# Working Paper

## Best Approximations of Control/Uncertain Differential Systems by Means of Discrete-Time Systems

*V.M. Veliov*

WP-91-45
November 1991

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

# Best Approximations of Control/Uncertain Differential Systems by Means of Discrete-Time Systems

*V.M. Veliov*

WP-91-45
November 1991



IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

# Foreword

The author studies the exact (best possible) rate of approximability of an uncertain or control system by means of $N$-stage discrete-time systems. An ultimate solution is presented in the linear case and an estimate of the rate of approximability is given for a broad class of nonlinear systems. Some applications for numerical treatment of optimal control problems and of uncertain systems are indicated.

Alexander B. Kurzhanski
Chairman,
Systems and Decision Sciences Program

# Contents

# Best Approximations of Control/Uncertain Differential Systems by Means of Discrete-Time Systems

V.M. Veliov

## 1 Introduction

The approximability of control (uncertain) systems by means of discrete-time systems is a central issue in several areas of applications like approximation of optimal control problems, discretization of the associated Hamilton-Jacobi equations, simulation or guaranteed control of uncertain systems. In the paper we address the following problem: how accurate (and in which sense) a continuous-time system can be approximated by means of $N$-stage discrete-time systems.

We consider a continuous-time system

$$\dot{x} = f(x,t,u), \quad x(t_0) = x_0, \tag{1}$$

$$u(t) \in U, \tag{2}$$

where $x \in \mathbf{R}^n$, $t \in [t_0, T]$, $u(\cdot)$ is a function with values in the compact set $U \subset \mathbf{R}^r$, $f : \mathbf{R}^n \times \mathbf{R}^1 \times \mathbf{R}^r \to \mathbf{R}^n$.

The above system can be interpreted as a differential equation containing a control (or uncertain) function $u(\cdot)$ which is allowed (resp. a priori known) to take values in $U$. A function $x(\cdot)$ is a solution of (1),(2) on $[t_0, \tau]$ if $x(\cdot)$ is absolutely continuous and satisfies (1) for some

$$u \in \mathcal{U}(t_0,\tau) = \{u(\cdot) \in L_1(t_0,\tau);\ \ u(t) \in U \text{ for a.e. } t\}.$$

The following two objects are of critical importance in the theory of control/uncertain systems:

$X$ - the set of all solutions of (1),(2) on $[t_0, T]$;

$R = \{x(T); x(\cdot) \in X\}$ - the set of all reachable points at time $T$ according to (1),(2).

A lot of work has been done for development of approximations either to the whole sets $X$ and $R$ or to particular subset or elements of these sets (for instance those, optimizing certain criterion, or satisfying some additional conditions). In this respect we refer to Kurzhanski [11], Kurzhanski and Vályi [12], Chernousko [1], Dontchev [4], and the bibliographies given there. Most of the approaches explicitly or implicitly exploit techniques for approximation of system (1),(2) by means of discrete-time control/uncertain systems. An overview on the last subject can be found in the forthcoming paper by Dontchev abd Lempio [6].

In the further lines we give a precise formulation of the problem. In order to avoid unessential technicalities, from now on we assume the following:

**Supposition 1.** There is a compact set $S \subset \mathbf{R}^n$ such that for every $u(\cdot) \in \mathcal{U}(t_0, T)$ (1) has a unique solution $x(\cdot)$ on $[t_0, T]$ and $x(t) \in S$ for $t \in [t_0, T]$.

Together with system (1),(2) consider an $N$-stage discrete-time system

$$x_{k+1} = f_N(x_k, k, u_k), \; x_0 - \text{ given in (1)}, \tag{3}$$

$$u_k \in U, \; k = 0, \ldots, N-1, \tag{4}$$

where $f_N(\cdot, k, u) : \mathbf{R}^n \to \mathbf{R}^n$. A sequence $x_0, \ldots, x_N$ that satisfies (3),(4) together with appropriate $u_0, \ldots, u_{N-1}$ is called ($N$-stage) solution of (3),(4). In order to make a link between systems (1),(2) and (3),(4) we introduce the uniform grid $t_k = t_0 + kh$, $h = (T - t_0)/N$ and associate the moment $t_k$ of the continuous time with the $k$-th stage of the discrete time. Moreover, with every sequence $u_0, \ldots, u_{N-1} \in U$ we associate the function $u[u_0, \ldots, u_{N-1}](t) = u_i$ for $t \in [t_i, t_{i+1})$.

We shall use two criteria for the quality of the approximation that (3),(4) provides to (1),(2). The first one, $\tilde{\rho}_N^x$ is connected with the set of trajectories.

$\tilde{\rho}_N^x$ is the infimum of those reals $\rho$ for which:

i) for every $u_0, \ldots, u_{N-1} \in U$ and corresponding solution $x_0, \ldots, x_N$ of (3), the solution $x(\cdot)$ of (1) corresponding to $u(\cdot) = u[u_0, \ldots, u_{N-1}](\cdot)$ satisfies

$$|x(t_k) - x_k| \leq \rho \text{ for } k = 0, \ldots, N; \tag{5}$$

ii) for every solution $x(\cdot)$ of (1),(2) there is a solution $x_0, \ldots, x_N$ of (3),(4) satisfying (5).

Similarly we define the criterion $\tilde{\rho}_N^R$ connected with the reachable set:

$\tilde{\rho}_N^R$ is the infimum of those reals $\rho$ that satisfy the above conditions i) and ii) with (5′) below instead of (5).

$$|x(t_N) - x_N| \leq \rho. \tag{5'}$$

The values of $\tilde{\rho}_N^x$ and $\tilde{\rho}_N^R$ depend on the particular discrete-time system (3),(4). Now we define

$$\rho_N^x = \inf \{\tilde{\rho}_N^x; \; f_N(\cdot,\cdot,k) \text{ runs over the functions } \mathbf{R}^n \times U \to \mathbf{R}^n, \; k = 0, \ldots, N-1\}$$

and similarly $\rho_N^R$ is the infimum of $\tilde{\rho}_N^R$ over all $N$-stage discrete-time systems of the form of (3),(4) (observe that the only relation between (1),(2) and a system of the type of (3),(4) is the presence of the same set $U$ and the same initial condition $x_0$ in the latter). The functions $N \to \rho_N^x$ and $N \to \rho_N^R$ will be called rates of $x$ (resp. R) - approximability of the system (1),(2) by means of discrete-time systems. The main goal of the paper is to investigate the approximability rates. In the single-valued case ($U = \{u\}$ - a singleton) the solution is trivial, namely $\rho_N^x = \rho_N^R = 0$ for every $N$. Actually, one can define $f_N(x, k, u)$ as the value at the moment $t_{k+1}$ of the solution of (1) starting from $x$ at $t_k$ (supposing existence and uniqueness). This definition of system (3) is not constructive from the numerical point of view, but provided that $f$ is sufficiently smooth one can use, say, a Runge-Kutta scheme in order to define constructively a discrete-time equation with a given accuracy. The situation is quite different in the multivalued case ($U$ being not a singleton). An intrinsic characterization of the rates of approximability is given in Section 2 together with some comments and examples, the aim of which is to show that the rate of $x$-approximability is typically const$/N$ and the rate of $R$-approximation is between const$/N$ and const$/N^2$.

An auxiliary result, providing a basis for the analysis of the $R$-approximability rate and being also of independent interest, is presented in Section 3. The principle results, saying that the rate of $R$-approximability of a linear system is const$/N^2$ and of a large class of nonlinear ones – not less than const$/N^{1.5}$, are stated in sections 4 and 5. Some applications are indicated in Section 6.

## 2 Characterization of the approximability rates.

The simple proposition below gives an intrinsic characterization of the approximability rates. Most of the present section will be devoted to discussions of its consequences.

Define

$$\mathcal{U}_N = \{u(\cdot) \in \mathcal{U}(t_0, T);\ u(\cdot) \text{ is constant on every } [t_i, t_{i+1}),\ i = 0, \ldots, N-1\}.$$

To this subclass of the set $\mathcal{U} = \mathcal{U}(t_0, T)$ there corresponds a set of solutions $X_N$ on $[t_0, T]$ and a reachable set $R_N = \{x(T);\ x(\cdot) \in X_N\}$. Define

$$\hat{\rho}_N^x = \sup_{x(\cdot)\in X} \inf_{\hat{x}(\cdot)\in X_N} \max_{k=0,\ldots,N} |x(t_k) - \hat{x}(t_k)|,$$

$$\hat{\rho}_N^R = \sup_{x\in R} \inf_{\hat{x}\in R_N} |x - \hat{x}|$$

($\hat{\rho}_N^R$ is the Hausdorff distance between $R$ and $R_N$, $\hat{\rho}_N^x$ is the Hausdorff "distance" between $X$ and $X_N$ corresponding to the semi-norm $\max_{k=0,\ldots,N} |x(t_k)|$ in $C[t_0, T]$).

