Proc. of the Steklov Institute of Mathematics, 2007, v. 256, p. 92–104

Approximation theorem for a nonlinear control system with sliding modes

A.V. Dmitruk [1]

We consider the question of validity of the extension of a nonlinear control system by introducing the so-called sliding modes (i.e., by convexifying the set of admissible velocities) under the presence of constraints imposed on the endpoints of trajectories. We prove that a trajectory of the extended system can be approximated by trajectories of the original system if the equality constraints of the extended system are nondegenerate in the first order. The proof is based on a nonlocal estimate for the distance to the zero set of the nonlinear operator corresponding to the extended system, and involves a specific iteration process of corrections.

1. Statement of the question

On a fixed time interval $[0,T]$, consider the control system

$$\dot{x} - f(x,u,t) = 0, \tag{1}$$

$$g(x,u,t) = 0, \tag{2}$$

$$K(x(0), x(T)) = 0, \tag{3}$$

where $x(t)$ is a state variable, $u(t)$ a control, $x \in AC^m[0,T]$ ($m-$ dimensional absolutely continuous), $u \in L^r_\infty[0,T]$ ($r-$ dimensional measurable and bounded). The trajectories of equation (1) are subjected to the mixed constraint (2) of dimension $dim\, g = q$ and to endpoint constraint (3) of dimension $dim\, K = s$. We assume that the function K of argument $p = (x_0,\, x_T) \in \mathbf{R}^{2m}$ is defined and continuously differentiable on an open set $\mathcal{P} \subset \mathbf{R}^{2m}$, and the functions f, g are defined and continuous together with their first derivatives in x, u on an open set $\mathcal{Q} \subset \mathbf{R}^{m+r+1}$.

Let $\mathcal{D} = \mathcal{D}(\mathcal{P}, \mathcal{Q})$ be the set of all pairs of functions $(x(t), u(t))$ from the space $AC^m[0,T] \times L^r_\infty[0,T]$, for each of which there exists a compactum $\Gamma \subset \mathcal{Q}$ such that $(x(t), u(t), t) \in \Gamma$ a.e. on $[0,T]$ and $(x(0), x(T)) \in \mathcal{P}$. By a solution to system (1)–(3) we will call any pair of functions $(x,u) \in \mathcal{D}$ satisfying a.e. on $[0,T]$ equalities (1), (2) and also satisfying (3).

[1]Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Optimal Control. E-mail: dmitruk@member.ams.org.

Along with system (1)–(3), we will also consider the following extended system with the so-called *sliding modes*:

$$\dot{x} - \sum_{i=1}^{N} \alpha^i(t) f(x, u^i, t) = 0, \tag{4}$$

$$g(x, u^i, t) = 0, \quad i = 1, \ldots, N, \tag{5}$$

$$\sum_{i=1}^{N} \alpha^i(t) - 1 = 0, \tag{6}$$

$$K(x(0), x(T)) = 0, \tag{7}$$

where N is a fixed natural number, all $u^i \in L^r_\infty$, $\alpha^i \in L^1_\infty$, and a.e. $\alpha^i(t) \geq 0$, $i = 1, \ldots, N$. Here, both the functions $u^i(t)$ and the "weight coefficients" $\alpha^i(t)$ play the role of controls. For the sake of brevity, we adopt the notation: $p = (x(0), x(T))$ and $\mathbf{u} = (u^1, \ldots, u^N)$, $\alpha = (\alpha^1, \ldots, \alpha^N)$, where all $u^i \in L^r_\infty$, $\alpha^i \in L_\infty$.

The concept of solution to this system is a natural modification of the concept of solution to system (1)–(3) (see below).

The vector-function $\alpha(t) = (\alpha^1(t), \ldots, \alpha^N(t)) \in L^N_\infty [0, T]$ takes its values in the simplex

$$A = \{ \alpha \in \mathbf{R}^N \mid \forall i \quad \alpha^i \geq 0, \quad \sum_{i=1}^{N} \alpha^i = 1 \}. \tag{8}$$

Denote the set of vertices of this simplex by $ex\, A$, so that $\alpha \in ex\, A$ means that α is a base vector e^i of the space $\mathbf{R}^N$ for some i.

By the Caratheodory theorem, if $N \geq n$, then the set of velocities in this system at any point (x, t) is the convexification of the set of velocities in system (1)–(3) (the convexity of the velocity set plays an important role, e.g., for the existence of solution in optimization problems for such systems), but for our purposes the condition $N \geq n$ is inessential, and we do not assume it.

It is clear that any trajectory $x(t)$ of system (1)–(3) generated by a control $u(t)$ can also be considered as a trajectory of system (4)–(7) generated by a tuple of controls with $u^1(t) = u(t)$, arbitrary $u^2(t)$, $\ldots$, $u^N(t)$, and the weight coefficients $\alpha(t) = (1, 0, \ldots, 0)$. In this sense, the set of solutions to system (1)–(3) is naturally embedded in the set of solutions to system (4)–(7). Of course, the reverse embedding does not hold: an arbitrary trajectory $x(t)$ of system (4)–(7) is not, in general, a trajectory of system (1)–(3).

So, the question arises: when the passage to the extended system is justified, i.e., when a solution to system (4)–(7) can, in a sense, be approximated by trajectories of the original system (1)–(3)?

A natural attempt to resolve this question is as follows. Consider first the case when the equalities (2), and respectively (5), are absent. Let be given a solution to system (4)–(7). Fix the controls $u^1(t),\, \ldots,\, u^N(t)$, and consider a sequence of tuples of weight coefficients $\alpha_n(t) \in ex\,A$ weakly-* converging to the given tuple of coefficients $\alpha(t) \in A$, i.e., $\alpha^i_n(t) \xrightarrow{\text{weak-*}} \alpha^i(t) \quad \forall\, i = 1,\ldots,N.$ (Such a sequence always exists and can be easily constructed explicitly.) Set $x_n(0) = x(0)$. Then, since equation (4) is linear with respect to α, the sequence of solutions $x_n(t)$ to this equation converges to the given solution $x(t)$ uniformly on the interval $[0,T]$. If the function K does not depend on $x(T)$ (i.e., the endpoint $x(T)$ is free), then, setting $u_n(t) = \sum_{i=1}^N \alpha^i_n(t)\, u^i_n(t)$, we get a sequence of pairs $(x_n(t), u_n(t))$ satisfying the system (1)–(3) (in engineering applications such sequences are called sliding modes), and so, the above question is successfully resolved. In the case when equalities (2) and (5) are present, and the derivative g'_u is nondegenerate along the given trajectory, then, using the implicit function theorem, some components of the control can be expressed in terms of the state variables and the remaining free control components, and so, these equalities can be excluded and the situation can be reduced to the above case. The described approach was actually used by N.N. Bogolyubov (see [4]), L. Young [2], E.J. McShane [3]; after the famous paper of R.V. Gamkrelidze [1] it became a standard tool in the control theory and has been repeatedly used by many authors (see e.g., [5] – [12]).

