International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

On the Model Simplification of Control/Uncertain Systems with Multiple Time Scales

Veliov, V.M.

IIASA Working Paper

WP-96-082

July 1996

Working Paper

On the Model Simplification of Control/Uncertain Systems with Multiple Time Scales

V.M. Veliov

WP-96-82
July 1996

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

On the Model Simplification of Control/Uncertain Systems with Multiple Time Scales

V.M. Veliov

WP-96-82
July 1996

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Contents

ON THE MODEL SIMPLIFICATION OF CONTROL/UNCERTAIN SYSTEMS WITH MULTIPLE TIME SCALES

Vladimir Veliov[1]

1 Introduction

Differential equations with multiple time scales naturally arise in the modelling of real systems in which "slow" and "fast" motions are involved, the later ones usually caused by presence of "small" masses, capacities, time constants, etc. (multiplying a part of the derivatives of the state variables). A common practice (at least as a first approximation step) is to simplify such a model by neglecting the small parameters. This *ad hoc* simplification is legitimized in many cases by the *singular perturbation theory* for differential equations.

The situation significantly complicates if the differential equation into consideration depends on some control or (deterministic) uncertain inputs, giving rise to a differential inclusion. Neglecting the small parameters in this case may lead to a completely distorted picture, even in cases where the model simplification is legitimate for any fixed value of the control/uncertain input. The reason is, in principle, that the control/uncertain input may vary with the time with a speed that is compatible with that of the "fast" variables. The interaction between the "fast" variables and a "fast" changing input may create trajectories that could not be explained within the simplified model.

The aim of the present paper is to obtain general conditions which justify the model simplification by formally neglecting the small parameter multiplying a part of the derivatives of a control/uncertain system. As a basic model of such a system we use a singularly perturbed differential inclusion, the simplified model taking, therefore, the form of a system of differential and static (algebraic) inclusions. The results extend (and generalize) the classical Tikhonov theorem for singularly perturbed differential equations and outline its scope of extendibility to differential inclusions.

A detailed description of the problem and of the organization of the paper are given in the next section, which can be considered as a continuation of the introduction.

[1]Institute of Statistics, Informatics and Operations Research, University of Vienna, Universitätsstrasse 5, A-1010 Vienna, Austria; Institute of Mathematics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

2 The main issue: Continuity of the trajectory bundle

In 1952 Tikhonov published a theorem [22] characterizing the limit behaviour of the solutions of a *singularly perturbed* system of differential equations

$$\dot{x}(t) = f_1(t,x,y), \;\; x(0) = x^0, \tag{1}$$

$$\varepsilon\dot{y}(t) = f_2(t,x,y), \;\; y(0) = y^0, \tag{2}$$

where $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$, ε is a "small" positive parameter. For obvious reasons y is called *fast* variable and the second group of equations is also called "fast". For readers convenience we briefly remind the Tikhonov theorem.

The limit case of the above system is the so-called *degenerate* system

$$\dot{x}(t) = f_1(t,x,y), \;\; x(0) = x^0, \tag{3}$$

$$0 = f_2(t,x,y). \tag{4}$$

The second equation is supposed in the Tikhonov theorem to possess an isolated solution $y = \xi(t,x)$. Substituting y in the first equation, the degenerate system takes the form

$$\dot{x}(t) = f_1(t,x,\xi(t,x)), \;\; x(0) = x^0, \tag{5}$$

$$y(t) = \xi(t,x(t)).$$

The principle *stability* assumption of the Tikhonov theorem is: for each t and x the equilibrium $\xi(t,x)$ of the so-called *associated system*

$$\frac{dz}{d\tau} = f_2(t,x,z(\tau))$$

is Lyapunov asymptotically stable (uniformly in (t,x)). It is supposed, in addition, that the initial state y^0 belongs to the domain of attraction of $\xi(0,x^0)$.

The claim of the Tikhonov theorem is the following: if the solution $(x_0(\cdot), y_0(\cdot))$ of the *reduced* equation (5) exists on $[0,T]$, then for all sufficiently small ε the solution $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ of (1),(2) also exists on $[0,T]$ and for each fixed $\alpha > 0$

$$\lim_{\varepsilon\to 0} \|x_\varepsilon(\cdot) - x_0(\cdot)\|_{C[0,T]} = 0, \quad \lim_{\varepsilon\to 0} \|y_\varepsilon(\cdot) - y_0(\cdot)\|_{C[\alpha,T]} = 0. \tag{6}$$

For proofs and refinements of the Tikhonov theorem see [18, 15, 24].

Notice that the convergence (6) is a metric convergence: with respect to the norm in C^m for x and with respect to the metric

$$\tau(y_1(\cdot), y_2(\cdot)) = \inf\{\alpha + \beta\,;\; |y_1(t) - y_2(t)| \leq \beta \text{ for each } t \in [\alpha, T]\}$$

in the space of all bounded functions on $[0,T]$. We shall refer to the product of these two metrics as *Tikhonov metric.*

The Tikhonov theorem stimulated a large number of investigations and was elaborated in many directions, including numerical aspects (see [25] for an recent overview of the development of the singular perturbation theory, mainly in the former USSR). A new impulse to the singular perturbation analysis was given by control theory (see [19] for numerous applications). It turned out that if the system into consideration depends on a control variable, then the Tikhonov theorem may fail to work. Indeed, consider now a control system

$$\dot{x}(t) = f_1(t, x, y, u), \quad x(0) = x^0, \tag{7}$$

$$\varepsilon \dot{y}(t) = f_2(t, x, y, u), \quad y(0) = y^0, \tag{8}$$

where $u \in U \subset \mathbf{R}^r$ is a (time-varying) control parameter. Even if the Tikhonov stability condition is (uniformly) satisfied for each admissible control value $u \in U$ one may fix a discontinuous control function $u(\cdot)$ and the resulting solution of (7),(8) typically fails to converge in the sense of (6) to the corresponding solution of the degenerate system. The reason is, that the right-hand side f_2 in the Tikhonov theorem should be continuous. On the other hand it is too restrictive from control point of view to deal with (equi-)continuous controls only. The situation complicates even more if the control function plugged in (7),(8) depends on ε. A number of *discontinuity effects* are known. For example, it may happen (even in the case of a linear stable "fast" subsystem [8]) that for the solution $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ corresponding to $u_\varepsilon(\cdot)$ the "slow" component $x_\varepsilon(\cdot)$ converges in $C[0,T]$ to some $x_0(\cdot)$ but the limit fails to be a trajectory of the degenerate control system (that is, the trajectory bundle is not upper semicontinuous at $\varepsilon = 0$). Other examples ([12]) show that even for an entirely linear control system with stable "fast" subsystem it may happen that for a sequence of trajectories $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ of (7),(8) the end points $(x_\varepsilon(T), y_\varepsilon(T))$ converge to a point that is far away from the set of all end points of trajectories of the degenerate system (that is, the reachable set is not upper semicontinuous at $\varepsilon = 0$).

One possibility to cope with such discontinuities is to define a *limit* system (corresponding to $\varepsilon = 0$) in a different way, so that the set of solutions of the perturbed system converges to the set of solutions of the limit system in a prescribed topology. This is a reach field of investigation as far as a variety of topologies are meaningful, as well as senses of "convergence", when speaking about set-valued mappings. This way was undertaken in [12, 13] and was developed in [14] and other papers by the same author in a more general setting (see also [23, 7, 1]). In the present paper, however, we address the "classical" issue of convergence of the set of trajectories of (7),(8) to the set of trajectories of the *formally obtained degenerate system*, that is, by setting $\varepsilon = 0$. The reason is that, on one hand this issue provides a ground for the widely used formal model simplification practice (just by neglecting the small parameters), and on the other hand that we establish such convergence (even in a rather strong topology like in (6)) under conditions that, despite being relatively strong, still are fulfilled in many meaningful situations.

We present the control system (7),(8) by a singularly perturbed differential inclusion

$$\begin{pmatrix} \dot{x}(t) \\ \varepsilon\dot{y}(t) \end{pmatrix} \in F(t,x,y), \quad t \in [0,T],\ x \in \mathbf{R}^m,\ y \in \mathbf{R}^n, \tag{9}$$

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} x^0 \\ y^0 \end{pmatrix}. \tag{10}$$

Denote by Z_ε the set of solutions, that is, the set of all absolutely continuous pairs of functions $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ starting from (x^0, y^0), that satisfy (9) for a.e. $t \in [0,T]$. Corresponding to $\varepsilon = 0$ is the set of solutions Z_0 of the degenerated inclusion

$$\begin{pmatrix} \dot{x}(t) \\ 0 \end{pmatrix} \in F(t,x,y), \quad x(0) = x^0, \tag{11}$$

consisting of all pairs of an absolutely continuous $x(\cdot)$ and a measurable $y(\cdot)$ satisfying (11) for a.e. $t \in [0,T]$. Then the basic question in the singular perturbation analysis is: does, and in what sense, Z_ε converge to Z_0, or equivalently, is the mapping $\varepsilon \to Z_\varepsilon$ continuous at $\varepsilon = 0$ and in which sense.

The upper semicontinuity of Z_ε with respect to the ($C \times L_1$-weak) topology is proven in [11] under a specific strong monotonicity condition (which turns out to be sufficient for our asymptotic stability condition – see Section 4 for a more detailed comparison) and for mappings F with a convex graph with respect to y. Under the same monotonicity condition [9] proves upper semicontinuity in the Tikhonov metric of the mapping $\varepsilon \longrightarrow Z_\varepsilon^L$, where Z_ε^L is the set of those $z_\varepsilon \in Z_\varepsilon$, which are Lipschitz with Lipschitz constant L. In the recent paper [20] the upper semicontinuity of Z_ε in the Tikhonov metric is investigated by a rather different technique (employing viability theory). The result applies to mappings F in a decomposed form $F = F_1 \times F_2$, F_1 and F_2 corresponding to x and y, respectively. A sort of stability of the fast subsystem is ensured by a condition involving the contingent derivative of the mapping F_2.

The first main result of this paper gives a sufficient condition for upper semicontinuity in the Tikhonov metric of the mapping $\varepsilon \longrightarrow Z_\varepsilon$ at $\varepsilon = 0$. The upper semicontinuity is meant in the metric sense (the so-called ε-δ upper semicontinuity [6]). Equivalently formulated, the result states that

$$\sup_{z_\varepsilon \in Z_\varepsilon} \operatorname{dist}(z_\varepsilon, Z_0) \longrightarrow 0 \quad \text{with} \quad \varepsilon \longrightarrow 0,$$

where the distance in the above relation is with respect to the Tikhonov metric. In the case of a differential equation, where the mapping Z_ε is single valued, both the suppositions and the claim of the main theorem (Section 3) reduce to those of the Tikhonov theorem.

The principle requirement for this result is the strong asymptotic stability of the set $K_0(t,x)$ of equilibrium points of the *associated inclusion*

$$\frac{dy}{d\tau} \in \hat{F}(t,x,y) \stackrel{def}{=} \{\eta \in \mathbf{R}^n\,;\ \begin{pmatrix} \psi \\ \eta \end{pmatrix} \in F(t,x,y) \ \text{ for some } \psi \in \mathbf{R}^m\} \tag{12}$$

for fixed t and x. Certain structural condition for F is also required. These condition has no counterpart in the single valued case (where it is automatically satisfied), but is shown by an example to be essential.

In Section 4 we elaborate the stability condition in terms of Lyapunov functions and discuss its relation with other conditions used in similar contexts by other authors, as well as some examples.

The requirement of strong asymptotic stability of the equilibrium set $K_0(t,x)$ with respect to the associated inclusion (12) may happen to be too restrictive in some applications. In Section 5 we replace it by strong asymptotic stability of the *invariance envelop* of $K_0(t,x)$, which is an essentially weaker condition. The price of this relaxation is, however, that $Z\varepsilon$ is upper semicontinuous only in the $C \times$ (L_1-weak) topology. Moreover, convexity of the graph of F with respect of y is required. The result extends that of [11].

Section 6 is devoted to the lower semicontinuity of Z_e in the $C \times L_1$ metric. The principle condition here is the *weak* asymptotic stability of each point of the equilibrium set $K_0(t,x)$ with respect to the associated inclusion (12). The result complements those in [28, 21, 10].

The somewhat longer proofs of the two main results—the upper semicontinuity in the Tikhonv metric and the lower semicontinuity in the $C \times L_1$ metric— are given in appendixes A and B, respectively.

3 Upper semicontinuity in the Tikhonov metric: General theorem

We start by introducing some notations. The closed unit ball in $\mathbf{R}^m$ (as well as in all linear normed spaces) will be denoted by $\mathbf{B}$. The distance (with respect to any metric – usually the Euclidean one) from a point x to a set S is denoted by $\mathrm{dist}(x,S)$. The Hausdorff distance between two compact sets P and Q in a metric space is defined as

$$H(P,Q) = \sup\{\mathrm{dist}(p,Q),\, \mathrm{dist}(q,P)\,;\; p \in P,\; q \in Q\}.$$

We also denote by $\mathcal{P}_S A$ the projection of the set $A \in \mathbf{R}^r$ on the closed set $S \subset \mathbf{R}^r$, that is,

$$\mathcal{P}_S A = \{x \in S\,;\; |x-a| = \mathrm{dist}(a,S) \text{ for some } a \in A\}.$$

In particular, $\mathcal{P}_x F$ (or $\mathcal{P}_y F$) will denote the projection of F on the x-space $\mathbf{R}^m$ (y-space $\mathbf{R}^n$, respectively). Thus $\hat{F}(t,x,y) = \mathcal{P}_y F(t,x,y)$ (see (12)).

Now we formulate the assumptions.

Supposition A1. $F : [0,T] \times \mathbf{R}^m \times \mathbf{R}^n \Rightarrow \mathbf{R}^m \times \mathbf{R}^n$ is non-empty convex compact valued, measurable in t, locally bounded and locally Lipschitz (with respect to the Hausdorff metric) in (x,y), uniformly with respect to t; $\mathcal{P}_x F(t,\cdot,\cdot)$ is continuous, uniformly in $t \in [0,T]$, on the compact sets in $\mathbf{R}^m \times \mathbf{R}^n$; $\hat{F}(\cdot,\cdot,\cdot)$ is continuous.

Supposition A2. There is a compact set $D \subset \mathbf{R}^m$ and a non-empty compact valued mapping $K : [0,T] \times D \Rightarrow \mathbf{R}^n$ which is continuous, Lipschitz in x and satisfies

$$K(t,x) \subset K_0(t,x) \stackrel{def}{=} \{y\,;\; 0 \in \mathcal{P}_y F(t,x,y)\} \quad \forall\,(t,x) \in [0,T] \times D. \tag{13}$$

The set $K(t,x)$ will play the role of the "isolated zero" in the Tikhonov theorem, while $K_0(t,x)$ is the set of all "zeros".

Supposition A3. The mapping $(t,x) \mapsto F_0(t,x) \stackrel{def}{=} \mathcal{P}_x F(t,x,K(t,x))$ is Lipschitz continuous with respect to $x \in D$ with integrable on $[0,T]$ Lipschitz constant; all solutions of the differential inclusion

$$\dot{x}(t) \in \operatorname{co} F_0(t,x), \quad x(0) = x^0 \tag{14}$$

remain in the interior of D on $[0,T]$.

Supposition A4. (Strong asymptotic stability of K in the sense of Lyapunov.) For every $\mu > 0$ there exists $\delta = \delta(\mu) > 0$ and for every $\nu > 0$ there exists $\tau_0 = \tau_0(\delta,\nu)$ such that for every $(t,x) \in [0,T] \times D$ every solution $\tilde{y}(\cdot)$ of the associated inclusion (12) for which $\operatorname{dist}(\tilde{y}(0), K(t,x)) \le \delta$ exists on $[0,+\infty)$ and satisfies

$$\operatorname{dist}(\tilde{y}(\tau), K(t,x)) < \mu \;\; \forall \tau \ge 0 \quad \text{and} \quad \operatorname{dist}(\tilde{y}(\tau), K(t,x)) < \nu \;\; \forall \tau \ge \tau_0.$$

Supposition A5. For every $\nu > 0$ there exists $\tau_0 = \tau_0(\nu)$ such that every solution $\tilde{y}^0(\cdot)$ of

$$\frac{d}{d\tau}\tilde{y}^0(\tau) \in \hat{F}(0,x^0,\tilde{y}^0(\tau)), \quad \tilde{y}^0(0) = y^0 \tag{15}$$

exists on $[0,+\infty)$ and satisfies

$$\operatorname{dist}(\tilde{y}^0(\tau), K(0,x^0)) \le \nu, \quad \forall \tau \ge \tau_0.$$

The proofs of the next proposition and of the theorem below will be given in Appendix A.

Proposition 1 *Suppose that* **A1** – **A5** *are fulfilled (here (13) and Lipschitz continuity of* $K(t,\cdot)$ *need not be required). Then there exist* $\varepsilon_0 > 0$ *such that for each* $\varepsilon \in (0,\varepsilon_0]$ *every solution* $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ *of (9),(10) is extendible to* $[0,T]$, $x_\varepsilon(t) \in \operatorname{int} D$ *for* $t \in [0,T]$ *and*

$$\alpha_\varepsilon \stackrel{def}{=} \sup_{x_\varepsilon \in \mathcal{P}_x Z_\varepsilon} \operatorname{dist}(x_\varepsilon(\cdot), \mathcal{P}_x Z_0) \quad \textit{(here the distance is in } C([0,T])\textit{)},$$

$$\gamma_\varepsilon(t) \stackrel{def}{=} \sup_{(x_\varepsilon,y_\varepsilon) \in Z_\varepsilon} \operatorname{dist}(y_\varepsilon(t), K(t,x_\varepsilon(t)))$$

have the properties a) $\lim_{\varepsilon \to 0} \alpha_\varepsilon = 0$*; b)* $\gamma_\varepsilon(\cdot)$ *are uniformly bounded; c)* $\lim_{\varepsilon \to 0} \gamma_\varepsilon(t) = 0$ *uniformly on each subinterval* $[\theta,T]$, $\theta > 0$.

