ILLPOSEDNESS OF STANDARD OPTIMAL CONTROL PROBLEMS BY PERTURBATION OF TIME DELAY

A. MATVEEV

REPORT No. 16, 2002/2003, spring

ISSN 1103-467X
ISRN IML-R- -16-02/03- -SE+spring

INSTITUT MITTAG-LEFFLER
THE ROYAL SWEDISH ACADEMY OF SCIENCES

ILLPOSEDNESS OF STANDARD OPTIMAL CONTROL PROBLEMS BY PERTURBATION OF TIME DELAY

ALEXEY S. MATVEEV*

Abstract. The paper considers a general optimal control problem with a small time delay Δ in both the state and control. The objective is to find a correct way of neglecting the delay or, in other words, to construct a well-posed instantaneous approximation of the delayed problem. The natural way to do so is by mere neglecting the delay in the model. It is shown that this way is almost always incorrect: it gives rise to an error that does not vanish as $\Delta \to +0$, and results in an ill-posed model. A proper approximation of the delayed problem is offered.

Key words. Optimal control, time delay, stability, well-posedness

AMS subject classifications. 49K40, 49K45, 49K25

1. Introduction.

Time delay optimal control problem is a good and realistic option to model a wide range of real-life phenomena. Unfortunately solution of such problems encounters considerable troubles in even the simplest cases. This impels to employ instantaneous models whenever the delay is small. The conventional way to do so is by mere ignoring the delay effects. Is this correct? To get an answer, we consider infinitely small (i.e., converging to zero) delays and analyze the error that accrues from their neglecting. At the first sight, the outcome of such an analysis can be easily foreseen and comes to the platitude: the error must be infinitely small as well. This is true indeed if the delay occurs only in the state. However in the case of the delayed control, this platitudinous property fails to be true not only in general but also "almost always". Then the natural way to ignore the delay results in an ill-posed model and so is not acceptable. The objective of this research is to justify these claims and offer well-posed instantaneous approximations for optimization problems with small delays.

The motivation of the issues addressed in this paper comes from many sources. For example due to the recent growth in communication technology, it is becoming more common to employ digital finite capacity channels for communicating information from controllers to actuators. Not only quantization but also time delay is inherent in transmission via such channels. Moreover serial digital networks has become popular to serve complex systems with spatially distributed components like advanced aircraft, spacecraft, automotive, industrial and defence systems, arrays of microactuators, power control in mobile communication, multi-agent mobile robots, etc. Typically, the delay effects are enhanced in networks. Indeed then transmission delays are combined with ones induced by the protocol and related to resolving conflicts between several data sources sharing a common channel [12]. The above communication delays are usually small and ignored at the theoretical stage of a controller design. There are at least two additional reasons to do so. First in many cases, these delays are irregular, a priory unknown, or hardly computable [12]. Second ironically, the smaller the delay the harder the delayed optimization problem from the computational point of view.

In this paper, we consider an optimal control problem in the situation where controls are transmitted to the actuators over two channels. One of them is instantaneous whereas the other provides a constant delay $\Delta \approx 0$. Modulo a shift of the time, this also includes the case where the delays in the both channels are nonzero and close to each other. All measurable controls are admissible. This situation is the simplest among those worth studying. However, it is enough to reveal the major points of principle and so is a good option to start with.

*Department of Mathematics and Mechanics, Saint Petersburg University, Bybliotechnaya 2, Petrodvoretz, St.Petersburg, 198904, Russia (almat@am1540.spb.edu)

We compare the delayed problem with its natural instantaneous approximation that results from putting $\Delta := 0$. We show that in general this problem is not a proper (well-posed) approximation of the delayed one: the optimal value of the cost functional in the latter does not converge as $\Delta \to +0$ to that in the former. What is more, this phenomenon is robust and occurs for almost all delayed problems: the corresponding problems constitute an open and dense subset in the space of all problems. (Any problem is identified with its data.)

Thus the natural way to neglect the delay is almost always incorrect. What is the correct one? As an answer to this question, a proper "limit" as $\Delta \to 0$ of the delayed problem is found. In brief, this is an instantaneous ($\Delta = 0$) optimal control problem well-posed by perturbation [4, 26] of the delay. Here the perturbed problem is the original Δ-delayed one. More specifically, the minimum values of the cost functionals in the limit and Δ-delayed problems, respectively, are close: the latter converges to the former as $\Delta \to 0$. Moreover, solution of the limit problem gives an exhaustive description of the asymptotically optimal Δ-parametric control sequences. For the Δ-delayed problem, the control from such a sequence provides the level of performance, which differs from the optimal one by a vanishing (as $\Delta \to 0$) quantity. The above description is given by the fact that a sequence is asymptotically optimal if and only if it converges to the set of the optimal controls in the limit problem.

The proper limit problem is not standard, in contrast with the natural instantaneous approximation of the delayed problem. This promotes the interest to the cases where this approximation is yet well-posed, despite of the fact that they are rarely encountered. Such a case occurs if and only if the optimal values of the cost functionals in the two concerned problems coincide. This criterion is exhaustive but not relevant in the current context, since it employs the limit problem. We show that the question about the wellposedness of the natural approximation is undecidable if one's knowledge about the system is confined to only this approximation. Moreover, then it is impossible to estimate to which extent this model may be ill-posed. Thus required is an information about the delayed model, though not the values of the delays. We also offer a simple sufficient criterion. It is far from being exhaustive but deals with a popular situation where the system's equations and the cost functional are linear and convex, respectively, in controls, and the domain of admissible controls is also convex.

Up to now, not much attention was given to stability (wellposedness) of optimal control problems with respect to perturbation of the delay. The research activity was focused on the case where the delay occurs in the state alone [3, 7, 8, 21]. It was shown that under nonrestrictive assumptions, a small perturbation of the delay causes a small disturbance of the optimum. The bounds for certain Dini derivatives of the cost functional optimal value with respect to the delay were also established [7]. By the author's knowledge, the current paper pioneers an analysis of the stability with respect to the delay in the control.

An introduction to delayed optimal control problems can be found in [17, 18, 23]. The study of the maximum principle for such problems was pioneered in [13] and continued in e.g. [1, 2, 5, 6, 14–16, 23]. Relaxation issues were addressed in [19, 20, 23, 24].

The paper is organized as follows. Section 2 offers the statement of the optimization problem. Section 3 shows that the natural limit problem is almost always ill-posed. In Section 4, we discuss when this problem is yet proper. The well-posed limit problem is presented in Section 5. The remainder of the paper is devoted to proofs. They employ certain limits of the set of admissible controls, which are introduced in Section 6. Section 7 reveals their general properties. Computation of these limits in Section 9 is based upon calculation of extremal points of a polygon in the space of matrices, which is done in Section 8. Finally, Sections 10, 11, and 12 contain the proofs of the results stated in Sections 5, 3, and 4, respectively.

2. The problem statement and assumptions. We consider time delay optimal control problems of the form:

$$\text{minimize} \quad \int_T \varphi\left[t, x(t), x(t-\Delta), u(t), u(t-\Delta)\right] \, dt \qquad \text{subject to} \tag{2.1}$$

$$\dot{x}(t) = f\left[t, x(t), x(t-\Delta), u(t), u(t-\Delta)\right] \quad \text{a.a.} \quad t \in T := [t_0, t_1]\,, \tag{2.2}$$

$$x(t) = a(t) \quad \text{for} \quad t \in [t_0 - \Delta, t_0]\,, \tag{2.3}$$

$$u(t) \in \Omega \quad \text{a.a.} \quad t \in [t_0 - \Delta, t_1]\,. \tag{2.4}$$

Here "a.a." means "for almost all". The state $x = x(t) \in \mathbb{R}^n$ and the control $u = u(t) \in \mathbb{R}^m$ are absolutely continuous and measurable functions of time $t \in [t_0 - \Delta, t_1]$, respectively. We interpret the delay $\Delta > 0$ as the parameter of the problem, which is denoted by $\mathfrak{P}(\Delta)$.

DEFINITION 2.1. *The* natural limit problem *is that* $\mathfrak{P}(0)$ *obtained by ignoring the delay* $\Delta := 0$ *in* (2.1)—(2.4)*:*

$$\text{minimize} \quad \int_{t_0}^{t_1} \overline{\varphi}\left[t, x(t), u(t)\right] dt \quad \text{subject to} \tag{2.5}$$

$$\dot{x}(t) = \overline{f}\left[t, x(t), u(t)\right], \quad u(t) \in \Omega \quad \text{a.a.} \quad t \in T, \quad x(t_0) = a_0, \quad \text{where} \tag{2.6}$$

$$\overline{\varphi}(t, x, u) := \varphi(t, x, x, u, u), \quad \overline{f}(t, x, u) := f(t, x, x, u, u), \quad a_0 := a(t_0). \tag{2.7}$$

Can this problem be considered as a proper (well-posed) instantaneous approximation of the delayed one $\mathfrak{P}(\Delta)$ with $\Delta \approx 0$? The answer is in the affirmative only if

$$\lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta) = \mathfrak{I}_{opt}(0), \tag{2.8}$$

where $\mathfrak{I}_{opt}(\Delta)$ is the infimum of the cost functional in the problem $\mathfrak{P}(\Delta)$. Our first objective is to check whether relation (2.8) is true or not.

For the definiteness, we consider delays $\Delta \in (0, 1)$. Suppose that, in (2.1)-(2.4),

(i) the set $\Omega \subset \mathbb{R}^m$ is compact and contains at least two points;

(ii) the functions $f(t, x_1, x_2, u_1, u_2) \in \mathbb{R}^n$ and $\varphi(t, x_1, x_2, u_1, u_2) \in \mathbb{R}$ of the variables $t \in T = [t_0, t_1], x_1, x_2 \in \mathbb{R}^n$, and $u_1, u_2 \in \Omega$ are continuous and continuously differentiable with respect to x_1 and x_2;

(iii) the function $a(\cdot) : [t_0 - 1, t_0] \to \mathbb{R}^n$ is absolutely continuous;

(iv) for any $\Delta \in [0, 1)$ and any measurable control $u(\cdot) : [t_0 - \Delta, t_1] \to \Omega$, the Cauchy problem (2.2),(2.3) has a solution $x(\cdot) : [t_0 - \Delta, t_1] \to \mathbb{R}^n$ such that $|x(t)| \le k < \infty$, where the constant k does not depend on $t \in [t_0 - \Delta, t_1]$, $u(\cdot)$, and $\Delta \in [0, 1)$.

It follows from (i)-(iv) that $\mathfrak{I}_{opt}(\Delta) > -\infty$.

3. Illposedness of the natural limit problem. We start with a useful technical fact.

LEMMA 3.1. *The limit* $\lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta)$ *in* (2.8) *exists and is finite.*

The proofs of the results presented in this section will be given in Section 11.

Our first theorem shows that in general, the natural limit problem is ill-posed.

THEOREM 3.2. *Given an interval* $T = [t_0, t_1]$ *and a set* Ω *satisfying* (i)*, there exists a problem* (2.1)-(2.4) *for which the properties* (ii)-(iv) *hold and relation* (2.8) *violates:*

$$\lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta) < \mathfrak{I}_{opt}(0). \tag{3.1}$$

This problem can be chosen so that the function $f(t, x_1, x_2, u_1, u_2)$ in (2.2) is linear in x_1, x_2, u_1, u_2 and independent of t, the function $\varphi(t, x_1, x_2, u_1, u_2)$ in (2.1) does not depend on t, x_1, x_2 and is a polynomial of the second degree in u_1, u_2, and in (2.3), $a(\cdot) \equiv$ const.

Inequality (3.1) means that the natural limit problem conceals certain possibilities to improve the performance.

Now we are going to analyze whether the property (3.1) is typical or not. In doing so, it is natural to compare the variety of all problems (2.1)-(2.4) with the set of those for which (3.1) holds. To this end, we fix the interval T and the set Ω and identify the problem (2.1)-(2.4) with the triplet of functions $\xi := [f(\cdot), \varphi(\cdot), a(\cdot)]$. We also equip the linear space Ξ of all such triplets ξ satisfying (ii) and (iii) with the locally convex topology generated by the following c-parametric ($c > 0$) family of hemi-norms

$$\mathfrak{H}_c(\xi) := \sum_{r = f, \partial f/\partial x_1, \partial f/\partial x_2, \varphi} \max |r(t, x_1, x_2, u_1, u_2)| + \max_{t \in [t_0 - 1, t_0]} |a(t)|, \tag{3.2}$$

where the first maximum is over $t \in T, |x_i| \leq c, u_i \in \Omega$. It can be shown that the subset $\Xi_{(iv)} \subset \Xi$ of all the triplets $\xi \in \Xi$ satisfying (iv) is open.

THEOREM 3.3. *Given an interval* $T = [t_0, t_1]$ *and a set* Ω *satisfying* (i)*, consider the set* $\Xi_{(3.1)}$ *of all the triplets* $\xi \in \Xi_{(iv)}$ *for which* (3.1) *is valid. The set* $\Xi_{(3.1)}$ *is nonempty and open. If the domain* Ω *is connected, then the set* $\Xi_{(3.1)}$ *is dense in* $\Xi_{(iv)}$.

The second claim means that the phenomenon (3.1) is typical.

4. Problems for which the natural limit model is proper. In this section, we discuss how to recognize problems for which (2.8) is true. This remains of interest despite of the fact that by Theorem 3.3, relation (2.8) is highly probable to be violated because of errors in our knowledge of ξ. However, the both quantities in it can be shown to depend on ξ continuously. So whenever (2.8) holds for the nominal model, small errors in our knowledge of ξ cannot violate (2.8) much. Furthermore the following lemma shows that in the case (2.8), the control optimal for the natural limit problem is suboptimal for the delayed one.

