Infinite-Horizon Optimal Control with Applications in Growth Theory

S. M. Aseev

Steklov Mathematical Institute
Moscow, Russia

Moscow
2009

УДК 517.977.5+519.86(075.8)
ББК 22.16я73
А90

Данное издание одобрено программным комитетом летней школы EGMD-2009 и прошло ограниченное рецензирование.

This publication was approved by the Program Committee of the EGMD-2009 Summer School but received only limited review.

Асеев С.М.

А90 **Задачи оптимального управления с бесконечным горизонтом и их приложения в теории экономического роста:** Учебное пособие. – М.: Издательский отдел факультета ВМК МГУ имени М.В. Ломоносова (лицензия ИД № 05899 от 24.09.2001 г.); МАКС Пресс, 2009. – 148 с. (на англ. яз.)

ISBN 978-5-89407-367-5
ISBN 978-5-317-02862-6

Настоящее учебное пособие посвящено изложению основ теории оптимального управления для задач с бесконечным горизонтом и их приложениям в экономике. Излагается новый, основанный на "конечно-временных аппроксимациях", подход к получению специализированных вариантов принципа максимума Понтрягина для таких задач. В качестве примера рассмотрена неоклассическая модель оптимального экономического роста с логарифмической функцией мгновенной полезности.

Для студентов, а также аспирантов и научных работников, интересующихся применением математической теории оптимального управления в экономике.

УДК 517.977.5+519.86(075.8)
ББК 22.16я73

S.M. Aseev, Infinite-Horizon Optimal Control with Applications in Growth Theory, Moscow: MSU CMC Publications Department; MAKS Press, 2009, 148 p.

These lecture notes provide an introduction to the optimal control theory with a focus on recent achievements in the area of infinite-horizon optimal control. In particular, they present the recently developed "finite-horizon approximation" approach to deriving a modified maximum principle targeted specifically to infinite-horizon optimal control problems. The neoclassical optimal economic growth model with the logarithmic instantaneous utility function is considered as an illustrating example.

The course is targeted on researches and graduated students who are interested in application of the mathematical optimal control theory in economics.

Учебное издание

АСЕЕВ Сергей Миронович

ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ С БЕСКОНЕЧНЫМ ГОРИЗОНТОМ
И ИХ ПРИЛОЖЕНИЯ В ТЕОРИИ ЭКОНОМИЧЕСКОГО РОСТА

Издательский отдел
Факультета вычислительной математики и кибернетики МГУ им. М.В. Ломоносова
Лицензия ИД N 05899 от 24.09.01 г.

119992, ГСП-2, Москва, Ленинские горы, МГУ имени М.В. Ломоносова, 2-й учебный корпус

Напечатано с готового оригинал-макета

Издательство ООО «МАКС Пресс»
Лицензия ИД N 00510 от 01.12.99 г. Подписано к печати 16.06.2009 г.
Печать офсетная. Бумага офсетная. Формат 70х100 1/16. Усл.печ.л. 12,0. Тираж 100 экз. Изд. № 326.

119992, ГСП-2, Москва, Ленинские горы, МГУ им. М.В. Ломоносова,
2-й учебный корпус, 627 к. Тел. 939-3890, 939-3891. Тел./Факс 939-3891

Отпечатано в ППП «Типография «Наука»
121099, Москва, Шубинский пер., 6
Заказ №

ISBN 978-5-89407-367-5
ISBN 978-5-317-02862-6

Contents

Foreword		4
1.	Optimal control problems in technics and economics	6
2.	General properties of control processes	26
3.	Attainability domains	46
4.	The Pontryagin maximum principle	54
5.	Optimal economic growth problem	70
6.	Finite-horizon approximations	75
7.	Maximum principle and transversality conditions for optimal economic growth problems	85
8.	Economic interpretation of the maximum principle	95
9.	Case of dominating discount	109
10.	Monotonous case	119
11.	Maximum principle and sufficient conditions of optimality	127
12.	Example	131
Bibliography		147

Foreword

The present notes are written to serve as complementary materials to my lectures on infinite-horizon optimal control at Summer School on Economic Growth: Mathematical Dimensions which will take place at the Faculty of Computational Mathematics and Cybernetics of Moscow State University on 5–26 July, 2009. They are compiled on the base of the courses on optimal control theory those I have been teaching at this faculty in different years starting from the 1980s, and, also, on the experience of application of optimal control in economics that I have got participating in several research projects of the Dynamic Systems Program at the International Institute for Applied Systems Analysis, Laxenburg, Austria in 2001–2008. The aim of these notes is to provide participants of the school an introduction to the optimal control theory with a focus on recent achievements in the area of infinite-horizon optimal control and their applications in economic growth theory.

Infinite-horizon optimal control problems arise naturally in economics when dealing with dynamical models of optimal allocation of resources, in particular, in studies on optimization of economic growth. The goal functional to be maximized in optimal economic growth problems has a special form. It is defined by an improper integral that contains an exponential discounting factor. Therefore, usually the problems of such type are called optimal economic growth problems.

An economic model that leads to an optimal control problem of this type was first studied by F. Ramsey in the 1920s (see [**33**]). Ramsey's model had a strong impact on the formation of the modern economic growth theory (see [**1**], [**8**]), and became a prototype for a variety of further models. In the 1960s–1970s, the development of this field in economics was given great impetus by the discovery of the famous Pontryagin maximum principle. In recent years the interest to application of infinite-horizon optimal control in growth theory considerably increased. This is associated both with a general upsurge of interest in dynamical models in economics, and the fact that infinite-horizon optimal control provides a suitable (and the right) framework for exposition and development of the modern economic growth theory.