However, if the function K does depend on both endpoints of the trajectory (e.g., the endpoints are fixed: $x(0) = a,\; x(T) = b$), then the above-constructed trajectories have, in general, $K(x_n(0), x_n(T)) \neq 0$, so, the pairs (x_n, u_n) do not satisfy system (1)–(3), and hence, this approach does not work; it needs a modification.

2. Preliminary facts and the main result

1. We propose the following approach. Consider the constraints (4)–(7) as the zero set of an operator F from the Banach space $W = AC^m \times (L^r_\infty)^N \times L^N_\infty\,[0,T]$ with elements $w = (x, \mathbf{u}, \alpha)$ and the norm $||w|| = ||x||_{AC} + ||\mathbf{u}||_\infty + ||\alpha||_\infty$, where $||x||_{AC} = |x(0)| + ||\dot{x}||_1$, to the space $Z = L^m_1 \times (L^q_\infty)^N \times L_\infty\,[0,T] \times \mathbf{R}^s$ with elements $z = (\xi, \eta, \nu, \kappa)$, where $\eta = (\eta^1, \ldots, \eta^N)$ with all $\eta^i \in L^q_\infty$, and the norm $||z|| = ||\xi||_1 + ||\eta||_\infty + ||\nu||_\infty + |\kappa|$.

In the space W, define a set $\mathcal{D}_N = \mathcal{D}_N(\mathcal{P}, \mathcal{Q})$ consisting of all triples $(x, \mathbf{u}, \alpha)$ such that $\forall\, i = 1,\ldots,N$ the pair $(x, u^i) \in \mathcal{D}(\mathcal{P}, \mathcal{Q})$. By a solution to system (4)–(7) we will call any triple of functions $(x, \mathbf{u}, \alpha) \in \mathcal{D}_N$ that satisfies equalities (4)–(6) almost everywhere on $[0,T]$ and also satisfies (7).

The operator F is given by the left hand parts of equalities (4)–(7). Obviously, it is Frechet differentiable at any point $(x, \mathbf{u}, \alpha) \in \mathcal{D}_N$, and its derivative

$$F'(x, \mathbf{u}, \alpha) = \; G[x, \mathbf{u}, \alpha]: \;\; W \longrightarrow Z$$

maps as follows: $(\bar{x}, \bar{\mathbf{u}}, \bar{\alpha}) \mapsto (\bar{\xi}, \bar{\eta}, \bar{\kappa}, \bar{\nu})$, where

$$\dot{\bar{x}} - \sum \alpha^i f'_x(x, u^i, t)\, \bar{x} - \sum \alpha^i f'_u(x, u^i, t)\, \bar{u}^i - \sum \bar{\alpha}^i f(x, u^i, t) \;=\; \bar{\xi},$$

$$
\begin{aligned}
g'_x(x,u^i,t)\,\bar x + g'_u(x,u^i,t)\,\bar u^i \; &= \bar\eta^i, \qquad i=1,\dots,N,\\
\sum \bar\alpha^i(t) \;\; &= \;\; \bar\nu.\\
K'_{x(0)}(p)\,\bar x(0) + K'_{x(T)}(p)\,\bar x(T) \; &= \bar\kappa,
\end{aligned}
$$

Clearly, the linear operator $G[x,\mathbf{u},\alpha]$ depends continuously on the point $(x,\mathbf{u},\alpha)$ with respect to the operator norm, i.e., F is continuously differentiable on $\mathcal{D}_N$.

Let $\mathcal{M}$ be the zero set of the operator $F,$ i.e., the set of all triples $(x,\mathbf{u},\alpha)\in W$ satisfying equalities (4)–(7). The basic fact, upon which we will rely, is the following estimate for the distance to the set $\mathcal{M}$.

Theorem 1. *Let a triple $w_0=(x_0,\mathbf{u}_0,\alpha_0)\in\mathcal{M}$ be such that the operator $F'(w_0)$ maps "onto". Let $\mathcal{A}\subset L^N_\infty[0,T]$ be an arbitrary bounded set. Then there exists numbers $\varepsilon>0,$ B and weak-* neighborhood $\mathcal{V}(\alpha_0)$ such that, for any triple $w=(x,\mathbf{u},\alpha)$ satisfying the conditions*

$$
||x-x_0||_C<\varepsilon, \qquad ||\mathbf{u}-\mathbf{u}_0||_\infty<\varepsilon, \tag{8}
$$

$$
\alpha\in\mathcal{V}(\alpha_0)\cap\mathcal{A}, \tag{9}
$$

the following estimate holds:

$$
\operatorname{dist}(w,\mathcal{M}) \;\le\; B\,||F(w)||\,. \tag{10}
$$

Note that this estimate is of nonlocal character: α is taken not from an ordinary neighborhood of α_0 , but from a broader weak-* neighborhood. The proof of Theorem 1 is given by the author in [19], [20], and [21, Ch. 5]; it is based on an abstract generalization of the classical Lyusternik theorem proposed by A.A. Milyutin (see [16, 20, 21]), which is also of intrinsic interest as a fact of functional analysis, being a quite efficient tool in dealing with nonlinear equality constraints.

We will use Theorem 1 in the case when $\mathcal{A}$ is the set of functions $\alpha(t)\in L^N_\infty[0,T]$ that take their values in the above simplex $A.$

2. Besides Theorem 1, we will also need two facts. The first one is based on a specific property of the norm in the space $L_1[0,T]$. Denote for brevity $\Delta=[0,T]$.

Let us assume here that A is an arbitrary closed and bounded polyhedron in $\mathbf{R}^N$. Denote by $\mathcal{A}$ the set of all vector-functions $\alpha(t)$ from $L^N_1(\Delta)$ with values $\alpha(t)\in A$ almost everywhere on Δ. By $ex\,A,$ denote the set of vertices of the polyhedron $A.$

Lemma 1 (ON THE L_1- DISTANCE TO ALMOST VERTICES OF A POLYHEDRON).
Let a measurable function $e(t)$ belong to $ex\,A$ a.e. on $\Delta,$ and a function $\alpha\in\mathcal{A}$ be such that $\int_\Delta|\alpha(t)-e(t)|\,dt\le\delta,$ where $\delta>0.$ Let a sequence of functions $\alpha_n\in\mathcal{A}$

weakly converge to α *(i.e.,* $\alpha_n^i \overset{weak}{\longrightarrow} \alpha^i \quad \forall\, i = 1, \ldots, N$ *w.r.t. the functions from* L_∞*). Then, for sufficiently large* n,

$$|| \, \alpha_n - \alpha \, ||_1 \;\le\; C\delta, \tag{11}$$

where the constant C *depends only on the polyhedron* A *and does not depend on the functions* $\alpha_n(t)$, $\alpha(t)$.

Proof. Since the number of vertices of any polyhedron is finite, we can assume that a given vertex is constant: $e(t) \equiv e \in ex\,A$. Consider first the simplest particular case in which the proof is completely clear.

