TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

Optimal control of age-structured systems with mixed state-control constraints

Nikolai P. Osmolovskii, Vladimir M. Veliov

Research Report 2016-12

November, 2016

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-4
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

Optimal control of age-structured systems with mixed state-control constraints*

N.P. Osmolovskii[†] V.M. Veliov[‡]

Abstract

The paper deals with a general optimal control problem for age-structured systems. A necessary optimality condition of Pontryagin type is obtained, where the novelty is in that mixed control-state constraints are present. The proof uses an abstract Lagrange multiplier theorem, and the main difficulty is to obtain regularity of the Lagrange multipliers in the particular problem at hand.

Keywords: age-structured systems, optimal control.
MSC: 49K20, 35Q93, 92D25

1 Introduction

The optimal control theory for age-structured systems has proved to be useful in numerous branches of science: demography (migration or fertility control), population economics (life-cycle investment/education models), health economics (investments in health care), economics (vintage capital models), forestry and fishery (harvesting control), epidemiology (prevention/treatment control), sociology (drug consumption or rumor dynamics and control), etc. Among the huge amount of publications in the area we refer to the books [18, 12, 1] for the theory of age-structured systems, to [4, 10, 7, 17] for optimal control theory of such systems, and to [2, 8, 9, 3, 15] for applications in some of the above mentioned areas. The references are scarce, but the bibliography therein is much reacher.

In the present paper we consider a general age-structured optimal control problem of the form

$$(1) \qquad \min \left\{ \int_0^\omega l(a, y(T,a))\,\mathrm{d}a + \int_0^T \int_0^\omega L(t,a,y(t,a),q(t),u(t,a),v(t),w(a))\,\mathrm{d}a\,\mathrm{d}t \right\}$$

on the set of all functions $(y(t,a), q(t), u(t,a), v(t), w(a))$ satisfying the age-structured system

$$(2) \qquad \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) y(t,a) \;=\; f(t,a,y(t,a),q(t),u(t,a)), \quad (t,a) \in D,$$

*The first author is supported by Russian Foundation for Basic Research (RFFI) under grant No 16-01-00585, the second author is supported by the Austrian Science Foundation (FWF) under grant No P26640-N25.

[†]Department of Informatics and Mathematics, Kazimierz Pułaski University of Technology and Humanities in Radom, Radom, Poland; Systems Research Institute, Polish Academy of Sciences, Warszawa, Poland; Department of Applied Mathematics, Moscow State University of Civil Engineering, Moscow, Russia, osmolovski@uph.edu.pl

[‡]Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria, veliov@tuwien.ac.at

$$
\begin{aligned}
(3) \qquad y(0,a) &= y_0(a, w(a)), \quad a \in [0,\omega], \\
(4) \qquad y(t,0) &= \varphi(t, q(t), v(t)), \quad t \in [0,T], \\
(5) \qquad q(t) &= \int_0^\omega h(t,a,y(t,a))\, \mathrm{d}a, \quad t \in [0,T],
\end{aligned}
$$

and the additional constraints

$$
(6) \qquad g_i(t,a,y(t,a),u(t,a),v(t),w(a)) \le 0, \quad i = 1,\dots,d(g), \quad (t,a) \in D.
$$

Here $t \in [0,T]$ is interpreted as time, $a \in [0,\omega]$ as age, $D := [0,T] \times [0,\omega]$. The functions $y(t,a)$, $q(t)$, $u(t,a)$, $v(t)$, $w(a)$ take values in finite-dimensional Euclidean spaces, where y is regarded as state function, q is sometimes called aggregated state, and $u(t,a)$, $v(t)$, $w(a)$ are distributed, boundary and initial control functions, correspondingly. Equation (2), together with the initial condition (3) and the boundary condition (4), defines the state dynamics, (5) defines the aggregated state q; the functions f, φ and h have appropriate dimensions. Conditions (6) pose inequality type constraints on the control variables (if g_i is independent of y for some indexes i), and mixed control-state constraints (for those i for which g_i depends also on y). The precise statement of the problem and the assumptions used are given in the next section.

In the present paper we obtain necessary optimality conditions of Pontryagin's type for the problem (1)–(6). The novelty is that the problem involves mixed control-state constraints.

Pure state constraints arise in many applications of age-structured optimal control, usually in the form $y(t,a) \ge 0$. This is due to the necessity to keep the size of populations or capital stocks nonnegative, although the controlled dynamics may formally allow to steer the state to negative values. For example, in many economic models selling (or buying) capital stock is a control variable that may formally drive the capital to negative values if this is not explicitly prohibited.

Pure state constraints create substantial difficulties in the context of optimal control of age-structured systems, due to the infinite-dimensionality of the state. There are only a few known results: [7] (where the dynamic programming approach is employed), [14] (where only the end-state is constraint) and a few more in which the derivations are not sound.

We stress that in the present paper we do not consider pure state constraints, rather, mixed control-state constraints; this is due to our assumption (assumption A below), which requires the control to be essentially present in the constraints (6). Nevertheless, mixed constraints are also important in the context of optimal control of age-structured systems. First, often such constraints are meaningful in particular applications. For example, the maximal size of the feasible bank credit (investment control) may depend on the present capital stock (state variable). Second, mixed constraints can often be used for approximating problems with pure sate constraints. This approach is systematically employed in [19], for example, and can also be used for approximate numerical solution, since the Lagrange multipliers in the case of mixed constraints have a certain regularity, in contrast to the case of pure state constraints.

In the derivation of the Pontryagin-type optimality conditions (maximum principle) we use a general Lagrange multiplier theorem presented in [5] and [16], which results in a local form of the maximum principle. It is known, e.g. from [4, 10], that in the case of only control constraints also the global form of the maximal principle holds true for the distributed control $u(t,a)$. Of course, the local maximum principle implies in a standard way the global one on additional linearity/convexity-type assumptions also for problems with mixed constraints. However, it is an open question to obtain a global maximum principle for problem (1)–(6) without such additional assumptions.

The paper is organized as follows. In the next section we give a strict formulation of the problem together with the necessary assumptions. Section 3 presents the main result – a complete set of Pontryagin type necessary optimality (a local maximum principle). The proof of the main result is given in Section 4, which is split in several subsection for reader's convenience.

2 Statement of the problem

In order to give a precise meaning to equations (2)–(5) (in line with, and partly repeating, the previous contributions, e.g. [18, 4, 1, 10]) and the subsequent considerations, we introduce some notations, spaces and definitions.

First of all, denote by $d(z)$ the dimension of a vector z, that is, $z \in \mathbb{R}^{d(z)}$. In particular, this applies to y, q, u, etc., so that $y \in \mathbb{R}^{d(y)}$, etc. Then $f : \mathbb{R}^{1+1+d(y)+d(q)+d(u)} \to \mathbb{R}^{d(y)}$, $y_0 : \mathbb{R}^{1+d(w)} \to \mathbb{R}^{d(y)}$, $\varphi : \mathbb{R}^{1+d(q)+d(v)} \to \mathbb{R}^{d(y)}$, and $h : \mathbb{R}^{1+1+d(y)} \to \mathbb{R}^{d(q)}$. Assume that these functions are continuous together with their partial derivatives w.r.t. y, q, u, v, w. The same assumption also applies to the functions l and L in the objective functional and to the functions g_i in the constraints.

Consider the family Σ of all maximal segments $S \subset D$ parallel to the vector $e = (1, 1)$ (in line with the literature on first order PDEs, these will be called *characteristic segments*). Each such segment corresponds to a (let-most) point $(0, a) \in \{0\} \times (0, \omega]$ or to a (lowest) point $(t, 0) \in [0, T] \times \{0\}$ (see Fig. 1). Set

$$\Gamma = (\{0\} \times (0, \omega]) \cup ([0, T] \times \{0\}).$$

Conversely, each point $\gamma \in \Gamma$ defines a unique $S_\gamma \in \Sigma$ emanating from γ, and $\gamma \mapsto S_\gamma$ is one-to-one correspondence between the points of Γ and the characteristic segments $S \in \Sigma$.

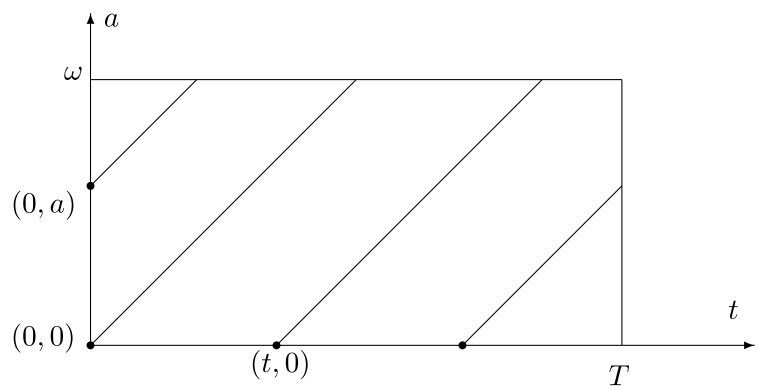

Fig. 1. Domain D and characteristic lines emanating from points in Γ.

If $\tilde{\Gamma}$ is a subset of Γ of full measure (in the sense of the Lebesgue measure), then we say that the corresponding subset $\tilde{\Sigma} \subset \Sigma$ consists of *almost all* segments in Σ or that it is a *full* subset of Σ.

For $S \in \Sigma$ and a Lipschitz continuous function $x : S \to \mathbb{R}^{d(y)}$ denote by $\frac{\partial x}{\partial e}$ its directional derivative in the direction $e := (1, 1)$, that is, $\frac{\partial x}{\partial e}(s) = \lim_{\sigma \to 0+}(x(s + \sigma e) - x(s))/\sigma$, $s \in S$.

