TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

Another View of the Maximum Principle for Infinite-Horizon Optimal Control Problems in Economics

(updated version of Research Report 2017-05)

Sergey M. Aseev, Vladimir M. Veliov

Research Report 2019-05

March 2019

ISSN 2521-313X

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

Another View of the Maximum Principle for Infinite-Horizon Optimal Control Problems in Economics

S. M. Aseev*   V. M. Veliov†

Abstract

We present a recently developed complete version of the Pontryagin maximum principle for a class of infinite-horizon optimal control problems arising in economics. The main peculiarity of the result is that the adjoint variable is explicitly specified by a formula which resembles the Cauchy formula for solutions of linear differential systems. In certain situations this formula implies the "standard" transversality conditions at infinity. Moreover, it can serve as their alternative. We provide examples demonstrating the advantage of the suggested version of the maximum principle. In particular, we consider its applications to Halkin's example, to Ramsey's optimal growth model and to a basic model of optimal extraction of a non-renewable resource. An economic interpretation of the developed characterization of the adjoint variable is also presented.

Keywords: optimal control, infinite horizon, Pontryagin maximum principle, transversality conditions, shadow prices, capital accumulation, Ramsey model, optimal resource extraction.

1 Introduction

To the best of our knowledge, for the first time L.S. Pontryagin has announced his celebrated maximum principle for problems of optimal control at Session of the USSR Academy of Sciences on Scientific Problems

*Steklov Mathematical Institute of RAS, Gubkina str. 8, 119991, Moscow, Russia; International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria; E-mail: aseev@mi-ras.ru

†Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Wiedner Hauptstrasse 8/E105-4, A-1040 Vienna, Austria; E-mail: vladimir.veliov@tuwien.ac.at

of Computer-Aided Manufacturing in 1956 (see [49]). Then the seminal monograph "The mathematical theory of optimal processes" by L.S. Pontryagin and his collaborators, presenting foundations of the new theory, was published in 1961 (see [50]). Starting from this time the theory of optimal control began its rapid development.

As it was indicated already in Pontryagin's lecture [49] the so-called "adjoint variable" plays a main role in the relations of the maximum principle. It allows to determine the values of an optimal control via the maximum condition. The behavior of the adjoint variable is governed by the adjoint system. However, the adjoint system has infinitely many solutions, and its particular solution that corresponds the reference optimal control is determined usually by additional boundary conditions known as "transversality conditions". This explains the role of transversality conditions and suggests a standard way to complete relations of the maximum principle.

In the last decades, the Pontryagin maximum principle was extended to various classes of problems. One of the important classes of optimal control problems to which numerous attempts to extend the maximum principle have been made is a class of infinite-horizon problems arising in economics. Typically, the initial state is fixed and the terminal state (at infinity) is free in such problems, while the utility functional to be maximized is given by an improper integral on the time interval $[0, \infty)$. The last circumstance gives rise to specific mathematical features of the problems. To be specific, let $x_*(\cdot)$ be an optimal trajectory and $(\psi^0, \psi(\cdot))$ be a pair of adjoint variables corresponding to $x_*(\cdot)$ according to the maximum principle. Although the state at infinity is not constrained, such problems could be abnormal (i.e. $\psi^0 = 0$) and the "standard" transversality conditions at infinity of the form $\lim_{t\to\infty} \psi(t) = 0$ or $\lim_{t\to\infty}\langle \psi(t), x_*(t)\rangle = 0$ may fail. The results justifying validity of these relations were obtained only under rather restrictive assumptions (see [21, 35, 43, 45, 53, 56, 64]) that make them inapplicable to many particular economic problems.

Notice, that additional characterization of the adjoint variable $\psi(\cdot)$ is critically important for the efficient use of the maximum principle, because in the general case, without complementary conditions, the set of extremals satisfying the maximum principle may be "too wide." Meanwhile, a number of known examples (see [11, 26, 38, 45]) clearly demonstrate that complementary conditions for the adjoint variable which differ from the standard transversality conditions have to be involved.

The aim of this paper is to present the recent results by the authors, developing another view of the Pontryagin optimality conditions for infinite-horizon optimal control problems arising in economics especially with regard to the correct determination of the adjoint variable $\psi(\cdot)$.

The main feature of the developed version of the maximum principle is that the adjoint variable is explicitly specified by a formula which resembles the Cauchy formula for solutions of linear differential systems. In certain situations this formula implies the standard transversality conditions at infinity. Moreover, it can serve as their alternative. Another important feature of the developed version of the maximum principle is that it is proved under weak regularity assumptions. This makes possible to directly apply it to some meaningful economic models. A third feature of the proposed approach is that it is also applicable to problems in which infinite objective values may appear. In this case the concept of "overtaking optimality" is employed, which is important for many economic considerations.

The paper is organized as follows. In Section 2 we give a strict formulation of the problem, introduce the notion of optimality used in the paper, and formulate and discuss our main result – the normal form version of the maximum principle with explicitly specified solution of the adjoint equation. Here we also present two illustrative examples. The first example is a classical Halkin's example in which the standard transversality conditions at infinity are violated while all optimal controls in the problem are determined by the developed explicit representation of the adjoint variable. The second example clarifies the role of our main growth assumption. Together, these two examples demonstrate the alternative character of the developed characterization of the adjoint variable against the standard transversality conditions at infinity.

In Section 3 we specialize the main result for several classes of problems in terms of growth rates of the functions involved. This allows us to formulate conditions implying the standard transversality conditions at infinity in terms of the growth rates.

Section 4 is devoted to applications of our main result in economics. Here the economic meaning of the developed Cauchy type formula is discussed in detail. In the case of autonomous problems with discounting we establish a link between the Cauchy type formula and Michel's asymptotic condition for the Hamiltonian (see [45]). Then we demonstrate applications of our main result to two meaningful economic problems – to Ramsey's optimal growth model and to a basic model of optimal extraction of a non-renewable resource. The Ramsey model is the most important theoretical construct in modern growth theory. The analysis of Ramsey's model presented in economic literature is usually based on the assumption that the standard transversality condition holds as a necessary condition for optimality. However, as a rule, this fact is not rigorously justified (see, for example [18, Section 2.6]). Here we present a rigorous analysis of the Ramsey model based on application of the developed version of the maximum principle. The basic model of extraction of a non renewable resource provides an example of an infinite-horizon optimal control problem in economics in which the standard transversality conditions at infinity fail while a correct characterization of the adjoint variable is provided by the Cauchy type formula.

We conclude the paper with a brief bibliographical survey provided in Section 5.

2 Statement of the problem and the main result

2.1 Statement of the problem

Let G be a nonempty open convex subset of $\mathbb{R}^n$ and let

$$f : [0,\infty) \times G \times \mathbb{R}^m \to \mathbb{R}^n \quad \text{and} \quad f^0 : [0,\infty) \times G \times \mathbb{R}^m \to \mathbb{R}^1$$

be given functions.

Consider the following optimal control problem (P):

$$J(x(\cdot), u(\cdot)) = \int_0^\infty f^0(t, x(t), u(t))\, dt \to \max\,, \tag{1}$$

$$\dot{x}(t) = f(t, x(t), u(t)), \qquad x(0) = x_0, \tag{2}$$

$$u(t) \in U(t). \tag{3}$$

Here $x(t) = (x^1(t), \dots, x^n(t)) \in \mathbb{R}^n$ is a state vector, $u(t) = (u^1(t), \dots, u^m(t)) \in \mathbb{R}^m$ is a control vector at time $t \geq 0$, $x_0 \in G$ is a fixed initial state and $U\colon [0,\infty) \rightrightarrows \mathbb{R}^m$ is a multivalued mapping with nonempty values.

Infinite-horizon optimal control problems of type (P) arise in different fields of economics, in particular, in the theory of economic growth [18]. Typically, in economic applications the components of vector $x(t)$ can be interpreted as values of various capital stocks while the components of vector $u(t)$ as values of different kinds of investments at time $t \geq 0$.

To the best of our knowledge Ramsey [51] was the first who presented, in the 1920s, the problem of optimization of economic growth as a variational problem of maximizing an integral functional on an infinite time horizon. This line of research was continued by Cass [27], Koopmans [44], Shell [58], Arrow and Kurz [3], and become the standard method of solving optimal economic growth models. Nevertheless, the theory of first order necessary optimality conditions for infinite-horizon problems is still less developed than that in the finite-horizon case.

It is well known, that the proper choice of the present value "shadow prices function" (adjoint variable) $\psi(\cdot)$ along the optimal trajectory $x_*(\cdot)$ plays a crucial role in the identification of the corresponding optimal investment policy $u_*(\cdot)$ in problem (P). Indeed, if such a function $\psi\colon [0,\infty)\mapsto \mathbb{R}^n$ is known[1], then the optimal investment policy $u_*(\cdot)$ can be determined by maximization of the instantaneous net present value utility on the time interval $[0,\infty)$:

$$\begin{aligned} &f^0(t,x_*(t),u_*(t)) + \langle \psi(t), f(t,x_*(t),u_*(t))\rangle \\ &\qquad \stackrel{\text{a.e.}}{=} \sup_{u\in U(t)} \Big\{ f^0(t,x_*(t),u) + \langle \psi(t), f(t,x_*(t),u)\rangle \Big\}. \quad (4) \end{aligned}$$

Here the first term $f^0(t,x_*(t),u_*(t))$ in the left-hand side represents the present value utility flow, while the second term $\langle \psi(t), f(t,x_*(t),u_*(t))\rangle$ is the present value increment of the capital stock $x_*(t)$ at instant $t\geq 0$. Thus, one can say that (P) is, in fact, a problem of determining an appropriate function of shadow prices $\psi(\cdot)$.

Notice, that for the finite-horizon counterpart of problem (P) the Pontryagin maximum principle provides a unique function $\psi(\cdot)$ for which the *maximum condition* (4) holds.

Let us recall this classical result in optimal control theory [50]. Define the Hamilton-Pontryagin function $\mathcal{H} : [0,\infty)\times G\times \mathbb{R}^m\times\mathbb{R}^1\times\mathbb{R}^n\to\mathbb{R}^1$ for problem (P) in the standard way:

$$\mathcal{H}(t,x,u,\psi^0,\psi) = \psi^0 f^0(t,x,u) + \langle \psi, f(t,x,u)\rangle,$$

$$t\in[0,\infty),\ x\in G,\ u\in\mathbb{R}^m,\ \psi^0\in\mathbb{R}^1,\ \psi\in\mathbb{R}^n.$$

In the normal case, i.e. when $\psi^0 = 1$, we will omit ψ^0 and write simply $\mathcal{H}(t,x,u,\psi)$ instead of $\mathcal{H}(t,x,u,1,\psi)$.

Now, consider the following problem (P_T) on a fixed finite time interval $[0,T]$, $T>0$:

$$J_T(x(\cdot),u(\cdot)) = \int_0^T f^0(t,x(t),u(t))\,dt \to \max\,,$$

$$\dot{x}(t) = f(t,x(t),u(t)), \qquad x(0)=x_0,$$

[1] Recall that in economics "shadow price" of capital is equal to the present discounted value of future marginal products (see [18, Chapter 2]).

$$u(t) \in U(t).$$

Here all the data in problem (P_T) are the same as in (P). The only difference is that the problem (P_T) is considered on the finite time interval $[0, T]$. Then the Pontryagin maximum principle asserts that, on suitable regularity assumptions, for any optimal admissible pair $(x_T(\cdot), u_T(\cdot))$ in problem (P_T) there is an absolutely continuous function (adjoint variable) $\psi_T \colon [0, T] \mapsto \mathbb{R}^n$ for which the maximum condition (4) is fulfilled. This function $\psi_T(\cdot)$ is defined uniquely as a solution of the normal form *adjoint system*

$$\dot{\psi}(t) = -\mathcal{H}_x(t, x_T(t), u_T(t), \psi(t)) \tag{5}$$

with boundary condition

$$\psi(T) = 0. \tag{6}$$

Condition (6) is known in the optimal control theory as the *transversalty condition for free terminal state.* This condition properly identifies the function $\psi_T(\cdot)$ among all functions $\psi(\cdot)$ which satisfy together with $(x_T(\cdot), u_T(\cdot))$ the *core conditions* of the maximum principle: adjoint system (5) and maximum condition (4) (with the subscript $*$ replaced with T) .

This result motivated numerous attempts to extend the maximum principle for problem (P_T) to the infinite-horizon problem (P), by involving "natural" counterparts of transversality condition (6), in particular of the form

$$\lim_{t\to\infty} \psi(t) = 0, \tag{7}$$

or

$$\lim_{t\to\infty} \langle \psi(t), x_*(t) \rangle = 0. \tag{8}$$

However, the positive results in this direction were obtained under additional conditions that make them inapplicable to many particular problems (see bibliographical comments in Section 5). Moreover, as it was pointed out by Halkin [38] by means of counterexamples, although the state at infinity is not constrained in problem (P), such problems could be *abnormal* (that is, $\psi^0 = 0$) and complementary conditions of the form (7) or (8) may fail to be fulfilled for the "right" adjoint function for which the core conditions of the maximum principle hold.

We should also mention another asymptotic condition on the adjoint variable of the form

$$\lim_{t\to\infty} H(t, x_*(t), \psi^0, \psi(t)) = 0, \tag{9}$$

proved by Michel (see [45]) in the specific case when the problem (P) is autonomous with discounting, i.e. $f(t,x,u) \equiv f(x,u)$, $f^0(t,x,u) = e^{-\rho t} g(x,u)$, $x \in G$, $u \in U(t) \equiv U$, $t \geq 0$, the discount rate ρ is an arbitrary real number (not necessary positive), and the optimal value of the functional is finite. Here $H(t, x_*(t), \psi^0, \psi(t)) = \sup_{u \in U} \mathcal{H}(t, x_*(t), u, \psi^0, \psi(t))$ is the Hamiltonian. Condition (9) is similar to the *transversality in time condition*

$$H(T, x_T(T), \psi^0, \psi(T)) = 0$$

which is well known for finite-horizon problems with free terminal time $T > 0$ (see [50]).

Let us return to conditions (5) and (6) for the adjoint variable in the finite-horizon problem (P_T). It is easy to see, that due to the Cauchy formula for linear differential systems (see [39]) the adjoint system (5) and the transversality condition (6) result in the following representation:

$$\psi(t) = Z_T(t) \int_t^T [Z_T(s)]^{-1} f_x^0(s, x_T(s), u_T(s))\, ds, \qquad t \in [0,T]. \tag{10}$$

Here $Z_T(\cdot)$ is the normalized at instant $t = 0$ fundamental matrix solution on $[0,T]$ of the linear system

$$\dot{z}(t) = -\left[f_x(t, x_T(t), u_T(t))\right]^* z(t). \tag{11}$$

This means that the columns of the matrix function $Z_T(\cdot)$ are linearly independent solutions of (11) on $[0,T]$ while $Z_T(0) = I$, where I is the identity matrix.

The pointwise representation (10) suggests a "natural" candidate for appropriate adjoint function $\psi(\cdot)$ in problem (P). Indeed, substituting $(x_T(\cdot), u_T(\cdot))$ with $(x_*(\cdot), u_*(\cdot))$ in (10) and formally passing to the limit with $T \to +\infty$ (which can be justified under appropriate conditions that guarantee convergence of the integral in the next line) we obtain the expression

$$\psi(t) = Z_*(t) \int_t^\infty [Z_*(s)]^{-1} f_x^0(s, x_*(s), u_*(s))\, ds, \qquad t \geq 0, \tag{12}$$

where now $Z_*(\cdot)$ is the normalized at instant $t = 0$ fundamental matrix solution of the linear system

$$\dot{z}(t) = -\left[f_x(t, x_*(t), u_*(t))\right]^* z(t), \qquad t \in [0, +\infty).$$

Clearly, in the infinite-horizon case, formula (12) is a straightforward analog of the Cauchy formula (10) for the adjoint function in problem (P_T).

However, the formula (12) can not be reduced to asymptotic conditions (7) or (8). This explicit formula does not assume or imply the standard transversality conditions (7) or (8), which may be inconsistent with the core conditions of the maximum principle. However, for particular classes of problems it may imply (7) and/or (8), as seen below in the paper. If the problem (P) is autonomous with discounting and the optimal value of the functional is finite then the explicit formula (12) can imply also the asymptotic condition (9). It turns out that the Cauchy type formula (12) can be justified as a part of necessary optimality condition for problem (P) on mild regularity and growth assumptions, and it can serve as an alternative to conditions (7) and (8).

Now let us specify the statement of problem (P).

The following assumption is standing throughout the paper and will not be explicitly mentioned in the text below.

Assumption $(A0)$**.** For almost every $t \in [0, \infty)$ the derivatives $f_x(t, x, u)$ and $f^0_x(t, x, u)$ exist for all $(x, u) \in G \times \mathbb{R}^m$ and the functions $f(\cdot, \cdot, \cdot)$, $f^0(\cdot, \cdot, \cdot)$, $f_x(\cdot, \cdot, \cdot)$, and $f^0_x(\cdot, \cdot, \cdot)$ are Lebesgue-Borel (LB) measurable in (t, u) for every $x \in G$, and continuous in x for almost every $t \in [0, \infty)$ and every $u \in \mathbb{R}^m$. The multivalued mapping $U(\cdot)$ is LB-measurable.

The LB-measurability in (t, u) [30, Definition 6.33] means that the functions (and the sets) to which the property applies are measurable in the σ-algebra generated by the Cartesian product of the Lebesgue σ-algebra on $[0, \infty)$ and the Borel σ-algebra on $\mathbb{R}^m$. An important property is that for any LB measurable function $g : [0, \infty) \times \mathbb{R}^m \to \mathbb{R}^n$, the superposition $t \mapsto g(t, u(t))$ with a Lebesgue measurable function $u : [0, \infty) \to \mathbb{R}^m$ is Lebesgue measurable [30, Proposition 6.34]. The LB-measurability of the multivalued mapping $U(\cdot)$ means that the set $\operatorname{graph} U(\cdot) = \{(t, u) \in [0, \infty) \times \mathbb{R}^m \colon u \in U(t)\}$ is a LB-measurable subset of $[0, \infty) \times \mathbb{R}^m$.

Remark 1. *In some situations it is natural (and convenient) to consider problems* (P) *with functions* $f(\cdot, \cdot, \cdot)$ *and* $f^0(\cdot, \cdot, \cdot)$ *which are defined only for* $(t, u) \in \operatorname{graph} U(\cdot)$, *where* $U(\cdot)$ *is a LB-measurable multivalued mapping. In this case the LB-measurability of* $f(\cdot, \cdot, \cdot)$ *and* $f^0(\cdot, \cdot, \cdot)$ *in* (t, u) *means that these functions are measurable in the relative* σ*-algebra in* $\operatorname{graph} U(\cdot)$ *induced by the* σ*-algebra of all LB-measurable subsets of* $[0, \infty) \times \mathbb{R}^m$*. It is equivalent to the LB-measurability of functions* $f(\cdot, \cdot, \cdot)$ *and* $f^0(\cdot, \cdot, \cdot)$ *extended as arbitrary constants from* $\operatorname{graph} U(\cdot)$ *to* $[0, \infty) \times \mathbb{R}^m$ *for all* $x \in G$.

We consider any Lebesgue measurable function $u \colon [0, \infty) \mapsto \mathbb{R}^m$ satisfying condition (3) for all $t \geq 0$ as a *control*. If $u(\cdot)$ is a control then the corresponding *trajectory* $x(\cdot)$ is a locally absolutely continuous solution of the initial value problem (2), which (if it exists) is defined in G on some (maximal) finite or infinite time interval $[0, \tau)$, $\tau > 0$. The local absolute continuity of $x(\cdot)$ means that $x(\cdot)$ is absolutely continuous on any compact subinterval $[0, T]$ of its domain of definition $[0, \tau)$.

