On the infinite-horizon optimal control of age-structured systems

Bernhard Skritek, Vladimir Veliov

Research Report 2014-03

March, 2014

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

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

On the infinite-horizon optimal control of age-structured systems*

B. Skritek † V. M. Veliov‡

Abstract

The paper presents necessary optimality conditions of Pontryagin's type for infinite-horizon optimal control problems for age-structured systems with state- and control-dependent boundary conditions. Despite the numerous applications of such problems in population dynamics and economics, a "complete" set of optimality conditions is missing in the existing literature, because it is problematic to define in a sound way appropriate transversality conditions for the corresponding adjoint system. The main novelty is that (building on recent results by S. Aseev and the second author) the adjoint function in the Pontryagin principle is explicitly defined, which avoids the necessity of transversality conditions. The result is applied to several models considered in the literature.

Keywords: age-structured systems, infinite-horizon optimal control, Pontryagin's maximum principle, population dynamics, vintage economic models

MSC Classification: 49K20, 93C20, 91B99, 35F15

1 Introduction

Age-structured first order PDEs [21] provide a main tool for modeling population systems and are recently employed also in economics, where age is involved in order to distinguish machines or technologies of different vintages (dates of production). Optimal control problems for such systems are also widely investigated (see e.g. [5, 6, 12] and the bibliography therein). Most of these problems are naturally formulated on an infinite time-horizon.

Infinite-horizon optimal control problems are still challenging even for ODE systems (see e.g. the recent contributions [2, 4]). The key issue is to define appropriate *transversality conditions*, which allow to select the right solution of the adjoint system for which the Pontryagin maximum principle holds. In the infinite dimensional case (including age-structured systems) this issue is

*This research was supported by the Austrian Science Foundation (FWF) under grant I 476-N13.

†ORCOS, Institute of Mathematical Methods in Economics, Vienna, University of Technology, Argentinierstrasse 8, A-1040 Vienna, Austria, e-mail: `bernhard.skritek@tuwien.ac.at`.

‡ORCOS, Institute of Mathematical Methods in Economics, Vienna University of Technology, Argentinierstrasse 8, A-1040 Vienna, Austria, e-mail: `veliov@tuwien.ac.at`.

open, especially in the case of non-local dynamics or boundary conditions as considered here.[1] This is one reason for which often optimal control problems are considered on a truncated time-horizon (see e.g. [1, 11, 12, 17, 18] and the examples in Section 6), although the natural formulation is on infinite horizon.

The main result in the paper gives necessary optimality conditions of Pontryagin's type for age-structure control systems, where the boundary conditions depend on the current state and on a control. This is a "complete set" of conditions, meaning that the solution of the adjoint system for which the maximization condition in the Pontryagin principle holds is defined in a unique way. The result is obtained by implementing an approach that does not require any transversality conditions, since the "right" solution of the adjoint system is explicitly defined. This approach was recently developed for ODE problems in [2, 3, 4]. The extension to age-structured systems is, however, not straightforward and requires substantial additional work, some of which is rather technical. For this reason we consider only systems with affine dynamics. The approach is implementable also in the non-linear case, where known additional arguments from the stability theory for (non-local) age-structured systems have to be involved.

The usual notion of optimality, in which the optimal solution maximizes the objective functional, is not always appropriate when considering infinite-horizon problems, especially for economic problems with endogenous growth. The reason is that the objective value can be infinite for many (even for all) admissible controls, while they may differ in their intertemporal performance. For this reason we adapt the notion of *weakly overtaking optimality* [8]. Of course, in case of a finite objective functional this notion coincides with the usual one. We mention that the dynamic programming approach (see e.g. [9, 10]), which does not involve any transversality conditions, is not applicable in the case of infinite value function. We also mention that, in contrast to many known results for infinite-horizon ODE problems, the obtained maximum principle is in *normal form*, that is, with the Lagrange multiplier of the objective functional equal to one.

The paper is organized as follows. Section 2 presents the problem and some basic assumptions. Section 3 introduces notations and reminds some known facts. The main result is formulated in Section 4. The proof follows in Section 5. Section 6 presents some selected applications. Some technical proofs are given in the Appendix.

2 Formulation of the problem

Consider the following optimization problem

$$\max_{u,v} \int_0^\infty \int_0^\omega g(t,a,y(t,a),z(t),u(t,a),v(t))\, \mathrm{d}a\, \mathrm{d}t, \tag{1}$$

[1] We do not mention some publications where transversality conditions are introduced ad hoc or based on non-sound arguments (see the recent paper [16] for more information). There are some exceptions, out of which we mention [5], where, however, the dynamics and the boundary conditions are local.

subject to

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)y(t,a)=F(t,a,u(t,a),v(t))\,y(t,a)+f(t,a,u(t,a),v(t)),\tag{2}$$

$$y(t,0)=\Phi(t,v(t))\,z(t)+\varphi(t,v(t)),\qquad y(0,a)=y_0(a),\tag{3}$$

$$z(t)=\int_0^\omega[H(t,a,u(t,a),v(t))\,y(t,a)+h(t,a,u(t,a),v(t))]\,\mathrm{d}a,\tag{4}$$

$$u(t,a)\in U,\tag{5}$$

$$v(t)\in V.\tag{6}$$

Here $(t,a)\in D:=[0,\infty)\times[0,\omega]$, $\omega>0$. The functions $y:D\to\mathbf{R}^n$ and $z:[0,\infty)\to\mathbf{R}^m$ represent the state of the system; $u:D\to U$ and $v:[0,\infty)\to V$ are control functions with values in the subsets U and V of finite-dimensional Euclidean spaces. The matrix- or vector-valued functions F, f, Φ, φ, H, h have corresponding dimensions. The considered system is affine in the states, while the integrand g in the objective functional (1) and the dependence on the controls can be non-linear.

The sets of *admissible controls*, $\mathcal{U}$ and $\mathcal{V}$, consist of all functions $u:D\to U$ and $v:[0,\infty)\to V$ belonging to the spaces $L_\infty^{\mathrm{loc}}(D)$ and $L_\infty^{\mathrm{loc}}(0,\infty)$ of measurable and locally bounded functions, respectively.

We use the classical PDE representation of the transport-reaction equation (2), although the left-hand side should be interpreted as the directional derivative of y in the direction $(1,1)$:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)y(t,a)=\mathcal{D}y(t,a):=\lim_{\varepsilon\to0+}\frac{y(t+\varepsilon,a+\varepsilon)-y(t,a)}{\varepsilon}$$

(further on we use the notation $\mathcal{D}$ for this directional derivative).

Denote by $\mathcal{A}(D)$ the set of all n-dimensional functions $y\in L_\infty^{\mathrm{loc}}(D)$ which are absolutely continuous on almost every characteristic line $t-a=$ const. For $y\in\mathcal{A}(D)$ the traces $y(t,0)$ and $y(0,a)$ in (3) are well-defined almost everywhere. Given $u\in\mathcal{U}$ and $v\in\mathcal{V}$, a couple of functions $y\in\mathcal{A}(D)$, $z\in L_\infty^{\mathrm{loc}}(0,\infty)$ is a solution of (2)–(4) if y satisfies (2) almost everywhere on almost every characteristic line intersecting D, and (3), (4) are also satisfied almost everywhere. For more detailed explanations of the notion of solution of (2)–(4) see e.g. [1, 12, 15, 21].

The following assumptions are standing in this paper.

Assumption (A1). The set V is convex. The functions F, f, Φ, φ, H, h and g, together with the partial derivatives g_y, g_z and the partial derivatives with respect to v of all the above functions, are locally bounded, measurable in (t,a) for every (y,z,u,v), and locally Lipschitz continuous in (y,z,u,v).[2]

[2]The last part of the assumption means that for every compact sets Y, Z, $\bar{U}\subset U$, $\bar{V}\subset V$ and $T>0$ there exists

Now we will clarify the notion of optimality employed in this paper. On assumptions (A1), for every $u \in \mathcal{U}$ and $v \in \mathcal{V}$ system (2)–(4) has a unique solution $(y, z) \in \mathcal{A}(D) \times L_\infty^{\rm loc}(0,\infty)$ (see e.g. [7, Lemma 5.3]). Moreover, for every $T > 0$ the corresponding integral

$$J_T(u,v) = \int_0^T \int_0^\omega g(t,a,y(t,a),z(t),u(t,a),v(t)) \,\mathrm{d}a\,\mathrm{d}t$$

is finite. The following definition follows [8].

Definition 1. *A control pair $(\hat u, \hat v) \in \mathcal{U} \times \mathcal{V}$ is* weakly overtaking optimal *(WOO) if for any $(u,v) \in \mathcal{U} \times \mathcal{V}$ and for every $\varepsilon > 0$ and $T > 0$ there exists $T' \geq T$ such that*

$$J_{T'}(\hat u, \hat v) \geq J_{T'}(u,v) - \varepsilon.$$

Clearly, if the outer integral in (1) is absolutely convergent, then the above definition of optimality coincides with the classical one.

In this paper we do not investigate the issue of existence of a WOO solution. We assume that such exists, and in what follows we fix a WOO solution $(\hat u, \hat v, \hat y, \hat z)$, for which we obtain necessary optimality conditions of Pontryagin's type.

Further on, we use the following *notational conventions*: we skip functions with a "hat" when they appear as arguments of other functions. For example $F(t,a) := F(t,a,\hat u(t,a),\hat v(t))$, $g(t,a,y,z) := g(t,a,y,z,\hat u(t,a),\hat v(t))$, etc.

In addition, we introduce the following simplifying assumption.

Assumption (A2). There exists a measurable function $\rho : [0,\infty) \to [0,\infty)$ such that

$$|g_y(t,a,y,z)| + |g_z(t,a,y,z)| \leq \rho(t) \quad \text{for every } (t,a) \in D \text{ and } (y,z).$$

The above condition is made only for simplification, and it is fulfilled for most applications in economics and population dynamics. It can be removed (following Assumption A2 in [4]) at the price of a more implicit assumption (A3) than the one introduced in Section 4.

---

a constant L such that for each of the functions listed above (take $g(t,a,y,z,u,v)$ as a representative)

$$|g(t,a,y_1,z_1,u_1,v_1) - g(t,a,y_2,z_2,u_2,v_2)| \leq L(|y_1-y_2| + |z_1-z_2| + |u_1-u_2| + |v_1-v_2|)$$

for every $(t,a,y_i,z_i,u_i,v_i) \in [0,T] \times [0,\omega] \times Y \times Z \times \bar U \times \bar V$, $i = 1,2$. This part of Assumption (A1) can be weakened to locally uniform continuity, but we prefer to avoid the resulting technicalities.

3 Preliminaries

In this section we introduce a few notations, remind some known facts, and provide some auxiliary material that will be used in the sequel.