The above proposition will be used in the proof of the main result, but it is also of independent interest. In fact, it readily implies the convergence claim of the Tikhonov theorem if $K(t,x)$ is single valued. This proposition "provokes" the idea to redefine the limit set Z_0 as

$$\tilde{Z}_0 = \{(x(\cdot), y(\cdot))\,;\; x(\cdot) \in \mathcal{P}_x Z_0,\; y(t) \in K(t, x(t)),\, t \in [0,T]\}.$$

Obviously with this definition for $\varepsilon = 0$ the mapping Z_ε would be upper semicontinuous in the Tikhonov metric. It is also clear, that $Z_0 \subset \tilde{Z}_0$, but the last set is (as a rule, in the set-valued case) strictly bigger than Z_0. In particular, Z_ε would not be lower semicontinuous if defined in the above modified way. Therefore, we stick to the formal definition of Z_0 given in the introduction.

The following simple example shows that Proposition 1 still do not imply that the Tikhonov distance from $(x_\varepsilon, y_\varepsilon)$ to Z_0 tends to zero with ε, that is, suppositions **A1**–**A5** do not imply upper semicontinuity in the Tikhonov metric.

Example 1. Consider the following control system:

$$\begin{aligned} \dot{x} &= u y_2, & x(0) = 0,\\ \varepsilon \dot{y_1} &= -y_1 + u, & y_1(0) = 0,\\ \varepsilon \dot{y_2} &= -y_2 + v, & y_2(0) = 0, \end{aligned}$$

$$u,\, v \in [-1, 1].$$

Obviously suppositions **A1**–**A5** are fulfilled for $K(t,x) = K_0(t,x) = \{(y_1, y_2)\,;\; y_1 \in [-1,1],\, y_2 \in [-1,1]\}$ – the set of equilibrium points of the associated system.

Define the function u_ε to be equal alternatingly to 1 or -1 on intervals with length $\varepsilon \ln 2$, starting with value 1. Denote by $t_1^\varepsilon, t_2^\varepsilon \ldots$ the jump points of u_ε. Then define v_ε to have exactly one jump in each interval $[t_i^\varepsilon, t_{i+1}^\varepsilon]$, chosen so that $y_2(t)$ has the sign of $u_\varepsilon(t)$. This is possible, as simple calculation shows. Denote by $(x^\varepsilon, y_1^\varepsilon, y_2^\varepsilon)$ the corresponding trajectory.

The degenerate system for this example is

$$\begin{aligned} \dot{x} &= u y_2, \quad x(0) = 0,\\ y_1 &= u,\\ y_2 &= v. \end{aligned}$$

Somewhat longer calculations show that the Tikhonov distance from $(x^\varepsilon, y_1^\varepsilon, y_2^\varepsilon)$ to Z_0 is bigger than a positive constant (0.2 is such) no matter how small is ε. The reason for the discontinuity in this example is that the control u and the "fast" variable y_2 interact nonlinearly in the "slow" equation.

As the above example indicates, in order to ensure that Z_ε converges to Z_0 in the Tikhonov metric one has to impose some additional condition. Below we formulate such a condition, concerning only the structure of the right-hand side F of (9).

Denote by $\text{Lins}\, V$ the support subspace to the convex set $V \subset \mathbf{R}^n$:

$$\text{Lins}\, V = \text{Lin}\,(V - v), \quad \text{where } v \in V.$$

Supposition A6. The set-valued mapping F in (9) has the form

$$F(t,x,y) = \begin{pmatrix} F_1(t,x,y) \\ F_2(t,x,y) \end{pmatrix} + \begin{pmatrix} B(t,x) \\ I_n \end{pmatrix} U(t,x), \tag{16}$$

where F_2 satisfies the same conditions as $\hat{F}$ in Supposition **A1**, $B : [0,T] \times \mathbf{R}^m \mapsto \mathbf{R}^{m \times n}$, I_n is the $(n \times n)$-unit matrix and $U : [0,T] \times \mathbf{R}^n \Rightarrow \mathbf{R}^n$ is a continuous, convex compact valued mapping. Moreover, let

i) $\text{Lins}\, U(t,x)$ have a constant dimension;

ii) $\mathcal{P}_{\text{Lins}\, U(t,x)} \text{Lins}\, F_2(t,x,y) \subset \text{Ker}\, B(t,x)$.

We mention that now

$$\mathcal{P}_x F = F_1 + BU, \quad \hat{F} = \mathcal{P}_y F = F_2 + U$$

and the corresponding requirements about F and $\mathcal{P}_y F$ in Supposition **A1** are still in effect.

In view of Example 1 the above condition is not that much restrictive: in control-theoretical terms it means that the controls u that influence directly both the "slow" and the "fast" equations should enter linearly, with multipliers that do not depend on the "fast" variables. On the other hand, no similar restrictions concern the control inputs (encapsulated in F_1 and F_2) that enter either in the "slow", or in the "fast" subsystem but not in both of them.

We remind that a set-valued mapping $[0,1] \ni \varepsilon \longrightarrow \Gamma(\varepsilon) \subset S$ (where S is a topological space) is *upper semicontinuous* at $\varepsilon = 0$ iff for every open set $\Omega \subset S$ such that $\Gamma(0) \subset \Omega$ it holds also $\Gamma(\varepsilon) \subset \Omega$ for all sufficiently small ε. If S is a metric space, then we say that Γ is *metrically upper semicontinuous* if the above property is fulfilled for each Ω of the form $\Omega = \{s \in S\,;\ \text{dist}(s, \Gamma(0)) < \alpha\}$, $\alpha > 0$. If the closure of $\Gamma(0)$ is compact, then metric upper semicontinuity implies upper semicontinuity, but not in general.

Theorem 1 *Suppose that* **A1–A6** *are satisfied. Then for all sufficiently small ε every solution of (16),(10) is extendible to $[0,T]$ and the mapping $\varepsilon \longrightarrow Z_\varepsilon$ is metrically upper semicontinuous at $\varepsilon = 0$ with respect to the Tikhonov metric.*

Metric upper semicontinuity is equivalent to the following property: for every sequence $z_\varepsilon \in Z_\varepsilon$ there are $\tilde{z}_\varepsilon \in Z_0$ such that $z_\varepsilon - \tilde{z}_\varepsilon \longrightarrow 0$ in the Tikhonov metric. It is also equivalent to

$$\sup_{z_\varepsilon \in Z_\varepsilon} \text{dist}(z_\varepsilon, Z_0) \longrightarrow 0 \quad \text{with } \varepsilon \longrightarrow 0,$$

where the distance is in the Tikhonov metric.

In the case of a differential equation one may just take $U(t,x) = \{0\}$ and F_1 and F_2 - single valued. In this case **A6** is automatically fulfilled, while **A1**–**A5** are implied by the assumptions of the Tikhonov theorem, if we take $K(t,x) = K_0(t,x)$ to be the *unique* zero of the "fast" equation. Since in this case Z_ε and Z_0 are also single valued we obtain the Tikhonov theorem.

In control theory it often happens that the controls influencing directly the "slow" and the "fast" equations are independent. In such a situation one also may take $U(t,x) = \{0\}$ and skip **A6**. On the other hand, as Example 1 shows, **A6** is essential in the more general case of controls that enter simultaneously in the "slow" and in the "fast" equations.

The stability assumption **A4** deserves a special attention. It will be discussed in the next section together with some examples and applications of Theorem 1.

4 Upper semicontinuity in the Tikhonov Metric: Particular cases and examples

First we shall elaborate the strong asymptotic stability condition **A4** giving a more practical sufficient condition. Since it concerns the associated inclusion (12), where t and x are fixed, we sometimes omit them in the notations. Thus, the question is under what conditions a set K consisting of equilibrium points of $\hat{F}$ is strongly asymptotically stable (in the sense of **A4**).

The following proposition gives a necessary and sufficient condition for the required type of stability in terms of Lyapunov functions. The proof adapts the ideas from the similar considerations in [6, Chapter 14] (where K consists of a single point) and [5, Sect.6] (where the stability – not necessarily asymptotic – of a *set* K is investigated).

Proposition 2 *Suppose that $G : \mathbf{R}^n \Rightarrow \mathbf{R}^n$ is convex compact valued and upper semicontinuous. Then strong asymptotic stability of the compact set K with respect to the differential inclusion*

$$\dot{y} \in G(y) \tag{17}$$

(in the sense employed in **A4***) is equivalent to the following: there exists a function $V : \mathbf{R}^n \times [0, +\infty) \mapsto [0, +\infty)$ which is lower semicontinuous, $V(x, \cdot)$ is monotone decreasing and*

a_1) $\forall \mu > 0 \; \exists \delta > 0 \; : \; K + \delta \mathbf{B} \subset \{y\, ; \; V(y,0) \leq \mu\}$*;*

a_2) $\forall \mu > 0 \; \exists \delta > 0 \; : \; \{y\, ; \; V(y,0) \leq \delta\} \subset K + \mu \mathbf{B}$*;*

b_1) $\lim_{\tau\to+\infty} V(y,\tau) = 0$ uniformly with respect to y in some neighborhood $K + \delta_0\mathbf{B}$ of K;

b_2) any trajectory $y(\cdot)$ of (17) which exists on some interval $[0,\theta)$ and for which $y(0) \in K + \delta_0\mathbf{B}$ satisfies

$$V(y(\tau),0) \le V(y(0),\tau) \ \ \forall\tau \in [0,\theta).$$

One may look for a Lyapunov function in the form $V(y,\tau) = V(y)e^{-2\rho\tau}$ (corresponding to exponential asymptotic stability), where $V(y)$ is locally Lipschitz and satisfies a_1) and a_2). Then b_1) is also satisfied and b_2) is implied by

$$V^+(y(\tau);\dot{y}(\tau)) \le -2\rho V(y(\tau)), \ \ \tau \ge 0$$

on the trajectories of (17), where $\rho > 0$ and $V^+(y;\eta)$ is the upper Dini derivative of V in the direction η. A standard argument shows that the last inequality is satisfied if

$$\sup_{\eta\in G(y)} V^+(y;\eta) \le -2\rho V(y) \tag{18}$$

for every y in a neighborhood of K. For example, one can take $V(y) = 0.5\text{dist}(y,K)^2$ (here *dist* could be taken with respect to any Hilbert metric in $\mathbf{R}^n$). To give a sufficient condition for (18) in this case, we need the notion of proximal normal cone.

By definition, the *proximal normal cone* $N_K^\perp(y)$ to a closed set $K \subset \mathbf{R}^n$ at $y \in K$ is the cone generated by the vectors $z - y$ for which $y \in \mathcal{P}_K(z)$.

Suppose that for every $z \in \partial K$ (the boundary of K) and $l \in N_K^\perp(z)$

$$\max_{\eta\in G(z+l)} \langle l,\eta\rangle \le -\rho|l|^2. \tag{19}$$

Then for each $y \in \mathbf{R}^n \setminus K$ and $z \in \mathcal{P}_K(y)$ we have $l = y - z \in N_K^\perp(z)$, therefore

$$\langle y-z,\eta\rangle \le -\rho|y-z|^2 = -\rho\text{dist}(y,K)^2.$$

Taking into account that $V^+(y;\eta) = \min_{z\in\mathcal{P}_K(y)}\langle y-z,\eta\rangle$ we have

$$\sup_{\eta\in G(y)} V^+(y;\eta) = \sup_{\eta\in G(y)} \min_{z\in\mathcal{P}_K(y)} \langle y-z,\eta\rangle$$

$$\le \min_{z\in\mathcal{P}_K(y)} \sup_{\eta\in G(z+l)} \langle l,\eta\rangle \le -\rho\text{dist}(y,K)^2 = -2\rho V(y).$$

Thus (18) is implied by (19), which appears to be a sufficient condition for *global* strong (exponential) asymptotic stability of K with respect to (17).

Returning to the associated inclusion (9) we come up with the following condition.

Condition B. There is a positive constant ρ such that for each fixed $(t,x) \in [0,T] \times D$, for every $y \in \partial K(t,x)$ and for every $l \in N_{K(t,x)}^\perp(y)$

$$\max_{\eta\in\hat{F}(t,x,y+l)} \langle l,\eta\rangle \le -\rho|l|^2. \tag{20}$$

We summarize the above conclusions in the next proposition.

Proposition 3 *Suppose that* **A1** *(Section 3) is fulfilled and that $K(t,x) \subset K_0(t,x)$ is non-empty, closed and bounded, uniformly in $t \in [0,T]$, $x \in D$. Then Condition* **B** *implies the stability conditions* **A4** *and* **A5**.

Condition **B**, being sufficient for strong stability, implies also strong invariance of $K(t,x)$ with respect to the associated inclusion (12). Notice, that the necessary and sufficient condition for strong invariance of $K(t,x)$ in "proximal" terms is

$$\max_{\eta \in \hat{F}(t,x,y)} \langle l, \eta \rangle \leq 0 \quad \forall l \in N^{\perp}_{K(t,x)}(y)$$

[17, Theorem 2.1], [5, Theorem 3.1]. Since $K(t,x)$ consists of equilibrium points only, in fact

$$\max_{\eta \in \hat{F}(t,x,y)} \langle l, \eta \rangle = 0.$$

Thus for a strongly invariant set $K(t,x)$ of equilibrium points of $\hat{F}(t,x,\cdot)$ Condition **B** is equivalent to

$$\max_{\eta \in \hat{F}(t,x,y+l)} \langle l, \eta \rangle - \max_{\eta \in \hat{F}(t,x,y)} \langle l, \eta \rangle \leq -\rho |l|^2 \quad \forall y \in \partial K(t,x),\ \forall l \in N^{\perp}_{K(t,x)}(y).$$

Clearly, the last inequality is implied by the following condition introduced in [11] for investigation of the upper semicontinuity of Z_ϵ: for every $t \in [0,T]$ and $x \in D$

$$\max_{\eta'' \in \hat{F}(t,x,y'')} \langle y'' - y', \eta'' \rangle - \max_{\eta' \in \hat{F}(t,x,y')} \langle y'' - y', \eta' \rangle \leq -\rho |y'' - y'|^2 \quad \forall y', y'' \in \mathbf{R}^n. \qquad (21)$$

Proposition 4 *Suppose that* **A1** *(Section 3) is fulfilled. Let condition (21) be fulfilled. Then $K_0(t,x)$ is non-empty and bounded (uniformly in $(t,x) \in [0,T] \times D$). If, in addition, $K_0(t,x)$ is strongly invariant with respect to the associated inclusion (12), then the stability conditions* **A4** *and* **A5** *are fulfilled.*

Proof. The proof uses ideas from [11]. Condition (21) applied for $y' = 0$, $y'' = y \in \mathbf{R}^n$ gives (suppressing (t,x) in the notations)

$$\max_{\eta \in \hat{F}(y)} \langle y, \eta \rangle - \max_{\eta' \in \hat{F}(0)} \langle y, \eta' \rangle \leq -\rho |y|^2,$$

$$\max_{\eta \in \hat{F}(y)} \langle y, \eta \rangle \leq \max_{\eta' \in \hat{F}(0)} |\eta'||y| - \rho |y|^2 \leq 0 \qquad (22)$$

for $|y| \geq C$ – sufficiently large. Since $N^{\perp}_{C\mathbf{B}}(y) = \text{cone}\{y\}$ for $|y| = C$, (22) implies

$$\max_{\eta \in \hat{F}(y)} \langle y, \eta \rangle \leq 0 \quad \forall\, l \in N^{\perp}_{C\mathbf{B}}(y),$$

which means that $C\mathbf{B}$ is weakly invariant (in fact, even strongly) with respect to $\hat{F}$ [27, Theorem 3.1] (see also [26, 5]. Then $\hat{F}$ has an equilibrium point in $C\mathbf{B}$ [2, Chapter 5.2] and $K_0 \neq \emptyset$. If $y \in K_0$, then (22) implies also

$$\rho|y|^2 \leq M|y| - \max_{\eta \in \hat{F}(y)} \langle y, \eta \rangle \leq M|y|$$

(M is a bound of $\hat{F}(0)$), hence $|y| \leq M/\rho$. Thus K_0 is non-empty and bounded. The last claim of the proposition is a consequence of Proposition 3 as far as it was shown that (21) implies Condition $\mathbf{B}$ if $K(t,x)$ is strongly invariant. Q.E.D.

Thus (21) together with strong invariance of the set $K_0(t,x)$ with respect to the associated inclusion implies upper semicontinuity of Z_ε in the Tikhonov metric. It was proven in [11] that (21) together with convexity of the graph of F with respect to y (a condition close to linearity, but essential) imply upper semicontinuity of Z_ε in the ($C \times L_1$-weak) topology. Thus, given (21), two quite different conditions arise in studying the upper semicontinuity of Z_ε in different topologies: either strong invariance of $K_0(t,x)$, or, alternatively, convexity of the graph of $F(t,x,\cdot)$.

We mention, that (20) is essentially weaker than (21) as the following example shows.

Example 2. Consider for $x \in \mathbf{R}^n$, $y \in \mathbf{R}^1$

$$\dot{x} \in F_1(t,x,y), \;\; x(0) = x^0,$$
$$\varepsilon\dot{y} \in -y^3 + y + [-d,d], \;\; y(0) = y^0.$$

Suppose that $d > 2/3\sqrt{3}$, so that, as trivial calculation show, $K_0(x) = K_0 = [-k(d), k(d)]$ is an interval and $k(d) > 1/\sqrt{3}$. In this case (21) is not fulfilled, say, for $y' = 0$ and for y'' – sufficiently small. Nevertheless, Supposition $\mathbf{B}$ is easily verifiable. Indeed, one has to check it only for $y \in \{-k(d), k(d)\}$ and, for example, for $y = k(d)$ we have $N^{\perp}_{K_0}(y) = \{l \in \mathbf{R}\,;\, l \geq 0\}$. For $l > 0$

$$\max_{\eta \in \hat{F}(t,x,y+l)} \langle l, \eta \rangle \leq -(3k(d)^2 - 1)l^2 = -\rho|l|^2, \;\; \rho > 0.$$

This example is remarkable in one more respect. Often in control theoretic considerations, where

$$\hat{F}(y) = \{f(y,u)\,;\, u \in U\},$$

in order to ensure certain continuity of the trajectory bundle one requires that $f(\cdot,u)$ has a unique *stable* equilibrium for each $u \in U$. This is not the case in Example 2, since for $u = 0$ (and for many others $u \in U$) the corresponding equation $\dot{y} = -y^3 + y$ has three equilibrium points, one of which — $u = 0$ — is unstable. Nevertheless, Z_ε is not only upper semicontinuous in the Tikhonov metric (according to Theorem 1), but also turns out to be lower semicontinuous in the ($C \times L_2$)-metric (see Section 6 or [28]).