By definition, a pair $(x(\cdot), u(\cdot))$, where $u(\cdot)$ is a control and $x(\cdot)$ is the corresponding trajectory, is an *admissible pair* in problem (P) if the trajectory $x(\cdot)$ is defined on the whole time interval $[0, \infty)$ and the function $t \mapsto f^0(t, x(t), u(t))$ is locally integrable on $[0, \infty)$ (i.e. integrable on any finite time interval $[0, T]$, $T > 0$). Thus, for any admissible pair $(x(\cdot), u(\cdot))$ and any $T > 0$ the integral

$$J_T(x(\cdot), u(\cdot)) := \int_0^T f^0(t, x(t), u(t))\, dt$$

is finite. If $(x(\cdot), u(\cdot))$ is an admissible pair we refer to $u(\cdot)$ as *admissible control* and to $x(\cdot)$ as the corresponding *admissible trajectory*.

Now we recall two basic concepts of optimality used in the literature (see e.g. [26]).

In the first one, the integral in (1) is understood in improper sense, i.e. for an arbitrary admissible pair $(x(\cdot), u(\cdot))$, by definition

$$J(x(\cdot), u(\cdot)) = \lim_{T\to\infty} \int_0^T f^0(t, x(t), u(t))\, dt,$$

if the limit exists.

Definition 2. *An admissible pair $(x_*(\cdot), u_*(\cdot))$ is called* strongly optimal *in problem (P) if the corresponding integral in (1) converges (to a finite number) and for any other admissible pair $(x(\cdot), u(\cdot))$ we have*

$$J(x_*(\cdot), u_*(\cdot)) \geq \limsup_{T\to\infty} \int_0^T f^0(t, x(t), u(t))\, dt.$$

In the second definition, the integral in (1) is not necessary finite.

Definition 3. *An admissible pair $(x_*(\cdot), u_*(\cdot))$ is called* finitely optimal *in problem (P) if for any $T > 0$ this pair (restricted to $[0, T]$) is optimal in the following optimal control problem (Q_T) with fixed initial and final states:*

$$J_T(x(\cdot), u(\cdot)) = \int_0^T f^0(t, x(t), u(t))\, dt \to \max,$$

$$\begin{aligned} \dot{x}(t) = f(t, x(t), u(t)), \quad & x(0) = x_0, \ \ x(T) = x_*(T), \\ u(t) \in U(t). & \end{aligned}$$

It is easy to see that the strong optimality implies the finite one.

The following weak regularity assumption plays a key role for the validity of the general version of the Pontryagin maximum principle for a finitely optimal pair $(x_*(\cdot), u_*(\cdot))$ in problem (P) (similar assumptions for problems with finite time-horizons one can find in [29, Chapter 5], [30, Hypothesises 22.25]).

Assumption $(A1)$**.** *There exist a continuous function* $\gamma\colon [0,\infty) \mapsto (0,\infty)$ *and a locally integrable function* $\varphi\colon [0,\infty) \mapsto \mathbb{R}^1$*, such that* $\{x\colon \|x - x_*(t)\| \leq \gamma(t)\} \subset G$ *for all* $t \geq 0$*, and for almost all* $t \in [0,\infty)$ *we have*

$$\max_{\{x\colon \|x-x_*(t)\|\leq\gamma(t)\}} \Big\{ \|f_x(t,x,u_*(t))\| + \|f^0_x(t,x,u_*(t))\| \Big\} \leq \varphi(t). \quad (13)$$

Notice, that if $(x_*(\cdot), u_*(\cdot))$ is an admissible pair and $(A1)$ holds, then $x_*(\cdot)$ is a unique trajectory that corresponds to $u_*(\cdot)$ (see [36, Chapter 1, Theorem 2]).

Remark 4. *Assumption* $(A1)$ *is automatically fulfilled under the usual regularity conditions that* $u_*(\cdot) \in L^\infty_{loc}[0,\infty)$*,* $U(t) \equiv U$*,* $t \geq 0$*, the functions* $f_x(\cdot,\cdot,\cdot)$ *and* $f^0_x(\cdot,\cdot,\cdot)$ *are measurable in* t*, continuous in* (x,u)*, and locally bounded. Here the local boundedness of these functions of* t*,* x *and* u *(take* $\phi(\cdot,\cdot,\cdot)$ *as a representative) means that for every* $T > 0$*, every compact* $D \subset G$ *and every bounded set* $V \subset U$ *there exists* M *such that* $\|\phi(t,x,u)\| \leq M$ *for almost all* $t \in [0,T]$*, and all* $x \in D$ *and* $u \in V$*.*

If Assumption $(A1)$ is fulfilled, then any finitely optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ satisfies the following general version of the maximum principle, which is proved in [38] under the *standard regularity conditions*. Namely, the proof given in [38] is valid if $u_*(\cdot) \in L^\infty_{loc}[0,\infty)$, $U(t) \equiv U$, $t \geq 0$, and the functions $f(\cdot,\cdot,\cdot)$, $f^0(\cdot,\cdot,\cdot)$, $f_x(\cdot,\cdot,\cdot)$ and $f^0_x(\cdot,\cdot,\cdot)$ are measurable in t, continuous in (x,u), and locally bounded.

Theorem 5. *Let* $(x_*(\cdot), u_*(\cdot))$ *be a finitely optimal admissible pair in problem* (P) *and let* $(A1)$ *be fulfilled. Then there is a non-vanishing pair of adjoint variables* $(\psi^0, \psi(\cdot))$*, with* $\psi^0 \geq 0$ *and a locally absolutely continuous* $\psi(\cdot) : [0,\infty) \to \mathbb{R}^n$*, such that the core conditions of the maximum principle hold, i.e.*

(i) $\psi(\cdot)$ *is a solution to the adjoint system*

$$\dot{\psi}(t) = -\mathcal{H}_x(t, x_*(t), u_*(t), \psi^0, \psi(t));$$

(ii) the maximum condition takes place:

$$\mathcal{H}(t, x_*(t), u_*(t), \psi^0, \psi(t)) \stackrel{a.e.}{=} \sup_{u \in U(t)} \mathcal{H}(t, x_*(t), u, \psi^0, \psi(t)).$$

The main points in the proof of this theorem are essentially the same as in Halkin's original result (see [38, Theorem 4.2]). The needed elaboration is given in [16].

Theorem 5 provides the most general known version of the maximum principle for problem (P). However, Theorem 5 establishes only the core conditions (i) and (ii) of the maximum principle without any additional characterizations of adjoint variables $(\psi^0, \psi(\cdot))$. Due to this circumstance, the relations of Theorem 5 are incomplete, and, as a rule, application of Theorem 5 to particular problems is ineffective (see discussion in Example 1 bellow). To complete the relations of Theorem 5 we need to employ a stronger concept of optimality and to use an additional growth assumption.

The following concept of optimality appears to be the most useful among the numerous alternative definitions proposed in the context of economics. It takes intermediate place between finite optimality and strong optimality (see [26]).

Definition 6. *The admissible pair* $(x_*(\cdot), u_*(\cdot))$ *is called* weakly overtaking optimal *if for arbitrary* $\varepsilon > 0$, $T > 0$ *and any other admissible pair* $(x(\cdot), u(\cdot))$ *there is a* $T' > T$ *such that*

$$\int_0^{T'} f^0(t, x_*(t), u_*(t))\, dt \geq \int_0^{T'} f^0(t, x(t), u(t))\, dt - \varepsilon.$$

The following growth assumption for an admissible pair $(x_*(\cdot), u_*(\cdot))$ was introduced in [15] as an extension of the so-called *dominating discount condition*, see [6, 9, 10, 11, 14, 17].

Assumption $(A2)$: *There exist a number* $\beta > 0$ *and an integrable function* $\lambda : [0, \infty) \mapsto \mathbb{R}^1$ *such that for every* $\zeta \in G$ *with* $\|\zeta - x_0\| < \beta$ *equation* (2) *with* $u(\cdot) = u_*(\cdot)$ *and initial condition* $x(0) = \zeta$ *(instead of* $x(0) = x_0$*) has a solution* $x(\zeta; \cdot)$ *in* G *on* $[0, \infty)$*, and*

$$\max_{x \in [x(\zeta;t), x_*(t)]} \left| \langle f^0_x(t, x, u_*(t)), x(\zeta; t) - x_*(t) \rangle \right| \stackrel{a.e.}{\leq} \|\zeta - x_0\| \lambda(t).$$

Here $[x(\zeta; t), x_*(t)]$ *is the line segment between the points* $x(\zeta; t)$ *and* $x_*(t)$.

Observe that according to the Lipschitz dependence of the solution $x(\zeta; t)$ on the initial condition ζ, we have an inequality $\|x(\zeta; t) - x_*(t)\| \leq l(t)\|\zeta -$

$x_0\|$, where $l(\cdot)$ is independent of ζ. The function $\lambda(\cdot)$ incorporates the growth of this Lipschitz constant and the growth of the derivative of the objective integrand around the reference pair. The real assumption is actually that $\lambda(\cdot)$ is integrable.

Notice, that the constant $\beta > 0$ and the integrable function $\lambda(\cdot)$ may depend on the reference admissible pair $(x_*(\cdot), u_*(\cdot))$ in $(A2)$. In some cases Assumption $(A2)$ can be a priory justified for all optimal (or even for all admissible) pairs $(x_*(\cdot), u_*(\cdot))$ in (P) together with their own constants β and functions $\lambda(\cdot)$ (see the examples in Section 4).

The following auxiliary result (see [16, Lemma 3.2]) implies that the integral in (12) is finite.

Lemma 7. *Let* $(A1)$ *and* $(A2)$ *be satisfied. Then the following estimation holds:*

$$\left\| [Z_*(t)]^{-1} f_x^0(t, x_*(t), u_*(t)) \right\| \leq \sqrt{n}\lambda(t) \quad \text{for a.e.} \quad t \geq 0. \tag{14}$$

Due to Lemma 7 and the integrability of $\lambda(\cdot)$, the function $\psi : [0, \infty) \to \mathbb{R}^n$ defined by formula (12) is locally absolutely continuous. By a direct differentiation we verify that the so defined function $\psi(\cdot)$ satisfies on $[0, \infty)$ the adjoint system

$$\dot{\psi}(t) = -\mathcal{H}_x(t, x_*(t), u_*(t), \psi(t)).$$

2.2 Main result

The next version of the Pontryagin maximum principle for the infinite-horizon problem (P) is the main result in this paper.

Theorem 8. *Let* $(x_*(\cdot), u_*(\cdot))$ *be a weakly overtaking optimal pair in problem* (P). *Assume that the regularity Assumption* $(A1)$ *and the growth Assumption* $(A2)$ *are satisfied. Then the vector function* $\psi : [0, \infty) \mapsto \mathbb{R}^n$ *defined by (12) is (locally) absolutely continuous and satisfies the core conditions of the normal form maximum principle, i.e.*

(i) $\psi(\cdot)$ *is a solution to the adjoint system*

$$\dot{\psi}(t) = -\mathcal{H}_x(t, x_*(t), u_*(t), \psi(t)), \tag{15}$$

(ii) the maximum condition takes place:

$$\mathcal{H}(t, x_*(t), u_*(t), \psi(t)) \overset{a.e.}{=} \sup_{u \in U(t)} \mathcal{H}(t, x_*(t), u, \psi(t)). \tag{16}$$

The rather technical proof of Theorem 8 is presented in detail in [16].

An important feature of Theorem 8 is that it is proved under weak regularity assumptions. This makes possible to directly apply it to some meaningful economic models. The admissible controls $u(\cdot)$ are not necessary bounded (even in a local sense), and the functions $f(\cdot,\cdot,\cdot)$ and $f^0(\cdot,\cdot,\cdot)$ are not necessary continuous in the variable u. Instead, we assume that the functions $f(\cdot,\cdot,\cdot)$, $f^0(\cdot,\cdot,\cdot)$ and the multivalued mapping $U(\cdot)$ are Lebesgue-Borel (LB) measurable in the variables (t,u), while their partial derivatives $f_x(\cdot,\cdot,u_*(\cdot))$ and $f_x^0(\cdot,\cdot,u_*(\cdot))$ are locally integrally bounded in a (t,x)-tube around the graph of the reference optimal trajectory $x_*(\cdot)$. From practical point of view (especially in economics) considering LB-measurable data $f(\cdot,\cdot,\cdot)$ and $f^0(\cdot,\cdot,\cdot)$ is important, since it allows for discontinuity with respect to the control. Such a discontinuity appears, for example, if fixed costs of the control are present whenever the control (i.e. maintenance) is positive, which jump down to zero if zero control is applied. The unboundedness of admissible controls $u(\cdot)$ allows one to treat some economic problems (such as problems of optimal exploitation of renewable or non-renewable resources) in their most natural settings when the rate of extraction of the resource satisfies only an integral constraint in an L-space.

Notice, that the maximum principle ("extended maximum principle") for finite horizon problems with LB-measurable data $f(\cdot,\cdot,\cdot)$, $f^0(\cdot,\cdot,\cdot)$ and not necessary bounded admissible controls $u(\cdot)$ was established by Clarke (see [29, 30]) by methods of the nonsmooth analysis. Our proof of Theorem 8 employs a modification of the classical needle variations technique and makes a substantial use of the Yankov-von Neumann-Aumann selection theorem [41, Theorem 2.14]. The usage of simple needle variations allows us to treat the case of the unbounded time interval $[0,\infty)$.

Another useful feature of our main result is that it applies to problems with infinite objective integral (1), where the notion of *overtaking optimality* is adopted.

It is easy to see that on the assumptions of Theorem 8 together with the additional assumption that $\|Z_*(t)\| \leq c$ for some constant $c \geq 0$ and all sufficiently large t, the formula (12) immediately implies the "standard" asymptotic condition (7). In Section 3 we consider some other situations when the formula (12) implies the "standard" asymptotic conditions (7) and (8). In Section 4 we establish a link between (12) and (9).

The convergence of the integral in (12) apparently implies that the adjoint function $\psi(\cdot)$ defined in (12) satisfies the asymptotic relation

$$\lim_{t\to+\infty} [Z_*(t)]^{-1} \psi(t) = 0. \tag{17}$$

Even more, it is straightforward to prove that on the assumptions of Theorem 8, the function $\psi(\cdot)$ defined by (12) is the only solution of the adjoint equation (15) that satisfies (17). Indeed, if $\psi(\cdot)$ and $\tilde{\psi}(\cdot)$ are two solutions of (15) satisfying (17) then

$$\frac{d}{dt}(\psi(t) - \tilde{\psi}(t)) = -\left[f_x(t, x_*(t), u_*(t))\right]^* (\psi(t) - \tilde{\psi}(t)).$$

From this for every $t \geq 0$ we get

$$\psi(0) - \tilde{\psi}(0) = [Z_*(t)]^{-1} (\psi(t) - \tilde{\psi}(t)).$$

Since the right-hand side converges to zero with $t \to +\infty$, we obtain that $\psi(0) - \tilde{\psi}(0) = 0$, which implies that $\psi(\cdot) = \tilde{\psi}(\cdot)$.

Another direct corollary of the convergence of the integral in (12) is the equality

$$\psi(0) = \int_0^\infty [Z_*(s)]^{-1} f_x^0(s, x_*(s), u_*(s))\, ds. \tag{18}$$

Then, since $\psi(\cdot)$ is a solution to the linear differential system (15) we have

$$\psi(t) = Z_*(t)\psi(0) - Z_*(t) \int_0^t [Z_*(s)]^{-1} f_x^0(s, x_*(s), u_*(s))\, ds, \quad t \geq 0. \tag{19}$$

Obviously, (18) and (19) imply (12). Thus, on the assumptions of Theorem 8, the function $\psi(\cdot)$ defined by (12) is the only solution of the adjoint equation (15) that satisfies the initial condition (18).

2.3 Two illustrative examples

Now, we present two illustrative examples. The first example is the original Halkin's example [38, Section 5]. It demonstrates the completeness of conditions of Theorem 8, and the advantage of formula (12) in comparison with asymptotic conditions (7), (8) and (9). The second example clarifies the role of Assumption $(A2)$ in Theorem 8. Together, these examples illustrate the alternative character of the Cauchy type formula (12) against the standard transversality conditions (7) and (8).

Example 1 (Halkin's example). Consider the following problem $(P1)$:

$$J(x(\cdot), u(\cdot)) = \int_0^\infty (1 - x(t))u(t)\, dt \to \max\,, \tag{20}$$

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

$$u(t) \in [0, 1].$$

This example is interesting, since it shows that the standard asymptotic conditions (7) and (8) are inconsistent with the core condition (i) and (ii) of the maximum principle (see Theorem 5), while the asymptotic condition (9) is not productive in this case. Let us clarify this statement.

For any $T > 0$ and for an arbitrary admissible pair $(x(\cdot), u(\cdot))$ we have

$$J_T(x(\cdot), u(\cdot)) = \int_0^T \dot{x}(t)\, dt = x(T) = 1 - e^{-\int_0^T u(t)\, dt}. \tag{21}$$

This implies that an admissible pair $(x_*(\cdot), u_*(\cdot))$ is weakly overtaking optimal (also strongly optimal)[2] if and only if $\int_0^\infty u_*(t)\, dt = +\infty$. Also note that $x_*(t) \to 1$ with $t \to +\infty$.

According to Halkin's Theorem 5, any optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ satisfies, together with the adjoint function $\psi(\cdot)$, the adjoint equation in (i) and the maximum condition in (ii), which read in the particular case as

$$\dot{\psi}(t) = (\psi(t) + \psi^0)u_*(t), \tag{22}$$

$$(1 - x_*(t))(\psi(t) + \psi^0)\, u_*(t) \overset{\text{a.e.}}{=} \max_{u \in [0,1]} \{(1 - x_*(t))(\psi(t) + \psi^0)\, u\} \tag{23}$$

with some $\psi^0 \geq 0$. From the adjoint equation (22) we obtain that

$$\psi(t) = (\psi(0) + \psi^0)\, e^{\int_0^t u_*(s)\, ds} - \psi^0.$$

Thus, for all $t \geq 0$ and $u \in [0, 1]$ we have

$$\mathcal{H}(t, x_*(t), u, \psi^0, \psi(t)) = (1 - x_*(t))(\psi(t) + \psi^0)u = (\psi^0 + \psi(0))u.$$

Since from (23) we have $(\psi^0 + \psi(0))u_*(t) \overset{\text{a.e.}}{=} \max_{u \in [0,1]} (\psi^0 + \psi(0))u$, and any strongly optimal control $u_*(\cdot)$ is not identically zero, we must have $\psi^0 + \psi(0) \geq 0$. If $\psi^0 + \psi(0) = 0$ then without loss of generality we can set $\psi^0 = 1$ and $\psi(0) = -1$. Then both conditions (7) and (8) are obviously violated. If $\psi^0 + \psi(0) > 0$ then $\psi(t) \to +\infty$ with $t \to +\infty$ and again both

[2] We mention that in this example every admissible control is obviously finitely optimal. Thus the concept of finite optimality is too week here.

conditions (7) and (8) are violated. Thus, for arbitrary optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ both conditions (7) and (8) are inconsistent with the core conditions of the maximum principle.

Notice, that due to [45] the asymptotic condition for the Hamiltonian (9) is a necessary optimality condition in Halkin's example. However, this condition does not give us any useful information in this case. Indeed, due to our analysis above this condition holds only in the case when $\psi^0 + \psi(0) = 0$ or equivalently $\psi^0 = 1$ and $\psi(0) = -1$. But in this case the core conditions (i) and (ii) of the maximum principle and condition (9) hold trivially along any admissible pair and hence they do not provide any useful information.

Now, we are going to apply Theorem 8 with $G = \mathbb{R}^1$. Let us fix an arbitrary admissible pair $(x_*(\cdot), u_*(\cdot))$. Assumptions $(A0)$ and $(A1)$ are obviously fulfilled; see Remark 4. In order to check Assumption $(A2)$ we notice that

$$x(\zeta;t) = 1-(1-\zeta)e^{-\int_0^t u_*(s)\,ds} \quad \text{and} \quad f_x^0(t,x,u_*(t)) = -u_*(t) \quad \text{for all} \quad t \geq 0.$$

Hence,

$$\max{}_{x\in[x(\zeta;t),x_*(t)]} \left|\langle f_x^0(t,x,u_*(t)), x(\zeta;t) - x_*(t)\rangle\right| \overset{a.e.}{=} |\zeta - x_0|\lambda(t),$$

where

$$\lambda(t) = u_*(t)e^{-\int_0^t u_*(s)\,ds} \quad \text{for all} \quad t \geq 0.$$

The function $\lambda(\cdot)$ is integrable on $[0,\infty)$. Hence, condition $(A2)$ is also satisfied and we can apply Theorem 8.