Let $N = 1$, $A = [0,1]$, and $e = 0$. Then, by the assumption, the function $\alpha(t) \in [0,1]$ is such that $\int \alpha(t)\,dt \le \delta$, and the sequence $\alpha_n(t)$ is such that

$$0 \le \alpha_n(t) \le 1 \qquad \text{and} \qquad \alpha_n \overset{\text{weak}}{\longrightarrow} \alpha.$$

The key fact is that, for the nonnegative functions $\bar\alpha(t) \ge 0$, the norm of the space $L_1(\Delta)$ is a linear functional:

$$|| \, \bar\alpha \, ||_1 \;=\; \int_\Delta \bar\alpha(t)\,dt.$$

Using this observation and the conditions of the lemma, we obtain

$$||\alpha_n - \alpha||_1 \;\le\; ||\alpha_n||_1 + ||\alpha||_1 \;=\; \int \alpha_n\,dt + \int \alpha\,dt =$$

$$= \int (\alpha_n - \alpha)\,dt \;+\; 2\int \alpha\,dt \;<3\delta$$

for sufficiently large n. (The first integral in the last row is smaller than δ by virtue of the convergence $\alpha_n - \alpha \overset{\text{weak}}{\longrightarrow} 0$.)

In order to prove the lemma in the general case, one should note that, for any polyhedron and any its vertex e, there always exist a number $\gamma > 0$ and a unit vector $p \in \mathbf{R}^N$ such that $|\, a - e\,| \le \gamma\,(p, a-e)$ for all $a \in A$ (this follows from the fact that the tangent cone to a polyhedron at any its vertex is pointed). Then,

$$||\alpha_n - \alpha||_1 \le ||\alpha_n - e||_1 + ||\alpha - e||_1 \le$$

$$\le \gamma \int (p, \alpha_n - e)\,dt + \gamma \int (p, \alpha - e)\,dt =$$

$$= \gamma \int (p, \alpha_n - a)\,dt + 2\gamma \int (p, \alpha - e)\,dt <$$

$$< \delta + 2\gamma ||\alpha - e||_1 < (1 + 2\gamma)\,\delta.$$

It remains to set $C = 1 + 2\gamma$. The lemma is proved. □

We will use this lemma in the case when A is the simplex (8).

Note a nice corollary from Lemma 1, although it will not be used in this paper.

Theorem 2. *Let, as above, A be a convex polyhedron in $\mathbf{R}^N$. Let $\alpha(t) \in ex\,A$, $\alpha_n(t) \in A$ almost everywhere, and $\alpha_n \overset{weak}{\longrightarrow} \alpha$. Then $||\alpha_n - \alpha||_1 \to 0$.*

Thus, if a sequence of functions takes values in a convex polyhedron and weakly converges to its vertices, then it converges to them in the norm of L_1.

For the proof, one should only note that the condition of Lemma 1 is satisfied here with any $\delta > 0$, and hence, estimate (11) implies the required convergence.

For an arbitrary convex compactum and $ex\,A$ being the set of its extreme points, Theorem 2 was proved in [14, 15].

3. Another fact that we need concerns the control systems linear in control:

$$\dot{x} = f(x,t) + G(x,t)\,v. \tag{12}$$

Assume here that $x \in \mathbf{R}^m$, $v \in \mathbf{R}^k$, the vector-function f and matrix G, together with their derivatives f_x, G_x, are defined and continuous in an open set $\mathcal{R} \subset \mathbf{R}^{m+1}$.

Let a pair of functions $(x_0, v_0) \in AC^m[0,T] \times L_1^k[0,T]$ be such that $(x_0(t), t) \in \mathcal{R}$ everywhere on $[0,T]$ and satisfy equation (12). Denote $x_0(0) = a_0$.

Lemma 2. *Let a sequence of functions $v_n \in L_1^k[0,T]$ be such that $v_n \overset{weak}{\longrightarrow} v_0$ (weakly with respect to $L_\infty^k[0,T]$), and a sequence of vectors $a_n \in \mathbf{R}^m$ be such that $a_n \to a_0$. Then, for sufficiently large n, equation (12) with $v = v_n(t)$ and initial condition $x_n(0) = a_n$ has a unique solution $x_n(t)$ on $[0,T]$, and moreover, $x_n(t) \Longrightarrow x_0(t)$ uniformly on $[0,T]$.*

In brief, this lemma can be stated as follows: if a system is linear in the control, and controls converge weakly, then the corresponding state variables converge uniformly. This fact is likely to be well known, but we can not indicate a concrete reference to its proof, so we give it here.

Proof. Since the graph of the trajectory $x_0(t)$ is a compact subset of $\mathcal{R}$, its closed $\varepsilon-$ neighborhood also contains in $\mathcal{R}$ for some $\varepsilon > 0$. Without loss of generality we assume that, in this $\varepsilon-$ neighborhood, the functions f, G are uniformly bounded and Lipschitz continuous in x with a constant K common for all t.

From the general theorems on the existence of solutions to differential equations, it follows that $\forall\, n$, in some neighborhood $\mathcal{O}_n$ of the point $t_0 = 0$, equation (12) with $v = v_n(t)$ and initial condition $x_n(0) = a_n$ has a solution $x_n(t)$, which is unique and can be extended to an interval $\Delta_n = [t_0,\, t_0 + \delta_n]$, at least while it remains within the $\varepsilon-$ tube around the trajectory $x_0(t)$. Let us estimate the length of this interval.

Set $x_n = x_0 + \bar{x}_n$, $v_n = v_0 + \bar{v}_n$, $a_n = a_0 + \bar{a}_n$.

By the assumption, $\bar{v}_n \overset{\text{weak}}{\longrightarrow} 0$, $\bar{a}_n \to 0$, and the following equations hold on each Δ_n :

$$\dot{x}_0 = f(x_0, t) + G(x_0, t)\, v_0\,, \qquad x_0(t_0) = a_0\,,$$

$$(x_0 + \bar{x}_n)^{\bullet} = f(x_0 + \bar{x}_n, t) + G(x_0 + \bar{x}_n, t)(v_0 + \bar{v}_n)\,, \qquad x_n(t_0) = a_0 + \bar{a}_n\,.$$

Then, $\bar{x}_n$ satisfies on Δ_n the equation

$$\dot{\bar{x}}_n = \Gamma(\bar{x}_n, t) + \Lambda(\bar{x}_n, t)\, v_n + G(x_0, t)\, \bar{v}_n\,, \qquad \bar{x}_n(t_0) = \bar{a}_n\,, \tag{13}$$

where the functions

$$\Gamma(\bar{x}, t) = f(x_0(t) + \bar{x}, t) - f(x_0(t), t), \qquad \Lambda(\bar{x}, t) = G(x_0(t) + \bar{x}, t) - G(x_0(t), t),$$

satisfy the estimates for $|\bar{x}| \le \varepsilon$: $\quad |\Gamma(\bar{x}, t)| \le K|\bar{x}|, \qquad |\Lambda(\bar{x}, t)| \le K|\bar{x}|.$

In order to estimate $|\bar{x}_n|$, let us pass to the integral form of equation (13). First, note that since $\bar{v}_n \overset{\text{weak}}{\longrightarrow} 0$ in L_1 with respect to L_∞, the Dunford–Pettis criterion [13] implies that the family of functions $\bar{v}_n$ has a common modulus of absolute continuity of the Lebesgue integral, i.e., there exists a function $\mu : (0, \infty) \to (0, \infty)$ with $\mu(\delta) \to 0$ as $\delta \to 0+$ such that, for any set $E \subset [0, T]$ with $mes\, E \le \delta$,

$$\forall\, n \qquad \int_E |\bar{v}_n|\, dt \le \mu(\delta).$$

We can assume that μ is nondecreasing and continuous, and that it is also suited for the function $v_0(t)$.