LEMMA 4.1. *Let the hypotheses* (i)-(iv) *be fulfilled and a measurable function* $u(\cdot) : [t_0 - 1, t_1] \to \Omega$ *be given. If the control* $u|_{[t_0, t_1]}$ *is optimal in the natural limit problem* $\mathfrak{P}(0)$*, the control* $u|_{[t_0 - \Delta, t_1]}$ *is* $\mu(\Delta)$*-suboptimal in the primal one* (2.1)-(2.4)*. More precisely, the corresponding value of the cost functional does not exceed* $\mathfrak{I}_{opt}(\Delta) + \mu(\Delta)$*, where*

$$\mu(\Delta) \leq \left| \lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta) - \mathfrak{I}_{opt}(0) \right| + \varepsilon(\Delta) \quad \text{and} \quad \varepsilon(\Delta) \to 0 \qquad \textit{as} \quad \Delta \to +0. \tag{4.1}$$

DEFINITION 4.2. *If* (2.8) *is true, the problem* (2.1)-(2.4) *is said to be* stable to the delay.

By the next theorem, one is unable to recognize such a stability and estimate the discrepancy between the left and right hand sides of (2.8) if only the natural limit model is known.

THEOREM 4.3. *Consider the problem* (2.5),(2.6)*, where the vector* $a_0 \in \mathbb{R}^n$*, the set* $\Omega \subset \mathbb{R}^m$*, and the functions* $\overline{\varphi}(\cdot), \overline{f}(\cdot)$ *are given and satisfy* (i), (ii), (iv) *(provided we put* $f(t, x_1, x_2, u_1, u_2) := \overline{f}(t, x_1, u_1)$ *and* $\varphi(t, x_1, x_2, u_1, u_2) := \overline{\varphi}(t, x_1, u_1)$ *in* (ii),(iv)*).*

Given $A > 0$*, there exists a problem* (2.1)-(2.4) *satisfying* (i)-(iv) *and such that the original problem* (2.5),(2.6) *is its natural limit case (i.e.,* (2.7) *is true) and*

$$\left| \lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta) - \mathfrak{I}_{opt}(0) \right| > A. \tag{4.2}$$

Necessary and sufficient conditions for the problem (2.1)-(2.4) to be stable to the delay will be given in the next section. So far as they are quite complicated, we adduce here simple but useful sufficient conditions for such a stability.

THEOREM 4.4. *Let the hypotheses* (i)-(iv) *be valid. Suppose that the set Ω is convex, the function $\varphi(t, x_1, x_2, u_1, u_2)$ is convex in (u_1, u_2) for any t, x_1, x_2, and the function $f(\cdot)$ disintegrates into the sum $f(t, x_1, x_2, u_1, u_2) = f_1(t, x_1, x_2, u_1) + f_2(t, x_1, x_2, u_2)$ with the continuous summands $f_1(\cdot)$ and $f_2(\cdot)$ that are linear in u_1 and u_2, respectively.*[1] *Then the problem* (2.1)-(2.4) *is stable with respect to the delay.*

5. The proper instantaneous (limit) model of the delayed optimal control problem. We recall that this is an instantaneous (undelayed) optimal control problem $\mathcal{P}$ such that first, the minimum of the cost functional in $\mathcal{P}$ equals the limit from (3.1) and second, solution of this problem provides a complete characterization of asymptotically optimal sequences of controls in the original Δ-parametric family of problems (2.1)-(2.4). These sequences $\{u_\Delta(\cdot)\}$ are defined to be those for which $u_\Delta(\cdot)$ is "almost optimal" in $\mathfrak{P}(\Delta)$ for $\Delta \approx 0$:

$$I_\Delta\left[u_\Delta(\cdot)\right] - \mathfrak{I}_{opt}(\Delta) \to 0 \quad \text{as} \quad \Delta \to +0. \tag{5.1}$$

Here $I_\Delta\left[u_\Delta(\cdot)\right]$ is the value of the cost functional from (2.1). More specifically, the sequence is asymptotically optimal if and only if it converges to the optimal control in the proper limit problem $\mathcal{P}$ if this control is unique, and to the set of all such controls otherwise.

To introduce $\mathcal{P}$, we need some notations. Denote by $\mathfrak{FM}(\Omega \times \Omega)$ the collection of all Radon (finite regular Borel) measures in $\Omega \times \Omega$ and put

$$\mathfrak{SPM}(\Omega \times \Omega) := \Big\{\eta \in \mathfrak{FM}(\Omega \times \Omega) : \eta(du_1, du_2) \geq 0, \quad \eta(\Omega \times \Omega) = 1, \\ \text{and} \quad \eta\left(du, \Omega\right) = \eta\left(\Omega, du\right)\Big\}. \tag{5.2}$$

(The last relation means that $\eta(E \times \Omega) = \eta(\Omega \times E)$ for any Borel set $E \subset \Omega$.) By the Riesz theorem, $\mathfrak{FM}(\Omega \times \Omega)$ is the topological dual to the space $\mathbb{C}(\Omega \times \Omega)$ of continuous functions $g : \Omega \times \Omega \to \mathbb{R}$. We endow $\mathfrak{FM}(\Omega \times \Omega)$ and, thereby, $\mathfrak{SPM}(\Omega \times \Omega)$ with the corresponding weak star topology. Denote by $\mathfrak{CM}$ the set of all measurable maps $\nu : T \to \mathfrak{SPM}(\Omega \times \Omega)$. They are controls in the **proper limit problem**, which looks as follows

$$\text{minimize} \quad \int_T dt \int_{\Omega\times\Omega} \varphi\left[t, x(t), x(t), w\right] \nu(t, dw) \quad \text{subject to} \quad \nu \in \mathfrak{CM}, \tag{5.3}$$

$$\dot{x}(t) = \int_{\Omega\times\Omega} f\left[t, x(t), x(t), w\right] \nu(t, dw), \qquad x(t_0) = a_0 := a(t_0). \tag{5.4}$$

THEOREM 5.1. *Let the hypotheses* (i)-(iv) *be fulfilled. In the problem* (5.3), (5.4)*, the minimum value of the cost functional is archived and equals* $\lim\limits_{\Delta\to+0} \mathfrak{I}_{opt}(\Delta)$.

The proofs of the results presented in this section will be given in Section 10.

We note that the use of measure-valued controls is a customary approach in optimal control theory [11, 23, 25]. In the case under consideration, the point of novelty comes to imposing the last restriction from (5.2). Theorem 5.1 gives rise to the following fact.

[1]More precisely, the last claim means that $f_i(t, x_1, x_2, u_i) = A_i(t, x_1, x_2)u_i$, where the $n \times m$-matrix-functions $A_i(t, x_1, x_2)$ are continuous.

COROLLARY 5.2. *The problem* (2.1)-(2.4) *is stable to the delay if and only if the infimum value of the cost functional in the problem* (2.5),(2.6) *equals that in* (5.3), (5.4).

The problem (5.3), (5.4) has several advantages over (2.1)-(2.4). For example, along the lines of the theory of functional differential equations [10], the state space in the latter should be viewed as infinite dimensional in contrast with the former. Though another approach [22] reduces the problem (2.1)-(2.4) to one without delay and with the state space of finite dimension N, this dimension is much greater $N \geq n(t_1 - t_0)/\Delta \ggg n$ than that in (5.3), (5.4) if $\Delta \approx 0$. The both interpretations reflect the point of the matter and affect the formulations of basic optimization results such as the maximum or dynamic programming principles [23].

The problem (5.3), (5.4) has deeper relations to the primal one (2.1)-(2.4) as compared with those indicated in Theorem 5.1. For example, the set U_Δ of admissible controls $u : [t_0 - \Delta, t_1] \to \Omega$ in the problem (2.1)-(2.4) converges to the set $\mathfrak{CM}$ from (5.3) as $\Delta \to 0$. To elucidate this claim, we need to embed both $\mathfrak{CM}$ and U_Δ into a common enveloping space. To this end, we denote by $\mathbb{L}_1\mathbb{C}(T, K)$ (where K is a compact topological space) the Banach space of the equivalence classes of maps $g : T \times K \to \mathbb{R}$ such that the function $g(\cdot, \varkappa)$ is measurable for all $\varkappa \in K$, the function $g(t, \cdot)$ is continuous for almost all $t \in T$, and

$$\|g\| := \int_T \max_{\varkappa \in K} |g(t, \varkappa)| \, dt < \infty. \tag{5.5}$$

The above equivalence $g_1 \backsim g_2$ holds if and only if $g_1(t, \cdot) \equiv g_2(t, \cdot)$ for almost all $t \in T$. The norm in $\mathbb{L}_1\mathbb{C}(T, K)$ is given by (5.5). We also denote by $\mathfrak{FM}(K)$ and $\mathfrak{PM}(K)$ the collections of all the Radon and Radon probability measures in K, respectively, and put

$$\|\nu\| := \sup \int g(\varkappa)\nu(d\varkappa) \qquad \forall\, \nu \in \mathfrak{FM}(K),$$

where sup is over $g(\cdot) \in \mathbb{C}(K)$ with $\max_\varkappa |g(\varkappa)| \leq 1$. Denote by $\mathfrak{F}(T, K)$ the space of equivalence classes of measurable functions $\nu : T \to \mathfrak{FM}(K)$ such that ess $\sup_t \|\nu(t)\| < \infty$. This space can be put in duality to $\mathbb{L}_1\mathbb{C}(T, K) = \{g(\cdot)\}$ by setting

$$\langle \nu, g \rangle := \int_T dt \int_K g(t, \varkappa)\nu(t, d\varkappa) \tag{5.6}$$

[23]. By the Bishop and Dunford-Pettis theorems, the relativization to $\mathfrak{P}(T, K) := \{\nu(\cdot) \in \mathfrak{F}(T, K) : \nu(t) \in \mathfrak{PM}(K) \text{ a.a. } t \in T\}$ of the weak topology generated on $\mathfrak{F}(T, K)$ by the above duality is induced on $\mathfrak{P}(T, K)$ by a certain norm $\|\cdot\|_w$ in $\mathfrak{F}(T, K)$ (weak norm). The space $\mathfrak{P}\,(T) := \mathfrak{P}(T, \Omega \times \Omega)$ will be of special interest further. We denote by dist(A, B) the Hausdorff distance between the sets $A, B \subset \mathfrak{P}\,(T)$, i.e.,

$$\mathrm{dist}(A, B) := \max \left\{ \sup_{\nu \in A} \inf_{\mu \in B} \|\nu - \mu\|_w ; \sup_{\mu \in B} \inf_{\nu \in A} \|\nu - \mu\|_w \right\}. \tag{5.7}$$

Given $\Delta \in (0, 1)$, we embed U_Δ into $\mathfrak{P}\,(T)$ by identifying $u(\cdot) \in U_\Delta$ with the measure-valued function $\nu : T \to \mathfrak{PM}(\Omega \times \Omega)$, where $\nu(t)$ is the Dirac measure at the point $[u(t), u(t - \Delta)] \in \Omega \times \Omega$, i.e.,

$$\int r(u', u'')\nu(t, du', du'') := r[u(t), u(t - \Delta)] \tag{5.8}$$

for any $r(\cdot) \in \mathbb{C}(\Omega \times \Omega)$. Thus, $U_\Delta \subset \mathfrak{P}\,(T)$. The inclusion $\mathfrak{CM} \subset \mathfrak{P}\,(T)$ is obvious.

LEMMA 5.3. *The Hausdorff distance between* $\mathfrak{CM}$ *and* U_Δ *converges to* 0 *as* $\Delta \to +0$.

It readily follows from this lemma that any $\nu^0(\cdot) \in \mathfrak{CM}$ can be approximated

$$u_\Delta(\cdot) \to \nu^0(\cdot) \quad \text{as} \quad \Delta \to +0 \tag{5.9}$$

by "ordinary" controls $u_\Delta(\cdot) \in U_\Delta$. A number of algorithms can be proposed to construct such a sequence $\{u_\Delta(\cdot)\}$ explicitly. Its importance is demonstrated by the following theorem.

THEOREM 5.4. *Let the hypotheses* (i)-(iv) *be fulfilled and $\nu^0(\cdot)$ be an optimal control in the problem* (5.3), (5.4). *Consider a sequence $\{u_\Delta(\cdot)\}_{\Delta\in(0,1)}$ of "ordinary" controls $u_\Delta(\cdot) \in U_\Delta$, the solution $x_\Delta(\cdot)$ of the Cauchy problem* (2.2),(2.3) *for $u(\cdot) := u_\Delta(\cdot)$, and the trajectory $x^0(\cdot)$ of the system* (5.4) *related to the control $\nu^0(\cdot)$.*

If (5.9) *is true, the sequence $\{u_\Delta(\cdot)\}$ is asymptotically optimal, i.e.,* (5.1) *holds, and*

$$\max_{t\in T} \left|x_\Delta(t) - x^0(t)\right| \to 0 \quad as \quad \Delta \to 0.$$

The last theorem shows that a sequence is asymptotically optimal only if (5.9) holds.

THEOREM 5.5. *Let the optimal control $\nu^0(\cdot)$ in the problem* (5.3), (5.4) *be unique. A sequence $\{u_\Delta(\cdot)\}_{\Delta\in(0,1)}$ of "ordinary" controls $u_\Delta(\cdot) \in U_\Delta$ is asymptotically optimal if and only if it converges to $\nu^0(\cdot)$, i.e.,* (5.9) *holds.*

In general, such a sequence is asymptotically optimal if and only if it converges to the set $\mathfrak{OP}$ of all optimal controls $\nu^0(\cdot)$ in the problem (5.3), (5.4)*, i.e.,*

$$\text{dist } [u_\Delta(\cdot), \mathfrak{OP}] := \inf_{\nu^0(\cdot)\in\mathfrak{OP}} \|u_\Delta(\cdot) - \nu^0(\cdot)\|_w \to 0 \quad \text{as} \quad \Delta \to +0. \tag{5.10}$$

Theorems 5.1 and 5.5 mean that the problem $\mathcal{P}$ is well-posed by perturbation of the delay [4, 26], provided (2.1)—(2.4) is interpreted as the perturbed problem.