We remind that due to the explicit formula $x_*(t) = 1 - e^{-\int_0^t u_*(s)\,ds}$, $t \geq 0$, the maximum condition (23) in the normal case $\psi^0 = 1$ has the form

$$(1 + \psi(0))u_*(t) \overset{\text{a.e.}}{=} \max{}_{u\in[0,1]} \big\{(1 + \psi(0))u\big\}. \tag{24}$$

Formula (12) for the adjoint variable gives

$$\begin{aligned}\psi(t) = e^{\int_0^t u_*(s)\,ds} &\int_t^\infty e^{-\int_0^s u_*(\tau)\,d\tau}(-u_*(s))ds \\ &= e^{\int_0^t u_*(s)\,ds}\Big[\lim_{T\to\infty} e^{-\int_0^T u_*(s)\,ds} - e^{-\int_0^t u_*(s)\,ds}\Big], \qquad t \geq 0.\end{aligned}$$

Now, we consider two cases. If $\int_0^\infty u_*(t)\,dt = \infty$ (that is, $u_*(\cdot)$ is optimal), then $\psi(t) = -1$ for all $t \geq 0$, thus the maximum condition (24) is apparently satisfied. If $\int_0^\infty u_*(t)\,dt$ is finite (that is, $u_*(\cdot)$ is not optimal), then $\psi(t) > -1$ for all $t \geq 0$ and (24) implies that $u_*(t) = 1$ for almost every $t \geq 0$, which contradicts the assumption $\int_0^\infty u_*(t)\, dt < \infty$.

Summarizing, Theorem 8 provides a complete characterization of all optimal controls in $(P1)$, while the core conditions of the maximum principle are inconsistent with the standard asymptotic conditions (7) and (8), while the asymptotic condition (9) is uninformative (satisfied by any admissible control) in Halkin's example. $\square$

As it can be easily seen, if for a reference admissible pair $(x_*(\cdot), u_*(\cdot))$ Assumption $(A1)$ holds and the integral

$$I_*(t) = \int_t^\infty Z_*^{-1}(s) f_x^0(s, x_*(s), u_*(s))\, ds, \qquad t \geq 0, \tag{25}$$

converges then the function $\psi(\cdot)$ (see (12)) is defined. If, moreover, the pair $(x_*(\cdot), u_*(\cdot))$ is weakly overtaking optimal and the stronger condition $(A2)$ (see Lemma 7) holds then all assumptions of Theorem 8 are satisfied, and due to Theorem 8 the normal form maximum principle holds with the adjoint variable $\psi(\cdot)$ given by (12). Since the convergence of the integral (25) is sufficient for defining the function $\psi(\cdot)$ it is natural to ask whether Assumption $(A2)$ in Theorem 8 could not be relaxed to condition of converges of the improper integral (25)?

The analysis of the problem below shows that the convergence of the integral (25) (together with $(A1)$) is not enough for validity of Theorem 8, although the function $\psi(\cdot)$, defined by formula (12), is locally absolutely continuous, and satisfies the adjoint equation (15) and transversality conditions (7) and (8) in this case.

Example 2. Consider the following problem $(P2)$:

$$J(x(\cdot), u(\cdot)) = \int_0^\infty e^{-t}\Big[u(t) - 5x(t)^2\Big]\, dt \to \max,$$

$$\dot{x}(t) = \Big[u(t) + x(t)\Big]\phi(x(t)), \qquad x(0) = 0,$$

$$u(t) \in [0, 1].$$

Here $\phi\colon \mathbb{R}^1 \mapsto [0,1]$ is a $C^\infty(\mathbb{R}^1)$ function such that

$$\phi(x) = \begin{cases} 1 & \text{if} \quad |x| \leq 1, \\ 0 & \text{if} \quad |x| \geq 2. \end{cases}$$

Set $G = (-\infty, \infty)$. Obviously, $(P2)$ is a particular case of problem (P), and condition $(A1)$ holds for any admissible pair $(x_*(\cdot), u_*(\cdot))$ in $(P2)$ (see Remark 4).

Let us show, that $(x_*(\cdot), u_*(\cdot))$, $x_*(t) \equiv 0$, $u_*(t) \overset{\text{a.e.}}{=} 0$, $t \geq 0$, is a unique optimal pair in $(P2)$. Indeed, due to [19, Theorem 3.6.] there is an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in $(P2)$. Assume that $u_*(\cdot)$ is non-vanishing on a set of positive measure. Then for the corresponding optimal trajectory $x_*(\cdot)$ there is a unique instant $\tau > 0$ such that $x_*(\tau) = 1$.

Consider the following auxiliary problem $(P2_\tau)$:

$$J_\tau(x(\cdot), u(\cdot)) = \int_0^\tau e^{-t}\Big[u(t) - 5x(t)^2\Big]\, dt \to \max\,,$$

$$\dot{x}(t) = \Big[u(t) + x(t)\Big]\phi(x(t)), \qquad x(0) = 0, \quad x(\tau) = 1,$$

$$u(t) \in [0, 1].$$

Here all data in $(P2_\tau)$ are the same as in $(P2)$, and the only difference is that $(P4_\tau)$ is considered on the fixed time interval $[0, \tau]$ with the terminal condition $x(\tau) = 1$.

As can be easily seen, the optimal in $(P2)$ pair $(x_*(\cdot), u_*(\cdot))$ is also optimal in problem $(P2_\tau)$. Hence, according to the classical Pontryagin maximum principle [50] for problems on finite time intervals with fixed endpoints there are non-vanishing simultaneously adjoint variables $\psi^0 \geq 0$ and $\psi(\cdot)$ such that the absolutely continuous function $\psi(\cdot)$ is a solution on $[0, \tau]$ of the adjoint system

$$\dot{\psi}(t) = 10\psi^0 e^{-t} x_*(t) - \psi(t), \tag{26}$$

and for a.e. $t \in [0, \tau]$ the maximum condition takes place:

$$u_*(t)\left(\psi^0 e^{-t} + \psi(t)\right) = \max\,\{0, \psi^0 e^{-t} + \psi(t)\}. \tag{27}$$

Here we employed the fact that $x_*(t) < 1$ for all $t < \tau$ and, hence, $\phi_x(x_*(t)) \equiv 0$ for all $t \in [0, \tau]$.

If $\psi^0 = 0$, then due to (26) we have $\psi(t) = \psi(0)e^{-t}$, $t \geq 0$. Due to the maximum condition (27) this implies either $u_*(t) \overset{\text{a.e.}}{=} 0$ (if $\psi(0) < 0$) or $u_*(t) \overset{\text{a.e.}}{=} 1$ (if $\psi(0) > 0$). By assumption $u_*(\cdot)$ is non-vanishing on a set of positive measure. Hence, $u_*(t) \overset{\text{a.e.}}{=} 1$, $t \in [0, \tau]$.

Substituting $u_*(t) \overset{\text{a.e.}}{=} 1$ in the control system we get $x_*(t) = e^t - 1$,

$t \in [0, \tau]$. This implies $\tau = \ln 2$. Further, by the direct calculation we get

$$J_\tau(x_*(\cdot), u_*(\cdot)) = \int_0^{\ln 2} e^{-t}\Big[1 - 5(e^{2t} - 2e^t + 1)\Big]\, dt$$
$$= -4 \int_0^{\ln 2} e^{-t}\, dt - 5 \int_0^{\ln 2} e^t\, dt + 10 \int_0^{\ln 2} dt = 10 \ln 2 - 7 < 0. \quad (28)$$

Since $x_*(\ln 2) = 1$ and $x_*(t) \geq 1$ for $t \geq \ln 2$, we get

$$\int_{\ln 2}^{\infty} e^{-t}\Big[u_*(t) - 5x_*(t)^2\Big]\, dt < 0.$$

Hence,

$$J(x_*(\cdot), u_*(\cdot)) < J_\tau(x_*(\cdot), u_*(\cdot)) < 0,$$

which contradicts the optimality of the pair $(x_*(\cdot), u_*(\cdot))$ in problem $(P2)$. Thus, either $\psi^0 > 0$ or $u_*(t) \overset{\text{a.e.}}{=} 0$, $t \geq 0$.

Consider the case $\psi^0 > 0$. In this case without loss of generality we can assume $\psi^0 = 1/10$. Due to (26) in this case we get

$$\psi(t) = e^{-t}\left[\psi(0) + \int_0^t x_*(s)\, ds\right], \qquad t \in [0, \tau].$$

This implies

$$\psi^0 e^{-t} + \psi(t) = e^{-t}\left[\frac{1}{10} + \psi(0) + \int_0^t x_*(s)\, ds\right], \qquad t \in [0, \tau].$$

If $\psi(0) > -1/10$, then due to the maximum condition (27) we obtain that $u_*(t) \overset{\text{a.e.}}{=} 1$, $t \in [0, \tau]$. But, as it is shown above, $J(x_*(\cdot), u_*(\cdot)) < 0$ in this case, which contradicts the optimality of the pair $(x_*(\cdot), u_*(\cdot))$.

If $\psi(0) \leq -1/10$, then due to the maximum condition (27) we obtain that the control $u_*(\cdot)$ is vanishing on some interval $[0, \tau_1]$, $\tau_1 < \tau$, and then $u_*(t) = 1$ for a.e. $t \in [\tau_1, \tau]$. In this case, $x_*(t) \equiv 0$ on time interval $[0, \tau_1]$ and $x_*(t) = e^{t-\tau_1} - 1$ for all $t \in [\tau_1, \tau]$. This implies, $\tau = \tau_1 + \ln 2$. Hence, in this case we obtain that (see (28))

$$\begin{aligned} J_\tau(x_*(\cdot), u_*(\cdot)) &= \int_{\tau_1}^{\tau_1 + \ln 2} e^{-t}\left[1 - 5(e^{2(t-\tau_1)} - 2e^{t-\tau_1} + 1)\right]\, dt \\ &= e^{-\tau_1} \int_0^{\ln 2} e^{-t}\left[1 - 5(e^{2t} - 2e^t + 1)\right]\, dt < 0. \end{aligned}$$

But this inequality again contradicts the optimality of pair $(x_*(\cdot), u_*(\cdot))$ in problem $(P2)$.

Thus, we have proved that $(x_*(\cdot), u_*(\cdot))$, $x_*(t) \equiv 0$, $u_*(t) \overset{\text{a.e.}}{=} 0$, $t \geq 0$, is the unique strongly optimal pair in $(P2)$.

Along the pair $(x_*(\cdot), u_*(\cdot))$ we have $f_x^0(t, x_*(t), u_*(t)) = -10x_*(t)e^{-t} \equiv 0$, $t \geq 0$. Thus, for any $t \geq 0$ the integral (25) converges absolutely, $I_*(t) \equiv 0$, $t \geq 0$, and the adjoint function $\psi(\cdot)$ defined by equality (12) is also vanishing: $\psi(t) \equiv 0$, $t \geq 0$. Thus $\psi^0 \neq 0$, and one can take $\psi^0 = 1$. However, the maximum condition (16) (that is, (27) with $\psi^0 = 1$ in the present example) does not hold for $u_*(t) \equiv 0$, $t \geq 0$, with the adjoint variable $\psi(t) \equiv 0$, $t \geq 0$. Thus, the assertion of Theorem 8 fails in the case of problem $(P2)$. The reason of this phenomenon is the violation of the growth condition $(A2)$ for the pair $x_*(t) \equiv 0$, $u_*(t) \overset{\text{a.e.}}{=} 0$, $t \geq 0$.

However, all assumptions of the general maximum principle (see Theorem 5) are satisfied for problem $(P2)$. In particular, the strongly optimal in $(P2)$ pair $(x_*(\cdot), u_*(\cdot))$, $x_*(t) \equiv 0$, $u_*(t) \overset{\text{a.e.}}{=} 0$, $t \geq 0$, satisfies conditions (15) and (16) of the general maximum principle with adjoint variables $\psi^0 = 1$ and $\psi(t) = -e^{-t}$, $t \geq 0$. Obviously, this adjoint variable satisfies the both asymptotic conditions (7) and (8). Thus the normal form of the maximum principle holds, although the correct adjoint function is not presented by the formula (12).

The explanation is, that Assumption $(A2)$ is not only used to ensure convergence of the integral in (25), but is also essential in the proof of Theorem 8.

Example 2 also shows that formula (12) is not implied by the asymptotic conditions (7) or (8). $\square$

3 Problems with dominating discount

As it is shown in Example 2, Assumption $(A2)$ plays an essential role in Theorem 8. In this section we present a class of infinite horizon problems (P) for which the sufficient conditions for validity of $(A2)$ can be expressed in the terms of the growth rates of the functions involved. This allows us also to describe in the terms of the growth rates some situations when the explicit formula (12) implies the asymptotic conditions (7) and (8).

In addition to $(A1)$ (see also Remark 4) we pose the following assumptions.

Assumption $(B1)$**.** *There exist numbers $\mu \geq 0$, $r \geq 0$, $\kappa \geq 0$, $\beta > 0$, $\rho \in \mathbb{R}^1$, $\nu \in \mathbb{R}^1$, and $c \geq 0$ such that for every $t \geq 0$*

(i) $\|x_*(t)\| \leq c\, e^{\mu t}$*;*

(ii) for every $\zeta \in G$ *with* $\|\zeta - x_0\| < \beta$ *equation* (2) *with* $u(\cdot) = u_*(\cdot)$ *and initial condition* $x(0) = \zeta$ *(instead of* $x(0) = x_0$*) has a solution* $x(\zeta; \cdot)$ *in* G *on* $[0, \infty)$*, and it holds that*

$$\|x(\zeta; t) - x_*(t)\| \leq c\|\zeta - x_0\|\, e^{\nu t}, \quad t \geq 0 \tag{29}$$

and

$$\|f^0_x(t, y, u_*(t))\| \leq c\,(1 + \|y\|^r)\, e^{-\rho t} \textit{ for every } \ t \geq 0 \textit{ and } \ y \in [x(\zeta; t), x_*(t)].$$

Some comments about the above assumptions follow. The fist inequality in Assumption $(B1)$(ii) specifies the known fact of Lipschitz dependence of the solution of an ODE on the initial condition, requiring additionally that the Lipschitz constant depends exponentially (with rate ν) on the time horizon. Notice that the number ν can be negative. The multiplier $e^{-\rho t}$ in the second inequality of Assumption $(B1)$(ii) indicates that the objective integrand may contain a "discount" factor with a rate ρ (possibly negative). Assumption $(B1)$(i) requires a priory information about the exponential growth rate of the optimal trajectory, which can often be obtained in economic contexts.

While Assumption $(B1)$ is needed, essentially, to define the constants ρ, r, μ and ν, the next one imposes a certain relation between them, called *dominating discount condition* [6, 9, 10, 11, 14, 17].

Assumption $(B2)$:

$$\rho > \nu + r \max\{\mu, \nu\}.$$

In [15, Lemma 5.1] it is proved that Assumptions $(B1)$, $(B2)$ imply $(A2)$. Thus, the following corollary of Theorem 8 holds.

Corollary 9. *The claims of Theorem 8 are valid on Assumptions* $(A1)$*,* $(B1)$ *and* $(B2)$*.*

We mention that although the dominating discount condition $(B2)$ may be easier to check than $(A2)$, its fulfillment depends on the time scale (see [47] or [15, Part 3, Section 5]). In contrast, Assumption $(A2)$ is invariant with respect to any diffeomorphic change of the time variable. Indeed, if the time variable is changed as $t = \xi(s)$, $s \geq 0$ (where ξ maps diffeomorphically $[0, \infty)$ to $[0, \infty)$), it could be directly checked that in the resulting problem Assumption $(A2)$ is fulfilled with $\tilde{\lambda}(s) = \lambda(\xi(s))\dot{\xi}(s)$, $s \geq 0$, which is integrable if and only if $\lambda(\cdot)$ is.

Bellow we consider two more specific classes of problems where the dominating discount condition can be verified in a more convenient way: problems for one-sided Lipschitz systems and problems for systems with regular linearization.

3.1 Problems with one-sided Lipschitz dynamics

Assumption $(B1)$, hence also $(A2)$, can be verified in a more convenient way for systems with one-sided Lipschitz right-hand sides. The result below considerably extends the one previously obtained in [14, Section 4], therefore we present it in some more details. First, we recall the following definition.

Definition 10. *A function $f(\cdot,\cdot,\cdot)$, with $f(t,x,u)$ defined for $x \in G$ and $(t,u) \in \operatorname{graph} U(\cdot)$ with values in $\mathbb{R}^n$, is called* one-sided Lipschitz *with respect to x (uniformly in $(t,u) \in \operatorname{graph} U(\cdot)$) if there exists a number $\nu \in \mathbb{R}^1$ such that*

$$\langle f(t,x,u)-f(t,y,u),\, x-y\rangle \le \nu\|x-y\|^2 \quad \textit{for all } x,y \in G \textit{ and } (t,u) \in \operatorname{graph} U(\cdot).$$

Notice that the constant ν can be negative.

The following is an important well-known property of the one-sided Lipschitz systems.

Lemma 11. *For any control $u(\cdot)$ and any two solutions $x_1(\cdot)$ and $x_2(\cdot)$ of the equation $\dot{x}(t) = f(t,x(t),u(t))$ that exists in G on an interval $[\tau,T]$ it holds that*

$$\|x_1(t) - x_2(t)\| \le e^{\nu(t-\tau)}\|x_1(\tau) - x_2(\tau)\| \quad \textit{for every } t \in [\tau,T].$$

This property allows to prove the following lemma.

Lemma 12. *If $f(\cdot,\cdot,\cdot)$ is one-sided Lipschitz, then*

$$\left\| Z_*(\tau)\,[Z_*(s)]^{-1} \right\| \le \sqrt{n}\, e^{\nu(s-\tau)} \quad \textit{for every } \tau,\ s \in [0,+\infty),\ \tau \le s.$$

Proof. Let us fix an arbitrary τ and s as in the formulation of the lemma. Let $x_i(\cdot)$ be the solution of the equation $\dot{x}(t) = f(t,x(t),u_*(t))$ with $x_i(\tau) = x_*(\tau)+\alpha e_i$ where e_i is the i-th canonical unit vector in $\mathbb{R}^n$ and α is a positive scalar. Clearly, $x_i(\cdot)$ exists in G on $[\tau,s]$ for all sufficient small $\alpha > 0$.

It is a known (see e.g. [2, Chapter 2.5.6]) that on our standing assumptions

$$x_i(t) = x_*(t) + \alpha y_i(t) + o(\alpha,t), \quad t \in [\tau,s],$$

where $\|o(\alpha,t)\|/\alpha \to 0$ with α, uniformly in $t \in [\tau, s]$, and $y_i(\cdot)$ is the solution of the equation $\dot y(t) = f_x(t, x_*(t), u_*(t))y(t)$ with $y(\tau) = e_i$. This solution, however, has the form $y_i(t) = [Z_*(\tau)^*]^{-1} Z_*(t)^* e_i$. That is, $y_i(t)^*$ is the i-th row of the matrix $Z_*(\tau)\,[Z_*(t)]^{-1}$. Hence,

$$\|Z_*(\tau)\,[Z_*(s)]^{-1}\| = \left(\sum_{i=1}^{n} \|y_i(s)\|^2\right)^{1/2} = \left(\sum_{i=1}^{n} (\|x_i(s) - x^*(s) - o(\alpha,t)\|/\alpha)^2\right)^{1/2}.$$

Using Lemma 11 and passing to the limit with $\alpha \to 0$ we obtain the desired inequality. □

Using this lemma and Assumption $(B2)$ we may estimate the norm of the adjoint vector $\psi(t)$, $t \geq 0$, defined by (12) as follows:

$$\begin{aligned} \|\psi(t)\| &\leq \int_t^{\infty} \|Z_*(t)\,[Z_*(s)]^{-1}\|\,\|f_x^0(s, x_*(s), u_*(s))\|\,ds \\ &\leq \int_t^{\infty} \sqrt{n}\,e^{\nu(s-t)}\kappa(1 + c_1^r e^{\mu r s})e^{-\rho s}\,ds \;\leq\; c_3 e^{-(\rho - r\mu)t}, \qquad (30) \end{aligned}$$

where $c_3 \geq 0$ is an appropriate constant. This estimation leads to the next corollary of Theorem 8. In the formulation we use the weighted space $L_\infty(e^{\gamma t}; [0,\infty))$, consisting of all measurable functions $\psi : [0,\infty) \mapsto \mathbb{R}^n$ for which the norm

$$\|\psi(\cdot)\|_{\infty,\gamma} := \operatorname*{ess\,sup}_{t \in [0,+\infty)} e^{\gamma t}\|\psi(t)\|$$

is finite.