For $S \in \Sigma$ we define as usual the space $W^{1,\infty}(S, \mathbb{R}^{d(y)})$ consisting of all Lipschitz continuous functions $x : S \to \mathbb{R}^{d(y)}$, with the norm

$$\|x\|_{W^{1,\infty}\left(S,\mathbb{R}^{d(y)}\right)} := \|x\|_{C\left(S,\mathbb{R}^{d(y)}\right)} + \left\|\frac{\partial x}{\partial e}\right\|_{L^\infty\left(S,\mathbb{R}^{d(y)}\right)}.$$

Let us introduce the space $\mathcal{Y}_\Sigma$ consisting of all measurable functions $y : D \to \mathbb{R}^{d(y)}$ for which the restriction $y_{|S}$ on almost every characteristic segment $S \in \Sigma$ is Lipschitz continuous, and in addition, the norm

$$\|y\|_{\mathcal{Y}_\Sigma} := \operatorname*{esssup}_{\gamma\in\Gamma} \| \, y_{|S_\gamma} \|_{W^{1,\infty}(S_\gamma,\mathbb{R}^{d(y)})}$$

is finite.

Clearly every $y \in \mathcal{Y}_\Sigma$ belongs to the space $L^\infty\left(D, \mathbb{R}^{d(y)}\right)$, as well as $\frac{\partial y}{\partial e}$. Moreover, $\mathcal{Y}_\Sigma$ is a Banach space.

Let $u \in L^\infty(D, \mathbb{R}^{d(u)})$, $v \in L^\infty([0, T], \mathbb{R}^{d(v)})$, $w \in L^\infty([0, \omega], \mathbb{R}^{d(w)})$ be any triple of control functions, and let $y \in \mathcal{Y}_\Sigma$, $q \in L^\infty([0, T], \mathbb{R}^{d(q)})$. Then the right-hand side of (2) is measurable and bounded, hence (by the Fubini theorem), its restriction on almost every characteristic segment $S \in \Sigma$ belongs to $L^\infty(S, \mathbb{R}^{d(y)})$. Due to the definition of the space $\mathcal{Y}_\Sigma$, the same applies to the left-hand side of (2), where the expression $\frac{\partial}{\partial t} + \frac{\partial}{\partial a}$ has to be interpreted as the directional derivative $\frac{\partial y}{\partial e}$. Therefore, (2) is understood as an equation that has to be fulfilled almost everywhere on almost every characteristic segment.

Moreover, for any $y \in \mathcal{Y}_\Sigma$ the restrictions $[0, \omega] \ni a \mapsto y(0, a)$ and $[0, T] \ni t \mapsto y(t, 0)$ are well defined elements of $L^\infty([0, \omega], \mathbb{R}^{d(y)})$ and $L^\infty([0, T], \mathbb{R}^{d(y)})$, respectively (notice that the same applies to the restriction of y on any horizontal or vertical segment in D). Then (3) and (4) should also be understood as equations in $L^\infty([0, \omega], \mathbb{R}^{d(y)})$ and $L^\infty([0, T], \mathbb{R}^{d(y)})$, respectively. Similarly, (5) is an equation in $L^\infty([0, T], \mathbb{R}^{d(q)})$.

Having in mind the above, we define that the pair $(y, q) \in \mathcal{Y}_\Sigma \times L^\infty([0, T], \mathbb{R}^{d(q)})$ is a solution of system (2)–(5) (for given (u, v, w) as above) if (2) is fulfilled on almost everywhere on almost every characteristic segment $S \in \Sigma$ and (3)–(5) hold almost everywhere in the respective intervals. Clearly, if a solution exists, then the objective functional (1) is well defined and finite. The right-hand sides of the constraints (6) are measurable and bounded, and have to be satisfied for a.e. $(t, a) \in D$.

Thus we consider problem (1)–(6) in the Banach space

$$X := \mathcal{Y}_\Sigma \times L^\infty([0, T], \mathbb{R}^{d(q)}) \times L^\infty(D, \mathbb{R}^{d(u)}) \times L^\infty([0, T], \mathbb{R}^{d(v)}) \times L^\infty([0, \omega], \mathbb{R}^{d(w)})$$

of functions (y, q, u, v, w), where the norm is defined as

$$\|(y, q, u, v, w)\| = \|y\|_{\mathcal{Y}_\Sigma} + \|q\|_\infty + \|u\|_\infty + \|v\|_\infty + \|w\|_\infty.$$

A local minimum of the problem in this norm is called a *weak local minimum*. Our aim in this paper is to obtain first order necessary conditions for a weak local minimum. This will be done by using an abstract Lagrange multipliers rule, recalled in Section 4.1. In order to prove regularity of the Lagrange multipliers in the context of our specific problem we need an additional assumption formulated in the next lines.

Denote the set of active constraints at a point (t, a, y, u, v, w) by

$$I(t, a, y, u, v, w) = \{i \in \{1, \ldots, d(g)\} : g_i(t, a, y, u, v, w) = 0\}.$$

Assumption A. At any point (t, a, y, u, v, w) such that $(t, a) \in D$ and $g_i(t, a, y, u, v, w) \le 0$, $i = 1, \dots, d(g)$, the gradients of the active constraints g_i with respect to u, that is the vectors

$$g_{iu}(t, a, y, u, v, w), \quad i \in I(t, a, y, u, v, w),$$

are positively linearly independent.

Recall that a system of vectors $p_1, \dots, p_m$ in $\mathbb{R}^r$ is said to be *positively linearly independent* if the equality $\alpha_1 p_1 + \ldots + \alpha_m p_m = 0$ with all $\alpha_i \ge 0$ implies that $\alpha_i = 0$ for all i.

3 Main result: a necessary optimality condition

We assume that problem (1)–(6) has a week local solution $(\bar{y}, \bar{q}, \bar{u}, \bar{v}, \bar{w}) \in \mathcal{X}$. In the sequel we use the following notational convention: we skip functions with "bar" whenever they play the role of arguments of other functions. For example, $f_y(t, a) := f_y(t, a, \bar{y}(t, a), \bar{q}(t), \bar{u}(t, a))$, $g_i(t, a) := g_i(t, a, \bar{y}(t, a), \bar{u}(t, a), \bar{v}(t), \bar{w}(a)))$, etc. Moreover, for convenience we consider the "adjoint" vectors ψ and μ that appear below as row-vectors, in contrast to the column-vectors y, q, etc. that were involved so far.

The next theorem gives a Pontryagin type necessary optimality condition.

Theorem 3.1 *Let assumption A be fulfilled and let $(\bar{y}(t, a), \bar{q}(t), \bar{u}(t, a), \bar{v}(t), \bar{w}(a))$ be a weak local solution of problem (1)–(6). Then there exist Lagrange multipliers $\alpha \in \{0, 1\}$ and $\lambda \in L^\infty(D, \mathbb{R}^{d(g)})$, with $\lambda_i(t, a) \ge 0$ for a.e. $(t, a) \in D$ and $\alpha + \sum_{i=1}^{d(g)} \|\lambda_i\|_\infty > 0$, such that:*

(i) the complementarity conditions hold:

$$\lambda_i(t, a)\, g_i(t, a) = 0, \quad \text{for a.e. } (t, a) \in D, \quad i = 1, \dots, d(g);$$

(ii) the adjoint system

$$(7) \quad -\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) \psi(t, a) = \psi(t, a) f_y(t, a) + \mu(t) h_y(t, a) + \alpha L_y(t, a) + \sum_{i=1}^{d(g)} \lambda_i(t, a) g_{iy}(t, a),$$

$$(8) \quad \mu(t) = \psi(t, 0)\varphi_q(t) + \int_0^\omega \Big(\psi(t, a) f_q(t, a) + \alpha L_q(t, a)\Big)\, \mathrm{d}a,$$

with end- and boundary (transversality) conditions

$$(9) \quad \psi(T, a) = \alpha l_y(a, \bar{y}(T, a)),$$

$$(10) \quad \psi(t, \omega) = 0,$$

has a unique solution $(\psi, \mu) \in \mathcal{Y}_\Sigma \times L^\infty([0, T], \mathbb{R}^{d(q)})$;

(iii) the local minimum principle holds:

$$(11) \quad \psi(t, a) f_u(t, a) + \alpha L_u(t, a) + \sum_{i=1}^{d(g)} \lambda_i(t, a) g_{iu}(t, a) = 0 \quad \text{for a.e. } (t, a) \in D,$$

$$(12) \quad \psi(t, 0)\varphi_v(t) + \int_0^\omega \Big(\alpha L_v(t, a) + \sum_{i=1}^{d(g)} \lambda_i(t, a) g_{iv}(t, a)\Big)\, \mathrm{d}a = 0 \quad \text{for a.e. } t \in [0, T],$$

$$(13) \quad \psi(0, a) y_{0w}(a) + \int_0^T \Big(\alpha L_w(t, a) + \sum_{i=1}^{d(g)} \lambda_i(t, a) g_{iw}(t, a)\Big)\, \mathrm{d}t = 0 \quad \text{for a.e. } a \in [0, \omega].$$

The meaning of solution of system (7)–(10) is similar as that of the primal system (2)–(5).

In the next paragraph we reformulate the result in Theorem 3.1 in a (point-wise) Hamiltonian form.

Let us introduce the *Hamiltonian*

$$H(t,a,y,q,u,v,w,\psi) = \psi f(t,a,y,q,u) + \alpha L(t,a,y,q,u,v,w) \tag{14}$$

and the *augmented Hamiltonian*

$$\begin{aligned} H^a(t,a,y,q,u,v,w,\psi,\lambda,\mu) &= H(t,a,y,q,u,v,w,\psi) \\ &+\mu h(t,a,y) + \sum_{i=1}^{d(g)} \lambda_i g_i(t,a,y,u,v,w), \end{aligned} \tag{15}$$

where $\psi^\top \in \mathbb{R}^{d(y)}$ and $\mu^\top \in \mathbb{R}^{d(q)}$ and "$\top$" means transposition. Then, for the solution $(\bar{y}(t,a), \bar{q}(t), \bar{u}(t,a), \bar{v}(t), \bar{w}(a))$, the adjoint equation can be written as

$$-\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\psi(t,a) = H^a_y(t,a),$$

where the multiplier $\mu(t)$ is given by the relation

$$\mu(t) = \psi(t,0)\,\varphi_q(t) + \int_0^\omega H_q(t,a)\,\mathrm{d}a.$$

The local minimum principle in (iii) takes the form

$$\begin{aligned} H^a_u(t,a) &= 0, \\ \psi(t,0)\,\varphi_v(t) + \int_0^\omega H_v(t,a)\,\mathrm{d}a &= 0, \\ \psi(0,a)\,y_{0w}(a) + \int_0^T H_w(t,a)\,\mathrm{d}t &= 0. \end{aligned} \tag{16}$$

4 Proof of Theorem 3.1

In this section we first recall an abstract Lagrange multiplier theorem, then in the subsequent subsections we verify that the assumptions of this theorem are fulfilled for our problem (1)–(6), and derive the statements of Theorem 3.1.