Typically, optimal economic growth problems assume that the growth process is endless, which gives rise to specific mathematical features of the Pontryagin maximum principle. The most characteristic feature is that the so-called adjoint variables (also treated as "shadow prices") may exhibit "pathological" behavior in the long run. This fact prevents application of standard versions of the Pontryagin maximum principle (associated with processes of finite durations) and appeals to developing new modifications of the Pontryagin technique, which pay special attention to the above-mentioned "pathology".

The present notes are concerned mainly with existence results and necessary optimality conditions in the form of the Pontryagin maximum principle.

The presentation of the material is organized as follows.

Sections 1–4 provide an introduction to the classical finite-horizon optimal control theory. First we study some general properties of control processes and attainability domains. Then we establish a few results on existence of an optimal admissible control. Finally, we develop the classical Pontryagin maximum principle for an optimal control problem with a free right endpoint. When studying general properties of control processes we recall also some standard facts of functional analysis. Here we follow [**28**], [**31**] and [**34**]. Proving existence results we follow mainly [**13**] and [**20**]. The exposition of the Pontryagin maximum principle is based on the classical technique of needle variations [**32**].

Sections 5–11 present a recently developed "finite-horizon approximation" approach to deriving a modified maximum principle targeted specifically to optimal economic growth problems (see [**4**]–[**6**]). The attention is focused on the characterization of the behavior of Pontryagin's adjoint variables and the Hamiltonian in a neighborhood of infinity. We prove also an existence result and discuss economic interpretation of the maximum principle. Finally, we present Arrow's theorem on sufficient conditions of optimality. For the sake of simplicity of presentation the consideration is restricted by the case when the control system is affine in control. This case is general enough but technically it is not so cumbersome as the general nonlinear one. Presentation of the finite-horizon approximation approach is based mainly on methodology developed in [**4**]–[**6**]. The economic interpretation of the maximum principle follows the standard understanding of adjoint variables as "shadow prices" (see [**16**], [**27**] and [**41**]). Presenting sufficient conditions of optimality we follow [**37**].

The neoclassical optimal economic growth model with the logarithmic instantaneous utility function is considered as an illustrating example in concluding Section 12.

These notes were written under partial support of the Russian Foundation for Basic Research (project no. 09-01-00624-a) and the Presidium of the Russian Academy of Sciences (program "Mathematical control theory").

I am grateful to A.I. Smirnov for his help in final preparation of the manuscript for publication.

Sergey Aseev
June, 2009

1. Optimal control problems in technics and economics

Originally, mathematical theory of optimal control was created in the 1950s –1960s due to the necessity to solve nonclassical dynamic optimization problems arising in technics when considering processes of control of various technical (mainly mechanical) systems, for example, such as an aircraft, a robot arm or a space station. Nevertheless, it was recognized soon, just after discovery of the famous Pontryagin maximum principle, that the new mathematical discipline provides a natural and an affective framework for treating numerous dynamic optimization problems arising not only in technics but also in other fields of science, first of all in economics.

From the formal point of view the mathematical optimal control theory is an extension and the strength of the classical calculus of variations.

The primary goal of this section is to introduce typical optimal control problems arising in technics and economics, and to discuss some of their peculiarities.

In the simplest case dynamics of a technical system can be described in terms of an ordinary differential equation with control parameter (control system) of the following form:

$$\dot{x} = f(x, u). \tag{1.1}$$

Here, $x = (x^1, \ldots, x^n) \in R^n$ and $u = (u^1, \ldots, u^m) \in R^m$ are the phase vector of the system and the vector of control parameters (control), respectively; $f : R^n \times R^m \mapsto R^n$ is a vector function of two vector variables $x \in R^n$ and $u \in R^m$. Usually, it is assumed that function $f(\cdot, \cdot)$ is continuous jointly in variables x and u, and continuously differentiable in variable x.

Assume that the control system (1.1) is considered on a time interval $[0, T]$, $T > 0$. If so, and if initial state of the system $x(0) = x_0 \in R^n$ and an admissible control $u(\cdot)$ on $[0, T]$ are given then substituting the control $u(\cdot)$ in the right-hand side of differential system (1.1) we arrive at the the following Cauchy problem:

$$\dot{x} = f(x, u(t)), \tag{1.2}$$

$$x(0) = x_0. \tag{1.3}$$

Under the standard conditions the Cauchy problem (1.2), (1.3) has a unique solution $x(\cdot)$ on some time interval $[0, \tau]$, $0 < \tau \leq T$ (see [**21**] and [**24**]). This solution $x(\cdot)$ is called an admissible trajectory of control system (1.1) corresponding to the admissible control $u(\cdot)$ on the time interval $[0,\tau]$. The pair $(x(\cdot),u(\cdot))$ is called an admissible pair (or a control process) of control system (1.1) on $[0,\tau]$.

As a rule, an admissible control $u(\cdot)$ of control system (1.1) is not an arbitrary m-dimensional vector function. In the general case, admissible control $u(\cdot)$ is a bounded (Lebesgue) measurable vector function which satisfies the following geometric constraint

$$u(t) \in U \quad \text{for all} \quad t \in [0,T]. \tag{1.4}$$

Here U is a given nonempty set in R^m. In many important applications this set U is closed and bounded (in this case U is a compactum in R^m).