**Proposition 1.** Let Supposition 1 hold, let $U$ be compact and let $f$ be Lipschitz continuous in $x$ in a compact neighborhood $\tilde{S}$ of $S$ (uniformly in $t \in [t_0, T]$, $u \in U$), measurable in $t$ and continuous in $u$ (uniformly in $x \in \tilde{S}$, $t \in [t_0, T]$). Then

$$\rho_N^x \le \hat{\rho}_N^x \le 2\rho_N^x, \tag{6}$$

$$\rho_N^R \le \hat{\rho}_N^R \le 2\rho_N^R \tag{7}$$

for every sufficiently large $N$.

**Proof.** The proofs of (6) and (7) are identical. Let us start with the first inequality in (6). From the suppositions it follows that there is $\delta > 0$ such that for every $x \in S$, $s \in [t_0, T)$ and $u \in U$ the equation

$$\dot{x} = f(x, t, u), \quad x(s) = x,$$

has a unique solution $x[x, s, u](\cdot)$ on $[s, \min\{s + \delta, T\}]$ and it does not abandon $\tilde{S}$. Given $N$, $k \in \{0, \ldots, N-1\}$, $x \in S$ and $u \in U$ we define $f_N(x, k, u) = x[x, t_k, u](t_{k+1})$. The definition is correct for $N > (T - t_0)/\delta$. Outside $\tilde{S}$ $f_N(\cdot, k, u)$ could be defined arbitrarily. Consider the discrete-time system (3),(4) with the so defined $f_N$. It is straightforward that point i) of the definition of $\tilde{\rho}_N^x$ is satisfied with $\rho = 0$ and point ii) – with $\rho = \hat{\rho}_N^x$. Hence $\rho_N^x \leq \hat{\rho}_N^x$.

Now, fix $N$ and take $\varepsilon > 0$. Let $x(\cdot) \in X$ be arbitrary. There is a discrete-time system (3),(4) such that the corresponding $\tilde{\rho}_N^x$ satisfy

$$\tilde{\rho}_N^x \leq \rho_N^x + \varepsilon. \tag{8}$$

By definition there is a sequence $u_0, \ldots, u_{N-1} \in U$ such that the corresponding solution $x_0, \ldots, x_N$ of (3) satisfies

$$|x(t_k) - x_k| \leq \tilde{\rho}_N^x + \varepsilon, \ \ k = 0, \ldots, N. \tag{9}$$

Let $\hat{x}(\cdot)$ be the solution of (1) corresponding to $u(\cdot) = u[u_0, \ldots, u_{N-1}](\cdot)$. Then by definition

$$|\hat{x}(t_k) - x_k| \leq \tilde{\rho}_N^x + \varepsilon. \tag{10}$$

Combining (8), (9) and (10), we obtain

$$|x(t_k) - \hat{x}(t_k)| \leq 2\rho_N^x + 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this means $\hat{\rho}_N^x \leq 2\rho_N^x$ and the proof is complete.

The proof of the first part of Proposition is nonconstructive in the numerical sense. However, it is clear that one can use some single-step discretization scheme (say, of a Runge-Kutta type) in order to define a sequence of discrete-time systems (3),(4) in terms of the values of $f$ (and possibly of its derivatives), provided that $f$ is sufficiently smooth. In contrast to the case of a single differential equation, however, the accuracy of the approximation in the sense of $\tilde{\rho}_N^x$ of $\tilde{\rho}_N^R$ cannot be better than proportional to $\hat{\rho}_N^x$ (resp. $\hat{\rho}_N^R$), independently of how smooth $f$ is.

It is quite clear that $\hat{\rho}_N^x$ and $\hat{\rho}_N^R$ are strictly positive, excluding some "degenerate" cases like a single-valued $U$. In the next lines we mention some positive and negative results that throw some light on the behaviour of $\hat{\rho}_N^x$ and $\hat{\rho}_N^R$.

Nikol'skii [13] proved the estimation $\hat{\rho}_N^R \leq$ const$/\sqrt{N}$ but short time later the results in Nikol'skii [14], Dontchev and Farkhi [5], Wolenski [22] implied that under quite general conditions $\hat{\rho}_N^x \leq$ const$/N$ (the same for $\hat{\rho}_N^R$). On the other hand, the following claim shows that this estimate for $\hat{\rho}_N^x$ cannot be improved in general.

**Claim 1.** Consider the system

$$\begin{aligned}\dot{x}^1 &= u^1, \\ \dot{x}^2 &= x^1, \\ \dot{x}^i &= f_i(x^1,\ldots,x^n,\ u^1,\ldots,u^r,t) \quad i = 3,\ldots,n,\end{aligned}$$

where $f_i$ are arbitrary, $u^1 \in [-1,1]$, $(u^2,\ldots,u^r) \in U_1$ - arbitrary, $t \in [0,1]$. Then

$$\hat{\rho}_N^x \geq 1/(24N).$$

In contrast to $\hat{\rho}_N^x$, it turns out that the sequence $\hat{\rho}_N^R$ can converge faster. The estimate

$$\hat{\rho}_N^R \leq \text{const}/N^2$$

was proven in Veliov [18] for linear systems and in Veliov [20] for a (very special) class of nonlinear systems. On the other hand, it holds

**Claim 2.** Consider the system

$$\begin{aligned}\dot{x} &= Ax + bu^1, \quad t \in [0,1], \\ \dot{y} &= f(x,y,u^1,\ldots,u^r),\end{aligned}$$

where $x \in \mathbf{R}^2$, $y \in \mathbf{R}^n$, $u^1 \in [-1,1]$, $(u^2,\ldots,u^r) \in U$ where $A, b, f$ and $U$ are arbitrary. Let rank $\{b, Ab\} = 2$. Then there is a constant $\bar{c}$ such that

$$\hat{\rho}_N^R \geq \bar{c}/N^2.$$

We shall mention that, in the case of a polyhedral set $U$, second order approximations with respect to the set of trajectories were introduced in [19], making use of an appropriate expansion of the set $U$ (and thus being outside the above framework).

# 3 An auxiliary result

We shall present a result that provides the basis for the estimates of the $R$-approximability rate given in the next sections. The essence of the result is that it estimates the variation of the derivative of the maximum of a family of functions by means of the variations of the derivatives of these functions. This is an important point in the context of the discrete approximations because, as known from Ivanov [8] (or Sendov and Popov [16]) the second order approximability of integrals by means of linear quadrature formulae is connected with boundedness of the variation of the derivative of the subintegral function (similarly, for differential equations).

We proceed with the exact formulation. Denote by $\bigvee_s^t p$ the variation of the function $p(\cdot) : [s,t] \longmapsto \mathbf{R}^n$. If $p(\cdot)$ is considered as an element of $L_\infty$ then the notion still has sense as the infimum of the variations of the functions equivalent to $p(\cdot)$.

Let $\{p(\cdot, v)\}_{v \in V}$ be a family of functions, everyone defined for a.e. $t \in [t_0, T]$, where $V$ is an arbitrary set. We say that the family is of bounded joint variation if there is a number $w$ such that

$$\sum_{i=1}^{s} |p(t_i, v_i) - p(t_{i+1}, v_i)| \leq w \tag{11}$$

for every finite collection $t_0 \leq t_1 < \ldots < t_{s+1} \leq T$, $v_1, \ldots v_s \in V$ for which $p(t_i, v_i)$ and $p(t_{i+1}, v_i)$ are defined, $i = 1, \ldots, s$. If $\{p(\cdot, v)\}_{v \in V}$ is of bounded joint variation, then the infimum of those numbers $w$ that satisfy (11) will be called joint variation of the family and will be denoted by $\mathcal{J}\bigvee_{t_0}^T p(\cdot, V)$. We shall mention that boundedness ot the joint variation of a family is a stronger property than uniform boundedness of the variations of the functions from this family.

**Proposition 2.** Let $\{p(\cdot, v) : [t_0, T] \longmapsto \mathbf{R};\ v \in V\}$ be a family of functions such that

(i) $|p(t_1, v) - p(t_2, v)| \leq L|t_1 - t_2|$ for every $t_1, t_2 \in [t_0, T]$ and $v \in V$;

(ii) the family $\{\frac{\partial p}{\partial t}(\cdot, v)\}_{v \in V}$ is of bounded joint variation.

Then the function $p(t) = \sup\{p(t, v);\ v \in V\}$ is Lipschitz continuous and

$$\bigvee_{t_0}^{T} \dot{p}(\cdot) \leq 2L + 2\mathcal{J}\bigvee_{t_0}^{T} \frac{\partial}{\partial t} p(\cdot, V).$$

**Proof.** From Clarke [2, Theorem 2.8.6] it follows that $p(\cdot)$ is Lipschitz continuous and the Clarke's generalized derivative $\partial p(t)$ satisfies

$$\partial p(t) \subset \mathrm{cl}\,\mathrm{co}\{\lim_{i \to +\infty} p'_t(t_i, v_i);\ t_i \to t,\ p(t, v_i) \to p(t)\} = [a(t), b(t)], \tag{12}$$

where in the right-hand side stands the closed convex hull of the condensation points of all sequence $p_t'(t_i, v_i))$, for which this derivative exists and $t_i \to t$, $p(t, v_i) \to p(t)$. Since the mapping $t \to \partial p(t)$ is u.s.c. [2, Proposition 2.1.5] we can extend for convenience $\dot{p}(\cdot)$ to the whole interval $[t_0, T]$ in such a way that $\bigvee_{t_0}^{T} \dot{p}(\cdot)$ does not change and $\dot{p}(t) \in [a(t), b(t)]$ for every $t \in [t_0, T]$. Actually, at a point $t$ where $\dot{p}(t)$ does not exist one can define it as an arbitrary condenzation point of a sequence $p(t_i)$ such that $t_i \to t$ and $\dot{p}(t_i)$ exists.