4.1 Lagrange multipliers rule for an abstract problem

Let X, Y, Z be Banach spaces, $\mathcal{D} \subset X$ be an open set, $K \subset Z$ be a closed convex cone with nonempty interior. Let $f : \mathcal{D} \to \mathbb{R}$, $F : \mathcal{D} \to Y$, $b_i : \mathcal{D} \to Z$, $i = 1, \ldots, \nu$, be given mappings. Consider the following problem:

$$f(x) \to \min, \qquad F(x) = 0, \qquad b_i(x) \in K, \;\; i = 1, \ldots, \nu. \tag{17}$$

Let $K^0 := \{\zeta^* \in Z^* : \langle \zeta^*, \kappa \rangle \le 0 \text{ for every } \kappa \in K\}$ be the polar cone to K. Here $\langle \zeta, \kappa \rangle$ is the duality pairing between Z and its dual space Z^*.

The following theorem gives necessary conditions for a point $\bar{x} \in \mathcal{D}$ to be a local minimizer for the problem (17) (see, e.g., [5] and [16]).

THEOREM 4.1 *Let $\bar{x}$ provides a local minimum in problem (17). Assume that the objective function f and the mappings b_i are Fréchet differentiable at $\bar{x}$, the operator F is continuously Fréchet differentiable at $\bar{x}$, and $F'(\bar{x})X = Y$ (the regularity of equality constraint). Then there exist Lagrange multipliers $\alpha \ge 0$, $\zeta_i^* \in K^0$, $i = 1, \ldots, \nu$, and $y^* \in Y^*$, satisfying the nontriviality condition*

$$\alpha + \sum_{i=1}^{\nu} \|\zeta_i^*\| > 0,$$

the complementary slackness conditions

$$\langle \zeta_i^*, \, b_i(\bar{x}) \rangle = 0, \qquad i = 1, \ldots, \nu,$$

and such that the Lagrange function

$$\mathcal{L}(x) = \; \alpha f(x) + \langle y^*, \, F(x) \rangle + \sum_{i=1}^{\nu} \langle \zeta_i^*, \, b_i(x) \rangle$$

is stationary at $\bar{x}$: $\mathcal{L}'(\bar{x}) = 0$.

4.2 Equality operator and its derivative

Later in this section we shall apply Theorem 4.1 with the following specifications for the spaces X and Y and the mapping F. The space X is as defined in Section 2,

$$Y = L^\infty(D, \mathbb{R}^{d(y)}) \times L^\infty([0, \omega], \mathbb{R}^{d(y)}) \times L^\infty([0, T], \mathbb{R}^{d(y)}) \times L^\infty([0, T], \mathbb{R}^{d(q)}),$$

and for $x = (y, q, u, v, w) \in X$ we define

$$F(x)(t, a) = \begin{pmatrix} f(t, a, y(t, a), q(t), u(t, a)) - \frac{\partial}{\partial e} y(t, a) \\ y_0(a, w(a)) - y(0, a) \\ \varphi(t, q(t), v(t)) - y(t, 0) \\ \int_0^\omega h(t, a, y(t, a)) \, \mathrm{d}a - q(t) \end{pmatrix}.$$

Clearly, F maps X to Y and equations (2)–(5) can be rewritten as $F(x) = 0$. The aim of this subsection is to prove that the surjectivity condition in Theorem 4.1 is fulfilled.

LEMMA 4.1 *At any point $\bar{x} := (\bar{y}, \bar{q}, \bar{u}, \bar{v}, \bar{w}) \in X$ the operator $F'(\bar{y}, \bar{q}, \bar{u}, \bar{v}, \bar{w}) : X \to Y$ is surjective, that is $F'(\bar{x})X = Y$.*

Proof. Below we use again our notational convention to skip arguments of functions with "bar". Then for any $x = (y, q, u, v, w) \in X$ we can represent

$$F'(\bar{x})x = \begin{pmatrix} f_y(t,a)y(t,a) + f_q(t,a)q(t) + f_u(t,a)u(t,a) - \frac{\partial}{\partial e} y(t,a) \\ y_{0w}(a)w(a) - y(0,a) \\ \varphi_q(t)q(t) + \varphi_v(t)v(t) - y(t,0) \\ \int_0^\omega h_y(t,a)y(t,a)\,\mathrm{d}a - q(t) \end{pmatrix}.$$

We have to prove that for every element $(m, n, r, p) \in Y$ there exists an element $x = (y, q, u, v, w) \in X$ such that $F'(\bar{x})x = (m, n, r, p)$. Setting $u(t,a) = 0,\ v(t) = 0,\ w(a) = 0$, this follows from the next lemma. □

LEMMA 4.2 *Consider the system*

$$\frac{\partial}{\partial e} y(t,a) = A(t,a)y(t,a) + B(t,a)q(t) + m(t,a), \tag{18}$$

$$y(0,a) = n(a), \quad y(t,0) = C(t)q(t) + r(t), \tag{19}$$

$$q(t) = \int_0^\omega Q(t,a)y(t,a)\,\mathrm{d}a + p(t), \tag{20}$$

where the matrices A, B, C, Q have measurable and essentially bounded entries. Then for every (m, n, r, p) in Y the above system has a unique solution (y, q) in $\mathcal{Y}_\Sigma \times L^\infty([0,T], \mathrm{I\!R}^{d(q)})$.

REMARK 4.1 The existence part of this lemma can be proved in a similar way as Lemma 5.1 in [4] by using the (linear version of the) Banach contraction theorem. The proof below provides, in addition, an explicit representation of the solution (y, q), given by the formulae (21), (22), (24), (25), which will be used in Section 4.6 below. This proof is also known, in principle, but we give it for completeness of the presentation.

Proof. For $a_0 \in [0, \omega]$ denote by $F_0(a_0; s, \tau),\ s, \tau \in [0, \min\{\omega - a_0, T\}]$, the fundamental matrix solution of the homogeneous ODE system

$$\dot{x}(s) = A(s, a_0 + s)x(s)$$

normalized at $s = \tau$, that is, $F_0(a_0; \tau, \tau) = I$ – the unit matrix. Similarly, for $t_0 \in [0, T]$ denote by $F(t_0; s, \tau),\ s, \tau \in [0, \min\{\omega, T - t_0\}]$, the fundamental matrix solution of

$$\dot{x}(s) = A(t_0 + s, s)x(s)$$

normalized at $s = \tau$. Then using the Cauchy formula one can obtain the following representation of solution of equation (18), where q is taken as given. For $(t, a) \in D$ with $t < a$,

$$y(t,a) = F_0(a-t; t, 0)n(a-t) + \int_0^t F_0(a-t; t, \tau)[B(\tau, a-t+\tau)q(\tau) + m(\tau, a-t+\tau)]\,\mathrm{d}\tau. \tag{21}$$

For $(t, a) \in D$ with $t \geq a$, we have

$$\begin{aligned} y(t,a) = {} & F(t-a; a, 0)[C(t-a)q(t-a) + r(t-a)] \\ & + \int_0^a F(t-a; a, \tau)[B(t-a+\tau, \tau)q(t-a+\tau) + m(t-a+\tau, \tau)]\,\mathrm{d}\tau. \end{aligned} \tag{22}$$

Substituting this expressions in (20) we obtain an integral equation for q as follows. For $t \in [0, \omega)$ (if $T > \omega$) we have

$$
\begin{aligned}
q(t) &= p(t) + \left(\int_0^t + \int_t^\omega \right) Q(t,a) y(t,a) \, \mathrm{d}a \\
&= p(t) + \int_0^t Q(t,a) F(t-a; a, 0)[C(t-a) q(t-a) + r(t-a)] \, \mathrm{d}a + \\
&\quad + \int_0^t Q(t,a) \int_0^a F(t-a; a, \tau)[B(t-a+\tau, \tau) q(t-a+\tau) + m(t-a+\tau, \tau)] \, \mathrm{d}\tau \, \mathrm{d}a \\
&\quad + \int_t^\omega Q(t,a) F_0(a-t; t, 0) n(a-t) \, \mathrm{d}a \\
&\quad + \int_t^\omega Q(t,a) \int_0^t F_0(a-t; t, \tau)[B(\tau, a-t+\tau) q(\tau) + m(\tau, a-t+\tau)] \, \mathrm{d}\tau \, \mathrm{d}a.
\end{aligned}
$$

For $t \in [\omega, T]$ we have the expression

$$
\begin{aligned}
q(t) &= p(t) + \int_0^\omega Q(t,a) F(t-a; a, 0)[C(t-a) q(t-a) + r(t-a)] \, \mathrm{d}a + \\
&\quad + \int_0^\omega \int_0^a Q(t,a) F(t-a; a, \tau)[B(t-a+\tau, \tau) q(t-a+\tau) + m(t-a+\tau, \tau)] \, \mathrm{d}\tau \, \mathrm{d}a.
\end{aligned}
$$

Consider the three terms containing q in the first case, $t \in [0, \omega)$. Changing the variable $t - a = s$ in the first integral, changing $t - a + \tau = s$ and the order of integration in the second integral, and changing the order of integration in the third integral, we obtain that the sum of these three integrals has the form

$$
\int_0^t K_0(t,s) q(s) \, \mathrm{d}s,
$$

where K_0 is a $d(q) \times d(q)$)-matrix function with measurable and essentially bounded entries.