In what follows we use capital letters to denote sets and matrices, lower-case Latin letters to denote numbers and column vectors, and Greek lower-case letters to denote numbers and row-vectors. There will be a few exceptions (such as δ and Δ), which will not lead to confusions.

3.1. Volterra equations of the second kind (see e.g. [14] for details).

Let $K(t,s)$, $t \geq 0$, $s \in [0,t]$ be a measurable and locally bounded $(m \times m)$-matrix function, considered as an integral kernel of a Volterra integral equation of the second kind. The kernel K defines a resolvent, $R(t,s)$, $t \geq 0$, $s \in [0,t]$, which is also measurable and locally bounded, and satisfies the equations

$$R(t,s) = K(t,s) + \int_s^t K(t,\theta)\, R(\theta,s)\, \mathrm{d}\theta = K(t,s) + \int_s^t R(t,\theta)\, K(\theta,s)\, \mathrm{d}\theta \qquad (7)$$

almost everywhere. It will be convenient to extend $R(t,s) = 0$ for $s > t$.

For an arbitrarily fixed $\tau > 0$, and an m-dimensional vector function $q(\tau,\cdot) \in L_\infty^{\mathrm{loc}}(\tau,\infty)$, the Volterra equation

$$p(t) = q(\tau,t) + \int_\tau^t K(t,s)\, p(s)\, \mathrm{d}s, \quad t \geq \tau \qquad (8)$$

has a unique solution $p \in L_\infty^{\mathrm{loc}}(\tau,\infty)$, and it is given by

$$p(t) = q(\tau,t) + \int_\tau^t R(t,s)\, q(\tau,s)\, \mathrm{d}s.$$

(This can be directly checked by inserting the last expression in (8).)

Similarly, if $\psi \in L_\infty^{\mathrm{loc}}(0,\infty)$, $T > 0$, and

$$\zeta(t) = \psi(t) + \int_t^T \psi(\theta)\, R(\theta,t)\, \mathrm{d}\theta, \quad t \in [0,T], \qquad (9)$$

then ζ satisfies the equation

$$\zeta(t) = \psi(t) + \int_t^T \zeta(\theta)\, K(\theta,t)\, \mathrm{d}\theta, \quad t \in [0,T]. \qquad (10)$$

In our considerations below $K(\theta,s) = 0$ for $\theta \notin [s, s+\omega]$, and the integral in (9) is convergent when $T \to \infty$ and is locally bounded in t. In this situation the implication (9) $\Longrightarrow$ (10) is true also for $T = \infty$.

3.2. Fundamental solution of equation (2).

Below we use the notational convention made in the end of Section 2.

Consider the homogeneous part of equation (2) with $(u, v) = (\hat{u}, \hat{v})$:

$$\mathcal{D}y(t, a) = F(t, a)\, y(t, a). \tag{11}$$

Define the set

$$\Gamma_0 := \{(t_0, a_0) \in D : \text{ either } t_0 = 0 \text{ or } a_0 = 0\},$$

that is, the lower left boundary of D. The *fundamental matrix solution* of (11), $X \in L_\infty^{\rm loc}(D)$, is defined as the $(n \times n)$- matrix solution of the equation

$$\mathcal{D}X(t, a) = F(t, a)\, X(t, a), \qquad X(t, a) = I \text{ for } (t, a) \in \Gamma_0 \tag{12}$$

(I is the identity matrix). The definition is correct, since the characteristic lines of (11) emanating from Γ_0 cover in a disjunctive way the domain D. For every $(t_0, a_0) \in \Gamma_0$ the function X can be defined on the characteristic line passing through (t_0, a_0) as $X(t_0 + s, a_0 + s) = Z(s)$, $s \in [0, \omega - a_0]$, where Z is determined by the equation

$$\dot{Z}(s) = F(t_0 + s, a_0 + s)\, Z(s), \quad Z(0) = I. \tag{13}$$

Thus, for given side conditions on the lower-left boundary Γ_0, one can represent the solution of (2) in terms of X by the Cauchy formula for ODEs.

Moreover,

$$\mathcal{D}X^{-1}(t, a) = -X^{-1}(t, a)\, F(t, a), \quad (t, a) \in D. \tag{14}$$

3.3. Lipschitz stability of system (2)–(4).

Below we present a simplified and adapted version of the stability result in [12, Proposition 1], splitting it in two parts. Remember the notational convention made in the end of Section 2. Also, we denote $D_T := [0, T] \times [0, \omega]$.

For a number $\tau \geq 0$ and a function $\delta \in L_\infty(0, \omega)$ we consider system (2)–(4) (with $(u, v) = (\hat{u}, \hat{v})$) on $[\tau, \infty)$, with a "disturbed initial condition"

$$\tilde{y}(\tau, a) = \hat{y}(\tau, a) + \delta(a).$$

Denote by $(\tilde{y}, \tilde{z})$ the corresponding solution on $[\tau, \infty)$.

Proposition 1. *For each $T > 0$ there exists a constant $c_0(T)$ such that for every $\tau \in [0, T)$ and $\delta \in L_\infty(0, \omega)$ it holds that*

$$\|\tilde{y} - \hat{y}\|_{L_k(D_T)} + \|\tilde{z} - \hat{z}\|_{L_k(0,T)} \leq c_0(T)\, \|\delta\|_{L_k(0,\omega)}, \quad k \in \{1, \infty\}.$$

Now, let $B(\tau, b; \alpha)$ denote the box $[\tau - \alpha, \tau] \times [b - \alpha, b]$, where $\tau > 0$, $b \in (0, \omega]$. If $0 < \alpha \leq \alpha_0 := \min\{\tau, b\}$, then $B(\tau, b; \alpha) \subset D$. Let $\bar{u} : B(\tau, b; \alpha_0) \to U$ and $\bar{v} : [\tau - \alpha_0, \tau] \to V$ be two measurable and bounded functions. Consider again system (2)–(4) for two pairs of admissible controls: $(\hat{u}, \hat{v})$ and

$$u_\alpha(t,a) = \begin{cases} \bar{u}(t,a) & \text{for } (t,a) \in B(\tau, b; \alpha), \\ \hat{u}(t,a) & \text{for } (t,a) \notin B(\tau, b; \alpha), \end{cases} \qquad v_\alpha(t) = \begin{cases} \bar{v}(t) & \text{for } t \in [\tau - \alpha, \tau], \\ \hat{v}(t) & \text{for } t \notin [\tau - \alpha, \tau]. \end{cases} \tag{15}$$

Denote by (y_α, z_α) the corresponding solution of (2)–(4).

Proposition 2. *For each $\tau > 0$, $b \in (0, \omega]$, and compact sets $\bar{U} \subset U$ and $\bar{V} \subset V$ there exists a constant c_0 such that for every $\alpha \in (0, \alpha_0]$ and measurable $\bar{u} : B(\tau, b; \alpha) \to \bar{U}$ and $\bar{v} : [\tau - \alpha, \tau] \to \bar{V}$ it holds that*

$$\begin{aligned} &\|z_\alpha - \hat{z}\|_{L_\infty(0,\tau)} \le c_0 \alpha, \\ &\|y_\alpha - \hat{y}\|_{L_\infty(D_\tau)} \le c_0(\alpha + \|\bar{v} - \hat{v}\|_{L_\infty(\tau-\alpha,\tau)}), \\ &\|y_\alpha(t,\cdot) - \hat{y}(t,\cdot)\|_{L_1(0,\omega)} \le c_0 \alpha \left(\alpha + \|\bar{v} - \hat{v}\|_{L_\infty(\tau-\alpha,\tau)}\right), \quad t \in [\tau - \alpha, \tau]. \end{aligned}$$

The first proposition follows directly from [12, Proposition 1], while the second one additionally takes into account the specific (needle variation) form of the disturbance (u_α, v_α). The local Lipschitz property assumed in (A1) is also used in this simplified reformulation of [12, Proposition 1].

3.4. Variation of the initial data in system (2)–(4).

Let us fix a number $\tau > 0$ and consider a variation $\delta(a)$ of the state $\hat{y}(\tau, \cdot)$. We shall study the propagation of this variation on the domain $[\tau, \infty) \times [0, \omega]$. That is, we consider on $[\tau, \infty) \times [0, \omega]$ the system

$$\mathcal{D}y(t,a) = F(t,a)\, y(t,a) + f(t,a), \tag{16}$$

$$y(t,0) = \Phi(t)\, z(t) + \varphi(t), \quad y(\tau, a) = \hat{y}(\tau, a) + \delta(a), \tag{17}$$

$$z(t) = \int_0^\omega [H(t,a)\, y(t,a) + h(t,a)]\, \mathrm{d}a, \tag{18}$$

where $\delta \in L_\infty(0, \omega)$. Denote the corresponding solution by $(y, z) \in \mathcal{A}([\tau, \infty) \times [0, \omega]) \times L_\infty^{\mathrm{loc}}(0, \omega)$. It exists and is unique [7, Lemma 5.3]. For the variations $\Delta y := y - \hat{y}$ and $\Delta z := z - \hat{z}$ we have

$$\begin{aligned} \mathcal{D}\Delta y(t,a) &= F(t,a)\, \Delta y(t,a), \\ \Delta y(t,0) &= \Phi(t)\, \Delta z(t), \quad \Delta y(\tau, a) = \delta(a), \\ \Delta z(t) &= \int_0^\omega H(t,a)\, \Delta y(t,a)\, \mathrm{d}a. \end{aligned}$$

Due to (12) and (13) we have

$$\Delta y(t,a) = \begin{cases} X(t,a)\, X^{-1}(\tau, a - t + \tau)\, \delta(a - t + \tau), & \text{if } a - t + \tau \ge 0, \\ X(t,a)\, \Phi(t-a)\, \Delta z(t-a), & \text{if } a - t + \tau < 0. \end{cases}$$

For convenience we extend the definition of δ and Δz setting $\delta(a) = 0$ for $a \notin [0, \omega]$ and $\Delta z(t) = 0$ for $t \in [0, \tau)$. Then

$$\Delta y(t,a) = X(t,a)\, X^{-1}(\tau, a - t + \tau)\, \delta(a - t + \tau) + X(t,a)\, \Phi(t-a)\, \Delta z(t-a). \tag{19}$$

Moreover, we set $H(t,a) = 0$ for $a \notin [0, \omega]$, and abbreviate $HX(t,a) := H(t,a)\, X(t,a)$. Using the equation for Δz and the above extensions and notation, and changing the variables we have

$$\begin{aligned}\Delta z(t) = \int_0^\omega H(t,a)\, \Delta y(t,a)\, \mathrm{d}a &= \int_0^\omega HX(t,a)\, X^{-1}(\tau, \tau - t + a)\, \delta(\tau - t + a)\, \mathrm{d}a \\ &+ \int_0^\omega HX(t,a)\, \Phi(t-a)\, \Delta z(t-a)\, \mathrm{d}a \\ &= \int_0^\omega HX(t, s + t - \tau)\, X^{-1}(\tau, s)\, \delta(s)\, \mathrm{d}s + \int_\tau^t HX(t, t-s)\, \Phi(s)\, \Delta z(s)\, \mathrm{d}s.\end{aligned}$$

With the notations

$$K(t,s) := HX(t, t-s)\, \Phi(s), \quad Q(\tau, t, s) := HX(t, s + t - \tau)\, X^{-1}(\tau, s), \tag{20}$$

and $q(\tau, t) := \int_0^\omega Q(\tau, t, s)\, \delta(s)\, \mathrm{d}s$, we obtain that Δz satisfies on $[\tau, \infty)$ equation (8) with the measurable and locally bounded kernel K. Notice that $K(t,s) = 0$ for $t > s + \omega$. Denote by $R(t,s)$ its resolvent (see Section 3.1). Then, changing the order of integration below, we obtain that

$$\begin{aligned}\Delta z(t) &= q(\tau, t) + \int_\tau^t R(t,s)\, q(\tau, s)\, \mathrm{d}s \\ &= \int_0^\omega \left[Q(\tau, t, s) + \int_\tau^t R(t,x)\, Q(\tau, x, s)\, \mathrm{d}x \right] \delta(s)\, \mathrm{d}s. \end{aligned} \tag{21}$$

Thus we have the explicit representations (21) and (19) of the variations Δy and Δz as linear functions of δ.