If a control $u(\cdot)$ is measurable then the corresponding solution $x(\cdot)$ to the Cauchy problem (1.2), (1.3) is understood in the Carathéodory sense (see [**21**]). In this case it is an absolutely continuous n-dimensional vector function $x(\cdot)$ satisfying almost everywhere (a.e.) on some time interval $[0,\tau]$, $0 < \tau \le T$, the differential equation (1.2) and initial condition (1.3), i.e., on $[0,\tau]$ we have

$$\dot{x}(t) \overset{a.e.}{=} f(x(t),u(t))$$

and

$$x(0) = x_0.$$

In some situations, it is natural to consider a more narrow class of admissible controls consisting of all piecewise continuous m-dimensional vector functions $u(\cdot)$ on some time interval $[0,T]$, $T > 0$, satisfying the geometric constraint (1.4). In that case the corresponding admissible trajectory $x(\cdot)$ is a piecewise continuously differentiable vector function.

Remind, that the class of piecewise continuous vector functions $u(\cdot)$ on time interval $[0,T]$, $T > 0$, consists of all vector functions $u(\cdot)$ on $[0,T]$ each of which has not more than a finite number points of discontinuity, continuous at the endpoints 0 and T, and it has also finite left and right limits at each its point of discontinuity. To be definite one can assume that piecewise continuous on $[0,T]$ vector function $u(\cdot)$ is continuous from the left at each its point of discontinuity, i.e., the following equality takes place at each point τ of discontinuity of $u(\cdot)$:

$$\lim_{t \to \tau - 0} u(t) = u(\tau).$$

Nevertheless, from the point of view of rigorous analysis the class of admissible controls consisting of piecewise continuous vector functions $u(\cdot)$ on $[0,T]$, $T > 0$, is not so suitable as the class of admissible controls consisting of all bounded measurable vector functions $u(\cdot)$ on $[0,T]$. The reason is that general theorems on existence of an optimal control are established (as we will see below) only in functional spaces of measurable controls $u(\cdot)$. The desirable regularity properties (for example, the piecewise continuity or even continuity) of an optimal control $u_*(\cdot)$ are derived usually a posteriori as a result of application of necessary optimality conditions in the form of the Pontryagin maximum principle. But application of necessary optimality conditions usually assumes that an optimal admissible control exists. It should be noted that application of necessary conditions of optimality without the corresponding existence result is ambiguous.

In the general case the vector function $f(\cdot,\cdot)$ on the right-hand side of control system (1.1) can be defined not necessary on the whole space $R^n \times R^m$ but just on the Cartesian product $G \times U \subset R^n \times R^m$, where G is a nonempty open set in R^n (in particular, G can coincide with R^n) and U is a set of geometric constraints on admissible controls $u(\cdot)$ (see (1.4)). It is assumed always in this case that $x_0 \in G$ and any admissible trajectory $x(\cdot)$ of control system (1.1) takes values $x(t)$ belonging to the set G for all instants t from the time interval $[0,\tau]$ where it is defined. So, in that follows we assume that the right-hand side $f(\cdot,\cdot)$ of control system (1.1) is defined on such Cartesian product $G \times U$.

Often, in technical applications, admissible trajectories $x(\cdot)$ have satisfy a terminal constraint at the final instant $T > 0$ (which can be fixed in advance or free)

$$x(T) \in M_1, \tag{1.5}$$

where M_1 is a given nonempty terminal set in R^n such that $M_1 \bigcap G \neq \varnothing$. Of course, to satisfy inclusion (1.5), the admissible trajectory $x(\cdot)$ must be defined on the whole time interval $[0,T]$.

For example, the terminal set M_1 can consist just from a given terminal point $x_1 \in G$: $M_1 = \{x_1\}$. In this case $x(T) = x_1$, and we have a problem with a fixed right endpoint. Situation when $M_1 = R^n$ (there are no constraints on the terminal state $x(T)$ of the control system (1.1)) is called the case of a free right endpoint.

The terminal set M_1 can be defined by different ways, for example, by a finite number of inequalities:

$$M_1 = \{x \in R^n : \zeta(x) \leq 0\}.$$

Here $\zeta(\cdot)$ is a given continuously differentiable scalar or vector function.

Note, that initial constraints can be introduced analogously.

As a rule, there are many admissible pairs $(x(\cdot),u(\cdot))$ of control system (1.1), satisfying given initial and terminal constraints (if any). In this case a natural problem arise about the choice of an admissible pair $(x_*(\cdot),u_*(\cdot))$ which minimize (or maximize) a certain goal functional $J(x(\cdot),u(\cdot))$. In that case we arrive at the so called optimal control problem.

For example, one can consider the following integral goal functional:

$$J(x(\cdot),u(\cdot)) = \int_0^T f^0(x(t),u(t))\,dt, \tag{1.6}$$

where $T > 0$ is a given final instant of time, $f^0(\cdot,\cdot)$ is a scalar function of two vector variables $x \in G$ and $u \in U$, and such that (Lebesgue) integral (1.6) exists for any admissible control process $(x(\cdot), u(\cdot))$. That is true, for instance, in the case when function $f^0(\cdot,\cdot)$ is continuous and bounded on $G \times U$.