Lemma 11 allows to simplify Assumption $(B1)$ in the following way.

Corollary 13. *Assume that the function $f(\cdot,\cdot,\cdot)$ is one-sided Lipschitz in the sense of Definition 10. Let $(x_*(\cdot), u_*(\cdot))$ be a weakly overtaking optimal pair in problem (P), and let Assumptions $(A1)$ and $(B1)$ without the requirement that (29) be fulfilled for this pair. Assume, in addition, that $(B2)$ is fulfilled with the number ν in Definition 10. Then the function $\psi : [0,\infty) \mapsto \mathbb{R}^n$ defined by (12) is locally absolutely continuous and conditions (15) and (16) in Theorem 8 are satisfied. Moreover, the function $\psi(\cdot)$ is the unique solution of the adjoint equation (15) belonging to the weighted space $L_\infty(e^{(\rho - r\mu)t}; [0,\infty))$.*

Proof. The inequality in Lemma 11 applied with $x_1(\cdot) = x_*(\cdot)$ and $x_2(\cdot) = x(\zeta;\cdot)$ and $\tau = 0$ implies inequality (29) in $(B1)$. Then according to Corollary 9, Assumption $(A2)$ is satisfied. Thus the first part of the corollary follows from Theorem 8.

In (30) we have established the inclusion $\psi(\cdot) \in L_\infty(e^{(\rho - r\mu)t}; [0, \infty))$. Assume that $\tilde{\psi}(\cdot)$ is another solution of (15) which belongs to $L_\infty(e^{(\rho - r\mu)t}; [0, \infty))$. Then for any $t \geq 0$ we have

$$\psi(0) - \tilde{\psi}(0) = [Z_*(t)]^{-1} \, (\psi(t) - \tilde{\psi}(t)).$$

Hence,

$$\begin{aligned} \|\psi(0) - \tilde{\psi}(0)\| &\leq \| \, [Z_*(t)]^{-1} \, \| \, (\|\psi(t)\| + \|\tilde{\psi}(t)\|) \\ &\leq \sqrt{n} \, e^{\nu t} e^{-(\rho - r\mu)t} \left(\|\psi(\cdot)\|_{\infty, \rho - r\mu} + \|\tilde{\psi}(\cdot)\|_{\infty, \rho - r\mu} \right) \\ &\leq c_4 e^{-(\rho - \nu - r\mu)t}, \qquad t \geq 0. \end{aligned}$$

for an appropriate constant $c_4 \geq 0$ (which may depend on $\tilde{\psi}(\cdot)$). Since the right-hand side converges to zero as $t \to +\infty$, we obtain that $\|\psi(0) - \tilde{\psi}(0)\| = 0$ which completes the proof. □

The next corollary links the membership $\psi(\cdot) \in L_\infty(e^{(\rho - r\mu)t}; [0, \infty))$ provided by Corollary 13 with the asymptotic conditions (7) and (8).

Corollary 14. *If the assumptions of Corollary 13 hold and, in addition, $\rho > r\mu$ then the asymptotic condition* (7) *is valid. Moreover, if in addition to the assumptions of Corollary 13 the stronger inequality $\rho > (r+1)\mu$ holds then both asymptotic conditions* (7) *and* (8) *are valid.*

Proof. Note first, that since $\mu \geq 0$ and $r \geq 0$ (see $(B1)$) both inequalities $\rho > r\mu$ and $\rho > (r+1)\mu$ imply $\rho > 0$. On the assumptions of Corollary 13 we have $\psi(\cdot) \in L_\infty(e^{(\rho - r\mu)t}; [0, \infty))$. This means that there is a constant $c_3 \geq 0$ such that the inequality (30) takes place. Hence, inequality $\rho > r\mu$ implies the validity of the asymptotic condition (7) in this case. Further, due to condition (i) in $(B1)$ we have $\|x_*(t)\| \leq ce^{\mu t}$, $t \geq 0$. Hence, the stronger inequality $\rho > (r+1)\mu$ implies the validity of both asymptotic conditions (7) and (8) in this case. □

3.2 Systems with regular linearization

Here we consider another special case in which Assumption $(B1)$ takes a more explicit form.

First we recall a few facts from the stability theory of linear systems (see e.g. [31, 33] for more details). Consider a linear differential system

$$\dot{y}(t) = A(t) y(t), \tag{31}$$

where $t \in [0, \infty)$, $y \in \mathbb{R}^n$, and all components of the real $n \times n$ matrix function $A(\cdot)$ are bounded measurable functions.

Let $y(\cdot)$ be a nonzero solution of system (31). Then, the number

$$\tilde{\lambda} = \limsup_{t \to \infty} \frac{1}{t} \ln \|y(t)\|$$

is called *characteristic Lyapunov exponent* or, briefly, *characteristic exponent* of the solution $y(\cdot)$. The characteristic exponent $\tilde{\lambda}$ of any nonzero solution $y(\cdot)$ of system (31) is finite. The set of characteristic exponents corresponding to all nonzero solutions of system (31) is called *spectrum* of system (31). The spectrum always consists of at most n different numbers.

The solutions of the system (31) form a finite-dimensional linear space of dimension n. Any basis of this space, i.e., any set of n linearly independent solutions $y_1(\cdot), \dots, y_n(\cdot)$, is called *fundamental system* of solutions of system (31). A fundamental system of solutions $y_1(\cdot), \dots, y_n(\cdot)$ is said to be *normal* if the sum of the characteristic exponents of these solutions $y_1(\cdot), \dots, y_n(\cdot)$ is minimal among all fundamental systems of solutions of (31).

It turns out that a normal fundamental system of solutions of (31) always exists. If $y_1(\cdot), \dots, y_n(\cdot)$ is a normal fundamental system of solutions, then the characteristic exponents of the solutions $y_1(\cdot), \dots, y_n(\cdot)$ cover the entire spectrum of system (31). This means that for any number $\tilde{\lambda}$ in the spectrum $\tilde{\lambda}_1, \dots, \tilde{\lambda}_l$ of system (31), there exists a solution from the set $y_1(\cdot), \dots, y_n(\cdot)$ that has this number as its characteristic exponent. Note, that different members $y_j(\cdot)$ and $y_k(\cdot)$ of the fundamental system $y_1(\cdot), \dots, y_n(\cdot)$ may have the same characteristic exponent. Denote by n_s the multiplicity of the characteristic exponent $\tilde{\lambda}_s$, $s = 1, \dots, l$, belonging to the spectrum of (31). Any normal fundamental system contains the same number n_s of solutions of (31) with characteristic number $\tilde{\lambda}_s$, $1 \leq s \leq l$, from the Lyapunov spectrum of (31).

Denote

$$\sigma = \sum_{s=1}^{l} n_s \tilde{\lambda}_s.$$

The linear system (31) is said to be *regular* if

$$\sigma = \liminf_{t \to \infty} \frac{1}{t} \int_0^t \operatorname{trace} A(s)\, ds,$$

where trace$A(s)$ is the sum of all elements of $A(s)$ that lie on the principal diagonal.

Note that differential system (31) with constant matrix $A(t) \equiv A$, $t \geq 0$, is always regular. In this case the maximal element $\bar{\lambda}$ of the spectrum of (31) equals the maximal real part of the eigenvalues of the matrix A. Another important class of regular differential systems consists of systems (31) with periodic components (with the same period) in matrices $A(\cdot)$.

It is known (see, for example, [33]), that if the system (31) is regular, then for any $\varepsilon > 0$ the following inequality holds:

$$\left\| Z_*(\tau)\,[Z_*(s)]^{-1} \right\| \leq c(\varepsilon)\, e^{\bar{\lambda}(s-\tau)+\varepsilon s} \qquad \text{for every } \tau,\ s \in [0,+\infty),\ \tau \leq s. \tag{32}$$

where $\bar{\lambda}$ is the maximal element of the spectrum and the constant $c(\varepsilon) \geq 0$ depends only on ε.

The inequality (32) is similar to the inequality that appears in the assertion of Lemma 12 above. Analogously to Lemma 12 the inequality (32) leads to the following corollary of Theorem 8.

Corollary 15. *Let $(x_*(\cdot), u_*(\cdot))$ be a weakly overtaking optimal pair in problem (P), and let Assumptions $(A1)$ and $(B1)$ without the requirement that (29) be fulfilled for this pair. Let the linearized system*

$$\dot{y}(t) = f_x(t, x_*(t), u_*(t))\, y \tag{33}$$

be regular. Assume, in addition, that $(B2)$ is fulfilled with the number ν taken equal to (or larger than) the maximal element $\bar{\lambda}$ of the spectrum of system (33). Then the function $\psi : [0,\infty) \mapsto \mathbb{R}^n$ defined by (12) is locally absolutely continuous and conditions (15) and (16) in Theorem 8 are satisfied. Moreover, the function $\psi(\cdot)$ is the unique solution of the adjoint equation (15) belonging to the weighted space $L_\infty(e^{(\rho-r\mu)t}; [0,\infty))$.

Essentially the proof repeats the argument in proof of Lemma 12 above (see also [10, Section 5] and [14, Corollary 2]).

Analogously to Corollary 14 the following result links the membership $\psi(\cdot) \in L_\infty(e^{(\rho-r\mu)t}; [0,\infty))$ provided by Corollary 15 with asymptotic conditions (7) and (8).

Corollary 16. *If the assumptions of Corollary 15 hold and, in addition, $\rho > r\mu$ then the asymptotic condition (7) is valid. Moreover, if in addition to the assumptions of Corollary 15 the stronger inequality $\rho > (r+1)\mu$ holds then both asymptotic conditions (7) and (8) are valid.*

4 Applications in economics

In this section we discuss economic interpretations of the adjoint variable $\psi(\cdot)$ provided by formula (12), in view of Theorem 8 and in comparison with the dynamic programming principle. Then we present applications of Theorem 8 to two basic optimal growth models.

4.1 Economic interpretations

First, notice that the traditional interpretation of the components of the adjoint vector $\psi(t)$, $t \geq 0$, as present value *shadow prices* of the corresponding components (types) of the optimal capital stock $x_*(t)$ is based on the identification of the net present value of the capital stock vector $x_*(t)$ with the value $V(t, x_*(t))$ of the optimal value function $V(\cdot, \cdot)$, and subsequent use of the dynamic programming method (see [34] and [1, Chapter 7]). We recall this standard construction in optimal control theory.

Consider the following family of optimal control problems, $\{(P(\tau, \zeta))\}_{\tau \geq 0, \zeta \in G}$:

$$J_\tau(x(\cdot), u(\cdot)) = \int_\tau^\infty f^0(s, x(s), u(s))\, ds \to \max\,,$$

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(\tau) = \zeta, \qquad u(t) \in U(t).$$

Here, the initial time $\tau \geq 0$ and the initial state $\zeta \in G$ are considered as parameters. Admissible pairs $(x(\cdot), u(\cdot))$ in problem $(P(\tau, \zeta))$ are defined as in problem (P), but with initial data (τ, ζ) at the place of $(0, x_0)$. Thus, $P(0, x_0)$ is identical with (P).

Let us assume for now, that problem $(P(\tau, \zeta))$ has a strongly optimal solution for any $(\tau, \zeta) \in [0, \infty) \times G$. Then we can define the corresponding optimal value function $V(\cdot, \cdot)$ of variables $\tau \in [0, \infty]$ and $\zeta \in G$ as follows:

$$V(\tau, \zeta) = \sup_{(x(\cdot), u(\cdot))} J_\tau(x(\cdot), u(\cdot)). \tag{34}$$

Here the maximum is taken with respect to all admissible pairs $(x(\cdot), u(\cdot))$ in $(P(\tau, \zeta))$.

Let $(x_*(\cdot), u_*(\cdot))$ be a strongly optimal pair in (P). If the function $V(\cdot, \cdot)$ is twice continuously differentiable in some open neighborhood of graph $x_*(\cdot)$ then, applying the dynamic programming approach, it is not difficult to show that all conditions of the maximum principle (Theorem 5) hold in the normal form ($\psi^0 = 1$) with adjoint variable $\psi(\cdot)$ defined along the optimal trajectory $x_*(\cdot)$ as

$$\psi(t) = \frac{\partial V(t, x_*(t))}{\partial x}, \qquad t \geq 0. \tag{35}$$

Due to the definition of the value function $V(\cdot,\cdot)$, one can identify the present value of the capital vector $\zeta \in G$ at time $\tau \geq 0$ with $V(\tau,\zeta)$. Then due to (35), at each instant $t \geq 0$ the components of $\psi(t)$ can be interpreted as the current value *marginal prices* (also called *shadow prices*) of the corresponding components of the capital vector $x_*(t)$. This observation gives an economic meaning of the relations of the maximum principle.

Notice, that the optimal value function $V(\cdot,\cdot)$ is not necessarily differentiable. However, the differentiability of $V(t,\cdot)$ at $x_*(t)$, $t \geq 0$, is of critical importance for the interpretation of the vector $\psi(t)$ that appears in the maximum principle (Theorem 5) as a vector of marginal prices. Indeed, $\psi(t)$ being the marginal price vector at $x_*(t)$ means that $V(t,x_*(t)+\Delta x) = V(t,x_*(t)) + \langle \psi(t),\, \Delta x\rangle + o(\|\Delta x\|)$ for every increment vector Δx, where $o(\|\Delta x\|)/\|\Delta x\| \to 0$ with $\Delta x \to 0$; this implies (Frechét) differentiability of $V(t,\cdot)$ at $x_*(t)$.

It turns out that on the assumptions of Theorem 8 the adjoint variable $\psi(\cdot)$ defined by formula (12) can be interpreted as a function of *integrated intertemporal prices*, without any a priory assumptions on the optimal value function $V(\cdot,\cdot)$. We explain this interpretation in the next paragraphs.

Let $(x(\cdot),u(\cdot))$ be an admissible pair (not necessary optimal) in (P), for which Assumption $(A1)$ is fulfilled (with $(x(\cdot),u(\cdot))$ at the place of $(x_*(\cdot),u_*(\cdot))$; see also Remark 4). Let us fix an arbitrary $s > 0$. Due to the theorems on continuous dependence and differentiability of solutions of the Cauchy problem with respect to the initial conditions (see [2, §2.5.5. and §2.5.6.]), for any $\tau \in [0,s)$ there is an open neighborhood $\mathcal{V}_s(\tau) \subset G$ of $x(\tau)$ such that for any $\tau \in [0,s)$ and $\zeta \in \mathcal{V}(\tau)$ the solution $x(\tau,\zeta;\cdot)$ of the Cauchy problem

$$\dot{x}(t) = f(t,x(t),u(t)), \qquad x(\tau) = \zeta$$

exists on $[\tau,s]$, lies in G, and the function $x(\tau,\cdot;s)\colon \mathcal{V}(\tau) \to \mathbb{R}^n$ is continuously (Frechét) differentiable. Moreover, the following equality holds:

$$x_\zeta(\tau,x(\tau);s) = \Big[Z(\tau)[Z(s)]^{-1}\Big]^*, \tag{36}$$

where $Z(t)$ (consistently with our previous notations) is the fundamental matrix solution normalized at $t=0$ of the linear differential equation

$$\dot{z}(t) = -[f_x(t,x(t),u(t))]^* z(t),$$

so that $\Big[Z(\tau)[Z(s)]^{-1}\Big]^*$ is the state transition (Cauchy) matrix of the linearized system

$$\dot{y}(t) = f_x(t,x(t),u(t))y.$$

Now let us define the *intertemporal utility function* $\pi(\tau,\cdot;s)$ on $\mathcal{V}(\tau)$ by the equality

$$\pi(\tau,\zeta,s) = f^0(s,x(\tau,\zeta;s),u(s)), \qquad \zeta\in\mathcal{V}(\tau). \tag{37}$$

Substantially, $\pi(\tau,\zeta,s)$ is the *intertemporal value* gained at instant s by the capital stock ζ at instant τ after transition of the system from the state ζ to the state $x(\tau,\zeta;s)$ via the reference control $u(\cdot)$ on the time interval $[\tau,s]$.

Due to (36), the properties of $f^0(\cdot,\cdot,\cdot)$ and the identity $x(\tau,x(\tau),s) = x(s)$, the function $\pi(\tau,\cdot,s)$ defined by (37) is differentiable at $x(\tau)$ and by the chain rule

$$\begin{aligned}
\pi_\zeta(\tau,x(\tau),s) &= \Big[\left[f^0_x(s,x(s),u(s))\right]^* x_\zeta(\tau,x(\tau);s))\Big]^*\\
&= \Big[\left[f^0_x(s,x(s),u(s))\right]^* \left[Z(\tau)[Z(s)]^{-1}\right]^*\Big]^*\\
&= Z(\tau)[Z(s)]^{-1}f^0_x(s,x(s),u(s)).
\end{aligned} \tag{38}$$

Thus the vector $\pi_\zeta(\tau,x(\tau),s)$ can be interpreted as the corresponding *intertemporal price* vector of the capital stock $x(\tau)$.

Notice that $s>0$ was arbitrarily chosen, therefore the function $(t,s)\mapsto \pi_\zeta(t,x(t),s)$ is defined for all $s>0$ and $t\in[0,s)$. Moreover, the representation (38) implies that this function is Lebesgue measurable. Thus we can define the function

$$\mu(t) = \int_t^\infty \pi_\zeta(t,x(t),s)\,ds, \qquad t\ge 0, \tag{39}$$

provided that the above integral converges for any $t\ge 0$. Thus, the *integrated intertemopral prices function* $\mu(\cdot)$ is defined by (39) along any (not necessary optimal) admissible trajectory $x(\cdot)$ in (P). Notice, that only Assumption $(A1)$ and convergence of the improper integral in (39) are needed to define the integrated intertemopral prices function $\mu(\cdot)$. No smoothness, Lipschitzness, continuity, and even finiteness assumptions on the corresponding optimal value function $V(\cdot,\cdot)$ in a neighborhood of the reference admissible trajectory $x(\cdot)$ are required.

Now, let $(x_*(\cdot),u_*(\cdot))$ be a weakly overtaking optimal admissible pair in (P) and let Assumption $(A1)$ be satisfied for this pair. The matrix function $Z(\cdot)$ and the function $\mu(\cdot)$ associated with the pair $(x_*(\cdot),u_*(\cdot))$ will be denoted by $Z_*(\cdot)$ and $\mu_*(\cdot)$, correspondingly. From (38) and (39) we obtain that

$$\mu_*(t) = Z_*(t)\int_t^\infty [Z_*(s)]^{-1}f^0_x(s,x_*(s),u_*(s))\,ds, \qquad t\ge 0. \tag{40}$$

If the above integral is finite for every $t \geq 0$, then $\mu_*(\cdot)$ coincides with the function $\psi(\cdot)$ defined in (12), which appears in the formulation of Theorem 8. Observe, that if also Assumption $(A2)$ is fulfilled for the pair $(x_*(\cdot), u_*(\cdot))$, then due to Lemma 7 the improper integral in (40) converges for any $t \geq 0$, thus $\mu_*(\cdot) = \psi(\cdot)$ is well-defined on $[0, \infty)$. Hence, we obtain that on Assumptions $(A1)$ and $(A2)$, the adjoint variable $\psi(\cdot)$ that appears in Theorem 8 coincides with the integrated intertemopral prices function $\mu_*(\cdot)$.