Condition B is a relatively strong one, especially the implicit requirement that $K_0(t,x)$ is strongly invariant, but that is the price for the upper semicontinuity in the Tikhonov metric. We shall elaborate in more details the case

$$\hat{F}(t,x,y) = f(t,x,y) + C(t,x)U, \quad U \subset \mathbf{R}^r \text{ is convex and compact.} \tag{23}$$

Suppose that f is continuously differentiable in y and that the derivative $(\partial f)/(\partial y)$ is invertible. Suppose also that the equation

$$0 = f(t,x,y) + C(t,x)u$$

is solvable for each $t \in [0,T]$, $x \in D$ and $u \in U$ and that the solution $\xi(t,x,u)$ is unique.

In order to interpret Condition B in this special case we remind some notions and properties from the set-valued analysis. For $y \in K_0$

$$T_{K_0}(y) = \{p \in \mathbf{R}^n \,;\, \liminf_{h \to 0+} \frac{1}{h} \operatorname{dist}(y + hp, K_0) = 0\}$$

is the *contingent cone* to K_0 at y. Its polar cone

$$N_{K_0}(y) = \{l \in \mathbf{R}^n \,;\, \langle l, p \rangle \le 0 \; \forall p \in T_{K_0}(y)\}$$

is known to satisfy

$$N^{\perp}_{K_0}(y) \subset N_{K_0}(y). \tag{24}$$

Using the obvious continuous differentiability of $\xi(u)$ (we suppress again (t,x) in the notations) and the relation $K_0 = \xi(U)$ one easily obtains that for $y = \xi(u)$

$$-\left[\frac{\partial f}{\partial y}(y)\right]^{-1} C T_U(u) \subset T_{K_0}(y).$$

Then in view of (24)

$$N^{\perp}_{K_0}(y) \subset N(y) \stackrel{def}{=} \{l \,;\, \langle l, -\left[\frac{\partial f}{\partial y}(y)\right]^{-1} C\zeta \rangle \le 0 \;\; \forall \zeta \in T_U(u)\}.$$

Then it is enough to ensure that (20) is fulfilled for every $l \in N(y)$.

Take arbitrary $y \in K_0$, $l \in N(y)$ and $\eta \in \hat{F}(y + l)$. For some $u, v \in U$ we have

$$\begin{aligned} 0 &= f(y) + Cu, \\ \eta &= f(y + l) + Cv. \end{aligned}$$

Subtracting and multiplying by l we obtain

$$\langle l, \eta \rangle = \langle l, f(y + l) - f(y) \rangle + \langle l, C(v - u) \rangle.$$

Then, in order to ensure (20) it suffices to suppose

$$\langle f(t,x,y_2) - f(t,x,y_1), y_2 - y_1 \rangle \le -\rho |y_2 - y_1|^2, \;\; \forall\, y_1 \in K_0(t,x),\, y_2 \in \mathbf{R}^n \tag{25}$$

and, in addition

$$\langle l, C(v-u)\rangle \le 0 \;\; \forall u, v \in U,\, \forall l \in N(\xi(u)).$$

Taking into account the definition of $N(y)$ and since $v - u \in T_U(u)$ we conclude that the last inequality is implied by

$$CT_U(u) \subset -\left[\frac{\partial f}{\partial y}(y)\right]^{-1} CT_U(u).$$

Thus we proved the following.

Proposition 5 *Suppose that $\hat{F}$ has the form (23) and that the suppositions formulated next to (23) are fulfilled. Then the stability suppositions* **A4** *and* **A5** *of Theorem 1 are implied by (25) together with the condition*

$$-\frac{\partial f}{\partial y}(t,x,\xi(t,x,u))C(t,x)T_U(u) \subset C(t,x)T_U(u) \tag{26}$$

for all $t \in [0,T]$, $x \in \mathbf{R}^m$ and $u \in U$.

We mention that condition (25) ensures stability of the drift term f, while (26) turns out to be a sufficient condition (together with (25)) for strong invariance of $K_0(t,x)$ with respect to the associated inclusion (cf. [17, Theorem 2.1], or [5, Theorem 3.1]).

If f is continuously differentiable in y then condition (25) is implied by strict negative definiteness of $\frac{\partial f}{\partial y}$ on the manifold $\{(t,x,\xi(t,x))\}$. The last condition is commonly used in the singular perturbation literature from its very beginning, as well as in the control theoretical context (see e.g. [16, 4] and Remark 1 below).

Remark 1. The stability condition (25) reflects the particular choice of the Lyapunov function $V(y) = 0.5\text{dist}(y,K)^2$. A weaker condition can be obtained by using the same Lyapunov function but with respect to a more general metric in $\mathbf{R}^n$. For example, for a linear fast differential inclusion

$$\varepsilon\dot{y} \in A(t,x)y + U(t,x).$$

condition (25) requires (uniformly in (t,x)) negative definiteness of $A(t,x)$ while, in fact, suppositions **A4** and **A5** are ensured by

$$\operatorname{re}\sigma(A(t,x)) \le \gamma < 0,$$

where $\sigma(A)$ are eigenvalues of A. To show this one may take as a Lyapunov function $V(y) = \langle P(t,x)y, y\rangle$ with appropriate P (a symmetric positive definite solution of the corresponding algebraic Lyapunov equation).

As a particular case of Proposition 5 we consider a set-valued mapping $\hat{F}$ defined as

$$\hat{F}(t,x,y) = \left\{ \begin{pmatrix} f^1(t,x,y_1) + c^1(t,x)u_1 \\ \vdots \\ f^n(t,x,y_n) + c^n(t,x)u_n \end{pmatrix} ;\ u_i \in [-1,1] \right\} \tag{27}$$

In this case condition (25) requires that $\partial f^i/\partial y_i \leq -\sigma < 0$ for each t, x, y, while (26) is automatically fulfilled as a consequence of the last inequality. Theorem 1 is applicable.

As an application one can interpret the fast system

$$\varepsilon \dot{y} = f(t,x,y) + C(t,x)u, \tag{28}$$

(f and C are defined in accordance with (27)) as a dynamic sensor model: $y_1, \ldots, y_n$ are dynamic sensors tracking different outputs under perturbations $u_1, \ldots, u_n$. The perturbations are *independent* of each other in the sense that $U = [0,1]^n$ is a Cartesian product. Consider a closed-loop system

$$\dot{x} = f_1(t,x,y,v), \quad v \in V, \tag{29}$$

fed back by the output y and disturbed by v. Denote by $x_\varepsilon[u,v], y_\varepsilon[u,v]$ the trajectory of (29),(28) corresponding to disturbances u and v. Let the task be to evaluate the performance

$$J_\varepsilon = \max_{u(\cdot),v(\cdot)} [g(x[u,v](T)) + \alpha \max_{t \in [t_0,T]} |y[u,v](s) - \hat{y}(s)|],$$

where $t_0 > 0$ and $\hat{y}(\cdot)$ is a given reference observation (instead of "max" one can take integral on $[0,T]$ in the performance function). Could one neglect the sensor dynamics in this problem? That is, could one approximate J_ε by the value J_0 solving the simpler problem corresponding to $\varepsilon = 0$. The answer will be positive if the trajectory bundle of (29),(28) is continuous at $\varepsilon = 0$ in the Tikhonov metric. The lower semicontinuity issue in the considered example is simpler. The upper semicontinuity follows (supposing (25)) from Theorem 1 and the above considerations.

We mention also that one can interpret v as a control variable in the above model and consider a min-max optimal control problem with performance like J_ε, but taking "min" with respect to v. The sensor dynamics can also be neglected without significant loss of performance (if ε is small). We stress the fact that this is not always the case in control problems even when the stability condition (25) is satisfied. Neglecting the dynamics in a sensor model like (28) may totally corrupt the result. As an example, which also illustrates the role of condition (26), we consider the following feedback tracking problem.

Example 3. Consider a two-dimensional control system

$$\begin{aligned} \dot{x}_1 &= 2ux_1 - 2ux_2 - u + 1, \\ \dot{x}_2 &= 2(u-1)x_1 - 2(1-u)x_2 + u, \quad u \in [0,1] \end{aligned}$$

for which two observations $y_1(t)$ and $y_2(t)$ are available at the current moment t. Let the sensor equations be

$$\begin{aligned} \varepsilon \dot{y}_1 &= -y_1 + x_1 + w, \\ 0.5\varepsilon \dot{y}_2 &= -y_2 + x_2 + w, \end{aligned}$$

where $w \in [-8, 8]$ is a perturbation (the value 8 is taken just for convenience). Suppose that the task is to design an output feedback regulator $u = u(y_1(t), y_2(t))$ such that the closed loop system has a prescribed behaviour. Namely, we want to ensure that any trajectory of the closed-loop system that starts from the set

$$M = \{(x_1, x_2)\,;\; x_2 \in [x_1 - 0.5, x_1 + 0.5]\}$$

remains in this set. A "good" idea is to simplify the problem by neglecting the sensors dynamics. Supposing instantaneous response of the sensor ($\varepsilon = 0$) we come to the static sensor model

$$y_1 = x_1 + w, \quad y_2 = x_2 + w.$$

It is easy to verify that supposing y_1, y_2 to satisfy the last (degenerate) equations, the set-valued output feedback low

$$u(y_1, y_2) = \begin{cases} 0 & \text{if } y_1 < y_2, \\ [0,1] & \text{if } y_1 = y_2, \\ 1 & \text{if } y_1 > y_2 \end{cases}$$

solves the problem. One may expect that if the actual observation comes from the dynamic sensor (with "fast" response ε), then the same feedback control will still "approximately work", namely, that any trajectory starting from M will stay $O(\varepsilon)$-near M at least on a finite time-interval. This is not the case. It can be shown that for an appropriately oscillating disturbance w (in the spirit of Example 1) the corresponding trajectory of the closed-loop system starting from the point $(0, -0.5) \in M$ has $\text{dist}((x_1(1), x_2(1)), M) \geq 0.25$, no matter how small is ε. The reason for this irregularity is that condition (26) is not satisfied ((25) is apparently satisfied).

5 Upper semicontinuity in the $C \times (L_1$-weak) topology

In this section we investigate the convergence of Z_ε in the $C \times (L_1$-weak) topology. Theorem 1 implies upper semicontinuity of Z_ε also in this topology, but under the condition of strong asymptotic stability of the equilibrium set K_0 with respect to the associated inclusion. Below we relax this condition requiring strong asymptotic stability of the *invariance envelope* of K_0. However, in addition we require (as in [11]) convexity of the graph of F with respect to y.

Consider (9), (10) supposing **A1**. As before we denote

$$K_0(t,x) = \{y\,;\ 0 \in \mathcal{P}_y F(t,x,y)\}.$$

Let $K(t,x)$ be the *strongly invariant envelope* of $K_0(t,x)$ with respect to the associated inclusion (12). That is, $K(t,x)$ is the minimal closed set containing $K_0(t,x)$ and such that every trajectory of (12) starting from it remains in it. In fact, $K(t,x)$ is the closure of the reachable set (with free end-time) of (12) starting from $K_0(t,x)$.

Supposition A2'. $K_0(t,x) \neq \emptyset$ for every $(t,x) \in [0,T] \times D$ and the invariance envelope $K : [0,T] \times D \Rightarrow \mathbf{R}^n$ is compact valued and continuous.

Furthermore, we suppose also that **A3**–**A5** are fulfilled for the so-defined K. Instead of the structural condition **A6** we assume now

Supposition A6'. For each $(t,x) \in [0,T] \times D$ the graph of $F(t,x,\cdot)$ is convex.

Theorem 2 *Let suppositions* **A1**, **A2'**, **A3**, **A4**, **A5** *and* **A6'** *be fulfilled. Then the mapping* $\varepsilon \longrightarrow Z_\varepsilon$ *is upper semicontinuous at* $\varepsilon = 0$ *with respect to the* $C \times (L_1$*-weak) topology.*

Proof. Upper semicontinuity is implied by the following property: for arbitrary sequences $\varepsilon_k \to 0$, $(x_k(\cdot), y_k(\cdot)) \in Z_{\varepsilon_k}$ there is a subsequence converging (in the specified topology) to some $(x_0(\cdot), y_0(\cdot)) \in Z_0$. We shall prove the last property.

Proposition 1 implies that $x_k(\cdot)$ and $y_k(\cdot)$ are bounded in C, hence the sequence $\dot{x}_k(\cdot)$ is also bounded. Thus one can extract a subsequence (we do not change the indexation) $(x_k(\cdot), y_k(\cdot))$ converging to some $(x_0(\cdot), y_0(\cdot))$ in the $C \times (L_1$-weak) topology. It remains to prove that $(x_0(\cdot), y_0(\cdot)) \in Z_0$.

The key point to do this is to observe that $\varepsilon_k \dot{y}_k(\cdot)$ converges to zero in the L_1-weak topology. Indeed, as far as the sequence $y_k(\cdot)$ is bounded in C we have for each $t \in [0,T]$

$$\int_0^T \varepsilon \dot{y}_k(s)\,ds = \varepsilon \int_0^T \dot{y}_k(s)\,ds = \varepsilon(y_k(T) - y_k(0)) \to 0.$$

The proof can be completed as that of Theorem 1 in [11], but here we give a somewhat simpler argument. Applying Mazur's theorem we find for every k a finite collection $\alpha_i^k \geq 0$, $i \geq k$, $\sum_{i \geq k} \alpha_i^k = 1$, such that (again for subsequences)

$$\psi_k(t) \stackrel{def}{=} \sum_{i \geq k} \alpha_i^k \dot{x}_i(t) \to \dot{x}_0(t), \quad \eta_k(t) \stackrel{def}{=} \sum_{i \geq k} \alpha_i^k \varepsilon_i \dot{y}_i(t) \to 0, \quad \tilde{y}_k(t) \stackrel{def}{=} \sum_{i \geq k} \alpha_i^k y_i(t) \to y_0(t)$$

for a.e. $t \in [0,T]$. Take an arbitrary $\delta > 0$. Then

$$\begin{pmatrix} \dot{x}_i(t) \\ \varepsilon \dot{y}_i(t) \end{pmatrix} \in F(t, x_0(t), y_i(t)) + \delta \mathbf{B}$$

for a.e. t and all sufficiently large i. This means that

$$(y_i(t), \dot{x}_i(t), \dot{y}_i(t))^* \in \text{graph}(F(t, x_0(t), \cdot) + \delta \mathbf{B})$$

(* stands for transposition) and since the set in the right-hand side is convex, also

$$(\tilde{y}_k(t), \psi_k(t), \eta_k(t))^* \in \text{graph}(F(t, x_0(t), \cdot) + \delta \mathbf{B}).$$

Hence,

$$\begin{pmatrix} \psi_k(t) \\ \eta_k(t) \end{pmatrix} \in F(t, x_0(t), \tilde{y}_k(t)) + \delta \mathbf{B}.$$

Passing to the limit and then taking into account that δ was arbitrarily chosen we complete the proof. Q.E.D.

Remark 2. The supposition that $K(t,x)$ is the strongly invariant envelope of $K_0(t,x)$ was not used in the proof and is not essential. However, the generality that this remark brings can be seen to be somewhat illusionary, taking into account the stability condition.

In order to elaborate the stability condition **A4** for the particular choice of K as the strongly invariant envelope of K_0 we need the following result, which extends [17, Theorem 1.2].

Lemma 1 *Let $P_0 \subset \mathbf{R}^n$ be closed and let P be its strongly invariant envelope with respect to the differential inclusion*

$$\dot{y} \in G(y), \tag{30}$$

where G is locally Lipschitz. Then for every $y \in P$

$$\max_{\eta \in G(y)} \langle l, \eta \rangle \leq 0 \quad \forall\, l \in N_P^{\perp}(y) \tag{31}$$

and for every $y \in P \setminus P_0$

$$\max_{\eta \in G(y)} \langle l, \eta \rangle = 0 \quad \forall\, l \in N_P^{\perp}(y) \tag{32}$$

As a consequence, if P_0 consists only of equilibrium points of G, then (32) is fulfilled for all $y \in P$.

Proof. The inequality (31) is claimed by Theorem 3.1 in [5] – it is a necessary and sufficient condition for strong invariance, obtained independently also in [17].

Let $y \in P \setminus P_0$. Suppose that (32) is not true, that is,

$$\max_{\eta \in G(y)} \langle l, \eta \rangle \leq -\delta < 0$$

for some $l \in N_P^{\perp}(y)$, $|l| = 1$. Since P is the closure of the reachable set of (30) there is a sequence of trajectories $y_k(\cdot)$ such that

$$y_k(0) \in P_0, \quad y_k = y_k(t_k) \longrightarrow y,$$

where $t_k \geq 0$. If $t_k \to 0$, then we easily obtain $y \in P_0$ since G is locally Lipschitz. This contradicts the assumption, therefore $t_k \geq \tau$ for some $\tau > 0$. Then $z_k = y_k(t_k - \tau_k) \in P$ for all sufficiently large k, where $\tau_k = \sqrt{h_k} \stackrel{def}{=} \sqrt{|y_k - y|}$. We have

$$y_k - z_k \in \int_{t_k-\tau_k}^{t_k} G(y_k(s))\, ds \subset \int_{t_k-\tau_k}^{t_k} G(y)\, ds + \tau_k L \max_{s \in [t_k-\tau_k, t_k]} |y_k(s) - y| \mathbf{B},$$

where L is the Lipschitz constant of G in a neighborhood of y. Multiplying by l we obtain for an appropriate constant c_1

$$\langle l, y_k - z_k \rangle \leq -\delta \tau_k + c_1(h_k + \tau_k)\tau_k.$$

and

$$\langle l, y - z_k \rangle \leq -\delta \tau_k + c_1(h_k + \tau_k)\tau_k + h_k \leq -\delta \tau_k + c_2 \tau_k^2. \tag{33}$$

Let q be a condensation point of $q_k \stackrel{def}{=} (z_k - y)/|z_k - y|$. Since

$$\alpha_k = |z_k - y| \leq h_k + c_3 \tau_k \leq c_4 \tau_k \to 0,$$

and $y + \alpha_k q_k \in P$, we have $q \in T_P(y)$ – the contingent cone to P at y defined in Section 4. According to (24) applied to the set P we have $\langle l, q \rangle \leq 0$. Hence,

$$\langle l, z_k - y \rangle = \alpha_k \langle l, q \rangle + o(\alpha_k) \leq o(\tau_k), \quad \frac{o(\tau_k)}{\tau_k} \to 0.$$

Combining with (33) we obtain

$$-o(\tau_k) \leq -\tau_k \delta + c_2 \tau_k^2,$$

which leads to the contradiction $\delta \leq 0$. Q.E.D.