In a similar way the sum of the two terms containing q in the case $t \in [\omega, T]$ can be represented in the form

$$
\int_{t-\omega}^t K_1(t,s) q(s) \, \mathrm{d}s,
$$

where K_1 is function like K_0 above. Concatenating K_0 and K_1 we obtain a measurable and bounded matrix function $K(t,s)$ such for any $t \in [0, T]$ the function q must satisfy the equation

$$
q(t) = \int_{\max\{0, \omega - t\}}^t K(t,s) q(s) \, \mathrm{d}s + \rho(t), \tag{23}
$$

where ρ represents the sum of all terms that appear in the representations of q above. Using similar manipulations of the integrals appearing in ρ we obtain a representation of the form

$$
\begin{aligned}
\rho(t) &= p(t) + \int_0^t Q^r(t,s) r(s) \, \mathrm{d}s + \int_0^\omega Q^n(t,a) n(a) \, \mathrm{d}a \\
&\quad + \int_0^\omega \int_0^{\min\{t,a\}} Q^m(t,a,\tau) m(t-\tau, a-\tau) \, \mathrm{d}\tau \, \mathrm{d}a,
\end{aligned} \tag{24}
$$

where Q^r, Q^n and Q^m are matrix functions of corresponding dimensions, defined, measurable and bounded on $[0,T] \times [0,T]$, $[0,T] \times [0,\omega]$, and $[0,T] \times [0,\omega] \times [0,\omega]$, correspondingly.

Notice that each of the terms in the definition of ρ is a measurable and bounded function. Then equation (23) regarded as a Volterra integral equation of the second kind has a unique solution $q \in L^\infty([0,T])$ (see Corollary 3.10 and Theorem 3.6 in [11]). This solution can be represented as

$$q(t) = \rho(t) + \int_0^t R(t,s)\rho(s)\,\mathrm{d}s, \tag{25}$$

where R is the resolvent kernel, which is also measurable and bounded on $[0,T]\times[0,T]$, [11, Theorem 3.6].

Using formulae (21) and (22) we define a function y. Due to $m \in L^\infty(D, \mathbb{R}^{d(y)})$, we have that $y \in L_\infty(D, \mathbb{R}^{d(y)})$. Differentiating equations (21) and (22) along the characteristic lines we obtain that equations (18)–(20) are fulfilled by the pair (y,q) and $y \in \mathcal{Y}_\Sigma$.

The last claim of the lemma follows from the representations (24), (25), and (21)–(22). □

4.3 Applying Theorem 4.1 to Problem (1)–(6)

In order to put problem (1)–(6) into the framework of Theorem 4.1 we set $Z := L^\infty(D,\mathbb{R})$ and rewrite the constraints (6) in the form

$$g_i(\cdot,\cdot,y(\cdot,\cdot),u(\cdot,\cdot),v(\cdot),w(\cdot)) \in K, \quad i = 1,\dots,d(g),$$

where $K = L^\infty_-(D,\mathbb{R})$ (the cone of non-positive functions in the space Z). It is known that K is a closed, convex cone with nonempty interior in Z.

In the sequel we will need the following lemma (see, e.g., [16] for the proof).

Lemma 4.3 *Let $\bar z \in K$ (that is $\bar z(t,a) \le 0$ a.e. in D). Then, for $\lambda \in Z^*$, the conditions $\lambda \in K^0$ and $\langle \lambda, \bar z\rangle = 0$ (i.e. λ is a support functional to K at the point $\bar z$) are equivalent to the following conditions: $\lambda \ge 0$, and for each $\delta > 0$ the element λ is concentrated on the set*

$$M_\delta = \{(t,a) \in D : \bar z(t,a) \ge -\delta\}.$$

We recall that $\lambda \ge 0$ means that $\langle \lambda, z\rangle \ge 0$ for every $z \in -K$, and "λ is concentrated on M_δ" means that $\langle \lambda, z\rangle = 0$ for every z which equals zero almost everywhere in M_δ.

For any given $i = 1,\dots,d(g)$, we will use the above lemma for the function

$$\bar z_i(t,a) = g_i(t,a,\bar y(t,a),\bar u(t,a),\bar v(t),\bar w(a)),$$

where from now on, $(\bar y,\bar q,\bar u,\bar v,\bar w)$ is a local minimizer in problem (1)–(6). According to the lemma, the support functional λ_i is concentrated on each set

$$M_{i\delta} = \{(t,a) \in D \,:\, g_i(t,a,\bar y(t,a),\bar u(t,a),\bar v(t),\bar w(a)) \ge -\delta\}, \qquad \delta > 0.$$

The Lagrange function in Theorem 4.1 reads for our problem as

$$\mathcal{L}(y,q,u,v,w) = \langle \psi, f(\cdot,\cdot,y(\cdot,\cdot),q(\cdot),u(\cdot,\cdot)) - \frac{\partial}{\partial e} y(\cdot,\cdot)\rangle$$
$$+\langle \nu_0, y_0(\cdot,w(\cdot) - y(0,\cdot))\rangle + \langle \nu_1, \varphi(\cdot,q(\cdot),v(\cdot)) - y(\cdot,0)\rangle$$
$$+\langle \mu, \int_0^\omega h(\cdot,a,y(\cdot,a))\,\mathrm{d}a - q(\cdot)\rangle$$

$$+\alpha \int_0^\omega l(a, y(T,a))\, \mathrm{d}a + \alpha \int_0^T \int_0^\omega L(t,a,y(t,a),q(t),u(t,a),v(t),w(a))\, \mathrm{d}a\, \mathrm{d}t$$

$$+\sum_{i=1}^{d(g)} \langle \lambda_i, g_i(\cdot,\cdot,y(\cdot,\cdot),u(\cdot,\cdot),v(\cdot),w(\cdot))\rangle,$$

where α is a number and all other multipliers are linear functionals from the dual spaces

$$\psi \in \Big(L^\infty(D, \mathbb{R}^{d(y)})\Big)^*, \quad \nu_0 \in \Big(L^\infty([0,\omega], \mathbb{R}^{d(y)})\Big)^*, \quad \nu_1 \in \Big(L^\infty([0,T], \mathbb{R}^{d(y)})\Big)^*,$$

$$\mu \in \Big(L^\infty([0,T], \mathbb{R}^{d(q)})\Big)^*, \quad \lambda_i \in Z^* = (L^\infty(D, \mathbb{R}))^*.$$

According to Theorem 4.1 and Lemma 4.3, there exists a tuple of multipliers $\alpha \geq 0, \psi, \mu, \nu_0, \nu_1, \lambda_i$ such that

$$\alpha + \sum_{i=1}^{d(g)} \|\lambda_i\| > 0,$$

$\lambda_i \geq 0$, λ_i is concentrated on each set $M_{i\delta}$, $\delta > 0$, $i = 1, \ldots, d(g)$, and

$$\mathcal{L}'(\bar{y}, \bar{q}, \bar{u}, \bar{v}, \bar{w}) = 0.$$

The latter means that for every $(y, q, u, v, w) \in X$

$$\langle \psi, f_y(\cdot,\cdot)y(\cdot,\cdot) + f_q(\cdot,\cdot)q(\cdot) + f_u(\cdot,\cdot)u(\cdot,\cdot) - \frac{\partial}{\partial e} y(\cdot,\cdot)\rangle$$

$$+\langle \nu_0, y_{0w}(\cdot)w(\cdot) - y(0,\cdot)\rangle + \langle \nu_1, \varphi_q(\cdot)q(\cdot) + \varphi_v(\cdot)v(\cdot) - y(\cdot,0)\rangle$$

$$+\langle \mu, \int_0^\omega h_y(\cdot,a)y(\cdot,a)\ \mathrm{d}a - q(\cdot)\rangle + \alpha \int_0^\omega l_y(a)y(T,a)\ \mathrm{d}a$$

$$+\alpha \int_0^T \int_0^\omega \Big(L_y(t,a)y(t,a) + L_q(t,a)q(t) + L_u(t,a)u(t,a) + L_v(t,a)v(t) + L_w(t,a)w(a)\Big)\, \mathrm{d}a\, \mathrm{d}t$$

$$+\sum_{i=1}^{d(g)} \langle \lambda_i, g_{iy}(\cdot,\cdot)y(\cdot,\cdot) + g_{iu}(\cdot,\cdot)u(\cdot,\cdot) + g_{iv}(\cdot,\cdot)v(\cdot) + g_{iw}(\cdot,\cdot)w(\cdot)\rangle = 0. \tag{26}$$

In the next subsections we will analyze in detail this equation.

4.4 Regularity of the functional ψ

First we will show that the functional $\psi \in \Big(L^\infty(D, \mathbb{R}^{d(y)})\Big)^*$ has a *regular integral representation* (or shortly, is *regular*). This means that it can be represented in the form

$$\langle \psi, z\rangle = \int_0^T \int_0^\omega \tilde{\psi}(t,a)z(t,a)\, \mathrm{d}a\, \mathrm{d}t, \quad \forall z \in L^\infty(D, \mathbb{R}^{d(y)}), \tag{27}$$

where $\tilde{\psi} \in L^1(D, \mathbb{R}^{d(y)})$.

Since $\langle \psi, z\rangle = \sum_{i=1}^{d(y)} \langle \psi_i, z_i\rangle$, it suffices to prove that the functional ψ_i is regular for an arbitrary fixed $i \in \{1, \ldots, d(y)\}$. To this end, we set in (26) $q = 0$, $u = 0$, $v = 0$, $w = 0$. Then we obtain from (26) the variational equation

$$-\langle \psi, \frac{\partial}{\partial e} y(\cdot,\cdot)\rangle + \langle \psi, f_y(\cdot,\cdot)y(\cdot,\cdot)\rangle - \langle \nu_0, y(0,\cdot)\rangle - \langle \nu_1, y(\cdot,0)\rangle$$

$$+\langle \mu, \int_0^\omega h_y(\cdot,a)y(\cdot,a)\,\mathrm{d}a\rangle + \alpha \int_0^\omega l_y(a)y(T,a)\,\mathrm{d}a$$

$$+\alpha \int_0^T \int_0^\omega L_y(t,a)y(t,a)\,\mathrm{d}a\,\mathrm{d}t + \sum_{i=1}^{d(g)} \langle \lambda_i, g_{iy}(\cdot,\cdot)y(\cdot,\cdot)\rangle = 0 \quad \forall\, y \in \mathcal{Y}_\Sigma. \tag{28}$$

By the Yosida–Hewitt theorem (see, e.g., [20] or [13, p. 382]), the functional ψ_i has a representation

$$\psi_i = \psi_i^r + \psi_i^s,$$

where ψ_i^r and ψ_i^s are the regular and the singular parts of the functional ψ_i, respectively. In its turn,

$$\psi_i^s = \psi_i^{s+} - \psi_i^{s-},$$

where ψ_i^{s+} and ψ_i^{s-} are the positive and the negative parts of the functional ψ_i^s, respectively. The latter means that the functionals ψ_i^{s+} and ψ_i^{s-} are nonnegative: $\psi_i^{s+} \geq 0$ and $\psi_i^{s-} \geq 0$, and $\|\psi_i^{s+}\| + \|\psi_i^{s-}\| = \|\psi_i^s\|$. Using equation (28), we will show that $\psi_i^{s+} = 0$.