Assumption $(A2)$ is sufficient, but not necessary for finiteness of the integral in (40) for all $t \geq 0$. Given also that the function $\mu_*(\cdot)$ has the economic meaning of integrated intertemporal prices function, it is natural to ask whether Assumption $(A2)$ in Theorem 8 could not be relaxed to condition of converges of the improper integral in (12) or (40). The answer to this question is negative, as shown in Example 4. It could happen (if $(A2)$ fails) that for a unique strongly optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem (P) Assumption $(A1)$ is satisfied, the corresponding improper integral in (12) converges absolutely, and the general maximum principle (Theorem 5) holds in the normal form with adjoint variable $\psi(\cdot)$ which is not equal to the integrated intertemporal prices function $\mu(\cdot)$, although function $\mu(\cdot)$ is well defined by equality (40). Thus, in general the adjoint variable $\psi(\cdot)$ that appears in the normal form conditions of the general maximum principle (Theorem 5) could be something different from the integrated intertemporal prices function $\mu(\cdot)$, while under conditions of Theorem 8 both these functions coincide. Assumption $(A2)$ is not only needed to ensure finiteness of $\mu_*(\cdot)$ via Lemma 7; it is also essential for the proof of Theorem 8.

Consider now a weakly overtaking optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in (P) for which the assumptions of Theorem 8 (i.e. Assumptions $(A1)$ and $(A2)$) are satisfied, and in addition, $J(x_*(\cdot), u_*(\cdot))$ in (1) is finite. In this case the assertion of Theorem 8 can be strengthened.

Due to $(A2)$, there is an open neighborhood Ω of the set graph $x_*(\cdot)$ such that the integral below converges for any $(\tau, \zeta) \in \Omega$ and, hence, the following *conditional value* function $W(\cdot, \cdot)\colon \Omega \mapsto \mathbb{R}^1$ is well defined:

$$W(\tau, \zeta) = \int_{\tau}^{\infty} \pi(\tau, \zeta; s)\, ds, \qquad (\tau, \zeta) \in \Omega.$$

Notice, that the optimal value function $V(\cdot, \cdot)$ (see (34)) is not necessarily defined in this case. Substantially, the value $W(\tau, \zeta)$, $(\tau, \zeta) \in \Omega$, has economic meaning of integrated intertemporal value of the capital vector ζ at time τ (under the condition that the given investment plan $u_*(\cdot)$ is realized for initial capital vector ζ at initial instant τ on the whole infinite time interval $[\tau, \infty)$).

The following result strengthens the assertion of Theorem 8 on the additional assumption of convergence of the improper integral in (1).

Theorem 17. *Let $(x_*(\cdot), u_*(\cdot))$ be a locally weakly overtaking optimal pair in problem (P) for which Assumptions $(A1)$ and $(A2)$ are fulfilled, and suppose that the integral in* (1) *converges to the finite value $J(x_*(\cdot), u_*(\cdot))$. Then*

(i) for any $t \geq 0$ the partial (Frechét) derivative $W_x(t, x_*(t))$ exists. Moreover, the vector function $\psi(\cdot)\colon [0, \infty) \mapsto \mathbb{R}^n$ defined by the equality

$$\psi(t) = W_x(t, x_*(t)), \qquad t \geq 0,$$

is locally absolutely continuous and satisfies the core conditions (15) *and* (16) *of the maximum principle in the normal form for problem (P);*

(ii) the partial derivative $W_t(t, x_*(t))$ exists for a.e. $t \geq 0$, and

$$W_t(t, x_*(t)) + \sup_{u \in U(t)} \big\{ \langle W_x(t, x_*(t)), f(t, x_*(t), u) \rangle + f^0(t, x_*(t), u) \big\} \stackrel{a.e.}{=} 0.$$

The proof of Theorem 17, given in [5, Section 2], is based on the theorem on differentiability of solutions of the Cauchy problem with respect to the initial conditions, Theorem 8, and the fact that under the conditions of Theorem 17 we have $W_x(t, x_*(t)) \equiv \mu(t)$, $t \geq 0$, (see (39)) and equality (40) takes place.

Substantially, assertion (i) of Theorem 17 is a reformulation of Theorem 8 in the economic terms of function $W(\cdot, \cdot)$ under additional assumption of convergence of $J(x_*(\cdot), u_*(\cdot))$. However, assertion (ii) is a complementary fact. In particular, this assertion allows to link the adjoint variable $\psi(\cdot)$ that appears in Theorem 17 with Michel's asymptotic condition (9).

Corollary 18. *Assume that the assumptions of Theorem 17 are fulfilled and that problem (P) is autonomous with exponential discounting, i.e. $f(t, x, u) \equiv f(x, u)$, $f^0(t, x, u) \equiv e^{-\rho t} g(x, u)$ and $U(t) \equiv U$ for all $t \geq 0$, $x \in G$, $u \in \mathbb{R}^m$, where $\rho \in \mathbb{R}^1$ is not necessarily positive. Then the following stationarity condition holds:*

$$\mathcal{H}(t, x_*(t), u_*(t), \psi(t)) \stackrel{a.e.}{=} \rho \int_t^\infty e^{-\rho s} g(x_*(s), u_*(s))\, ds, \qquad t \geq 0. \tag{41}$$

Proof. Indeed, for all $t \geq 0$ we have

$$W(t, x_*(t)) = e^{-\rho t} \int_t^\infty e^{-\rho(s-t)} g(x_*(s), u_*(s))\, ds = e^{-\rho t} W(0, x_*(t)).$$

Hence,

$$W_t(t, x_*(t)) = -\rho e^{-\rho t} W(0, x_*(t)) = -\rho \int_t^\infty e^{-\rho s} g(x_*(s), u_*(s))\, ds, \qquad t \geq 0.$$

By virtue of assertions (i) and (ii) of Theorem 17, this implies (41). □

Finally, note that if problem (P) is autonomous with discounting, and the usual regularity assumptions concerning the weakly overtaking optimal control $u_*(\cdot)$, and functions $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$ hold (see Remark 4), then the core conditions (15) and (16) of the normal form maximum principle imply that the function $h(\cdot)$: $h(t) = H(t, x_*(t), \psi(t)) = \sup_{u \in U} \mathcal{H}(t, x_*(t), u, \psi(t))$, $t \geq 0$, is locally absolutely continuous and $\dot{h}(t) \overset{\text{a.e.}}{=} \partial H(t, x_*(t), \psi(t))/\partial t = -\rho g(x_*(t), u_*(t))$, $t \geq 0$ (see [50, Chapter 2]). Since the functional $J(x_*(\cdot), u_*(\cdot))$ converges in Theorem 17, we obtain that conditions (9) and (41) are equivalent in this case.

4.2 Two economic examples

Example 3 (Ramsey's model). This example demonstrates the applicability of Theorem 8 to the Ramsey model of optimal economic growth (see [18, Chapter 2]). This model is the most important theoretical construct in modern growth theory. It was first developed by Ramsey [51] in 1928 and then adapted by Cass [27] and Koopmans [44] in the 1960th. It is known also in the literature as the Ramsey-Cass-Koopmans model. Here, for simplicity of presentation we restrict our consideration to the canonical setting of the model with the Cobb-Douglas technology and the isoelastic utility. For the case of the general neoclassical production function see [23].

Below we present a rigorous analysis of the Ramsey model based on Theorem 8. We show that Theorem 8 is applicable, hence, the normal form core conditions of the maximum principle hold with the adjoint variable $\psi(\cdot)$ specified by formula (12). In this case formula (12) directly implies the validity of asymptotic condition (8).

Consider a closed aggregated economy that at each instant of time $t \geq 0$ produces a single homogeneous product $Y(t) > 0$ in accordance with the Cobb-Douglas production function (see [18, Chapter 1]):

$$Y(t) = AK(t)^\alpha L(t)^{1-\alpha}. \tag{42}$$

Here $A > 0$ is the technological coefficient, $0 < \alpha < 1$ is the output elasticity of capital, $K(t) > 0$ and $L(t) > 0$ are the capital stock and the labour force available at instant $t \geq 0$ respectively.

In the closed economy, at each instant $t \geq 0$, a part $I(t) = u(t)Y(t)$, $u(t) \in [0,1)$, of the final product is invested, while the remaining (nonvanishing) part $C(t) = (1 - u(t))Y(t)$ is consumed. Therefore, the capital dynamics can be described by the following differential equation:

$$\dot{K}(t) = u(t)Y(t) - \tilde{\delta}K, \qquad K(0) = K_0 > 0, \tag{43}$$

where $\tilde{\delta} > 0$ is the capital depreciation rate.

Assume, that the labor resource $L(\cdot)$ satisfies the exponential growth condition, i.e.,

$$\dot{L}(t) = \mu L(t), \qquad L(0) = L_0 > 0, \tag{44}$$

where $\mu \geq 0$ is a constant. Assume also, that the instantaneous utility function $g\colon (0,\infty) \mapsto \mathbb{R}^1$ is isoelastic (see [18, Chapter 2]). In this case

$$g(c) = \begin{cases} \frac{c^{1-\sigma}-1}{1-\sigma}, & \sigma > 0, \quad \sigma \neq 1, \\ \ln c, & \sigma = 1, \end{cases} \tag{45}$$

where $c > 0$ is the per-capita consumption. Then introducing a new (capital-labor ratio) variable $x(t) = K(t)/L(t)$, $t \geq 0$, in view of (42)-(45), and due to homogeneity of the Cobb-Doulglas production function (42), we arrive to the following optimal control problem $(P3)$:

$$J(x(\cdot),u(\cdot)) = \int_0^\infty e^{-\rho t} g((1-u(t))Ax(t)^\alpha)\, dt \to \max\,, \tag{46}$$

$$\dot{x}(t) = u(t)Ax(t)^\alpha - \delta x(t), \quad x(0) = x_0 = \frac{K_0}{L_0}, \tag{47}$$

$$u(t) \in [0,1). \tag{48}$$

Here $\rho > 0$ is a social discount rate, $(1 - u(t))Ax(t)^\alpha = C(t)/L(t)$ is the per capita consumption at instant $t \geq 0$, and $\delta = \tilde{\delta} + \mu > 0$ is the adjusted depreciation rate.

Set $G = (0,\infty)$. Then any measurable function $u\colon [0,\infty) \mapsto [0,1)$ is an admissible control in problem $(P3)$. Indeed, due to (47) the corresponding to $u(\cdot)$ trajectory $x(\cdot)$ is defined on $[0,\infty)$ in G, and the integrand in (46), i.e. the function $t \mapsto e^{-\rho t} g((1-u(t))Ax(t)^\alpha))$, is locally integrable on $[0,\infty)$. Thus, the trajectory $x(\cdot)$ is admissible. Moreover, due to (45), (47) and (48) the integrand in (46) is bounded from above by an exponentially declining function (uniformly in all admissible pairs $(x(\cdot), u(\cdot))$). Hence, there is a decreasing nonnegative function $\omega\colon [0,\infty) \mapsto \mathbb{R}^1$, $\lim_{t\to\infty} \omega(t) = 0$, such that for any $0 \le T < T'$ the following inequality holds (see [6, Section 2, Assumption $(A3)$]):

$$\int_T^{T'} e^{-\rho t} g((1-u(t))Ax(t)^\alpha)\, dt \le \omega(T).$$

This implies that for any admissible pair $(x(\cdot), u(\cdot))$ the improper integral in (46) either converges to a finite number or diverges to $-\infty$, and $J(x(\cdot), u(\cdot)) \le \omega(0)$ (see [6, Section 2]). Hence, in the case of problem $(P3)$ the concepts of strong optimality and weak overtaking optimality coincide. So, everythere below in this example we understand optimality of an admissible pair $(x_*(\cdot), u_*(\cdot))$ in the problem $(P3)$ in the strong sense. In particular, if an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ exists then $J(x_*(\cdot), u_*(\cdot))$ is a finite number.

Let us define an auxiliary state variable $y(\cdot)$ via the Bernoulli transformation: $y(t) = x(t)^{1-\alpha}$, $t \ge 0$. Then it can be easily seen, that in terms of the state variable $y(\cdot)$ the problem $(P3)$ takes the following equivalent form $(\widetilde{P3})$:

$$\tilde{J}(y(\cdot), u(\cdot)) = \int_0^\infty e^{-\rho t} g((1-u(t))Ay(t)^{\frac{\alpha}{1-\alpha}})\, dt \to \max,$$

$$\dot{y}(t) = (1-\alpha)Au(t) - (1-\alpha)\delta y(t), \quad y(0) = y_0 = x_0^{1-\alpha}, \tag{49}$$

$$u(t) \in [0,1). \tag{50}$$

For the problem $(\widetilde{P3})$ we set again $G = (0,\infty)$. Since $f(t,y,u) := (1-\alpha)Au - (1-\alpha)\delta y$ and $f^0(t,y,u) := e^{-\rho t} g((1-u)Ay^{\frac{\alpha}{1-\alpha}})$, $t \ge 0$, $x \in G$ and $u \in [0,1)$ in $(\widetilde{P3})$ Assumption $(A0)$ is fulfilled for these functions and the multivalued mapping $U(\cdot)$: $U(t) \equiv [0,1)$, $t \ge 0$ (see Remark 1). Obviously, arbitrary measurable function $u\colon [0,\infty) \mapsto \mathbb{R}^1$ satisfying the pointwise constraint (50) is an admissible control in $(\widetilde{P3})$. Thus, $(\widetilde{P3})$ is a particular case of problem (P).

Further, in $(\widetilde{P3})$ for all $t \ge 0$, $y > 0$ and $u \in [0,1)$ we have $f_y(t,y,u) \equiv -(1-\alpha)\delta$, and

$$\begin{aligned} f_y^0(t,y,u) &= e^{-\rho t} \frac{d\, g((1-u)Ay^{\frac{\alpha}{1-\alpha}})}{d\, c} \frac{(1-u)A\alpha y^{\frac{\alpha}{1-\alpha}-1}}{(1-\alpha)} \\ &= \frac{(1-u)Ae^{-\rho t}\alpha y^{\frac{\alpha}{1-\alpha}-1}}{(1-\alpha)} \left[(1-u)Ay^{\frac{\alpha}{1-\alpha}}\right]^{-\sigma} = \frac{\alpha e^{-\rho t}}{(1-\alpha)y} \left[(1-u)Ay^{\frac{\alpha}{1-\alpha}}\right]^{1-\sigma} \end{aligned}$$

if $\sigma > 0$, $\sigma \neq 1$, and

$$f_y^0(t, y, u) = \frac{\alpha e^{-\rho t}}{(1-\alpha)y}$$

if $\sigma = 1$.

Thus, for any $\sigma > 0$ and for all $t \geq 0$, $y > 0$ and $u \in [0, 1)$ we have

$$f_y^0(t, y, u) = \frac{\alpha e^{-\rho t}}{(1-\alpha)y} \left[(1-u)Ay^{\frac{\alpha}{1-\alpha}}\right]^{1-\sigma}. \tag{51}$$

Hence, Assumption $(A1)$ is satisfied for any admissible pair $(y_*(\cdot), u_*(\cdot))$ in $(\widetilde{P3})$.

Notice, that the control system (49) is linear in problem $(\widetilde{P3})$. Hence, for arbitrary admissible control $u_*(\cdot)$ and any initial state $y(0) = \zeta > 0$ the corresponding admissible trajectory $y(\zeta, \cdot)$ is given by the Cauchy formula

$$y(\zeta, t) = e^{-(1-\alpha)\delta t}\zeta + (1-\alpha)Ae^{-(1-\alpha)\delta t}\int_0^t e^{(1-\alpha)\delta s}u_*(s)\, ds, \quad t \geq 0. \tag{52}$$

Let us show that Assumption $(A2)$ is also satisfied for any optimal admissible pair $(y_*(\cdot), u_*(\cdot))$ in $(\widetilde{P3})$ (if such exists).

Take an arbitrary admissible pair $(y_*(\cdot), u_*(\cdot))$ such that $\tilde{J}(y_*(\cdot), u_*(\cdot)) > -\infty$, and set $\beta = y_0/2$. Then due to (51) and (52) for any $\zeta : |\zeta - y_0| < \beta$ and all $t \geq 0$ we get

$$\begin{aligned}
&\max{}_{y\in[y(\zeta;t),y_*(t)]} \left| f_y^0(t, y, u_*(t))(y(\zeta;t) - y_*(t)) \right| \\
&= \frac{\alpha}{1-\alpha} \max{}_{y\in[y(\zeta;t),y_*(t)]} \frac{e^{-\rho t}e^{-(1-\alpha)\delta t}|\zeta - y_0| \left[(1-u_*(t))Ay^{\frac{\alpha}{1-\alpha}}\right]^{1-\sigma}}{y} \\
&\leq \frac{\alpha|\zeta - y_0|}{1-\alpha} \max{}_{y\in[y(\zeta;t),y_*(t)]} \frac{e^{-\rho t}\left[(1-u_*(t))Ay^{\frac{\alpha}{1-\alpha}}\right]^{1-\sigma}}{\frac{y_0}{2} + (1-\alpha)A\int_0^t e^{(1-\alpha)\delta s}u_*(s)\, ds} \\
&\leq \frac{2\alpha|\zeta - y_0|}{y_0(1-\alpha)} \max{}_{y\in[y(\zeta;t),y_*(t)]} \left\{ e^{-\rho t}\left[(1-u_*(t))Ay^{\frac{\alpha}{1-\alpha}}\right]^{1-\sigma} \right\} = |\zeta - y_0|\lambda(t),
\end{aligned}$$

where

$$\lambda(t) = \frac{2\alpha}{y_0(1-\alpha)} \max{}_{y\in[y(\zeta;t),y_*(t)]} \left\{ e^{-\rho t}\left[(1-u_*(t))Ay^{\frac{\alpha}{1-\alpha}}\right]^{1-\sigma} \right\}. \tag{53}$$

Notice that for all $t \geq 0$ and $\zeta \in [y_0 - \beta, y_0 + \beta]$ formula (52) (where the integral term is non-negative) implies the chain of inequalities

$$\frac{1}{2}y_*(t) \leq y\left(\frac{1}{2}y_0; t\right) \leq y(\zeta; t) \leq y\left(\frac{3}{2}y_0; t\right) \leq \frac{3}{2}y_*(t). \tag{54}$$

Due to the choice of β we have $\zeta \in [\frac{1}{2}y_0, \frac{3}{2}y_0]$. The monotonicity of $\zeta \mapsto y(\zeta;t)$ implies $y(\zeta;t) \in [y(\frac{1}{2}y_0;t), y(\frac{3}{2}y_0;t)]$ which together with (54) gives $[y(\zeta;t), y_*(t)] \subset [\frac{1}{2}y_*(t), \frac{3}{2}y_*(t)]$. Thus

$$\lambda(t) \le \frac{2\alpha}{y_0(1-\alpha)} \max_{y \in [\frac{1}{2}y_*(t), \frac{3}{2}y_*(t)]} \Big\{ e^{-\rho t} \Big[(1-u_*(t)) A y^{\frac{\alpha}{1-\alpha}} \Big]^{1-\sigma} \Big\}, \quad t \ge 0.$$

Due to the monotonicity with respect to y of the function in the braces (which is non-increasing for $\sigma \in (0,1]$ and non-decreasing for $\sigma \ge 1$) we obtain that

$$0 \le \lambda(t) \le \frac{2\alpha}{y_0(1-\alpha)} e^{-\rho t} \max \left\{ \left(\frac{1}{2}\right)^{\frac{\alpha(1-\sigma)}{1-\alpha}}, \left(\frac{3}{2}\right)^{\frac{\alpha(1-\sigma)}{1-\alpha}} \right\} \Big[(1-u_*(t)) A y_*(t)^{\frac{\alpha}{1-\alpha}} \Big]^{1-\sigma}. \tag{55}$$

Since $\tilde{J}(y_*(\cdot), u_*(\cdot)) > -\infty$ the function $t \mapsto e^{-\rho t} g\Big((1-u_*(t)) A y_*(t)^{\frac{\alpha}{1-\alpha}} \Big)$ is integrable on $[0,\infty)$. Then the function

$$t \mapsto e^{-\rho t} \Big[(1-u_*(t)) A y_*(t)^{\frac{\alpha}{1-\alpha}} \Big]^{1-\sigma} = e^{-\rho t} \Big[1 + (1-\sigma) g\Big((1-u_*(t)) A y_*(t)^{\frac{\alpha}{1-\alpha}} \Big) \Big]$$

is also integrable. Due to (55) this implies that the function $\lambda(\cdot)$ defined in (53) is integrable on $[0,\infty)$. Thus, Assumption $(A2)$ is satisfied for arbitrary $\sigma > 0$ and all admissible pairs $(y_*(\cdot), u_*(\cdot))$ with $\tilde{J}(y_*(\cdot), u_*(\cdot)) > -\infty$.

