$ by piecing together $\bar{u}_1, \ldots, \bar{u}_k$ for some k. Of course, the idea is to adaptively allocate the nodes for the numerical integration scheme so that a relatively "small" step is used near the switching points. This point shall be discussed in detail below.

3.2. *Convergence of approximate trajectories*

For $\Delta \in \mathfrak{P}_0$, $q \in Q$, define

$$\bar{\mathfrak{E}}_1(\Delta, \cdot, q) : \delta(q) \to (\mathbb{R}^m)^\Delta : x \mapsto \bar{u},$$
$$\bar{\mathfrak{E}}_2(\Delta, \cdot, q) : \delta(q) \to Q^{\Delta \cap [0,1[} : x \mapsto \bar{\sigma},$$

where $(\bar{u}, \bar{\sigma})$ is the $(\mathfrak{M}, \Delta)$-approximate trajectory of the initial value problem (A, x, q). Further, define $\bar{\mathfrak{E}}(\Delta, x, q) = (\bar{u}, \bar{\sigma})$.

Because $\mathfrak{E}(\cdot, q)$ and $\bar{\mathfrak{E}}(\Delta, \cdot, q)$ have different codomains, to measure the distance between exact and approximate solutions, we introduce the "distance" function d^Δ as follows, Define

$$Q^{\Delta \cap [0,1[} \to 2^{PRC([0,1[,Q)} : \mu \mapsto [\mu]$$

by $\mu^* \in [\mu] \Leftrightarrow \mu^* | \Delta \cap [0, 1[= \mu$. Define

$$\begin{aligned} \rho^\Delta &: PRC([0, 1[, Q) \times Q^{\Delta \cap [0,1[} \to \mathbb{R}_+ \\ &: (\sigma, \mu) \mapsto \sup\{\rho_1(\sigma, \mu^*) | \mu^* \in [\mu]\}, \\ d^\Delta &: C([0, 1], \mathbb{R}^m) \times PRC([0, 1[, Q) \times (\mathbb{R}^m)^\Delta \times Q^{\Delta \cap [0,1[} \to \mathbb{R}_+ \\ &: ((u, \sigma), (v, \mu)) \mapsto \max\{|u(t) - v(t)| \, | t \in \Delta\} + \rho^\Delta(\sigma, \mu). \end{aligned}$$

We shall need the following technical lemma.

PROPOSITION 3.3. Consider $u \in C^\infty([a, b], \mathbb{R}^m)$, $g \in C^\infty(\mathbb{R}^m, \mathbb{R})$. Let $g \circ u$ change sign at the point $t_0 \in]a, b[$ and be nonzero at all other points in the interval $[a, b]$. If $D(g \circ u)(t_0) \neq 0$, then there is an $\varepsilon > 0$ such that given any $v \in C([a, b], \mathbb{R}^m)$ with $|u(t) - v(t)| < \varepsilon$, for all $t \in [a, b]$, the function $g \circ v$ also changes sign in $]a, b[$ and at all such points t the inequality $|t - t_0| = O(|u(t) - v(t)|)$ holds.

Proof. We have $|D(g \circ u)(t)| \geqslant c > 0$ for some c and for all t in some open interval N containing t_0. Thus, for any $\eta \in C([a, b], \mathbb{R})$ small enough, $g \circ u + \eta$ also changes sign in $]a, b[$. Moreover, η can be chosen so small that all sign changes of $g \circ u + \eta$ are in N. This gives $|t - t_0| = O(|\eta(t)|)$ at all points where $g \circ u + \eta$ changes sign. The rest follows by taking $\eta = g \circ v - g \circ u$ and then using the chain rule.

PROPOSITION 3.4. If $q_0 \in Q$ and $x_0 \in \delta^0(q_0)$, we then have

$$d^\Delta(\mathfrak{E}(x_0, q_0), \bar{\mathfrak{E}}(\Delta, x_0', q_0)) \to 0,$$

as $|\Delta| + |x_0 - x_0'| \to 0, \Delta \in \mathfrak{P}_0, x_0' \in \mathbb{R}^m$.