Let us return to the stability conditions **A4** and **A5**. In Section 4 we showed that a sufficient condition for **A4** and **A5** is the existence of a positive ρ such that

$$\max_{\eta \in \hat{F}(t,x,y+l)} \langle l, \eta \rangle \leq -\rho |l|^2. \tag{34}$$

for every $(t, x) \in [0, T] \times D$, $y \in K(t, x)$ and $l \in N_{K(t,x)}^{\perp}(y)$. As before, $\hat{F} = \mathcal{P}_y F$. (See also Remark 1 in Section 4 for possible use of more general Lyapunov functions.)

Since the set $K(t,x)$ is the strongly invariant envelope of the set of equilibriums of (12), the above lemma gives

$$\max_{\eta\in\hat{F}(t,x,y)} \langle l,\eta\rangle = 0$$

for every t, x, y and l as in (34). Thus, denoting $\sigma(P|l) = \max_{p\in P}\langle l,p\rangle$ (the support function of the closed set P), condition (34) turns out to be equivalent to

$$\sigma(F_2(t,x,y'')|y''-y') - \sigma(F_2(t,x,y')|y''-y') \leq -\rho|y''-y'|^2 \tag{35}$$

for every $(t,x)\in[0,T]\times D$, $y'\in K(t,x)$ and $y''\in\mathbf{R}^n$ such that $y''-y'\in N^{\perp}_{K(t,x)}(y')$. This is an essentially weaker form of the condition introduced in [11], the latter requiring that (35) is fulfilled for all $y', y''\in\mathbf{R}^n$ (see Example 2 in Section 4).

We outline the difference between the stability conditions in this and in the preceding sections by a "fast" inclusion in the form of (23). For upper semicontinuity in the Tikhonov metric (Theorem 1) we require asymptotic stability of $K_0(t,x)$, which in this case amounts of stability of the drift term (25) and strong invariance of $K_0(t,x)$ (26). In the present section the set $K(t,x)$ is strongly invariant by definition and **A4**, **A5** are implied by (25) alone. Some examples follow.

Example 4. In the next three systems $y=(y_1,y_2)\in\mathbf{R}^2$:

$\dot{x} \in F_1(x,y)$	$\dot{x} \in F_1(x,y)$	$\dot{x} \in F_1(x)+C(x)y+D(x)u$
$\varepsilon\dot{y}_1 = -y_1+g_1(x)+u$	$\varepsilon\dot{y}_1 = -y_1+g_1(x)+u$	$\varepsilon\dot{y}_1 = -y_1+g_1(x)+u$
$\varepsilon\dot{y}_2 = -y_2+g_2(x)+u$	$\varepsilon\dot{y}_2 = -2y_2+g_2(x)+v$	$\varepsilon\dot{y}_2 = -2y_2+g_2(x)+u$
$u\in[-1,1]$	$u,v\in[-1,1]$	$u\in[-1,1]$.

The drift terms in all of these examples are obviously stable in the sense of (25). Proposition 5 is applicable to the first two systems – the second is in the form of (27), while (26) can easily be checked for the second one. Thus one can claim upper semicontinuity of Z_ε in the Tikhonov metric. The third system does not satisfy **A4** since $K_0=\{(y_1,y_2);\ y_1 = 2y_2\in[-1,1]\}$ is not strongly invariant with respect to the associated system

$$\begin{aligned}\dot{y}_1 &= -y_1+g_1(x)+u\\ \dot{y}_2 &= -2y_2+g_2(x)+u\\ &\quad u\in[-1,1],\ t\in[0,+\infty).\end{aligned}$$

The invariant envelope of K_0 for this system can be explicitly found, but what is actually essential is that it is bounded. Then Theorem 2 can be applied as far as the overall system is linear in y (thus has a convex graph). Here only L_1-weak convergence of the "fast" variables can be claimed, but upper semicontinuity of Z_ε in the Tikhonov metric really does not take place.

We mention that in Example 1 considered in Section 2 the trajectory bundle Z_ε is not upper semicontinuous at $\varepsilon=0$ even in the $C\times(L_1$-weak) topology. The reason is that the corresponding differential inclusion does not have a convex graph. The last condition is essential. On the other hand, in Example 1 the family of the slow trajectories $X_\varepsilon = \mathcal{P}_x Z_\varepsilon$ is upper semicontinuous in C. In the next example (analysed in detail in [8]) even X_ε is not upper semicontinuous. in C at $\varepsilon = 0$.

Example 5.

$$\begin{aligned} \dot{x} &= |y_1 - 2y_2| \\ \varepsilon \dot{y}_1 &= -y_1 + u \\ \varepsilon \dot{y}_2 &= -2y_2 + u \\ & \quad u \in [-1, 1]. \end{aligned}$$

Here the right-hand side of the first equation can be replaced with the smooth $(y_1 - 2y_2)^2$ or even by the bilinear $u(y_1 - 2y_2)$. X_ε is still not upper semicontinuous and the reason is again the nonconvexity of the graph of F.

6 Lower semicontinuity in the $C \times L_1$ metric

In this section we investigate the lower semicontinuity of the set of trajectories Z_ε of a singularly perturbed differential inclusion. Similarly as in Supposition **A6** (Section 3), in the case of a singularly perturbed control system it is reasonable to make difference between the controls acting independently either in the slow or in the fast subsystem and those acting in both of them. For this reason we consider the following more detailed representation of (7),(8):

$$\begin{aligned} \dot{x}(t) &= f_1(t, x, y, u, v_1), \;\; x(0) = x^0, \\ \varepsilon \dot{y}(t) &= f_2(t, x, y, u, v_2), \;\; y(0) = y^0, \end{aligned}$$

where $u \in U$, $v_1 \in V_1$, $v_2 \in V_2$. In other terms, we suppose in advance that the right-hand side of differential inclusion (9) has the form

$$F(t, x, y) = \left\{ \begin{pmatrix} \xi \\ \eta \end{pmatrix} \; ; \; \xi \in F_1(t, x, y, u), \; \eta \in F_2(t, x, y, u), \; u \in U \right\},$$

where $F_1 : [0, T] \times \mathbf{R}^m \times \mathbf{R}^n \times U \Rightarrow \mathbf{R}^m$, $F_2 : [0, T] \times \mathbf{R}^m \times \mathbf{R}^n \times U \Rightarrow \mathbf{R}^n$, $U \subset \mathbf{R}^r$. Formally, taking set-valued F_1 and F_2 instead of single-valued ones (like in (7),(8)) does not increase the generality, but will allow to formulate the stability requirement below in an essentially weaker way. In control terms, certain stability of f_2 will be supposed for all fixed values of u, but not for all values of v_1 and v_2.

The following will be required.

Supposition C1. F_1 and F_2 are non-empty convex compact valued, locally bounded and locally Lipschitz with respect to (x, y, u), uniformly in t; F_1 is measurable in t, F_2 is continuous; the set U is convex and compact.

Supposition C2. There is a compact set $D \subset \mathbf{R}^m$ such that for each $t \in [0,T]$, $x \in D$ and $u \in U$ the set

$$\hat{K}_0(t,x,u) = \{y\,;\; 0 \in F_2(t,x,y,u)\}$$

is non-empty; the mapping $\hat{K}_0$ is convex compact valued, continuous and Lipschitz continuous in x.

Supposition C3. All solutions of the differential inclusion

$$\dot{x}(t) \in \mathrm{co}\,\{F_1(t,x,\hat{K}_0(t,x,u),u)\,;\; u \in U\}, \;\; x(0) = x^0$$

remain in the interior of D on $[0,T]$.

Supposition C4. (Weak asymptotic stability of each point of $\hat{K}_0$.) For every $\mu > 0$ there exists $\delta = \delta(\mu) > 0$ and for every $\nu > 0$ there exists $\tau_0 = \tau_0(\delta,\nu)$ such that for every $(t,x,u) \in [0,T] \times D \times U$, for every $y \in \hat{K}_0(t,x,u)$ and for every $y_0 \in y + \delta\mathbf{B}$ there exists a solution $\tilde{y}(\cdot)$ of the associated inclusion

$$\dot{\tilde{y}}(\tau) \in F_2(t,x,\tilde{y}(\tau),u), \;\; \tilde{y}(0) = y^0, \tag{36}$$

on $[0,+\infty)$ such that

$$|\tilde{y}(\tau) - y| \le \mu \;\; \forall \tau \ge 0 \quad \text{and} \quad |\tilde{y}(\tau) - y| \le \nu \;\; \forall \tau \ge \tau_0.$$

Supposition C5. There is a compact set $G_0 \subset \mathbf{R}^n$ such that for every $\nu > 0$ there is $\tau_0 = \tau_0(\nu)$ such that for each $u \in U$ and $y \in \hat{K}_0(0,x^0,u)$ there is a solution $\tilde{y}^0(\cdot)$ of

$$\frac{d}{d\tau}\tilde{y}^0(\tau) \in \hat{F}(0,x^0,\tilde{y}^0(\tau)), \;\; \tilde{y}^0(0) = y^0$$

on $[0,\tau_0)$ which satisfies

$$\tilde{y}^0(\tau) \in G_0 \;\; \forall \tau \ge 0 \quad \text{and} \quad |\tilde{y}^0(\tau_0) - y| \le \nu.$$

Theorem 3 *Suppose that* **C1–C5** *are fulfilled and that the set Z_0 is non-empty. Then the set Z_ε is non-empty for all sufficiently small $\varepsilon > 0$ and the mapping $\varepsilon \longrightarrow Z_\varepsilon$ is lower semicontinuous in $C \times L_1$.*

The proof will be given in Appendix B.

We remind that the lower semicontinuity claim of the above theorem means that for every $z_0 \in Z_0$ and for every sequence $\varepsilon_k \to 0$ there are corresponding $z_k \in Z_{\varepsilon_k}$ such that

$$x_k(\cdot) \to x_0(\cdot) \;\; \text{in } C[0,T]$$

$$y_k(\cdot) \to y_0(\cdot) \;\; \text{in } L_1[0,T].$$

Equivalently, for every $z_0 \in Z_0$

$$\text{dist}(z_0, Z_\varepsilon) \to 0 \quad \text{with} \quad \varepsilon \to 0,$$

where dist is in the $C \times L_1$ metric.

Similarly as in Section 4 the stability conditions **C4** and **C5** can be elaborated in terms of Lyapunov functions. In particular, taking the same Lyapunov function $V(y,\tau) = 0.5e^{-2\rho\tau}\text{dist}(y, \hat{K}_0(t,x,u))$ one comes up with the condition

$$\max_{\eta \in F_2(t,x,y,u)} \langle y - \bar{y}, \eta \rangle \le -\rho |y - \bar{y}|^2 \quad \forall \bar{y} \in \hat{K}_0(t,x,u)\, \forall y \in \mathbf{R}^n,$$

which is sufficient for **C4** and **C5**.

The convexity assumption about $\hat{K}_0$ might be somewhat superfluous, as the result in [28] indicates. However, the nonconvex case requires a more profound set-analytic techniques and could be a subject of further investigation.

Appendix A

In the subsequent results and proofs we use the notations introduced in the main text. The proofs of the next lemma and proposition go along the line of proof of the Tikhonov theorem given in [24]. However, the set-valued case requires corresponding changes (like the use of the Filippov theorem) and we present the detailed proofs for completeness.

Lemma 2 *Suppose that* **A1**, **A2** *and* **A4** *are fulfilled (except that (13) and Lipschitz continuity of $K(t,\cdot)$ need not be required). Then for every $\mu > 0$ there exists $\varepsilon_0 = \varepsilon_0(\mu) > 0$ such that for every $t_0 \in [0,T]$, $x_0 \in \text{int}\, D$ and y_0 for which $\text{dist}(y_0, K(t_0,x_0)) < \delta(\mu/2)$ (see Supposition* **A4**) *and for every $\varepsilon \in (0, \varepsilon_0]$, each solution $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ of (9) with initial condition $x_\varepsilon(t_0) = x_0$, $y_\varepsilon(t_0) = y_0$ is extendible and satisfies*

$$\text{dist}(y_\varepsilon(t), K(t, x_\varepsilon(t))) < \mu \tag{37}$$

at least as long as $x_\varepsilon(t) \in \text{int}\, D$ and $t \le T$.

Proof. Supposition **A1** guarantees local extendibility of the solutions of (9). Therefore, it is enough to prove that for any solution $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ starting from (x_0, y_0) at t_0, the inequality (37) holds as long as $x_\varepsilon(t) \in \text{int}\, D$ and $t \le T$.

Assume that for some $\mu > 0$ a number $\varepsilon_0 = \varepsilon_0(\mu) > 0$ for which the last property holds does not exist. Then there are sequences $\{\varepsilon_k\} \to 0$, $t_0^k \in [0,T)$, $x_0^k \in \text{int}\, D$, $y_0^k \in K(t_0^k, x_0^k) + \delta\mathbf{B}$

(we set for brevity $\delta = \delta(\mu/2)$) and corresponding solutions $(x_{\varepsilon_k}(\cdot), y_{\varepsilon_k}(\cdot))$ of (9) starting from (t_0^k, x_0^k, y_0^k) such that for each k the inequality (37) fails at some moment t while $x_{\varepsilon_k}(\cdot)$ is still in int D. Because of the continuity of K, $x_{\varepsilon_k}(\cdot)$ and $y_{\varepsilon_k}(\cdot)$ this implies existence of $\bar{t}_k > t_0^k$ such that

$$\operatorname{dist}(y_{\varepsilon_k}(t), K(t, x_{\varepsilon_k}(t)) < \mu, \quad t \in [t_0^k, \bar{t}_k) \tag{38}$$

(notice that $\delta(\mu/2) \leq \mu/2 < \mu$ by the sense of Supposition **A4**, therefore $\bar{t}_k > t_0^k$) and

$$\operatorname{dist}(y_{\varepsilon_k}(\bar{t}_k), K(\bar{t}_k, x_{\varepsilon_k}(\bar{t}_k)) = \mu, \tag{39}$$

while $x_{\varepsilon_k}(t) \in \operatorname{int} D$, $t \in [t_0^k, \bar{t}_k]$. Again a continuity argument (and $\delta < \mu$) implies that there is a maximal t in $(t_0^k, \bar{t}_k)$ (denoted by t_k) such that

$$\operatorname{dist}(y_{\varepsilon_k}(t_k), K(t_k, x_{\varepsilon_k}(t_k)) \leq \delta. \tag{40}$$

Then for $t \in (t_k, \bar{t}_k)$

$$\delta < \operatorname{dist}(y_{\varepsilon_k}(t), K(t, x_{\varepsilon_k}(t)) < \mu. \tag{41}$$

The relations $t_k \in [0,T]$, $x_{\varepsilon_k}(t_k) \in D$, $y_{\varepsilon_k}(t_k) \in K(t_k, x_{\varepsilon_k}(t_k)) + \delta\mathbf{B}$ imply that the sequence $\{t_k, x_{\varepsilon_k}(t_k), y_{\varepsilon_k}(t_k))\}$ has a cluster point (t', x', y'). It can be supposed that the whole sequence converges to this point.

Define

$$\begin{aligned} &\bar{x}_{\varepsilon_k}(\tau) = x_{\varepsilon_k}(t_k + \varepsilon_k \tau), \\ &\bar{y}_{\varepsilon_k}(\tau) = y_{\varepsilon_k}(t_k + \varepsilon_k \tau), \quad \tau \in [0, \tfrac{\bar{t}_k - t_k}{\varepsilon_k}] = [0, \tau_k]. \end{aligned}$$

Obviously $(\bar{x}_{\varepsilon_k}(\cdot), \bar{y}_{\varepsilon_k}(\cdot))$ solves on $[0, \tau_k]$ the inclusion

$$\begin{pmatrix} \frac{d}{d\tau}\bar{x}_{\varepsilon_k} \\ \frac{d}{d\tau}\bar{y}_{\varepsilon_k} \end{pmatrix} \in J_{\varepsilon_k} F(t_k + \varepsilon_k \tau, \bar{x}_{\varepsilon_k}, \bar{y}_{\varepsilon_k}), \quad \begin{aligned} \bar{x}_{\varepsilon_k}(0) = x_{\varepsilon_k}(t_k), \\ \bar{y}_{\varepsilon_k}(0) = y_{\varepsilon_k}(t_k), \end{aligned} \tag{42}$$

where

$$J_\varepsilon = \begin{pmatrix} \varepsilon I_m & 0 \\ 0 & I_n \end{pmatrix}$$

and I_r is the $(r \times r)$-unit matrix. Consider differential inclusion

$$\begin{aligned} &\dot{\tilde{x}} = 0, & &\tilde{x}(0) = x', \\ &\dot{\tilde{y}} \in \mathcal{P}_y F(t', \tilde{x}, \tilde{y}), & &\tilde{y}(0) = y' \end{aligned} \tag{43}$$

on the interval $[0, \tau_0]$, where $\tau_0 = tau_0(\delta, \delta/2)$ (see Supposition **A4**) is fixed in such a way that $\operatorname{dist}(\tilde{y}(\tau_0) < \delta/2$.