Take arbitrarily $t_0 \le t_1 < \ldots < t_{s+1} \le T$ and consider

$$w = \sum_{i=1}^{s} |\dot{p}(t_i) - \dot{p}(t_{i+1})|.$$

Without any restriction we can assume that $\dot{p}(t_i) - \dot{p}(t_{i+1})$ changes its sign alternatively with $i$, since otherwise one can remove some of the points $\{t_i\}$ ensuring this property without changing $w$. Thus

$$w = |\sum_{i=1}^{s} (-1)^i (\dot{p}(t_i) - \dot{p}(t_{i+1}))|. \tag{13}$$

Since $\dot{p} \in [a(t), b(t)]$ for every $t \in [t_0, T]$ we have

$$a(t_i) - b(t_{i+1} \le \dot{p}(t_i) - \dot{p}(t_{i+1}) \le b(t_i) - a(t_{i+1}).$$

Hence, using (13) we estimate

$$w \le \sum_{i=1}^{s} |c_i - c_{i+1}|,$$

where $c_i$ is either $a(t_i)$ or $b(t_i), i = 1, \ldots, s+1$.

Denote

$$\delta_1 = 0.5(t_2 - t_1),\ \delta_{s+1} = 0.5(t_{s+1} - t_s),$$

$$\delta_i = 0.5 \min\{t_i - t_{i-1},\ t_{i+1} - t_i\},\ i = 2, \ldots s.$$

Since $c_i$ are extreme points of the right-hand side of(12), for every $\varepsilon > 0$ and each $i = 1, \ldots, s+1$ there exist $\theta_i \in [t_0, T]$ and $v_i \in V$ such that $p_t'(\theta_i, v_i)$ exists and

$$|\theta_i - t_i| \le \varepsilon \delta_i^2, \tag{14}$$

$$p(t_i, v_i) \ge p(t_i) - \varepsilon \delta_i^2, \tag{15}$$

$$|p_t'(\theta_i, v_i) - c_i| \le \varepsilon \delta_i. \tag{16}$$

Using (16) we obtain

$$w \;\le\; \sum_{i=1}^{s} |p_t'(\theta_i, v_i) - p_t'(\theta_{i+1}, v_{i+1})| + 2\varepsilon \sum_{i=1}^{s+1} \delta_i \;\le\; w_1 + 2\varepsilon(T - t_0). \qquad (17)$$

Consider the function

$$\psi(t) = -p(t, v_i) + p(t, v_{i+1}) \text{for} t \in [t_i, t_{i+1}].$$

From (15) we have

$$\psi(t_i) \le -p(t_i) + \varepsilon\delta_i^2 + p(t_i) = \varepsilon\delta_i^2,$$

$$\psi(t_{i+1}) \ge -p(t_{i+1}) + p(t_{i+1}) - \varepsilon\delta_{i+1}^2 = -\varepsilon\delta_{i+1}^2.$$

Hence

$$-\varepsilon(\delta_i^2 + \delta_{i+1}^2) \le \int_{t_i}^{t_{i+1}} \dot\psi(t)\, dt \le 4\varepsilon\delta_i^2 L + \int_{t_i+\varepsilon\delta_i^2}^{t_{i+1}-\varepsilon\delta_i^2} \dot\psi(t)\, dt$$

$$\le 4\varepsilon\delta_i^2 L + (t_{i+1} - t_i - 2\varepsilon\delta_i^2)\,\text{essup}\,\{\dot\psi(t);\; t \in [t_i + \varepsilon\delta_i^2, t_{i+1} - \varepsilon\delta_i^2]\}$$

and there is $\tau_i \in (t_i + \varepsilon\delta_i^2, t_{i+1} - \varepsilon\delta_i^2)$ at which $\dot\psi(\tau_i)$ exists and

$$\dot\psi(\tau_i) \ge -\frac{2\varepsilon(\delta_i^2 + \delta_{i+1}^2 + 4\delta_i^2 L)}{t_{i+1} - t_i} \ge -2\varepsilon(\delta_i + \delta_{i+1} + 4L\delta_i).$$

Thus we can estimate

$$w_1 \le \sum_{i=1}^{s} |\dot\psi(\tau_i)| + \sum_{i=1}^{s} |p_t'(\theta_i, v_i) - p_t'(\tau_i, v_i)| + \sum_{i=1}^{s} |p_t'(\theta_{i+1}, v_{i+1}) - p_t'(\tau_i, v_{i+1})|$$

and since $\theta_i \le \tau_i \le \theta_{i+1}$ (see (14))we have

$$w_1 \le \sum_{i=1}^{s} \dot\psi(\tau_i) + \varepsilon(T - t_0)(4 + 8L) + \mathcal{J}\bigvee_{t_0}^{T} p_t'(\cdot, V) \qquad (18)$$

we can express

$$\sum_{i=1}^{s} \dot\psi(\tau_i) = \sum_{i=1}^{s-1} (p_t'(\tau_i, v_{i+1}) - p_t'(\tau_{i+1}, v_{i+1})) + p_t'(\tau_s, v_{s+1}) - p_t'(\tau_1, v_1)$$

$$\le \mathcal{J}\bigvee_{t_0}^{T} p_t'(\cdot, V) + 2L.$$

The last inequality combined with (17) and (18) implies the claim of the proposition, since $\varepsilon > 0$ is arbitrary. Q.E.D.

We shall apply the above proposition for $p(t,v) = \langle r(t), v\rangle, v \in U$ where $r(\cdot)$ : $[t_0, T] \longmapsto \mathbf{R}^r$, $U \subset \mathbf{R}^r$. Then $p(t) = sup_{v\in U}\langle r(t), v\rangle = \rho(r(t)|U)$ is the support function of $U$ in the direction $r(t)$.

**Corollary 1.** Let $r(\cdot)$ be Lipschitz continuous with Lipschitz constant $L_r$ and let $\dot{r}(\cdot)$ be of bounded variation. Let $U$ be convex and compact. Then

$$\bigvee_{t_0}^{T} \frac{d}{dt}\rho(r(t)|U) \le 2|U|\,(L_r + 2\bigvee_{t_0}^{T} \dot{r}),$$

where $|U| = \max\{|u|;\ u \in U\}$.

The claim follows directly from Proposition 2, since $\langle r(\cdot), v\rangle$ is Lipschitz continuous with Lipschitz constant $L_r|U|$ and

$$\mathcal{J}\bigvee_{t_0}^{T} \le |U|\bigvee_{t_0}^{T} \dot{r}.$$

# 4 The linear case

Consider a linear system

$$\dot{x} = A(t)x + B(t)u, \quad u \in U \subset \mathbf{R}^r, \quad t \in [t_0, T]. \tag{19}$$

Suppose that $A$ is of bounded variation, $B$ is Lipschitz continuous, $\dot{B}$ is of bounded variation, $U$ is convex and compact.

As mentioned in Section 2, for every solution $x(\cdot)$ of (19), there is $u_N(\cdot) \in \mathcal{U}_N$ such that the corresponding solution $x_N(\cdot)$ satisfies

$$|x_N(t_k) - x(t_k)| \le \text{const}/N, \quad k = 0, \dots, N.$$

On the other hand, from Veliov [18] if follows that there is $\tilde{u}_N(\cdot) \in \mathcal{U}_N$ such that the corresponding $\tilde{x}_N(\cdot)$ satisfies

$$|\tilde{x}_N(T) - x(T)| \le \text{const}/N^2.$$

The next theorem proves that one can assume $x_N(\cdot) = \tilde{x}_N(\cdot)$ in the above two claims.

**Theorem 1.** There are constants $C_1$ and $C_2$ such that for every solution $x(\cdot)$ of (19) on $[t_0, T]$ corresponding to some $u(\cdot) \in \mathcal{U}$ and for every $N$ there exists a trajectory $x_N(\cdot)$ of (19) corresponding to some $u_N(\cdot) \in \mathcal{U}_N$ such that

$$\begin{aligned} \| x(\cdot) &- x_N(\cdot) \|_C \leq C_1(T-t_0)/N, \\ |x(T) &- x_N(T)| \leq C_2(T-t_0)^2/N^2. \end{aligned}$$

As in Section 3 we denote by $\bigvee_s^t q$ the variation of a function $q$ on $[s,t]$.

In the proof of Theorem 1 we shall use the following lemmas.