Since ψ_i^{s+} is a singular functional, there exists a sequence of measurable sets $E_n \subset D$, $n = 1, 2, \ldots$, such that

$$\operatorname{meas} E_n \to 0 \ (n \to \infty) \tag{29}$$

and ψ_i^{s+} is concentrated on every E_n, i.e.

$$\langle \psi_i^{s+}, z\rangle = \langle \psi_i^{s+}, z\chi_{E_n}\rangle \quad \forall\, z \in L^\infty(D, \mathbb{R}), \quad n = 1, 2, \ldots,$$

where χ_{E_n} is the characteristic function of the set E_n. Thus

$$\langle \psi_i^{s+}, \chi_{E_n}\rangle = \langle \psi_i^{s+}, \chi_D\rangle = \|\psi_i^{s+}\|. \tag{30}$$

For every $n = 1, 2, \ldots$, define the function $y_n(t,a)$ as follows: its ith component $y_{ni}(t,a)$ satisfies the equation

$$\frac{\partial}{\partial e} y_{ni}(t,a) = \chi_{E_n}(t,a)$$

along almost every characteristic segment $S \in \Sigma$, with zero conditions on Γ:

$$y_{ni}(0,a) = 0, \quad y_{ni}(t,0) = 0.$$

All other components of $y_n(t,a)$ are taken identically equal to zero.

The convergence (29) implies

$$\|y_n\|_\infty \to 0 \quad (n \to \infty). \tag{31}$$

Consider the first term of equation (28) for $y = y_n$. We obviously have

$$\langle \psi, \frac{\partial}{\partial e} y_n \rangle = \langle \psi_i, \frac{\partial}{\partial e} y_{ni} \rangle = \langle \psi_i, \chi_{E_n} \rangle = \langle \psi_i^r, \chi_{E_n} \rangle + \langle \psi_i^{s+}, \chi_{E_n} \rangle - \langle \psi_i^{s-}, \chi_{E_n} \rangle.$$

Here $\langle \psi_i^r, \chi_{E_n} \rangle \to 0$, $(n \to \infty)$, since the functional ψ_i^r is regular and condition (29) holds for the sequence E_n. Moreover, $\langle \psi_i^{s-}, \chi_{E_n} \rangle = 0$ since the functionals ψ_i^{s+} and ψ_i^{s-} are mutually singular. Then, taking into account relations (30), we obtain that

$$\langle \psi, \frac{\partial}{\partial e} y_n(\cdot, \cdot) \rangle \to \|\psi_i^{s+}\|. \tag{32}$$

In view of condition (31), all other terms in (28) for $y = y_n$ tend to zero. Thus, substituting $y_n(t, a)$ in equation (28) and passing to the limit for $n \to \infty$, we obtain that $\|\psi_i^{s+}\| = 0$, and hence $\psi_i^{s+} = 0$. Similarly, we prove that $\psi_i^{s-} = 0$. Then $\psi^s = 0$. The latter means that $\psi = \psi^r$, i.e. the functional ψ is regular. Consequently, condition (27) holds with some function $\tilde{\psi} \in L^1(D, \mathbb{R}^{d(y)})$, representing the functional ψ^r. In what follows, we omit the wave in the notation $\tilde{\psi}(t, a)$.

4.5 Stationarity with respect to u and regularity of λ_i

Set in (26) $y = 0$, $q = 0$, $v = 0$, $w = 0$. Then, taking into account the regularity of the functional ψ, proved in the preceding subsection (see (27)), we obtain that for every $u \in L^\infty(D, \mathbb{R}^{d(u)})$

$$\int_0^T \int_0^\omega \Big(\psi(t, a) f_u(t, a) + \alpha L_u(t, a) \Big) u(t, a) \,\mathrm{d}a \,\mathrm{d}t + \sum_{i=1}^{d(g)} \langle \lambda_i, g_{iu}(\cdot, \cdot) u(\cdot, \cdot) \rangle = 0. \tag{33}$$

Now we use the regularity assumption (Assumption A) for the mixed constraints. According to Theorem 7.1 in [5], equation (33) implies that all functionals λ_i have regular integral representation, i.e. there exist integrable functions $\tilde{\lambda}_i(t, a)$ such that

$$\langle \lambda_i, z \rangle = \int_0^T \int_0^\omega \tilde{\lambda}_i(t, a) z(t, a) \,\mathrm{d}a \,\mathrm{d}t \qquad \forall\, z \in L^\infty(D, \mathbb{R}). \tag{34}$$

Moreover, from the properties of λ_i stated in Theorem 4.1 (λ_i corresponds to the multiplier ζ_i^* in that theorem), it follows that the functions $\tilde{\lambda}_i$ are nonnegative and satisfy the complementary slackness conditions

$$\tilde{\lambda}_i(t, a) \geq 0, \quad \tilde{\lambda}_i(t, a)\, g_i(t, a) = 0, \quad i = 1, \ldots, d(g). \tag{35}$$

In view of (34), relation (33) implies

$$\psi(t, a) f_u(t, a) + \alpha L_u(t, a) + \sum_{i=1}^{d(g)} \tilde{\lambda}_i(t, a) g_{iu}(t, a) = 0. \tag{36}$$

This is a stationarity condition with respect to the control u (see (11)). It is fulfilled for a.e. $(t, a) \in D$. In the sequel, we omit the "tilde" in the notation $\tilde{\lambda}_i$.

4.6 Regularity of ν_0 and ν_1

Take any $n \in L^\infty([0,\omega]$ and $r \in L^\infty([0,T],\mathbb{R}^{d(y)})$. According to Lemma 4.2 the system

$$\begin{aligned}
\frac{\partial}{\partial e} y(t,a) &= f_y(t,a)y(t,a) + f_q(t,a)q(t), \\
y(0,a) &= -n(a), \quad y(t,0) - \varphi_q(t)q(t) = -r(t), \\
q(t) &= \int_0^\omega h_y(t,a)y(t,a)\,\mathrm{d}a
\end{aligned}$$

has a unique solution $(y,q) \in \mathcal{Y}_\Sigma \times L^\infty([0,T],\mathbb{R}^{d(q)})$. Set $u(t,a) = 0$, $v(t) = 0$, and $w(a) = 0$. Then equation (26) becomes

$$\langle \nu_0, n(\cdot)\rangle + \langle \nu_1, r(\cdot)\rangle + \alpha \int_0^\omega l_y(a)y(T,a)\,\mathrm{d}a$$

$$+ \int_0^T \int_0^\omega \Big(\alpha L_y(t,a)y(t,a) + \alpha L_q(t,a)q(t) + \sum_{i=1}^{d(g)} \lambda_i(t,a)g_{iy}(t,a)y(t,a)\Big)\,\mathrm{d}a\,\mathrm{d}t = 0. \tag{37}$$

Since n and r can be selected arbitrarily and independently of each other and y and q have integral representations via n and r (see Remark 4.1) we obtain from this equation that the functionals ν_0 and ν_1 are regular, i.e. there exist integrable functions $\tilde{\nu}_0$ and $\tilde{\nu}_1$ such that

$$\langle \nu_0, n(\cdot)\rangle = \int_0^\omega \tilde{\nu}_0(a)n(a)\,\mathrm{d}t \quad \forall n, \quad \langle \nu_1, r(\cdot)\rangle = \int_0^\omega \tilde{\nu}_1(t)r(t)\,\mathrm{d}t \quad \forall r. \tag{38}$$

In the sequel we omit the "tilde" in the notations $\tilde{\nu}_0$ and $\tilde{\nu}_1$.

4.7 Stationarity with respect to v

Set in (26) $y = 0$, $q = 0$, $u = 0$, $w = 0$. Then, taking into account the regularity of the functionals λ_i and ν_1 (see (34),(38)), we obtain that for every $v \in L^\infty([0,T],\mathbb{R}^{d(v)})$

$$\int_0^\omega \nu_1(t)\varphi_v(t)v(t)\,\mathrm{d}t$$

$$+ \int_0^T \int_0^\omega \alpha L_v(t,a)v(t)\,\mathrm{d}a\,\mathrm{d}t + \sum_{i=1}^{d(g)} \int_0^T \int_0^\omega \lambda_i(t,a)g_{iv}(t,a)v(t)\,\mathrm{d}a\,\mathrm{d}t = 0.$$

This equation can be rewritten as

$$\int_0^\omega \nu_1(t)\varphi_v(t)v(t)\,\mathrm{d}t + \int_0^T \Big[\int_0^\omega \Big(\alpha L_v(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a)g_{iv}(t,a)\Big)\,\mathrm{d}a\Big]v(t)\,\mathrm{d}t = 0$$

for all v. Hence, we obtain

$$\nu_1(t)\varphi_v(t) + \int_0^\omega \Big(\alpha L_v(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a)g_{iv}(t,a)\Big)\,\mathrm{d}a = 0. \tag{39}$$

4.8 Stationarity with respect to w

Set in (26) $y = 0$, $q = 0$, $u = 0$, $v = 0$. Then, quite similarly to the derivation of equation (39), we obtain

$$\nu_0(a) y_{0w}(a) + \int_0^T \Big(\alpha L_w(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a) g_{iw}(t,a)\Big)\,\mathrm{d}t = 0. \tag{40}$$