The resolvent R and the fundamental matrix solution X defined above will be involved in all the subsequent analysis.

4 Main result: the maximum principle

Papers [7, 12, 20] contain necessary optimality conditions in the form of the Pontryagin maximum principle for general for age-structured systems on a finite time-horizon $[0, T]$. These conditions involve adjoint functions $\xi : D_T \to \mathbf{R}^n$ and $\zeta : [0, T] \to \mathbf{R}^m$ corresponding to the state variables y and z. These functions satisfy the following *adjoint system*:

$$-\mathcal{D}\xi(t,a) = \xi(t,a)\, F(t,a) + \zeta(t)\, H(t,a) + g_y(t,a), \tag{22}$$

$$\zeta(t) = \xi(t,0)\, \Phi(t) + \int_0^\omega g_z(t,a)\, \mathrm{d}a, \tag{23}$$

where we use the notational convention made in the end of Section 2. This system is complemented by the boundary condition $\xi(t,\omega)=0$ and an appropriate transversality condition at $t=T$. In the infinite-horizon case the adjoint equations are the same, but the transversality condition at $t=\infty$ is problematic due to several reasons (some of them are present also for ODE control problems). In the result below we avoid the necessity of transversality conditions, since we explicitly define a unique solution of the adjoint system for which the maximum principle holds.

It will be convenient to define the pre-Hamiltonian

$$\mathcal{H}(t,a,y,z,u,v,\xi,\zeta) := g(t,a,y,z,u,v) + \xi\,[F(t,a,u,v)\,y + f(t,a,u,v)] \\ + \zeta\,[H(t,a,u,v)\,y + h(t,a,u,v)]. \tag{24}$$

With this notation the adjoint equation (22) can be written in the shorter form

$$\mathcal{D}\xi(t,a) = -\mathcal{H}_y(t,a,\xi(t,a),\zeta(t)). \tag{25}$$

Now, using the notations introduced in Section 3, we define the following functions

$$\hat{\xi}(t,a) := \left[\int_a^\omega g_y X(t-a+x,x)\,\mathrm{d}x + \int_t^{t+\omega-a} \hat{\zeta}(\theta)\,HX(\theta,\theta-t+a)\,\mathrm{d}\theta\right] X^{-1}(t,a), \quad (t,a)\in D, \tag{26}$$

$$\hat{\zeta}(t) := \psi(t) + \int_t^\infty \psi(\theta)\,R(\theta,t)\,\mathrm{d}\theta, \quad t\ge 0, \tag{27}$$

where

$$\psi(t) := \int_0^\omega [g_y X(t+s,s)\,\Phi(t) + g_z(t,s)]\,\mathrm{d}s, \tag{28}$$

and as before we shorten $g_yX(t,a)=g_y(t,a)X(t,a)$ and $HX(t,a):=H(t,a)X(t,a)$.

In order to justify the utilization of the infinite-horizon integral in the definition of $\hat{\zeta}$ we introduce the following additional assumption.

Assumption (A3). There exists a measurable function $\lambda(t,\theta)$, $\lambda:[0,\infty)\times[0,\infty)\to[0,\infty)$, such that

$$\int_0^\omega [\rho(t+s)\,|X(t+s,s)|\,|\Phi(t)| + \rho(t)]\,\mathrm{d}s\;|R(t,\theta)| \le \lambda(t,\theta), \quad \forall\, t\ge\theta\ge 0,$$

and the integral $\int_\theta^\infty \lambda(t,\theta)\,\mathrm{d}t$ is finite and locally bounded as a function of θ.

Essentially, the above assumption poses some restriction on the combined growth of the resolvent R, the fundamental matrix solution X, and the data of the problem. It can be formulated in different ways, out of which we chose the one that is most convenient in the proof. Sufficient conditions for (A3) that are easier to check are given in the end of this section.

Lemma 1. *On assumptions (A1)–(A3) the integral in (27) is absolutely convergent and the functions $\hat{\xi}$ and $\hat{\zeta}$ defined by (26), (27) (regarding (28)) belong to the spaces $\mathcal{A}(D)$ and $L_\infty^{\text{loc}}(0,\infty)$, correspondingly, and satisfy the adjoint system (22), (23).*

This lemma will be proved in Appendix.

Now, we are ready to formulate the main result in this paper.

Theorem 1. *Let assumptions (A1), (A2) be satisfied. Let* $(\hat{u}, \hat{v}, \hat{y}, \hat{z})$ *be a WOO solution of problem* (1)–(6)*, and let Assumption (A3) be fulfilled for this solution.*

Then the functions $\hat{\xi}$ *and* $\hat{\zeta}$ *defined in (26) and (27) (with regard of (28)) satisfy the adjoint system (22), (23), and the following maximization conditions are fulfilled:*

$$\mathcal{H}(t, a, \hat{u}(t, a)) = \sup_{u \in U} \mathcal{H}(t, a, u),$$

$$\left[\int_0^\omega \mathcal{H}_v(t, a, \hat{v}(t)) \, \mathrm{d}a + \xi(t, 0)(\Phi_v(t, \hat{v}(t))\hat{z}(t) + \varphi_v(t, \hat{v}(t))) \right] (v - \hat{v}(t)) \leq 0, \quad \forall\, v \in V.$$

A few remarks follow. First, we mention that the above maximum principle is of *normal form*, that is, the objective integrand appears in the definition of the pre-Hamiltonian with a multiplier equal to 1. This is typical for finite-horizon problems without state constraints, but not for infinite-horizon problems (notice that our definition of optimality goes even beyond the classical one). Second, the maximization condition with respect to v is local (in contrast to that for u). This is the case also for finite-horizon problems, and it is an open question if a global maximum principle with respect to the boundary control really holds. The formal reason for this localness is in Proposition 2, where the L_∞-norm of the disturbance of v appears in the estimation, rather the L_1-norm, as it is for u.

It is important to mention that the adjoint variable $\hat{\xi}(t, \cdot)$ defined in (26), (27) does not necessarily converge to zero when $t \to \infty$. That is, the "classical" transversality condition $\lim_{t\to\infty} \hat{\xi}(t, \cdot) = 0$ does not work in general. The same applies to the condition $\lim_{t\to\infty} \int_0^\omega \hat{\xi}(t, a)\, \hat{y}(t, a) \, \mathrm{d}a = 0$, the ODE-counterpart of which is also used in the literature.

The above theorem is formulated and proved for affine systems. The extension to nonlinear systems, where the aggregate state z does not appear in the state equation (2) is a matter of technicality. However, if z appears in (2), then the problem becomes substantially more difficult.

Now, we elaborate Assumption (A3). If the growth estimations

$$|g_y(t, a, y, z)| \leq c e^{\lambda_1 t}, \quad |X(t, a)| \leq c e^{\lambda_2 t}, \quad |\Phi(t)| \leq c e^{\lambda_3 t}, \quad |g_z(t, a, y, z)| \leq c e^{\lambda_4 t} \tag{29}$$

hold for some constants c and λ_i, then the integral in (A3) is majored by

$$\tilde{c}\, e^{\lambda_0 t} |R(t, \theta)|, \quad \text{where } \lambda_0 := \max\{\lambda_1 + \lambda_2 + \lambda_3, \, \lambda_4\}, \tag{30}$$

and $\tilde{c}$ is another constant. Then (A3) will be satisfied if for $\lambda(t, \theta) := \tilde{c}\, e^{\lambda_0 t}\, |R(t, \theta)|$ the integral $\int_\theta^\infty \lambda(t, \theta) \, \mathrm{d}t$ is finite and locally bounded as a function of θ. The following is a sufficient condition for that, which does not involve the resolvent R.

Assumption (A3'). The inequalities (29) hold for any $(t, a, y, z) \in D \times \mathbf{R}^n \times \mathbf{R}^m$ and

$$\operatorname*{ess\,sup}_{t \in [0, \infty)} \int_t^{t+\omega} e^{\lambda_0 (s-t)} |K(s, t)| \, \mathrm{d}s < 1, \tag{31}$$

where λ_0 is as in (30).

Let us denote by $L_\infty^{-\lambda_0}(0,\infty)$ the weighted L_∞-space with the weight $e^{-\lambda_0 t}$. In order to show that (A3') implies (A3) with $\lambda(t,\theta) := \tilde{c}\, e^{\lambda_0 t} |R(t,\theta)|$ we use Proposition 3.10 in [14, Chapter 9], according to which the operator $\mathcal{R}$ defined as $(\mathcal{R}\mu)(\theta) := \int_\theta^\infty \mu(t) R(t,\theta)\,\mathrm{d}t$ maps $L_\infty^{-\lambda_0}(0,\infty)$ into itself and is bounded. Applying this fact for $\mu(t) = \tilde{c} e^{-\lambda_0 t}$ we obtain that the function $\theta \mapsto \int_\theta^\infty \lambda(t,\theta)\,\mathrm{d}t$ is finite and bounded in $L_\infty^{-\lambda_0}(0,\infty)$, thus locally bounded.

Thus (A3') implies (A3). Inequality (31) in (A3') has a clear interpretation for population models, as it will be indicated in Section 6.2.

We also mention that λ_1 and λ_4 are usually negative due to discounting (which is implicitly included in g), and $\lambda_3 = 0$. Then λ_0 can happen to be negative, which helps for the validity of (31). The models mentioned in Section 6.3 crucially employ this fact.