**Lemma 1.** Let $g(t,x)$ be defined on $[t_0, T] \times \mathbf{R}^n$ and be measurable in $t$ and Lipschitz continuous in $x$ with Lipschitz constant $L$ uniform in $t$. Then for every function $q$ with bounded variation and for every $\alpha \in (0, T-t_0)$ there is an absolutely continuous function $q_\alpha$ such that

$$\int_{t_0}^{T} |g(t, \alpha q_\alpha(t)) - g(t, \alpha q(t))| \, dt \leq 2\alpha^2 L \bigvee_{t_0}^{T} q$$

and

$$\bigvee_{t_0}^{T} \dot{q}_\alpha \leq \frac{2}{\alpha} \bigvee_{t_0}^{T} q.$$

**Proof.** Define

$$q_\alpha(t) = \begin{cases} \frac{1}{\alpha} \int_t^{t+\alpha} q(s) \, ds, & t \in [t_0, T-\alpha], \\ \\ \frac{1}{\alpha} \int_{T-\alpha}^{T} q(s) \, ds, & t \in [T-\alpha, T] \end{cases}$$

Obviously $q_\alpha(\cdot)$ is Lipschitz continuous and

$$\bigvee_{t_0}^{T} \dot{q}_\alpha \leq \frac{1}{\alpha} \Big( \bigvee_{t_0+\alpha}^{T} q + \bigvee_{t_0}^{T} q \Big) \leq \frac{2}{\alpha} \bigvee_{t_0}^{T} q.$$

Moreover,

$$\int_{t_0}^{T} |g(t, \alpha q_\alpha(t)) - g(t, \alpha q(t))| \, dt \leq \alpha L \int_{t_0}^{T} |q_\alpha(t) - q(t)| \, dt$$

$$\leq \alpha L \int_{t_0}^{T-\alpha} |q(t) - \frac{1}{\alpha}\int_{t}^{t+\alpha} q(s)\,ds|\,dt + \alpha L \int_{T-\alpha}^{T} |q(t) - \frac{1}{\alpha}\int_{T-\alpha}^{T} q(s)\,ds|\,dt$$

$$\leq L \int_{t_0}^{T-\alpha} \int_{t}^{t+\alpha} \bigvee_{t}^{t+\alpha} q\,ds\,dt + L \int_{T-\alpha}^{T} \int_{T-\alpha}^{T} \bigvee_{T-\alpha}^{T} q\,ds\,dt$$

$$\leq \alpha L\Big(\int_{t_0}^{T-\alpha} \bigvee_{t}^{t+\alpha} q\,dt + \int_{T-\alpha}^{T} \bigvee_{T-\alpha}^{T} q\,dt\Big) \leq \alpha L\Big(\int_{T-\alpha}^{T} \bigvee_{t_0}^{t} q\,dt + \int_{T-\alpha}^{T} \bigvee_{T-\alpha}^{T} q\,dt\Big) \leq 2\alpha^2 L \bigvee_{t_0}^{T} q.$$

Q.E.D.

**Lemma 2.** Let the function $f : \mathbf{R}^n \to \mathbf{R}$ be Lipschitz continuous with Lipschitz constant $L$ and let it have the property that for every Lipschitz continuous function $r : [t_0, T] \to \mathbf{R}^n$ with $\dot{r}$ of bounded variation, the superposition $f(r(\cdot))$ satisfies

$$\bigvee_{t_0}^{T} \frac{d}{dt} f(r(\cdot)) \leq d_1 L_r + d_2 \bigvee_{t_0}^{T} \dot{r}$$

where $L_r$ is the Lipschitz constant of $r$ and $d_1$ and $d_2$ do not depend on $r$. Let $p, q : [t_0, T] \to \mathbf{R}^n$ are such that $p$ is absolutely continuous and $\dot{p}$ and $q$ are of bounded variations. For an integer $N$ consider

$$I_N = \int_{t_0}^{T} f(p(t) + \frac{T-t_0}{N} q(t))\,dt$$

and

$$\tilde{I}_N = \frac{T-t_0}{N} \sum_{i=1}^{N} f(p(\tilde{t}_i) + \frac{T-t_0}{N} q(\tilde{t}_i)),$$

where $\tilde{t}_i = t_0 + (i-0.5)(T-t_0)/N$

Then

$$|\tilde{I}_N - \tilde{I}_N| \leq \frac{(T-t_0)^2}{N^2}(0.25 d_1 L_p + 0.25 d_2 \bigvee_{t_0}^{T} \dot{p} + (4L + 0.25(d_1 + 2d_2)) \bigvee_{t_0}^{T} q).$$

**Proof.** Let $q_N$ be the function from Lemma 1 corresponding to $g(t,x) = f(p(t) + x)$, $\alpha = (T-t_0)/N$ and the function $q$ above. Denote

$$J_N = \int_{t_0}^{T} f(p(t) + \frac{T-t_0}{N} q_N(t))\,dt,$$

$$\tilde{J}_N = \frac{T-t_0}{N} \sum_{i=1}^{N} f(p(\tilde{t}_i) + \frac{T-t_0}{N} q_N(\tilde{t}_i)).$$

Then

$$
\begin{aligned}
|\tilde{I}_N - I_N| &\leq |\tilde{J}_N - \tilde{I}_N| + |\tilde{J}_N - J_N| + |J_N - I_N| \\
&\leq \frac{(T-t_0)^2}{N^2} L \sum_{i=1}^{N} |q_N(\tilde{t}_i) - q(\tilde{t}_i)| + |\tilde{J}_N - J_N| + \frac{2L(T-t_0)^2}{N^2} \bigvee_{t_0}^{T} q \\
&\leq \frac{(T-t_0)^2}{N^2} L \Big( \sum_{i=1}^{N-1} \bigvee_{\tilde{t}_i}^{\tilde{t}_{i+1}} q + \bigvee_{T-\alpha}^{T} q \Big) + |\tilde{J}_N - J_N| + \frac{2L(T-t_0)^2}{N^2} \bigvee_{t_0}^{T} q \\
&\leq \frac{4(T-t_0)^2 L}{N^2} \bigvee_{t_0}^{T} q + |\tilde{J}_N - J_N|. \qquad (20)
\end{aligned}
$$

Here we have used the particular form of $q_N$ as defined in Lemma 1.

According to K. Ivanov [8] (see also B. Sendov and V. Popov [16]) we estimate

$$
\begin{aligned}
|\tilde{J}_N - J_N| &\leq \frac{(T-t_0)^2}{4N^2} \bigvee_{t_0}^{T} \frac{d}{dt} f(p(\cdot)) + \frac{(T-t_0)}{N} q_N(\cdot)) \\
&\leq \frac{(T-t_0)^2}{4N^2} (d_1 L_r + d_2 \bigvee_{t_0}^{T} (\dot{p}(\cdot) + \frac{(T-t_0)}{N} \dot{q}_N)) \\
&\leq \frac{(T-t_0)^2}{4N^2} (d_1 L_p + d_2 \bigvee_{t_0}^{T} \dot{p} + (d_1 + 2d_2) \bigvee_{t_0}^{T} q).
\end{aligned}
$$

The Lipschitz constant of $\frac{(T-t_0)}{N} q_N$ is estimated above by $\bigvee_{t_0}^{T} q$.

Combining this with (20) we obtain the desired result.

**Lemma 3.** Let $p(\cdot)$ and $q(\cdot)$ be as in Lemma 2. Then

$$
\Big| \int_{t_0}^{T} \rho\Big(p(t) + \frac{(T-t_0)}{N} q(t) | U\Big) dt - \frac{T-t_0}{N} \sum_{i=1}^{N} \rho\left(p(\tilde{t}_i) + \frac{T-t_0}{N} q(\tilde{t}_i) | U\right) \Big|
$$

$$
\leq \frac{(T-t_0)^2}{N^2} |U| \, (0.5 L_p + 0.5 \bigvee_{t_0}^{T} \dot{p} + 5.5 \bigvee_{t_0}^{T} q), \qquad (21)
$$

where $\rho(l|U) = \max_{u \in U} < l, u >$ is the support function, $|U| = \max\{|u|;\ u \in U\}$, $N$ is an integer and $\tilde{t}_i$ are as in Lemma 2.

**Proof.** The support function of $U$ is Lipschitz continuous with Lipschitz constant $|U|$. According to Corollary 1

$$
\bigvee_{to}^{T} \frac{d}{dt} \rho(r(t)|U) \leq 2L_r + 2|U| \bigvee_{to}^{T} \dot{r}.
$$

Thus we can apply Lemma 2 with $d_1 = 2|U|$, $d_2 = 2|U|$. This gives estimation (21).