Proof. Let (x_0, q_0) be given, with $x_0 \in \delta^0(q_0)$ and $q_0 \in Q$. Let $(u, \sigma) = \mathfrak{E}(x_0, q_0)$, $(\bar{u}, \sigma) = \mathfrak{E}(\Delta, x_0', q_0)$, with $|\Delta| + |x_0 - x_0'| < \varepsilon$, where $\varepsilon > 0$ is yet to be determined. Define $t_0 = t_0' = 0$, and let $t_1 < \cdots < t_k$ and $t_1' < \cdots < t_l'$ denote the switching points of (u, σ) and of $(\bar{u}, \bar{\sigma})$, respectively. Define $q_i = \sigma(t_i)$, $i = 1, \ldots, k$, and $q_i' = \bar{\sigma}(t_i')$, $i = 1, \ldots, l$. Further, let $t_{k+1} = t_{l+1}' = 1$.

Thanks to (NA3), it is easy to see that we have $|\bar{u}(t)| \leqslant |x_0'| \, e^\gamma + c\gamma^{-1} (e^\gamma - 1)$ for $t \in [0, 1]$, i.e.: $\bar{u}(t)$ remains bounded independently of the partition Δ and of the number l of switching points. Therefore, there is a ball B in $\mathbb{R}^m$ which contains $u(t)$ and $\bar{u}(t)$, $t \in [0, 1]$, for all sufficiently small $\varepsilon > 0$. The constant λ and the functions ζ and ξ which appear below depend on B.

Consider $\phi_0 \in C^\infty([0, b_0], \mathbb{R})$ and $\psi_0 \in C([0, b_0], \mathbb{R})$ given by $\phi_0(t) = g_{q_0,q_1}(E_{q_0}(t)x_0)$ and $\psi_0(t) = g_{q_0,q_1}(\bar{E}_{q_0}(\Delta, 0, t)x_0')$. For $b_0 > t_1$, but sufficiently close to t_1, and for sufficiently small $\varepsilon > 0$, thanks to (NA2), the hypotheses of proposition 3.3 are satisfied. Therefore, denoting by t_1'' the least t such that ψ_0 changes sign at t, we have $|t_1'' - t_1| = O(\xi(|\Delta|) + |x_0 - x_0'|)$ as $\varepsilon \to 0$. We can make $\phi_0 - \psi_0$ as small as we please in the topology of $C([0, b_0], \mathbb{R})$ by choosing ε small enough relative to some choice of $b_0 > t_1$, but close enough to t_1. Therefore, by continuity we can exclude the possibility that $q_{q_0,p}(\bar{E}_{q_0}(\Delta, 0, t)x_0') \geqslant 0$ for any $t \in \Delta$, $t \leqslant b_0$, $p \in I(q_0) \backslash \{q_1\}$. Thus, we have the bounds $|t_1'' - t_1'| \leqslant |\Delta|_{t_1}^-$ and

$$|t_1' - t_1| \leqslant |\Delta|_{t_1}^- + O(\xi(|\Delta|) + |x_0 - x_0'|).$$

Moreover, we have $q_1' = q_1$. For $t \in \Delta$, $0 \leqslant t \leqslant \min\{t_1, t_1'\}$, we have the bound

$$|u(t) - \bar{u}(t)| = |E_{q_0}(t)x_0 - \bar{E}_{q_0}(\Delta, 0, t)x_0'| \leqslant \zeta(|\Delta|) + \lambda|x_0 - x_0'|,$$

thanks to (NA1), where λ is a bound for the first derivative of the flows on $[0, 1] \times B$.