5 Proof of the maximum principle

The proof of Theorem 1 is somewhat long and technical, therefore first we briefly present the idea, which builds on [4]. We follow the general understanding that the adjoint function $\xi(t,\cdot)$ evaluated at time t gives the principle term of the effect of a disturbance $\delta = \delta(a)$ of the state $\hat{y}(t,\cdot)$ on the objective value. Therefore, for an arbitrary $\tau \in (0,\infty)$ we consider a disturbance $\delta(a)$ of $\hat{y}(\tau,\cdot)$. Then, by linearization, one can represent the objective value on an interval $[\tau, T]$, $T > \tau$, corresponding to the perturbed $\hat{y}(\tau,\cdot) + \delta(\cdot)$ (with the same controls $\hat{u}$ and $\hat{v}$) as

$$\int_\tau^T \xi^T(\tau,a)\,\delta(a)\,\mathrm{d}a \;+\; \text{“rest terms”}.$$

We shall obtain a representation of $\xi^T(\tau,\cdot)$ in terms of the fundamental matrix solution X and the resolvent R. Then utilizing assumption (A3) we prove that $\xi^T(\tau,\cdot)$ converges to the adjoint function ξ defined in (26), and Lemma 1 holds. This is the first part of the proof.

In the second part, we apply a needle-type variation of the controls on $[\tau-\alpha,\tau]$, which results in a specific disturbance δ of $\hat{y}(\tau,\cdot)$. Then we represent the direct effect of this variation on the objective value (that is, on $[\tau-\alpha,\tau]$) and the indirect effect (resulting from δ) in terms of the pre-Hamiltonian $\mathcal{H}$. Finally, we use the definition of WOO to obtain the maximization conditions in Theorem 1.

Now we begin with the detailed proof, in which we use the notational convention made in the end of Section 2.

Part 1. Let us fix an arbitrary $\tau > 0$ and consider any two numbers $T > \tau + \omega$ and $T' > T + \omega$. For any $\delta \in L_\infty(0,\omega)$ we consider the disturbed system (16)–(18). Using the same notation (y,z)

and $(\Delta y, \Delta z)$ as in Part 3.4 of Section 3, we obtain in a standard way the representation

$$\Delta_{T'}(y,z) := \int_\tau^{T'} \int_0^\omega [g(t,a,y(t,a),z(t)) - g(t,a)] \, \mathrm{d}a \, \mathrm{d}t$$
$$= \int_\tau^{T'} \int_0^\omega [\bar{g}_y(t,a)\, \Delta y(t,a) + \bar{g}_z(t,a)\, \Delta z(t)] \, \mathrm{d}a \, \mathrm{d}t,$$

where $\bar{g}_y(t,a) := g_y(t,a,\bar{y}(t,a),\bar{z}(t,a))$, $\bar{g}_z(t,a) := g_z(t,a,\bar{y}(t,a),\bar{z}(t,a))$, and $\bar{y}$, $\bar{z}$ are measurable functions satisfying

$$(\bar{y}(t,a),\bar{z}(t,a)) \in \mathrm{co}\,\{(y(t,a),z(t)),(\hat{y}(t,a),\hat{z}(t))\}. \tag{32}$$

Now, we use the representation (21) of Δz, and the representation (19) of Δy, with (21) inserted in (19). After some elementary calculus (changing variables and order of integration) we obtain that

$$\Delta_{T'}(y,z) = \int_0^\omega \xi^{T'}(\tau,s)\, \delta(s) \, \mathrm{d}s,$$

where

$$\xi^{T'}(\tau,s) = \int_s^\omega \bar{g}_y X(\tau+a-s,a) \, \mathrm{d}a \, X^{-1}(\tau,s) \tag{33}$$
$$+ \int_\tau^{T'} \left[\int_0^\omega \big(\chi(T'-a-\theta)\, \bar{g}_y X(\theta+a,a)\, \Phi(\theta) + \bar{g}_z(\theta,a)\big) \, \mathrm{d}a \right] \tag{34}$$
$$\times \left[Q(\tau,\theta,s) + \int_\tau^\theta R(\theta,x)\, Q(\tau,x,s) \, \mathrm{d}x \right] \mathrm{d}\theta, \tag{35}$$

and χ is the Heaviside-function: $\chi(s)$ equals 0 for $s<0$ and equals 1 for $s \geq 0$.

In the above expression for $\xi^{T'}(\tau,s)$ we shall split $\int_\tau^{T'} = \int_\tau^T + \int_T^{T'}$ and will investigate the two appearing terms separately. We shall use the symbols c_1, c_2, ... for numbers that are independent of δ, and also of T and T', unless otherwise indicated by an argument of c_i. However, these numbers may depend on τ and s.

According to Proposition 1,

$$\|\Delta y\|_{L_\infty(D_T)} + \|\Delta z\|_{L_\infty(0,T)} \leq c_0(T)\, \|\delta\|_{L_\infty(0,\omega)}.$$

Then both (y,z) and $(\hat{y},\hat{z})$ remain in a bounded domain when $t \leq T$, and in this domain g_y and g_z are Lipschitz continuous (with a constant depending on T, of course). Moreover, X, Φ, and the term in the brackets in (35) are bounded when $\theta \leq T$ (again by a constant depending on T). Therefore, having in mind (32), we can replace the functions $\bar{g}_y(t,a)$ and $\bar{g}_z(t,a)$ with $g_y(t,a)$ and $g_z(t,a)$ in the term (33), and also in the term (34), where the integration is taken only to T. For the resulting residual, $e_1(\tau,s;T)$, we have

$$|e_1(\tau,s;T)| \leq c_1(T)\, \|\delta\|_{L_\infty(0,\omega)}.$$

The integral on $[T,T']$ in (34), (35) will be estimated differently, using assumption (A3). We obtain the following estimation of this integral, denoted by $e_2(\tau,s;T,T')$. First of all, we notice that $Q(\tau,\theta,s) = 0$ if $\theta + s - \tau > \omega$ (see Section 3.4) and this is the case if $\theta > T$. Thus the remaining integral on $[T, T']$ in (34), (35) can be estimated by

$$|e_2(\tau,s;T,T')| \le \int_T^{T'} \int_0^\omega (\rho(\theta+a)|X(\theta+a,a)\,\Phi(a)| + \rho(\theta))\,\mathrm{d}a \int_\tau^\theta |R(\theta,x)|\,|Q(\tau,x,s)|\,\mathrm{d}x\,\mathrm{d}\theta$$

$$= \int_\tau^{\tau+\omega} \int_T^{T'} \int_0^\omega (\rho(\theta+a)|X(\theta+a,a)\,\Phi(a)| + \rho(\theta))\,\mathrm{d}a\,|R(\theta,x)|\,\mathrm{d}\theta\,|Q(\tau,x,s)|\,\mathrm{d}x$$

$$\le \int_\tau^{\tau+\omega} \int_T^{T'} \lambda(\theta,x)\,\mathrm{d}\theta\,\|Q(\tau,\cdot,s)\|_{L_\infty(\tau,\tau+\omega)} \le c_2 \int_\tau^{\tau+\omega} \int_T^\infty \lambda(\theta,x)\,\mathrm{d}\theta\,\mathrm{d}x.$$

The last term converges to zero when $T \to \infty$, due to the assumption about λ in (A3) and the Lebesgue dominated convergence theorem.

As a result of the above considerations we obtain that

$$\xi^{T'}(\tau,s) = \int_s^\omega g_y X(\tau+a-s,a)\,\mathrm{d}a\,X^{-1}(\tau,s) + \int_\tau^T \left[\int_0^\omega (g_y X(\theta+a,a)\,\Phi(\theta) + g_z(\theta,a))\,\mathrm{d}a\right]$$
$$\times \left[Q(\tau,\theta,s) + \int_\tau^\theta R(\theta,x)\,Q(\tau,x,s)\,\mathrm{d}x\right]\mathrm{d}\theta + e_1(\tau,s;T) + e_2(\tau,s;T,T')$$

(we used that $\chi(T'-a-\theta) = 1$ for $\theta \le T$, since $T' > T+\omega$). Rearranging the terms, substituting Q from (20), and using that $Q(\tau,\theta,s) = 0$ for $\theta > \tau+\omega-s$ and that $T > \tau+\omega$, the above expression for $\xi^{T'}$ becomes

$$\xi^{T'}(\tau,s) = \left[\int_s^\omega g_y X(\tau-s+a,a)\,\mathrm{d}a + \int_\tau^{\tau+\omega-s} \zeta^T(\theta)\,H X(\theta,\theta-\tau+s)\,\mathrm{d}\theta\right] X^{-1}(\tau,s)$$
$$+ e_1(\tau,s;T) + e_2(\tau,s;T,T'),$$

where

$$\zeta^T(\theta) := \psi(\theta) + \int_\tau^T \psi(x)\,R(x,\theta)\,\mathrm{d}x,$$

and ψ is given by (28). Due to Assumption (A3)

$$\int_T^\infty |\psi(x)\,R(x,\theta)|\,\mathrm{d}x \le \int_T^\infty \lambda(x,\theta)\,\mathrm{d}x$$

and the right-hand side is locally bounded in θ. Then we obtain that

$$\xi^{T'}(\tau,s) = \hat{\xi}(\tau,s) + e_1(\tau,s;T) + e_2(\tau,s;T,T') + e_3(\tau;T), \quad |e_3(\tau;T)| \le c_3 \int_\tau^{\tau+\omega} \int_T^\infty \lambda(x,\theta)\,\mathrm{d}x\,\mathrm{d}\theta$$

and the last term converges to zero when $T \to \infty$ due to the dominated convergence theorem. Hence, for the variation of the objective value we have

$$\Delta_{T'}(y,z) = \int_0^\omega \hat{\xi}(\tau,s)\,\delta(s)\,\mathrm{d}s + e_4(\tau;T,T',\delta), \tag{36}$$

with

$$|e_4(\tau;T,T',\delta)| \le \left(c_4\bar{\varepsilon}(T) + c_5(T)\|\delta\|_{L_\infty(0,\omega)}\right)\|\delta\|_{L_1(0,\omega)}, \tag{37}$$

where $\bar{\varepsilon}(T) \to 0$ when $T \to \infty$. Clearly, the constants c_4 and c_5, and the function $\bar{\varepsilon}$ may depend also on τ.

In order to shorten the notations, further on we abbreviate $F^\sharp(t,a,y,u,v) := F(t,a,u,v)\,y + f(t,a,u,v)$, and similarly $H^\sharp = H\,y + h$, $\Phi^\sharp = \Phi\,y + \varphi$. Moreover, we apply the notational convention (to skip arguments with "hat"-s) made in the end of Section 2.

Part 2 for u.