4.9 Stationarity with respect to q and regularity of μ

Setting in (26) $y = 0$, $u = 0$, $v = 0$, $w = 0$ and taking into account the regularity of the functionals ψ and ν_1, we obtain that for every $q \in L^\infty([0,T], \mathbb{R}^{d(q)})$

$$\int_0^T \int_0^\omega \psi(t,a) f_q(t,a) q(t)\,\mathrm{d}a\,\mathrm{d}t + \int_0^T \nu_1(t)\varphi_q(t) q(t)\,\mathrm{d}t - \langle \mu, q(\cdot)\rangle$$

$$+ \int_0^T \int_0^\omega \alpha L_q(t,a) q(t)\,\mathrm{d}a\,\mathrm{d}t = 0.$$

That is,

$$\langle \mu, q\rangle = \int_0^T \Big(\int_0^\omega (\psi(t,a) f_q(t,a) + \alpha L_q(t,a))\,\mathrm{d}a + \nu_1(t)\varphi_q(t)\Big)\, q(t)\,\mathrm{d}t$$

for all q. This implies, in particular, that μ is a regular functional with the representation

$$\mu(t) := \int_0^\omega \Big(\psi(t,a) f_q(t,a) + \alpha L_q(t,a)\Big)\,\mathrm{d}a + \nu_1(t)\varphi_q(t), \tag{41}$$

so that

$$\langle \mu, q\rangle = \int_0^T \mu(t) q(t)\,\mathrm{d}t \quad \forall\, q \in L^\infty([0,T], \mathbb{R}^{d(q)}). \tag{42}$$

4.10 Stationarity with respect to y

Now, we set in (26) $q = 0$, $u = 0$, $v = 0$, $w = 0$, taking into account the functionals ψ, λ_i, ν_0, ν_1, and μ have regular integral representations (see (27), (34), (38), and (42)). We obtain that

$$\int_0^T \int_0^\omega \psi(t,a)\Big(f_y(t,a) y(t,a) - \frac{\partial}{\partial e} y(t,a)\Big)\,\mathrm{d}a\,\mathrm{d}t$$

$$- \int_0^\omega \nu_0(a) y(0,a)\,\mathrm{d}a - \int_0^T \nu_1(t) y(t,0)\,\mathrm{d}t + \int_0^\omega \alpha l_y(a) y(T,a)\,\mathrm{d}a$$

$$+ \int_0^T \int_0^\omega \Big(\mu(t) h_y(t,a) + \alpha L_y(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a) g_{iy}(t,a)\Big) y(t,a)\,\mathrm{d}a\,\mathrm{d}t = 0 \tag{43}$$

for every $y \in \mathcal{Y}_\Sigma$.

Now, our aim is to prove that the function $\psi(t,a)$ is absolutely continuous along almost every characteristic, and also to obtain the adjoint equation for ψ and the transversality conditions in Theorem 3.1.

Let $p(t,a)$ be a function which is absolutely continuous along almost every characteristic S_γ and satisfies, along almost every S_γ, the equation

$$-\frac{\partial}{\partial e}p(t,a) = p(t,a)f_y(t,a) + \mu(t)h_y(t,a) + \alpha L_y(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a)g_{iy}(t,a) \tag{44}$$

and also satisfies the following transversality conditions

$$p(T,a) = \alpha l_y(a), \quad p(t,\omega) = 0. \tag{45}$$

Recall that S_γ is the characteristic segment starting at the point $\gamma = (t_\gamma, a_\gamma) \in \Gamma$, where as before Γ is the union of the bottom and left segments of the boundary of D. Let τ_γ be the terminal value of the parameter τ such that $\sigma_\gamma := \gamma + \tau_\gamma e$ is the end of S_γ, opposite to γ. That is, $\sigma_\gamma \in \tilde{\Gamma} := \{(T,a) : a \in [0,\omega]\} \cup \{(t,\omega) : t \in [0,T]\}$ and

$$S_\gamma = [\gamma, \gamma + \tau_\gamma e] = [\gamma, \sigma_\gamma].$$

Then equation (44) means that along almost every characteristic S_γ we have

$$\begin{aligned} -\frac{d}{d\tau}p(\gamma+\tau e) \quad &= \quad p(\gamma+\tau e)f_y(\gamma+\tau e) + \mu(t_\gamma+\tau)h_y(\gamma+\tau e) \\ &\qquad + \alpha L_y(\gamma+\tau e) + \sum_{i=1}^{d(g)} \lambda_i(\gamma+\tau e)g_{iy}(\gamma+\tau e). \end{aligned}$$

In particular, this implies existence and uniqueness of the function p, since on almost every characteristic S_γ it is defined by a linear ODE with the end-point condition (45) at $\tau = \tau_\gamma$, and all the functions (not counting p) in the right-hand side are integrable, while f_y is measurable and bounded.

Let $z \in L^\infty(D, \mathbb{R}^{d(y)})$, $y^0 \in L^\infty([0,\omega], \mathbb{R}^{d(y)})$ and $y^1 \in L^\infty([0,T], \mathbb{R}^{d(y)})$ be arbitrarily chosen. For $\gamma \in \Gamma$ define

$$y^\Gamma(\gamma) = \begin{cases} y^0(a) & \text{if } \gamma = (0,a), \;\; a \in [0,\omega], \\ y^1(t) & \text{if } \gamma = (t,0), \;\; t \in [0,T]. \end{cases}$$

Let $y \in \mathcal{Y}_\Sigma$ be defined by the family (parameterized by $\gamma \in \Gamma$) of solutions along almost every characteristic S_γ of the linear ODEs

$$\frac{d}{d\tau}y(\gamma+\tau e) = f_y(\gamma+\tau e)y(\gamma+\tau e) + z(\gamma+\tau e)$$

with the initial condition $y(\gamma) = y^\Gamma(\gamma)$ at $\tau = 0$. In other words, y satisfies almost everywhere the equation

$$\frac{\partial}{\partial e}y(t,a) = f_y(t,a)y(t,a) + z(t,a). \tag{46}$$

Along almost every characteristic S_γ we have

$$p(\sigma_\gamma)\,y(\sigma_\gamma) - p(\gamma)\,y(\gamma) = \int_0^{\tau_\gamma} \frac{\mathrm{d}}{\mathrm{d}\tau}\Big(p(\gamma+\tau e)y(\gamma+\tau e)\Big)\,\mathrm{d}\tau$$

$$= \int_0^{\tau_\gamma} \Big(\frac{d}{d\tau}p(\gamma+\tau e)\Big)y(\gamma+\tau e)\,\mathrm{d}\tau + \int_0^{\tau_\gamma} p(\gamma+\tau e)\frac{\mathrm{d}}{\mathrm{d}\tau}y(\gamma+\tau e)\,\mathrm{d}\tau$$

$$= \int_0^{\tau_\gamma} \Big(-p(\gamma+\tau e)f_y(\gamma+\tau e) - \mu(t_\gamma+\tau)h_y(\gamma+\tau e)$$

$$-\alpha L_y(\gamma+\tau e) - \sum_{i=1}^{d(g)} \lambda_i(\gamma+\tau e)g_{iy}(\gamma+\tau e)\Big)y(\gamma+\tau e)\,\mathrm{d}\tau$$

$$+ \int_0^{\tau_\gamma} p(\gamma+\tau e)\Big(f_y(\gamma+\tau e)y(\gamma+\tau e) + z(\gamma+\tau e)\Big)\,\mathrm{d}\tau$$

$$= \int_0^{\tau_\gamma} \Big(-\mu(t_\gamma+\tau)h_y(\gamma+\tau e) - \alpha L_y(\gamma+\tau e)$$

$$-\sum_{i=1}^{d(g)} \lambda_i(\gamma+\tau e)g_{iy}(\gamma+\tau e)\Big)y(\gamma+\tau e)\,\mathrm{d}\tau + \int_0^{\tau_\gamma} p(\gamma+\tau e)z(\gamma+\tau e)\,\mathrm{d}\tau.$$

Integrating along Γ, we obtain that

$$\int_{\gamma\in\Gamma} p(\sigma_\gamma)y(\sigma_\gamma)\,\mathrm{d}\gamma - \int_{\gamma\in\Gamma} p(\gamma)y(\gamma)\,\mathrm{d}\gamma = \int_{\gamma\in\Gamma}\int_0^{\tau_\gamma} \Big(-\mu(t_\gamma+\tau)h_y(\gamma+\tau e) - \alpha L_y(\gamma+\tau e)$$

$$-\sum_{i=1}^{d(g)} \lambda_i(\gamma+\tau e)g_{iy}(\gamma+\tau e)\Big)y(\gamma+\tau e)\,\mathrm{d}\tau\,\mathrm{d}\gamma + \int_{\gamma\in\Gamma}\int_0^{\tau_\gamma} p(\gamma+\tau e)z(\gamma+\tau e)\,\mathrm{d}\tau\,\mathrm{d}\gamma. \tag{47}$$

The left-hand side of the last equality can be represented as

$$\int_{\gamma\in\Gamma} p(\sigma_\gamma)y(\sigma_\gamma)\,\mathrm{d}\gamma - \int_{\gamma\in\Gamma} p(\gamma)y(\gamma)\,\mathrm{d}\gamma = \int_{\gamma\in\widetilde{\Gamma}} p(\gamma)y(\gamma)\,\mathrm{d}\gamma - \int_{\gamma\in\Gamma} p(\gamma)y(\gamma)\,\mathrm{d}\gamma$$

$$= \int_0^T p(t,\omega)y(t,\omega)\,\mathrm{d}t + \int_0^\omega p(T,a)y(T,a)\,\mathrm{d}a - \int_0^T p(t,0)y(t,0)\,\mathrm{d}t - \int_0^\omega p(0,a)y(0,a)\,\mathrm{d}a.$$

Substituting this expression in (47), noticing that $\int_{\gamma\in\Gamma}\int_0^{\sigma_\gamma} = \int_0^T\int_0^\omega$ and rearranging the terms we obtain that

$$\int_0^T\int_0^\omega \Big(\mu(t)h_y(t,a) + \alpha L_y(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a)g_{iy}(t,a)\Big)y(t,a)\,\mathrm{d}a\,\mathrm{d}t$$

$$= \int_0^T \int_0^\omega p(t,a) z(t,a)\,\mathrm{d}a\,\mathrm{d}t - \int_0^T p(t,\omega) y(t,\omega)\,\mathrm{d}t - \int_0^\omega p(T,a) y(T,a)\,\mathrm{d}a$$

$$+ \int_0^T p(t,0) y(t,0)\,\mathrm{d}t + \int_0^\omega p(0,a) y(0,a)\,\mathrm{d}a. \tag{48}$$