For $t_1 \leqslant t_1'$, $t \in \Delta$, $t_1 \leqslant t \leqslant t_1'$, we have the bound

$$\begin{aligned}|u(t) - \bar{u}(t)| &= |E_{q_1}(t - t_1)u(t_1) - \bar{E}_{q_0}(\Delta, 0, t)x_0'| \\ &\leqslant |E_{q_1}(t - t_1)E_{q_0}(t_1)x_0 - E_{q_1}(0)E_{q_0}(t_1)x_0| \\ &\quad + |E_{q_0}(0)E_{q_0}(t_1)x_0 - E_{q_0}(t - t_1)E_{q_0}(t_1)x_0'| \\ &\quad + |E_{q_0}(t)x_0' - \bar{E}_{q_0}(\Delta, 0, t)x_0'| \\ &\leqslant 2\lambda|t - t_1'| + \lambda|x_0 - x_0'| + \zeta(|\Delta|).\end{aligned}$$

For $t_1' \leqslant t_1, t \in \Delta, t_1' \leqslant t \leqslant t_1$, we have the bound

$$\begin{aligned}|u(t) - \bar{u}(t)| &= |E_{q_0}(t)x_0 - \bar{E}_{q_1}(\Delta, t_1', t)\bar{u}(t_1')| \\ &\leqslant |E_{q_0}(t - t_1')u(t_1') - E_{q_0}(0)u(t_1')|\end{aligned}$$

$$\begin{aligned}&+|E_{q_1}(0)u(t_1') - E_{q_1}(t - t_1')\bar{u}(t_1')|\\ &+|E_{q_1}(t - t_1')\bar{u}(t_1') - \bar{E}_{q_1}(\Delta, t_1', t)\bar{u}(t_1')|\\ \leqslant\ &2\lambda|t_1 - t_1'| + \lambda|u(t_1') - u(t_1')| + \zeta(|\Delta|).\end{aligned}$$

We have obtained the bound

$$|u(t) - \bar{u}(t)| = O(|\Delta|_{t_1}^{-} + \xi(|\Delta|) + \zeta(|\Delta|) + |x_0 - x_0'|),$$

for $t \in [0, c_1]$, where $c_1 = \max\{t_1, t_1'\}$.

Proceeding by induction, suppose that we have, for some $j < \min\{k, l\}$, that $q_i = q_i'$ and

$$|t_i - t_i'| \leqslant |\Delta|_{t_i}^{-} + O(\Sigma_{1 \leqslant r \leqslant i-1}|\Delta|_{t_r}^{-} + \xi(|\Delta|) + \zeta(|\Delta|) + |x_0 - x_0'|),$$

for $i = 1, \ldots, j$. Further suppose that we also have

$$|u(t) - \bar{u}(t)| = O(\Sigma_{1 \leqslant i \leqslant j}|\Delta|_{t_i}^{-} + \xi(|\Delta|) + \zeta(|\Delta|) + |x_0 - x_0'|)$$

for $t \in [0, c_j]$, where $c_j = \max\{t_j, t_j'\}$.

Consider $\phi_j \in C^\infty([a_j, b_j], \mathbb{R})$ and $\psi_j \in C([a_j, b_j], \mathbb{R})$ given by

$$\begin{aligned}\phi_j(t) &= g_{q_j, q_{j+1}}(E_{q_j}(t - t_j)u(t_j)),\\ \psi_j(t) &= g_{q_j, q_{j+1}}(\bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')),\end{aligned}$$

where $a_j = \min\{t_j, t_j'\}$. For $b_j > t_{j+1}$, but sufficiently close to t_{j+1}, thanks to proposition 3.3 we conclude that $q_{j+1} = q_{j+1}'$, and obtain the bound

$$|t_{j+1} - t_{j+1}'| \leqslant |\Delta|_{t_{j+1}}^{-} + O(\Sigma_{1 \leqslant i \leqslant j}|\Delta|_{t_i}^{-} + \xi(|\Delta|) + \zeta(|\Delta|) + |x_0 - x_0'|),$$

as $\varepsilon \to 0$. For $c_j \leqslant t \leqslant \min\{t_{j+1}, t_{j+1}'\}$, $t \in \Delta$, we have

$$|u(t) - \bar{u}(t)| = |E_{q_j}(t - c_j)u(c_j) - \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')|.$$