From (40) we have $\operatorname{dist}(y', K(t', x')) \leq \delta$ and according to the definitions of δ and τ_0

$$\operatorname{dist}(\tilde{y}(\tau), K(t', x')) < \mu/2, \quad \tau \in [0, \tau_0], \tag{44}$$

$$\operatorname{dist}(\tilde{y}(\tau_0), K(t', x')) < \delta/2. \tag{45}$$

Now the well known Filippov theorem (see e.g. [3, Chapter 10] will be applied to (43) and the reference function $(\bar{x}_{\varepsilon_k}(\cdot), \bar{y}_{\varepsilon_k}(\cdot))$ on the interval $[0, \bar{\tau}_k]$, where $\bar{\tau}_k = \min\{\tau_k, \tau_0\}$. As far as $\bar{x}_{\varepsilon_k}(\tau) \in D$ and

$$\bar{y}_{\varepsilon_k}(\tau) \in K(t_k + \varepsilon_k \tau, \bar{x}_{\varepsilon_k}(\tau)) + \mu \mathbf{B} \subset K([0,T], D) + \mu \mathbf{B} = E$$

(see (41)) one can estimate on $[0, \bar{\tau}_k]$

$$\beta_k \stackrel{def}{=} \operatorname{dist}\left(\begin{pmatrix} \dot{\bar{x}}_{\varepsilon_k}(\tau) \\ \dot{\bar{y}}_{\varepsilon_k}(\tau) \end{pmatrix}, \begin{pmatrix} 0 \\ \mathcal{P}_y F(t', \bar{x}_{\varepsilon_k}(\tau), \bar{y}_{\varepsilon_k}(\tau)) \end{pmatrix} \right)$$

$$\leq H\left(J_{\varepsilon_k} F(t_k + \varepsilon_k \tau, \bar{x}_{\varepsilon_k}(\tau), \bar{y}_{\varepsilon_k}(\tau)), \begin{pmatrix} 0 \\ \mathcal{P}_y F(t', \bar{x}_{\varepsilon_k}(\tau), \bar{y}_{\varepsilon_k}(\tau)) \end{pmatrix} \right)$$

$$\leq \varepsilon_k M + \delta_y(|t_k + \varepsilon_k \tau - t'|),$$

where M is a bound of $|\mathcal{P}_x F(t,x,y)|$ on $[0,T] \times D \times E$ and $\delta_y(\cdot)$ is the modulus of continuity of $\mathcal{P}_y F$ with respect to t on the same set. As far as $|t_k + \varepsilon_k \tau - t'|$ converges to zero uniformly on $[0, \bar{\tau}_k]$ (notice that $\bar{\tau}_k$ is bounded by τ_0) we obtain

$$\beta_k \longrightarrow 0.$$

According to the Filippov theorem (43) has a solution $(\tilde{x}_k(\cdot), \tilde{y}_k(\cdot))$ on $[0, \bar{\tau}_k]$ such that

$$|\tilde{x}_k(\tau) - \bar{x}_{\varepsilon_k}(\tau)| + |\tilde{y}_k(\tau) - \bar{y}_{\varepsilon_k}(\tau)| \longrightarrow 0 \tag{46}$$

uniformly in $\tau \in [0, \bar{\tau}_k]$. Having in mind that $(\tilde{x}_k(\cdot), \tilde{y}_k(\cdot))$ satisfies (44) we obtain that

$$\operatorname{dist}(\bar{y}_{\varepsilon_k}(\tau), K(t', x')) < \frac{\mu}{2}, \quad \tau \in [0, \bar{\tau}_k]$$

for all sufficiently large k. Moreover, $\tilde{x}_k(\tau) \equiv x'$, therefore taking $t \in [t_k, t_k + \varepsilon_k \bar{\tau}_k]$ and $\tau = (t - t_k)/\varepsilon_k$ we have

$$\operatorname{dist}(y_{\varepsilon_k}(t), K(t, x_{\varepsilon_k}(t)) = \operatorname{dist}(\bar{y}_{\varepsilon_k}(\tau), K(t_k + \varepsilon_k \tau, \bar{x}_{\varepsilon_k}(\tau))) \leq \frac{\mu}{2}$$

if k is sufficiently large. The last inequality compared with (39) implies that $\bar{t}_k > t_k + \varepsilon_k \bar{\tau}_k$, which means that $\bar{\tau}_k = \tau_0$. But then (46) can be applied for $\tau = \tau_0$ and (45) gives

$$\operatorname{dist}(\bar{y}_{\varepsilon_k}(\tau_0), K(t', x')) < \frac{\delta}{2}$$

for all sufficiently large k. As above one concludes also that

$$\operatorname{dist}(y_{\varepsilon_k}(t_k + \varepsilon_k \tau_0), K(t_k + \varepsilon_k \tau_0, x_{\varepsilon_k}(t_k + \varepsilon_k \tau_0)) < \frac{\delta}{2}$$

for all sufficiently large k. But this contradicts the left inequality in (41), since $t_k + \varepsilon_k \tau_0 \in (t_k, \bar{t}_k)$. This contradiction completes the proof of the lemma. Q.E.D.

Proof of Proposition 1. We start with some preliminary technical remarks and notations. Define

$$\delta_x(\mu)$$

$$= \sup\{H(\mathcal{P}_x F((t,x,y), \mathcal{P}_x(t,x,y+\xi))\,;\ t \in [0,T],\ x \in D,\ y \in K(t,x),\ |\xi| \le \mu\}, \quad (47)$$

which monotonically converges to zero with μ, because of the continuity of F and K. Moreover, **A3** implies that the set of solutions of (14) is compact in $C[0,T]$, hence there is a compact set $S_x \subset \operatorname{int} D$ containing the values of the trajectories of (14). Let $\beta_0 > 0$ be such that still

$$S_x + \beta_0 \mathbf{B} \subset \operatorname{int} D.$$

The compactness of $K(0,x^0)$, **A1** and **A5** imply that there is a compact set $S_y \subset \mathbf{R}^m$ such that

$$\tilde{y}^0(\tau) \in S_y, \quad \tau \ge 0,$$

for every solution $\tilde{y}^0(\cdot)$ of (15). Denote also

$$\delta_y(\alpha) = \sup\{H(\hat{F}((0,x,y), \hat{F}(s,x,y))\,;\ s \in [0,\alpha],\ x \in x^0 + \mathbf{B},\ y \in S_y + \mathbf{B}\}, \quad (48)$$

which monotonically converges to zero with α, because of the continuity of $\hat{F}$.

Finally, denote

$$C = \exp \int_0^T L_x(t)\,dt, \quad (49)$$

where $L_x(t)$ is the Lipschitz constant of $F_0(t,\cdot)$ on D (so that C is finite according to **A3**) and fix $\mu_0 > 0$ such that

$$TC\delta_x(\mu_0) \le \frac{\beta_0}{2}. \quad (50)$$

Now we proceed with the proof of the proposition. Local extendibility of the solutions of (9) is implied by **A1**. Uniform boundedness of the solutions (which would give expendability till T) is claimed in the second assertion. Moreover, $x_\varepsilon(t) \in \operatorname{int} D$ follows from a) if $\varepsilon_0 > 0$ is chosen so small that $\alpha_\varepsilon < \beta_0$. Thus it remains to prove b) and c).

Fix an arbitrary $\mu \in (0,\mu_0]$ and denote $\delta = \delta(\mu/2)$ (see Supposition **A4**). Let $(x_\varepsilon(\cdot), y_\varepsilon(\cdot))$ be a solution of (9) on $[0,T]$. Then

$$\bar{x}_\varepsilon(\tau) = x_\varepsilon(\varepsilon\tau)$$
$$\bar{y}_\varepsilon(\tau) = y_\varepsilon(\varepsilon\tau), \quad \tau \in [0, T/\varepsilon]$$

satisfy the inclusion

$$\begin{pmatrix} \dot{\bar{x}}_\varepsilon(\tau) \\ \dot{\bar{y}}_\varepsilon(\tau) \end{pmatrix} \in J_\varepsilon F(\varepsilon\tau, \bar{x}_\varepsilon(\tau), \bar{y}_\varepsilon(\tau)), \quad \begin{matrix} \bar{x}_\varepsilon(0) = x^0, \\ \bar{y}_\varepsilon(0) = y^0 \end{matrix}$$

on $[0, T/\varepsilon]$, where J_ε is defined as in the proof of Lemma 2. Consider the differential inclusion

$$\begin{aligned} &\dot{\tilde{x}}^0(\tau) = 0, && \tilde{x}^0(0) = x^0, \\ &\dot{\tilde{y}}^0(\tau) \in \mathcal{P}_y F(0, \tilde{x}^0(\tau), \tilde{y}^0(tau)), && \tilde{y}^0(0) = y^0 \end{aligned} \tag{51}$$

on the interval $[0, \tau_0]$, where $\tau_0 = \tau_0(\delta/3)$ (see Supposition **A5**). Denote $\Gamma = (x^0 + \mathbf{B}) \times (S_y + \mathbf{B})$ and assume that $(\bar{x}_\varepsilon(\tau), \bar{y}_\varepsilon(\tau)) \in \Gamma$ until some (maximal) moment $\bar{\tau}_\varepsilon \leq \tau_0$. Then the deviation

$$\beta_\varepsilon = \operatorname{dist}\left(\begin{pmatrix} \dot{\bar{x}}_\varepsilon(\tau) \\ \dot{\bar{y}}_\varepsilon(\tau) \end{pmatrix}, \begin{pmatrix} 0 \\ \mathcal{P}_y F(0, \bar{x}_\varepsilon(\tau), \bar{y}_\varepsilon(\tau)) \end{pmatrix} \right)$$

$$\leq \varepsilon |\mathcal{P}_x F(\varepsilon\tau, \bar{x}_\varepsilon(\tau), \bar{y}_\varepsilon(\tau))| + H(\mathcal{P}_y F(\varepsilon\tau, \bar{x}_\varepsilon(\tau), \bar{y}_\varepsilon(\tau)), \mathcal{P}_y F(0, \bar{x}_\varepsilon(\tau), \bar{y}_\varepsilon(\tau)))$$

$$\leq \varepsilon M + \delta_y(\varepsilon\tau) \longrightarrow 0,$$

where M is a bound of $|\mathcal{P}_x(t, x, y)|$ on $[0, T] \times \Gamma$ and δ_y is defined by (48). According to the Filippov theorem there is a solution $(\tilde{x}_\varepsilon^0(\cdot), \tilde{y}_\varepsilon^0(\cdot))$ of (51) such that

$$|\bar{x}_\varepsilon(\tau) - \tilde{x}_\varepsilon^0(\tau)| + |\bar{y}_\varepsilon(\tau)) - \tilde{y}_\varepsilon^0(\tau)| \leq C_1 \beta_\varepsilon, \quad \tau \in [0, \bar{\tau}_\varepsilon], \tag{52}$$

where C_1 is independent of μ and ε. Since $\tilde{x}_\varepsilon^0(\tau) \equiv x^0$ and $\tilde{y}_\varepsilon^0(\tau) \in S_y$, the last inequality shows that for all sufficiently small ε (so that $C\beta_\varepsilon < 1$) $\bar{\tau}_\varepsilon = \tau_0$. From **A5** and (52) we obtain successively

$$\operatorname{dist}(\tilde{y}_\varepsilon^0(\tau_0), K(0, x^0)) < \frac{\delta}{3}$$

$$\operatorname{dist}(\bar{y}_\varepsilon(\tau_0), K(0, x^0)) < \frac{2\delta}{3}$$

$$\operatorname{dist}(y_\varepsilon(\varepsilon\tau_0), K(\varepsilon\tau_0, x_\varepsilon(\varepsilon\tau_0)) < \delta$$

for all sufficiently small ε. Applying Lemma 2, (now for $\varepsilon < \varepsilon_0(\mu)$) we obtain

$$\operatorname{dist}(y_\varepsilon(t), K(t, x_\varepsilon(t)) \leq \mu, \quad t \in [\varepsilon\tau_0, \bar{t}_\varepsilon], \tag{53}$$

where either $\bar{t}_\varepsilon = T$ or $x_\varepsilon(\bar{t}_\varepsilon) \in \partial D$, but in both cases $x_\varepsilon(t) \in D$ for $t \in [0, \bar{t}_\varepsilon]$. We remind that according to (52) $(x_\varepsilon(t), y_\varepsilon(t)) \in \Gamma$ for $t \in [0, \varepsilon\tau_0]$, hence

$$|x_\varepsilon(t) - x^0| \leq 1, \quad t \in [0, \varepsilon\tau_0], \tag{54}$$

$$\operatorname{dist}(y_\varepsilon(t), K(t, x_\varepsilon(t)) \leq C_2, \quad t \in [0, \varepsilon\tau_0], \tag{55}$$

where C_2 is independent of μ and ε.

Now we shall apply the Filippov theorem to differential inclusion (14) and reference function $x_\varepsilon(\cdot)$. Using (55), (53) and (54) we obtain

$$\gamma_\varepsilon = \int_0^T \operatorname{dist}(\dot{x}_\varepsilon(t), \mathcal{P}_x F(t, x_\varepsilon(t), K(t, x_\varepsilon(t)))) \, dt$$

$$\leq \int_0^T H^+(\mathcal{P}_x F(t, x_\varepsilon(t), y_\varepsilon(t)), \mathcal{P}_x F(t, x_\varepsilon(t), K(t, x_\varepsilon(t)))) \, dt$$

$$\leq \int_0^{\varepsilon\tau_0} \delta_y(C_2)\,dt + \int_{\varepsilon\tau_0}^{T} \delta_x(\mu)\,dt = \delta_x(C_2)\tau_0\varepsilon + T\delta_y(\mu),$$

where $H^+(P,Q) \stackrel{def}{=} \sup\{\text{dist}(p,Q)\,;\ p \in P\}$ and δ_x is defined in (47). According to the Filippov theorem there is a solution $x_\varepsilon^0(\cdot)$ of (14) such that

$$|x_\varepsilon(t) - x_\varepsilon^0(t)| \leq C\delta_x(C_2)\tau_0\varepsilon + CT\delta_x(\mu) \leq C\delta_x(C_2)\tau_0\varepsilon + \frac{\beta_0}{2}, \tag{56}$$

where C is defined in (49) and the last inequality uses (50). Thus the right-hand side is less than β_0 for all sufficiently small ε. Hence, $\bar{t}_\varepsilon = T$. Moreover,

$$\text{dist}_{C[0,T]}(x_\varepsilon(\cdot), \mathcal{P}_x Z_0) \longrightarrow C\delta_x(C_2)\tau_0\varepsilon + CT\delta_x(\mu). \tag{57}$$

The properties a) and b) and c) of α_ε and $\gamma_(\cdot)$ are obvious consequences of (55), (53) and (57), since $\mu > 0$ can be taken arbitrarily small. Q.E.D.

In the proof of Theorem 1 we shall use a stronger form of the theorem of Filippov. We formulate this stronger modification in a somewhat loose way and skip the proof which is standard.

Lemma 3 *Let $H : [0,T] \times \mathbf{R}^r \Rightarrow \mathbf{R}^r$ be convex compact valued and measurable in t. Let $x(\cdot)$ be a solution of the differential inclusion*

$$\dot{x}(t) \in H(t, x(t)) + q(t), \quad x(0) = x_0,$$

where $q(\cdot)$ is integrable. Let $H(t,\cdot)$ be Lipschitz continuous with Lipschitz constant $\lambda(t)$ in the δ-neighborhood of $x(t)$, $\delta > 0$. Suppose that $\lambda(\cdot)$ is integrable and that

$$\rho(q)(t) = |\int_0^t q(s)\,ds| + e^{\int_0^t \lambda(s)\,ds} \int_0^t \lambda(s) |\int_0^s q(\tau)\,d\tau|\,ds < \delta.$$

Then there exists a solution $\tilde{x}(\cdot)$ of

$$\dot{\tilde{x}}(t) \in H(t, \tilde{x}(t)), \quad \tilde{x}(0) = x_0$$

such that

$$|\tilde{x}(t) - x(t)| \leq \rho(q)(t), \quad t \in [0,T].$$

Proof of Theorem 1. The proof will make use of the above auxiliary results and will consist of several steps.

1. Preliminaries. Upper semicontinuity of Z_ε in the Tikhonov metric is equivalent to the following: for every sufficiently small $\varepsilon > 0$ and for every $z_\varepsilon = (x_\varepsilon, y_\varepsilon) \in Z_\varepsilon$ there exists $\bar{z}_\varepsilon = (\bar{x}_\varepsilon, \bar{y}_\varepsilon) \in Z_0$ such that $|x_\varepsilon - \bar{x}_\varepsilon|_C \to 0$ and $\tau(y_\varepsilon, \bar{y}_\varepsilon) \to 0$ when $\varepsilon \longrightarrow 0$.

Let $z_\varepsilon \in Z_\varepsilon$ and let $u_\varepsilon(t) \in U(t, x_\varepsilon(t))$ be the corresponding selection in the representation (16) of F. Proposition 1 implies that y_ε (as well as x_ε) are uniformly bounded and **A1** implies that there is a constant $\bar{K}$ such that

$$|\varepsilon \dot{y}_\varepsilon(t)| \leq \bar{K}, \quad t \in [0,T],$$

for all sufficiently small $\varepsilon > 0$.

We shall define a trajectory $(\bar{x}_\varepsilon, \bar{y}_\varepsilon) \in Z_0$ by an iteration procedure. (This iteration procedure can be avoided in the case of a convex set K and $U(t,x) = \{0\}$ by using the Steiner point selection technique [3, Chapter 9]). At the k-th step we shall define appropriate $x_\varepsilon^k(\cdot), y_\varepsilon^k(\cdot)$ (where ε is fixed, but sufficiently small) and then we shall take the limit with respect to k to obtain the desired $(\bar{x}_\varepsilon, \bar{y}_\varepsilon) \in Z_0$.