**Proof of Theorem 1.** Consider the space $E = \mathbf{R}^n \times C^n[t_0, T]$ with the norm

$$\|(x, x(\cdot)\| = \max\left\{|x|, \frac{T-t_0}{N}\|x(\cdot)\|_C\right\},$$

where the norms in the right-hand side are taken with respect to the Euclidean norm in $\mathbf{R}^n$. Without any restriction we can suppose $x(t_0) = 0$. Denote

$$Z = \{(x(T), x(\cdot)); x(\cdot) \text{ solves (11) for some } u(\cdot) \in \mathcal{U}\},$$

$$Z_N = \{(x(T), x(\cdot)); x(\cdot) \text{ solves (11) for some } u(\cdot) \in \mathcal{U}_N\}.$$

Obviously $Z_N \subset Z \subset E$. Since both $Z_N$ and $Z$ are convex and compact in $E$, we need to estimate the difference

$$\rho = \sup_{\substack{\mu \in E^* \\ \|\mu\|_* \leq 1}} \{\sup_{z \in Z} < \mu, z > - \sup_{z_N \in Z_N} < \mu, z_N >\}, \tag{22}$$

where $E^*$ is the conjugate to $E$, $<,>$ is a duality functional and $\| \quad \|_*$ is the corresponding norm in $E^*$. One can identify $E^*$ with $\mathbf{R}^n \times (C^n[t_0, T])^*$ with

$$\| (l, \lambda) \|_* = |l| + \| \lambda \|_{(C^n)^*} \ , l \in \mathbf{R}^n, \ \lambda \in (C^n)^*,$$

$$< (x, x(\cdot)), \ (l, \lambda) > = < l, x > + \frac{T-t_0}{N} < \lambda, x(\cdot) >,$$

where $<,>$ in the right-hand side means the scalar product in $\mathbf{R}^n$ and the duality functional in $C^n \times (C^n)^*$, respectively. Taking into account the Riesz representation of the elements of $(C^n)^*$ we can reduce the estimation of (22) to estimation of

$$\sup_{|l| + \bigvee_T^{t_0} \lambda \leq 1} \left\{ \sup_{(x, x(\cdot)) \in Z} (< l, x > + \frac{T-t_0}{N} \int_{t_0}^{T} x(t) d\lambda(t)) - \sup_{(x_N, x_N(\cdot)) \in Z_N} (< l, x_N > + \right.$$

$$\left. + \frac{T-t_0}{N} \int_{t_0}^{T} x_N(t) d\lambda(t)) \right\}, \tag{23}$$

where $l \in \mathbf{R}^n$ and $\lambda(\cdot)$ runs over the functions $[t_0, T] \to \mathbf{R}^n$ with bounded variation.

Denote by $\phi(t, x)$ the fundamental matrix solution of (19),

$$p(t) = B^* \phi^*(T, t) l,$$

$$q(t) = \int_t^T B^*(t)\phi^*(s,t)\,d\lambda(s).$$

Then for fixed $l$ and $\lambda(\cdot)$ the first term in the braces in (23) is

$$\sup_{u(\cdot)\in\mathcal{U}} (< l, \int_{t_0}^T \phi(T,t)B(t)u(t)\,dt > + \frac{T-t_0}{N} \int_{t_0}^T \int_{t_0}^s \phi(s,t)B(t)u(t)\,dt\,d\lambda(s))$$

$$= \sup_{u(\cdot)\in\mathcal{U}} \int_{t_0}^T < p(t) + \frac{T-t_0}{N} q(t), u(t) > dt = \int_{t_0}^T \rho\Big(p(t) + \frac{T-t_0}{N} q(t)|U\Big)\,dt.$$

Similarly we express the second term in (23). Hence the expression in the braces in (23) can be estimated by

$$\int_{t_0}^T \rho\Big(p(t) + \frac{T-t_0}{N} q(t)\,\Big|U\Big)\,dt - \sum_{i=0}^{N-1} \rho\Big(\int_{t_i}^{t_{i+1}} (p(t) + \frac{T-t_0}{N} q(t))\,dt\Big|U\Big)$$

$$\le \int_{t_0}^T \rho\Big(p(t) + \frac{T-t_0}{N} q(t)\Big|U\Big)\,dt - \frac{T-t_0}{N} \sum_{i=0}^{N-1} \rho\Big(p(\tilde{t}_i) + \frac{T-t_0}{N} q(\tilde{t}_i)\Big|U\Big)$$

$$+ \sum_{i=0}^{N-1} |U| \Big(\frac{1}{4} \bigvee_{t_i}^{t_{i+1}} \dot{p} + \frac{1}{2} \bigvee_{t_i}^{t_{i+1}} q\Big) \frac{(T-t_0)^2}{N^2}.$$

Here we use that

$$\Big|\frac{T-t_0}{N} p(\tilde{t}_i) - \int_{t_i}^{t_{i+1}} p(t)dt\Big| \le \frac{(T-t_0)^2}{4N^2} \bigvee_{t_i}^{t_{i+1}} \dot{p},$$

$$\Big|\frac{T-t_0}{N} q(\tilde{t}_i) - \int_{t_i}^{t_{i+1}} q(t)\,dt\Big| \le \frac{T-t_0}{2N} \bigvee_{t_i}^{t_{i+1}} q,$$

the first of which follows again from the result in Sendov and Popov [16].

Now we can apply (21) to estimate the expression in the braces in (23) by

$$\frac{(T-t_0)^2}{N^2} |U|\,(0.5L_p + 0.5 \bigvee_{t_0}^T \dot{p} + 5.5 \bigvee_{t_0}^T q\Big) + \frac{(T-t_0)^2}{N^2} |U| \Big(\frac{1}{4} \bigvee_{t_0}^T \dot{p} + \frac{1}{2} \bigvee_{t_0}^T q\Big).$$

It remains to estimate $\bigvee_{t_0}^T \dot{p}$ and $\bigvee_{t_0}^T q$. We have

$$\bigvee_{t_0}^T \dot{p} \le \bigvee_{t_0}^T \frac{d}{dt} B(t)\phi^*(T,t)l \le a_1$$

for some constant $a_1$, since $\dot{B}$ and $A$ are of bounded variations and $|l| \leq 1$. Finally, in order to estimate $\bigvee_{t_0}^{T} q$ one can use the estimation

$$\bigvee_{t_0}^{T} \int_{\cdot}^{T} r(\cdot, s)\, d\lambda(s) \;\leq\; (\| r \|_{L_\infty} + L_r) \bigvee_{t_0}^{T} \lambda,$$

where $L_r$ is the Lipschitz constant of $r$ with respect to the first variable. The proof of the theorem is complete.

As a consequence of the above theorem, we obtain the estimation

$$\rho_N^R \;\leq\; \text{const}(T - t_0)/N^2$$

know from [18] while the claim for simultaneous approximation of any solution of (19) and its end point will be essential for the analysis of the nonlinear case presented in the next section.

We shall stress the fact that the constants $C_1$ and $C_2$ in Theorem 1 depend on $|U|$, on $[t_0, T]$, on the Lipschitz constant of $B$ and on the variations of $\dot{B}$ and $A$ and the norms of these matrices, but not on the particular set $U$ and particular matrices $A$ and $B$.

We shall mention also that the "error" in the reachable set (in Hausdorff sense), that one makes when replacing the set of all selections of $U$ on the interval $[t_k, t_{k+1}]$ with all constant selections, is (in general) proportional to $1/N^2$, whatever is the compact initial set at $t_k$. Moreover, this "error" propagates over time and contributes to the final "error" at $t_N = T$. Nevertheless, the final "error", as the above inequality shows, remains proportional to $1/N^2$. That is, the "errors" arising in the intervals $[t_k, t_{k+1}]$ do not accumulate. This nonaccumulation effect is explained in more details in [21] and (in the context of the discrete approximations) in [18].

# 5 The nonlinear case.

In this section we consider the nonlinear system (1),(2), supposing linearity in $u$. That is, we suppose that the system is in the form of

$$\dot{x} \;=\; f(x,t) \;+\; \sum_{i=1}^{r} u_i g_i(x,t), \;\; x(t_0) \;=\; x_0, \tag{24}$$

$$u = (u_1, \dots, u_r) \in U. \tag{25}$$

It was proven in [20] that $\hat{\rho}_N^R \le \mathrm{const}/N^2$, provided that $f$ and $g_i$ are sufficiently smooth and (what is quite restrictive for many applications) $r = n$, $\mathrm{rank}\,[g_1, \dots, g_n] = n$ for every $(x,t)$ and $U$ is strongly convex (like a nondegenerate ellipsoid in $\mathbf{R}_n$). The proof is not based on the nonaccumulation effect and, in fact, implies also that $\hat{\rho}_N^x \le \mathrm{const}/N^2$. Here we present a result suggesting that the nonaccumulation effect (at least partly) takes place for more general nonlinear systems. Despite that the next theorem does not give a complete solution to the problem of estimation of $\hat{\rho}_N^R$, it shows that under quite unrestrictive conditions $\hat{\rho}_N^R$ converges to zero faster than $1/N$.

**Theorem 2.** Let Supposition 1 (from the introduction) hold and let $f$ and $g_i$ be twice differentiable in $(x,t)$ and the second derivations are Lipschitz continuous in the set $S$. Suppose, in addition that

$$\frac{\partial g_i}{\partial x} g_j - \frac{\partial g_j}{\partial x} g_i = 0 \text{ for } i, j, = 1, \dots, r \tag{26}$$

and for every $x \in S$, $t \in [t_0, T]$. Then there is a constant $c$ such that

$$\hat{\rho}_N^R \le c/N^{1.5}$$

for every integer $N$. Moreover, the constant $c$ depends on $[t_0, T], x_0, U, S$ the bounds and the Lipschitz constants of $f, g_i$ and their derivatives up to second order in $S \times [t_0, T]$, but not on the particular functions $f$ and $g_i$.