If $c_j = t_j'$, the above is bounded by $\zeta(|\Delta|) + \lambda|u(c_j) - \bar{u}(c_j)|$. If $t_j' < t_j$, we have the bound

$$\begin{aligned}|u(t) - u(t)| &= |E_{q_j}(t - t_j)E_{q_{j-1}}(t_j - t_j')u(t_j') - \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')|\\ &\leqslant |E_{q_j}(t - t_j)E_{q_{j-1}}(t_j - t_j')u(t_j') - E_{q_j}(t - t_j)E_{q_{j-1}}(0)u(t_j')|\\ &\quad + |E_{q_j}(t - t_j)u(t_j') - E_{q_j}(t - t_j')\bar{u}(t_j')|\\ &\quad + |E_{q_j}(t - t_j')\bar{u}(t_j') - \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')|\\ &\leqslant \lambda|E_{q_{j-1}}(t_j - t_j')u(t_j') - E_{q_{j-1}}(0)u(t_j')| + \lambda|t_j - t_j'|\\ &\quad + \lambda|u(t_j') - \bar{u}(t_j')| + \zeta(|\Delta|)\\ &\leqslant \lambda(1 + \lambda)|t_j - t_j'| + \lambda|u(t_j') - \bar{u}(t_j')| + \zeta(|\Delta|).\end{aligned}$$

We have obtained a bound which holds for $t \in \Delta$, where $c_j \leqslant t \leqslant \min\{t_{j+1}, t_{j+q}'\}$. Below we compute a bound for the case m in$\{t_{j+1}, t_{j+1}'\} \leqslant t \leqslant \max\{t_{j+1}, t_{j+1}'\}$.

For $t_{j+1} \leqslant t_{j+1}'$, $t_{j+1} \leqslant t \leqslant t_{j+1}'$, $t_j' \leqslant t_j$, we have

$$|u(t) - \bar{u}(t)| = |E_{q_{j+1}}(t - t_{j+1})E_{q_j}(t_{j+1} - t_j)E_{q_{j-1}}(t_j - t_j')u(t_j')$$

$$
\begin{aligned}
&- \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')| \\
&\leqslant |E_{q_{j+1}}(t - t_{j+1})E_{q_j}(t_{j+1} - t_j)E_{q_{j-1}}(t_j - t_j')u(t_j') \\
&\quad - E_{q_{j+1}}(t - t_{j+1})E_{q_j}(t_{j+1} - t_j)E_{q_{j-1}}(0)u(t_j')| \\
&\quad + |E_{q_{j+1}}(t - t_{j+1})E_{q_j}(t_{j+1} - t_j)u(t_j') - E_{q_{j+1}}(0)E_{q_j}(t_{j+1} - t_j)u(t_j')| \\
&\quad + |E_{q_j}(t_{j+1} - t_j)u(t_j') - E_{q_j}(t - t_j')\bar{u}(t_j')| \\
&\quad + |E_{q_j}(t - t_j')\bar{u}(t_j') - \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')| \\
&\leqslant \lambda^2(1 - \lambda)|t_j - t_j'| + \lambda|t_{j+1} - t_{j+1}'| \\
&\quad + \lambda|u(t_j') - \bar{u}(t_j')| + \zeta(|\Delta|).
\end{aligned}
$$

For $t_{j+1} \leqslant t_{j+1}', t_{j+1} \leqslant t \leqslant t_{j+1}', t_j \leqslant t_j'$, we have

$$
\begin{aligned}
|u(t) - \bar{u}(t)| &= |E_{q_{j+1}}(t - t_{j+1})E_{q_j}(t_{j+1} - t_j')u(t_j') \\
&\quad - \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')| \\
&\leqslant |E_{q_{j+1}}(t - t_{j+1})E_{q_j}(t_{j+1} - t_j')u(t_j') \\
&\quad - E_{q_{j+1}}(0)E_{q_j}(t_{j+1} - t_j')u(t_j')| \\
&\quad + |E_{q_j}(t_{j+1} - t_j')u(t_j') - E_{q_j}(t - t_j')\bar{u}(t_j')| \\
&\quad + |E_{q_j}(t - t_j')\bar{u}(t_j') - \bar{E}_{q_j}(\Delta, t_j', t)\bar{u}(t_j')| \\
&\leqslant 2\lambda|t_{j+1} - t_{j+1}'| + \lambda|u(t_j') - \bar{u}(t_j')| + \zeta(|\Delta|).
\end{aligned}
$$