6. Limits of the set of admissible controls. In the remainder of the paper, we prove the results stated above. The key to the proofs is studying some limits of the Δ-parametric set U_Δ of admissible controls as $\Delta \to 0$. In this section, we introduce these limits and reveal their role. For the sake of further generalizations, we consider a general Δ-parametric set here.

Let we be given a Δ-parametric (here and throughout $\Delta \in (0,1)$) set $\mathfrak{U}_\Delta$ of measurable controls $u(\cdot) : [t_0 - \Delta, t_1] \to \Omega$. Further we employ the notations introduced before Lemma 5.3 and embed $\mathfrak{U}_\Delta$ into $\mathfrak{P}(T)$ by identifying $u(\cdot) \in \mathfrak{U}_\Delta$ with the measure-valued function $\nu : T \to \mathfrak{PM}(\Omega\times\Omega)$ given by (5.8). We also introduce the following "limits"

$$\text{LIM}_{\Delta\to 0}\mathfrak{U}_\Delta := \big\{\nu \in \mathfrak{P}(T) : \text{ there exists a sequence } \{u_\Delta(\cdot)\} \text{ of controls } u_\Delta(\cdot) \in \mathfrak{U}_\Delta \ \forall\Delta \text{ such that } u_\Delta(\cdot) \to \nu \text{ as } \Delta \to 0\big\}, \tag{6.1}$$

$$\text{LIM}^{\text{seq}}_{\Delta\to 0}\mathfrak{U}_\Delta := \big\{\nu \in \mathfrak{P}(T) : \text{ there exist sequences } \{\Delta_i\}_{i=1}^\infty \subset (0,1) \text{ and } \{u_i(\cdot)\}_{i=1}^\infty \text{ such that } \quad u_i(\cdot) \in U_{\Delta_i} \ \forall i \text{ and } \Delta_i \to 0, u_i(\cdot) \to \nu \text{ as } i \to \infty\big\}. \tag{6.2}$$

(The space $\mathfrak{P}(T)$ is endowed with the week topology generated by the duality (5.6).)

DEFINITION 6.1. Modified *problems* (2.1)-(2.4) *and* (5.3), (5.4) *are defined to be those resulting from the original problems* (2.1)-(2.4) *and* (5.3), (5.4) *by substituting the inclusion $u(\cdot) \in \mathfrak{U}_\Delta$ in place of* (2.4) *and the set* $\text{LIM}_{\Delta\to 0}\mathfrak{U}_\Delta$ *in place of* $\mathfrak{CM}$ *in* (5.3), *respectively.*

In this section, we show that under natural assumptions, the proper limit as $\Delta \to 0$ of the first modified problem is the second one. To this end, we start with two technical facts.

LEMMA 6.2. *Both sets* (6.1) *and* (6.2) *are closed in* $\mathfrak{P}(T)$.

Proof. Let $\mathfrak{S}$ denote the set (6.1) and $\nu \in \overline{\mathfrak{S}}$. Then $\|\nu - \nu_j\|_w < j^{-1}$ with some $\nu_j \in \mathfrak{S}$ for $j = 1, 2, \ldots$. By (6.1), there exist sequences $\{u_\Delta^{(j)}\}$ of controls $u_\Delta^{(j)} \in \mathfrak{U}_\Delta$ such that $u_\Delta^{(j)} \to \nu_j$ as $\Delta \to +0$ for $j = 1, 2, \ldots$. Hence $\|\nu_j - u_\Delta^{(j)}\|_w < j^{-1}$ whenever $\Delta < \Delta_j$ and $\Delta_j > 0$ is small enough. Without the loss of generality, we can assume that $\Delta_{j+1} < \Delta_j \ \forall j$ and $\Delta_j \to 0$ as $j \to \infty$. Putting $u_\Delta := u_\Delta^{(j)} \in \mathfrak{U}_\Delta$ for $\Delta \in (\Delta_{j+1}, \Delta_j]$, we get

$$\|\nu - u_\Delta\|_w \leq \|\nu - \nu_j\|_w + \|\nu_j - u_\Delta\|_w \leq 2j^{-1} \to 0 \quad \text{as} \quad \Delta \to +0.$$

Therefore, $\nu \in \mathfrak{S}_{\lim}$ and $\overline{\mathfrak{S}} = \mathfrak{S}$. The closedness of the set (6.2) is established likewise. □

REMARK 6.3. *Since $\mathfrak{P}(T)$ is compact, so are the sets* (6.1), (6.2) *by Lemma* 6.2.

LEMMA 6.4. *For any $\nu \in \mathfrak{S} := \mathrm{LIM}^{seq}_{\Delta \to 0} U_\Delta$, the Cauchy problem* (5.4) *has a unique solution $x(\cdot|\nu) : T \to \mathbb{R}^n$. This solution obeys the estimation $|x(\cdot|\nu)| \leq k \ \forall t \in T$, where k is taken from* (iv) *in Section 3. Let $\{\Delta_i\}_{i=1}^\infty \subset (0, 1)$ and $\{u_i\}_{i=1}^\infty$ be two sequences such that $u_i \in \mathfrak{U}_{\Delta_i} \ \forall i$ and $\Delta_i \to 0, u_i \to \nu$ as $i \to \infty$. Then*

$$\max_{t \in T} |x_i(t) - x(t|\nu)| \to 0 \quad \textit{and} \quad I_{\Delta_i}(u_i) \to I(\nu) \quad \textit{as} \quad i \to \infty. \tag{6.3}$$

Here $x_i(\cdot)$ is the solution of the Cauchy problem (2.2), (2.3) *with $\Delta := \Delta_i$, $u := u_i$, whereas $I_\Delta(u)$ and $I(\nu)$ are the values of the cost functionals generated by $u \in U_\Delta$ and $\nu \in \mathfrak{S}$ in the problems* (2.1)–(2.4) *and* (5.3), (5.4)*, respectively. The map $\nu \in \mathfrak{S} \mapsto I(\nu)$ is continuous.*

Proof. By (6.2), any $\nu \in \mathfrak{S}$ is related to sequences $\{\Delta_i\}$ and $\{u_i\}$ with the properties given by the third sentence of the lemma. The Arzela's theorem and (iv) from Section 3 imply that $x_i(\cdot)$ is defined on $[t_0 - \Delta_i, t_1]$ and $|x_i(t)| \leq k \ \forall t \in T, i = 1, 2, \ldots$ and also that the sequence $\{x_i|_T\}$ is pre-compact in $\mathbb{C}(T, \mathbb{R}^n)$. It is easy to check (see, e.g., [23, Theorem VII.1.2]) that each of its limit points satisfies (5.4). Thus relations (5.4) do have a solution such that $|x(t|\nu)| \leq k \ \forall t$. Its uniqueness results from (ii) in Section 3. In particular, all the limit points of the pre-compact sequence $\{x_i(\cdot)\}$ equal $x(\cdot|\nu)$, which implies the first relation from (6.3). We recall that the control $u_i(\cdot) \in \mathfrak{U}_{\Delta_i} \subset U_{\Delta_i}$ is identified with the element $\nu_i(t, du_1, du_2) = \delta_{[u_i(t), u_i(t-\Delta_i)]}(du_1, du_2)$ of $\mathfrak{P}(T)$, where δ_ω is the Dirac measure at the point ω. Denote $\widetilde{\Delta}_i := \Delta_i$. By (2.1) and (5.3), the quantity $I_{\Delta_i}(u_i) - I(\nu)$ amounts to

$$\int\limits_T dt \int\limits_{\Omega \times \Omega} \left\{ \varphi\left[t, x_i(t), x_i(t - \widetilde{\Delta}_i), w\right] - \varphi\left[t, x(t|\nu), x(t|\nu), w\right] \right\} \nu_i(t, dw)$$
$$+ \int\limits_T dt \int\limits_{\Omega \times \Omega} \underbrace{\varphi\left[t, x(t|\nu), x(t|\nu), w\right]}_{\eta(t,w)} \left[\nu_i(t, dw) - \nu(t, d\omega)\right]. \tag{6.4}$$

The integrand in the first integral converges to 0 uniformly over $t \in T, w \in \Omega \times \Omega$ by the first relation from (6.3). By (5.6), the second summand equals $\langle \nu_i - \nu, \eta \rangle$. Hence (6.3) does hold.

Now let $\{\nu_i\} \subset \mathfrak{S}$ and $\nu_i \to \nu \in \mathfrak{S}$ as $i \to \infty$. Denote $x_i(\cdot) := x(\cdot|\nu_i)$. Retracing the above arguments ensures consecutively that the sequence $\{x_i(\cdot)\}$ is pre-compact in $\mathbb{C}(T, \mathbb{R}^n)$, each of its limit points satisfies (5.4) and so equals $x(\cdot|\nu)$ and hence $\max_{t \in T} |x_i(t) - x(t|\nu)| \to 0$ as $i \to \infty$. Then the last claim of the lemma is proved like the second relation from (6.3) since the quantity $I(\nu_i) - I(\nu)$ has the form (6.4) with $\widetilde{\Delta}_i := 0$. □

The following three theorems in fact mean that one of the modified problems introduced by Definition 6.1 is the proper limit as $\Delta \to +0$ of the other.

THEOREM 6.5. *Suppose that $\mathfrak{S} := \mathrm{LIM}^{seq}_{\Delta \to 0} \mathfrak{U}_\Delta = \mathrm{LIM}_{\Delta \to 0} \mathfrak{U}_\Delta$ for a Δ-parametric set $\mathfrak{U}_\Delta \subset U_\Delta$ of controls, where the limits are defined by* (6.1) *and* (6.2)*. Denote by $\mathfrak{I}^m_{opt}(\Delta)$ the infimum of the cost functional in the modified problem* (2.1)-(2.4)*. The minimum of the cost functional in the modified problem* (5.3), (5.4) *is attained and equals $\lim_{\Delta \to +0} \mathfrak{I}^m_{opt}(\Delta)$.*

Proof. By Remark 6.3, the set $\mathfrak{S}$ is compact. So thanks to the last claim from Lemma 6.4, $\inf_{\nu\in\mathfrak{S}} I(\nu)$ is attained at a control $\nu^0 \in \mathfrak{S}$. Due to (6.1), there exists a sequence $\{u_\Delta\}$ such that $u_\Delta \in \mathfrak{U}_\Delta$ and $u_\Delta \to \nu^0$ as $\Delta \to +0$. By (6.3), $\mathfrak{I}^m_{opt}(\Delta) \leq I_\Delta(u_\Delta) \to I(\nu^0)$ and so

$$\overline{\lim}_{\Delta\to+0}\mathfrak{I}^m_{opt}(\Delta) \leq I(\nu^0).$$

On the other hand, there apparently exist sequences $\{\Delta_i\}$ and $\{v_i\}$ such that $v_i \in \mathfrak{U}_{\Delta_i}$ $\forall i$ and $\Delta_i \to 0$, $I_{\Delta_i}(v_i) \to \underline{\lim}_{\Delta\to+0}\mathfrak{I}^m_{opt}(\Delta)$ as $i\to\infty$. So far as the space $\mathfrak{P}(T) \supset \{v_i\}$ is compact, passing to a subsequence ensures that $v_i \to \nu$ as $i\to\infty$, where $\nu \in \text{LIM}^{\text{seq}}_{\Delta\to+0}\mathfrak{U}_\Delta = \mathfrak{S}$ in correspondence with (6.2). By invoking (6.3), we get

$$\underline{\lim}_{\Delta\to+0}\mathfrak{I}^m_{opt}(\Delta) = \lim_{i\to\infty} I_{\Delta_i}(v_i) = I(\nu) \geq I(\nu^0).$$

Thus $\mathfrak{I}^m_{opt}(\Delta) \to I(\nu^0)$ as $\Delta \to +0$, which completes the proof. □

THEOREM 6.6. *Suppose that the assumption of Theorem* 6.5 *is fulfilled. The Hausdorff distance* (5.7) *between the sets* $\mathfrak{S} := \text{LIM}_{\Delta\to+0}\mathfrak{U}_\Delta$ *and* $\mathfrak{U}_\Delta$ *vanishes as* $\Delta \to +0$.

Proof. Suppose the contrary. Then due to (5.7), there exist $\delta > 0$ and a sequence $\{\Delta_i\} \subset (0,1)$ such that $\Delta_i \to 0$ as $i\to\infty$ and either

(i) $\inf_{u\in\mathfrak{U}_{\Delta_i}}\|\nu_i - u\|_w \geq \delta$ $\forall i$ for some sequence $\{\nu_i\} \subset \mathfrak{S}$ or

(ii) $\inf_{\mu\in\mathfrak{S}}\|\mu - u_i\|_w \geq \delta$ $\forall i$ for some sequence $\{u_i\}$, $u_i \in \mathfrak{U}_{\Delta_i}$ $\forall i$.