Now we investigate the effect of a needle variation of the control (u,v) on the objective value, starting with u (keeping $v = \hat{v}$). Let us fix an arbitrary $u \in U$, and denote by $\Omega(u)$ the set of all points $(\tau,b) \in \operatorname{int} D$ which are Lebesgue points of each of the following functions

$$(t,a) \mapsto g(t,a,u) - g(t,a), \qquad (t,a) \mapsto \int_0^\omega g_z(t,s)\,\mathrm{d}s\,(H^\sharp(t,a,u) - H^\sharp(t,a)),$$
$$(t,a) \mapsto \hat{\xi}(\tau,\tau-t+a)\,\left[F^\sharp(t,a,u) - F^\sharp(t,a)\right], \qquad (t,a) \mapsto \hat{\xi}(t,0)\,\Phi(t)\,(H^\sharp(t,a,u) - H^\sharp(t,a)).$$

This means that, taking $p = p(t,a)$ as a representative of the above functions,

$$\lim_{\alpha\to 0+} \frac{1}{\alpha^2}\int_{B(\tau,b;\alpha)} p(t,a)\,\mathrm{d}a\,\mathrm{d}t = p(\tau,b),$$

where $B(\tau,b;\alpha) := [\tau-\alpha,\tau]\times[b-\alpha,b]$, as in Section 3.3.

Let us arbitrarily fix $(\tau,b) \in \Omega(u)$ and let $\alpha > 0$ be such that $2\alpha < \tau$, $2\alpha < b$, and $2\alpha < \omega - b$. Define the control u_a as in (15) with $\bar{u}(t,a) = u$. Let (y_α, z_α) be the solution of (2)–(4) corresponding to $(u_\alpha,\hat{v})$. According to Proposition 2, we have for the resulting differences $\Delta y = y_\alpha - \hat{y}$, $\Delta z = z_\alpha - \hat{z}$

$$\|\Delta y\|_{L_\infty(D_\tau)} + \|\Delta z\|_{L_\infty(0,\tau)} \le c_0(\tau)\,\alpha, \qquad \|\Delta y(\tau,\cdot)\|_{L_1(0,\omega)} \le c_0(\tau)\,\alpha^2. \tag{38}$$

From the first estimation and the absolute continuity of y along the characteristic lines $t-a=$ const, it follows that $\delta := \Delta y(\tau,\cdot)$ satisfies

$$\|\delta\|_{L_\infty(0,\omega)} \le c_0(\tau)\,\alpha. \tag{39}$$

Hence, from (37) we have

$$|e_4(\tau;T,T',\delta)| \le (c_4\bar{\varepsilon}(T) + c_5(T)\,c_0(\tau)\,\alpha)\,c_0(\tau)\,\alpha^2 = c_6\bar{\varepsilon}(T)\,\alpha^2 + c_7(T)\,\alpha^3. \tag{40}$$

According to Proposition 1, the second inequality in (38), and (39), we obtain that for every $T > \tau + \omega$

$$\|\Delta y\|_{L_\infty(D_T\setminus D_\tau)} + \|\Delta z\|_{L_\infty(\tau,T)} \le c_0(T)^2\,\alpha, \quad \|\Delta y\|_{L_1(D_T\setminus D_\tau)} + \|\Delta z\|_{L_1(\tau,T)} \le c_0(T)^2\,\alpha^2. \tag{41}$$

Now, for arbitrarily fixed $T > \tau + \omega$ and $T' > T + \omega$, we consider the variation

$$\begin{aligned} J_{T'}(u_\alpha, \hat{v}) - J_{T'}(\hat{u}, \hat{v}) &= \int_{\tau-\alpha}^{\tau} \int_0^\omega [g(t, a, y_\alpha(t, a), z_\alpha(t), u_\alpha(t, a)) - g(t, a)] \,\mathrm{d}a \,\mathrm{d}t + \Delta_{T'}(y_\alpha, z_\alpha) \\ &= \Delta_\tau(u_\alpha) + \int_0^\omega \hat{\xi}(\tau, s)\, \delta(s) \,\mathrm{d}s + e_4(\tau; T, T', \delta), \end{aligned} \tag{42}$$

where $\Delta_\tau(u_\alpha)$ is a notation for the above double integral, and the last term results from (36) with $\delta(a) = \Delta y(\tau, a)$.

In the sequel $o(\varepsilon)$ denotes any function (independent of T and T' but possibly depending on τ) such that $|o(\varepsilon)|/\varepsilon \to 0$ with $\varepsilon \to 0$.

Lemma 2. *The term $\Delta_\tau(u_\alpha)$ in (42) has the representation*

$$\Delta_\tau(u_\alpha) = \alpha^2 \left(g(\tau, b, u) - g(\tau, b)\right) + \alpha^2 \int_0^\omega g_z(\tau, a) \,\mathrm{d}a \left(H^\sharp(\tau, b, u) - H^\sharp(\tau, b)\right) + o(\alpha^2),$$

Lemma 3. *The term $\int_0^\omega \hat{\xi}(\tau, a)\, \delta(a) \,\mathrm{d}a$ in (42), with $\delta(a) = \Delta y(\tau, a)$ has the representation*

$$\begin{aligned} \int_0^\omega \hat{\xi}(\tau, a)\, \delta(a) \,\mathrm{d}a &= \alpha^2\, \hat{\xi}(\tau, b) \left(F^\sharp(\tau, b, u) - F^\sharp(\tau, b)\right) \\ &\quad + \alpha^2\, \hat{\xi}(\tau, 0)\, \Phi(\tau) \left(H^\sharp(\tau, b, u) - H^\sharp(\tau, b)\right) + o(\alpha^2). \end{aligned}$$

Combining the representations in the last two lemmas and (42), and taking into account the definition of the pre-Hamiltonian $\mathcal{H}$ we obtain that

$$J_{T'}(u_\alpha, \hat{v}) - J_{T'}(\hat{u}, \hat{v}) = \alpha^2 \left[\mathcal{H}(\tau, b, u) - \mathcal{H}(\tau, b)\right] + e_4(\tau; T, T', \delta) + o(\alpha^2). \tag{43}$$

Part 3 for u.

In the above parts of the proof we have fixed an arbitrary $u \in U$, and arbitrary $(\tau, b) \in \Omega(u)$, and for all sufficiently small $\alpha > 0$ we have defined the variation u_α of $\hat{u}$. Now, let us take an arbitrary $\varepsilon_0 > 0$ and an arbitrary $T > \tau + \omega$ such that $\bar{\varepsilon}(T) \le \varepsilon_0$ (see the line below (37)). We shall apply Definition 1 for $(u_\alpha, \hat{v})$, $\varepsilon = \alpha^3$ and T. It says that there exists $T' > T$ (without any restriction we may assume $T' > T + \omega$) such that $J_{T'}(\hat{u}, \hat{v}) \ge J_{T'}(u_\alpha, \hat{v}) - \varepsilon$. Then, according to (43), (37) and (40),

$$\begin{aligned} \alpha^3 = \varepsilon \ge J_{T'}(u_\alpha, \hat{v}) - J_{T'}(\hat{u}, \hat{v}) &= \alpha^2 \left[\mathcal{H}(\tau, b, u) - \mathcal{H}(\tau, b)\right] + e_4(\tau; T, T', \delta) + o(\alpha^2) \\ &\ge \alpha^2 \left[\mathcal{H}(\tau, b, u) - \mathcal{H}(\tau, b)\right] - c_6\, \varepsilon_0\, \alpha^2 - c_7(T)\alpha^3 - |o(\alpha^2)|. \end{aligned}$$

Dividing by α^2 we obtain that

$$\alpha \ge \mathcal{H}(\tau, b, u) - \mathcal{H}(\tau, b) - c_6\, \varepsilon_0 - c_7(T)\alpha - \theta(\alpha),$$

where $\theta(\alpha) \to 0$ with $\alpha \to 0$. Passing to a limit with $\alpha \to 0$ we have $0 \geq \mathcal{H}(\tau, b, u) - \mathcal{H}(\tau, b) - c_6\, \varepsilon_0$, and since $\varepsilon_0 > 0$ was arbitrarily fixed, we obtain the inequality

$$\mathcal{H}(\tau, b, u) \leq \mathcal{H}(\tau, b, \hat{u}(\tau, b))$$

for the considered $u \in U$ and $(\tau, b) \in \Omega(u)$.

Now, consider a countable and dense subset $U^d \subset U$. Since $\Omega(u)$ has full measure in D for every $u \in U^d$, then $D' := \cap_{u \in U^d} \Omega(u)$ also has full measure in D. Thus the inequality $\mathcal{H}(t, a, u) \leq \mathcal{H}(t, a, \hat{u}(t, a))$ is fulfilled for every $(t, a) \in D'$ and every $u \in U^d$. Since U^d is dense in U and $\mathcal{H}$ is continuous in u, this inequality holds for every $u \in U$ and $(t, a) \in D'$. This implies the first relation in Theorem 1.

Part 2 for v. Now, we fix $u = \hat{u}$ and consider a variation v_α as in (15) with $\bar{v}(t) = v_\alpha(t) := \hat{v}(t) + \alpha(v - \hat{v}(t))$. Here $v \in V$ is arbitrarily chosen, $\tau \in \Omega(v)$, $\alpha \in (0, \tau)$, where now $\Omega(v)$ is the set of all τ that are Lebesgue points of the following functions:

$$t \mapsto \int_0^\omega g_z(t, s)\, \mathrm{d}s \int_0^\omega H_v^\sharp(t, a)(v - \hat{v}(t))\, \mathrm{d}a, \quad t \mapsto \hat{\xi}(t, 0)\Phi(t) \int_0^\omega H_v^\sharp(t, a)(v - \hat{v}(t))\, \mathrm{d}a,$$
$$t \mapsto \hat{\xi}(t, 0)\Phi_v^\sharp(t)(v - \hat{v}(t)), \quad t \mapsto \int_0^\omega \hat{\xi}(t, a) F_v^\sharp(t, a)(v - \hat{v}(t))\, \mathrm{d}a, \quad t \mapsto \int_0^\omega g_v(t, a)(v - \hat{v}(t))\, \mathrm{d}a.$$

Let (y_α, z_α) be the solution of (2)–(4) corresponding to $(\hat{u}, v_\alpha)$. Similarly as in "Part 2 for u", one can obtain estimations (38)–(41), thanks to the inequality $\|v_\alpha - \hat{v}\|_{L_\infty(0,\infty)} \leq c\,\alpha$.

Now, for any $T > \tau + \omega$ and $T' > T + \omega$ we consider the objective value

$$\begin{aligned} J_{T'}(\hat{u}, v_\alpha) - J_{T'}(\hat{u}, \hat{v}) &= \int_{\tau-\alpha}^{\tau} \int_0^\omega [g(t, a, y_\alpha(t, a), z_\alpha(t), v_\alpha(t, a)) - g(t, a)]\, \mathrm{d}a\, \mathrm{d}t + \Delta_{T'}(y_\alpha, z_\alpha) \\ &= \Delta_\tau(v_\alpha) + \int_0^\omega \hat{\xi}(\tau, a)\, \delta(a)\, \mathrm{d}a + e_4(\tau; T, T', \delta), \end{aligned} \tag{44}$$

where $\Delta_\tau(v_\alpha)$ is a notation for the above double integral.