For $t_{j+1}' \leqslant t_{j+1}, t_{j+1}' \leqslant t \leqslant t_{j+1}, t_j' \leqslant t_j$, we have

$$
\begin{aligned}
|u(t) - \bar{u}(t)| &= |E_{q_j}(t - t_{j+1}')u(t_{j+1}') - \bar{E}_{q_{j+1}}(\Delta, t_{j+1}', t)\bar{u}(t_{j+1}')| \\
&\leqslant |E_{q_j}(t - t_{j+1}')u(t_{j+1}') - E_{q_j}(0)u(t_{j+1}')| \\
&\quad + |E_{q_{j+1}}(0)u(t_{j+1}') - E_{q_{j+1}}(t - t_{j+1}')\bar{u}(t_{j+1}')| \\
&\quad + |E_{q_{j+1}}(t - t_{j+1}')\bar{u}(t_{j+1}') - \bar{E}_{q_{j+1}}(\Delta, t_{j+1}', t)\bar{u}(t_{j+1}')| \\
&\leqslant 2\lambda|t_{j+1} - t_{j+1}'| + \lambda|u(t_{j+1}') - \bar{u}(t_{j+1}')| + \zeta(|\Delta|) \\
&\leqslant 2\lambda|t_{j+1} - t_{j+1}'| + 2\lambda^2|t_j - t_j'| + \lambda^2|u(t_j') - \bar{u}(t_j')| \\
&\quad + (1 + \lambda)\zeta(|\Delta|).
\end{aligned}
$$

For $t_{j+1}' \leqslant t_{j+1}, t_{j+1}' \leqslant t \leqslant t_{j+1}, t_j \leqslant t_j'$, we have

$$
\begin{aligned}
|u(t) - u(t)| &= |E_{q_j}(t - t_{j+1}')u(t_{j+1}') - \bar{E}_{q_{j+1}}(\Delta, t_{j+1}', t)\bar{u}(t_{j+1}')| \\
&\leqslant |E_{q_j}(t - t_{j+1}')u(t_{j+1}') - E_{q_j}(0)u(t_{j+1}')| \\
&\quad + |E_{q_{j+1}}(0)u(t_{j+1}') - E_{q_{j+1}}(t - t_{j+1}')\bar{u}(t_{j+1}')| \\
&\quad + |E_{q_{j+1}}(t - t_{j+1}')\bar{u}(t_{j+1}') - \bar{E}_{q_{j+1}}(\Delta, t_{j+1}', t)\bar{u}(t_{j+1}')|
\end{aligned}
$$

$$\leqslant 2\lambda|t_{j+1} - t'_{j+1}| + \lambda|u(t'_{j+1}) - \bar{u}(t'_{j+1})| + \zeta(|\Delta|)$$
$$\leqslant 2\lambda|t_{j+1} - t'_{j+1}| + 2\lambda^2|t_j - t'_j| + \lambda^2|u(t'_j) - \bar{u}(t'_j)|$$
$$+ (1 + \lambda)\zeta(|\Delta|).$$

In all of the above cases, thanks to the induction hypothesis, we obtain the bound

$$|u(t) - \bar{u}(t)| = O(\Sigma_{1\leqslant i\leqslant j+1}|\Delta|^{-}_{t_i} + \xi(|\Delta|) + \zeta(|\Delta|) + |x_0 - x'_0|),$$

for $t \in [0, c_{j-1}]$, where $c_{j+1} = \max\{t_{j+1}, t'_{j+1}\}$.