Assuming that initial and terminal states of control system (1.1) are fixed and the goal is to minimize the value of functional (1.6), we have the following statement of the optimal control problem on a fixed time interval $[0,T]$, $T > 0$, with an integral functional:

$$\dot{x} = f(x,u), \quad u \in U, \tag{1.7}$$

$$x(0) = x_0, \quad x(T) = x_1, \tag{1.8}$$

$$J(x,u) = \int_0^T f^0(x,u)\,dt \to \min. \tag{1.9}$$

Here $x_0, x_1 \in G$, vector function $f(\cdot,\cdot)$, the scalar function $f^0(\cdot,\cdot)$, the matrix function

$$\frac{\partial f(\cdot,\cdot)}{\partial x} = \left(\frac{\partial f^i(\cdot,\cdot)}{\partial x^j}\right)_{i,j=1,\dots,n},$$

and the gradient

$$\frac{\partial f^0(\cdot,\cdot)}{\partial x} = \left(\frac{\partial f^0(\cdot,\cdot)}{\partial x^1},\dots,\frac{\partial f^0(\cdot,\cdot)}{\partial x^n}\right)$$

are continuous on the Cartesian product $G \times U$. The class of admissible controls of system (1.7) consists of all (Lebesgue) measurable vector functions $u : [0,T] \mapsto U$. The measurable vector functions $u_* : [0,T] \mapsto U$ is an optimal admissible control in problem (1.7)–(1.9) if the corresponding admissible trajectory $x_*(\cdot)$ is defined on the whole time interval $[0,T]$, satisfies initial and terminal constraints (1.8), and provide the minimal possible value of the goal functional (1.9). In this case $(x_*(\cdot), u_*(\cdot))$ is called an optimal admissible pair (or an optimal control process) in problem (1.7)–(1.9).

The formulated optimal control problem (1.7)–(1.9) is a simplest typical optimal control problems arising in technics.

Note, that optimal control problem (1.7)–(1.9) can be considered also in the situation when the final instant of time $T > 0$ is free. In this case instant T became a parameter of optimization as well. The only constraint on the choice of $T > 0$ is that the boundary condition $x(T) = x_1$ (see (1.8)) must be satisfied.

In the situation when the right endpoint $x(T)$ of admissible trajectories $x(\cdot)$ is not fixed (for example, it is free or it satisfies a given terminal constraint (1.5)), one can consider a slightly more general goal functional $J(x(\cdot), u(\cdot))$ of the form

$$J(x(\cdot), u(\cdot)) = \int_0^T f^0(x(t), u(t))\,dt + V(x(T)), \tag{1.10}$$

where $V(\cdot)$ is a continuously differentiable function on G.

It is easy to see, that in the case when $V(\cdot)$ is a twice continuously differentiable function on G the goal functional (1.10) can be reduced to the functional of the form (1.9). Indeed, in this case for any admissible pair $(x(\cdot), u(\cdot))$ of control system (1.7) on $[0, T]$ the following equality takes place:

$$V(x(T)) = V(x_0) + \int_0^T \Big\langle \frac{\partial V(x(t))}{\partial x}, f(x(t), u(t)) \Big\rangle \, dt.$$

Hence, as far as initial state x_0 (see (1.8)) of the control system (1.7) is fixed we get that the optimal control problem of minimization of functional (1.10) is equivalent to the problem of minimization of the following integral functional of the form (1.9):

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

where

$$\hat{f}^0(x, u) = f^0(x, u) + \Big\langle \frac{\partial V(x)}{\partial x}, f(x, u) \Big\rangle \quad \text{for all} \quad x \in G, \ u \in U.$$

Obviously, $\hat{f}^0(\cdot, \cdot)$ is continuous in variables x and u, and continuously differentiable in variable x.

From the other hand the goal functional (1.9) can be easily reduced to the terminal one by introducing an extended state variable $\tilde{x} = (x^0, x) \in R^1 \times R^n$ and an extended control system

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u) = (f^0(x, u), f(x, u)), \quad u \in U$$

with initial condition

$$\tilde{x}(0) = (0, x_0).$$

Indeed, in this case for any control process $(x(\cdot), u(\cdot))$ of control system (1.8) we have

$$x^0(t) = \int_0^t f^0(x(s), u(s)) \, ds \quad \text{for all} \quad t \in [0, T]$$

and the goal functional (1.9) can be represented as a terminal one as follows:

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

Sometimes, in applications, it is natural to consider non autonomous optimal control problems. In this case the control system cam explicitly depend on time, i.e. it can be of the form

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

Here t is time and as in the autonomous case (see (1.1)) vector function $f : G \times [0, T] \times U \to R^n$ is continuous jointly in variables $x \in G$, $t \in [0, T]$ and $u \in U$, and continuously differentiable in variable x. In the non autonomous case both the integrand and the terminal term can also depend on time in the goal functional.

For example, a non autonomous optimal control problem on a fixed time interval $[0,T]$, $T>0$, and with a free terminal state can be formulated as follows:

$$\dot{x} = f(x,t,u), \quad u \in U, \tag{1.11}$$

$$x(0) = x_0, \tag{1.12}$$

$$J(T,x,u) = \int_0^T f^0(x,t,u)\,dt + V(x(T),T) \to \min. \tag{1.13}$$

It is easy to see that if the vector function $f(\cdot,\cdot,\cdot)$ and the scalar functions $f^0(\cdot,\cdot,\cdot)$ and $V(\cdot,\cdot)$ are continuously differentiable in (x,t) then the non autonomous optimal control problem (1.11)–(1.13) can be reduced to an autonomous one by introducing of an auxiliary phase variable $x^{n+1} \in R^1$ such that

$$\dot{x}^{n+1} = 1, \quad x^{n+1}(0) = 0.$$

In that case $x^{n+1}(t) \equiv t$, $t \geq 0$; so, time is treated just as an additional phase variable.