By Remark 6.3, both sets $\mathfrak{P}(T)$ and $\mathfrak{S}$ are compact. So passing to subsequences ensures that $\nu_i \to \nu \in \mathfrak{S}$, $u_i \to \xi \in \mathfrak{P}(T)$ as $i\to\infty$. Due to (6.2), $\xi \in \mathfrak{S}$, which implies that (ii) cannot be true and (i) thereby holds. By (6.1), $\nu = \lim_{i\to\infty} w_i$ for some $w_i \in \mathfrak{U}_{\Delta_i}$, $i = 1,2,\ldots$. So

$$\inf_{u\in\mathfrak{U}_{\Delta_i}}\|\nu_i - u\|_w \leq \|\nu_i - w_i\|_w \leq \|\nu_i - \nu\|_w + \|\nu - w_i\|_w \to 0 \quad \text{as} \quad i\to\infty$$

in violation of (i). The contradiction obtained proves the theorem. □

THEOREM 6.7. *Let the assumption of Theorem* 6.5 *be true. Employ the notation* $\mathfrak{I}^m_{opt}(\Delta)$ *from it. A sequence* $\{u_\Delta(\cdot)\}$ *of controls* $u_\Delta(\cdot) \in \mathfrak{U}_\Delta$ *is asymptotically optimal in the modified problem* (2.1)—(2.4)*, i.e.,* $I_\Delta[u_\Delta] - \mathfrak{I}^m_{opt}(\Delta) \to 0$ *as* $\Delta \to +0$*, if and only if it converges to the set* $\mathfrak{O}\mathfrak{P}$ *of all optimal controls in the modified problem* (5.3), (5.4)*, i.e.,* (5.10) *holds.*

Proof. Let (5.10) be true. Suppose that the sequence $\{u_\Delta\}$ is not asymptotically optimal. With regard to Theorem 6.5, we see that there exist $\delta > 0$ and a sequence $\{\Delta_i\}$ for which $I_{\Delta_i}(u_{\Delta_i}) \geq \delta + \min_{\mu\in\mathfrak{S}} I(\mu)$ $\forall i$ and $\Delta_i \to 0$ as $i\to\infty$. Since the space $\mathfrak{P}(T) \supset \{u_{\Delta_i}\}$ is compact, passing to a subsequence ensures that $u_{\Delta_i} \to \nu$ as $i\to\infty$, where $\nu\in\mathfrak{S}$ by (6.2). In view of (5.10), dist$[\nu,\mathfrak{O}\mathfrak{P}] \leq \|\nu - u_{\Delta_i}\|_w + $ dist$[u_{\Delta_i},\mathfrak{O}\mathfrak{P}] \to 0$ as $i\to\infty$, where the set $\mathfrak{O}\mathfrak{P}$ is closed due to the last claim from Lemma 6.4. Hence $\nu\in\mathfrak{O}\mathfrak{P}$. Due to (6.3), we have

$$\delta + \min_{\mu\in\mathfrak{S}} I(\mu) \leq I_{\Delta_i}(u_{\Delta_i}) \to I(\nu) = \min_{\mu\in\mathfrak{S}} I(\mu) \quad \text{as} \quad i\to\infty.$$

The contradiction obtained proves that the sequence $\{u_\Delta\}$ is asymptotically optimal.

Conversely, consider such a sequence $\{u_\Delta\}$. Suppose that (5.10) is violated. Then there exist $\delta > 0$ and a sequence $\{\Delta_i\}$ such that dist$[u_{\Delta_i},\mathfrak{O}\mathfrak{P}] \geq \delta$ $\forall i$ and $\Delta_i \to 0$ as $i\to\infty$. As above, one can assume that $u_{\Delta_i} \to \nu \in \mathfrak{S}$ as $i\to\infty$. By (6.3), $\lim_{\Delta\to+0}\mathfrak{I}^m_{opt}(\Delta) = \lim_{i\to\infty} I_{\Delta_i}(u_{\Delta_i}) = I(\nu)$. So Theorem 6.5 yields $\nu\in\mathfrak{O}\mathfrak{P}$ in violation of the above relation dist$[u_{\Delta_i},\mathfrak{O}\mathfrak{P}] \geq \delta$ $\forall i$. Thus (5.10) does hold. □

7. General properties of the limit (6.1)**.** As follows from the previous section, proofs of the results from Section 5 are reduced to calculating the limits (6.1) and (6.2) for $\mathfrak{U}_\Delta := U_\Delta$. In doing so, certain general properties of the limit (6.1) are helpful. In this section, we reveal them. They follow from the single property introduced by the following definition.

DEFINITION 7.1. *A subset $\mathfrak{S} \subset \mathfrak{P}(T)$ is said to be* permutable *if for any $\nu_1, \nu_2 \in \mathfrak{S}$ and $\tau \in (t_0, t_1)$, it contains the element given by the following formula (where $E := [t_0, \tau)$)*

$$\nu_E(t, du_1, du_2) := \begin{cases} \nu_1(t, du_1, du_2) & \text{whenever } t \in E \\ \nu_2(t, du_1, du_2) & \text{otherwise.} \end{cases} \tag{7.1}$$

The next lemma shows that this property holds in the case $\mathfrak{U}_\Delta := U_\Delta$ of our primal interest.

LEMMA 7.2. *If $\mathfrak{U}_\Delta := U_\Delta \ \forall \Delta$, the set* (6.1) *is permutable.*

Proof. Let $\mathfrak{S}$ denote the set (6.1) and $\nu_1, \nu_2 \in \mathfrak{S}$, $\tau \in (t_0, t_1)$. For $i = 1, 2$, consider a sequence $\{u_\Delta^{(i)}\}$ associated with ν_i by (6.1). Setting $u_\Delta(t) := u_\Delta^{(1)}(t)$ for $t < \tau$ and $u_\Delta(t) := u_\Delta^{(2)}(t)$ for $t \geq \tau$ gives a control $u_\Delta \in U_\Delta$. For $\eta \in \mathbb{L}_1\mathbb{C}(T, \Omega \times \Omega)$, we put $\eta_1(t, w) := \eta(t, w), \eta_2(t, w) := 0$ if $t < \tau$ and $\eta_2(t, w) := \eta(t, w), \eta_1(t, w) := 0$ if $t \geq \tau$. Then (5.6) and (5.8) yield

$$\langle \eta; u_\Delta \rangle = \int_T \eta\left[t, u_\Delta(t), u_\Delta(t - \Delta)\right] dt = \int_{\tau+\Delta}^{t_1} \eta\left[t, u_\Delta^{(2)}(t), u_\Delta^{(2)}(t - \Delta)\right] dt$$
$$+ \underbrace{\int_\tau^{\tau+\Delta} \eta\left[t, u_\Delta^{(2)}(t), u_\Delta^{(1)}(t - \Delta)\right] dt}_{\alpha^{(1)}(\Delta)} + \int_{t_0}^{\tau} \eta\left[t, u_\Delta^{(1)}(t), u_\Delta^{(1)}(t - \Delta)\right] dt$$
$$= \int_\tau^{t_1} \eta\left[t, u_\Delta^{(2)}(t), u_\Delta^{(2)}(t - \Delta)\right] dt - \underbrace{\int_\tau^{\tau+\Delta} \eta\left[t, u_\Delta^{(2)}(t), u_\Delta^{(2)}(t - \Delta)\right] dt}_{\alpha^{(2)}(\Delta)} + \alpha^{(1)}(\Delta)$$
$$+ \int_{t_0}^{\tau} \eta\left[t, u_\Delta^{(1)}(t), u_\Delta^{(1)}(t - \Delta)\right] dt = \left\langle \eta_2; u_\Delta^{(2)} \right\rangle + \left\langle \eta_1; u_\Delta^{(1)} \right\rangle + \alpha^{(1)}(\Delta) - \alpha^{(2)}(\Delta).$$

In view of (5.5),

$$\left|\alpha^{(i)}(\Delta)\right| \leq \int_\tau^{\tau+\Delta} \max_{u_1, u_2 \in \Omega} \left|\eta\left[t, u_1, u_2\right]\right| dt \to 0 \quad \text{as} \quad \Delta \to 0.$$

So letting $\Delta \to 0$ gives $\langle u_\Delta; \eta \rangle \to \langle \nu_1; \eta_1 \rangle + \langle \nu_2; \eta_2 \rangle =: \chi$, where

$$\chi = \int_T dt \int_{\Omega \times \Omega} \eta_1(t, w)\nu_1(t, dw) + \int_T dt \int_{\Omega \times \Omega} \eta_2(t, w)\nu_2(t, dw) = \langle \nu_E; \eta \rangle$$

and $\eta \in \mathbb{L}_1\mathbb{C}(T, \Omega \times \Omega)$ is arbitrary. Thus $u_\Delta \to \nu_E$, and so $\nu_E \in \mathfrak{S}$ by (6.1). □

In the remainder of the section, we reveal consequences of permutability and closedness.

LEMMA 7.3. *Suppose that the set $\mathfrak{S} \subset \mathfrak{P}(T)$ is closed and permutable. Given $\nu_1, \nu_2 \in \mathfrak{S}$, the element* (7.1) *belongs to $\mathfrak{S}$ for any Borel set E.*

Proof. The claim is immediate from Definition 7.1 in the case where E is the union of a finite number of intervals. As is well known, any Borel set E can be approximated by such unions $E^{(1)}, E^{(2)}, \dots$ in the sense that $\text{mes}(E^{(i)} \triangle E) \to 0$ as $i \to \infty$. Here and throughout, the symbol **mes** stands for the Lebesgue measure and $A \triangle B := A \setminus B \cup B \setminus A$. For any function $\eta \in \mathbb{L}_1\mathbb{C}(T, \Omega \times \Omega)$, we have due to (7.1),

$$\langle \nu_{E^{(i)}}; \eta \rangle - \langle \nu_E; \eta \rangle = \int_T dt \int_{\Omega\times\Omega} \eta(t, u_1, u_2)\, [\nu_{E^{(i)}}(t, du_1, du_2) - \nu_E(t, du_1, du_2)]$$

$$= \underbrace{\int\limits_{E\setminus E^{(i)}} dt \int\limits_{\Omega\times\Omega} \eta(t, u_1, u_2)\nu_{21}(t, du_1, du_2)}_{\beta(1,i)} + \underbrace{\int\limits_{E^{(i)}\setminus E} dt \int\limits_{\Omega\times\Omega} \eta(t, u_1, u_2)\nu_{12}(t, du_1, du_2)}_{\beta(2,i)}.$$

Here $\nu_{sl} := \nu_s - \nu_l$ for $s, l = 1, 2$ and thanks to (5.5),

$$|\beta(j, i)| \leq \int_{E^{(i)}\triangle E} dt \int_{\Omega\times\Omega} |\eta(t, u_1, u_2)|\, [\nu_1(t, du_1, du_2) + \nu_2(t, du_1, du_2)]$$
$$\leq 2 \int_{E^{(i)}\triangle E} \max_{u_1,u_2\in\Omega} |\eta(t, u_1, u_2)|\, dt \to 0 \quad \text{as} \quad i \to \infty.$$

Hence $\langle \nu_{E^{(i)}}; \eta \rangle \to \langle \nu_E; \eta \rangle$ as $i \to \infty$ for all $\eta \in \mathbb{L}_1\mathbb{C}(T, \Omega \times \Omega)$ or in other words, $\nu_{E^{(i)}} \to \nu_E$ as $i \to \infty$. Thus $\nu_E \in \overline{\mathfrak{S}} = \mathfrak{S}$. □

Lemma 7.3 implies easily the following claim.

COROLLARY 7.4. *Suppose that the assumptions of Lemma 7.3 are true, $\nu_1, \ldots, \nu_s \in \mathfrak{S}$, and $E_1, \ldots, E_s$ are Borel sets constituting a partition of T. Put $\nu(t, dw) := \nu_i(t, dw)$ whenever $t \in E_i$ and $i = 1, \ldots, s$. Then $\nu \in \mathfrak{S}$.*

Though the next fact is well known, we offer its proof for the convenience of the reader.

LEMMA 7.5. *Any closed and permutable set $\mathfrak{S} \subset \mathfrak{P}(T)$ is convex.*

Proof. Let $\nu_1, \nu_2 \in \mathfrak{S}, \theta_1, \theta_2 > 0$, and $\theta_1 + \theta_2 = 1$. We put $\nu_{(r)}(t, dw) := \nu_i(t, dw)$ whenever $t \in T_i(j, r)$ for some $i = 1, 2$ and $j = 0, \ldots, 2^r - 1$, where $r = 1, 2, \ldots$,

$$T_1(j, r) := \left[t_0 + j2^{-r}|T|, t_0 + j2^{-r}|T| + \theta_1 2^{-r}|T|\right),$$
$$T_2(j, r) := \left[t_0 + j2^{-r}|T| + \theta_1 2^{-r}|T|, t_0 + (j + 1)2^{-r}|T|\right),$$

and $|T| := t_1 - t_0$. For any $\eta(\cdot) \in \mathbb{L}_1\mathbb{C}(T, \Omega \times \Omega)$, relation (5.6) yields

$$\langle \nu_{(r)}; \eta \rangle = \int_T dt \int_{\Omega\times\Omega} \eta(t, w)\nu_{(r)}(t, dw) = \mathcal{A}_r(\mathfrak{f}_1, \mathfrak{f}_2) := \sum_{j=0}^{2^r-1} \sum_{i=1,2} \int_{T_i(j,r)} \mathfrak{f}_i(t)\; dt,$$

where $\mathfrak{f}_i(t) := \int_{\Omega\times\Omega} \eta(t, w)\nu_i(t, dw)$. For any $\chi_1(\cdot), \chi_2(\cdot) \in \mathbb{L}_1(T)$,

$$|\mathcal{A}_r(\chi_1, \chi_2)| \leq \sum_{j=0}^{2^r-1} \sum_{i=1,2} \int_{T_i(j,r)} |\chi_i(t)|\; dt \leq \sum_{i=1,2} \int_T |\chi_i(t)|\; dt = |(\chi_1, \chi_2)|.$$

Denote by S_r the set of functions $\chi(\cdot) : T \to \mathbb{R}$ that are constant on any interval $T(j, r) := T_1(j, r) \cup T_2(j, r)$. For $\chi(\cdot) \in S_r$ and $r' \geq r$, it is easy to see that $\int_{T_i(j,r')} \chi(t)dt = \theta_i \int_{T(j,r')} \chi(t)dt,\ i = 1, 2$. Hence $\mathcal{A}_r(\chi_1, \chi_2) = \mathcal{A}(\chi_1, \chi_2) := \sum_{i=1,2} \theta_i \int_T \chi_i(t)dt$ for large $r \geq \overline{r}(\chi_1, \chi_2)$ whenever $\chi_1(\cdot), \chi_2(\cdot) \in S := \bigcup_{r=1}^{\infty} S_r$. The set S is dense in $\mathbb{L}_1(T)$. So given $\varepsilon > 0$, there exist $\chi_1(\cdot), \chi_2(\cdot) \in S$ such that $|\mathfrak{f}_i - \chi_i| < \varepsilon$ for $i = 1, 2$. Whence

$$\overline{\lim}_{r\to\infty} |(\mathcal{A}_r - \mathcal{A})(\mathfrak{f}_1, \mathfrak{f}_2)| = \overline{\lim}_{r\to\infty} |(\mathcal{A}_r - \mathcal{A})[(\mathfrak{f}_1, \mathfrak{f}_2) - (\chi_1, \chi_2)]|$$
$$\leq \overline{\lim}_{r\to\infty} |\mathcal{A}_r[(\mathfrak{f}_1, \mathfrak{f}_2) - (\chi_1, \chi_2)]| + |\mathcal{A}[(\mathfrak{f}_1, \mathfrak{f}_2) - (\chi_1, \chi_2)]|$$
$$\leq (1 + \|\mathcal{A}\|)\, |(\mathfrak{f}_1, \mathfrak{f}_2) - (\chi_1, \chi_2)| \leq 2\varepsilon\, (1 + \|\mathcal{A}\|) \quad \forall \varepsilon > 0.$$

Thus we see that $\lim_{r\to\infty} \langle \nu_{(r)}; \eta\rangle = \mathcal{A}(\mathfrak{f}_1, \mathfrak{f}_2) = \theta_1 \int_T \mathfrak{f}_1(t)dt + \theta_2 \int_T \mathfrak{f}_2(t)dt = \int_T dt \int_{\Omega\times\Omega} \eta(t,w)[\theta_1\nu_1(t,dw)+\theta_2\nu_2(t,dw)] = \langle \theta_1\nu_1 + \theta_2\nu_2; \eta\rangle$ $\forall\eta$, i.e., $\nu_{(r)} \to \theta_1\nu_1 + \theta_2\nu_2$ as $r\to\infty$. Since $\nu_{(r)} \in \mathfrak{S}$ $\forall r$ by Lemma 7.3, we get $\theta_1\nu_1 + \theta_2\nu_2 \in \overline{\mathfrak{S}} = \mathfrak{S}$. □

Now we are in a position to prove the main result of the section.