Lemma 4. *The term $\Delta_\tau(v_\alpha)$ in (44) has the representation*

$$\Delta_\tau(v_\alpha) = \alpha^2 \int_0^\omega g_z(\tau, a)\, \mathrm{d}a \int_0^\omega H_v^\sharp(\tau, a)(v - \hat{v}(\tau))\, \mathrm{d}a + \alpha^2 \int_0^\omega g_v(\tau, a)(v - \hat{v}(\tau))\, \mathrm{d}a + o(\alpha^2),$$

Lemma 5. *The term $\int_0^\omega \hat{\xi}(\tau, s)\, \delta(s)\, \mathrm{d}s$ in (44), with $\delta(a) = \Delta y(\tau, a)$ has the representation*

$$\begin{aligned} \int_0^\omega \hat{\xi}(\tau, a)\, \delta(a)\, \mathrm{d}a = \alpha^2 \int_0^\omega \hat{\xi}(\tau, a)\, F_v^\sharp(\tau, a)(v - \hat{v}(\tau))\, \mathrm{d}a + \alpha^2\, \hat{\xi}(\tau, 0)\, \Phi_v^\sharp(\tau)(v - \hat{v}(\tau)) \\ + \alpha^2\, \hat{\xi}(\tau, 0)\, \Phi(\tau) \int_0^\omega H_v^\sharp(\tau, a)(v - \hat{v}(\tau))\, \mathrm{d}a + o(\alpha^2). \end{aligned}$$

Part 3 for v. Thanks to the above lemmas, and using that $\hat{\zeta}$ satisfies (23) we obtain the following representation for $T > \tau + \omega$ and $T' > T + \omega$:

$$J_{T'}(\hat{u}, v_\alpha) - J_{T'}(\hat{u}, \hat{v}) = \alpha^2 \left[\int_0^\omega \mathcal{H}_v(t, a, \hat{v}(t))\,\mathrm{d}a + \xi(t, 0)\,\Phi_v^\sharp(t, \hat{v}(t)) \right] (v - \hat{v}(t)) + e_4(\tau; T, T', \delta) + o(\alpha^2).$$

Then the prove the second inequality in Theorem 1 goes in essentially the same way as in "Part 3 for u".

The proof is complete.

6 Selected Applications

In this section we apply the obtained result to a few models from the economic literature, and, in particular, we shed some light on Assumption (A3). The first two have been analyzed on a finite horizon although the natural formulation is on the infinite horizon. We show that our Assumption (A3) is satisfied in these examples.

6.1 A Problem of Optimal Investment

In many cases, e.g. [11, 17, 18], the boundary condition (3) does not involve the integral state z. In these cases, Assumption (A3) is trivially fulfilled because the resolvent $R = 0$ (see (7) and also (20), where $\Phi = 0$).

Consider for example the optimal investment problem in [11] (where we change notations to fit to our general model). The objective is to maximize the discounted net profit,

$$\max_{u,v} \int_0^\infty e^{-rt} \Big(p(z(t)) - (b_0\, v(t) + c_0\, v(t)^2) - \int_0^\omega (b(a)\, u(t, a) + c(a)\, u(t, a)^2)\,\mathrm{d}a \Big) \mathrm{d}t,$$

subject to

$$\mathcal{D}y(t, a) = u(t, a) - \mu(a)\, y(t, a), \quad y(t, 0) = v(t), \quad y(0, a) = y_0(a),$$
$$z(t) = \int_0^\omega H(t - a)\, y(t, a)\,\mathrm{d}a,$$

where $y(t, a)$ is the capital stock of machines of age a at time t, u and v are investments in old and current vintages, respectively, $H(s)$ is the productivity of technologies (machines) of vintage s.

Here the fundamental solution X has the form

$$X(t, a) = e^{-\int_0^{\min\{t,a\}} \mu(a-s)\,\mathrm{d}s}.$$

The adjoint functions $\hat{\xi}$ and $\hat{\zeta}$ defined in (26), (27), (28) take the explicit forms

$$\psi(\theta) = e^{-rt} p'(\hat{z}(\theta)), \quad \hat{\zeta}(\theta) = \psi(\theta),$$
$$\hat{\xi}(t, a) = \int_t^{t+\omega-a} e^{-rt} p'(\hat{z}(\theta))\, H(t - a)\, X(\theta, \theta - t + a)\,\mathrm{d}\theta\, X^{-1}(t, a).$$

A few remarks follow. Clearly X is bounded, as well as X^{-1}, since the depreciation rate μ can be assumed bounded in the life-span of the machines. The revenue function $p(z)$ is defined on $(0, \infty)$ and is non-negative, increasing and concave. Then $p'(\hat{z}(t))$ is bounded, provided that $\hat{z}(t)$ does not approach zero, which is the case in the considered model. Consequently,

$$|\hat{\xi}(t,a)| \leq c \int_t^{t+\omega-a} e^{-rt} H(t-a)\,\mathrm{d}\theta.$$

If for the productivity function it holds that $H(t) < c_1 e^{\rho t}$ with $\rho < r$, then $\hat{\xi}(t,\cdot) \to 0$ when $t \to \infty$. Thus $\hat{\xi}$ satisfies the usual transversality condition. This result is consistent with that in [5], where $r > 0$ and $\rho = 0$. However, if $\rho \geq r$ the considered problem still has a WOO solution, and Theorem 1 holds with the above defined adjoint functions $\hat{\xi}$ and $\hat{\zeta}$, although neither of the standard transversality conditions $\hat{\xi}(t,\cdot) \to 0$ and $\hat{\xi}(t,\cdot)\,\hat{y}(t,\cdot) \to 0$ is fulfilled.

6.2 Optimal Harvesting

Consider the model of optimal harvesting in [1, p. 75]. The problem reads as

$$\max_{v(t)} \int_0^\infty e^{-rt} \int_0^\omega v(t)\, p(a)\, y(t,a)\,\mathrm{d}a\,\mathrm{d}t,$$

subject to

$$\begin{aligned}
\mathcal{D}y(t,a) &= -(\mu(a) + v(t))\, y(t,a),\\
y(0,a) &= y_0(a) > 0, \quad y(t,0) = z(t),\\
z(t) &= \int_0^\omega \beta(a)\, y(t,a)\,\mathrm{d}a,\\
v(t) &\in [0, \overline{v}].
\end{aligned}$$

Here $y(t,a)$ is interpreted as a stock of biological resource of age a and $v(t)$ is the harvesting effort. The mortality rate $\mu(a)$, fertility rate $\beta(a)$, and profit function $p(a)$ are all non-negative, measurable and bounded. The discount rate r is non-negative.

Assumption (A3) is not trivially fulfilled for this model because $\Phi(t,v) = 1$. However, below we show that it is generically non-restrictive.

The fundamental solution X reads as

$$X(t,a) = \exp\left(-\int_0^{\min\{t,a\}} (\mu(a-s) + \hat{v}(t-s))\,\mathrm{d}s\right).$$

Regarding (23), the adjoint functions defined in (26)–(28) take the form

$$\begin{aligned}
\hat{\xi}(t,a) &= \int_a^\omega \Big[X(t+s-a,s)\,\hat{v}(t+s-a)\,p(s) + \hat{\zeta}(t+s-a)\,\beta(s)\Big]\,\mathrm{d}s\,X(t,a)^{-1},\\
\hat{\zeta}(t) &= \xi(t,0).
\end{aligned}$$

Observe that the standard transversality condition $\hat{\xi}(t,\cdot) \to 0$ is not fulfilled unless $\hat{v}(t) \to 0$ as $t \to \infty$.

Consider the function

$$\Theta(\nu) := \int_0^\omega e^{-\nu a} e^{-\int_0^a \mu(s)\mathrm{d}s} \beta(a)\,\mathrm{d}a, \quad \nu \in \mathbf{R}.$$

Lemma 6. *If $\Theta(r) < 1$, Assumption (A3) is fulfilled. If $\Theta(r) > 1$, no WOO solution exists.*

Proof. Case $\Theta(r) < 1$. The functions $X(t,a)$, $\hat{v}(t)$ and $p(a)$ are essentially bounded, thus, $\psi \in L^r_\infty$. Obviously, for any admissible control v it holds that

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} \int_0^\omega |e^{-ra} e^{-\int_0^a (\mu(s)+v(t+s))\mathrm{d}s} \beta(a)|\,\mathrm{d}a < 1,$$

which is condition (31). Thus, Assumption (A3') is fulfilled, therefore Theorem 1 is applicable.

Case $\Theta(r) > 1$. According to Theorem 1.3 and equation (1.13) in [15, Chapter 2], the population grows in the long run with rate ν, for which $\Theta(\nu) = 1$. Since $\Theta(\cdot)$ is decreasing, $\nu > r$. Therefore, there exists a constant control $v > 0$, such that the population is growing with a rate $\nu > r$. This implies (see again [15, Chapter 2]), that for sufficiently large t, we have $y(t,a) \geq \frac{1}{2} y(0,a) e^{\nu t}$. Hence, $\lim_{T\to\infty} J_T(v) \to \infty$. Thus, the objective value is infinite for any WOO control (if such exists).

However, we show now that perpetual "postponement" of harvesting is beneficial in terms of the WOO criterion of optimality, thus no WOO control exists. Indeed, assume that $\hat{v}(t)$ is optimal, and denote by $\hat{y}(t,a)$ the corresponding trajectory. Clearly, $\hat{v}$ is not a.e. identical to zero, since otherwise the objective value will also be zero. Let $\hat{v}$ be not identically zero on the interval $[0,\tau]$. We modify $\hat{v}$ as $v(t) = 0$ for $t \in [0,\tau]$ and $v(t) = \hat{v}(t)$ for $t > \tau$. Then there is a constant c such that $y(\tau,a) \geq (1+c)\,\hat{y}(\tau,a)$ and due to the linear homogeneous structure of the system, this inequality is preserved for all $t \geq \tau$.

For any $T > \tau$

$$J_T(v) - J_T(\hat{v}) = c\int_\tau^T e^{-rt} p(a)\,\hat{v}(t)\,\hat{y}(t,a)\,\mathrm{d}a\,\mathrm{d}t - \int_0^\tau e^{-rt} p(a)\,\hat{v}(t)\,\hat{y}(t,a)\,\mathrm{d}a\,\mathrm{d}t.$$

Since the first integral above tends to infinity with T, the above difference also tends to infinity, which contradicts the WOO of $\hat{v}$. This completes the proof. □

In the borderline case, $\Theta(r) = 1$, it is not clear whether Assumption (A3) holds, but this case involves a relation between the intrinsic population vital rates and the economic discount, which is non-generic.