3.5. *Order of convergence*

Let k denote the number of switching points of the solution $\mathfrak{E}(x, q)$ of the initial value problem (A, x, q), where $x \in \delta^0(q)$. Then, thanks to proposition 3.4, there is a number $\varepsilon_0(x, q) > 0$ such that the number of switching points of $\bar{\mathfrak{E}}(\Delta, x', q)$ is also k whenever $|\Delta| + |x - x'| < \varepsilon_0(x, q)$. Fix x and q, $x \in \delta^0(q)$. Define the set

$$\mathfrak{S}(x, q) = \{(\Delta, x') \in \mathfrak{P}_0 \times \delta^0(q) \,|\, |\Delta| + |x - x'| < \varepsilon_0(x, q)\}.$$

There are k functions $\tau_1, \ldots, \tau_k : \mathfrak{S}(x, q) \to \,]0, 1[$ such that $\tau_i(\Delta, x') \to t_i$ as $|\Delta| + |x - x'| \to 0$, *where* t_i is the ith switching point of $\mathfrak{E}(x, q)$ and where $\tau_i(\Delta, x')$ is the ith switching point of $\bar{\mathfrak{E}}(\Delta, x', q)$.

We can restate proposition 3.4.

COROLLARY 3.6. If $q \in Q$ and $x \in \delta^0(q)$, we then have

$$d^\Delta(\mathfrak{E}(x, q), \bar{\mathfrak{E}}(\Delta, x', q)) = O(\Sigma_{1\leqslant i\leqslant k}|\Delta|^{-}_{\tau_i(\Delta, x')} + \xi(|\Delta|) + \zeta(|\Delta|) + |x - x'|),$$

for $(\Delta, x') \in \mathfrak{S}(x, q)$, as $|\Delta| + |x - x'| \to 0$, where k denotes the number of switching points $\bar{\mathfrak{E}}(\Delta, x', q)$.

3.7. *Step-size control*

The above result can be used to give a precise meaning to the idea of choosing Δ adaptively as the numerical integration is carried out: If in (NA1) and (NA2) we have a method of order κ, i.e. $\zeta(|\Delta|) = \xi(|\Delta|) = O(|\Delta|^\kappa)$, any sequence $\Delta_1, \Delta_2, \ldots$ of partitions in $\mathfrak{P}_0$, with $|\Delta_j| \to 0$ as $j \to \infty$, will yield $O(|\Delta_j|^\kappa)$ convergence as $j \to \infty$, provided we do the following.

The switching points of the approximate trajectory being computed can be detected during the computation. We do not integrate past a switching point until we have reduced the integration step (locally) to satisfy the inequality $|\Delta_j|^{-}_{\tau_i(\Delta_j, x')} \leqslant c|\Delta_j|^\kappa$, where $c > 0$ is any constant. $O(|\Delta_j|^\kappa + |x - x'|)$ convergence as $j \to \infty$ follows.

Alternatively, we may proceed as follows. As we carry out the numerical integration, suppose that we have a grid-point t such that the inequalities $g_{q,p}(\bar{u}(t - |\Delta|^{-}_t)) < 0$ and $g_{q,p}(\bar{u}(t)) \geqslant 0$ hold for some $p \in I(q)$, where q is the current discrete state. We can use $\bar{E}_q$ as an interpolation formula to define the function $r \mapsto g_{q,p}(\bar{u}(r))$ on $[t - |\Delta|^{-}_t, t]$. (The evaluation of this function requires no further evaluation of the derivatives $f(\cdot, q)$.) If we can solve for r the equation $g_{q,p}(\bar{u}(r)) = 0$, we can then use any solution as a grid point. If we do this for every switching point of the approximate solution, the error is

$$|u(t) - \bar{u}(t)| = O(\xi(|\Delta|) + \zeta(|\Delta|) + |x - x'|), \qquad t \in [0, 1].$$

Remarks 3.8. Because the notion of genericity, in the context used here, is perhaps new to the numerical analysis of initial value problems, a few remarks appear to be in order.