THEOREM 7.6. *Let a set $\mathfrak{S} \subset \mathfrak{P}\,(T)$ be closed and permutable. Whenever*

$$\mu_j \in \mathfrak{S}, \alpha_j(\cdot) \in \mathbb{L}_\infty(T)\ j = 1,\dots,l, \text{ and } \alpha_1(t) \geq 0,\dots,\alpha_l(t) \geq 0,\ \sum_{j=1}^{l} \alpha_j(t) = 1 \tag{7.2}$$

for almost all $t \in T$, the set $\mathfrak{S}$ contains the following element

$$\nu(t, du_1, du_2) := \sum_{j=1}^{l} \alpha_j(t)\mu_j(t, du_1, du_2).$$

Proof. The functions $\alpha_j(\cdot)$ can be uniformly approximated by step functions, i.e., measurable functions that take a finite number of values. Moreover, these step functions can be chosen to satisfy (7.2). Since $\overline{\mathfrak{S}} = \mathfrak{S}$, this implies that it suffices to prove Lemma 7.6 in the case where $\alpha_1(t),\dots,\alpha_l(t)$ are step functions. In such a case, there exists a Borel partition $T = E_1 \cup \dots \cup E_s$ of T such that $\alpha_j(t) = \alpha_j^{(i)}$ a.a. $t \in E_i$ and all $j = 1,\dots,l, i = 1,\dots,s$. By Lemma 7.5, $\nu_i := \sum_{j=1}^{l} \alpha_j^{(i)}\mu_j \in \mathfrak{S}$ for $i = 1,\dots,s$. It remains o note that $\nu(t,dw) = \nu_i(t,dw)$ whenever $t \in E_i$, and apply Corollary 7.4. □

Lemmas 6.2, 7.2 and Theorem 7.6 give rise to the following claim.

COROLLARY 7.7. *The conclusion of Theorem* 7.6 *is true for the limit* (6.1) *of the* Δ-*parametric set* $\mathfrak{U}_\Delta := U_\Delta$.

8. Extremal points of a polygon in the space of matrices. Our next objective is to show that the limits (6.1), (6.2) equal $\mathfrak{CM}$ if $\mathfrak{U}_\Delta = U_\Delta$ $\forall\Delta$. We recall that the symbol $\mathfrak{CM}$ stands for the collection of all measurable maps $\nu : T \to \mathfrak{SPM}(\Omega\times\Omega)$. Here $\mathfrak{SPM}(\Omega\times\Omega)$ is the set of all probability measures satisfying (5.2). To prove that any element ν of $\mathfrak{CM}$ belongs to the sets (6.1), (6.2), it is convenient to start with the elements of the form

$$\nu(t, du', du'') = \sum_{i,j=1}^{l} p_{ij}(t)\delta_{u_i}(du')\delta_{u_j}(du''), \tag{8.1}$$

where $u_1,\dots,u_l \in \Omega$, $u_i \neq u_j$ whenever $i \neq j$, and $\delta_v(du)$ is the Dirac measure at the point $v \in \Omega$. It is easy to see that $\nu \in \mathfrak{CM}$ if and only if for almost all t, the matrix

$$P(t) = (p_{ij}(t))_{i,j=1}^{l} \tag{8.2}$$

belongs to the set $\mathcal{M}_l$ of all $l \times l$-matrices $P = (p_{ij})_{i,j=1}^{l}$ such that

$$p_{i,j} \geq 0 \quad \forall i,j, \qquad \sum_{i,j=1}^{l} p_{i,j} = 1, \qquad \sum_{j=1}^{l} p_{i,j} = \sum_{j=1}^{l} p_{j,i} \quad \forall\ i = 1,\dots l. \tag{8.3}$$

The set $\mathcal{M}_l$ is clearly a polygon in the space of $l \times l$ matrices. So it is the convex hull of its extremal points. By Corollary 7.7, this means that any element $\nu \in \mathfrak{CM}$ of the form (8.1) belongs to the sets (6.1), (6.2) if so do elements (8.1) with the matrix $P(t)$ constant and from the set $\mathcal{M}_l^{\text{ext}}$ of the extremal points of $\mathcal{M}_l$. In this section, we calculate $\mathcal{M}_l^{\text{ext}}$.

LEMMA 8.1. *Suppose that* $\widetilde{P} = (\widetilde{p}_{ij}) \in \mathcal{M}_l^{ext}$, $P = (p_{ij}) \in \mathcal{M}_l$, *and* $\widetilde{p}_{ij} \geq \theta p_{ij}$ $\forall i, j$, *where* $\theta > 0$. *Then* $\widetilde{P} = P$.

Proof. Reducing $\theta > 0$ yields that $\widetilde{p}_{i'j'} > \theta p_{i'j'}$ for some i', j'. By (8.3), $1 = \sum_{i,j=1}^{l} \widetilde{p}_{ij} > \theta \sum_{i,j=1}^{l} p_{ij} = \theta$, i.e., $0 < \theta < 1$. Put $p'_{ij} := (1-\theta)^{-1}(\widetilde{p}_{ij} - \theta p_{ij})$. It is easy to see that $P' := (p'_{ij})_{i,j=1}^{l} \in \mathcal{M}_l$ and $\widetilde{P} = \theta P + (1-\theta)P'$. This implies $\widetilde{P} = P$ by the definition of the extremal point. □

DEFINITION 8.2. *A* $l \times l$*-matrix* P *is said to be* cyclic *if there exists a sequence of integers* $1 \leq i_1, i_2, \dots, i_s \leq l$ *such that* $i_h \neq i_d$ *for* $h \neq d$ *and*

$$\begin{aligned} p_{hd} = 0 \quad \textit{whenever} \quad (h,d) \neq (i_\sigma, i_{\sigma-1}) \quad \forall \sigma = 1, \dots, s \quad (i_0 := i_s), \\ p_{i_\sigma, i_{\sigma-1}} = s^{-1} \quad \forall \sigma = 1, \dots, s. \end{aligned} \tag{8.4}$$

Note that any such matrix belongs to $\mathcal{M}_l$.

LEMMA 8.3. *Any extremal point of* $\mathcal{M}_l$ *is a cyclic matrix.*

Proof. Let $\widetilde{P} = (\widetilde{p}_{ij}) \in \mathcal{M}_l^{\text{ext}}$. If $\widetilde{p}_{ii} > 0$ for some i, then Lemma 8.1 (where $\theta := \widetilde{p}_{ii}, p_{hd} := 0$ whenever $h \neq d$ or $h = d \neq i$, and $p_{ii} := 1$) ensures that $\widetilde{P} = P := (p_{hd})_{h,d=1}^{l}$, where the matrix P is cyclic (with $s := 1$ and $i_1 := i$ in (8.4)). Suppose that $\widetilde{p}_{11} = \dots = \widetilde{p}_{ll} = 0$. By (8.3), $\sum_{j=1}^{l} \widetilde{p}_{ij} > 0$ for some $i = d_1$. Hence $\widetilde{p}_{d_1 d_2} > 0$ for some index $d_2 \neq d_1$. Thanks to (8.3), $\sum_{j=1}^{l} \widetilde{p}_{d_2,j} = \sum_{j=1}^{l} \widetilde{p}_{j,d_2} > 0$ and so $\widetilde{p}_{d_2 d_3} > 0$ for some $d_3 = 1, \dots, l$. Apparently, $d_3 \neq d_2$. By continuing likewise, we get a sequence $d_1, d_2, d_3, \dots$ such that $d_{j+1} \neq d_j$ and $\widetilde{p}_{d_j d_{j+1}} > 0 \forall j$. We terminate it when some index is repeated, i.e., $d_i \neq d_j$ whenever $i < j \leq q$ and $d_{q+1} = d_r$ for some $r = 1, \dots, q$. Here $r \leq q-1$ since $d_q \neq d_{q+1}$. Put $s := q - r + 1 \geq 2$ and $i_j := d_{s+r-j}$ for $j = 1, \dots, s$. Then $i_j \neq i_{j'}$ whenever $j < j' \leq s$ and $\widetilde{p}_{i_j i_{j-1}} > 0$ for $j = 1, \dots, s$, where $i_0 := d_{s+r} = d_{q+1} = d_r = i_s$. Define the cyclic matrix $P = (p_{ij})_{i,j=1}^{l}$ in correspondence with (8.4) and denote $\theta := s \min_{j=1,\dots,s} \widetilde{p}_{i_j, i_{j-1}}$. Then $\widetilde{p}_{ij} \geq \theta p_{ij}$ $\forall i, j$, and Lemma 8.1 completes the proof. □

Now we are in a position to state the main result of the section.

THEOREM 8.4. *The extremal points of* $\mathcal{M}_l$ *constitute the set of all cyclic* $l \times l$*-matrices.*

Proof. Thanks to Lemma 8.3, it suffices to show that any such a matrix $P = (p_{ij})$ is an extremal point. Suppose that $P = \theta_1 P^{\langle 1 \rangle} + \theta_2 P^{\langle 2 \rangle}$, where $\theta_h > 0$, $P^{\langle h \rangle} = (p_{ij}^{\langle h \rangle}) \in \mathcal{M}_l$, and $\theta_1 + \theta_2 = 1$. Consider the sequence $i_1, \dots, i_s$ associated with P by Definition 8.2. Since $p_{ij} = \theta_1 p_{ij}^{\langle 1 \rangle} + \theta_2 p_{ij}^{\langle 2 \rangle}$, we have $p_{ij} = 0 \Rightarrow p_{ij}^{\langle h \rangle} = 0$ $h = 1, 2$. Hence (8.4) yields

$$p_{ij}^{\langle h \rangle} = 0 \quad \text{whenever} \quad (i,j) \neq (i_{\sigma-1}, i_\sigma) \quad \forall \sigma = 1, \dots, s. \tag{8.5}$$

So $s = 1 \Rightarrow P^{\langle 1 \rangle} = P^{\langle 2 \rangle} = P$, i.e., the point P is extremal. Now suppose that $s \geq 2$. Then $p_{ii}^{\langle h \rangle} = 0 \forall h = 1, 2, i = 1, \dots, l$ by (8.5). Whence $p_{ij}^{\langle h \rangle} = 0$ $\forall i \neq j \Rightarrow p_{ij}^{\langle h \rangle} = 0$ $\forall i, j$ in violation of (8.3). Thus $p_{ij}^{\langle h \rangle} > 0$ for some $i \neq j$. Then retracing the arguments from the proof of Lemma 8.3 shows that $p_{\widehat{i}_{\sigma-1} \widehat{i}_\sigma}^{\langle h \rangle} > 0$ $\forall \sigma = 1, \dots, \widehat{s} \geq 2$ for a sequence $\{\widehat{i}_\sigma\}_{\sigma=1}^{\widehat{s}}, \widehat{i}_{\sigma+1} \neq \widehat{i}_\sigma, \widehat{i}_0 := \widehat{i}_{\widehat{s}}$. Then (8.5) implies that $\widehat{s} = s$ and this sequence equals $\{i_\sigma\}$ up to a cyclic shift of the index. Thus $p_{i_{\sigma-1} i_\sigma}^{\langle h \rangle} > 0$ $\forall \sigma$. The last relation from (8.3) and (8.5) yield that $\zeta^{\langle h \rangle} := p_{i_{\sigma-1} i_\sigma}^{\langle h \rangle}$ does not depend on σ. Then the second relation from (8.3) gives $s\zeta^{\langle h \rangle} = 1$. Thus $P^{\langle h \rangle} = P$ for $h = 1, 2$ in view of (8.4), which completes the proof. □

9. Calculation of the limits (6.1), (6.2) **in the case where** $\mathfrak{U}_\Delta = U_\Delta$ $\forall \Delta$**.** From now on, we consider the Δ-parametric set U_Δ. In this section, we show that

$$\mathfrak{S} = \mathfrak{S}^{\text{seq}} = \mathfrak{CM}, \quad \text{where} \quad \mathfrak{S} := \text{LIM}_{\Delta \to 0} U_\Delta \quad \text{and} \quad \mathfrak{S}^{\text{seq}} := \text{LIM}_{\Delta \to 0}^{\text{seq}} U_\Delta. \tag{9.1}$$

LEMMA 9.1. *The following inclusions hold:* $\mathfrak{S} \subset \mathfrak{S}^{seq} \subset \mathfrak{CM}$.