Thus, for arbitrary $\sigma > 0$ and any optimal admissible pair $(y_*(\cdot), u_*(\cdot))$ in $(\widetilde{P3})$ all assumptions of Theorem 8 are satisfied. Hence, for any optimal admissible pair $(y_*(\cdot), u_*(\cdot))$ in $(\widetilde{P3})$ the core conditions (15) and (16) of the normal form maximum principle hold with adjoint variable $\psi(\cdot)$ specified by formula (12) (see (51) and (52)):

$$\begin{aligned}
\psi(t) &= \frac{\alpha e^{(1-\alpha)\delta t}}{1-\alpha} \int_t^\infty \frac{e^{-(1-\alpha)\delta s} e^{-\rho s} \Big[(1-u_*(s)) A y_*(s)^{\frac{\alpha}{1-\alpha}} \Big]^{1-\sigma}}{y_*(s)} \, ds \\
&= \frac{\alpha e^{(1-\alpha)\delta t}}{1-\alpha} \int_t^\infty \frac{e^{-\rho s} \Big[(1-u_*(s)) A y_*(s)^{\frac{\alpha}{1-\alpha}} \Big]^{1-\sigma}}{y_0 + (1-\alpha) A \int_0^s e^{(1-\alpha)\delta \tau} u_*(\tau) \, d\tau} \, ds, \qquad t \ge 0.
\end{aligned}$$

Replacing $\int_0^s$ in the right-hand side with $\int_0^t$ (which is not larger) and using (52), we obtain the following relations:

$$
\begin{aligned}
0 < \psi(t) y_*(t) &\leq \frac{\alpha}{(1-\alpha)} \int_t^\infty e^{-\rho s} \left[(1 - u_*(s)) A y_*(s)^{\frac{\alpha}{1-\alpha}} \right]^{1-\sigma} ds \\
&= \frac{\alpha}{(1-\alpha)} \int_t^\infty e^{-\rho s} \left[1 + (1-\sigma) g\Big((1 - u_*(s)) A y_*(s)^{\frac{\alpha}{1-\alpha}} \Big) \right] ds \\
&= \frac{\alpha e^{-\rho t}}{(1-\alpha)\rho} + \frac{\alpha(1-\sigma)}{1-\alpha} \int_t^\infty e^{-\rho s} g\Big((1 - u_*(s)) A y_*(s)^{\frac{\alpha}{1-\alpha}} \Big) ds. \qquad (56)
\end{aligned}
$$

Notice, that condition (56) is a stronger fact than the asymptotic condition (8).

Now, introducing the current value adjoint variable $p(\cdot)$, $p(t) = e^{\rho t}\psi(t)$, $t \geq 0$ we arrive at the current value adjoint system (see point (i) in Theorem 8 and (51))

$$
\dot{p}(t) = \left((1-\alpha)\delta + \rho \right) p(t) - \frac{\alpha}{(1-\alpha) y(t)} \left[(1 - u_*(t)) A y(t)^{\frac{\alpha}{1-\alpha}} \right]^{1-\sigma}, \qquad (57)
$$

and the current value maximum condition (see (ii))

$$
\begin{aligned}
&(1-\alpha) A u_*(t) p(t) + g\Big((1 - u_*(t)) A y_*(t)^{\frac{\alpha}{1-\alpha}} \Big) \\
&\qquad \stackrel{\text{a.e.}}{=} \max_{u \in [0,1)} \Big\{ (1-\alpha) A u p(t) + g\Big((1-u) A y_*(t)^{\frac{\alpha}{1-\alpha}} \Big) \Big\}. \qquad (58)
\end{aligned}
$$

Since for any $\sigma > 0$ the isoelastic function $g(\cdot)$ is strictly concave (see (45)) the current value maximum condition (58) implies that $u_*(t) \stackrel{\text{a.e.}}{=} u_*(y_*(t), p(t))$ where for any $y > 0$ and $p > 0$ the feedback $u_*(y,p)$ is defined via the unique solution of the equation

$$
(1-\alpha) A p + \frac{d}{du} g\Big((1-u) A y^{\frac{\alpha}{1-\alpha}} \Big) = (1-\alpha) A p - \frac{\Big(A y^{\frac{\alpha}{1-\alpha}} \Big)^{1-\sigma}}{(1-u)^\sigma} = 0,
$$

i.e.

$$
u_*(y,p) = \begin{cases} 1 - \dfrac{y^{\frac{\alpha(1-\sigma)}{\sigma(1-\alpha)}}}{A(1-\alpha)^{\frac{1}{\sigma}} p^{\frac{1}{\sigma}}}, & \text{if} \quad p > \dfrac{1}{A(1-\alpha)} y^{\frac{\alpha(1-\sigma)}{1-\alpha}}, \\[2ex] 0, & \text{if} \quad p \leq \dfrac{1}{A(1-\alpha)} y^{\frac{\alpha(1-\sigma)}{1-\alpha}}. \end{cases} \qquad (59)
$$

Substituting $u_*(y(t), p(t))$ defined in (59) in the control system (49) and in the adjoint system (57) instead of $u_*(t)$ we arrive at the following normal form current value Hamiltonian system of the maximum principle:

$$
\dot{y}(t) = (1-\alpha) A u_*(y(t), p(t)) - (1-\alpha)\delta y(t), \qquad (60)
$$

$$\dot{p}(t) = ((1-\alpha)\delta + \rho)\, p(t) - \frac{\alpha}{(1-\alpha)y(t)} \left[(1 - u_*(y(t), p(t))) A y(t)^{\frac{\alpha}{1-\alpha}} \right]^{1-\sigma}. \tag{61}$$

Due to Theorem 8 an optimal admissible trajectory $y_*(\cdot)$ (if any) together with the corresponding current value adjoint variable $p(\cdot)$ must satisfy to the system (60), (61), as well as the initial condition $y(0) = y_0 = x_0^{1-\alpha}$, and the estimate (56).

Due to the linearity of equation (49) and the concavity of the isoelastic function $g(\cdot)$ for any $\sigma > 0$ (see (45)), the Hamiltonian in problem $(\widetilde{P3})$ is a concave function of the state variable $y > 0$. This fact together with estimate (56) imply that all conditions of Arrow's sufficient conditions of optimality (see [55, Theorem 10]) are satisfied. Thus, any solution $(y_*(\cdot), p(\cdot))$ of the system (60), (61) on $[0, \infty)$ which satisfies the initial condition $y(0) = y_0 = x_0^{1-\alpha}$, and the estimate (56), corresponds to the optimal admissible pair $(y_*(\cdot), u_*(\cdot))$, where $u_*(t) = u_*(y_*(t), p(t))$, $t \geq 0$. Thus, the assertion of Theorem 8 is a necessary and sufficient condition (a criterion) of optimality of an admissible pair $(y_*(\cdot), u_*(\cdot))$ in $(\widetilde{P3})$.

The direct analysis (which we omit here) shows that for any $\sigma > 0$ and arbitrary initial state $y_0 > 0$ there is a unique solution $(y_*(\cdot), p(\cdot))$ of the system (60), (61) which satisfies both the initial condition $y(0) = y_0$ and the estimate (56). Hence, for any initial state $y_0 > 0$ there is a unique optimal admissible pair $(y_*(\cdot), u_*(\cdot))$ in $(\widetilde{P3})$. It can also be shown, that the solution $(y_*(\cdot), p(\cdot))$ approaches asymptotically the unique equilibrium $(\hat{y}, \hat{p})$ (of saddle type) of system (60), (61). Since $p(t) \to \hat{p}$ and $y_*(t) \to \hat{y}$ with $t \to \infty$ obviously both standard asymptotic conditions (7) and (8) hold in this example.

Finally, returning to the initial state variable $x_*(t) = y_*(t)^{\frac{1}{1-\alpha}}$, $t \geq 0$, we get a unique optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem $(P3)$. $\square$

Example 4 (Model of optimal extraction of a non-renewable resource). In this example we apply Theorem 8 to a basic model of optimal extraction of a non-renewable resource. Notice, that the issue of optimal use of an exhaustible resource was raised first by Hotelling in 1931 [40]. Then a model involving both the man-made capital and an exhaustible resource was developed in a series of papers now commonly referred to as the Dasgupta-Heal-Solow-Stiglitz (DHSS) model (see [32, 61, 62]). A complete analysis of the DHSS model in the case of constant return to scale and no capital depreciation was presented in [22]. Application of Theorem 8 to the DHSS model with logarithmic instantaneous utility, any return to scale and capital depreciation can be found in [7]. Here, we consider the case of a non-renewable (not necessarily completely extractable) resource. For the sake of simplicity we do not involve any man-made capital into consideration.

Consider the following problem $(P4)$:

$$J(x(\cdot),u(\cdot)) = \int_0^\infty e^{-\rho t} g(u(t)(x(t)-a))\,dt \to \max\,, \tag{62}$$

$$\dot{x}(t) = -u(t)(x(t)-a), \quad x(0) = x_0 > a, \tag{63}$$

$$u(t) \in (0,\infty). \tag{64}$$

Here $x(t)$ is the stock of a non-renewable resource at instant $t \geq 0$ and $a \geq 0$ is the non-extractable part of the stock. In the case $a = 0$ the resource can be asymptotically exhausted, while in the case $a > 0$, due to technological (or some other) reasons, it can be depleted only up to the given minimal level $a > 0$. Further, $u(t)$ is the (non-vanishing) rate of extraction of the available for exploitation part $x(t) - a$ of the total stock $x(t)$ of the resource at instant $t \geq 0$. All the extracted amount $u(t)(x(t)-a)$ of the resource is consumed at each instant $t \geq 0$. Thus, $c(t) = u(t)(x(t)-a)$, $t \geq 0$, is the corresponding consumption. As in Example 2, we assume that $\rho > 0$ is a social discount rate and the instantaneous utility function $g(\cdot)$ is isoelastic (see (45)).

Set $G = (a,\infty)$. Obviously Assumption $(A0)$ holds for the corresponding functions $f(\cdot,\cdot,\cdot)$ and $f^0(\cdot,\cdot,\cdot)$): $f(t,x,u) = -u(x-a)$ and $f^0(t,x,u) = e^{-\rho t} g(u(x-a))$, $t \geq 0$, $x > a$, $u \in (0,\infty)$, and the multivalued mapping $U(\cdot)$: $U(t) \equiv (0,\infty)$, $t \geq 0$, in $(P4)$ (see Remark 1). Thus, $(P4)$ is a particular case of problem (P).

Due to (63), for any locally integrable function $u\colon [0,\infty) \mapsto (0,\infty)$ and arbitrary initial state $\zeta > a$ the corresponding solution $x(\zeta,\cdot)$ of the Cauchy problem (63) is defined by the formula

$$x(\zeta,t) = (\zeta - a)e^{-\int_0^t u(s)\,ds} + a, \qquad t \geq 0. \tag{65}$$

Hence, any locally integrable function $u(\cdot)$ satisfying the pointwise constraint (64) is an admissible control in $(P4)$.

Since, for any $\sigma > 0$ and all $t \geq 0$, $x > a$, $u \in (0,\infty)$ in $(P4)$ we have

$$f_x(t,x,u) \equiv -u, \qquad f_x^0(t,x,u) = e^{-\rho t} u^{1-\sigma}(x-a)^{-\sigma},$$

Assumption $(A1)$ is also satisfied in $(P4)$ for any admissible pair $(x_*(\cdot),u_*(\cdot))$.

Let $(u(\cdot),x(\cdot))$ be an arbitrary admissible pair. In the case $\sigma = 1$ due to (63) we have $g(u(t)(x(t)-a)) = \ln(-\dot{x}(t)) \leq -\dot{x}(t)$ for a.e. $t \geq 0$, hence, for every $0 \le T < T'$

$$\int_T^{T'} e^{-\rho t} \ln(u(t)(x(t)-a))\, dt \le -\int_T^{T'} e^{-\rho t} \dot{x}(t)\, dt \le (x_0 - a)e^{-\rho T}. \quad (66)$$

In the case $\sigma < 1$ we have

$$g(c) = \frac{c^{1-\sigma} - 1}{1-\sigma} \le \frac{c}{1-\sigma}, \qquad c > 0,$$

hence, for every $0 \le T < T'$

$$\int_T^{T'} e^{-\rho t} \frac{(u(t)(x(t)-a))^{1-\sigma} - 1}{1-\sigma} \le \int_T^{T'} e^{-\rho t} \frac{-\dot{x}(t)}{1-\sigma}\, dt \le \frac{x_0 - a}{1-\sigma} e^{-\rho T}. \quad (67)$$

In the case $\sigma > 1$ we have $g(u(t)(x(t)-a)) \le 1/(\sigma - 1)$ for a.e. $t \ge 0$, hence, for every $0 \le T < T'$

$$\int_T^{T'} e^{-\rho t} \frac{(u(t)(x(t)-a))^{1-\sigma} - 1}{1-\sigma} \le \frac{1}{\sigma - 1} e^{-\rho T}. \quad (68)$$

Due to (66), (67) and (68) for any $\sigma > 0$ there is a decreasing nonnegative function $\omega\colon [0,\infty) \mapsto \mathbb{R}^1$, $\lim_{t\to\infty} \omega(t) = 0$, such that for arbitrary admissible pair $(x(\cdot), u(\cdot))$ the following estimate holds:

$$\int_T^{T'} e^{-\rho t} g(u(t)x(t))\, dt \le \omega(T), \qquad 0 \le T < T'. \quad (69)$$

As in Example 3 above, the estimate (69) implies that for any admissible pair $(x(\cdot), u(\cdot))$ the improper integral in (62) either converges to a finite number or diverges to $-\infty$, and $J(x(\cdot), u(\cdot)) \le \omega(0)$. Hence, in problem $(P4)$ for any $\sigma > 0$ the concepts of strong optimality and weak overtaking optimality coincide. So, everythere below in this example we understand optimality of an admissible pair $(x_*(\cdot), u_*(\cdot))$ in the problem $(P4)$ in the strong sense. In particular, if an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ exists then $J(x_*(\cdot), u_*(\cdot))$ is a finite number. This fact will be used later.

Below we focus our analysis on the case $\sigma \ne 1$, since the case of logarithmic instantaneous utility ($\sigma = 1$) was considered in [5, Example 3] and [7, Section 5].

Let us fix an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in $(P4)$ (if such exists). We shall show that Assumption $(A2)$ is satisfied for this pair. Let us set $\beta = (x_0 - a)/2$. Taking into account that the function $\zeta \mapsto x(\zeta; t)$ is monotone increasing and $x \mapsto (x-a)^{-\sigma}$ is monotone decreasing, we obtain that for $\zeta \in [x_0-\beta, x_0+\beta] = [(x_0+a)/2, (3x_0-a)/2]$

$$
\begin{aligned}
&\max{}_{x\in[x(\zeta;t),x_*(t)]} \left| f^0_x(t,x,u_*(t))(x(\zeta;t)-x_*(t)) \right| \\
&\leq |\zeta - x_0| e^{-\int_0^t u_*(s)\,ds} e^{-\rho t} u_*(t)^{1-\sigma} \max{}_{x\in[x(\zeta;t),x_*(t)]}(x-a)^{-\sigma} \\
&\leq |\zeta - x_0| e^{-\int_0^t u_*(s)\,ds} e^{-\rho t} u_*(t)^{1-\sigma} (x((x_0+a)/2;t)-a)^{-\sigma} \\
&= |\zeta - x_0| \Big(\frac{x_0+a}{2}\Big)^{-\sigma} e^{-\rho t} \left[u_*(t) e^{-\int_0^t u_*(s)\,ds}\right]^{1-\sigma} \\
&\leq |\zeta - x_0| \Big(\frac{x_0+a}{2}\Big)^{-\sigma} e^{-\rho t} \left(\frac{-\dot{x}_*(t)}{x_0-a}\right)^{1-\sigma} = |\zeta - x_0|\lambda(t), \qquad t \geq 0,
\end{aligned}
$$

where

$$\lambda(t) = \Big(\frac{x_0+a}{2}\Big)^{-\sigma} e^{-\rho t} \left(\frac{-\dot{x}_*(t)}{x_0-a}\right)^{1-\sigma}, \qquad t \geq 0,$$

is integrable on $[0,\infty)$ due to the integrability of the function

$$t \mapsto e^{-\rho t}(-\dot{x}_*(t))^{1-\sigma} = e^{-\rho t}(1+(1-\sigma)g(u_*(t)(x_*(t)-a)).$$

Thus Assumption $(A2)$ is fulfilled with $\beta = (x_0-a)/2$ and above defined function $\lambda(\cdot)$.

Then, due to Theorem 8 for an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ (if any) the core conditions (15) and (16) of the normal form maximum principle holds with adjoint variable $\psi(\cdot)$ such that

$$\psi(t) \quad = \quad e^{\int_0^t u_*(s)\,ds} \int_t^\infty e^{-\rho s} u_*(s)^{1-\sigma}(x_0-a)^{-\sigma} e^{-(1-\sigma)\int_0^s u_*(\xi)\,d\xi}\, ds. \tag{70}$$

If an optimal control $u_*(\cdot)$ does exist then according to (16) we have $\mathcal{H}_u(t, x_*(t), u_*(t), \psi(t)) \overset{\text{a.e.}}{=} 0$ on $[0,\infty)$, where the Hamilton-Pontryagin function has the form

$$\mathcal{H}(t,x,u,\psi) = \frac{e^{-\rho t}(u^{1-\sigma}(x-a)^{1-\sigma}-1)}{1-\sigma} - \psi\, u\,(x-a), \quad t \geq 0,\ x > a,\ u > 0.$$

Differentiating in u we obtain that the following equality holds:

$$e^{-\rho t} u_*(t)^{-\sigma}(x_*(t)-a)^{-\sigma} - \psi(t) \overset{\text{a.e.}}{=} 0, \qquad t \geq 0.$$

Substituting the expressions (65) (with $\zeta = x_0$) and (70) for $x_*(t)$ and $\psi(t)$, we obtain that

$$u_*(t)^{-\sigma} \overset{\text{a.e.}}{=} e^{\rho t} e^{(1-\sigma)\int_0^t u_*(s)\,ds} \int_t^\infty e^{-\rho s} u_*(s)^{1-\sigma} e^{-(1-\sigma)\int_0^s u_*(\xi)\,d\xi}\, ds. \tag{71}$$

Due to the absolute convergence of the above integral, the last expression implies that $u_*(\cdot)$ is (equivalent to) a locally absolutely continuous function on $[0,\infty)$.

Further, due to (65) equality (71) implies

$$\begin{aligned} u_*(t)^{-\sigma} &= \frac{e^{\rho t} e^{(1-\sigma)\int_0^t u_*(s)\,ds}}{(x_0-a)^{1-\sigma}} \int_t^\infty e^{-\rho s}(u_*(s)(x_*(s)-a))^{1-\sigma}\,ds \\ &= \frac{e^{\rho t}}{(x_*(t)-a)^{1-\sigma}} \int_t^\infty e^{-\rho s}(u_*(s)(x_*(s)-a))^{1-\sigma}\,ds \qquad t \geq 0. \end{aligned}$$

Hence,

$$u_*(t) = \frac{e^{-\rho t}(u_*(t)(x_*(s)-a))^{1-\sigma}}{\int_t^\infty e^{-\rho s}(u_*(s)(x_*(s)-a))^{1-\sigma}\,ds} = -\dot{z}(t), \qquad t \geq 0, \tag{72}$$

where the locally absolutely continuous function $z(\cdot)$ is defined by equality

$$z(t) = \ln \int_t^\infty e^{-\rho s}\left(u_*(s)(x_*(s)-a)\right)^{1-\sigma}\,ds, \qquad t \geq 0.$$

Integrating equality (72) on arbitrary time interval $[0,T]$, $T>0$, we get

$$\begin{aligned} \int_0^T u_*(s)\,ds = \ln \int_0^\infty e^{-\rho s}\left(u_*(s)(x_*(s)-a)\right)^{1-\sigma}\,ds \\ - \ln \int_T^\infty e^{-\rho s}\left(u_*(s)(x_*(s)-a)\right)^{1-\sigma}\,ds. \end{aligned}$$

Since $(x_*(\cdot), u_*(\cdot))$ is an optimal admissible pair, the first term in the right-hand side is finite, while the second term converges to $-\infty$. Thus, we obtain that $\int_0^\infty u_*(s)\,ds = \infty$.