If (x, q) is an admissible state, but not a regular state, of the automaton A, we have said nothing about the well-posedness of the initial value problem (A, x, q) nor about the convergence of $\bar{\mathfrak{E}}(\Delta, x, q)$ to $\mathfrak{E}(x, q)$. However, we have shown that there are arbitrarily small perturbations y of x such that (y, q) is a regular state of A. On the other hand, a regular state (x, q) remains a regular state under all sufficiently small perturbations of x.

If the automaton A is supposed to model a physical device, it would be physically meaningless to look at the initial value problem (A, x, q) for $x \in \delta(q) \backslash \delta^0(q)$, because this set is nowhere dense in $\delta(q)$. Physical considerations aside, the fixed-precision numerical evaluation of $\bar{\mathfrak{E}}(\Delta, x, q)$, for $x \in \delta(q) \backslash \delta^0(q)$, would be rather meaningless for the same reason. In other words, only the behavior for $x \in \delta^0(q)$ is generic for fixed-precision numerical computations.

4. CODING THE DERIVATIVES

Numerical integrators for the ODE $\dot{x} = f(x)$ are usually implemented to invoke a user-supplied procedure for the evaluation of f. The analogous thing to do for the differential automaton (S, f, ν) is to have the numerical integrator invoke a user-supplied procedure to evaluate f and ν. For example, consider the system shown in Fig. 3, where we suppose that α_1, α_2, and α_3 are C^∞ functions. Here, we have $S = \mathbb{R}^2 \times Q$, where $Q = \{(i, j) | i, j = 1, 2, 3\}$. The map $f: (x_1, x_2, i, j) \mapsto (\dot{x}_1, \dot{x}_2)$ is given by

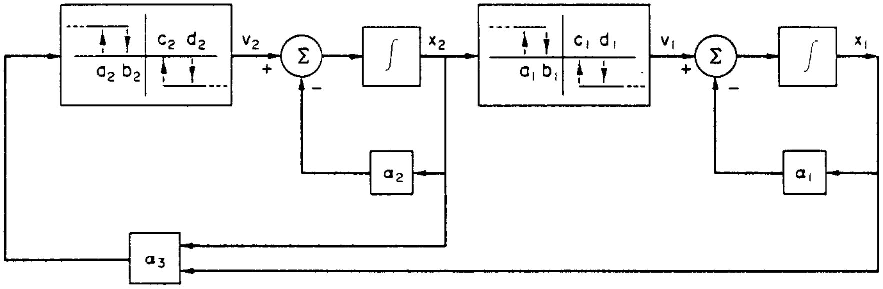

Fig. 3.

$$\dot{x}_1 = v_1 - \alpha_1(x_1),$$

$$\dot{x}_2 = v_2 - \alpha_2(x_2),$$

where v_1 and v_2 are given by

$$v_1 = \begin{cases} 1, & \text{if } j = 1, \\ 0, & \text{if } j = 2, \\ -1, & \text{if } j = 3, \end{cases}$$

$$v_2 = \begin{cases} 1, & \text{if } i = 1, \\ 0, & \text{if } i = 2, \\ -1, & \text{if } i = 3. \end{cases}$$

The map $\nu:(x_1, x_2, i, j) \mapsto (i', j')$ is given in Fig. 4. (The arrows corresponding to the cases $i' = i$ and $j' = j$ are not shown.)