2. Description of the first iteration step. We start the iteration procedure with $x_\varepsilon^0 = x_\varepsilon$, $y_\varepsilon^0 = y_\varepsilon$, $u_\varepsilon^0 = u_\varepsilon$. The first iteration step is somewhat different from the subsequent ones, therefore we describe it in detail and then we formulate the common rule.

According to Proposition 1 there exists a measurable function $y_\varepsilon^1(t) \in K(t, x_\varepsilon^0(t))$ such that

$$|y_\varepsilon^1(t) - y_\varepsilon^0(t)| \leq \gamma_\varepsilon(t). \tag{58}$$

Since $K(t,x) \subset K_0(t,x)$ we have

$$0 \in F_2(t, x_\varepsilon^0(t), y_\varepsilon^1(t)) + u_\varepsilon^1(t) \tag{59}$$

for some measurable selection

$$u_\varepsilon^1(t) \in U(t, x_\varepsilon^0(t)). \tag{60}$$

Thus for some measurable $\bar{\xi}_\varepsilon(t) \in F_2(t, x_\varepsilon^0(t), y_\varepsilon^1(t))$

$$0 = \bar{\xi}_\varepsilon(t) + u_\varepsilon^1(t). \tag{61}$$

Let λ be the Lipschitz constant of $F(t,\cdot,\cdot)$ corresponding to the compact set $G = D \times (K([0,T], D) + \nu \mathbf{B})$, where ν is an upper bound of $\gamma_\varepsilon(\cdot)$. According to Proposition 1 the values of $(x_\varepsilon^0(t), y_\varepsilon^0(t))$ and $(x_\varepsilon^0(t), y_\varepsilon^1(t))$ belong to this set. The same will concern the values of the other pairs $(x_\varepsilon^i(t), y_\varepsilon^{i+1}(t))$ defined iteratively below as it will become clear in the end.

There is a measurable selection $\eta_\varepsilon(t) \in F_2(t, x_\varepsilon^0(t), y_\varepsilon^0(t))$ such that

$$\varepsilon \dot{y}_\varepsilon^0(t) = \eta_\varepsilon(t) + u_\varepsilon^0(t). \tag{62}$$

Then the projection

$$\xi_\varepsilon(t) = \mathcal{P}_{F_2(t, x_\varepsilon^0(t), y_\varepsilon^1(t))} \eta_\varepsilon(t)$$

is measurable and

$$|\xi_\varepsilon(t) - \eta_\varepsilon(t)| \leq \lambda \gamma_\varepsilon(t). \tag{63}$$

Subtracting (61) from (62) we obtain

$$u_\varepsilon^0(t) - u_\varepsilon^1(t) = -(\xi_\varepsilon(t) - \bar{\xi}_\varepsilon(t)) - (\eta_\varepsilon(t) - \xi_\varepsilon(t)) + \varepsilon \dot{y}_\varepsilon^0(t). \tag{64}$$

Condition i) in Supposition **A6** implies that the projection map $\mathcal{P}_{\mathrm{Lins}\, U(t,x)}(\cdot)$ is continuous in (t,x), therefore there exists a continuous $n \times n$-matrix $C(t,x)$ such that $\mathcal{P}_{\mathrm{Lins}\, U(t,x)}(l) = C(t,x)l$ for each $l \in \mathbf{R}^n$. Multiplying (64) by $C(t, x_\varepsilon^0(t))$ and then by $B(t, x_\varepsilon^0(t))$ and taking into account that $u_\varepsilon^0(t) - \bar{u}_\varepsilon^0(t) \in \mathrm{Lins}\, U(t, x_\varepsilon^0(t))$ we obtain that

$$B(t, x_\varepsilon^0(t))(u_\varepsilon^0(t) - u_\varepsilon^1(t))$$

$$= -BC(t, x_\varepsilon^0(t))(\xi_\varepsilon(t) - \bar{\xi}_\varepsilon(t)) - BC(t, x_\varepsilon(t))(\eta_\varepsilon(t) - \xi_\varepsilon(t)) + BC((t, x_\varepsilon^0(t))\varepsilon \dot{y}_\varepsilon^0(t).$$

Since $\xi_\varepsilon(t) - \bar{\xi}_\varepsilon(t) \in \mathrm{Lins}\, F_2(t, x_\varepsilon^0(t), y_\varepsilon^1(t))$ Supposition **A6** ii) implies that the first term in the right-hand side is zero, hence

$$B(t, x_\varepsilon^0(t))(u_\varepsilon^0(t) - u_\varepsilon^1(t)) = BC(t, x_\varepsilon(t))\omega_\varepsilon(t),$$

where

$$\omega_\varepsilon(\cdot) = (\xi_\varepsilon(\cdot) - \eta_\varepsilon(\cdot)) + \varepsilon \dot{y}_\varepsilon^0(t).$$

Having in mind also (63) we obtain that x_ε^0 satisfies the differential inclusion

$$\dot{x}_\varepsilon^0(t) \in F_1(t, x_\varepsilon^0(t), y_\varepsilon^0(t)) + B(t, x_\varepsilon^0(t))u_\varepsilon^1(t) + BC(t, x_\varepsilon^0(t))\omega_\varepsilon(t)$$

$$\subset F_1(t, x_\varepsilon^0(t), y_\varepsilon^1(t)) + B(t, x_\varepsilon^0(t))u_\varepsilon^1(t) + BC(t, x_\varepsilon^0(t))\omega_\varepsilon(t) + \lambda\gamma_\varepsilon(t)\mathbf{B}, \quad x_\varepsilon^0(0) = x^0. \tag{65}$$

We shall apply Lemma 3 to the differential inclusion (65) for

$$q(t) = q_\varepsilon(t) \in BC(t, x_\varepsilon^0(t))\omega_\varepsilon(t) + \lambda\gamma_\varepsilon(t)\mathbf{B}.$$

Having in mind (63) and the properties of γ_ε given in Proposition 1 we can represent

$$\int_0^t q_\varepsilon(s)\, ds = \varepsilon \int_0^t BC(s, x_\varepsilon(s))\dot{y}_\varepsilon^0(s)\, ds + \alpha_\varepsilon(t).$$

where $\alpha_\varepsilon(t) \to 0$ uniformly. In order to estimate the first term in the right-hand side we take one row $d_\varepsilon(t)$ of $BC(s, x_\varepsilon(t))$ and keep in mind that $x_\varepsilon(\cdot)$ are uniformly bounded and equi-Lipschitz, hence $d_\varepsilon(\cdot)$ are uniformly bounded (by a constant M) and equi-continuous. We know also that y_ε and $\varepsilon\dot{y}_\varepsilon$ are uniformly bounded (by constants $\bar{M}$ and $\bar{K}$, respectively).

Take an arbitrary positive number δ and $t \in (0, T]$. Denote $t_j = jt/p$, $j = 1, \ldots, p$, where p is fixed so large that for each two numbers $t', t'' \in [0, t]$ for which $|t' - t''| \le T/p$ it holds

$$|d_\varepsilon(t') - d_\varepsilon(t'')| \le \frac{\delta}{2\bar{K}T} \quad \forall \varepsilon > 0.$$

Without any restriction we may assume that $\varepsilon < \delta/(4M\bar{M}p)$, since all the constants in the right-hand side are independent of ε. We have

$$|\int_0^t \varepsilon\langle \dot{y}_\varepsilon^0(s), d_\varepsilon(s)\rangle\, ds| = |\sum_{j=0}^{p-1} \int_{t_j}^{t_{j+1}} [\langle\, \varepsilon\dot{y}_\varepsilon^0(s), d_\varepsilon(t_j) + (d_\varepsilon(s) - d_\varepsilon(t_j)\rangle\, ds|$$

$$\leq |\sum_{j=0}^{p-1} \varepsilon \int_{t_j}^{t_{j+1}} \langle \dot{y}_\varepsilon^0(s), d_\varepsilon(t_j)\rangle \, ds| + \sum_{j=0}^{p-1} \int_{t_j}^{t_{j+1}} |\varepsilon \dot{y}_\varepsilon^0(s)||d_\varepsilon(s) - d_\varepsilon(t_j)| \, ds$$

$$= \varepsilon |\sum_{j=0}^{p-1} \langle y_\varepsilon^0(t_{j+1}) - y_\varepsilon^0(t_j), d_\varepsilon(t_j)\rangle| + \sum_{j=0}^{p-1} \frac{T}{p} \bar{K} \frac{\delta}{2\bar{K}T} \leq 2\varepsilon p \bar{M} M + \frac{\delta}{2} \leq \delta.$$

Thus we obtain that $\rho(q_\varepsilon)(\cdot)$ (we use the notation from Lemma 3) converges uniformly to zero. Then Lemma 3 implies the existence of a solution $x_\varepsilon^1(\cdot)$ of

$$\dot{x}_\varepsilon^1 \in F_1(t, x_\varepsilon^1(t), y_\varepsilon^1(t)) + B(t, x_\varepsilon^1(t)) u_\varepsilon^1(t) \tag{66}$$

for which

$$\beta_\varepsilon = \|x_\varepsilon^1(\cdot) - x_\varepsilon^0(\cdot)\|_C \to 0. \tag{67}$$

3. Description of the $(k+1)$-st iteration step. Now we begin with the description of the k-th iteration step, supposing that absolutely continuous $x_\varepsilon^i(\cdot)$ and measurable $y_\varepsilon^i(\cdot), u_\varepsilon^i(\cdot)$ are already defined for $i = 0, \ldots, k$ in such a way that

$$\dot{x}_\varepsilon^k \in F_1(t, x_\varepsilon^k(t), y_\varepsilon^k(t)) + B(t, x_\varepsilon^k(t)) u_\varepsilon^k(t), \quad x_\varepsilon^k(0) = x^0, \tag{68}$$

$$0 \in F_2(t, x_\varepsilon^{k-1}(t), y_\varepsilon^k(t)) + u_\varepsilon^k(t), \tag{69}$$

$$u_\varepsilon^k(t) \in U(t, x_\varepsilon^{k-1}(t)), \tag{70}$$

$$|x_\varepsilon^k(t) - x_\varepsilon^{k-1}(t)| \leq a_\varepsilon^k(t) \stackrel{def}{=} \left(e^{\lambda T} C_0\right)^{k-1} \frac{t^{k-1}}{(k-1)!} \beta_\varepsilon, \tag{71}$$

$$|y_\varepsilon^k(t) - y_\varepsilon^{k-1}(t)| \leq L a_\varepsilon^{k-1}(t) = L \left(e^{\lambda T} C_0\right)^{k-2} \frac{t^{k-2}}{(k-2)!} \beta_\varepsilon, \tag{72}$$

where $C_0 = \lambda(L + M'(1+L))$, L is the Lipschitz constant of $K(t, \cdot)$ in D and M' is a bound of the norm of BC on $[0,T] \times D$. From (66), (59), (60) and (67) we see that (68)–(71) are fulfilled for $k = 1$. Only (72) is not fulfilled because of the specificity of the first step (for $k = 1$ we have (58), where $\gamma_\varepsilon(\cdot)$ is only point-wise convergent). But, in fact, we shall not use the validity of (72) for k when proving that it holds for $k+1$.

In the end it will become clear that all $(x_\varepsilon^k(t), y_\varepsilon^k(t))$, $k \geq 1$ belong to $D \times K([0,T], D)$ for all sufficiently small ε, therefore we skip checking this at every step.

From (69) and (71) we obtain

$$y_\varepsilon^k(t) \in K(t, x_\varepsilon^{k-1}(t)) \subset K(t, x_\varepsilon^k(t)) + L a_\varepsilon^k(t) \mathbf{B}.$$

Then there exists a measurable $y_\varepsilon^{k+1}(\cdot)$ such that

$$y_\varepsilon^{k+1}(t) \in K(t, x_\varepsilon^k(t)), \quad |y_\varepsilon^{k+1}(t) - y_\varepsilon^k(t)| \leq L a_\varepsilon^k(t). \tag{73}$$

The second relation is just (72) for $k+1$. The first of the above relations means that there is a measurable $u_\varepsilon^{k+1}(t) \in U(t, x_\varepsilon^k(t))$, such that

$$0 \in F_2(t, x_\varepsilon^k(t), y_\varepsilon^{k+1}(t)) + u_\varepsilon^{k+1}(t).$$

Thus (70) and (69) are also fulfilled for $k+1$.

On the other hand from (69), (71) and the second relation in (73) we obtain

$$0 \in F_2(t, x_\varepsilon^k(t), y_\varepsilon^{k+1}(t)) + u_\varepsilon^k(t) + \lambda(|x_\varepsilon^k(t) - x_\varepsilon^{k-1}(t)| + |y_\varepsilon^{k+1}(t) - y_\varepsilon^k(t)|)\mathbf{B}$$
$$\subset F_2(t, x_\varepsilon^k(t), y_\varepsilon^{k+1}(t)) + u_\varepsilon^k(t) + \lambda a_\varepsilon^k(t)(1+L)\mathbf{B}.$$

Thanks to Supposition **A6** ii) we obtain exactly as in the first step

$$B(t, x_\varepsilon^k(t))(u_\varepsilon^k(t) - u_\varepsilon^{k+1}(t)) = BC(t, x_\varepsilon^k(t))\omega_\varepsilon(t),$$

where

$$|\omega_\varepsilon(\cdot)| \le \lambda a_\varepsilon^k(t)(1+L).$$

Using the above two relations and the second estimation in (73) we get

$$\dot{x}_\varepsilon^k(t) \in F_1(t, x_\varepsilon^k(t), y_\varepsilon^k(t)) + B(t, x_\varepsilon^k(t))u_\varepsilon^{k+1}(t) + BC(t, x_\varepsilon^k(t))\omega_\varepsilon(t) \subset$$
$$F_1(t, x_\varepsilon^k(t), y_\varepsilon^{k+1}(t)) + B(t, x_\varepsilon^k(t))u_\varepsilon^{k+1}(t) + \lambda(M' a_\varepsilon^k(t)(1+L) + La_\varepsilon^k(t))\mathbf{B}$$
$$= F_1(t, x_\varepsilon^k(t), y_\varepsilon^{k+1}(t)) + B(t, x_\varepsilon^k(t))u_\varepsilon^{k+1}(t) + C_0 a_\varepsilon^k(t)\mathbf{B}.$$

According to the (classical) Filippov theorem there exists a solution $x_\varepsilon^{k+1}(\cdot)$ of

$$\dot{x}_\varepsilon^{k+1} \in F_1(t, x_\varepsilon^{k+1}(t), y_\varepsilon^{k+1}(t)) + B(t, x_\varepsilon^{k+1}(t))u_\varepsilon^{k+1}(t), \quad x_\varepsilon^{k+1}(0) = x^0,$$

such that

$$|x_\varepsilon^{k+1}(t) - x_\varepsilon^k(t)| \le e^{\lambda T} \int_0^t C_0 a_\varepsilon^k(s)\, ds =$$
$$(e^{\lambda T} C_0)\left(e^{\lambda T} C_0\right)^{k-1} \beta_\varepsilon \int_0^t \frac{s^{k-1}}{(k-1)!}\, ds = \left(e^{\lambda T} C_0\right)^k \frac{t^k}{k!}\beta_\varepsilon = a_\varepsilon^{k+1}.$$

The above two relations shaw that (68) and (71) are fulfilled also for $k+1$. This completes the $(k+1)$-st iteration step.

4) Passing to a limit. Notice that sum of all $a_\varepsilon^k(t)$ is finite and can be estimated by

$$\sum_{i=0}^{+\infty} a_\varepsilon^i(t) \le C_1 \beta_\varepsilon, \tag{74}$$

where C_1 is independent of ε and t. According to Proposition 1 $x_\varepsilon^0(t) + \kappa\mathbf{B} \subset \operatorname{int} D$ for some $\kappa > 0$. Since $\beta_\varepsilon \to 0$ with ε one may chose $\varepsilon_0 > 0$ so small that $C_1\beta_\varepsilon < \kappa$ for $\varepsilon \in (0, \varepsilon_0]$. Then $x_\varepsilon^k(t) \in D$ for all $\varepsilon < \varepsilon_0$ which legitimates the use of the constants λ, L and M' as they where defined above.

The convergence of the sum in (74) together with (71) and (72) implies also uniform convergence of $x_\varepsilon^k(\cdot)$ and $y_\varepsilon^k(\cdot)$ to respective $\bar{x}_\varepsilon(\cdot)$ and $\bar{y}_\varepsilon(\cdot)$. Since $u_\varepsilon^k(t) \in U([0,T], D)$ which is bounded, one can take an L_2-weakly convergent subsequent with a limit $\bar{u}_\varepsilon(\cdot)$ (we shall use the same indexes).

5. End of the proof. Thanks to the uniform convergence of $x_\varepsilon^k(\cdot)$, the convexity and the continuity of $U(t,x)$, and the Mazur theorem we obtain in a standard way

$$\bar{u}_\varepsilon(t) \in U(t, \bar{x}_\varepsilon(t)).$$

By the same argument we have also

$$0 \in F_2(t, \bar{x}_\varepsilon(t), \bar{y}_\varepsilon(t)) + \bar{u}_\varepsilon(t).$$

Choosing a L_2-weakly convergent subsequence of $\dot{x}_\varepsilon^k(\cdot)$ and employing for third time the Mazur theorem we obtain also

$$\dot{\bar{x}}_\varepsilon \in F_1(t, \bar{x}_\varepsilon(t), \bar{y}_\varepsilon(t)) + B(t, \bar{x}_\varepsilon(t))\bar{u}_\varepsilon(t), \quad \bar{x}_\varepsilon(0) = x^0.$$

The last three inclusions mean that $(\bar{x}_\varepsilon(\cdot), \bar{y}_\varepsilon(\cdot)) \in Z_0$. Moreover,

$$|\bar{x}_\varepsilon(t) - x_\varepsilon(t)| = |\bar{x}_\varepsilon(t) - x_\varepsilon^0(t)| \le \sum_{i=0}^{+\infty} a_\varepsilon^i(t) \le C_1\beta_\varepsilon \to 0,$$

$$|\bar{y}_\varepsilon(t) - y_\varepsilon(t)| = |\bar{y}_\varepsilon(t) - y_\varepsilon^0(t)| \le |y_\varepsilon^1(t) - y_\varepsilon^0(t)| + L\sum_{i=1}^{+\infty} a_\varepsilon^i(t)$$

$$\le \gamma_\varepsilon(t) + LC_1\beta_\varepsilon.$$

Because of the properties of the function $\gamma_\varepsilon(\cdot)$ from Proposition 1 we have

$$\tau(\bar{y}_\varepsilon(\cdot), y_\varepsilon(\cdot)) \to 0.$$

This completes the proof of the theorem. Q.E.D.