Consider the integral of the last term in (13):

$$h_n(t) = \int_0^t G(x_0(\tau), \tau)\, \bar{v}_n(\tau)\, d\tau, \qquad t \in [0, T].$$

Since $|G(x_0(\tau), \tau)| \le K$, and $\bar{v}_n \overset{\text{weak}}{\longrightarrow} 0$, we have $h_n(t) \to 0 \quad \forall\, t \in [0, T]$.
Moreover, since $\forall\, t',\, t'' \in [0, T]$

$$|h_n(t'') - h_n(t')| \le\; K \int_{t'}^{t''} |\bar{v}_n(\tau)|\, d\tau \;\le\; K\mu(t'' - t'),$$

the function $K\mu(\delta)$ is a common modulus of continuity for all functions $h_n(t)$ on $[0, T]$, hence the pointwise convergence $h_n(t) \to 0$ implies the uniform convergence $h_n(t) \Longrightarrow 0$ on $[0, T]$.

The integral form of equation (13) is as follows:

$$\bar{x}_n(t) =\; \bar{a}_n + \int_0^t \left(\Gamma(\bar{x}_n, \tau) + \Lambda(\bar{x}_n, \tau)\, v_n\right)\, d\tau \;+ h_n(t).$$

Therefore, $\max|\bar{x}_n(t)|$ on $\Delta_n = [t_0,\, t_0 + \delta_n]$ can be estimated as

$$||\bar{x}_n||_C \le\; K\, ||\bar{x}_n||_C\, \delta_n + ||\bar{x}_n||_C\; 2\mu(\delta_n) +\; |\bar{a}_n| + ||h_n||_C\,,$$

so

$$||\bar{x}_n||_C\,(1 - K\delta_n - 2\mu(\delta_n)) \le\; |\bar{a}_n| + ||h_n||_C \;\to 0. \tag{14}$$

It remains to show that the coefficient at $||\bar{x}_n||_C$ can be bounded away from zero. To this end, set all $\delta_n = \hat{\delta}$, where $\hat{\delta} > 0$ is a solution to the equation $K\hat{\delta} + \mu(\hat{\delta}) = 1/2$. Obviously, such a solution exists (and is unique). Then, the estimate (14) guarantees that for sufficiently large n, on the common fixed interval $\hat{\Delta} = [t_0,\, t_0 + \hat{\delta}]$, we have

$$\frac{1}{2}\,||\bar{x}_n||_C \;\le\; |\bar{a}_n| + ||h_n||_C \;\to 0,$$

hence, a solution to equation (13) exists on $\hat{\Delta}$, remains in the tube $|\bar{x}| \le \varepsilon$, and moreover, uniformly converges to zero. In particular, $\bar{x}(t_0 + \hat{\delta}) \to 0$.

Now, take a new initial point $t_1 = t_0 + \hat{\delta}$ and consider the next interval $[t_1,\, t_1 + \hat{\delta}]$ of the same length $\hat{\delta}$. Since all the above estimates are still valid on this new interval, equation (13) still has a solution here, which uniformly converges to zero. Further, take the initial point at $t_2 = t_1 + \hat{\delta}$, and so on. In a finite number of such steps we will cover the whole interval $[0, T]$. The lemma is proved. □

Remark. One can see from the proof, that the smoothness in x and continuity in t of the functions $f,\, G$ are not necessary; it is sufficient to assume that they are uniformly bounded on $\mathcal{R}$, measurable in t and uniformly Lipschitz continuous in x with a constant common to almost all t.

Lemma 2 is proved here for the weakest convergence (among the "standard" ones) of the functions v_n. For example, if v_n belong to L_∞ and converge to v_0 weakly-* (i.e., with respect to the functions from L_1), then certainly these v_n belong to L_1 and converge to v_0 weakly with respect to the functions from L_∞, hence Lemma 2 is also valid in this case. This is the case in which we will use this lemma below.

4. In addition, note one more interesting fact related to the linear operator $F'(w)$. Suppose it maps "onto". Then its second component (i.e., the mapping $(\bar{x}, \bar{u}) \longmapsto g_x\,\bar{x} + g_u\,\bar{u})$ is also "onto". What can be said in this case about the matrices g_x and g_u (assuming them to be measurable and bounded)? It turns out that the matrix g_x can be absolutely arbitrary, whereas $g_u(x(t), u(t), t)$ must necessarily be of full rank uniformly in t. This is just the condition which is always assumed to hold for the constraint $g(x, u, t) = 0$ when it is included in an optimization problem.

Let us formulate this fact in the general form. Consider the operator $P:\; C^m(\Delta) \times L^r_\infty(\Delta) \to L^q_\infty(\Delta)$ acting by the rule

$$(\bar{x}, \bar{u}) \mapsto\; \Gamma(t)\,\bar{x}(t) + \Lambda(t)\,\bar{u}(t) \;=\; \bar{\eta} \in L^q_\infty(\Delta)\,, \tag{15}$$

where the matrices Γ and Λ of dimensions $q \times m$ and $q \times r$, respectively, are measurable and essentially bounded. (Here, $C^m(\Delta)$ is the space of $m-$ dimensional continuous functions on an interval Δ.) Let B_ρ denotes the ball of radius ρ centered at the origin in the corresponding space.

Lemma 3. *The operator P is surjective if and only if the matrix $\Lambda(t)$ is of full rank uniformly in t, i.e., it has an essentially bounded right inverse, or, which is the same: $\exists\,\delta > 0$ such that $\Lambda(t)B_1 \supset B_\delta$ a.e. on Δ.*

One more equivalent condition: $\;\;det\,(\Lambda(t)\,\Lambda^*(t)) \ge const > 0.$

Proof. The sufficiency of this condition for the surjectivity is obvious. What is nontrivial here, is necessity. Consider first the case $m = r = q = 1$, i.e., when

$$P(\bar{x}, \bar{u}) = \gamma(t)\,\bar{x} + \lambda(t)\,\bar{u} \,\in L_\infty(\Delta),$$

where γ and λ are scalar functions. Since the operatot P is "onto", for some $a > 0$ we have

$$P\,(B_1^C \times B_1^{L_\infty}) \,\supset\, B_a^{L_\infty}\,. \tag{16}$$

We must show that $vraimin\,|\lambda(t)| > 0$. Suppose the contrary: $vraimin\,|\lambda(t)| = 0$. Then $|\lambda(t)| \le a/3$ on a set E of positive measure. Now, consider the function γ. According to the Luzin's $C-$ property, E contains a closed set M of positive measure, on which this function is continuous. Restrict the spaces C and L_∞ to this set M. Obviously, inclusion (16) still remains valid for these restricted spaces.