Proof. The first inclusion is immediate from (6.1) and (6.2). To prove the second one, consider an element $\nu \in \mathfrak{S}^{\mathrm{seq}}$ and the corresponding sequences $\{\Delta_i\}$ and $\{u_i\}$ from (6.2). Given two functions $\zeta(\cdot) \in \mathbb{L}_1(T)$ and $\rho(\cdot) \in \mathbb{C}(\Omega)$, we put $\eta_i(t, u_1, u_2) := \zeta(t)\rho(u_i)$ $(i = 1, 2)$ for all $t \in T$ and $u_1, u_2 \in \Omega$. By (5.6) and (5.8),

$$\begin{aligned}\langle u_i; \eta_2\rangle - \langle u_i; \eta_1\rangle &= \int_T \rho\left[u_i(t-\Delta_i)\right]\zeta(t)\,dt - \int_T \rho\left[u_i(t)\right]\zeta(t)\,dt \\ &= \int_{t_0-\Delta_i}^{t_1-\Delta_i} \rho\left[u_i(t)\right]\zeta(t+\Delta_i)\,dt - \int_{t_0}^{t_1} \rho\left[u_i(t)\right]\zeta(t)\,dt.\end{aligned}$$

Letting $i \to \infty$ gives

$$0 = \langle \nu; \eta_2\rangle - \langle \nu; \eta_1\rangle = \int_T \zeta(t)\,dt \int_{\Omega\times\Omega} \left[\rho(u_2) - \rho(u_1)\right]\nu(t, du_1, du_2).$$

So far as the function $\zeta(\cdot) \in \mathbb{L}_1(T)$ is arbitrary, we have for almost all $t \in T$

$$0 = \int_{\Omega\times\Omega} \left[\rho(u_2) - \rho(u_1)\right]\nu(t, du_1, du_2) = \int_\Omega \rho(u)\nu(t, du, \Omega) - \int_\Omega \rho(u)\nu(t, \Omega, du)$$

for all $\rho(\cdot) \in \mathbb{C}(\Omega)$. So $\nu(t, du, \Omega) = \nu(t, \Omega, du)$ and thus $\nu \in \mathfrak{CM}$. $\Box$

Thus to justify (9.1), it suffices to show that any element $\nu \in \mathfrak{CM}$ belongs to $\mathfrak{S} = \mathrm{LIM}_{\Delta\to+0} U_\Delta$. We start with the elements of the form (8.1).

LEMMA 9.2. *Suppose that the matrix* (8.2) *is cyclic and does not depend on t, i.e.,* $P(t) = P = (p_{ij})$ *a.a.* $t \in T$. *Then the element* (8.1) *belongs to* $\mathfrak{S} = \mathrm{LIM}_{\Delta\to+0} U_\Delta$.

Proof. Consider the sequence $1 \leq i_1, \ldots, i_s \leq l$ associated with P by (8.4) and denote

$$\begin{gathered} t_\Delta(r,j) := t_0 - \Delta + sr\Delta + j\Delta, \quad T_\Delta(r,j) := \Big[t_\Delta(r, j-1)\,;t_\Delta(r,j)\Big) \cap [t_0 - \Delta, t_1], \\ u_\Delta(t) := u_{i_j} \quad \text{whenever} \quad t \in T_\Delta(r,j) \quad \text{for some} \quad r = 0, 1, \ldots,\ j = 1, \ldots, s. \end{gathered}$$

Then $u_\Delta \in U_\Delta$. So due to (6.1), it suffices to show that $u_\Delta \to \nu$ as $\Delta \to +0$, i.e., $\langle u_\Delta; \eta\rangle \to \langle \nu; \eta\rangle$ as $\Delta \to +0$ for any $\eta \in \mathbb{L}_1\mathbb{C}(T, \Omega\times\Omega)$. By [23, Lemma IV.2.4], one can focus on the functions η of the form $\eta(t, u_1, u_2) = \chi_E(t)\rho(u_1, u_2)$, where $\rho(\cdot) \in \mathbb{C}(\Omega\times\Omega), E = [\tau_1, \tau_2], t_0 < \tau_1 < \tau_2 < t_1$, and $\chi_E(t) := 1$ for $t \in E$ whereas $\chi_E(t) := 0$ if $t \notin E$. Put

$$A_\Delta := \Big\{\alpha = 0, 1, \cdots : T_\Delta(\alpha) := \Big[t_\Delta(0,0) + s\alpha\Delta ; t_\Delta(0,0) + s(\alpha+1)\Delta\Big] \subset [\tau_1, \tau_2]\Big\},$$

$\alpha_\Delta^- := \min\{\alpha : \alpha \in A_\Delta\}$, $\alpha_\Delta^+ := \max\{\alpha : \alpha \in A_\Delta\}$. Then by (5.6) and (5.8),

$$\begin{aligned}\langle u_\Delta; \eta\rangle &= \int_T \chi_E(t) \underbrace{\rho\left[u_\Delta(t), u_\Delta(t-\Delta)\right]}_{\varkappa(t)} dt = \int_{\tau_1}^{\tau_2} \varkappa(t)\,dt \\ &= \underbrace{\int_{T_\Delta(\alpha_\Delta^- - 1)\cap[\tau_1,\tau_2]} \varkappa(t)\,dt}_{\omega^-(\Delta)} + \underbrace{\int_{T_\Delta(\alpha_\Delta^+ + 1)\cap[\tau_1,\tau_2]} \varkappa(t)\,dt}_{\omega^+(\Delta)} + \underbrace{\sum_{r=\alpha_\Delta^-}^{\alpha_\Delta^+} \sum_{j=1}^{s} \int_{T_\Delta(r,j)} \varkappa(t)\,dt}_{\omega(\Delta)}.\end{aligned}$$

Here $\mathrm{mes}T_\Delta(\alpha^\pm_\Delta \pm 1) = s\Delta$ and so $|\omega^\pm(\Delta)| \le c_\rho s\Delta$, where $c_\rho := \max\limits_{u',u''\in\Omega} |\rho(u',u'')|$. Thus $\omega^\pm(\Delta) \to 0$ as $\Delta \to +0$. For the indices r, j involved in the last sum and $\Delta \approx 0$,

$$\int_{T_\Delta(r,j)} \varkappa(t)dt = \int_{t_\Delta(r,j-1)}^{t_\Delta(r,j-1)+\Delta} \rho\big(u_{i_j}, u_{i_{j-1}}\big)dt = \Delta\rho\big(u_{i_j}, u_{i_{j-1}}\big).$$

Note that $\Delta = \mathrm{mes}T_\Delta(r,j)$. Hence

$$\omega(\Delta) = \frac{1}{s}\sum_{j=1}^{s} \rho\big(u_{i_j}, u_{i_{j-1}}\big) \sum_{r=\alpha^-_\Delta}^{\alpha^+_\Delta} \sum_{j=1}^{s} \mathrm{mes}T_\Delta(r,j).$$

The last double sum equals $\mathrm{mes}\mathcal{T}_\Delta$, where $\mathcal{T}_\Delta := \bigcup_{r=\alpha^-_\Delta}^{\alpha^+_\Delta} \bigcup_{j=1}^{s} T_\Delta(r,j) \subset [\tau_1,\tau_2]$, $[\tau_1,\tau_2]\setminus \mathcal{T}_\Delta \subset T_\Delta(\alpha^-_\Delta - 1) \cup T_\Delta(\alpha^+_\Delta + 1)$, and $\mathrm{mes}\big[T_\Delta(\alpha^-_\Delta - 1) \cup T_\Delta(\alpha^+_\Delta + 1)\big] \to 0$ as $\Delta \to +0$ by the foregoing. It follows that

$$\langle u_\Delta; \eta\rangle \to \Gamma := s^{-1}(\tau_2 - \tau_1)\sum_{j=1}^{s} \rho\big(u_{i_j}, u_{i_{j-1}}\big).$$

At the same time, (5.6), (8.1), and (8.4) yield $\Gamma = \int_{\tau_1}^{\tau_2} dt \sum_{i,j=1}^{l} p_{ij}\rho\,(u_i, u_j) = \langle \nu; \eta\rangle$, which completes the proof. $\square$

LEMMA 9.3. *The set* $\mathop{\mathrm{LIM}}\limits_{\Delta\to+0} U_\Delta$ *contains any element* (8.1) *with* $P(t) \in \mathcal{M}_l$ *a.a.* $t \in T$.

Proof. Put $\mathfrak{A} := \{\alpha : \mathcal{M}_l^{\mathrm{ext}} \to [0,\infty) : \sum_{P\in\mathcal{M}_l^{\mathrm{ext}}} \alpha(P) = 1\}$ and $\Lambda\big[\alpha(\cdot)\big] := \sum_{P\in\mathcal{M}_l^{\mathrm{ext}}} \alpha(P)P$, where $\mathcal{M}_l^{\mathrm{ext}}$ is the set of the extremal points of $\mathcal{M}_l$. By the Krein-Milman theorem, $\Lambda(\mathfrak{A}) \supset \mathcal{M}_l$. Then Theorems I.7.6 and I.7.7 [23] imply that $P(t) = \Lambda\big[\alpha(t,\cdot)\big]$ for a measurable function $t \mapsto \alpha(t,\cdot) \in \mathfrak{A}$. Lemma 9.2 and Corollary 7.7 complete the proof. $\square$

Now we are in a position to state the main result of the section.

THEOREM 9.4. *The relationships* (9.1) *are true.*

Proof. By Lemma 9.1, it suffices to show that $\mathfrak{CM} \subset \mathfrak{S}$. Let $\nu \in \mathfrak{CM}$. Given $\delta > 0$, we introduce a Borel partition $\Omega = \Omega^\delta(1) \cup \cdots \cup \Omega^\delta(l_\delta)$ such that $\sup_{u\in\Omega^\delta(i)} |u - u_i^\delta| < \delta$ for some $u_i^\delta \in \Omega^\delta(i)$ and all $i = 1,\ldots,l_\delta$. It follows from (5.2) that (8.3) is valid for $l := l_\delta$ and

$$p_{ij} := p_{ij}^\delta(t) := \nu\left[t, \Omega^\delta(i), \Omega^\delta(j)\right] \qquad i,j = 1,\ldots,l_\delta.$$

Thus $P(t) := \big(p_{ij}^\delta(t)\big)_{i,j=1}^{l_\delta} \in \mathcal{M}_{l_\delta}$ a.a. t, where the set $\mathcal{M}_{l_\delta}$ consists of all matrices satisfying (8.3). So Lemma 9.3 ensures that the element ν_δ defined by (8.1) (where $l := l_\delta, p_{ij} := p_{ij}^\delta(t), u_i := u_i^\delta$) belongs to $\mathfrak{S}$. For any $\rho \in \mathbb{C}(\Omega\times\Omega)$ and a.a. $t \in T$,

$$\begin{aligned}
&\left|\int_{\Omega\times\Omega} \rho(u',u'')\nu_\delta(t,du',du'') - \int_{\Omega\times\Omega} \rho(u',u'')\nu(t,du',du'')\right| \\
&\quad = \left|\sum_{i,j=1}^{l_\delta} \rho\left(u_i^\delta, u_j^\delta\right) p_{ij}^\delta(t) - \sum_{i,j=1}^{l_\delta} \int_{\Omega^\delta(i)\times\Omega^\delta(j)} \rho(u',u'')\nu(t,du',du'')\right| \\
&\quad \le \sum_{i,j=1}^{l_\delta} \int_{\Omega^\delta(i)\times\Omega^\delta(j)} \left|\rho\left(u_i^\delta, u_j^\delta\right) - \rho(u',u'')\right| \nu(t,du',du'') \le q_\delta[\rho(\cdot)]. \quad (9.2)
\end{aligned}$$

Here $q_\delta[\rho(\cdot)] := \max|\rho(v_1, v_2) - \rho(u_1, u_2)|$, where max is over $(v_1, v_2), (u_1, u_2) \in \Omega \times \Omega$ such that $|v_1 - u_1| \leq \delta, |v_2 - u_2| \leq \delta$. Pick $\eta \in \mathbb{C}(T)$. Then

$$\left| \int_T \eta(t)\, dt \int_{\Omega\times\Omega} \rho(u', u'')\nu_\delta(t, du', du'') - \int_T \eta(t)\, dt \int_{\Omega\times\Omega} \rho(u', u'')\nu(t, du', du'') \right| \\ \leq q_\delta[\rho(\cdot)] \int_T |\eta(t)|\, dt.$$

Let $\delta \to +0$. Since $q_\delta[\rho(\cdot)] \to 0$, then $\nu_\delta \to \nu$ by [23, Theorem IV.2.4]. As was shown, $\nu_\delta \in \mathfrak{S}$. Then Lemma 6.2 yields $\nu \in \mathfrak{S}, \mathfrak{CM} \subset \mathfrak{S}$, which completes the proof of (9.1). □

10. Proofs of the statements from Section 5. Theorem 5.1 is immediate from Theorems 6.5 and 9.4. Lemma 5.3 follows from Theorems 6.6 and 9.4. Theorem 5.4 results from Theorems 6.7, 9.4 and the first relation from (6.3). Theorem 5.5 is straightforward from Theorems 6.7 and 9.4.