Appendix B

In the proof of the lower semicontinuity of the trajectory set we shall use the following lemma, which we prove for completeness.

Lemma 4 *Let $(t,u) \longrightarrow M(t,u)$ be a continuous mapping from $[0,T] \times U$ to the convex compact subsets of $\mathbf{R}^n$. Then, given measurable functions $u_0(\cdot)$ and $y_0(\cdot)$ satisfying*

$$y_0(t) \in M(t, u_0(t)), \quad u_0(t) \in U$$

and a positive number α, there exist continuous $\hat{y}_0(\cdot)$ and Lipschitz continuous $\hat{u}_0(\cdot)$ such that

$$\hat{y}_0(t) \in M(t, \hat{u}_0(t)), \quad \hat{u}_0(t) \in U$$

and

$$\|u_0(\cdot) - \hat{u}_0(\cdot)\|_{L_1} \leq \alpha, \quad \|y_0(\cdot) - \hat{y}_0(\cdot)\|_{L_1} \leq \alpha.$$

Proof. The proof consists of three steps. First we approximate $u_0(\cdot)$ by an appropriate Lipschitz selection $\hat{u}_0(\cdot)$ of U. Then we approximate $y_0(\cdot)$ by a continuous function that is not necessary a selection of $M(t, \hat{u}_0(t))$, finaly we "project" it on $M(t, \hat{u}_0(t))$ preserving the continuity.

1. First we shall define an appropriate $\hat{u}_0(\cdot)$. Obviously $u_0(\cdot)$ can be approximated in L_1 with any accuracy $1/k$ by a Lipschitz continuous selection $u_k(\cdot)$ of U, since U is convex. Denote $\rho_k(t) = |u_0(t) - u_k(t)|$. Since $\rho_k \to 0$ in L_1, there is a subsequence (we keep the same indexes) such that $\rho_k(t) \to 0$ almost everywhere. Moreover, $\rho_k(\cdot)$ is uniformly bounded, since U is compact. The modulus of continuity $\omega_M(\cdot)$ of M is a monotone increasing function and $\lim_{h \to 0} \omega_M(h) = \omega_M(0) = 0$. Therefore the function $\omega_M(\rho_k(\cdot))$ is measurable and uniformly bounded, hence integrable. By the dominated convergence theorem

$$\int_0^T \omega_M(\rho_k(s))\, ds \to 0,$$

thus

$$\|\omega_M(\rho_k(\cdot))\|_{L_1} \leq \alpha$$

for $k_0 \geq \alpha^{-1}$ for which $1/k < \alpha$. We define $\hat{u}_0(\cdot) = u_{k_0}(\cdot)$.

2. Now we take an arbitrary continuous approximation $\tilde{y}_0(\cdot)$ of $y_0(\cdot)$ such that

$$\|\tilde{y}_0(\cdot) - y_0(\cdot)\|_{L_1} \leq \alpha.$$

We have

$$\int_0^T \mathrm{dist}(\tilde{y}_0(t), M(t, \hat{u}_0(t))\, dt \leq \|\tilde{y}_0(\cdot) - y_0(\cdot)\|_{L_1} + \int_0^T \mathrm{dist}(y_0(t), M(t, \hat{u}_0(t))\, dt$$

$$\leq \alpha + \int_0^T H(M(t, u_0(t)), M(t, \hat{u}_0(t)))\, dt \leq \alpha + \int_0^T \omega_M(\rho_k(t))\, dt \leq 2\alpha.$$

3. Finally we approximate $\tilde{y}_0(\cdot)$ by a selection of M. Namely, we consider the mapping

$$M(t, \hat{u}_0(t)) \cap [\tilde{y}_0(t) + 2\mathrm{dist}(\tilde{y}_0(t), M(t, \hat{u}_0(t)))\mathbf{B}].$$

Lemma 9.4.2 in [3] claims that for every two convex compact sets M_1 and M_2 in $\mathbf{R}^n$ and every two points y_1 and y_2 in $\mathbf{R}^n$ it holds

$$H(M_1 \cap (y_1 + 2\mathrm{dist}(y_1, M_1)), M_2 \cap (y_2 + 2\mathrm{dist}(y_2, M_2))) \leq 5(H(M_1, M_2) + |y_1 - y_2|). \quad (75)$$

This implies that the above mapping is continuous (and convex valued), therefore has a continuous selection $\hat{y}_0(\cdot)$. We have

$$\|\hat{y}_0(\cdot) - y_0(\cdot)\|_{L_1} \leq \|\hat{y}_0(\cdot) - \tilde{y}_0(\cdot)\|_{L_1} + \|\tilde{y}_0(\cdot) - y_0(\cdot)\|_{L_1}$$

$$\leq 2 \int_0^T \operatorname{dist}(\tilde{y}_0(t), M(t, \hat{u}_0(t)))\, dt + \alpha \leq 5\alpha$$

The proof is complete. Q.E.D.

Proof of Theorem 3. We split the proof in a number of steps.

1. First we introduce some notations. By a compactness argument Supposition **C3** is fulfilled also for some smaller compact set D' such that $D' + \rho\mathbf{B} \subset D$ for some $\rho > 0$.

Let $G' \subset \mathbf{R}^n$ be a compact set such that

$$G_0 \cup \hat{K}_0([0,T], D, U) \subset G'.$$

(see suppositions **C2** and **C5**). Denote $G = G' + \mathbf{B}$. Let L be a Lipschitz constant of F_1 and F_2 on $D \times G \times U$ (uniformly in $t \in [0,T]$) and let M be a bound of $|F_1|$ and $|F_2|$ in the same set. The modulus of continuity of F_2 with respect to t (uniformly in $(x, y, u) \in D \times G \times U$) will be denoted by $\omega_t(\cdot)$. Similarly, $\hat{L}$ will denote a Lipschitz constant of $\hat{K}_0$ with respect to $x \in D$ (uniformly in (t, u)) and $\hat{\omega}(\cdot)$ will denote the modulus of continuity $\hat{K}_0$ on $[0,T] \times D \times U$.

Consider

$$\dot{x} \in F_1(t, x, G, U), \quad x(0) = x^0. \tag{76}$$

Since $x^0 \in D'$, there is $\beta_0 > 0$ such that all solutions remain in $D' + 0.5\rho\mathbf{B}$ on $[0, \beta_0]$.

2. Suppose that the claim of the theorem is not true. Then there exist $z_0 = (x_0(\cdot), y_0(\cdot)) \in Z_0$, a sequence $\varepsilon_k \to 0$ and $\sigma > 0$ such that for each $z_k = (x_k(\cdot), y_k(\cdot)) \in Z_{\varepsilon_k}$ either

$$\|x_k(\cdot) - x_0(\cdot)\|_{C[0,T]} > \sigma, \tag{77}$$

or

$$\|y_k(\cdot) - y_0(\cdot)\|_{L_1[0,T]} > \sigma. \tag{78}$$

Let $u_0(\cdot)$ be the selection of U corresponding to $z_0(\cdot)$.

3. We continue with some preparatory work.

3.1. According to Lemma 9.4.2 [3] (see also (75)) the mapping

$$(t, x, y, u) \longrightarrow F_1(t, x, y, u) \cap [\dot{x}_0(t) + 2\operatorname{dist}(\dot{x}_0(t), F_1(t, x, y, u))\mathbf{B}]$$

is locally Lipschitz with respect to (x, y, u) (uniformly in t) and measurable in t. Obviously it is also convex and compact valued. Then its Steiner point $\xi_1(t, x, y, u)$ has the same properties (Theorem 9.4.1 in [3]). We keep the notation L also for its Lipschitz constant in the set $D \times G \times U$ (uniformly in $t \in [0, T]$).

3.2. According to Lemma 4 applied to $M(t, u) = \hat{K}_0(t, x_0(t), u)$ for every $\alpha > 0$ there are continuous $\hat{u}_0(\cdot)$ and $\hat{y}_0(\cdot)$ such that

$$\hat{y}_0(t) \in \hat{K}_0(t, x_0(t), \hat{u}_0(t)), \quad \hat{u}_0(t) \in U,$$

$$\|u_0(\cdot) - \hat{u}_0(\cdot)\|_{L_1} \leq \alpha, \quad \|y_0(\cdot) - \hat{y}_0(\cdot)\|_{L_1} \leq \alpha.$$

Denote by $\omega_\alpha(\cdot)$ the modulus of continuity of $\hat{y}_0(\cdot)$ and by $L(\alpha)$ – the Lipschitz constant of $\hat{u}_0(\cdot)$.

The value of α will be considered as fixed, but the way we fix it will be specified later on.

3.3 Denote by $y^0[x, t]$ the Steiner point of the mapping

$$(t, x) \longrightarrow \hat{K}_0(t, x, \hat{u}_0(t)) \cap [\hat{y}_0(t) + 2\text{dist}(\hat{y}_0(t), \hat{K}_0(t, x, \hat{u}(t)))\mathbf{B}]$$

As above, we can prove that y^0 is continuous. Its modulus of continuity in $[0, T] \times D$ will be denoted by $\omega_0^\alpha(\cdot)$.

4. In the next step we derive some implications of (77),(78), which will lead to a contradiction in the end.

4.1. Consider the system

$$\dot{x}(t) = \xi_1(t, x, y, \hat{u}_0(t)), \quad x(0) = x^0, \tag{79}$$

$$\varepsilon_k \dot{y}(t) \in F_2(t, x, y, \hat{u}_0(t)), \quad y(0) = y^0. \tag{80}$$

Each solution (obviously locally such exist) will be considered on a maximal interval of extendibility $[0, \theta_k(z_k)] \subset [0, T]$ in the set $D \times G$. The set of all (maximal in the above sense) solutions will be denoted by $\tilde{Z}_k$.

For fixed α, k and $z_k(\cdot) = (x_k(\cdot), y_k(\cdot)) \in \tilde{Z}_k$ we denote for brevity $y_k^0(t) = y^0[x_k(t), t]$. Obviously this is a continuous function with modulus $\omega_k^\alpha(h) = \omega_0^\alpha((M+1)h)$ and

$$y_k^0(t) \in \hat{K}_0(t, x_k(t), \hat{u}_0(t)).$$

Moreover,

$$|y_k^0(t) - \hat{y}_0(t)| \leq 2\text{dist}(\hat{y}_0(t), \hat{K}_0(t, x_k(t), \hat{u}_0(t)))$$

$$\leq 2\text{dist}(\hat{y}_0(t), \hat{K}_0(t, x_0(t), \hat{u}_0(t))) + 2\hat{L}|x_k(t) - x_0(t)| = 2\hat{L}|x_k(t) - x_0(t)|.$$

4.2. Denote

$$\rho_k \stackrel{def}{=} \tau_{[0, \theta_k(z_k)]}(y_k(\cdot), y_k^0(\cdot))$$

(here τ is the Tikhonov metric defined in Section 2, taken on the interval $[0, \theta_k(z_k)]$). By the definition of $\tau(\cdot,\cdot)$ we have $\rho_k \leq \theta_k(z_k)$. Since $y_k(t)$, $\hat{y}_0(t) \in G$, we have

$$|y_k(t) - \hat{y}_0(t)| \leq \text{diam}(G) \quad \text{for } t < \rho_k. \tag{81}$$

By the definition of the metric τ for every two bounded functions y' and y'' it holds $|y'(t) - y''(t)| \leq \tau(y', y'')$ for $t > \tau(y', y'')$. Then for $t \in (\rho_k, \theta_k(z_k)]$ (if it has happened that $\theta_k(z_k) > \rho_k$) we have

$$|y_k(t) - \hat{y}_0(t)| \leq |y_k^0(t) - y_k(t)| + |y_k^0(t) - \hat{y}_0(t)| \leq \rho_k + 2\hat{L}|x_k(t) - x_0(t)|. \tag{82}$$

Taking into account the differential equations that $x_k(\cdot)$ and $x_0(\cdot)$ satisfy and using (81) and (82) we obtain

$$|x_k(t) - x_0(t)| \leq L \int_0^t (|x_k(s) - x_0(s)| + |y_k(s) - y_0(s)| + |\hat{u}_0(s) - u_0(s)|)\, ds$$

$$\leq L \int_0^t |x_k(s) - x_0(s)|\, ds + L \int_0^T |y_0(s) - \hat{y}_0(s)|\, ds + L \int_0^{\rho_k} |y_k(s) - \hat{y}_0(s)|\, ds$$

$$+ L \int_{\rho_k}^t |y_k(s) - \hat{y}_0(s)|\, ds + L \int_0^T |\hat{u}_0(s) - u_0(s)|)\, ds$$

$$\leq L \int_0^t |x_k(s) - x_0(s)|\, ds + L\alpha + L\text{diam}(G)\rho_k + LT\rho_k + 2L\hat{L} \int_0^t |x_k(s) - x_0(s)|\, ds + L\alpha.$$

The same inequality obviously holds also for $t \in [0, \rho_k]$. Using the Grunwall inequality we estimate

$$|x_k(t) - x_0(t)| \leq C_1\rho_k + C_2\alpha, \tag{83}$$

where C_1 and C_2 do not depend on α and k. Then (82) implies

$$|y_k(t) - \hat{y}_0(t)| \leq \rho_k + 2\hat{L}(C_1\rho_k + C_2\alpha) \leq C_3\rho_k + C_4\alpha \quad \text{for } t > \rho_k.$$

In particular

$$\|y_k(\cdot) - \hat{y}_0(\cdot)\|_{L_1[0,\theta_k(z_k)]} \leq \rho_k \,\text{diam}(G) + T(C_3\rho_k + C_4\alpha) = C_5\rho_k + C_6\alpha. \tag{84}$$

Since C_2, C_4 and C_6 are independent of α and k, we may suppose that α is fixed in advance so small that

$$C_2\alpha < \frac{\rho}{2}, \quad C_4\alpha < \frac{1}{2}, \quad C_6\alpha < \frac{\sigma}{2}.$$

Then for all $t > \rho_k$ (in case of $\theta_k(z_k) > \rho_k$) we have

$$x_k(t) \in D' + (\frac{\rho}{2} + C_1\rho_k)\mathbf{B}, \quad y_k(t) \in G' + (\frac{1}{2} + C_3\rho_k)\mathbf{B}. \tag{85}$$

Let

$$\mu \stackrel{def}{=} \min\{\frac{1}{2}, \beta_0, \frac{\rho}{4C_1}, \frac{1}{4C_3}, \frac{\sigma}{4C_1}, \frac{\sigma}{4C_5}\}.$$

Assume that $\rho_k \leq 2\mu$ for some k and in the same time $\theta_k(z_k) > 2\mu$. Then (85) implies that $z_k(t) \in \text{int}\,(D \times G)$ for $t \geq \rho_k$, which means that $\theta_k(z_k) = T$. Since the right-hand sides of (83) and (84) are not bigger than σ we obtain a contradiction with (77),(78). Thus we conclude that if (77),(78) are fulfilled, then for every k and for every $z_k \in \tilde{Z}_k$ either

$$\theta_k(z_k) < 2\mu, \tag{86}$$

or $\rho_k > 2\mu$, which implies that

$$|y_k(t) - y^0[x_k(t), t]| > \mu \quad \text{for some } \ t \geq \mu. \tag{87}$$

5. In the next step we prove that the set $\tilde{Z}_k$ contains an element $z_k = (x_k(\cdot), y_k(\cdot))$ for which $|y_k(t) - y^0[x_k(t), t]|$ "quickly" decreases towards zero right from $t = 0$.

5.1. Denote

$$\delta = \delta(\mu/2), \quad \nu = \delta/2, \quad \tau_0 = \tau_0(\nu),$$

where $\delta(\cdot)$ and $\tau_0(\cdot)$ are defined in suppositions **C4** and **C5**, respectively.