Take an arbitrary discontinuous function $\hat{\eta} \in L_\infty(M)$ with $||\hat{\eta}||_\infty \le a$ having an oscillation $> a$ at a point $\theta \in M$ (i.e., $\limsup_{t\to\theta} \hat{\eta}(t) - \liminf_{t\to\theta} \hat{\eta}(t) > a$; hence, in particular, θ is not isolated in M). Then the ball $B_{a/3}(\hat{\eta})$ obviously contains discontinuous functions only (since their oscillations at θ are greater than $a/3$), and therefore, has no common points with the set

$$\Phi = \,\{\,\bar{\varphi}(t) = \gamma(t)\,\bar{x}(t) \mid \;\bar{x} \in C(M),\;\; ||\bar{x}||_C \le 1\},$$

since the latter entirely contains of continuous functions. But, in view of (16),

$$\hat{\eta} = \bar{\varphi} + \lambda(t)\bar{u} \quad \text{for some } \bar{\varphi} \in \Phi,\;\; ||\bar{u}||_\infty \le 1,$$

hence $||\bar{\varphi} - \hat{\eta}||_\infty \le ||\lambda\bar{u}||_\infty \le a/3$, and then $\bar{\varphi} \in B_{a/3}(\hat{\eta})$, a contradiction.

The general case can be reduced to the considered one-dimensional case. We leave it as an exercise to the reader. □

5. Let us now state the main result of this paper. In the space W, along with its natural norm-topology, we will also consider the $(C, L_\infty, \sigma^*)-$ topology, which is the product of the $C-$ topology for x, the $L_\infty-$ topology for $\mathbf{u}$, and the weak-* topology for α.

Theorem 3. *Let a triple $(x_0, \mathbf{u}_0, \alpha_0) \in W$ satisfy the system (4)–(7), i.e., $F(x_0, \mathbf{u}_0, \alpha_0) = 0$. Suppose that*

a) $\alpha_0^i(t) \ge const > 0$ *almost everywhere on* $[0, T]$ $\;\forall\, i = 1, \ldots, N,$

i.e., the point $\alpha_0(t) = (\alpha_0^1(t), \ldots, \alpha_0^N(t))$ is located uniformly inside the simplex A;

b) the derivative $F'(x_0, \mathbf{u}_0, \alpha_0)$ *maps* W *onto* Z.

Then, in any $(C, L_\infty, \sigma^*)-$ *neighborhood of the triple* $(x_0, \mathbf{u}_0, \alpha_0)$ *there exists a triple* $(\hat{x}, \hat{\mathbf{u}}, \hat{\alpha})$ *still satisfying system (4)–(7) and such that each function* $\hat{\alpha}^i(t)$ *takes only two values: 0 and 1 (i.e.,* $\hat{\alpha}^i(t)$ *is the characteristic function of a measurable set* $E^i \subset [0, T]$ *).*

Introducing an "ordinary" control $\hat{u} = \sum \hat{\alpha}^i(t)\, \hat{u}^i(t)$, one obtains a pair $(\hat{x}, \hat{u})$ satisfying the original system (1)–(3). It is this sense in which Theorem 3 allows one to approximate the trajectory $(x_0, \mathbf{u}_0, \alpha_0)$ of the extended system (4)–(7) by trajectories of the original system (1)–(3).

In fact, this theorem justifies the validity of the passage from the original control system to its extension by means of sliding modes. In the Western literature such theorems are often called relaxation theorems.

3. Proof of Theorem 3

Recall the notation $\mathbf{u} = (u^1, \ldots, u^N)$ and $\alpha = (\alpha^1, \ldots, \alpha^N)$, where all $u^i \in L^r_\infty$, $\alpha^i \in L_\infty$, and also recall that for $w = (x, \mathbf{u}, \alpha) \in W$ the norm is $||w|| = ||x||_{AC} + ||\mathbf{u}||_\infty + ||\alpha||_\infty$, where $||\mathbf{u}||_\infty = \sum ||u^i||_\infty$. To simplify the notation, we do not further write the tuple $\mathbf{u}$ in the bold face; hopefully, this will not cause confusion.

1) Let be given an arbitrary $(C, L_\infty, \sigma^*)-$ neighborhood of the point $w_0 = (x_0, u_0, \alpha_0)$. According to Theorem 1, we can assume that the estimate (10) holds there in. Moreover, this neighborhood always contains a closed neighborhood $\Omega = \Omega(w_0, \varepsilon)$ of the form

$$||x - x_0||_C \le \varepsilon, \qquad ||u - u_0||_\infty \le \varepsilon,$$

$$\alpha \in \mathcal{V}(\alpha_0, \varepsilon) = \{\alpha \in L^N_\infty : \ |\langle l_j, \alpha - \alpha_0\rangle| \le \varepsilon, \ \ j = 1, \ldots, \sigma\},$$

where $\varepsilon > 0$, σ is a natural number, and all $l_j(t) \in L^N_1[0, T]$ with $||l_j||_1 \le 1$.

Fix for what follows the functions $l_j(t)$, and, for brevity in notation, introduce a tuple of functionals $l = (l_1, \ldots, l_\sigma)$, so that $\langle l, \alpha\rangle$ is a vector $\langle l, \alpha\rangle = (\langle l_1, \alpha\rangle, \ldots \langle l_\sigma, \alpha\rangle)$.

The required point $\hat{w} = (\hat{x}, \hat{u}, \hat{\alpha}) \in \Omega \cap \mathcal{M}$ will be obtained as the limit of a sequence of points $w_k = (x_k, u_k, \alpha_k) \in \Omega \cap \mathcal{M}$ which will be now constructed starting from the point w_0.

For any $\delta \in [0, 1)$ denote by $A(\delta)$ the simplex in the space $\mathbf{R}^N$ obtained by contracting the original simplex A with coefficient $(1 - \delta)$ with respect to its center. Thus, $A(0) = A$, and for any $\delta > 0$ the simplex $A(\delta)$ lies inside A "at the depth δ". The condition (a) of the theorem means that, for some $\delta > 0$, we have $\alpha_0(t) \in A(\delta)$ a.e.