11. Proofs of the statements from Section 3.

Proof of Lemma 3.1. This lemma is immediate from Theorem 5.1. □

Proof of Theorem 3.2. Put in (2.1)–(2.4) $f(t, x_1, x_2, u_1, u_2) := a(t) := 0, \varphi(t, x_1, x_2, u_1, u_2) := -|u_1 - u_2|^2$, where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^m$. Then due to (2.7), $\overline{\varphi}(t, x, u) = 0$ and so $\mathfrak{I}_{opt}(0) = 0$, where $\mathfrak{I}_{opt}(0)$ is the infimum of the cost functional in the problem (2.5), (2.6). By (i) from Section 3, the set Ω contains two points $u_1 \neq u_2$. Put

$$\nu(t, dw) := \eta(dw) := \frac{1}{2(t_1 - t_0)} \Big[\delta_{(u_1,u_2)}(dw) + \delta_{(u_2,u_1)}(dw)\Big], \tag{11.1}$$

where $w = (u', u'') \in \Omega \times \Omega$ and $\delta_\omega(dw)$ is the Dirac measure at the point $\omega \in \Omega \times \Omega$. Then η satisfies the relations from (5.2) and hence $\nu \in \mathfrak{CM}$. As a result, Theorem 5.1 yields

$$\lim_{\Delta\to+0} \mathfrak{I}_{opt}(\Delta) \leq -\int_T dt \int_{\Omega\times\Omega} |u' - u''|^2 \nu(t, du', du'') = -|u_1 - u_2|^2 < 0 = \mathfrak{I}_{opt}(0).$$

The remainder of the section is devoted to the proof of Theorem 3.3. Further we employ the notations introduced before and in this theorem. By (iv) from Section 3, the Cauchy problem (2.2), (2.3) has a unique solution $x(\cdot) = x_\Delta[\cdot|u(\cdot), \xi]$ for any $\xi = [f(\cdot), \varphi(\cdot), a(\cdot)] \in \Xi_{(iv)}$, $\Delta \in [0, 1)$, and $u(\cdot) \in U := \{u(\cdot) : [t_0 - 1, t_1] \to \Omega, u(\cdot) \text{ is measurable }\}$.

LEMMA 11.1. *For any $\xi^\# \in \Xi_{(iv)}$, the following convergence holds*

$$\varkappa[\xi, \xi^\#] := \sup_{\substack{\Delta\in[0,1)\\ u(\cdot)\in U}} \max_{t\in T} \left|x_\Delta[t|u(\cdot), \xi] - x_\Delta[t|u(\cdot), \xi^\#]\right| \to 0 \; as \left\{ \begin{array}{l} \xi \to \xi^\# \\ \xi \in \Xi_{(iv)} \end{array} \right. . \tag{11.2}$$

Proof. The proof is based on the implicit function theorem [23, Th.II.3.8]. To employ it, put $\Upsilon := [0, 1) \times U$, consider the Banach space $\mathfrak{X}$ of all bounded maps $\upsilon = [\Delta, u(\cdot)] \in \Upsilon \xrightarrow{\mathfrak{x}} \mathfrak{x}_\upsilon(\cdot) \in \mathbb{C}(T, \mathbb{R}^n)$ equipped by the norm $|\mathfrak{x}| := \sup_{\upsilon\in\Upsilon} \max_{t\in T} |\mathfrak{x}_\upsilon(t)|$, and introduce the map $F : \mathfrak{X} \times \Xi_{(iv)} \to \mathfrak{X}$ by putting $F[\mathfrak{x}, \xi] := \widetilde{\mathfrak{x}}$, where $\xi = [f(\cdot), \varphi(\cdot), a(\cdot)]$ and

$$\widetilde{\mathfrak{x}}_{[\Delta,u(\cdot)]}(t) := a(t_0) + \\ \int_{t_0}^t f\left[\theta, \mathfrak{x}_{[\Delta,u(\cdot)]}(\theta), \left\{ \begin{array}{ll} \mathfrak{x}_{[\Delta,u(\cdot)]}(\theta - \Delta) & \text{if } \theta \geq t_0 + \Delta \\ a(\theta - \Delta) & \text{if } \theta < t_0 + \Delta \end{array} \right\}, u(\theta), u(\theta - \Delta)\right] d\theta.$$

By (i)—(iv) from Section 3, the element $\mathfrak{x}_{[\Delta,u(\cdot)]}(t) := x_\Delta[t|u(\cdot),\xi]$ belongs to $\mathcal{X}$ and is the unique solution for the equation $F(\mathfrak{x},\xi)=\mathfrak{x}$. Theorem II.3.8 [23] completes the proof. ☐

LEMMA 11.2. *The infima* $\mathcal{J}_{opt}^{\Delta}(\xi), \mathcal{J}_{opt}^{\lim}(\xi), \mathcal{J}_{opt}(\xi)$ *of the cost functionals in the problems* (2.1)—(2.4)*;* (5.3), (5.4)*; and* (2.5), (2.6)*, respectively, depend on* $\xi \in \Xi_{(iv)}$ *continuously.*

Proof. Denote by $I_\Delta[u(\cdot)|\xi]$ the value of the functional from (2.1) on the process $\big(x_\Delta[\cdot|u(\cdot),\xi],u[\cdot]\big)$. Let $\xi \to \xi^{\#} = [f^{\#}(\cdot),\varphi^{\#}(\cdot),a^{\#}(\cdot)] \in \Xi_{(iv)}, \xi \in \Xi_{(iv)}$. Then

$$\begin{aligned}\big|I_\Delta[u(\cdot)|\xi]-I_\Delta[u(\cdot)|\xi^{\#}]\big| \leq & \int_T \Big|\varphi^{\#}\left\{t,x_\Delta[t|u(\cdot),\xi],x_\Delta[t-\Delta|u(\cdot),\xi],u(t),u(t-\Delta)\right\} \\ & -\varphi^{\#}\left\{t,x_\Delta[t|u(\cdot),\xi^{\#}],x_\Delta[t-\Delta|u(\cdot),\xi^{\#}],u(t),u(t-\Delta)\right\}\Big|\,dt \\ & +\int_T \Big|\Big(\varphi-\varphi^{\#}\Big)\left\{t,x_\Delta[t|u(\cdot),\xi],x_\Delta[t-\Delta|u(\cdot),\xi],u(t),u(t-\Delta)\right\}\Big|\,dt. \quad (11.3)\end{aligned}$$

Due to (iv) from Section 3, $\big|x_\Delta[t|u(\cdot),\xi^{\#}]\big| \leq k$ for all $\Delta \in [0,1), u(\cdot) \in U$, and $t \in [t_0-\Delta,t_1]$. So the first integral on the right does not exceed

$$(t_1-t_0)\max\Big|\varphi^{\#}\big[t,y',y'',w\big]-\varphi^{\#}\big[t,x',x'',w\big]\Big|, \quad (11.4)$$

where $\max$ is over $t \in T, w \in \Omega\times\Omega, |x'|,|x''| \leq k, |y'-x'|,|y''-x''| \leq \varkappa[\xi,\xi^{\#}]$, and $\varkappa[\xi,\xi^{\#}]$ is given by (11.2). So owing to (ii) from Section 3 and (11.2), this integral converges to 0 uniformly over $u(\cdot) \in U$ and $\Delta \in [0,1)$. By (11.2), one can assume that $\big|x_\Delta[t|u(\cdot),\xi]\big| \leq c := k+1$ for all $u(\cdot) \in U, \Delta \in [0,1)$, and $t \in [t_0-\Delta,t_1]$. Then the second integral in (11.3) does not exceed $(t_1-t_0)\mathfrak{H}_c(\xi-\xi^{\#}) \to 0$, where $\mathfrak{H}_c(\cdot)$ is given by (3.2). It follows that

$$\sup_{\Delta\in[0,1)}\sup_{u(\cdot)\in U}\big|I_\Delta[u(\cdot)|\xi]-I_\Delta[u(\cdot)|\xi^{\#}]\big| \to 0. \quad (11.5)$$

Since $\mathcal{J}_{opt}^{\Delta}(\xi)=\inf_{u(\cdot)\in U} I_\Delta[u(\cdot)|\xi]$, we have

$$\sup_{\Delta\in[0,1)}\big|\mathcal{J}_{opt}^{\Delta}(\xi)-\mathcal{J}_{opt}^{\Delta}(\xi^{\#})\big| \to 0. \quad (11.6)$$

Whence for any $\Delta \in [0,1)$, the function $\xi \in \Xi_{(iv)} \mapsto \mathcal{J}_{opt}^{\Delta}(\xi)\big(=\mathcal{J}_{opt}(\xi) \text{ for } \Delta=0\big)$ is continuous. So is the map $\xi \mapsto \mathcal{J}_{opt}^{\lim}(\xi)$ thanks to (11.6) and Theorem 5.1. ☐

Proof of Theorem 3.3. By Theorem 3.2, $\Xi_{(3.1)} \neq \emptyset$. The set $\Xi_{(3.1)}$ is open due to Lemma 11.2 since $\Xi_{(3.1)} = \{\xi : \mathcal{J}_{opt}^{\lim}(\xi) \neq \mathcal{J}_{opt}(\xi)\}$ thanks to Theorem 5.1. Now suppose that the domain Ω is connected and show that the set $\Xi_{(3.1)}$ is dense in $\Xi_{(iv)}$, i.e., $\xi = [f(\cdot),\varphi(\cdot),a(\cdot)] \in \Xi_{(iv)}\setminus\Xi_{(3.1)} \Rightarrow \xi \in \overline{\Xi_{(3.1)}}$. Note that $\xi \notin \Xi_{(3.1)} \Rightarrow \mathcal{J}_{opt}(\xi)=\mathcal{J}_{opt}^{\lim}(\xi)$. As is well known, $\mathcal{J}_{opt}(\xi)$ amounts to the infimum of the cost functional in the problem (2.5), (2.6) over the set $S=\{u(\cdot)\}$ of step controls, i.e., $\mathcal{J}_{opt}(\xi)=\inf_{u(\cdot)\in S} I_{(2.5)}[u(\cdot)|\xi]$, where $I_{(2.5)}[u(\cdot)|\xi]$ is the functional from (2.5). So for any $\varepsilon>0$, there exists $u(\cdot)\in S$ such that

$$I_{(2.5)}[u(\cdot)|\xi] < \mathcal{J}_{opt}(\xi)+\varepsilon. \quad (11.7)$$

Here $u(\cdot)\in S \Rightarrow u(t)=u^{(i)}$ a.a. $t\in E^{(i)}$, where $E^{(1)},\dots,E^{(N)}$ is a Borel partition of T and $u^{(1)},\dots,u^{(N)} \in \Omega, u^{(p)} \neq u^{(q)}$ whenever $p\neq q$. So far as the set Ω is connected and contains at least two points, there exist sequences $\{u_j^{(i)}\}_{j=1}^{\infty} \subset \Omega$ such that

$$\lim_{j\to\infty} u_j^{(i)} = u^{(i)} \quad \forall i=1,\dots,N \quad \text{and} \quad \varkappa_j := \min_{p,q=1,\dots,N}\Big|u_j^{(p)}-u^{(q)}\Big| > 0 \quad \forall j. \quad (11.8)$$

Now introduce the controls $\nu, \nu_j \in \mathfrak{CM}$ in the problem (5.3), (5.4) by putting

$$\nu_j\big(t, du', du''\big) := \frac{1}{2}\Big[\delta_{\big[u^{(i)}, u_j^{(i)}\big]}\big(du', du''\big) + \delta_{\big[u_j^{(i)}, u^{(i)}\big]}\big(du', du''\big)\Big] \quad \forall t \in E^{(i)},$$
$$\nu\big(t, du', du''\big) := \delta_{[u(t), u(t)]}\big(du', du''\big). \quad (11.9)$$

Here $i = 1, \dots, N$. For any $\zeta(\cdot) \in \mathbb{C}(T)$ and $\rho(\cdot) \in \mathbb{C}(\Omega \times \Omega)$,

$$\int_T \zeta(t)\, dt \int_{\Omega\times\Omega} \rho\big(w\big)\nu_j\big(t, dw\big) = \sum_{i=1}^N \frac{1}{2}\int_{E^{(i)}} \zeta(t)\Big[\rho\big(u^{(i)}, u_j^{(i)}\big) + \rho\big(u_j^{(i)}, u^{(i)}\big)\Big]$$
$$\xrightarrow[j\to\infty]{(11.8)} \sum_{i=1}^N \int_{E^{(i)}} \zeta(t)\rho\big(u^{(i)}, u^{(i)}\big) = \int_T \zeta(t)\, dt \int_{\Omega\times\Omega} \rho\big(u', u''\big)\nu\big(t, du', du''\big).$$

So $\nu_j \to \nu$ in $\mathfrak{CM}$ as $j \to \infty$ due to Lemma IV.2.4 [23]. By the last claim from Lemma 6.4, $I_{(5.3)}\big(\nu_j|\xi\big) \to I_{(5.3)}\big(\nu|\xi\big)$ as $j \to \infty$, where $I_{(5.3)}\big(\nu|\xi\big)$ is the functional from (5.3). Evidently, $I_{(5.3)}\big(\nu|\xi\big) = I_{(2.5)}\big[u(\cdot)|\xi\big]$. Hence (11.7) yields

$$I_{(5.3)}\big(\nu_j|\xi\big) \le \mathfrak{I}_{opt}(\xi) + \varepsilon \quad (11.10)$$

for some j. From now on, we fix such a j. Pick a continuous function $\chi(\cdot) := [0, +\infty) \to [0, 1]$ such that $\chi(0) = 0$ and $\chi(r) = 1$ whenever $r \ge \varkappa_j$, where $\varkappa_j$ is taken from (11.8). Set

$$\varphi_\varepsilon(t, x_1, x_2, u_1, u_2) := \varphi(t, x_1, x_2, u_1, u_2) - \frac{2\varepsilon}{t_1 - t_0}\chi(|u_2 - u_1|) \quad (11.11)$$

and consider the perturbation $\xi_\varepsilon := [f(\cdot), \varphi_\varepsilon(\cdot), a(\cdot)] \in \Xi_{(iv)}$ of ξ. For ξ and ξ_ε, the natural limit problems (2.5), (2.6) are identical and so

$$\mathfrak{I}_{opt}(\xi_\varepsilon) = \mathfrak{I}_{opt}(\xi). \quad (11.12)$$

At the same time, $x(\cdot) := x(\cdot|\nu_j, \xi) = x(\cdot|\nu_j, \xi_\varepsilon)$ and

$$\mathfrak{I}_{opt}^{\lim}(\xi_\varepsilon) \le I_{(5.3)}(\nu_j|\xi_\varepsilon) \overset{(5.3)}{=\!=} \int_T dt \int_{\Omega\times\Omega} \varphi_\varepsilon\left[t, x(t), x(t), u_1, u_2\right]\nu_j(t, du_1, du_2)$$
$$\overset{(11.9)}{=\!=} \sum_{i=1}^N \frac{1}{2}\int_{E^{(i)}} \Big\{\varphi_\varepsilon\big[t, x(t), x(t), u^{(i)}, u_j^{(i)}\big] + \varphi_\varepsilon\big[t, x(t), x(t), u_j^{(i)}, u^{(i)}\big]\Big\}dt$$
$$\overset{(11.11)}{=\!=} \sum_{i=1}^N \frac{1}{2}\int_{E^{(i)}} \Big\{\varphi\big[t, x(t), x(t), u^{(i)}, u_j^{(i)}\big] + \varphi\big[t, x(t), x(t), u_j^{(i)}, u^{(i)}\big]\Big\}dt$$
$$- \sum_{i=1}^N \frac{2\varepsilon}{t_1 - t_0}\int_{E^{(i)}} \underbrace{\chi\big(\left|u^{(i)} - u_j^{(i)}\right|\big)}_{=1} \overset{(5.3),\,(11.9)}{=\!=} I_{(5.3)}(\nu_j|\xi) - 2\varepsilon \overset{(11.10)}{\le} \mathfrak{I}_{opt}(\xi) - \varepsilon$$
$$\overset{(11.12)}{=\!=} \mathfrak{I}_{opt}(\xi_\varepsilon) - \varepsilon.$$

This means that $\xi_\varepsilon \in \Xi_{(3.1)}$. It remains to note that $\xi_\varepsilon \to \xi$ as $\varepsilon \to 0 + 0$ and so $\xi \in \overline{\Xi_{(3.1)}}$.