Let $\hat{Z}_k$ be the set of those $z_k \in \tilde{Z}_k$ for which

$$|x_k(t) - x_0(t)| \leq \frac{\rho}{2} \quad \text{for all } \ t \in [0, \varepsilon_k \tau_0], \tag{88}$$

$$y_k(t) \in G \quad \text{for all } \ t \in [0, \varepsilon_k \tau_0], \tag{89}$$

$$|y_k(\varepsilon_k \tau_0) - y^0[x_k(\varepsilon_k \tau_0), \varepsilon_k \tau_0]| \leq \delta. \tag{90}$$

5.2. We shall prove that $\hat{Z}_k$ is non-empty for all sufficiently large k. According to Supposition **C5** and in view of the inclusion $y^0[x_k(0), 0] \in \hat{K}_0(0, x^0, \hat{u}_0(0))$ the system

$$\dot{x}(\tau) = 0, \quad x(0) = x^0,$$
$$\dot{y}(\tau) \in F_2(0, x(\tau), y(\tau), \hat{u}_0(0)), \quad y(0) = y^0$$

has a solution $x(\tau) \equiv x^0$, $y(\cdot)$, for which $y(\tau) \in G_0$ for each $\tau \geq 0$ and

$$|y(\tau_0) - y^0[x_k(0), 0]| \leq \nu.$$

Consider on $[0, \tau_0]$ the system

$$\dot{x}(\tau) = \xi_1(\varepsilon_k \tau, x(\tau), y(\tau), \hat{u}_0(\varepsilon_k \tau)), \quad x(0) = x^0, \tag{91}$$

$$\dot{y}(\tau) \in F_2(\varepsilon_k \tau, x(\tau), y(\tau), \hat{u}_0(\varepsilon_k \tau)), \quad y(0) = y^0. \tag{92}$$

We shall apply the Filippov theorem to the last system with the reference trajectory $x(\cdot), y(\cdot)$. For this reason we estimate

$$\Delta(\tau) = |\dot{x}(\tau) - \varepsilon_k \xi_1(\varepsilon_k \tau, x(\tau), y(\tau), \hat{u}_0(\varepsilon_k \tau))| + \text{dist}(\dot{y}(\tau), F_2(\varepsilon_k \tau, x(\tau), y(\tau), \hat{u}_0(\varepsilon_k \tau)))$$

$$\leq \varepsilon_k M + \omega_t(\varepsilon_k\tau_0) + LL(\alpha)\tau_0\varepsilon_k \stackrel{def}{=} \gamma(\alpha, \varepsilon_k).$$

The number α being fixed, the quantity in the right-hand side is arbitrarily small for sufficiently large k. According to the Filippov theorem there is a solution $(\tilde{x}_k(\cdot), \tilde{y}_k(\cdot))$ of (91), (92) on $[0, \tau_0]$ such that

$$|\tilde{x}_k(\tau) - x^0| \leq e^{L\tau_0} \int_0^{\tau_0} \Delta(\tau)\, d\tau \leq e^{L\tau_0}\tau_0\gamma(\alpha, \varepsilon_k) \stackrel{def}{=} d\gamma(\alpha, \varepsilon_k),$$

$$|\tilde{y}_k(\tau) - y(\tau)| \leq d\gamma(\alpha, \varepsilon_k).$$

Hence, $x_k(t) = \tilde{x}_k(t/\varepsilon_k)$, $y_k(t) = \tilde{y}_k(t/\varepsilon_k)$ satisfy (79),(80) on $[0, \varepsilon_k\tau_0]$ and

$$|x_k(t) - x_0(t)| \leq |\tilde{x}_k(t/\varepsilon_k) - x^0| + Mt \leq d\gamma(\alpha, \varepsilon_k) + M\tau_0\varepsilon_k$$

$$|y_k(\varepsilon_k\tau_0) - y^0[x_k(\varepsilon_k\tau_0), \varepsilon_k\tau_0]|$$

$$\leq |\tilde{y}_k(\tau_0) - y(\tau_0)| + |y(\tau_0) - y^0[x_k(0), 0]| + |y^0[x_k(0), 0] - y^0[x_k(\varepsilon_k\tau_0), \varepsilon_k\tau_0]$$

$$\leq d\gamma(\alpha, \varepsilon_k) + \nu + \omega_0^\alpha((M+1)\varepsilon_k\tau_0).$$

Moreover, $x_0(t) \in D'$, $y_k^0(t) \in G'$ and $y(\tau) \in G_0 \subset G'$. Thus, the following inequalities ensure (88)-(90):

$$d\gamma(\alpha, \varepsilon_k) + M\tau_0\varepsilon_k < \frac{\rho}{2},$$

$$d\gamma(\alpha, \varepsilon_k) < 1,$$

$$d\gamma(\alpha, \varepsilon_k) + \nu + \omega_0^\alpha((M+1)\varepsilon_k\tau_0) \leq 2\nu \ \ (= \delta).$$

The number α being fixed, the above inequalities are obviously fulfilled for all sufficiently large k, hence $\hat{Z}_k \neq \emptyset$.

6. The next step will be to prove that for all sufficiently large k and for each $z_k \in \hat{Z}_k$ there exists $t \geq \varepsilon_k\tau_0$ for which

$$|y_k(t) - y^0[x_k(t), t]| > \mu. \tag{93}$$

Take an arbitrary $z_k \in \hat{Z}_k$ and denote for brevity $y_k^0(t) = y^0[x_k(t), t]$. Suppose first that $\theta_k(z_k) > 2\mu$. Then according to the alternative (86),(87) we have (93) for some $t \geq \mu$, hence for some $t \geq \varepsilon_k\tau_0$, provided that k is sufficiently large.

Now we consider the second possibility: $\theta_k(z_k) \leq \mu$. If it occurs that $y_k(\theta_k(z_k)) \in \partial G$, then (93) is also fulfilled at $t = \theta_k(z_k) \geq \varepsilon_k\tau_0$, since $y^0[x_k(t), t] \in G'$ and $\mu < 1$.

If $y_k(\theta_k(z_k)) \in \text{int}\, G$, then $x_k(\theta_k(z_k))$ must belong to the boundary of D, according to the definition of $\theta_k(z_k)$. But this contradicts the inequality $\mu \leq \beta_0$, since $x_k(\cdot)$ satisfies the inclusion (76) on $[0, \theta(z_k)]$, thus $x_k(\theta_k(z_k)) \in D' +$
$rho/2 \subset \text{int}\, D$.

7. In the next step we prove existence of an "extremal" (in a specific sense) element $\bar{z}_k \in \hat{Z}_k$.

7.1 According to Step 6 for each $z_k \in \hat{Z}_k$ (k is sufficiently large) there is a maximal $t = t_k(z_k) \in [\varepsilon_k \tau_0, \theta_k(z_k))$ such that

$$|y_k(t) - y^0[x_k(t), t]| \leq \mu \quad \text{for all} \quad t \in [\varepsilon_k \tau_0, t_k(z_k)].$$

Since according to the definition of $\hat{Z}_k$

$$|y_k(\varepsilon_k \tau_0) - y^0[x_k(\varepsilon_k \tau_0), \varepsilon_k \tau_0]| \leq \delta < \mu$$

there is a last moment $t = t_k^0(z_k) \in [\varepsilon_k \tau_0, t_k(z_k))$ at which still

$$|y_k(t) - y^0[x_k(t), t]| \leq \delta.$$

Recapitulating, we have

$$|y_k(t_k^0(z_k)) - y^0[x_k(t_k^0(z_k)), t_k^0(z_k)]| = \delta,$$

$$\delta < |y_k(t) - y^0[x_k(t), t]| \leq \mu$$

on $(t_k^0(z_k), t_k(z_k)]$ and

$$|y_k(t) - y^0[x_k(t), t]| > \mu$$

for numbers $t > t_k(z_k)$ arbitrarily close to $t_k(z_k)$.

7.2 We shall prove that there exists $\bar{z}_k \in \hat{Z}_k$ such that

$$t_k^0(\bar{z}_k) = t_k^0 \overset{def}{=} \sup_{z_k \in \hat{Z}_k} t_k^0(z_k).$$

For a fixed k (large enough so that the consideration in Step 7.1 is valid) we take a sequence $z_k^i \in \hat{Z}_k$ such that

$$t_k^0(z_k^i) \to t_k^0.$$

Each z_k^i is defined at least on $[0, t_k(z_k^i)]$. Denote

$$\bar{t}_k = \liminf_i t_k(z_k^i)$$

and consider z_k^i on $[0, \bar{t}_k]$ (extending it, if necessary, to the right by $z_k^i(t_k(z_k^i))$). By a standard compactness argument there is a convergent subsequence of $\{z_k^i\}_i$ with limit $\bar{z}_k(\cdot)$ and $\bar{z}_k(\cdot)$ is a solution of (79),(80) on $[0, \bar{t}_k]$. Since (88)–(90) are fulfilled for z_k^i and $y^0[\cdot, \cdot]$ is continuous, the same relations are fulfilled also for $\bar{z}_k$. Thus $\bar{z}_k \in \hat{Z}_k$.

Since

$$|y_k^i(t) - y^0[x_k^i(t), t]| = \delta$$

for $t = t_k^0(z_k^i)$ and $y^0[\cdot, \cdot]$ is continuous, the same equality is fulfilled also for $\bar{z}_k$ and $t = t_k^0$. Since obviously $\varepsilon_k \tau_0 \leq t_k^0 \leq t_k(\bar{z}_k)$ we have $t_k^0(\bar{z}_k) \geq t_k^0$. Thus the supremum of $t_k^0(\cdot)$ on $\hat{Z}_k$ is attained at $\bar{z}_k$.

8. In the last step we shall modify the maximal element $\bar{z}_k = (\bar{x}_k(\cdot), \bar{y}_k(\cdot))$ in a way that contradicts its maximality. This contradiction completes the proof, since it is caused by our assumption that the claim of the theorem is not true.

For the maximal element $\bar{z}_k$ denote $\bar{t}_k = t_k(\bar{z}_k)$, $\bar{y}_k^0(\cdot) = y^0[\bar{x}_k(\cdot), \cdot]$. Denote $\bar{\tau} = \tau_0(\delta, \nu)$, where $\tau_0(\cdot, \cdot)$ is defined in Supposition **C4**.

We have

$$\bar{y}_k^0(t_k^0) \in \hat{K}_0(t_k^0, \bar{x}_k(t_k^0), \hat{u}_0(t_k^0))$$
$$|\bar{y}_k(t_k^0) - \bar{y}_k^0(t_k^0)| = \delta,$$

thus, according to Supposition **C4** the system

$$\begin{aligned} \dot{x}(\tau) &= 0, \quad x(0) = \bar{x}_k(t_k^0), \\ \dot{y}(\tau) &\in F_2(t_k^0, x(\tau), y(\tau), \hat{u}_0(t_k^0)), \quad y(0) = \bar{y}_k(t_k^0) \end{aligned}$$

has a solution $x(\tau) \equiv \bar{x}_k(t_k^0)$, $y(\cdot)$ for which

$$|y(\tau) - \bar{y}_k^0(t_k^0)| \le \frac{\mu}{2} \quad \text{for every} \quad \tau \ge 0,$$

$$|y(\bar{\tau}) - \bar{y}_k^0(t_k^0)| \le \nu.$$

Moreover,

$$x(\tau) = \bar{x}_k(t_k^0) \in D, \quad y(\tau) \in G' + \mu\mathbf{B} \subset G.$$

Then repeating the same Filippov's argument as in Step 5.2 we prove existence of a solution $(\tilde{x}_k(\cdot), \tilde{x}_k(\cdot))$ of

$$\begin{aligned} \dot{x}(t) &= \xi_1(t_k^0 + \varepsilon_k\tau, x(\tau), y(\tau), \hat{u}_0(t_k^0 + \varepsilon_k\tau)), \quad x(0) = \bar{x}_k(t_k^0), \\ \dot{y}(\tau) &\in F_2(t_k^0 + \varepsilon_k\tau, x(\tau), y(\tau), \hat{u}_0(t_k^0 + \varepsilon_k\tau)), \quad y(0) = \bar{y}_k(t_k^0) \end{aligned}$$

such that

$$|\tilde{x}_k(\tau) - x(\tau)| \le \bar{d}\bar{\gamma}(\alpha, \varepsilon_k),$$
$$|\tilde{y}_k(\tau) - y(\tau)| \le \bar{d}\bar{\gamma}(\alpha, \varepsilon_k),$$

where

$$\bar{\gamma}(\alpha, \varepsilon) = \varepsilon M + \omega_t(\varepsilon\bar{\tau}) + LL(\alpha)\varepsilon\bar{\tau}, \quad \bar{d} = e^{L\bar{\tau}}\bar{\tau}.$$

Now we extend $\bar{z}_k$ on $[t_k^0, t_k^0 + \varepsilon_k\bar{\tau}]$ as $\bar{\bar{x}}_k(t) = \tilde{x}_k((t - t_k^0)/\varepsilon_k)$, $\bar{\bar{y}}_k(t) = \tilde{y}_k((t - t_k^0)/\varepsilon_k)$. Obviously it satisfies (79),(80) on $[t_k^0, t_k^0 + \varepsilon_k\bar{\tau}]$. Moreover, for all sufficiently large k

$$|\bar{\bar{y}}_k(t) - y^0[\bar{\bar{x}}_k(t), t]| \le |\bar{\bar{y}}_k(t) - y^0[\bar{\bar{x}}_k(t_k^0), t_k^0]| + \omega_0^\alpha((M+1)\varepsilon_k\bar{\tau})$$

$$= |\bar{\bar{y}}_k(t) - \bar{y}_k^0(t_k^0)| + \omega_0^\alpha((M+1)\varepsilon_k\bar{\tau}) = |\tilde{y}_k((t - t_k^0)/\varepsilon_k) - \bar{y}_k^0(t_k^0)| + \omega_0^\alpha((M+1)\varepsilon_k\bar{\tau})$$

$$\le |\tilde{y}_k((t - t_k^0)/\varepsilon_k) - \bar{y}_k^0(t_k^0)| + |\tilde{y}_k((t - t_k^0)/\varepsilon_k) - y((t - t_k^0)/\varepsilon_k)| + \omega_0^\alpha((M+1)\varepsilon_k\bar{\tau})$$

$$\leq \frac{\mu}{2} + \bar{d\gamma}(\alpha, \varepsilon_k) + \omega_0^{\alpha}((M+1)\varepsilon_k\bar{\tau}) \leq \mu. \tag{94}$$

Similarly, for $t = t_k^0 + \varepsilon_k\bar{\tau}$

$$|\bar{\bar{y}}_k(t) - y^0[\bar{\bar{x}}_k(t), t]| \leq \nu + \bar{d\gamma}(\alpha, \varepsilon_k) + \omega_0^{\alpha}((M+1)\varepsilon_k\bar{\tau}) \leq 2\nu = \delta.$$

The inequality (94) implies that $t_k^0 + \varepsilon_k\bar{\tau} \leq t_k(\bar{\bar{z}}_k)$ and then the last inequality gives $t_k^0(\bar{\bar{z}}_k) \geq t_k^0 + \varepsilon_k\bar{\tau}$. This is a contradiction with the maximality of t_k^0. The proof is complete. Q.E.D.

References

[1] Z. Artstein and A. Vigodner. Singularly perturbed ordinary differential equations with dynamic limits. Preprint.

[2] J.-P. Aubin and A. Cellina. *Differential Inclusions.* Springer Verlag, Berlin, 1984.

[3] J.-P. Aubin and H. Frankowska. *Set-valued Analysis.* Birkhäuser, Boston, Basel, Berlin, 1990.

[4] A. Bensoussan. *Perturbatiom Methods in Optimal Control Problems.* Wiley, New York, 1988.

[5] F. Clarke, Y. Ledyaev, R. Stern, and P. Wolenski. Qualitative properties of trajectories of control systems: a survey. *J. of Dynamical and Control Systems*, 1(1):1–48, 1995.

[6] K. Deimling. *Multivalued Differential Equations.* Walter de Gruyter, Berlin, New York, 1992.

[7] Tz. Donchev and I. Slavov. Averaging method for one side Lipschitz differential inclusions with generalized solutions. To appear.

[8] Tz. Donchev and I. Slavov. Singularly perturbed functional-differential equations. *Set-Valued Analysis*, 3:113–128, 1995.

[9] A. Dontchev, Tz. Donchev, and I. Slavov. On the upper semicontinuity of the set of solutions of differential inclusions with a small parameter in the derivative. IMA Preprint 1169, Institute for Mathematics and Its Applications, University of Minnesota, September 1993.

[10] A. Dontchev, Tz. Donchev, and I. Slavov. A Tikhonov-type theorem for singularly perturbed differential inclusions. *Nonlinear Analysis, TMA*, 26:1547–1554, 1996.

[11] A. Dontchev and I. Slavov. Singular perturbation in a class of nonlinear differential inclusions. In H.-J. Sebastian and K. Tammer, editors, *System Modelling and Optimization*, Lecture Notes in Control and Inf. Sci., 143, pages 273–280. Springer-Verlag, Berlin, Heidelberg, New York, 1991.

[12] A. Dontchev and V. Veliov. Singular perturbations in Mayer's problem for linear systems. *SIAM J. on Control and Optimization*, 21(4):566–581, 1983.

[13] A. Dontchev and V. Veliov. Singular perturbation in linear differential inclusions - critical case. In *Parametric Optimization and Approximation.* Birghäuser, 1985.

[14] V. Gaitsgory. Suboptimization of singularly perturbed control systems. *SIAM J. Contr. Optim.*, 30:1228–1249, 1992. (Translated from Russian).

[15] F. Hoppensteadt. Properties of solutions of ordinary differential equations with small parameters. *Comm. Pure Appl. Math.*, 34:807–840, 1971.

[16] P. Kokotovic. *Singular Perturbation Techniques in Control Theory*, volume 90 of *Lect. Notes in Control and Inf. Sci.*, pages 1–50. Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[17] M. Krastanov. Forward invariant sets, homogeneity and small-time controllability. In *Geometry in Nonlinear Control and Differential Inclusions*, pages 287–300, Warsaw, 1995. Banach Center Publications, vol. 32.

[18] N. Levinson. Perturbations of discontinuous solutions of non-linear systems of differential equations. *Acta. Math.*, 82:71–106, 1950.

[19] H. Khalil P. Kokotovic and J. O'Reilly. *Singular Perturbation Method in Control: Analysis and Design.* Academic Press, 1986.

[20] M. Quincampoix. Singular perturbations for differential equations and inclusions: An approach through constrained systems. Report 9540, CEREMADE (URA CNRS 749), 11 1995. Journal version is to appear.

[21] M. Quincampoix. Singular perturbations for systems of differential inclusions. In *Geometry in Nonlinear Control and Differential Inclusions*, pages 341–348, Warsaw, 1995. Banach Center Publications.

[22] A. Tikhonov. Systems of differential equations containing a small parameter in the derivatives. *Mat Sbornik*, 31(73):572–586, 1952. In Russian.

[23] H. Tuan. Averaging theorem for differential inclusions with slow and fast variables. *Differential Equations*, pages 360–363, 1991.

[24] A.B. Vasil'eva and B.F. Butuzov. *Asymptotic Expansions of Solutions of Syngularly Perturbed Equations.* Nauka, Moscow, 1973. (Russian).

[25] A. Vasilieva. On the development of singular perturbation theory at Moscow State University and elsewhere. *SIAM Review*, 36(3), 1994.

[26] V. Veliov. Sufficient conditions for viability under imperfect measurement. *Set-Valued Analysis*, 1:305–317, 1993.

[27] V. Veliov. Attractiveness and invariance: The case of uncertain measurement. In A.B. Kurzhanski and V.M. Veliov, editors, *Modeling Techniques for Uncertain Systems*. Birkhäuser, Boston, Basel, Berlin, 1994.

[28] V. Veliov. Differential inclusions with stable subinclusions. *Nonlinear Analysis, TMA*, 23(8):1027–1038, 1994.