Differentiating (71) in t and utilizing the same expression for $u_*(t)$ we conclude that for a.e. $t \geq 0$ the function $u_*(\cdot)$ satisfies the equality

$$-\sigma u_*(t)^{-\sigma-1}\dot{u}_*(t) = \rho u_*(t)^{-\sigma} - u_*(t)^{1-\sigma} + u_*(t)(1-\sigma)u_*(t).$$

Dividing by $-\sigma u_*(t)^{-\sigma-1}$ we obtain that $u_*(\cdot)$ is a locally absolutely continuous solution of the following differential equation:

$$\dot{u}(t) = u(t)^2 - \frac{\rho}{\sigma}u(t).$$

The general solution of this simple Riccati equation is

$$u_*(t) = \frac{e^{-\frac{\rho}{\sigma}t}}{c - \frac{\sigma}{\rho}\left(1 - e^{-\frac{\rho}{\sigma}t}\right)},$$

where c is a constant (equal to $u_*(0)^{-1}$).

Since $u_*(\cdot)$ takes only positive values this expression for $u_*(\cdot)$ implies $c \geq \sigma/\rho$, and due to equality $\int_0^\infty u_*(s)\,ds = \infty$ we finally get $c = \sigma/\rho$. Thus, we conclude that application of Theorem 8 determines a unique admissible pair $(x_*(\cdot), u_*(\cdot))$ which is suspectable for optimality in problem $(P4)$ (see (65)):

$$x_*(t) = (x_0 - a)e^{-\frac{\rho}{\sigma}t} + a, \qquad u_*(t) \equiv \frac{\rho}{\sigma}, \qquad t \geq 0. \tag{73}$$

Notice, that the explicit formula (70) for the corresponding adjoint variable $\psi(\cdot)$ gives

$$\psi(t) \equiv \left(\frac{\rho(x_0 - a)}{\sigma}\right)^{-\sigma}. \tag{74}$$

Let us show that the admissible pair $(x_*(\cdot), u_*(\cdot))$ defined in (73) is indeed optimal in $(P4)$. For this consider the function $\Phi\colon [0,\infty) \times (0,\infty) \mapsto \mathbb{R}^1$ defined as follows:

$$\Phi(t, y) = e^{-\rho t} g(y) - \psi(t) y, \qquad t \geq 0, \quad y > 0,$$

where $\psi(\cdot)$ is defined by (74). It is easy to see that for any $t \geq 0$ due to the strict concavity of the isoelastic function $g(\cdot)$ (see (45)) there is a unique point $y_*(t)$ of global maximum of function $\Phi(t, \cdot)$ on $(0, \infty)$. For any $t \geq 0$ the point $y_*(t)$ is a unique solution of the equation $e^{-\rho t} g_y(y) = \psi(t)$. Since $g_y(y) = y^{-\sigma}$, $y > 0$, (see (45)) solving this equation we obtain that (see (74))

$$y_*(t) = \left(e^{\rho t}\psi(t)\right)^{-\frac{1}{\sigma}} = \frac{\rho(x_0 - a)e^{-\frac{\rho}{\sigma}t}}{\sigma}, \qquad t \geq 0.$$

However, due to (73) we get

$$(x_*(t) - a)u_*(t) = \frac{\rho(x_0 - a)e^{-\frac{\rho}{\sigma}t}}{\sigma} = y_*(t), \qquad t \geq 0.$$

Thus, we have proved that the following global inequality takes place:

$$\begin{aligned} &e^{-\rho t} g(u_*(t)(x_*(t) - a)) - \psi(t)u_*(t)(x_*(t) - a) \\ &\qquad \geq e^{-\rho t} g(u(x - a)) - \psi(t)u(x - a), \qquad t \geq 0, \quad x > 0, \quad u > 0. \end{aligned} \tag{75}$$

Now, let $(x(\cdot), u(\cdot))$ be an arbitrary other admissible pair. Then due to (75) we have (see (74))

$$\begin{aligned} &e^{-\rho t} g(u_*(t)(x_*(t) - a)) - e^{-\rho t} g(u(t)(x(t) - a)) \\ &\qquad \geq \left(\frac{\rho(x_0 - a)}{\sigma}\right)^{-\sigma} \left[u_*(t)(x_*(t) - a) - u(t)(x(t) - a)\right], \qquad t \geq 0. \end{aligned}$$

Integrating the last inequality on an arbitrary time interval $[0,T]$, $T>0$, we get

$$\int_0^T e^{-\rho t} g(u_*(t)(x_*(t)-a))\,dt - \int_0^T e^{-\rho t} g(u(t)(x(t)-a))\,dt$$
$$\geq \left(\frac{\rho(x_0-a)}{\sigma}\right)^{-\sigma} \int_0^T \left[u_*(t)(x_*(t)-a) - u(t)(x(t)-a)\right]\,dt$$
$$= \left(\frac{\rho(x_0-a)}{\sigma}\right)^{-\sigma} \left(x(T)-x_*(T)\right), \qquad T \geq 0.$$

Passing to the limit as $T\to\infty$, and taking into account that $\lim_{T\to\infty} x_*(T) = a$ and $x(T)\geq a$, we obtain that

$$\int_0^\infty e^{-\rho t} g(u_*(t)(x_*(t)-a))\,dt \geq \limsup_{T\to\infty} \int_0^T e^{-\rho t} g(u(t)(x(t)-a))\,dt.$$

Hence, $(x_*(\cdot),u_*(\cdot))$ is indeed (the unique) optimal admissible pair in problem $(P4)$.

Thus, we conclude, that for any $\sigma>0$, $\sigma\neq 1$ due to (73) and (74) the asymptotic condition (7) always fails in this example. Moreover, since $x_*(t)\to a$ as $t\to\infty$ (see (73)) the asymptotic condition (8) also fails if $a>0$, while it holds if $a=0$. Therefore, if $0<\sigma<1$ and $a>0$ then the explicit formula (12) plays a role of an alternative to asymptotic conditions (7) and (8) which are both inconsistent with the core conditions (15) and (16) of the maximum principle. The same phenomenon, i.e. the simultaneous violation of both the standard asymptotic conditions (7) and (8) can be observed also in the case $\sigma=1$ and $\rho>0$ (see [5, Example 3] and [7, Section 5]). In the last case the optimal extraction rate $u_*(\cdot)$ coincides with the classical Hotelling rule (see [40]), i.e. $u_*(t)\equiv\rho$, $t\geq 0$. In the case $\sigma>0$, $\sigma\neq 1$, considered above, the optimal extraction rate $u_*(\cdot)$ represents the adjusted Hotelling rule (according with the value of σ), i.e. $u_*(t)\equiv\rho/\sigma$, $t\geq 0$ (see (73)). □

5 Bibliographical comments

To the best of our knowledge, an optimal control problem with infinite time horizon was considered first in the seminal monograph [50, Chapter 4]. The problem considered in [50] is completely autonomous, involves no discounting, satisfies the usual regularity assumptions (see Remark 4), and involves additional asymptotic terminal condition $\lim_{t\to\infty} x(t) = x_1$, where x_1 is a given asymptotic terminal state in $\mathbb{R}^n$. Potentially, the approach suggested in [50] is applicable to a broad scope of infinite-horizon optimal control problems, in particular to problem (P) with free terminal state that is in the focus of the present paper. This approach is based on the construction of an "initial cone" K_{t_0} at the initial time t_0 instead of the "limiting cone" at the terminal time t_1 (which does not exist in the infinite-horizon case). The initial cone K_{t_0} is constructed in the same way as the limiting cone K_{t_1} at the terminal time t_1 in the case of a finite-horizon problem on the time interval $[t_0, t_1]$, $t_0 < t_1$. This construction is based on the classical needle variations technique (see [50]). The only difference with the finite-horizon case is that the increment of the principal linear part of the varied trajectory is transmitted (by solving a system of variational equations) to the initial moment t_0 rather than to the terminal time t_1 (which does not exist). All other points of this construction are essentially the same as in the finite-horizon case. When the initial cone K_{t_0} is constructed, the subsequent application of a topological result and the separation theorem provide a corresponding version of the maximum principle (see [50] for details). Notice, that this construction employees only the property of finite optimality of the reference optimal control $u_*(\cdot)$. Therefore, when applied to problem (P), this construction leads to exactly the same result as the general version of the maximum principle for problem (P) that was developed later by Halkin (see [38]).

The paper [38] by Halkin considers problem (P) with free terminal state at infinity under the usual regularity assumptions (see Remark 4). The integral functional (1) is not assumed to be finite in [38]. The approach employed in [38] is based on the consideration of the family of auxiliary optimal control problems (P_T) on finite time intervals $[0, T]$, $T > 0$, appearing in Definition 3. The finite optimality of the admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem (P) implies that on any finite time interval $[0, T]$, $T > 0$, the core conditions of the Pontryagin maximum principle for the pair $(x_*(\cdot), u_*(\cdot))$ hold with a corresponding non-vanishing pair of adjoint variables $\psi_T^0 \geq 0$, $\psi_T(\cdot)$. This implies the validity of the core conditions of the infinite-horizon maximum principle after taking a limit in the conditions of the maximum principle for these auxiliary problems (P_T) as $T \to \infty$ (see details in [26, 38]). No additional characterizations of the adjoint variables ψ^0 and $\psi(\cdot)$ such as normality of the problem and/or some boundary conditions at infinity are provided in [26]. Moreover, the paper [38] suggests two counterexamples demonstrating possible pathologies in the relations of the general version of the maximum principle for problem (P), namely possible abnormality of problem (P) ($\psi^0 = 0$ in this case) and possible violation of the standard asymptotic conditions (7) and (8).

Apparently, [58] and [38] were the first papers in which the authors demonstrated by means of counterexamples that abnormality is possible, and the "natural" asymptotic conditions (7) and (8) may be violated in the case of infinite-horizon problems with free terminal state at infinity. Since the discount rate ρ is equal to zero in these examples, the doubts that such pathologies are attributed only to problems without time discounting were rather common in economic literature for a along time (see for example [18, Section A.3.9] and [28, Chapter 9]). However, many "pathological" examples with positive discount rate are known nowadays (see for example [11, Chapter 1, Section 6] and [45, Section 2]), including models developed quite recently and having clear economic interpretations (see [16, Section 4], [5, Example 3], [7, Section 2.2] and Example 4 in Section 4).

After the publication of paper [38] many authors attempted to develop normal versions of the maximum principle for problem (P) and characterize, on various additional assumptions, the asymptotic behavior of the adjoint variable for which the maximum condition (4) is fulfilled. The first positive results in this direction were obtained in [17] and [21].

In [17], a particular case of problem (P) is investigated, where the control system is linear and autonomous:

$$\dot{x}(t) = Fx(t) + u(t), \qquad x(0) = x_0,$$

The constraining set $U \subset \mathbb{R}^n$ is convex and compact, the instantaneous utility function has the form $f^0(t,x,u) = e^{-\rho t} g(x,t)$, $t \geq 0$, $x \in \mathbb{R}^n$, $u \in U$, with a positive discount rate ρ and a locally Lipschitz with respect to both variables x and u function $g(\cdot,\cdot)$.[3] The authors assume that the following dominating discount condition holds:

$$\rho > (r+1)\lambda_F. \tag{76}$$

Here, λ_F is the maximal real part of the eigenvalues of the $n \times n$ matrix F and r is a nonnegative number that characterizes the growth of the function $g(\cdot,\cdot)$ in terms of its generalized gradient $\partial g(\cdot,\cdot)$ (in the sense of Clarke [29]; see [17] for more detail):

$$\|\zeta\| \leq \kappa(1 + \|(x,u)\|^r) \qquad \text{for any} \quad \zeta \in \partial g(x,u), \quad x \in G, \quad u \in U.$$

Note, that the generalized gradient is taken here with respect to both variables x and u.

[3] Here and below for the reasons of unification we use notations sometimes slightly different from the ones used in the original papers under discussion.

Since $\rho > 0$ it is easy to see that condition (76) guarantees convergence of the functional $J(x(\cdot), u(\cdot))$ for any admissible pair $(x(\cdot), u(\cdot))$. Accordingly, the concept of strong optimality is employed in [17].

In the case of $r > 0$, the authors of [17] proposed a version of the Pontryagin maximum principle in the normal form; this version contains a characterization of the behavior of the adjoint variable $\psi(\cdot)$ in terms of the convergence of the improper integral

$$\int_0^\infty e^{(q-1)\rho t} \|\psi(t)\|^q \, dt < \infty. \tag{77}$$

Here, the constant $q > 1$ is defined by the equality $1/q + 1/(r+1) = 1$. As pointed out in [17], condition (77) implies that the asymptotic conditions (7) holds.

Later the result obtained in [17] was generalized and strengthened by different methods (which differ also from the method employed in [17]) in the series of papers [6, 9, 10, 11, 14, 23]. In these papers a few different (not equivalent) extensions of the condition (76) to the case of nonlinear problems (P) were suggested and various normal form versions of the maximum principle with adjoint variable $\psi(\cdot)$ specified explicitly by formula (12) were developed. Here we mention only that in the linear case (considered in [17]) the dominating discount condition (76) implies validity of Assumption $(B2)$ in Section 3 and, hence, $(A2)$. In this case due to Theorem 8 the core conditions (15) and (16) of the normal form maximum principle hold with adjoint variable $\psi(\cdot)$ specified by formula (12) which directly implies the estimate (77) (see [11, Section 16] and [12]). Moreover, since $\rho > 0$, formula (12) implies validity of both asymptotic conditions (7) and (8) in this case (see [11, Section 12] and Corollary 16).

In [21] a version of first-order necessary optimality conditions that contains the asymptotic condition at infinity (8) was developed for the infinite-horizon dynamic optimization problem of the form

$$J(x(\cdot)) = \int_0^\infty f^0(t, x(t), \dot{x}(t)) \, dt \to \max, \tag{78}$$

$$(x(t), \dot{x}(t)) \in K, \qquad x(0) = x_0. \tag{79}$$

Here the set $K \subset \mathbb{R}^{2n}$ is assumed to be convex, closed with a nonempty interior; the function $f^0 : [0, \infty) \times K \mapsto \mathbb{R}^1$ is jointly concave in the variables x, $\dot{x}$ for any $t \geq 0$, and the optimal trajectory $x_*(\cdot)$ is assumed to lie in the interior of the set $\operatorname{dom} V(\cdot, t)$ for any $t \geq 0$, where

$$\operatorname{dom} V(t, \cdot) = \{x_0 \in \mathbb{R}^n : V(t, x_0) < \infty\}$$

is the effective set of the optimal value function $V(\cdot,\cdot)$:

$$V(t,x_0)=\sup\left\{\int_t^\infty f^0(s,x(s),\dot{x}(s))\,ds\colon\ (x(s),\dot{x}(s))\in K \text{ for } s\geq t;\ x(t)=x_0\right\}.$$

Under the assumptions made, the problem (78), (79) is "completely convex." In particular, the optimal value function $V(\cdot,\cdot)$ is concave in the variable x_0 for any $t\geq 0$, and the set of all admissible trajectories is convex in the space $C([0,T],\mathbb{R}^n)$ for any $T>0$.

The main result of [21] states that there exists an adjoint variable $\psi(\cdot)$ corresponding to the optimal trajectory $x_*(\cdot)$ such that

$$\psi(t)\in\partial_x V(t,x_*(t)),\qquad t\geq 0. \tag{80}$$

Here, $\partial_x V(t,x_*(t))$ is the partial subdifferential (in the sense of convex analysis) of the concave function $V(t,\cdot)$ at the point $x_*(t)$ for fixed t. Next, a certain generalized Euler equation and the asymptotic condition (8) were derived from the inclusion (80) in [21] under some additional assumptions. In particular, it was assumed that the phase vector x is nonnegative and the function $f^0(\cdot,\cdot,\cdot)$ is monotone in the variable $\dot{x}$ (see [21] for more detail).

The question of whether asymptotic conditions of the form (8) hold for the problem (78), (79) was considered in [43] without the convexity assumptions in the situation when the optimal trajectory $x_*(\cdot)$ is regular and interior, and the control system satisfies a homogeneity condition.

The next step in developing of complementary necessary conditions characterizing the asymptotic behavior of the adjoint variable $\psi(\cdot)$ has been done by Michel in 1982 (see [45]). In the special case when the problem (P) is autonomous with exponential discounting (i.e. $f(t,x,u)\equiv f(x,u)$, $f^0(t,x,u)\equiv e^{-\rho t}g(x,u)$ and $U(t)\equiv U$, where $\rho\in\mathbb{R}^1$ is not necessarily positive) and under the assumption that the optimal value $J(x_*(\cdot),u_*(\cdot))$ is finite, the author established validity of the asymptotic condition (9) along any strongly optimal admissible path $x_*(\cdot)$. This asymptotic condition is analogous to the transversality condition with respect to time in problems with free final time [50]. Since the standard regularity conditions are employed (see Remark 4), in this case condition (9) is equivalent to the stationarity condition

$$H(t,x_*(t),\psi^0,\psi(t))=\psi^0\rho\int_t^\infty e^{-\rho s}g(x_*(s),u_*(s))\,ds,\qquad t\geq 0. \tag{81}$$

Notice, that the adjoint variable ψ^0 can be equal to zero here, and an example of an autonomous problem (P) with positive discounting ($\rho>0$) in which the equality $\psi^0 = 0$ necessarily holds is presented in [45, Section 2]. The complementary character of condition (9) is demonstrated in [11, Example 6.6]. A generalization of Michel's result to the case when the instantaneous utility $f^0(\cdot,\cdot,\cdot)$ depends on the variable t in more general way was developed in [57], using a slightly modified argument. A normal form version of the maximum principle with adjoint variable $\psi(\cdot)$ having all positive components which involves condition (41) was developed also in [13] under some monotonicity type assumptions.

In some cases, in particular, in the situation when the function $g(\cdot,\cdot)$ is nonnegative and there exists a neighborhood $\mathcal{V}$ of 0 in $\mathbb{R}^n$ such that $\mathcal{V} \subset f(x_*(t), U)$ for all large enough instants t, the asymptotic condition (9) implies condition (7). Nevertheless, being one dimensional, condition (9) (as well as (81)) cannot provide a full set of complementary conditions for the adjoint variable $\psi(\cdot)$ in the general multidimensional case.

The relationship between the explicit formula (12) and the asymptotic condition (9) is discussed in Corollary 18 (see also [5, 4]).

In [64], Ye obtained the stationarity condition (81) in the case of a nonsmooth problem (P) with discounting (provided that the autonomous functions $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$ are Lipschitz in the phase variable x uniformly in u and Borel measurable in u, and the function $g(\cdot,\cdot)$ is bounded).

The main result in [64] provided a version of the maximum principle with the asymptotic conditions (7) and (9) under the additional assumption

$$\rho > \max\left\{0, \sup_{x,y\in G,\ u\in U} \frac{\langle x-y, f(x,u)-f(y,u)\rangle}{\|x-y\|^2}\right\}. \tag{82}$$

This assumption means that ρ is positive and $\rho > \nu$, where ν is the one-sided Lipschitz constant of the vector function $f(\cdot,\cdot)$ with respect to the phase variable x (see Definition 10). It is easy to see that if the functions $f(\cdot,\cdot)$, $f_x(\cdot,\cdot)$, $g(\cdot,\cdot)$, $g_x(\cdot,\cdot)$ are continuous in the variable x, the function $g(\cdot,\cdot)$ is Lipschitz in variable x uniformly in u (this implies $r = 0$ in $(B1)$) and condition (82) holds then due to Theorem 8 the core conditions (15) and (16) of the normal form maximum principle hold together with the adjoint variable $\psi(\cdot)$ specified via formula (12) (see Lemma 13). In this case the formula (12) implies validity of the asymptotic condition (7).