Using relation (48) in (43) and taking into account (46), (45) and the initial condition $y(\gamma) = y^\Gamma$, we obtain that

$$\int_0^T \int_0^\omega \Big(p(t,a) - \psi(t,a)\Big) z(t,a)\,\mathrm{d}a\,\mathrm{d}t$$
$$+ \int_0^\omega (p(0,a) - \nu_0(a)) y_0(a)\,\mathrm{d}a + \int_0^T (p(t,0) - \nu_1(t)) y_1(t)\,\mathrm{d}t = 0.$$

Since the functions z, y^0 and y^1 can be selected arbitrarily and independently of each other, this variational equation implies that

$$p(t,a) = \psi(t,a), \quad p(0,a) = \nu_0(a), \quad p(t,0) = \nu_1(t). \tag{49}$$

Consequently, the function ψ is absolutely continuous along almost every characteristic, satisfies the adjoint equation (7), and the transversality conditions (9). Moreover, due to (49) we have

$$\nu_0(a) = \psi(0,a), \quad \nu_1(t) = \psi(t,0). \tag{50}$$

Using these equalities in (39), (40), and (41), we obtain (12), (13), and (8), respectively.

4.11 Boundedness of the Lagrange multipliers

In order to prove boundedness of the Lagrange multiplier ψ we need some auxiliary material presented below.

Lemma 4.4 *Let $\mathcal{K} \subset \mathbb{R}^m$ be a compact set and let $a_i : \mathcal{K} \to \mathbb{R}^s$ be continuous functions, $i = 1, \dots, k$. Let $I(z) \subset \{1, \dots, k\}$ be a set of indices, defined at each point $z \in \mathcal{K}$. Assume that the mapping $I(z)$ has the following property: if $z_n \to z$ $(n \to \infty)$ and $i \in I(z_n)$ $\forall n$, then $i \in I(z)$ (that is, the mapping $z \to I(z)$ is upper semicontinuous on $\mathcal{K}$). Assume, in addition, that for any $z \in \mathcal{K}$ the system of vectors $a_i(z)$, $i \in I(z)$ is positively linearly independent. Then there exists a constant $C > 0$ such that for every $z \in \mathcal{K}$ and any numbers $\alpha_i \geq 0$, $i \in I(z)$, the following inequality holds:*

$$\sum_{i \in I(z)} \alpha_i \leq C \left| \sum_{i \in I(z)} \alpha_i a_i(z) \right|. \tag{51}$$

Proof. Contrary to the statement of the lemma, assume that there exist a sequence $\{z_n\}$ in $\mathcal{K}$ and sequences of numbers $\alpha_{in} \geq 0$, $i \in I(z_n)$, such that

$$\sum_{i \in I(z_n)} \alpha_{in} > n \left| \sum_{i \in I(z_n)} \alpha_{in} a_i(z_n) \right| \quad \forall\, n. \tag{52}$$

Without loss of generality (taking if necessary a subsequence) we assume that $I(z_n)$ does not depend on n, i.e. $I(z_n) = I \ \forall\, n$, and $z_n \to z$ $(n \to \infty)$. Note that $z \in \mathcal{K}$ and $I \subset I(z)$. Set

$$\beta_{in} = \frac{\alpha_{in}}{\sum_{i \in I} \alpha_{in}}, \quad i \in I. \tag{53}$$

Then

$$\beta_{in} \geq 0, \quad i \in I, \quad \sum_{i \in I} \beta_{in} = 1, \quad \forall\, n, \tag{54}$$

and from (52)-(54) it follows that

$$\left| \sum_{i \in I} \beta_{in} a_i(z_n) \right| < \frac{1}{n} \quad \forall\, n. \tag{55}$$

Again, without loss of generality we assume that $\beta_{in} \to \beta_i$, $i \in I$ $(n \to \infty)$. Then passing to the limit in (54) and (55) we obtain

$$\beta_i \geq 0, \quad i \in I, \quad \sum_{i \in I} \beta_i = 1, \quad \sum_{i \in I} \beta_i a_i(z) = 0.$$

Since $I \subset I(z)$, the latter contradicts to the positive linear independence of the system of vectors $a_i(z)$, $i \in I(z)$ at the point $z \in \mathcal{K}$. $\square$

Note that estimates, similar to (51), were obtained earlier in [16, p.107] and [5, Lemma 5].

In what follows, we set for brevity

$$x = (y, u, v, w), \quad z = (t, a, x) = (t, a, y, u, v, w).$$

Suppose that there exists a compact set $\mathcal{K} \subset \mathbb{R}^{d(z)}$ such that

$$g(z) \leq 0 \quad \forall\, z \in \mathcal{K}. \tag{56}$$

Set

$$I(z) = \{ i \in \{1, \ldots, d(g)\} : g_i(z) = 0 \}.$$

Then, obviously, $I(z)$ satisfies the condition of Lemma 4.4: if $z_n \in \mathcal{K} \ \forall\, n$, $z_n \to z$ $(n \to \infty)$ and $i \in I(z_n) \ \forall\, n$, then $i \in I(z)$. Further, note that, by Assumption A, for any $z \in \mathcal{K}$ the gradients

$$g_{iu}(z), \quad i \in I(z)$$

are positively linearly independent. Hence Lemma 4.4 implies the following corollary.

Corollary 4.1 *Let a compact set $\mathcal{K} \subset \mathbb{R}^{d(z)}$ satisfies condition (56). Then there exists a constant $C > 0$ such that, for any $z \in \mathbb{R}^{d(z)}$ and any numbers $\lambda_i \geq 0$, $i \in I(z)$ we have*

$$\sum_{i \in I(z)} \lambda_i \leq C \left| \sum_{i \in I(z)} \lambda_i g_{iu}(z) \right|. \tag{57}$$

Now, consider the measurable and essentially bounded function

$$\bar{x}(t,a) : Q \to \mathbb{R}^{d(x)}, \quad \bar{x}(t,a) = (\bar{y}(t,a), \bar{u}(t,a), \bar{v}(t), \bar{w}(a)).$$

Let $\mathcal{K} \subset \mathbb{R}^{d(z)}$ be a compact set satisfying condition (56) and such that

$$(t, a, \bar{x}(t,a)) \in \mathcal{K} \quad \text{for a.e.} \quad (t,a) \in Q. \tag{58}$$

For example, the role of $\mathcal{K}$ can be assigned to the set called the *closure of the function* $\bar{x}$ *w.r.t. the measure* [6]. Its definition is as follows. Suppose that $\mathcal{K}$ is such a set. Then, $\mathcal{K}$ consists of all points $(t', a', x') \in D \times \mathbb{R}^{d(x)}$ such that for any $\varepsilon > 0$,

$$\operatorname{meas}\{(t,a) \in D : |t - t'| < \varepsilon, \ |a - a'| < \varepsilon, \ |\bar{x}(t,a) - x'| < \varepsilon\} > 0.$$

Then the set $\mathcal{K}$ is compact and satisfies conditions (56) and (58).

Continuing with the auxiliary material, let us rewrite equation (11) in the form

$$\sum_{i=1}^{d(g)} \lambda_i(t,a) g_{iu}(t,a) = -\psi(t,a) f_u(t,a) - \alpha L_u(t,a), \quad \text{for a.e. } (t,a) \in D. \tag{59}$$

Then conditions (57)–(59) imply

$$\sum_{i \in I(t,a,\bar{x}(t,a))} \lambda_i(t,a) \le C|\psi(t,a) f_u(t,a) + \alpha L_u(t,a)| \quad \text{for a.e. } (t,a) \in D.$$

In view of the complementarity conditions,

$$\sum_{i \in I(t,a,\bar{x}(t,a))} \lambda_i(t,a) = \sum_{i=1}^{d(g)} \lambda_i(t,a).$$

Consequently, the following estimate holds:

$$\sum_{i=1}^{d(g)} \lambda_i(t,a) \le C|\psi(t,a) f_u(t,a) + \alpha L_u(t,a)| \quad \text{for a.e. } (t,a) \in D. \tag{60}$$

Let us define the function

$$\varphi(t,a) = \alpha L_y(t,a) + \sum_{i=1}^{d(g)} \lambda_i(t,a) g_{iy}(t,a). \tag{61}$$

In view of (60) and since $\alpha \ge 0$ and $\lambda_i(t,a) \ge 0$, we get from (61)

$$\begin{aligned}
|\varphi(t,a)| &\le \alpha \|L_y\|_\infty + \max_i \|g_{iy}\|_\infty \cdot \sum_{i=1}^{d(g)} \lambda_i(t,a) \\
&\le \alpha \|L_y(\cdot,\cdot)\|_\infty + \max_i \|g_{iy}\|_\infty \cdot C|\psi(t,a) f_u(t,a) + \alpha L_u(t,a)| \\
&\le \alpha\left(\|L_y\|_\infty + C_1 \|L_u\|_\infty\right) + C_1 \cdot \|f_u\|_\infty |\psi(t,a)| \quad \text{for a.e. } (t,a) \in D,
\end{aligned}$$

where $C_1 := C \max_i \|g_{iy}\|_\infty$. Thus we obtain that there exist constants $c > 0$ and $d > 0$ such that the function $\varphi \in L^1(D, \mathbb{R}^{d(y)})$, as in (61), and the adjoint function $\psi \in L^1(D, \mathbb{R}^{d(y)})$ satisfy the inequality

$$|\varphi(t,a)| \le c|\psi(t,a)| + d \quad \text{for a.e. } (t,a) \in D. \tag{62}$$

We shall use the following lemma.