Note, that due to the obvious equality

$$\min_{(x(\cdot),u(\cdot))} J(x(\cdot),u(\cdot)) = -\max_{(x(\cdot),u(\cdot))} \{-J(x(\cdot),u(\cdot))\}$$

any optimal control problem of minimization of a goal functional $J(x(\cdot),u(\cdot))$ is equivalent to the corresponding problem of maximization of the goal functional

$$\tilde{J}(x(\cdot),u(\cdot)) = -J(x(\cdot),u(\cdot))$$

and vice versa.

Traditionally, optimal control problems of minimization are considered in technics when studying problems of optimal control for mechanical systems, while optimal control problems of maximization are more common in economics.

Non autonomous optimal control problem of form (1.7)–(1.9) is a direct generalization of the simplest problem of classical calculus of variations:

$$J(x) = \int_0^T f^0(x,t,\dot{x})\,dt \to \min, \tag{1.14}$$

$$x(0) = x_0, \quad x(T) = x_1. \tag{1.15}$$

Here $T > 0$ is a fixed final time.

Indeed, introducing a control system

$$\dot{x} = u, \quad u \in R^n,$$

we immediately transform the problem of calculus of variations (1.14), (1.15) in the following non autonomous optimal control problem of form (1.7)–(1.9) on the fixed time interval $[0, T]$, $T > 0$:

$$\dot{x} = u, \quad u \in R^n, \tag{1.16}$$

$$x(0) = x_0, \quad x(T) = x_1, \tag{1.17}$$

$$J(x, u) = \int_0^T f^0(x, t, u)\, dt \to \min . \tag{1.18}$$

In this case $G = R^n$ and $U = R^n$.

In the classical calculus of variations, an optimal trajectory is sought, usually, in the class of functions $x(\cdot)$ which are (at least) continuously differentiable on $[0, T]$ and satisfy initial and terminal constraints (1.15), while admissible trajectories $x(\cdot)$ of control system (1.16) are just Lipschitz continuous if the class of admissible controls in problem (1.16)–(1.17) consists of all bounded measurable vector functions $u : [0, T] \mapsto U$.

Example 1. Let us consider a trolley of mass m moving along a coordinate line l under an external force F (see Figure 1). We assume that friction is negligible. Then according to second Newton's law the net force $F(t)$ on the trolley at any instant of time $t \geq 0$ is equal to the mass of the trolley multiplied by its acceleration:

$$F(t) = m\ddot{\xi}(t). \tag{1.19}$$

Here $\xi(t)$ denote the coordinate of the trolley on the line l at the instant t.

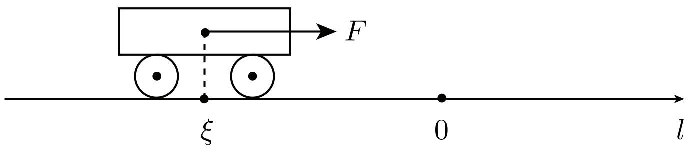

Figure 1: *Trolley.*

Assume, $m = 1$ and bounded external force can be chosen as a function of time $F(\cdot)$ on arbitrary time interval $[0, T]$, $T > 0$, such that $|F(t)| \leq 1$ for all $t \in [0, T]$. If so, and if initial coordinate $\xi(0)$ and starting velocity $\dot{\xi}(0)$ of the trolley are known then one can consider the problem of steering the trolley in a minimal time T to a given position $\xi(T)$ at which it should have a given velocity $\dot{\xi}(T)$. This problem can be easily formulated as an optimal control problem of form (1.7)–(1.9) with a free time T.

Indeed, introducing notations

$$x^1(t) = \xi(t), \quad x^2(t) = \dot{\xi}(t), \quad u(t) = F(t)$$

we can rewrite differential equation (1.19) as the following control system of form (1.7):

$$\dot{x}^1 = x^2,$$
$$\dot{x}^2 = u, \quad u \in U = [-1, 1].$$

Introducing initial and terminal conditions

$$x(0) = x_0 = (\xi(0), \dot{\xi}(0)), \quad x(T) = x_1 = (\xi(T), \dot{\xi}(T)),$$

and the goal functional

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

with $f^0(x, u) \equiv 1$, $x = (x^1, x^2) \in R^2$ and $u \in U$, we arrive at the following optimal control problem with free time $T > 0$ of form (1.7)–(1.9):

$$\dot{x}^1 = x^2, \tag{1.20}$$
$$\dot{x}^2 = u, \quad u \in [-1, 1], \tag{1.21}$$
$$x(0) = x_0, \quad x(T) = x_1, \tag{1.22}$$
$$T \to \min. \tag{1.23}$$

Here $G = R^2$ and $U = [-1, 1]$. The class of admissible controls of system (1.20), (1.21) consists of all measurable vector functions $u : [0, T] \mapsto [-1, 1]$, $T > 0$. The formulated problem (1.20)–(1.23) is a linear time-optimal problem.

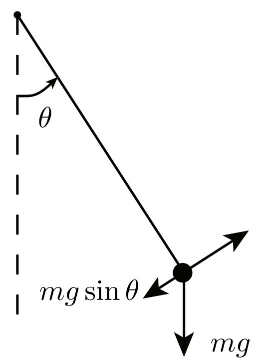

Figure 2: *Pendulum.*

Example 2. A simple pendulum consists of a particle mass m attached to a rod length l of negligible weight. We assume that friction is negligible, that all of the mass is concentrated at the end, and that the rod has unit length (see Figure 2). From Newton's law for rotating objects, there results, in terms of the variable $\theta(\cdot)$ that describes the counterclockwise angle with respect to the vertical, the following equality at any instant of time $t \geq 0$:

$$m\ddot{\theta}(t) + mg\sin\theta(t) = F(t) \tag{1.24}$$

where m is the mass, g is the acceleration due to gravity, and $F(t)$ is the value of the external torque at time t (counterclockwise being positive).