6.3 Populations of fixed size and optimal age-patterns of immigration

The application potential of the approach presented in this paper goes beyond the particular class of problems considered here.

The papers [13, 19] investigate the issue of optimal age-structured recruitment/immigration policies of organizations/countries, where the goal is to keep the size of the population constant, while optimizing combinations of certain demographic characteristics (such as average age, size of the inflow, dependency ratio). The involved optimal control problems are not covered by the consideration in the present paper due to fact that the aggregated state variable z appears in the distributed state equation (2). As mentioned in the discussions after Theorem 1, this case is technically more difficult. However, the approach utilized for the models in [13, 19] is essentially the same as the one in the present paper. In particular, the verification of an appropriate analog of Assumption (A3) was a key issue accomplished for the specific models in these papers.

Appendix

Below we give the proofs of Lemmas 1–5.

Proof of Lemma 1. For a fixed $t > 0$ the integrand in (27) can be estimated using (28), (A2) and (A3) as follows

$$\begin{aligned} |\psi(\theta)\, R(\theta,t)| &= \left| \int_0^\omega [g_y X(\theta+s,s)\Phi(\theta) + g_z(\theta,s)]\, \mathrm{d}s\, R(\theta,t) \right| \\ &\le \int_0^\omega [\rho(\theta+s)\, |X(\theta+s,s)|\, |\Phi(\theta)| + \rho(\theta)]\, \mathrm{d}s\, |R(\theta,t)| \le \lambda(\theta,t). \end{aligned} \tag{45}$$

The first claim of Lemma 1 follows from the integrability of $\lambda(\cdot,t)$. The local boundedness of $\hat{\zeta}$ follows from the local boundedness of ψ and $\int_t^\infty \lambda(\theta,t)\,\mathrm{d}\theta$.

Now consider $\hat{\xi}$. Denote $\xi_1(t,a) := \int_a^\omega g_y X(t-a+s,s)\,\mathrm{d}s X^{-1}(t,a)$, which is the first term in (22), including the multiplication with X^{-1}. First, we show that ξ_1 is Lipschitz continuous along the characteristic lines $t - a = const$:

$$\hat{\xi}_1(t+\varepsilon,a+\varepsilon) - \hat{\xi}_1(t,a) = \int_{a+\varepsilon}^\omega g_y X(t-a+s,s)\,\mathrm{d}s\, X^{-1}(t+\varepsilon,a+\varepsilon) - \int_a^\omega g_y X(t-a+s,s)\,\mathrm{d}s\, X^{-1}(t,a)$$

$$= \int_{a+\varepsilon}^\omega g_y X(t-a+s,s)\,\mathrm{d}s \left[X^{-1}(t+\varepsilon,a+\varepsilon) - X^{-1}(t,a)\right] + \int_a^{a+\varepsilon} g_y X(t-a+s,s)\,\mathrm{d}s\, X^{-1}(t,a).$$

The Lipschitz continuity follows from the Lipschitz continuity of X^{-1} along the characteristic lines and the local boundedness of $g_y X$ and X^{-1}. Applying the differentiation $\mathcal{D}$ to ξ_1 and using the expression (14) for $\mathcal{D}X^{-1}$, we obtain

$$-\mathcal{D}\xi_1(t,a) = \int_a^\omega g_y X(t-a+s,s)\,\mathrm{d}s\,(X^{-1}(t,a)F(t,a)) + g_y(t,a) = \xi_1(t,a)F(t,a) + g_y(t,a).$$

The proof that the second term, ξ_2, in the definition of $\hat{\xi}$ in (22) is Lipschitz continuous along the characteristic lines $t-a = const$ and satisfies the equation $-\mathcal{D}\xi_2(t,a) = \xi_2(t,a)F(t,a)+\hat{\zeta}(t)H(t,a)$ is similar and therefore omitted. Then $\hat{\xi} = \xi_1 + \xi_2$ belongs to the space $\mathcal{A}(D)$ and satisfies (22).

The definition (27) of $\hat{\zeta}$ has the form (9), with $T = \infty$. We know that $K(t,s) = 0$ for $t \notin [s, s+\omega]$. Moreover, from (45) and (A3) we obtain that the integral in (9) with $T = \infty$ is locally bounded in t. Then we may apply the implication in the end of Section 3.1, which claims that

$$\hat{\zeta}(t) = \psi(t) + \int_t^{\infty} \hat{\zeta}(\theta)\, K(\theta,t)\, \mathrm{d}\theta = \psi(t) + \int_t^{t+\omega} \hat{\zeta}(\theta)\, K(\theta,t)\, \mathrm{d}\theta.$$

Inserting the expressions (20) and (28) for ψ and K, respectively, we obtain that

$$\begin{aligned}\hat{\zeta}(t) &= \left[\int_0^{\omega} g_y X(t+x,x)\, \mathrm{d}x + \int_t^{t+\omega} \hat{\zeta}(\theta)\, HX(\theta,\theta-t)\, \mathrm{d}\theta\right] \Phi(t) + \int_0^{\omega} g_z(t,a)\, \mathrm{d}a \\ &= \hat{\xi}(t,0)\, \Phi(t) + \int_0^{\omega} g_z(t,a),\end{aligned}$$

that is, (23) is fulfilled by $(\hat{\xi}, \hat{\zeta})$. The proof is complete. □

Proof of Lemma 2. We represent

$$\begin{aligned}\Delta_\tau(u_\alpha) &= \int_{\tau-\alpha}^{\tau} \int_0^{\omega} [(g(t,a,y_\alpha,z_\alpha,u_\alpha) - g(t,a,z_\alpha,u_\alpha)) + (g(t,a,z_\alpha,u_\alpha) - g(t,a,u_\alpha)) \\ &\quad + (g(t,a,u_\alpha) - g(t,a))]\, \mathrm{d}a\, \mathrm{d}t \;=:\; I_1 + I_2 + I_3.\end{aligned}$$

We remind that $\|\Delta y\|_{L_\infty(D_\tau)} + \|\Delta z\|_{L_\infty(0,\theta)} \le c_0\, \alpha$ (see (41)). Moreover, $u_\alpha(t,a) \neq \hat{u}(t,a)$ only on the set $B(\tau,b;\alpha)$ (where $u_\alpha(t,a) = u$), and $\tilde{y}_\alpha(t,a) = \hat{y}(t,a)$ except on a set of measure $2\alpha^2$. Using in addition that g is Lipschitz continuous with respect to (y,z) in the domain where (y_α, z_α) and $(\hat{y}, \hat{z})$ take values for $(t,a) \in D_\tau$ we apparently have $|I_1| = o(\alpha^2)$. For I_3 we have

$$I_3 = \iint_{B(\tau,b;\alpha)} (g(t,a,u) - g(t,a))\, \mathrm{d}a\, \mathrm{d}t = \alpha^2 (g(\tau,b,u) - g(\tau,b)) + o(\alpha^2)$$

due to the Lebesgue property of (τ, b), see the beginning of "Part 2 for u" in Section 5. Finally,

$$I_2 = \int_{\tau-\alpha}^{\tau} \int_0^{\omega} (g_z(t,a,u_\alpha(t,a))\, \Delta z(t) + o(\alpha))\, \mathrm{d}a\, \mathrm{d}t = \int_{\tau-\alpha}^{\tau} \int_0^{\omega} g_z(t,a)\, \Delta z(t)\, \mathrm{d}a\, \mathrm{d}t + o(\alpha^2).$$

For Δz we have

$$\begin{aligned}\Delta z(t) &= \int_0^{\omega} \left(H^\sharp(t,a,y_\alpha(t,a),u_\alpha(t,a)) - H^\sharp(t,a)\right)\, \mathrm{d}a = \int_0^{\omega} \left(H^\sharp(t,a,u_\alpha(t,a)) - H^\sharp(t,a)\right)\, \mathrm{d}a + o(\alpha) \\ &= \int_{b-\alpha}^{b} \left(H^\sharp(t,a,u) - H^\sharp(t,a)\right)\, \mathrm{d}a + o(\alpha). \qquad (46)\end{aligned}$$

Inserting this in the expression for I_2 and changing the order of integration we obtain

$$\begin{aligned} I_2 &= \iint_{B(\tau,b;\alpha)} \int_0^\omega g_z(t,s)\,\mathrm{d}s\,(H^\sharp(t,a,u) - H^\sharp(t,a))\,\mathrm{d}a\,\mathrm{d}t + o(\alpha^2) \\ &= \int_0^\omega g_z(\tau,s)\,\mathrm{d}s\,(H^\sharp(\tau,b,u) - H^\sharp(\tau,b)) + o(\alpha^2), \end{aligned}$$

where for the last equality we use the Lebesgue property of (τ, b), see the beginning of "Part 2 for u" in Section 5.

Summing the obtained expressions for I_1, I_2 and I_3 we obtain the claim of the lemma. □

Proof of Lemma 3. We remind the inequalities $2\alpha < \tau$, $2\alpha < b$, and $2\alpha < \omega - b$ posed for α in "Part 2 for u" in Section 5. Observe that $\Delta y(\tau, a) = 0$ for all a except for $a \in [0, \alpha] \cup [b-\alpha, b+\alpha]$. Therefore, we consider the integral in the formulation of the lemma separately on these two intervals.

Beginning with $[0, \alpha]$, we represent

$$\begin{aligned} &\int_0^\alpha \hat\xi(\tau,a)\,\Delta y(\tau,a)\,\mathrm{d}a \\ &\quad= \int_0^\alpha \left[\hat\xi(\tau-a,0) + \int_0^a \mathcal{D}\hat\xi(\tau-a+s,s)\,\mathrm{d}s\right]\left[\Delta y(\tau-a,0) + \int_0^a \mathcal{D}\Delta y(\tau-a+s,s)\,\mathrm{d}s\right]\mathrm{d}a. \end{aligned}$$

We remind that $\|\Delta y\|_{L_\infty(D_\tau)} \le c_0\,\alpha$. For any $t \ge 0$ and $s \in [0, \alpha]$ we have $u_\alpha(t,a) = \hat u(t,a)$, hence $|\mathcal{D}\Delta y(t,s)| = |F(t,s)\,\Delta y(t,s)| \le c\,\alpha$. Moreover, due to Lemma 1 and the local boundedness of F, $\hat\xi$, $\hat\zeta$, H, and g_y we have $\|\mathcal{D}\xi\|_{L_\infty(D_\tau)} \le c$. Hence,

$$\int_0^\alpha \hat\xi(t,a)\,\Delta y(\tau,a)\,\mathrm{d}a = \int_0^\alpha \xi(\tau-a,0)\,\Delta y(\tau-a,0)\,\mathrm{d}a + o(\alpha^2).$$

Since $\Delta y(\tau - a, 0) = \Phi(\tau - a)\,\Delta z(\tau - a)$, using representation (46) and changing the integration variable we obtain that

$$\int_0^\alpha \hat\xi(t,a)\,\Delta y(\tau,a)\,\mathrm{d}a = \iint_{B(\tau,b;\alpha)} \xi(t,0)\,\Phi(t)\,(H^\sharp(t,a,u) - H^\sharp(t,a))\,\mathrm{d}a\,\mathrm{d}t + o(\alpha^2).$$

Due to the Lebesgue point property of (τ, b), the expression in the right-hand side equals the second term in the right hand side in the assertion of the lemma.