LEMMA 4.5 *For $c \geq 0$ and $d \geq 0$, denote*

$$\Omega := \{(\varphi, \psi) \in \mathbb{R}^{d(y)} \times \mathbb{R}^{d(y)} : |\varphi| \leq c|\psi| + d\}.$$

Then there exist a matrix function $M : \Omega \to \mathbb{R}^{d(y)\cdot d(y)}$ and a vector function $\xi : \Omega \to \mathbb{R}^{d(y)}$, both Borel measurable, such that for every $(\varphi, \psi) \in \Omega$ it holds that $\|M(\varphi, \psi)\| \leq c$, $|\xi(\varphi, \psi)| \leq d$, and

$$\varphi = M(\varphi, \psi)\psi + \xi(\varphi, \psi). \tag{63}$$

Proof. Define the subsets

$$\begin{aligned}
\Omega_2 &:= \{(\varphi, \psi) \in \Omega : \operatorname{rank}\{\varphi, \psi\} = 2\}, \\
\Omega_1 &:= \{(\varphi, \psi) \in \Omega : \operatorname{rank}\{\varphi, \psi\} = 1,\ \varphi \neq 0, \psi \neq 0\}, \\
\Omega_0 &:= \{(\varphi, \psi) \in \Omega : \varphi = 0 \ \text{ or } \ \psi = 0\}.
\end{aligned}$$

For any $(\varphi, \psi) \in \Omega_2 \cup \Omega_1$ denote by $P(\varphi, \psi)$ the projection operator of $\mathbb{R}^{d(y)}$ on $\operatorname{Lin}\{\varphi, \psi\}$. Clearly, the restriction of the mapping $(\varphi, \psi) \mapsto P(\varphi, \psi)$ to any of the sets Ω_2 and Ω_1 is continuous.

For $(\varphi, \psi) \in \Omega_2$ define $R(\varphi, \psi)$ as the clockwise rotation defined in $\operatorname{Lin}\{\varphi, \psi\}$ that sends ψ to a vector positively proportional to φ. For $(\varphi, \psi) \in \Omega_1$ we have $\psi = \beta\varphi$ for some number $\beta = \beta(\varphi, \psi) \neq 0$. In this case we define $R(\varphi, \psi)$ on $\operatorname{Lin}\{\varphi, \psi\} = \operatorname{Lin}\{\varphi\}$ by $R(\varphi, \psi)\eta = \operatorname{sign}(\beta(\varphi, \psi))\eta$. Then in both cases $R(\varphi, \psi)P(\varphi, \psi)$ is a linear mapping $\mathbb{R}^{d(y)} \to \mathbb{R}^{d(y)}$ with the norm equal to one. Denote by $\tilde{M}(\varphi, \psi)$ the matrix of this mapping in the canonical base. Clearly the restriction of $\tilde{M}(\varphi, \psi)$ on each of the sets Ω_2 and Ω_1 is a continuous matrix function.

From the definition of $\tilde{M}(\varphi, \psi)$ we have the identity

$$\tilde{M}(\varphi, \psi)\psi = \frac{\varphi}{|\varphi|}\,|\psi|, \quad (\varphi, \psi) \in \Omega_2 \cup \Omega_1.$$

Then for $(\varphi, \psi) \in \Omega_2 \cup \Omega_1$ we obtain the representation

$$\begin{aligned}
\varphi &= \frac{c|\psi|}{c|\psi| + d}\,\varphi + \frac{d}{c|\psi| + d}\,\varphi = \frac{c|\varphi|}{c|\psi| + d}\tilde{M}(\varphi, \psi)\psi + \frac{d}{c|\psi| + d}\,\varphi \\
&= M(\varphi, \psi)\psi + \xi(\varphi, \psi),
\end{aligned}$$

where

$$M(\varphi, \psi) = \frac{c|\varphi|}{c|\psi| + d}\tilde{M}(\varphi, \psi), \quad \xi(\varphi, \psi) = \frac{d}{c|\psi| + d}\,\varphi.$$

Obviously $\|M\| \leq c\|\tilde{M}\| = c$ and $|\xi| \leq d$ on the set $\Omega_2 \cup \Omega_1$. Moreover, the restrictions of M and ξ on Ω_2 and on Ω_1 are continuous.

We extend the definitions of M and ξ to Ω_0 by setting $M(\varphi, \psi) = 0$ and $\xi(\varphi, \psi) = \varphi$ for $(\varphi, \psi) \in \Omega_0$. Clearly, the restrictions of M and ξ on Ω_0 are also continuous. The identity (63) is evidently fulfilled. Moreover, if $\varphi \neq 0$, then $\psi = 0$ and $|\xi| = |\varphi| \leq d$ is fulfilled.

Finally, we notice that the set Ω_2 is open in Ω, $\Omega_1 \cup \Omega_0$ is closed, and Ω_0 is closed. Since the restrictions of M and ξ on the sets Ω_2, Ω_1 and Ω_0 are continuous, we have the claimed Borel measurability. □

Now we are ready to prove the boundedness of the Lagrange multiplier ψ. Observe that due to (62) we have $(\varphi(t,a),\psi(t,a)) \in \Omega$ for a.e. $(t,a) \in D$. According to Lemma 4.5, the following identity holds:

$$\varphi(t,a) = \bar{M}(t,a)\psi(t,a) + \bar{\xi}(t,a), \tag{64}$$

where $\bar{M}(t,a) := M(\varphi(t,a),\xi(t,a))$ and $\bar{\xi}(t,a) := \xi(\varphi(t,a),\psi(t,a))$. Notice that both $\bar{M}$ and $\bar{\xi}$ are measurable (due to the Borel measurability of M and ξ in Lemma 4.5) and are bounded.

We have obtained that ψ satisfies (7), (9) and (10). Inserting (64) in (7), the latter takes the form

$$-\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\psi(t,a) = \psi(t,a)(f_y(t,a) + \bar{M}^\top(t,a)) + \mu(t)h_y(t,a) + \bar{\xi}(t,a), \tag{65}$$

where $\bar{M}$ is transposed, since in Lemma 4.5 the vector ψ was regarded as a column, while ψ is considered as a row in the adjoint equation. Since all functions involved as data in (65) and in (9) and (10) are bounded, Lemma 4.2 implies that ψ and μ are essentially bounded. From (65) we also obtain that ψ is Lipschitz continuous along almost every characteristic and the norm

$$\operatorname*{esssup}_{\gamma\in\Gamma} \| \psi_{|S_\gamma} \|_{W^{1,\infty}(S_\gamma,\mathbb{R}^{d(y)})}$$

is finite.

The boundedness of ψ, nonnegativeness of λ_i, and estimate (60) imply the boundedness of λ_i. Further, conditions (50) imply that ν_0 and ν_1 are essentially bounded. Finally note that, without loss of generality, we can take $\alpha \in \{0,1\}$. This completes the proof of Theorem 3.1.

References

[1] S. Anita. *Analysis and control of age-depended population dynamics.* Kluwer Academic Publishers, 2000.

[2] E. Barucci and F. Gozzi. Investment in a vintage capital model. *Research in Economics* **52**:159–188, 1998.

[3] R. Boucekkine, N. Hritonenko, Y. Yatsenko (Eds.). *Optimal control of age-structured populations in economy, demography, and the environment.* Routledge, 2013.

[4] M. Brokate. Pontryagin's principle for control problems in age-dependent population dynamics. *J. Math. Biology*, **23**:75–101, 1985.

[5] A. V. Dmitruk and N. P. Osmolovskii. Necessary conditions for a weak minimum in optimal control problems with integral equations subject to state and mixed constraints, *SIAM J. on Control and Optimization*, **52**: 3437–3462, 2014.

[6] A.Ya. Dubobitskii, A.A. Milyutin. Necessary conditions for a weak extremum in general optimal control problem [in Russian], *Zh. Vychisl. Mat. Mat. Fiz.*, **8**(4):725–779, 1968.

[7] S. Faggian. Hamilton-Jacobi equations arising from boundary control problems with state constraints. *SIAM J. Control Optim.*, **47**(4):2157–2178, 2008.

[8] G. Feichtinger, R.F. Hartl, P.M. Kort, and V.M. Veliov. Anticipation effects of technological progress on capital accumulation: a vintage capital approach. *J. Econom. Theory*, **126**:143–164, 2006.

[9] G. Feichtinger, R.F. Hartl, P.M. Kort, and V.M. Veliov. Financially constrained capital investments: the effects of disembodied and embodied technological progress. *Journal of Mathematical Economics*, **44**:459–483, 2008.

[10] G. Fiechtinger, G. Tragler, and V. Veliov. Optimality conditions for age-structured control systems, *J. Math. Anal. Appl.*, **288**:47–68, 2003.

[11] G. Gripenberg, S.O. Londen, and O. Staffans. *Volterra integral and functional equations.* Cambridge Univ. Press, 1990.

[12] M. Iannelli. *Mathematical theory of age-structured population dynamics*, Giardini Editori, Pisa, 1995.

[13] L.V. Kantorovich and G.P. Akilov. *Funkcionalny Analiz (Functional Analysis)* [in Russian]. Nauka, Moscow, 1984.

[14] M.I. Krastanov, N.K. Ribarska, Ts.Y. Tsachev. Pontryagin maximum principle for infinite-dimensional problems. *SIAM J. Control Optim.*, **49**(5):2155–2182, 2011.

[15] M. Kuhn, S. Wrzaczek, A. Prskawetz, G. Feichtinger. Optimal choice of health and retirement in a life-cycle model. *Journal of Economic Theory*, **158**:186–212, 2015.

[16] A.A. Milyutin, A.V. Dmitruk, and N.P. Osmolovskii. *Maximum principle in optimal control* (in Russian). Moscow State University, Faculty of Mechanics and Mathematics, Moscow, 2004.

[17] V.M. Veliov. Optimal Control of Heterogeneous Systems: Basic Theory. *J. Math. Anal. Appl.*, **346**:227–242, 2008.

[18] G.F. Webb. *Theory of Nonlinear Age-dependent Population Dynamics.* Marcel Dekker, 1985.

[19] M.L. Weitzman. *Income, Wealth, and the maximum principle.* Harvard University Press, 2003.

[20] K. Yosida and E. Hewitt. Finitely additive measures. *Trans. Amer. Math Soc.*, **72**:46–66, 1952.