Assume, $m = 1$, $g = 1$ and bounded external torque can be chosen as a function of time $F(\cdot)$, $|F(t)| \leq 1$, on arbitrary time interval $[0, T]$, $T > 0$. If so, and if initial angle $\theta(0)$ and the starting velocity $\dot{\theta}(0)$ are known then one can consider the problem of steering the pendulum to the rest point where $\theta(T) = 0$ and $\dot{\theta}(T) = 0$ in a minimal time T.

If initial angle $\theta(0)$ and the starting velocity $\dot{\theta}(0)$ are small this time-optimal problem can be easily formulated as a linear optimal control problem of form (1.7)–(1.9).

Indeed, introducing notations

$$x^1(t) = \theta(t), \quad x^2(t) = \dot{\theta}(t), \quad u(t) = F(t)$$

we can rewrite differential equation (1.24) as the following control system of form (1.7):

$$\dot{x}^1 = x^2,$$

$$\dot{x}^2 = -\sin x^1 + u, \quad u \in U = [-1, 1].$$

Linearizing right-hand side of the control system around the rest point $(0, 0)$ and introducing initial and terminal conditions

$$x(0) = x_0 = (\theta(0), \dot{\theta}(0)), \quad x(T) = x_1 = (0, 0),$$

and the goal functional

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

with $f^0(x, u) \equiv 1$, $x = (x^1, x^2) \in R^2$ and $u \in [-1, 1]$, we arrive at the following linear time-optimal control problem of form (1.7)–(1.9):

$$\dot{x}^1 = x^2, \tag{1.25}$$

$$\dot{x}^2 = -x^1 + u, \quad u \in [-1, 1], \tag{1.26}$$

$$x(0) = x_0, \quad x(T) = (0, 0),$$

$$T \to \min.$$

Here $G = R^2$ and $U = [-1, 1]$. The class of admissible controls of system (1.25), (1.26) consists of all measurable vector functions $u : [0, T] \mapsto [-1, 1]$, $T > 0$.

Examples 1 and 2 are very famous in optimal control. Example 1 was considered by A. A. Feldbaum [**19**] and Example 2 was studied by D. W. Bushaw [**11**] in 1950s before creating the optimal control theory. The Pontryagin maximum principle provides a unified framework for treating these and various other more complex optimal control problems (see [**32**]).

Now, let us pass to optimal control problems arising in economics. In this case statements of optimal control problems have some important features.

First note that in economic models, the coordinates of the phase vector $x(t)$ are often treated as values of various factors of production that take part in the production process, such as capital (production assets) $K(t)$, materials $M(t)$, labor resources $L(t)$, etc. at the instant of time $t \geq 0$; the values of coordinates of control $u(t)$ usually characterize the current investments in the factors of production at time t. In optimal control problems arising in technics coordinates of the phase vector $x(t)$ are usually treated as generalized states of the system at time t, while values of coordinates of control vector $u(t)$ characterize forces on the system.

In economics the maximized goal functional $J(x(\cdot), u(\cdot))$ is interpreted usually as utility functional or utility index; it characterize the utility of an admissible pair $(x(\cdot), u(\cdot))$. The integral utility functional has a special form; typically it contains an exponential discounting factor $e^{-\rho t}$, $t \geq 0$, with a given discount parameter $\rho \geq 0$, i.e., in that case the integrand is the following:

$$f^0(x, t, u) = e^{-\rho t} g(x, u), \quad x \in G,\ t \geq 0,\ u \in U.$$

If $\rho = 0$ then there is no discounting in the problem.

As a rule, in economic problems the initial state of the system $x(0) = x_0$ is known, while the final state $x(T)$ is free. Indeed, typically in economic applications an optimal final state of the system can not be defined in advance and the determining of this optimal final state can constitute an important part of the problem.

For example, if the control process of an economic system is considered on a given fixed time interval $[0, T]$, $T > 0$, then the typical economic optimal control problem can be formulated as follows (compare with (1.7)–(1.9) and (1.11)–(1.13)):

$$\dot{x} = f(x, u), \quad u \in U, \tag{1.27}$$

$$x(0) = x_0, \tag{1.28}$$

$$J_T(x, u) = \int_0^T e^{-\rho t} g(x, u)\, dt + e^{-\rho T} V(X(T), T) \to \max. \tag{1.29}$$

Here $g(x(t), u(t))$ is the instantaneous utility of an admissible pair $(x(\cdot), u(\cdot))$ at the instant of time $t \geq 0$, while $V(X(T), T)$ is the current value of the final capital stock $x(T)$ at time $T > 0$. The discount factor $e^{-\rho t}$, $\rho \geq 0$, $t \in [0, T]$, in the functional (1.29), characterizes a subjective utility time preference and admits various interpretations.

Optimal control problems in economics can be formulated either at a microlevel (at the level of a firm or an enterprise) or at macrolevel (at the level of an industry, state, or at the level of world economy as a whole). In models of an enterprise, one often considers the profit gained by the enterprise in the unit time following the moment $t \geq 0$ (current profit) as the instantaneous utility $g(x(t), u(t))$. In macroeconomic models of economic growth optimization, the instantaneous utility can be defined by consumption per capita per unit of time, by the increment of consumption per capita per unit of time, etc.