In [60], Smirnov characterized the asymptotic behavior of the adjoint variable $\psi(\cdot)$ in terms of the Lyapunov exponents (see [31] and 3 for the relevant definitions of stability theory). The main assumption of [60] is that the system of variational equations considered along a reference optimal pair

$(x_*(\cdot), u_*(\cdot))$ is regular. In this case, under some additional assumptions,[4] it is proved in [60] that the characteristic Lyapunov exponent of the adjoint variable $\psi(\cdot)$ corresponding to the reference optimal pair $(x_*(\cdot), u_*(\cdot))$ is nonpositive (see [60, Theorem 3.1]). However, this result does not guarantee either normality of the optimal control problem under consideration or fulfillment of asymptotic conditions (7) or (8). As pointed out in [60] by means of a counterexample, regularity of the system of variational equations plays an important role here.

In [53], Seierstad considered an infinite-horizon optimal control problem (as a minimum problem) that is more general than (P). The statement of this problem includes a nonautonomous nonsmooth control system

$$\dot{x}(t) = f(t, x(t), u(t)), \qquad u(t) \in U,$$

an initial condition $x(0) = x_0 \in \mathbb{R}^n$, and equality- and inequality-type terminal boundary constraints at infinity that are imposed on some of the phase coordinates $x^i(\infty)$, $i = 1, \dots, m$ (it is assumed that $x(t) \in \mathbb{R}^n$, $m < n$, and the corresponding limits $x^i(\infty) = \lim_{t\to\infty} x^i(t)$, $i = 1, \dots, m$, exist). The problem considered in [53] consists in the minimization of a terminal functional of the form

$$J(\pi x(\infty)) = \sum_{i=1}^{m} \nu_i x^i(\infty).$$

Here, π is the $m \times n$ matrix of the operator of projection onto the subspace of the first m coordinates in $\mathbb{R}^n$ and ν_i, $i = 1, \dots, m$, are real numbers.

In [53], a version of the maximum principle that contains a full set of asymptotic conditions at infinity is developed, however, under rather restrictive growth and some other assumptions. In application to problem (P) with free right endpoint, this result implies normality of the maximum principle and validity of the asymptotic condition (7). For a discussion of the growth conditions employed in [53] see [11, Section 16].

Notice also, that a version of the maximum principle that contains a full set of asymptotic conditions at infinity is developed in [54] for a smooth infinite-horizon optimal control problem with unilateral state constraints and with terminal conditions on the states at infinity by employing the

[4] Note that in [60] an important condition that is used in the proof is missing in the formulation of the main result. Namely, the gradient of the integrand must be bounded: $\|b(t)\| \le K$ for a.e. $t \ge 0$ (see [60, Theorem 3.1]). Example 10.4 in [53] shows that this condition is essential.

needle variation technique. However, the growth conditions employed in [54] are rather demanding.

The approach based on approximations of the problem (P) by a specially constructed sequence of finite-horizon problems $\{(P_k)\}_{k=1}^{\infty}$ on time intervals $[0, T_k]$, $T_k > 0$, $\lim_{k\to\infty} T_k = \infty$, was developed in [9, 10, 11, 13] in the case of autonomous problems (P) with discounting. In this case, on any finite interval $[0, T]$, $T > 0$, the sequence of optimal controls $\{u_k(\cdot)\}$ (which exist) in the approximating problems (P_k), $k = 1, 2, \dots$, converges weakly in $L_1([0, T], \mathbb{R}^m)$ (or in other appropriate sense) to the reference optimal control $u_*(\cdot)$ in problem (P). The necessary optimality conditions for problem (P) are obtained by passing to the limit as $k \to \infty$ in the relations of the Pontryagin maximum principle for the approximating problems (P_k). It was proved in [9, 10, 11] that the maximum principle holds in normal form with the adjoint variable $\psi(\cdot)$ specified by the formula (12) under some dominating discount type conditions. Similar characterization of the adjoint variable $\psi(\cdot)$ was obtained also in the so called "monotone case" (see [13] and [11] for details). Recently, the main constructions and results in [9, 10, 11] were extended in [6, 23].

Although the finite-horizon approximation based approach enables one to develop different normal form versions of the maximum principle that contain full sets of necessary conditions for problem (P), there are inherent limits for the applicability of this approach. In particular, application of this approximation technique assumes validity of the conditions guaranteeing existence of solutions in the corresponding finite-horizon approximate problems. Moreover, some uniformity of convergence of the improper integral utility functional (1) for all admissible controls (see e.g. condition $(A3)$ in [11]) is employed. In many cases of interest assumptions of this type either fail or cannot be verified a priory. For instance, in problems without discounting and in models of endogenous economic growth (especially with declining discount rates) the corresponding integral utility functionals may diverge to infinity.

The approach for derivation of first order necessary optimality conditions for infinite-horizon optimal control problems, which is based on the methods of general theory of extremal problems (see [42]), was developed recently by Pickenhein in [47, 48] (in the linear-quadratic case) and Tauchnitz [63] (in the general nonlinear case). The key idea of this approach is to introduce weighted Sobolev spaces as state spaces and weighted Lebesgue spaces as control spaces into the problem setting. The value of the functional is assumed to be bounded and the optimality of an admissible control $u_*(\cdot) \in L_\infty([0, \infty), \mathbb{R}^m)$ is understood in the strong sense. The developed by this approach general version of the maximum principle (see [63, Theorem 4.1])) is not necessarily normal (the case $\psi^0 = 0$ is not excluded). It involves the adjoint variable $\psi(\cdot)$ belonging to an appropriate weighted functional space (i.e., satisfying a certain exponential growth condition). In this sense this result extends [17]. The membership of the adjoint variable $\psi(\cdot)$ to the appropriate weighted functional space implies validity of both asymptotic conditions (7) and (8). However, it does not necessarily identify an adjoint function for which the maximum condition in the maximum principle is satisfied.

In the linear-quadratic case the corresponding maximum principle holds in the normal form (see [48, Theorem 5]. The normal form version of the maximum principle with adjoint variable specified by Cauchy type formula (12) is obtained in [63, Theorem 6.1] under additional "stability" type condition (see condition $(A3)$ in [63].) It is also shown in [63, Example 6.2] that all assumptions of the last result can be satisfied for an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in this example while the assumptions of Theorem 8 fail. However, the employed weighted Sobolev and Lebesgue spaces are constructed making use of a priory known optimal pair $(x_*(\cdot), u_*(\cdot))$ in this case.

The methods of general theory of extremal problems were employed to develop a variant of the maximum principle for an infinite-horizon problem with terminal and mixed control-state constraints also in the earlier paper by Brodskii [24]. The result is obtained on restrictive growth assumptions, and it does not imply normality of the problem. In the case of a free terminal state at infinity, the result involves the asymptotic condition (7).

The relationship between the maximum principle with infinite-horizon and the dynamic programming was studied in [25, 52]. In the case when the optimal value function $V(\cdot, \cdot)$ is locally Lipschitz in the state variable x, some normal versions of the maximum principle together with sensitivity relations involving generalized gradients of $V(\cdot, \cdot)$ were developed.

The normal form version of the Pontryagin maximum for problem (P) with adjoint variable specified by the Cauchy type formula (12) that presented in the present paper (see Theorem 8) was developed recently by the authors in [14, 15, 16]. The approach employed in these papers is based on the classical needle variations technique and usage of the Yankov-von Neumann-Aumann selection theorem [41, Theorem 2.14]. The main results obtained by this approach are presented in the present paper, including economic applications (Section 4). The advantage of this approach is that it can be applied under less restrictive regularity and growth assumptions than the ones akin to the approximations based technique and the methods of general theory of extremal problems. In particular, this approach can justify the Cauchy type formula (12) as a part of the Pontryagin necessary optimality condition even in the case when the optimal value is infinite. The notion of weakly overtaking optimality (see [26]) can be employed in this case.

Notice, that the importance of the Cauchy type formula (12) is justified by the fact that in some cases it can imply the standard asymptotic conditions (7), (8) and (9), or provide even more complete information about the adjoint variable $\psi(\cdot)$. In other cases, when the asymptotic conditions (7), (8) are inconsistent with the core conditions (15) and (16) of the maximum principle, the formula (12) can serve as their alternative. As it is shown in [4, 5] and Section 4, the formula (12) also provides a possibility to treat the adjoint variable $\psi(\cdot)$ as the integrated intertemporal price function.

We mention that the same approach also proved to be productive for distributed control systems, as shown in [59] for a class of age-structured optimal control problems, and for discrete-time problems with infinite horizons, as shown in [8].

Acknowledgement

The second author is financially supported by the Austrian Science Foundation (FWF) under Grant No P26640-N25.

References

[1] D. Acemoglu, *Introduction to modern economic growth*, Princeton Univ. Press, Princeton, NJ, 2008.

[2] V.M. Alexeev, V. M. Tikhomirov, S. V. Fomin, *Optimal control*, Nauka, Moscow, 1979 (Plenum, New York, 1987).

[3] K.J. Arrow, M. Kurz, *Public Investment, the Rate of Return, and Optimal Fiscal Policy*, J. Hopkins Univ. Press, Baltimore, MD, 1970.

[4] S.M. Aseev, *On some properties of the adjoint variable in the relations of the Pontryagin maximum principle for optimal economic growth problems*, Proc. Steklov Inst. Math., **287**, Suppl. 1 (2014), pp. S11–S21.

[5] S.M. Aseev, *Adjoint variables and intertemporal prices in optimal control problems with infinite time horizon*, Proc. Steklov Inst. Math., **291**, (2015), pp. 223–237.

[6] S.M. Aseev, K.O. Besov, and A.V. Kryazhimskii, *Infinite-horizon optimal control problems in economics*, Russ. Math. Surv., **67**:2 (2012), pp. 195–253.

[7] S. Aseev, K. Besov, and S. Kaniovski, *The optimal use of exhaustible resources under nonconstant returns to scale*, WIFO Working Paper, **525**, Austrian Institute of Economic Research (WIFO), Vienna, 2016, 48 pp.

[8] S.M. Aseev, M.I. Krastanov, and V.M. Veliov, *Optimality conditions for discrete-time optimal control on infinite horizon*, Pure Appl. Funct. Anal., **2**:3 (2017), pp. 395–409.

[9] S.M. Aseev, A.V. Kryazhimskii, *The Pontryagin maximum principle for an optimal control problem with a functional specified by an improper integral*, Dokl. Math. **69**:1 (2004), pp. 89–91.

[10] S.M. Aseev, A.V. Kryazhimskii, *The Pontryagin Maximum principle and transversality conditions for a class optimal control problems with infinite time horizons*, SIAM J. Control Optim., **43** (2004), pp. 1094–1119.

[11] S.M. Aseev, A.V. Kryazhimskii, *The Pontryagin maximum principle and optimal economic growth problems*, Proc. Steklov Inst. Math., **257** (2007), pp. 1–255.

[12] S.M. Aseev, A.V. Kryazhimskii, *On a class of optimal control problems arising in mathematical economics*, Proc. Steklov Inst. Math., **262** (2008), pp. 1–16.

[13] S.M. Aseev, A.V. Kryazhimskii, A.M. Tarasyev, *The Pontryagin maximum principle and transversality conditions for an optimal control problem with infinite time interval*, Proc. Steklov Inst. Math. **233** (2001), pp. 64–80.

[14] S.M. Aseev, V.M. Veliov, *Maximum principle for infinite-horizon optimal control problems with dominating discount*, Dynamics of Continuous, Discrete and Impulsive Systems. Ser. B: Applications & Algorithms, **19**:1-2 (2012), pp. 43–63.

[15] S.M. Aseev, V.M. Veliov, *Needle variations in infinite-horizon optimal control*, Variational and Optimal Control Problems on Unbounded Domains / Ed. by G. Wolansky, A. J. Zaslavski. Contemporary Mathematics, **619**, Amer. Math. Soc., Providence, RI, 2014, pp. 1–17.

[16] S.M. Aseev, V.M. Veliov, *Maximum principle for infinite-horizon optimal control problems under weak regularity assumptions*, Proc. Steklov Inst. Math., **291**, Suppl. 1 (2015), pp. S22–S39 (Trudy Inst. Mat. i Mekh. UrO RAN, **20**:3 (2014), pp. 41–57)

[17] J.-P. Aubin, F.H. Clarke, *Shadow prices and duality for a class of optimal control problems*, SIAM J. Control Optim., **17** (1979), pp. 567–586.

[18] R.J. Barro, X. Sala-i-Martin, *Economic growth*, McGraw Hill, New York, 1995.

[19] E.J. Balder, *An existence result for optimal economic growth problems*, Journal of Mathematical Analysis and Applications, **95**: 1 (1983), pp. 195–213.

[20] R. Bellman, *Dynamic programming*, Princeton Univ. Press, Princeton, NJ, 1957.

[21] L. M. Benveniste, J. A. Scheinkman, *Duality theory for dynamic optimization models of economics: the continuous time case*, J. Econ. Theory, 27 (1982), pp. 1–19.

[22] H. Benchekroun, C. Withhagen, *The optimal depletion of exhaustible resources: A complete characterization*, Resource and Energy Economics, **33**: 3 (2011), pp. 612–636.

[23] K.O. Besov, *On necessary optimality conditions for infinite-horizon economic growth problems with locally unbounded instantaneous utility function*, Proc. Steklov Inst. Math. 284 (2014), pp. 50–80.

[24] Yu.I. Brodskii, *Necessary conditions for a weak extremum in optimal control problems on an infinite time interval*, Mathematics of the USSR-Sbornik, **34**: 3 (1978), pp. 327–343.

[25] P. Cannarsa, H. Frankowska, *Value function, relaxation, and transversality conditions in infinite horizon optimal control*, Journal of Mathematical Analysis and Applications, **457** (2018), pp. 1118–1217.

[26] D.A. Carlson, A.B. Haurie, and A. Leizarowitz, *Infinite horizon optimal control. Deterministic and stochastic systems*, Springer, Berlin, 1991.

[27] D. Cass, *Optimum growth in an aggregative model of capital accumulation*, The Review of Economic Studies, **32**, no. 3 (1965), pp. 233–240.

[28] A.I. Chiang, *Elements od dynamic optimization*, MacGraw-Hill, Singapore, 1992.

[29] F.H. Clarke, *Optimization and nonsmooth analysis*, New York, A Wilye, 1983.

[30] F. Clarke, *Functional analysis, calculus of variations and optimal control*, Graduate Texts in Mathematics 264, Springer-Verlag, London, 2013.

[31] L. Cesari, *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations*, Springer, Berlin, 1959.

[32] P. Dasgupta, G.M. Heal, *The optimal depletion of exhaustible resources*, The Review of Economic Studies (Symposium on the Economics of Exhaustible Resources), **41** (1974), pp. 3–28.

[33] B.P. Demidovich, *Lectures on the mathematical theory of stability*, Nauka, Moscow, 1967 (in Russian).

[34] R. Dorfman, *An economic interpretation of optimal control theory*, Am. Econ. Rev., **59** (1969), pp. 817–831.

[35] I. Ekeland, *Some variational problems arising from mathematical economics*, Mathematical Economics, Springer, Berlin (Lect. Notes Math. **1330**), 1988, pp. 1–18.

[36] A.F. Filippov, *Differential equations with discontinuous right-hand sides*, Nauka, Moscow, 1985 (Kluwer, Dordrecht, 1988).

[37] D. Grass, J.P. Caulkins, G. Feichtinger, G. Tragler, and D.A. Behrens, *Optimal control of nonlinear processes: with applications in drugs, corruption and terror*, Springer, Berlin, 2008.

[38] H. Halkin, *Necessary conditions for optimal control problems with infinite horizons*, Econometrica, **42** (1974), pp. 267–272.

[39] P. Hartman, *Ordinary differential equations*, J. Wiley & Sons, New York, 1964.

[40] H. Hotelling, *The economics of exhaustible resources*, J. Polit. Econ., **39**:2 (1931), pp. 137–175.

[41] S. Hu, N.S. Papageorgiou, *Handbook of multivalued analysis. Volume I: Theory*, Kluwer, Dordrecht, 1997.

[42] A.D. Ioffe, V.M. Tihomirov, *Theory of extremal problems*, North-Holland Pub. Co., Amsterdam, 1979

[43] T. Kamihigashi, *Necessity of transversality conditions for infinite horizon problems*, Econometrica, 69 (2001), pp. 995–1012.

[44] T.C. Koopmans, *On the concept of optimal economic growth*, in The Econometric Approach to Development Planning, North Holland Publ. Co. and Rand McNally, a reissue of Pontificiae Academiae Scientiarum scripta varia, **28** (1965), pp. 225–300.

[45] P. Michel, *On the transversality condition in infinite horizon optimal problems*, Econometrica, **50** (1982), pp. 975–985.

[46] I.P. Natanson, *Theory of functions of a real variable*, Frederick Ungar Publishing Co., New York, 1955.

[47] S. Pickenhain, *Hilbert space treatment of optimal control problems with infinite horizon.* In: Bock, H. G.; Phu, H. X.; Rannacher R.; Schloeder, J.P.: Modeling, Simulation and Optimization of Complex Processes, HPSC 2012, Springer (2014), pp. 169–182.

[48] S. Pickenhain, *Infinite horizon optimal control problems in the light of convex analysis in Hilbert spaces.* J. Set-Valued and Variational Analysis, **23**: 1 (2015), pp. 169–189.

[49] L.S. Pontryagin, *Some mathematical problems arising in relation to the theory of optimal systems of automatic control.* In: Session of the USSR Academy of Sciences on Scientific Problems of Computer-Aided Manufacturing, October 15-20, 1956, **2**: Basic Problems of Automatic Regulation and Control (Akad. Nauk SSSR, Moscow, 1957), pp. 107–117 (in Russian).

[50] L.S. Pontryagin, V.G. Boltyanskij, R.V. Gamkrelidze, E.F. Mishchenko, *The mathematical theory of optimal processes*, Fizmatgiz, Moscow, 1961 (Pergamon, Oxford, 1964).

[51] F.P. Ramsey, *A mathematical theory of saving*, Econ. J., **38** (1928), pp. 543–559.

[52] N. Sagara, *Value functions and transversality conditions for infinite-horizon optimal control problems*, Set-Valued Analysis, **18** (2010), pp. 1–28.

[53] A. Seierstad, *Necessary conditions for nonsmooth, infinite-horizon optimal control problems*, J. Optim. Theory Appl., **103**: 1 (1999), pp. 201–229.

[54] A. Seierstad, *A maximum principle for smooth infinite horizon optimal control problems with state constraints and with terminal constraints at infinity*, Open Journal of Optimization, **4** (2015), pp. 100-130.

[55] A. Seierstad, K. Sydsæter, *Sufficient conditions in optimal control theory*, International Economic Review, **18**: 2 (1977), pp. 367–391.

[56] A. Seierstad, K. Sydsæter, *Optimal control theory with economic applications*, North-Holland, Amsterdam, 1987.

[57] A. Seierstad, K. Sydsæter, *Conditions implying the vanishing of the Hamiltonian at infinity in optimal control problems*, Optimization Letters, **3**: 4 (2009), pp. 507–512.

[58] K. Shell, *Applications of Pontryagin's maximum principle to economics*, Mathematical Systems Theory and Economics 1, Springer, Berlin, 1969, pp. 241–292 (Lect. Notes Oper. Res. Math. Econ. **11**).

[59] B. Skritek, V. M. Veliov, *On the infinite-horizon optimal control of age-structured systems*, J. Optim. Theory Appl., **167** (2015), pp. 243–271.

[60] G.V. Smirnov, *Transversality condition for infinite-horizon problems*, J. Optim. Theory Appl., **88**:3 1996, pp. 671–688.

[61] R.M. Solow, *Intergenerational equity and exhaustible resources*, The Review of Economic Studies (Symposium on the Economics of Exhaustible Resources), **41** (1974), pp. 29–45.

[62] J. Stiglitz, *Growth with exhaustible natural resources: Efficient and optimal growth paths*, The Review of Economic Studies (Symposium on the Economics of Exhaustible Resources), **41** (1974), pp. 123–137.

[63] N. Tauchnitz, *The Pontryagin maximum principle for nonlinear optimal control problems with infinite horizon*, J. Optim. Theory and Appl., **167**: 1 (2015), pp. 27–48.

[64] J. J. Ye, *Nonsmooth maximum principle for infinite-horizon problems*, J. Optim. Theory Appl., **76**: 3 (1993), pp. 485–500.