Now, fix the obtained ε, δ, and define the sequences

$$\varepsilon_n = \varepsilon/2^n, \qquad \delta_n = \delta/3^n, \qquad n = 0, 1, 2, \ldots,$$

(so, $\varepsilon_0 = \varepsilon$, $\delta_0 = \delta$), and then set

$$\gamma_n = \frac{1}{B} \cdot \min\left\{ \frac{\varepsilon_n}{4}, \frac{\delta_n}{3} \right\},$$

where B is the constant from (10), so that $B\gamma_n \le \varepsilon_n/4$ and $B\gamma_n \le \delta_n/3$.

2) Make the first step. By the assumption, we have $F(w_0) = 0$, i.e., the point w_0 satisfies equalities (4)–(7). Consider equality (4) as an equation with respect to x for the fixed $u = u_0$ and $\alpha \xrightarrow{\text{weak-*}} \alpha_0$ with the initial condition $x(0) = x_0(0)$. Since this equation is linear in α, we have by Lemma 2 that the corresponding solutions satisfy $||x - x_0||_C \to 0$.

Since $\alpha_0(t) \in A(\delta_0)$ and $A(\delta_0) \subset A(2\delta_0/3)$, the above α with $\alpha \xrightarrow{\text{weak-*}} \alpha_0$ can be chosen to take its values in the vertices of the simplex $A(2\delta_0/3)$, and hence, there exists $\tilde\alpha(t) \in ex\, A(2\delta_0/3)$ such that

$$|\langle l,\, \tilde\alpha - \alpha_0 \rangle| \;<\; \varepsilon_0/4,$$

i.e., $\tilde\alpha \in \mathcal{V}(\alpha_0,\, \varepsilon_0/4)$, and moreover, $||\tilde x - x_0||_C \;<\; \varepsilon_0/4$,

$$||F(\tilde x, u_0, \tilde\alpha) - F(x_0, u_0, \alpha_0)|| \;<\; \gamma_0\,.$$

(The first and third components of $F(\tilde x, u_0, \tilde\alpha)$ vanish by the definition of $\tilde x, \tilde\alpha$, and the remaining two do not depend on α; therefore, they are close to the corresponding components of the operator $F(x_0, u_0, \alpha_0)$ for $\tilde x(t)$ uniformly close to $x_0(t)$.)

The obtained triple $(\tilde x, u_0, \tilde\alpha)$ obviously lies in Ω; hence, it satisfies estimate (10), according to which there exists a point $w_1 = (x_1, u_1, \alpha_1) \in \mathcal{M}$ such that $||w_1 - (\tilde x, u_0, \tilde\alpha)|| < B\gamma_0$. We then have

$$||x_1 - x_0||_C \le ||x_1 - \tilde x|| + ||\tilde x - x_0|| < B\gamma_0 + \varepsilon_0/4 \;\le\; \varepsilon_0/4 + \varepsilon_0/4 = \varepsilon_0/2 = \varepsilon_1\,,$$

$$||u_1 - u_0||_\infty \;<\; B\gamma_0 \le \varepsilon_0/4 \;< \varepsilon_0/2 = \varepsilon_1\,,$$

and

$$|\langle l,\, \alpha_1 - \alpha_0\rangle| \;\le\; |\langle l,\, \alpha_1 - \tilde\alpha\rangle| + |\langle l,\, \tilde\alpha - \alpha_0\rangle| \le$$
$$\le ||l||_1 \cdot ||\alpha_1 - \tilde\alpha||_\infty \;+ \varepsilon_0/4 \le\; 1 \cdot B\gamma_0 + \varepsilon_0/4 <\; \varepsilon_0/4 + \varepsilon_0/4 = \varepsilon_1\,,$$

i.e., the new point $w_1 = (x_1, u_1, \alpha_1) \in \mathcal{M}$ satisfies the conditions:

$$||x_1 - x_0||_C < \varepsilon_1, \qquad ||u_1 - u_0||_\infty < \varepsilon_1\,, \qquad \alpha_1 \in \mathcal{V}(\alpha_0, \varepsilon_1),$$

and so $w_1 \in \Omega(w_0, \varepsilon_1)$, which, in view of relation $\varepsilon_1 = \varepsilon_0/2$, implies that $\Omega(w_1, \varepsilon_1) \subset \Omega(w_0, \varepsilon_0) = \Omega$.

Moreover, since $\tilde{\alpha} \in ex\, A(2\delta_0/3)$ and $||\alpha_1 - \tilde{\alpha}||_\infty < B\gamma_0 \le \delta_0/3$, we have, on the one hand,

$$\alpha_1(t) \in A(2\delta_0/3 - B\gamma_0) \subset A(\delta_0/3),$$

and on the other, $\quad \operatorname{dist}(\alpha_1(t),\, ex\, A) \;\le\; 2\delta_0/3 + B\gamma_0 \le \delta_0\,.$

Thus, the following conditions also hold:

$$\text{a.e.} \qquad \alpha_1(t) \in A(\delta_1),$$

$$\text{a.e.} \qquad \operatorname{dist}(\alpha_1(t),\, ex\, A) \le\; 3\delta_1\,. \tag{17}$$

The first step is accomplished. It is, in a sense, a preliminary one and is slightly different from the others. (We obtain that α_1 satisfies estimate (17).) The following steps are exactly iterations of one and the same procedure.

3) Suppose that, for $k = 1, \dots, n$, we obtained points $w_k = (x_k, u_k, \alpha_k) \in \mathcal{M}$ satisfying the conditions:

$$||x_k - x_{k-1}||_C < \varepsilon_k, \qquad ||u_k - u_{k-1}||_\infty < \varepsilon_k\,, \tag{18}$$

$$\alpha_k \in \mathcal{V}(\alpha_{k-1}, \varepsilon_k), \tag{19}$$

$$\Omega(w_k, \varepsilon_k) \quad \text{is contained in} \quad \Omega, \tag{20}$$

$$\text{a.e.} \qquad \alpha_k(t) \in A(\delta_k), \tag{21}$$

$$\text{a.e.} \qquad \operatorname{dist}(\alpha_k(t),\, ex\, A) \le 3\delta_k\,. \tag{22}$$

Starting from the point w_n, let us construct a point w_{n+1}. To this end, consider equality (4) as an equation with respect to x for the fixed $u = u_n$ and $\alpha \xrightarrow{\text{weak-*}} \alpha_n$ with the initial condition $x(0) = x_n(0)$. By Lemma 2, the corresponding solutions satisfy $||x - x_0||_C \to 0$.