**Proof.** From $t, \tau \in [t_0, T]$, $t \le \tau$ we denote by $R(t, \tau; X)$ the reachable set of (24),(25) on $[t, \tau]$ starting from the set $X$ in the class $\mathcal{U}(t, \tau)$ of measurable selections $u(\cdot)$ of $U$. Clearly, if $X \subset R(t_0, t; x_0)$ for some $t$, then $R(t, \tau; X)$ is nonempty and is contained in $S$ for every $\tau \in [t, T]$. The proof of the following lemma is standard and therefore it will be omitted.

**Lemma 4.** For every $t, \tau \in [t_0, T]$, $t < \tau$, and $X \subset R(t_0, t; x_0)$ it holds

$$H(R(t_0, \tau; x_0), \ R(t, \tau; X)) \le e^{L_x(\tau - t)} H(X, R(t_0, t; x_0)),$$

where $L_x$ is the Lipschitz constant of $f$ and $g_i$ with respect to $x \in S$ and $H$ is the Hausdorff distance between sets.

The scheme of the proof of Theorem 2 is the following. Take an arbitrary integer $N$ and let $M$ be the largest integer such that $M^2 \le N$. Split the internal $[t_0, T]$ into $M$

equal subintervals by the points $t_i = t_0 + ih$, where $h = (T - t_0)/M$. Suppose that we have known a function $\varphi(\Delta, M)$ such that for every $t \in [t_0, T)$, and $X \subset R(t_0, t; x_0)$ it holds

$$H\Big(R_M(t, t+\Delta; X),\ R(t, t+\Delta; X)\Big) \ \le\ \varphi(\Delta, M), \tag{27}$$

where $R_M(t, t+\Delta; X)$ is the reachable set on $[t, t+\Delta]$ starting from $X$ and corresponding to the set $\mathcal{U}_M(t, t+\Delta)$ of selections $u(\cdot)$ that are constant on every subinterval $[t + k\Delta/M,\ t + (k+1)\Delta/M)$, $k = 0, \dots, M-1$. Then it is also standard to prove that

$$H\Big(R_{M \cdot M}(t_0, T; x_0),\ R(t_0, T; x_0)\Big) \ \le\ \varphi\Big(\frac{T - t_0}{M}, M\Big)\ \frac{M e^{L_x (T - t_0)}}{L_x (T - t_0)} \tag{28}$$

and clearly (28) holds also for $R_N$ instead of $R_{M \cdot M}$ in the left-hand side.

The main part of the proof will consist of the following proposition.

**Proposition 3.** Under the conditions of Theorem 2, the inequality (27) holds with

$$\varphi(\Delta, M) \ =\ c_3(\Delta^4 \ +\ \frac{\Delta^3}{M} \ +\ \frac{\Delta^2}{M^2}), \tag{29}$$

where $c_3$ is independent of $\Delta$ and $M$ and, moreover, is as claimed in the end of Theorem 2.

Using this proposition and (28) one obtains

$$\begin{aligned} H(R_N, R) \ &\le\ c_3\Big(\frac{(T - t_0)^3}{M^3} \ +\ \frac{(T - t_0)^2}{M^3} \ +\ \frac{T - t_0}{M^3}\Big)\ \frac{e^{L_x (T - t_0)}}{L_x} \\ &\le\ c_4 (T - t_0)/M^3 \ \le\ 8 c_4 (T - t_0)/N^{1.5} \end{aligned}$$

for $N \ge 4$. The theorem is proved.

It remains to prove Proposition 2, where we essentially use the following auxiliary result.

**Lemma 5.** Let $\varphi \in L_\infty(\theta, \tau)$ and

$$\Big| \int_\theta^s \varphi(t) dt \Big| \ \le\ \delta \quad \text{for every } s \in [\theta, \tau].$$

Let $v(0) = 0$ and $v(\cdot)$ be Lipschitz continuous with constant $L$. Then

$$\Big| \int_\theta^\tau \varphi(t) v(t) dt \Big| \ \le\ 2 L \delta (\tau - \theta).$$

**Proof.** Take an integer $M$ and denote $\theta_i = \theta + i(\tau - \theta)/M$. Then

$$\left|\int_{\theta_i}^{\tau} \varphi(t)dt\right| \leq \left|\int_{\theta}^{\tau} \varphi(t)dt\right| + \left|\int_{\theta}^{\theta_i} \varphi(t)dt\right| \leq 2\delta.$$

Define the functions $v_i(\cdot) : [\theta_{i-1}, \tau] \to \mathbf{R}^n$, $i = 1, \ldots, M$, as

$$v_i(t) = \begin{cases} v(t), \ t \in [\theta_{i-1}, \theta_i] \\ v(\theta_i), \ t \in (\theta_i, \tau] \end{cases}$$

and set $v_0(t) \equiv 0$. Then

$$\int_{\theta}^{\tau} \varphi(t)v(t)dt = \sum_{i=1}^{M} \int_{\theta_{i-1}}^{\tau} \varphi(t)(v_i(t) - v_{i-1}(t))dt. \tag{30}$$

Actually, we shall prove by induction in $k$ that

$$\int_{\theta}^{\theta_k} \varphi(t)v(t)dt = \sum_{i=1}^{k} \int_{\theta_{i-1}}^{\theta_k} \varphi(t)(v_i(t) - v_{i-1}(t))\, dt. \tag{31}$$

This is apparently true for $k = 1$. Suppose that (23) holds and consider

$$\int_{\theta}^{\theta_{k+1}} \varphi(t)v(t)dt = \sum_{i=1}^{k} \int_{\theta_{i-1}}^{\theta_k} \varphi(t)v(t) + \int_{\theta_k}^{\theta_{k+1}} \varphi(t)v(t)dt$$

$$= \sum_{i=1}^{k} \int_{\theta_{i-1}}^{\theta_k} \varphi(t)(v_i(t) - v_{i-1}(t))dt + \int_{\theta_k}^{\theta_{k+1}} \varphi(t)v(t)dt$$

$$= \sum_{i=1}^{k} \int_{\theta_{i-1}}^{\theta_{k+1}} \varphi(t)(v_i(t) - v_{i-1}(t))dt - \sum_{i=1}^{k} \int_{\theta_k}^{\theta_{k+1}} \varphi(t)(v_i(t) - v_{i-1}(t))dt$$

$$+ \int_{\theta_k}^{\theta_{k+1}} \varphi(t)v(t)dt = \sum_{i=1}^{k} \int_{\theta_{i-1}}^{\theta_{k+1}} \varphi(t)(v_i(t) - v_{i-1}(t))dt$$

$$- \int_{\theta_k}^{\theta_{k+1}} \varphi(t)v_k(t)dt + \int_{\theta_k}^{\theta_{k+1}} \varphi(t)v(t)dt$$

$$= \sum_{i=1}^{k} \int_{\theta_{i-1}}^{\theta_{k+1}} \varphi(t)(v_i(t) - v_{i-1}(t))dt + \int_{\theta_k}^{\theta_{k+1}} \varphi(t)(v_{k+1}(t) - v_k(t))dt$$

$$= \sum_{i=1}^{k+1} \int_{\theta_{i-1}}^{\theta_{k+1}} \varphi(t)(v_i(t) - v_{i-1}(t))dt.$$

Thus (30) is proved. Hence, denoting $c = \| \varphi \|_{L_\infty(\theta,\tau)}$, we have

$$\left| \int_{\theta}^{\tau} \varphi(t)v(t)dt \right| \leq \sum_{i=1}^{M} \left| \int_{\theta_{i-1}}^{\tau} \varphi(t)(v_i(t) - v_{i-1}(t))dt \right|$$

$$\leq \sum_{i=1}^{M} \left| \int_{\theta_{i-1}}^{\theta_i} \varphi(t)(v_i(t) - v_{i-1}(t))dt \right| + \sum_{i=1}^{M} \left| \int_{\theta_i}^{\tau} \varphi(t)(v_i(t) - v_{i-1}(t))dt \right|$$

$$\leq \sum_{i=1}^{M} \int_{\theta_{i-1}}^{\theta_i} CL(t-\theta_{i-1})dt + \sum_{i=1}^{M} \left| (v(\theta_i) - v(\theta_{i-1})) \int_{\theta_i}^{\tau} \varphi(t)dt \right|$$

$$\leq \sum_{i=1}^{M} \frac{CL}{2} (\theta_i - \theta_{i-1})^2 + \sum_{i=1}^{M} 2L(\theta_i - \theta_{i-1})\, \delta$$

$$\leq \sum_{i=1}^{M} \frac{CL}{2} (\theta_i - \theta_{i-1}) \frac{\tau-\theta}{M} + \sum_{i=1}^{M} 2L(\theta_i - \theta_{i-1})\, \delta = \frac{CL}{2}\frac{(\tau-\theta)^2}{M} + 2L\delta(\tau-\theta).$$

Since $M$ is arbitrary and the left-hand side does not depend on $M$ we get the desired inequality. The lemma is proved.