Now, we consider $E := \int_{b-\alpha}^{b+\alpha} \hat\xi(\tau,a)\,\Delta y(\tau,a)\,\mathrm{d}a$. For a in the interval of integration

$$\Delta y(\tau,a) = \Delta y(\tau-\alpha,a-\alpha) + \int_0^\alpha \mathcal{D}\Delta y(\tau-x,a-x)\,\mathrm{d}x$$

and the first term in the right-hand side is zero (due to $a - \alpha \ge 0$). Then

$$\begin{aligned} E &= \int_{b-\alpha}^{b+\alpha}\int_0^\alpha \hat\xi(\tau,a)\left[F^\sharp(\tau-x,a-x,y_\alpha(\tau-x,a-x),u_\alpha(\tau-x,a-x)) - F^\sharp(\tau-x,a-x)\right]\mathrm{d}x\,\mathrm{d}a \\ &= \int_{\tau-\alpha}^{\tau}\int_{b-\tau+t-\alpha}^{b-\tau+t+\alpha} \hat\xi(\tau,\tau-t+s)\left[F^\sharp(t,s,u_\alpha(t,s)) - F^\sharp(t,s)\right]\mathrm{d}s\,\mathrm{d}t + o(\alpha^2), \end{aligned}$$

where we passed to the new variables $t = \tau - x$ and $s = a - x$ and used that $\|\Delta y\|_{L_\infty(D_\tau)} \le c_0\,\alpha$. Notice that if $s < b - \alpha$ or $s > b$ the last integrand is zero, since $u_\alpha(t,s) = \hat{u}(t,s)$. Otherwise $u_\alpha(t,s) = u$. Then

$$E = \iint_{B(\tau,b;\alpha)} \hat{\xi}(\tau, \tau - t + s)\ \Big[F^\sharp(t,s,u) - F^\sharp(t,s)\Big]\ \mathrm{d}s\,\mathrm{d}t + o(\alpha^2),$$

Using the Lebesgue property of (τ, b), (see the beginning of "Part 2 for u" in Section 5) we obtain the first term in the right hand side in the assertion of the lemma. □

Proof of Lemma 4. By the definition of v_α we have $|v_\alpha(t) - \hat{v}(t)| \le c\,\alpha$. According to Proposition 2, $\|\Delta y(t,\cdot)\|_{L_1(0,\omega)} \le c\alpha^2$. Moreover,

$$\Delta z(t) = \int_0^\omega [\alpha H_v^\sharp(t,a)(v - \hat{v}(t)) + o(\alpha)]\,\mathrm{d}a.$$

Then we obtain

$$\Delta_\tau(v_\alpha) = \int_{\tau-\alpha}^{\tau}\int_0^\omega [g(t,a,z_\alpha(t)) - g(t,a,v) + g(t,a,v) - g(t,a) + o(\alpha)]\,\mathrm{d}a\,\mathrm{d}t \tag{47}$$

$$= \int_{\tau-\alpha}^{\tau}\int_0^\omega [g_z(t,a)\Delta z(t) + g(t,a,v) - g(t,a) + o(\alpha)]\,\mathrm{d}a\,\mathrm{d}t \tag{48}$$

$$= \int_{\tau-\alpha}^{\tau}\left\{\int_0^\omega g_z(t,s)\,\mathrm{d}s \int_0^\omega \alpha H_v^\sharp(t,a)(v - \hat{v}(t))\,\mathrm{d}a + \int_0^\omega g_v(t,a,v)\alpha(v - \hat{v}(t))\,\mathrm{d}a\right\}\,\mathrm{d}t + o(\alpha^2), \tag{49}$$

which proves the claim of the lemma due to the Lebesgue property of τ (see the beginning of "Part 2 for v" in Section 5). □

Proof of Lemma 5. The proof uses similar arguments as that of Lemma 3 and therefore is somewhat shortened. From (3) we have

$$\Delta y(t,0) = \Phi(t,\hat{v})\Delta z(\tau) + \alpha\Phi_v^\sharp(t)(v - \hat{v}(t)) + o(\alpha).$$

For $t = \tau$ and $a < \alpha$,

$$\Delta y(\tau,a) = \Delta y(\tau - a, 0) + \int_{-a}^0 \mathcal{D}\Delta y(\tau + s, a + s)\,\mathrm{d}s = \Delta y(\tau - a, 0) + o(\alpha).$$

while for $a > \alpha$,

$$\Delta y(\tau,a) = \int_{-\alpha}^0 [F^\sharp(t+s, a+s, y_\alpha, v_\alpha) - F^\sharp(t+s, a+s)]\,\mathrm{d}s$$

$$= \int_{-\alpha}^0 [\alpha F_v^\sharp(t+s, a+s)(v - \hat{v}(t+s)) + o(\alpha)]\,\mathrm{d}s.$$

Consider the integral $\int_0^\omega \hat\xi(\tau,a)\Delta y(\tau,a)\,\mathrm{d}a$ on $[0,\alpha)$. Because $|\Delta y(\tau,a)| \le c\alpha$, and the above representation,

$$\int_0^\alpha \hat\xi(\tau,a)\Delta y(\tau,a)\,\mathrm{d}a = \int_0^\alpha \xi(\tau-a,0)\Delta y(\tau-a,0)\,\mathrm{d}a + o(\alpha^2).$$

Since τ is a Lebesgue point, using the representation for $\Delta y(\tau,0)$ and $\Delta z(\tau)$ from the proof of the Lemma 4 we obtain the expression

$$\alpha^2\hat\xi(\tau,0)\Phi(\tau)\int_0^\omega [H_v^\sharp(\tau,a)(v-\hat v(\tau)) + o(\alpha)]\,\mathrm{d}a + \alpha^2\hat\xi(\tau,0)\Phi_v^\sharp(\tau)(v-\hat v(\tau)) + o(\alpha^2).$$

With the representation for $\Delta y(\tau,a)$ for $a > \alpha$, and the absolute continuity of $\hat\xi$ along the characteristic lines, we obtain

$$\begin{aligned}\int_\alpha^\omega \hat\xi(\tau,a)\Delta y(\tau,a)\,\mathrm{d}a &= \int_\alpha^\omega \int_{\tau-\alpha}^\tau \hat\xi(t,a-\tau+t)\alpha F_v^\sharp(t,a-\tau+t)(v-\hat v(t))\,\mathrm{d}a + o(\alpha^2)\\ &= \alpha\int_\alpha^{\omega-\alpha}\int_{\tau-\alpha}^\tau \hat\xi(t,a)\,F_v^\sharp(t,a)(v-\hat v(t))\,\mathrm{d}t\,\mathrm{d}a + o(\alpha^2)\\ &= \alpha\int_{\tau-\alpha}^\tau\int_0^\omega \hat\xi(t,a)\,F_v^\sharp(t,a)(v-\hat v(t))\,\mathrm{d}a\,\mathrm{d}t + o(\alpha^2).\end{aligned}$$

Since τ is a Lebesgue point, this implies the claim of the lemma. □

References

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

[2] S. M. Aseev and A. V. Kryazhimskii. The pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons. *SIAM J. Control Optim.*, 43:1094–1119, 2004.

[3] S. M. Aseev and V. M. Veliov. Maximum principle for infinite-horizon optimal control problems with dominating discount. *Dynamics of Continuous, Discrete and Impulsive Systems, Series B*, 19:43–63, 2012.

[4] S. M. Aseev and V. M. Veliov. Needle variations in infinite-horizon optimal control. *Proceedings of the Contemporary Mathematics series of the AMS (IMCP/CONM)*, to appear, 2014.

[5] E. Barucci and F. Gozzi. Technology adoption and accumulation in a vintage capital model. *J. of Economics*, 74 No. 1:1–30, 2001.

[6] R. Boucekkine, D. de la Croix, and O. Licandro. *Vintage capital growth theory: Three breakthroughs.* Emerald Group Publishing Limited, 2011.

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

[8] D. A. Carlson, A. B. Haurie, and A. Leizarowitz. *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems*. Springer, Berlin, 1991.

[9] S. Faggian. Applications of dynamic programming to economic problems with vintage capital. *Dynamics of Continuous, Discrete and Impulsive Systems, Series A: Mathematical Analysis*, 15:527–553, 2008.

[10] S. Faggian and F. Gozzi. Dynamic programming for infinite horizon boundary control problems of pdes with age structure. *Journal of Mathematical Economics*, 46 (4):416–437, 2010.

[11] G. Feichtinger, R. F. Hartl, P. M. Kort, and V. M. Veliov. Anticipation effects of technological progress on capital accumulation: a vintage capital approach. *Journal of Economic Theory*, 126(1):143 – 164, 2006.

[12] G. Feichtinger, G. Tragler, and V. M. Veliov. Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications*, 288:47–68, 2003.

[13] G. Feichtinger and V. M. Veliov. On a distributed control problem arising in dynamic optimization of a fixed-size population. *SIAM Journal on Optimization*, 18:980–1003, 2007.

[14] G. Gripenberg, S. O. Londen, and O. Staffans. *Volterra Integral and Functional Equations*. Cambride University Press, UK, 1990.

[15] M. Iannelli. *Mathematical theory of age-structured population dynamics*. Giardini Editori, Pisa, 1994.

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

[17] A. Prskawetz, T. Tsachev, and V. M. Veliov. Optimal education in an age-structured model under changing labor demand and supply. *Macroeconomic Dynamics*, 16:159–183, 2012.

[18] A. Prskawetz and V. M. Veliov. Age-specific dynamic labor demand and human capital investment. *Journal of Economic Dynamics and Control*, 31(12):3741 – 3777, 2007.

[19] C. Simon, B. Skritek, and V. Veliov. Optimal immigration age-patterns in populations of fixed size. *Journal of Mathematical Analysis and Applications*, 405(1):71 – 89, 2013.

[20] V. M. Veliov. Optimal control of heterogeneous systems: Basic theory. *J. Math. Anal. Appl.*, 346:227–242, 2008.

[21] G. Webb. *Theory of nonlinear age-dependent population dynamics*. New York: Marcel Dekker, 1985.