12. Proofs of the statements from Section 4. *Proof of Theorem* 4.3 Consider the problem (2.1)–(2.4) with $f(t,x_1,x_2,u_1,u_2) := \overline{f}(t,x_1,u_1), a(t) := a_0, \varphi(t,x_1,x_2,u_1,u_2) := \overline{\varphi}(t,x_1,u_1) - B|u_1 - u_2|$, where B will be specified later. This problem meets the requirements (i)–(iv) from Section 3 and its natural "limit" case equals the original problem (2.5), (2.6) (i.e., (2.7) is true). By invoking Theorem 5.1 and employing the element (11.1), we get

$$\lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta) \le \int_T dt \int_{\Omega\times\Omega} (\overline{\varphi}[t,x(t),u'] - B|u'-u''|)\ \nu(t,du',du'')$$
$$= \frac{1}{2(t_1-t_0)} \int_T (\overline{\varphi}[t,x(t),u_1] + \overline{\varphi}[t,x(t),u_2])\ dt - B|u_1-u_2|.$$

Here $x(\cdot)$ is the solution of (5.4). Due to Lemma 6.4, $|x(t)| \le k\ \forall t \in [t_0,t_1]$. Hence

$$\lim_{\Delta \to +0} \mathfrak{I}_{opt}(\Delta) \le \overbrace{\max_{t\in[t_0,t_1],u\in\Omega,|x|\le k} |\overline{\varphi}(t,x,u)|}^{\varkappa} - B\overbrace{|u_1-u_2|}^{\delta}.$$

By picking $B > \delta^{-1}(\varkappa - \mathfrak{I}_{opt}(0) + A)$, we arrive at (4.2) and complete the proof. □

Proof of Theorem 4.4. Let $[x(\cdot),\nu]$ be a process in the problem (5.3), (5.4). Due to (5.2),

$$u(t) := \int_\Omega u\,\nu(t,du,\Omega) = \int_\Omega u\,\nu(t,\Omega,du).$$

Since the set Ω is convex and compact, $u(t) \in \Omega$. Furthermore, (2.7) and (5.4) imply that

$$\dot{x}(t) = \int_{\Omega\times\Omega} \{A_1[t,x(t),x(t)]u' + A_2[t,x(t),x(t)]u''\}\ \nu(t,du',du'')$$
$$= \sum_{i=1}^{2} A_i[t,x(t),x(t)]u(t) = f[t,x(t),x(t),u(t),u(t)] = \overline{f}[t,x(t),u(t)],$$

and $x(0) = a(0)$. Thus (2.6) holds, and $[x(\cdot),u(\cdot)]$ is a process in the problem (2.5), (2.6). Owing to the Ensen inequality [9, p.402] and (2.7),

$$\int_T dt \int_{\Omega\times\Omega} \varphi[t,x(t),x(t),u',u'']\,\nu(t,du',du'')$$
$$\ge \int_T \varphi\left\{t,x(t),x(t),\int_{\Omega\times\Omega}[u',u'']\ \nu(t,du',du'')\right\}\ dt$$
$$= \int_T \varphi[t,x(t),x(t),u(t),u(t)]\,dt = \int_T \overline{\varphi}[t,x(t),u(t)]\,dt.$$

Summing up, we see that the infimum of the cost functional in the problem (5.3), (5.4) is not less than that in the problem (2.5), (2.6). At the same time, considering the elements ν of the form $\nu(t,du',du'') = \delta_{[u(t),u(t)]}(du',du'')$ in (5.3), (5.4) assures that the first infimum does not exceed the second one. Thus these infima coincide. Theorem 5.1 complete the proof. □

The proof of Lemma 4.1 is prefaced by the following technical fact.

LEMMA 12.1. *Given* $u(\cdot) \in \mathbb{L}_\infty([t_0-1,t_1] \to \Omega)$ *and* $\Delta \in [0,1]$*, the symbol* ν_Δ *stands for the element given by* (5.8)*. The function* $\Delta \mapsto \nu_\Delta \in \mathfrak{P}(T)$ *is continuous.*

Proof. By [23, Theorem IV.2.4], it suffices to show that for any $\zeta(\cdot) \in \mathbb{C}(T), \rho(\cdot) \in \mathbb{Y} := \mathbb{C}(\Omega\times\Omega)$, and $\Delta \in [0,1]$,

$$\Lambda_{\Delta',\Delta}[\rho(\cdot)] := \int_T \zeta(t)\,\{\rho[u(t),u(t-\Delta')] - \rho[u(t),u(t-\Delta)]\}\ dt \to 0 \qquad (12.1)$$

as $\Delta' \to \Delta$. Fix Δ and $\zeta(\cdot)$. All $\rho(\cdot) \in \mathbb{Y}$ such that (12.1) is true constitute a linear subspace $\mathcal{R}$ of $\mathbb{Y}$. We are going to show that it is closed. Indeed, let $\rho_*(\cdot) \in \overline{\mathcal{R}}$. For any $\varepsilon > 0$, there exists $\rho(\cdot) \in \mathcal{R}$ such that $\varkappa := \max_{\omega \in \Omega\times\Omega} |\rho(\omega) - \rho_*(\omega)| \leq \varepsilon$. Then

$$|\Lambda_{\Delta',\Delta}[\rho_*(\cdot)]| \leq |\Lambda_{\Delta',\Delta}[\rho_*(\cdot) - \rho(\cdot)]| + |\Lambda_{\Delta',\Delta}[\rho(\cdot)]| \leq 2\varkappa \int_T |\zeta(t)|\, dt + |\Lambda_{\Delta',\Delta}[\rho(\cdot)]|;$$
$$\overline{\lim}_{\Delta'\to\Delta} |\Lambda_{\Delta',\Delta}[\rho_*(\cdot)]| \leq 2\varepsilon \int_T |\zeta(t)|\, dt + \lim_{\Delta'\to\Delta} |\Lambda_{\Delta',\Delta}[\rho(\cdot)]| = 2\varepsilon \int_T |\zeta(t)|\, dt.$$

Letting $\varepsilon \to +0$ shows that $\rho_*(\cdot) \in \mathcal{R}$, i.e., $\mathcal{R} = \overline{\mathcal{R}}$. Therefore it suffices to prove (12.1) for $\rho(\cdot) \in \mathcal{E}$, where $\mathcal{E}$ is an arbitrary set whose linear hull is dense in $\mathbb{Y}$. This permits us to focus on the functions $\rho(\cdot)$ of the form $\rho(u', u'') = \rho_1(u')\rho_2(u'')$, $\rho_1(\cdot), \rho_2(\cdot) \in \mathbb{C}(\Omega)$. For them,

$$|\Lambda_{\Delta',\Delta}[\rho(\cdot)]| = \left| \int_T \underbrace{\zeta(t)\rho_1[u(t)]}_{\chi(t)} \{\rho_2[u(t-\Delta')] - \rho_2[u(t-\Delta)]\}\, dt \right|$$
$$\leq \operatorname{esssup}_{t\in T} |\chi(t)| \int_T |r(t-\Delta') - r(t-\Delta)|\, dt, \to 0 \quad \text{as} \quad \Delta' \to \Delta,$$

where $r(t) := \rho_2[u(t)]$. Thus (12.1) is valid. $\square$

Proof of Lemma 4.1. We denote $u_\Delta(\cdot) := u|_{[t_0-\Delta,t_1]}(\cdot) \in U_\Delta, \mu(\Delta) := I_\Delta(u_\Delta) - \mathfrak{I}_{opt}(\Delta)$, and employ the notation ν_Δ introduced in Lemma 12.1. Then $I_\Delta(u_\Delta) = \mathfrak{I}_{opt}(\Delta) + \mu(\Delta)$. i.e., the control u_Δ is $\mu(\Delta)$-suboptimal. Furthermore thanks to Lemmas 6.4 and 12.1,

$$\lim_{\Delta\to 0} \mu(\Delta) = \lim_{\Delta\to 0} I_\Delta(u_\Delta) - \lim_{\Delta\to 0} \mathfrak{I}_{opt}(\Delta) = I(\nu_0) - \lim_{\Delta\to 0} \mathfrak{I}_{opt}(\Delta) = \mathfrak{I}_{opt}(0) - \lim_{\Delta\to 0} \mathfrak{I}_{opt}(\Delta),$$

which implies (4.1) and completes the proof. $\square$

REFERENCES

[1] A. Arutunov and M. Mardanov, *On the theory of the maximum principle in problems with delays*, Differential Equations, 29 (1989), pp. 2048–2058.

[2] H. Banks, *Necessary condition for problems with variable timelags*, SIAM J. Control Optim., 6 (1968), pp. 9–47.

[3] R. Bellman and K. Cooke, *On the limit of solutions of differential-difference equations as the retardation approaches zero*, Proceedings of National Academy of Sciences, 45 (1959), pp. 1026–1028.

[4] M. Bennati, *Wellposedness by perturbation in optimization problems and metric characterization*, Rend. Met. e Appl., 16 (1996), pp. 613–623.

[5] G. Bunce, *A maximum principle for time lag control problems with bounded state*, J. Optim. Theory and Appl., 22 (1977), pp. 563–606.

[6] F. Clarke and G. Watkins, *Necessary conditions, controlability and the value function for differential-difference inclusions*, Nonlinear Anal. Theory Meth. Appl., 10 (1986), pp. 1155–1179.

[7] F. Clarke and P. Wolensky, *The sensitivity of optimal control problems to time delay*, SIAM J. Control and Optimization, 29 (1991), pp. 1176–1215.

[8] R. Driver, *Some harmless delays*, in Delay and Functional Differential Equations, K. Schmitt, ed., Academic Press, London, 1972.

[9] R. Edwards, *Functional Analysis. Theory and Applications*, Holt, Rinehart and Winston, New York, London, 1965.

[10] J. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1977.

[11] A. Ioffe and V. Tihomirov, *Theory of extremal problems*, in Studies in Mathematics and its Applications, 6, North-Holland, Amsterdam, 1979.

[12] H. Ishii and B. Francis, *Limited data rate in control systems with networks*, in Lecture Notes in Control and Information Sciences, M. Thoma and M. Morari, eds., Springer, Berlin, 2002.

[13] G. Kharatishvili, *Maximum principle in the theory of optimum time-delay processes*, Dokl. Akad. Nauk USSR, 136 (1961), pp. 39–42.

[14] G. KHARATISHVILI AND T. TADUMADZE, *Nonlinear optimal control systems with varying delays*, Matem. Sbornik, 107 (1973), pp. 613–633. in Russian.

[15] A. MATVEEV, *Optimal control problems with delays of general form and phase constraints*, Math. USSR Izv., 33 (1989), pp. 1200–1229.

[16] ———, *On nonequivalence of the pointwise and integral maximum principles for systems with delays in the controls*, St. Petersburg Math. J., 4 (1993), pp. 749–775.

[17] M. OĞUZTORELI, *Time-Lag Control Systems*, Academic Press, New-York, London, 1966.

[18] L. PONTRAYGIN, V. BOLTAYNSKII, R. GAMKRELIDZE, AND E. MISCHENKO, *The Mathematical Theory of Optimal Processes*, John Wiley, New York, 1962.

[19] J. ROSENBLUETH, *Strongly and weakly relaxed controls for time delay systems*, SIAM J. Control and Optim., 30 (1992), pp. 856–866.

[20] J. ROSENBLUETH AND R. VINTER, *Relaxation procedures for time delay systems*, Journal of Mathematical Analysis and Applications, 162 (1991), pp. 542–563.

[21] S. SUGIYAMA, *Continuity properties on the retardation in the theory of differential-difference equations*, Proceedings of Japan Academy of Sciences, 37 (1961), pp. 179–182.

[22] J. WARGA, *The reduction of certain control problems to an "ordinary differential" form*, SIAM Review, 10 (1968), pp. 219–222.

[23] ———, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

[24] J. WARGA AND Q. ZHU, *A proper relaxation of shifted and delayed controls*, J. Math. Anal. Appl., 169 (1992), pp. 546–561.

[25] L. YOUNG, *Lectures in the Calculus of Variations and Optimal Control Theory*, W. B. Saunders, Phil., PA, 1969.

[26] T. ZOLEZZI, *Extended wellposedness of optimization problems*, Journal of Optimization Theory and Applications, 91 (1996), pp. 257–266.