**Corollary 2.** Let $u$ and $\hat{u}$ be measurable bounded functions defined on $[\theta, \theta+h]$. Let $v(\cdot)$ be a Lipschitz continuous function with Lipschitz constant $Lh$ in $[\theta, \theta+h]$ and let

$$\left| \int_{\theta}^{s} u(t)dt - \int_{\theta}^{s} \hat{u}(t)dt \right| \leq ch \quad \text{for every } s \in [\theta, \theta+h].$$

Then

$$\left| \int_{\theta}^{\theta+h} u(t)v(t)dt - \int_{\theta}^{\theta+h} \hat{u}(t)v(t)dt \right| \leq 2cLh^3.$$

Now, let us prove Proposition 3. Let $t, \Delta$ and $X$ be as in the formulation. Then according to Supposition 1 every solution of (24) starting from $X$ at the moment $t$ exists up to the moment $T$ and takes values in S.

Given an arbitrary $u(\cdot) \in \mathcal{U}(t, t+\Delta)$ and a point $x = x(t) \in R(t_0, t; x_0)$ the corresponding solution $x(\cdot)$ of (24) can be presented as

$$x(t+\Delta) = x(t) + a_1\Delta + a_2\Delta^2 + a_3\Delta^3$$

$$+\sum_{i=1}^{r}\int_{t}^{t+\Delta}(b_0+b_1(s-t)+b_2(s-t)^2)u_i(s)ds \;+\; \sum_{i,j=1}^{r} d_{ij}\int_{t}^{t+\Delta}u_i(s)ds\int_{t}^{t+\Delta}u_j(s)ds$$

$$+\sum_{i=1}^{r}\int_{t}^{t+\Delta}u_i(s)\Psi_i(s)ds \;+\; \int_{t}^{t+\Delta}\Psi_0(s)ds \;+\; \zeta(\Delta), \tag{32}$$

where $|\zeta(\Delta)| \leq c_5\Delta^4$ and $c_5$ is as in the last claim of Theorem 2, $a_i, b_j, d_{ij}$ are vectors that can be expressed by means of $f, g_i$ and their derivatives up to second order at the point $x(t)$ and, finally, $\Psi_i(\cdot)$, $i = 0,\dots,r$ are functions that are linear combinations with vector coefficients like $a_i, b_j$ and $d_{ij}$ of integrals in the form of

$$\int_{t}^{s}\alpha(\theta)\int_{t}^{\theta}\beta(\tau)d\tau d\theta \;\text{ or }\; \int_{t}^{s}\alpha(\tau)d\tau\int_{t}^{s}\beta(\tau)d\tau, \tag{33}$$

where $\alpha(\cdot)$ and $\beta(\cdot)$ are some of the functions $u_j(\cdot)$, $j = 1,\dots,r$ or the constant 1. The above representation can be obtained by expressing explicitly the members of the Volterra series, corresponding to (24), that contain up to three iterated integrals, by means of $f, g_i$ and their derivatives up to second order. Here we essentially exploit the supposition (26) ($[g_i, g_j] = 0$) because otherwise terms containing

$$\int_{t}^{t+\Delta}u_i(s)\int_{t}^{s}u_j(\tau)d\tau ds$$

would appear in (32) for $i \neq j$.

Now we shall prove that there is $\tilde{u}(\cdot) \in \mathcal{U}_M(t, t+\Delta)$ such that

$$\left|\int_{t}^{t+\Delta}p(s)(u(s) \;-\; \tilde{u}(s))ds\right| \;\leq\; c_6\Delta^2/M^2 \tag{34}$$

for $p(s) = 1$, $p(s) = (t-s)$ and $p(s) = (s-t)^2$ and also that

$$\left|\int_{t}^{t+\Delta}[u_i(s)\Psi_i(s) \;-\; \tilde{u}_i(s)\tilde{\Psi}_i(s)]ds \;\leq\; c_7\Delta^3/M,\; i = 0,\dots r, \tag{35}$$

where by definition $u_0(s) = \tilde{u}_0(s) = 1$. Here $\tilde{\Psi}_i$ corresponds to $\Psi_i$ according to (32) for the selection $\tilde{u}(\cdot)$ instead of $u(\cdot)$ (see (33)). Namely, we define $\tilde{u}(\cdot) \in \mathcal{U}_M(t, t+\Delta)$ to be the selection that is claimed to exist (corresponding to $u(\cdot)$) by Theorem 1 applied for the linear system

$$\begin{aligned} \dot{x}_i(s) &= u_i(s), \qquad i = 1,\dots,r,\; s \in [t, t+\Delta], \\ \dot{y}_i(s) &= (s-t)u_i()), \quad x_i(t) = y_i(t) = z_i(t) = 0, \\ \dot{z}_i(s) &= (s-t)^2u_i(t). \end{aligned}$$

Apparently $\tilde{u}(\cdot)$ satisfies (34) for $c_6 = c_2$ (from Theorem 1). Moreover, if $x_i(\cdot)$ and $\tilde{x}_i(\cdot)$ are the solutions of the above linear system, corresponding to $u(\cdot)$ and $\tilde{u}(\cdot)$ then

$$|x_i(s) - \tilde{x}_i(s)| = |\int_t^s (u_i(\tau) - \tilde{u}_i(\tau))d\tau| \leq c_2\Delta/M, \tag{36}$$

where $c_2$ comes from Theorem 1. From (36) one can estimate

$$|\int_t^s (\alpha(\theta)\int_t^\theta \beta(\tau)d\tau - \tilde{\alpha}(\theta)\int_t^\theta \tilde{\beta}(\tau)d\tau)d\theta| \leq 2|U|c_2\Delta^2/M,$$

$$|\int_t^s \alpha(\tau)d\tau \int_t^s \beta(\tau)d\tau - \int_t^s \tilde{\alpha}(\tau)d\tau \int_t^s \tilde{\beta}(\tau)d\tau| \leq 2|U|c_2\Delta^2/M.$$

Hence

$$|\Psi_i(s) - \tilde{\Psi}_i(s)| \leq c_7 2m|U|\Delta^2/M = c_8\Delta^2/M, \tag{37}$$

where $m$ is the number of integrals of the type of (33) that are included in $\Psi_i$ (depending only on $n$ and $r$) and $c_7$ is a bound of the norms of the vectors multiplying the integrals (33) in $\Psi_i(\cdot)$ (also dependent only on the bounds on $f, g_i$ and their derivatives up to second order).

On the other hand, $\Psi_i$ and $\tilde{\Psi}_i$ are Lipschitz continuous with Lipschitz constant $2c_7|U|^2\Delta$. Hence using (36) and Corollary we obtain

$$|\int_t^{t+\Delta} [u_i(s)\Psi_i(s) - \tilde{u}_i(s)\Psi_i(s)]ds| \leq 4c_2c_7|U|^2\Delta^3/M = c_3\Delta^3/M. \tag{38}$$

Combining (37) and (38) one immediately obtains (35). (34) and (35) together with (32) give that the trajectory $\tilde{x}(\cdot)$ of (24) corresponding to $\tilde{u}(\cdot) \in \mathcal{U}_M(t, t+\Delta t)$ satisfies

$$|\tilde{x}(t+\Delta) - x(t+\Delta)| \leq c_{10}(\Delta^4 + \frac{\Delta^3}{M} + \frac{\Delta^2}{M^2})$$

and this proves the proposition and Theorem 2.

# 6 Some Applications

In this section we present in details an application of Theorem 2 for discretization of optimal control problems and then we only indicate some other application of the previous results.

6.1.Consider the optimal control problem

$$\min\{p(x(T)) + \int_{t_0}^{T} (q_1(x(t),t) + \varphi(u(t))q_2(x(t),t))ds\} \tag{39}$$

subject to

$$\dot{x} = f(x,t) + G(x,t)u, \quad x(t_0) = x_0 \in \mathbf{R}^n, \tag{40}$$

$$u \in U \subset \mathbf{R}^r, \tag{41}$$

where $G = (g_1, \ldots, g_r)$, $f, g_i : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}^n$.

Let Supposition 1 be fulfilled together with the following conditions:

i) $f, g_i, q_1, q_2$ are twice differentiable and all second derivatives are Lipschitz continuous in $S \times [t_0, T]$;

ii) $p$ is Lipschitz continuous in $S \times [t_0, T]$, $q_2$ is nonnegative, $U$ is convex and compact, $\varphi$ is convex and bounded on $U$;

iii) $[g_i, g_j] = 0$ on $S \times [t_0, T], i, j = 1, \ldots, 5$.

In order to obtain a correct discretization of the problem (39)–(41) one can proceed in the following way.

First reformulate the problem as

$$\min\{p(x(T)) + y\}$$

subject to (40) and

$$\dot{y} = q_1(x,t) + q_2(x,t)v, \; y(t_0) = 0,$$

$$(u,v) \in \tilde{U} = \{(u,v); v \in [\varphi(u), M]\} \subset \mathbf{R}^r \times \mathbf{R}^1,$$

where $M$ is a sufficiently large number. This is an optimization problem on a reachable set. Then consider the problem in the class $\mathcal{U}_N$ of admissible control functions $(u(\cdot), v(\cdot))$. Apply a single-step discretization scheme with (at least) third order local accuracy (we use a second order Runge-Kutta scheme below). Finally, the so obtained discrete optimization problem can be reformulated back in a form corresponding to the original problem (39)–(41). The global error consists in the error of the change of $\mathcal{U}$ with $\mathcal{U}_N$ and of the error of discretization, which is locally $0(1/N^3)$ and globally

$0(1/N^2)$ since the right-hand side is twice smooth in every interval between two consequent points $t_k = t_0 + k(T - t_0)/N$, $k = 0, \ldots, N$. As a result of application of the above scheme one obtains the following result.