Since $\alpha_n(t) \in A(\delta_n)$ and $A(\delta_n) \subset A(2\delta_n/3)$, there exists

$$\tilde{\alpha}(t) \in ex\, A(2\delta_n/3), \tag{23}$$

such that $|\langle l,\, \tilde{\alpha} - \alpha_n \rangle| \;<\; \varepsilon_n/4$, i.e., $\tilde{\alpha} \in \mathcal{V}(\alpha_n,\, \varepsilon_n/4)$, and moreover,

$$||\tilde{x} - x_n||_C \;<\; \varepsilon_n/4,$$

$$||F(\tilde{x}, u_n, \tilde{\alpha}) - F(x_n, u_n, \alpha_n)|| \;<\; \gamma_n\,. \tag{24}$$

In addition, by Lemma 1 and in view of estimate (22), we can assume that

$$||\tilde{\alpha} - \alpha_n||_1 \;\le\; 9N\,\delta_n\,. \tag{25}$$

Obviously, $(\tilde{x}, u_n, \tilde{\alpha}) \in \Omega(w_n, \varepsilon_n) \subset \Omega$, and hence, (24) implies that there exists a point $w_{n+1} \in \mathcal{M}$ such that

$$||w_{n+1} - (\tilde{x}, u_n, \tilde{\alpha})|| < B\gamma_n\,. \tag{26}$$

(The functions $\tilde{x}$ and $\tilde{\alpha}$ are not marked by indices; they are just intermediate auxiliary points at each step.) Then,

$$||x_{n+1} - x_n||_C \le ||x_{n+1} - \tilde{x}|| + ||\tilde{x} - x_n|| < B\gamma_n + \varepsilon_n/4 \le \varepsilon_n/4 + \varepsilon_n/4 = \varepsilon_n/2 = \varepsilon_{n+1}\,,$$

$$||u_{n+1} - u_n||_\infty < \; B\gamma_n \le \varepsilon_n/4 \; < \varepsilon_n/2 = \varepsilon_{n+1}\,,$$

and

$$|\langle l,\, \alpha_{n+1} - \alpha_n\rangle| \le \; |\langle l,\, \alpha_{n+1} - \tilde{\alpha}\rangle| + |\langle l,\, \tilde{\alpha} - \alpha_n\rangle| \le$$
$$\le ||l||_1 \cdot ||\alpha_{n+1} - \tilde{\alpha}||_\infty \; + \varepsilon_n/4 \; \le \; 1 \cdot B\gamma_n + \varepsilon_n/4 \; < \varepsilon_n/4 + \varepsilon_n/4 \; = \varepsilon_{n+1}\,,$$

i.e., $\alpha_{n+1} \in \mathcal{V}(\alpha_n, \varepsilon_{n+1})$. Thus, $w_{n+1} \in \Omega(w_n, \varepsilon_{n+1})$ and, taking into account (20) and the relation $\varepsilon_{n+1} = \varepsilon_n/2$, we get $\Omega(w_{n+1}, \varepsilon_{n+1}) \subset \Omega(w_n, \varepsilon_n) \subset \Omega$.

Next, in view of (26) we have $||\alpha_{n+1} - \tilde{\alpha}||_\infty < B\gamma_n \le \delta_n/3$, and then (23) implies the relations

$$\alpha_{n+1}(t) \in A(2\delta_n/3 - B\gamma_n) \subset A(\delta_n/3 = \delta_{n+1}),$$
$$\operatorname{dist}(\alpha_{n+1}(t),\, ex\, A) \le \; 2\delta_n/3 + B\gamma_n \; \le \delta_n = 3\delta_{n+1}\,.$$

We see that conditions (18)–(22) hold for $k = n+1$, and therefore, it is possible to make the next step, from w_{n+1} to w_{n+2}.

Finally, (25) and (26) yield one more important estimate

$$||\alpha_{n+1} - \alpha_n||_1 \; \le \; ||\alpha_{n+1} - \tilde{\alpha}||_\infty + ||\tilde{\alpha} - \alpha_n||_1 \; \le$$
$$\le \; B\gamma_n + 9N\,\delta_n \; \le \; \frac{1}{3}\,\delta_n + 9N\,\delta_n \; < (9N+1)\,\delta_n\,,$$

that holds at each step of our process.

4) Thus, we obtain a sequence of points $w_k = (x_k, u_k, \alpha_k) \in \mathcal{M}$, $k = 1, 2, \ldots,$ each of which satisfies conditions (18)–(22) and the additional estimate

$$||\alpha_{k+1} - \alpha_k||_1 \; \le \; (9N+1)\,\delta_k\,.$$

Due to this estimate and to (18), the sequence w_k is fundamental (i.e., is a Cauchy sequence) with respect to the norm $||x||_C + ||u||_\infty + ||\alpha||_1$, and hence, it has a limit $\hat{w} = (\hat{x}, \hat{u}, \hat{\alpha})$ in the space $C \times L_\infty \times L_1$ (broader than our space $W = AC \times L_\infty \times L_\infty$). Thus,

$$||x_k - \hat{x}||_C \to 0, \qquad ||u_k - \hat{u}||_\infty \to 0, \qquad ||\alpha_k - \hat{\alpha}||_1 \to 0. \tag{27}$$

Let us now see what conditions hold for the limit triple.

Since the equation $F^1(w_k) = 0$ holds for all k, where $F^1(w)$ is the first component of operator $F(w)$, then, considering the integral form of this equation, we conclude that the limit $\hat{x}(t)$ also satisfies this integral equation, which implies that actually $\hat{x}$ belongs to $AC(\Delta)$ and satisfies the equation $F^1(\hat{w}) = 0$ (hence, $||x_k - \hat{x}||_{AC} \to 0$).

The components F^2 and F^4 do not involve $\dot{x}$, α, and hence, the limit equalities $F^2(\hat{x}, \hat{u}) = 0$ and $F^4(\hat{x}, \hat{u}) = 0$ are obviously satisfied.

Now, for any closed set $V \subset \mathbf{R}^N$, the set of functions $\alpha \in L_1^N(\Delta)$ such that $\alpha(t) \in V$ a.e., is obviously closed in $L_1^N(\Delta)$. This implies that $\hat{\alpha}(t) \in A$ a.e., so $\hat{\alpha}(t)$ is bounded, i.e., $\hat{\alpha} \in L_\infty^N(\Delta)$ and $F^3(\hat{\alpha}) = \sum \hat{\alpha}^i(t) - 1 = 0$.

Thus, the limit point $\hat{w}$ belongs to the space W and satisfies the equality $F(\hat{w}) = 0$, i,e., $\hat{w} \in \mathcal{M}$. One can easily see that the set Ω is closed with respect to the convergence (27) (here one should use the fact that, if a sequence $\bar{\alpha}_k$ is bounded in $L_\infty^N(\Delta)$ and $||\bar{\alpha}_k||_1 \to 0$, then $\int_\Delta l\, \bar{\alpha}_k\, dt \to 0$ for all $l \in L_1^N(\Delta)$); therefore (20) implies that $\hat{w} \in \Omega$. (This can be also obtained by summing estimates (18) and (19).)