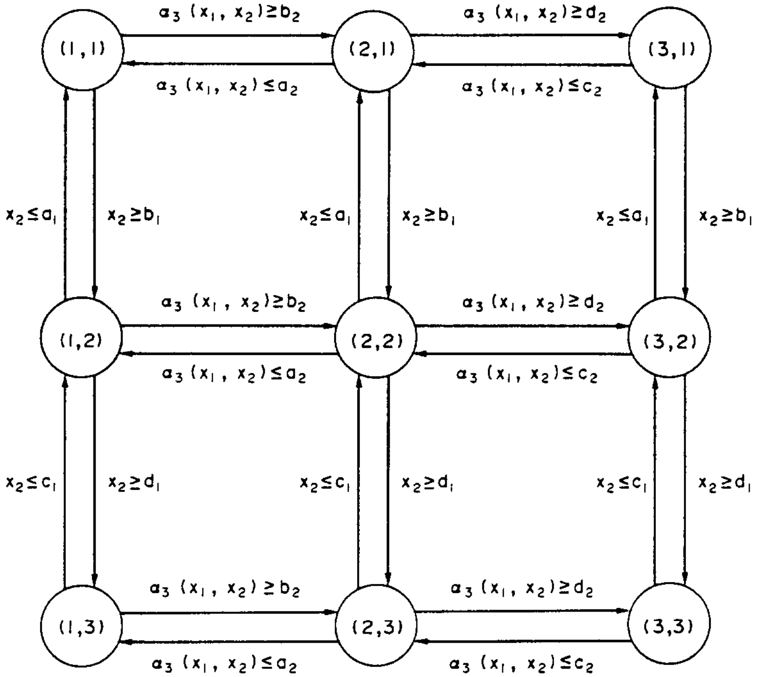

Fig. 4.

The corresponding pseudo-code is given below.

```
procedure Automaton ($x_1$, $x_2$: real; $i$, $j$: integer;
                     var $\dot{x}_1$, $\dot{x}_2$: real; var $i'$, $j'$: integer);
  var $v_1$, $v_2$, $z$: real;
  begin
    $i' := i$; $j' := j$; $z := \mathfrak{A}_3(x_1, x_2)$;
    case i of
      1: begin
        $v_2 := 1$;
        if $z \geqslant b_2$ then $i' := 2$
        end;
      2: begin
        $v_2 := 0$;
        if $z \leqslant a_2$ then $i' := 1$
        else if $z \geqslant d_2$ then $i' := 3$
        end;
  3: begin
    $v_2 := -1$;
    if $z \leqslant c_2$ then $i' := 2$
    end
  end;
case $j$ of
  1: begin
    $v_1 := 1$;
    if $x_2 \geqslant b_1$ then $j' := 2$
    end;
  2: begin
    $v_1 := 0$;
    if $x_2 \leqslant a_1$ then $j' := 1$
    else if $x_2 \geqslant d_1$ then $j' := 3$
    end;
  3: begin
    $v_1 := -1$;
    if $x_2 \leqslant c_1$ then $j' := 2$
    end
end;
  $\dot{x}_1 := v_1 - \mathfrak{A}_1(x_1)$:
  $\dot{x}_2 := v_2 - \mathfrak{A}_2(x_2)$;
  end;
```

Here $\mathfrak{A}_1$, $\mathfrak{A}_2$ and $\mathfrak{A}_3$ represent expressions which in the interpretation whose domain is the real line are assigned to the functions α_1, α_2, and α_3, respectively.

REFERENCES

1. CRYER C. W., Numerical methods for functional differential equations, in *Delay and Functional Differential Equations* (Edited by K. SCHMITT), pp. 17–101, Academic Press, New York (1972).
2. GILMORE R., *Catastrophy Theory for Scientists and Engineers*, Wiley-Interscience, New York (1981).
3. LANG S., *Real Analysis*, 2nd Edn, Addison-Wesley, Reading, MA (1983).
4. TAKENS F., Constrained equations; a study of implicit differential equations and their discontinuous solutions, in *Structural Stability, the Theory of Catastrophies, and Applications in the Sciences* (Edited by P. HILTON), *Lecture Notes in Mathematics* **525**, Springer, Berlin (1976).
5. TAVERNINI L., The approximate solution of Volterra differential systems with state-dependent time lags, *SIAM J. Numer. Analysis* **15**, 1039–1052 (1978).
6. TAVERNINI L., Linear multistep methods for the numerical solution of Volterra functional differential equations, *Applicable Analysis* **1**, 169–185 (1973).
7. TAVERNINI L., One-step methods for the numerical solution of Volterra functional differential equations, *SIAM J. Numer. Analysis* **8**, 786–795 (1971).
8. TAVERNINI L., On UNIP and the construction of digital simulation programs, *Simulation* **5**, 263–268 (1966).