**Theorem 3.** Let the suppositions listed at the beginning of the section hold. Consider the following mathematical programming problem (with respect to $u_0, \ldots, u_{N-1}$, $x_1, \ldots x_N$):

$$\min\Big\{p(x_N) + \sum_{k=0}^{N-1}(q_1(x_k, t_k) + q_1(x_k + hF(x_k, t_k, u_k), t_{k+1}))$$

$$+ \sum_{k=0}^{N-1}(q_2(x_k, t_k) + q_2(x_k + hF(x_k, t_k, u_k), t_{k+1}))\varphi(u_k)\Big\} \tag{42}$$

subject to

$$x_{k+1} = x_k + 0.5h(F(x_k, t_k, u_k) + F(x_k + hF(x_k, t_k, u_k), \ t_{k+1}, u_k)), \tag{43}$$

$$u_k \in U, \quad k = 0, \ldots, N-1, \tag{44}$$

where for brevity $F(x, t, u) = f(x, t) + G(x, t)u$. Let $V$ and $V_N$ denote the minimum value of the problems (39)–(41) and (42)–(44), respectively. Then there is a constant $c$ such that

1. $$|V_N - V| \leq c/N^{1.5}, \tag{45}$$

2. if $u_0, \ldots, u_{N-1}$ is an optimal control sequence of the discrete-time problem (42)–(44) and $u(t) = u_i$ for $t \in [t_i, t_{i+1})$, $i = 0, \ldots, N-1$ then the control $u(\cdot)$ provides a value $\tilde{V}_N$ of the objective function (39)–(41) that satisfies

$$|\tilde{V}_N - V_N| \leq c/N^{1.5}. \tag{46}$$

**Remark.** If the functions $q_1$ and $f$ are linear in $x$ and $q_2$ and $G$ are independent of $x$ then Theorem 1 is directly applicable. Under the corresponding suppositions the exponent 1.5 in (45)–(46) can be replaced with 2 (see also [18]).

According to Theorem 2, the constant $c$ in (45),(46) depends on $S$, $|U|$, $\max\{\varphi(u); u \in U\}$ and the Lipschitz constants and the bounds of $p, q_1, q_2, f, G$ and the derivatives up to second order of the last four functions. The situation is similar in the linear case as well (see the last sentence of Section 4). This means that the existence of the constant $c$ and its value are not connected with the properties of the optimal control (like Riemann integrability, boundedness of the variation or Lipschitz continuity). The results in Silin [17] show that even in the case of a time-invariant linear system with objective functions of the type of

$$|x(1) - \bar{x}|^2 \rightarrow \min$$

($\bar{x}$- given) the optimal control can be nonintegrable in the Riemann sense (and this is in a sense generic property when $n \geq 3$) but nevertheless the above remark is applicable.

**6.2.** Consider the optimal control problem

$$\min g(x(T)) \quad ,$$
$$\dot{x} = f(x,t,u), \quad x(t_0) = x_0, \quad t \in [t_0, T],$$
$$u \in U.$$

Let $V(t_0, x_0)$ be the optimal value as a function of the initial time and the initial state. Then $V$ is a viscosity solution of the Hamilton-Jacobi-Bellmann equation

$$\frac{\partial}{\partial t} V(t,x) + \inf_{u \in U} < \frac{\partial}{\partial x} V(t,x), \; f(x,t,u) > = 0 \tag{47}$$

(see Grandall and Lions [7]).

In spite of the fact that $V$ does not satisfy (47) in the classical sense (since $V$ is nondifferentiable, in general) it is known to admit approximations by means of functions satisfying appropriate discretized (in the time) versions of (47) (Dolcetta and Ishii [3]).

Such a discrete version of (47) reflects from the Belmann equation associated with a discrete-time optimal control problem that approximates the above one in the sense of the optimal value. Thus one can obtain a discretization of (47) using Theorem 3. If the suppositions are fulfilled the accuracy of the solution will be proportional to the step-length to degree 1.5. Details concerning the linear case are given in [18].

**6.3** Another field of possible applications of the results of sections 2 and 3 will be illustrated by the following estimation problem (see Krasowskii [10], Kurzhanski [11], Schweppe [15] for more general considerations).Consider the system

$$\dot{x} = A(t)x + B(t)u, \quad x(t_0) = 0, \tag{48}$$

where $u$ is an uncertain parameter, the only knowledge about which is that it takes values in a given set $U \subset \mathbf{R}^r$. If no observation is available, then the reachable set $R(t_0, T; x_0)$ is the best (i.e. the minimal in the sense of inclusion) guaranteed estimation of the state of the system at the moment $T$.

Let the suppositions of Theorem 1 be fulfilled. Given $N$ and $t_k$, $k = 0, \ldots, N$ as above, define

$$D_k = \int_{t_k}^{t_{k+1}} \phi(T, s) B(s) ds, \quad k = 0, \ldots, N-1,$$

where $\phi(t, s)$ is the fundamental matrix solutions of (48) normalized at $t = s$. Then Theorem 1 implies that

$$H(R(t_0, T; 0), \sum_{k=0}^{N-1} D_k U) \leq \text{const}/N^2.$$

In particular, if $U$ is an ellipsoid, then $R(t_0, T; 0)$ can be constructively approximated with accuracy const$/N^2$ by a sum of $N$ ellipsoids. If $U$ is a zonotope (that is a sum of, say, $m$ segments) then $R(t_0, T, 0)$ can be constructively approximated with accuracy const$/N^2$ by a zonotope generated by $mN$ segments. A computer realization of this technique that takes into account even the effects of the computational errors is present in Kirov and Krastanov [9].

# References


1. F. Chernousko, *Estimation of the Phase State of Dynamical Systems*, Nauka, Moscow, 1988 (in Russian).

2. F. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, 1983.

3. I. Dolcetta and H. Ishii, Approximate solutions of the Bellman Equation of deterministic control theory, *Applied Mathematics and Optimization*, **11** (1984), 161–181.

4. A. Dontchev, *Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems,* Lecture Notes in Control and Information Sciences, v. **52**, Springer-Verlag, 1983.

5. A. Dontchev and E. Farkhi, Error estimates for discretized differential inclusions, *Computing,* 41 (1989), 349-358.

6. A. Dontchev and F. Lempio, Difference methods for differential inclusions: a survey, Preprint of the University of Bayreuth, 1990. (to appear elsewhere).

7. M. Grandall and P.L. Lions, Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.*, **277** (1983), 1-42.

8. K. Ivanov, New estimates of error of quadrature formulae, formulae of numerical differentiation and integration, *Anal. Mathematica,* **6** (1980), 281-303.

9. N. Kirov and M. Krastanov, Dialog system INCLIN, to appear.

10. N. Krasovskii, *Control of Dynamic Systems*, Nauka, Moscow, 1986 (in Russian).

11. A. Kurzhanski, *Control and Observation under Conditions of Uncertainty,* Nauka, Moscow, 1977 (in Russian).

12. A. Kurzhanski, and I. Vályi, Set valued solutions to control problems and their approximations, in: A. Bensoussan and J.L. Lions Eds. *Analysis and Optimization of Systems,* Lecture Notes in Control and Information Systems, v. 111, Springer-Verlag, 1988, pp. 775-785.

13. M. Nikol'skii, On an approximation of the reachable set of a control process, *Mat. Zametki*, v. 41 (1987), 71-76 (in Russian).

14. M. Nikol'skii, On a method for approximation of the reachable set of a differential inclusion, *J. Vytchisl. Mat. i Mat. Phis.,* 28 (1988), 1252-1254 (in Russian).

15. F. Schweppe, *Uncertain Dynamic Systems*, Prentice Hall (1973).

16. B. Sendov and V. Popov, *Averaged Moduli of Smoothness*, John Wiley and Sons, New York, 1988.

17. D. Silin, On variation and Riemann integrability of optimal control in linear systems. *Dokl. Akad. Nauk. SSSR* **257** (1981) 548-550 (in Russian).

18. V. Veliov, Second order discrete approximations to linear differential inclusions, to appear in *SIAM J. on Numerical Analysis.*

19. V. Veliov, Approximations to differential inclusions by discrete inclusions, *Working Paper*, IIASA, WP-89-017, 1989.

20. V. Veliov, Second order discrete approximations to strongly convex differential inclusions, *Systems & Control Letters*, **13** (1989), 263-269.

21. V. Veliov, Parametric and functional uncertainties in dynamical systems: local and global relationship, to appear.

22. P. Wolenski, The exponential formula for the reachable set of Lipschitz differential inclusion, *SIAM J. Control and Optimization*, v. **28** (1990), 1148–1161.