Finally, estimate (22) implies that $\forall \rho > 0$, for large enough n, we have $\text{dist}\,(\alpha_k(t),\, ex\, A) \le \rho$ almost everywhere on Δ, i.e., $\alpha_k(t)$ belongs to the closed $\rho-$ extension of the set of vertices of simplex A. Then, the limit $\hat{\alpha}(t)$ also belongs almost everywhere to this $\rho-$ extension, and since $\rho > 0$ is arbitrary, it belongs to the intersection of all $\rho-$ extensions, which coincides with the set of vertices $ex\, A$, since the last one is closed. Thus, $\hat{\alpha}(t) \in ex\, A$ almost everywhere, which means that each component $\hat{\alpha}^i(t)$ can take only two values: 0 or 1. The theorem is proved. $\Box$

4. Comments

1. Condition (b) of Theorem 3 is essential — without it this theorem is not valid. A counterexample is easily obtained from the classical example of Bolza; namely, consider the system

$$\dot{x} = u, \qquad \dot{y} = x^2 + (u^2 - 1)^2, \qquad y(0) = y(1) = 0.$$

The set of its solutions is empty, since for any $x(t)$ and $y(0)$, always $y(1) > y(0)$. However, the corresponding extended (convexified) system

$$\dot{x} = \alpha_1\, u_1 + \alpha_2\, u_2, \qquad \dot{y} = x^2 + \alpha_1\,(u_1^2 - 1)^2 + \alpha_2\,(u_2^2 - 1)^2,$$

$$\alpha_1 + \alpha_2 = 1, \qquad y(0) = y(1) = 0$$

(here, we use the subscript indices) has a solution $x(t) \equiv y(t) \equiv 0$ generated by the controls $\alpha_1 = \alpha_2 = 1/2$, $u_1 = -1$, $u_2 = 1$, and this solution cannot be approximated by solutions of the original system, because the latter are just missing. The cause of this situation is that the given solution does not satisfy condition (b): the equality constraints of the extended system are degenerate.

2. As to condition (a), it plays a rather technical role, and probably can be excluded. For example, one can proceed as follows. For any $\delta > 0$, consider the set

$$E_\delta = \{ t \mid \forall i \quad \alpha^i(t) \ge \delta \quad \text{a.e. on} \quad [0,T] \},$$

outside of which all the functions u^i, α^i should be fixed, and vary the controls u^i and α^i only on this set, i.e., consider them as elements of the spaces $L^r_\infty(E_\delta)$ and $L_\infty(E_\delta)$ respectively. As $\delta \to 0$, the measure of E_δ tends to the full measure of the interval $[0,T]$. This implies that, if the original operator $F'(w_0) : W \to Z$ is surjective, then for small $\delta > 0$ its restriction to the corresponding space W_δ, defined on E_δ, will also remain surjective, and therefore, all the above procedure of iterative corrections should also work in this case. Of course, these preliminary considerations require a thorough study.

3. Theorem 3 can be applied to the proof of the Maximum principle for optimal control problems with terminal, state, and regular mixed constraints by introducing sliding modes. The author got to know the idea of such proof from A.A. Milyutin as early as in the mid 1970s; at that time he implemented it and later published in [17] and [21, Ch. 4]. A similar theorem was proved by S.V. Chukanov [18] for control systems governed by integral equations, and he also applied it to the proof of the corresponding Maximum principle.

5. Acknowledgments

This work was supported by the Russian Foundation for Basic Research (project 04-01-00482) and by the Program for Leading Scientific Schools (project NSh-5813.2006.1).

References

[1] R.V. Gamkrelidze, Optimal sliding states, *Soviet Math. Doklady,* 1962, v. 3, p. 559–562.

[2] L.C. Young, *Lectures on the Calculus of Variations and Optimal Control Theory,* Saunders, Philadelphia, 1969.

[3] E.J. McShane, Relaxed controls and variational problems, *SIAM J. on Control,* 1967, v. 5, p. 438–485.

[4] A.D. Ioffe and V.M. Tikhomirov, *Theory of Extremal Problems*, Moscow, Nauka, 1974; English translation: Amsterdam, North-Holland, 1979.

[5] J. Warga, *Optimal Control of Differential and Functional Equations,* New York, Academic Press, 1972.

[6] C. Olech, Existence theory in optimal control, in *"Control theory and topics in functional analysis",* v. I, Vienna, 1976, p. 291–328.

[7] E.J. Balder, A general denseness result for relaxed control theory, *Bull. of Australian Math. Society,* 1984, v. 30, p. 463–475.

[8] Z. Artstein, Rapid oscillations, chattering systems, and relaxed controls, *SIAM J. on Control and Optimization,* 1989, v. 27, no. 5, p. 940–948.

[9] J.F. Rosenblueth and R.B. Vinter, Relaxation procedures for time delay systems, *J. Math. Anal. and Appl.,* 1991, v. 162, no. 2, p. 542–563.

[10] T. Roubicek, *Relaxation in Optimization Theory and Variational Calculus,* de Gruyter, Berlin, 1997.

[11] A.A. Tolstonogov, *Differential Inclusions in a Banach Space,* Dordrecht, Kluwer, 2000.

[12] A.I. Bulgakov and V.V. Vasilyev, On the theory of functional-differential inclusions of neutral type, *Georgian Math. J.*, 2002, v. 9, no. 1, p. 33–52.

[13] N. Dunford and J.T. Schwartz, *Linear Operators. Part I: General Theory,* Interscience, London, 1958; reissued in New York, Wiley-Interscience, 1988.

[14] A. Visintin, Strong convergence results related to strict convexity, *Commun. Partial Diff. Equations,* 1984, v. 9, No. 5, p. 439–466.

[15] T. Rzezuchowski, Strong convergence of selections implied by weak, *Bull. Austral. Math. Soc.,* 1989, v. 39, p. 201–214.

[16] A.V. Dmitruk, A.A. Milyutin, and N.P. Osmolovskii, Luysternik's theorem and the theory of extrema, *Russian Math. Surveys,* 1980, v. 35, no. 6, p. 11–51.

[17] A.V. Dmitruk, Maximum principle for a general optimal control problem with state and regular mixed constraints, in *"Optimality of Control Dynamical Systems",* Moscow, All-Union Inst. for Systems Research, 1990, issue 14; English translation in *Computational Mathematics and Modeling*, 1993, v. 4, no. 4, p. 364–377.

[18] S.V. Chukanov, Maximum principle for optimal control problems with integral equations, in *"Necessary Condition in Optimal Control"* (A.A.Milyutin ed.), Moscow, Nauka, 1990, Ch. 6 (in Russian).

[19] A.V. Dmitruk, A nonlocal Lyusternik estimate and its application to control systems with sliding modes, in *"Nonlinear Control Systems 2001"* (ed. A.B.Kurzhanski and A.L.Fradkov), Elsevier, 2002, vol. 2, p. 1061–1064.

[20] A.V. Dmitruk, On a nonlocal metric regularity of nonlinear operators, *Control and Cybernetics*, 2005, v. 34, no. 3, p. 723–746.

[21] A.A. Milyutin, A.V. Dmitruk, and N.P. Osmolovskii, *The Maximum Principle in Optimal Control*, Mech-Math. Faculty of the Moscow State University, 2004 (in Russian).