In microeconomic models, the discounting factor $e^{-\rho t}$, $t \geq 0$, in the integral (1.29) is sometimes associated with inflation, the rate of devaluation of a "real" value of currency.

Indeed, suppose (1.27)–(1.29) is a problem of optimal control of an enterprise. Let $z(t)$ be the "real" value of currency at time $t \geq 0$. Without loss of generality, we assume that the real value of currency at the initial moment $t_0 = 0$ is equal to 1; i.e., we set $z(0) = 1$. Suppose that on each small time interval $[t, t + \Delta t]$, the currency is "devalued" so that the real value $z(t + \Delta t)$ of currency at time $t + \Delta t$ becomes less than its value $z(t)$ at time t by $\rho z(t)\Delta t$ up to terms of higher order of smallness than Δt. Thus, the devaluation of currency $z(t) - z(t + \Delta t)$ during a small time interval $[t, t + \Delta t]$ is approximately proportional to the value $z(t)$ at the beginning of the interval and to its length:

$$z(t) - z(t + \Delta t) = \rho z(t)\Delta t + o(\Delta t);$$

here, $\rho > 0$ is a constant "rate of inflation." Then, the real value $z(\cdot)$ of currency is given by the solution of the differential equation

$$\dot{z} = -\rho z$$

with the initial condition

$$z(0) = 1.$$

Hence we obtain

$$z(t) = e^{-\rho t}, \quad t \geq 0.$$

Let $g(x(t), u(t))$ be the instantaneous profit of the enterprise at time $t \geq 0$ (i.e., the profit of the enterprise per unit of time following the moment t) and $V(X(T), T)$ be the current value (in money terms) of the capital stock $x(T)$ of the enterprise at final instant $T > 0$. After allowing for inflation (when recalculated into prices at the initial moment 0), this instantaneous profit $g(x(t), u(t))$ and the price $V(X(T), T)$ decrease to the values $e^{-\rho t} g(x(t), u(t))$ and $e^{-\rho T} V(X(T), T)$ respectively. Thus, in this case, the value of the utility functional (1.29) gives an expression for the total profit of the enterprise at the final instant of time T with allowance for inflation (in prices at the initial moment 0).

Example 3. Let us consider an enterprise producing a homogeneous product and selling it in the market. The production assets (capital) of the enterprise are assumed to be homogeneous. The dynamics of the enterprise's capital, $x(\cdot)$, is given by

$$\dot{x} = u - \delta x, \quad u \in U = [0, u_{\max}].$$

Here $\delta > 0$ is a constant depreciation rate of capital; and variable u represents the amount of capital bought by the enterprise in one unit of time, which follows an instant t. The function $u(\cdot)$ represents an admissible capital accumulation policy of the enterprise. The velocity at which new equipment can be put into operation is bounded from above by $u_{\max} > 0$. We neglect the time needed for the installation of the new equipment. At the initial time, which we set to be zero, the value of the enterprise's capital is given: $x(0) = x_0 > 0$.

For simplicity we assume that in the enterprise, one unit of capital produces one unit of product per one unit of time. Thus, if the enterprise possesses capital $x(t)$ at time $t \geq 0$, it produces $y(t) = x(t)$ units of goods in one time unit following t.

Let $N(t)$ be the maximum amount of goods that can be sold in the market by the enterprise in one time unit following an instant t. Using a simplest model of the market, we put

$$N(t) = \frac{\bar{\pi} - \pi(t)}{b};$$

here $\pi(t) > 0$ is the price for the produced good, set by the enterprise for the time unit following t; $\bar{\pi} > 0$ is the maximal price, at which the good can still be sold in the market; $\bar{\pi}/b > 0$ is the market capacity, i.e., the maximum number of goods, which can be sold in the market in one unit of time.

We assume that all goods the enterprise is produces in the time unit following t are sold by the enterprise in the market during that unit of time (we rule out the situation where products are accumulated in stock). Thus,

$$N(t) = y(t) = x(t), \quad \pi(t) = \bar{\pi} - by(t) = \bar{\pi} - bx(t).$$

The enterprise's return from selling $y(t) = x(t)$ units of the produced good is given by

$$\pi(t)y(t) = \pi(t)x(t) = \bar{\pi}x(t) - bx^2(t).$$

Let $\zeta > 0$ be the cost of producing one unit of good and $c > 0$ be the cost of one unit of the new equipment. We assume that $a = \bar{\pi} - \zeta$ representing the maximum income per unit of good is positive. Now the current profit $g(x(t), u(t))$ of the enterprise, i.e., the profit the enterprise gains in one time unit following an instant t, is given by

$$g(x(t), u(t)) = \pi(t)y(t) - \zeta x(t) - cu(t) = ax(t) - bx^2(t) - cu(t).$$

Finally, suppose that the goal of the enterprise is to choose its admissible capital accumulation policy $u(\cdot)$ so as to maximize its total discounted profit (with rate $\rho > 0$) gained over a given finite time interval $[0, T]$, $T > 0$. Its formal setting is the following optimal capital accumulation problem:

$$\dot{x} = u - \delta x, \quad u \in U = [0, u_{\max}], \tag{1.30}$$

$$x(0) = x_0,$$

$$J_T(x(\cdot), u(\cdot)) = \int_0^T e^{-\rho t}\big[ax - bx^2 - cu\big]\, dt + e^{-\rho T} V(X(T), T) \to \max. \tag{1.31}$$

Here, $x_0 > 0$ is the given initial enterprise's capital amount, $u_{\max} > 0$ is the maximum capital allocation per unit of time, $\delta > 0$ is the capital depreciation rate, $\rho > 0$ is discount rate, $a > 0$ is the maximum income per unit of good, $\bar{\pi}/b$ is the market capacity, c is the cost of a unit of equipment and $V(x(T), T)$ represents the current value of the capital stock $x(T) > 0$ at the final time $T > 0$.

Now, let us look at the statement of problem (1.27)–(1.29) a little bit more carefully.

The first question arising here is how one should determine the final instant of time $T > 0$ (the planning horizon), and what will happen after the instant T? In economics, the control process of an industry or a state's economy is usually considered under the implicit assumption that the controlled object will exist forever, or at least indefinitely long. If so, then the optimal admissible control is aimed not only at maximizing current profit (interests of today's generation) but also at taking into account the entire future profit (interests of all future generations). This leads us to the necessity to consider the integral utility functional of the form

$$J(x(\cdot), u(\cdot)) = \int_0^\infty e^{-\rho t} g(x(t), u(t))\, dt.$$

The second question arising in consideration of problem (1.27)–(1.29) is how one should define in advance the current value $V(X(T), T)$ of the future terminal capital stock $x(T)$ at time T? If the object of control will exist forever then it is quite natural to associate with the current value $V(X(T), T)$ of $x(T)$ at time T all aggregated properly discounted future utility that can be produced by the capital stock $x(T)$ on the rest infinite time interval $[T, \infty)$, under condition that this control will be performed optimally. For instance, in the context of the model of capital accumulation considered in Example 3 the current value $V(\xi, T)$ of the capital stock $\xi > 0$ at time $T > 0$ can be defined as

$$V(\xi, T) = \max_{u(\cdot)} \int_T^\infty e^{-\rho(t-T)} \big[ax - bx^2 - cu\big] dt, \tag{1.32}$$

where $u(\cdot)$ is the enterprise capital accumulation policy on $[T, \infty)$ and value $\xi > 0$ defines initial condition

$$x(T) = \xi \tag{1.33}$$

and the process of the capital accumulation on $[0, \infty)$ is described by (1.30) (see Example 3 above). Note, that introducing a new integration variable $\tau = t - T$, $t \geq T$, in (1.32) one can show that in fact the current value $V(\xi, T)$ is time invariant; it does not depend on the second variable T, i.e. $V(\xi, T) \equiv V(\xi)$ for all $T > 0$, $\xi \in G$ where

$$V(\xi) = \max_{u(\cdot)} \int_0^\infty e^{-\rho t} \big[ax - bx^2 - cu\big] dt,$$

$u(\cdot)$ is the enterprise capital accumulation policy on $[0, \infty)$ and $x(0) = \xi$.

Substituting the current value $V(X(T), T)$ defined according with (1.32), (1.33) in (1.31) we get

$$J(x(\cdot), u(\cdot)) = \int_0^\infty e^{-\rho t} \big[ax(t) - bx^2(t) - cu(t)\big] dt.$$

So, both questions related to the statement of problem (1.27)–(1.29) leads naturally to the following infinite-horizon optimal control problem (P):

$$\dot{x} = f(x, u), \qquad u(t) \in U, \tag{1.34}$$

$$x(0) = x_0, \tag{1.35}$$

$$J(x, u) = \int_0^\infty e^{-\rho t} g(x, u) dt \to \max. \tag{1.36}$$

Here, $x = (x^1, \dots, x^n) \in R^n$ and $u = (u^1, \dots, u^m) \in R^m$ are the phase vector of the system and the vector of control parameters (control), respectively; $f : G \times U \mapsto R^n$ is a vector function of two vector variables $x \in G$ and $u \in U$. Instantaneous utility $g : G \times U \mapsto R^1$ is a scalar function of two vector variables $x \in G$ and $u \in U$; G is an open subset of R^n such that $x^0 \in G$ and all admissible trajectories take their values $x(t)$, $t \geq 0$, in G; $\rho \geq 0$ is a discount parameter. As usual we assume that both vector function $f(\cdot, \cdot)$ and scalar function $g(\cdot, \cdot)$ are continuous jointly in variables x and u, and continuously differentiable in variable x.

The formulated problem (P) can arise both in micro- and macroeconomic studies on optimization of economic growth. So, in that follows we refer to (P) as a typical economic growth problem. Below, in Sections 5–11 we specify our assumptions on problem (P) and develop the corresponding optimal control methodology.

Note that infinite horizon problem (P) (see (1.34)–(1.36)) can arise also in the situation when the control process is considered on an uncertain (random) time interval $[0,T]$, $T>0$.

Example 4. Optimal control problem (1.27)–(1.29) is formulated under assumption that the planing horizon $[0,T]$, $T>0$, is finite and known in advance. Consider the situation when is planing horizon $[0,T]$, $T>0$, is a priori unknown, and suppose that the length of this interval is a random variable. More precisely, assume that for any $t>0$, the conditional probability $\mathsf{P}(T<t+\Delta t\,|\,T\ge t)$, where Δt is small, that the control process of the system terminates at time $T\in(t,t+\Delta t]$ given that $T\ge t$, satisfies the equality

$$\mathsf{P}(T<t+\Delta t\,|\,T\ge t)=\nu\Delta t+\mathrm{o}(\Delta t), \tag{1.37}$$

where $\nu>0$ and $\mathrm{o}(\Delta t)/\Delta t\to 0$ as $\Delta t\to 0$. According to the definition of conditional probability (see [**22**]), we have

$$\begin{aligned}\mathsf{P}(T<t+\Delta t)&=\mathsf{P}(T<t)+\mathsf{P}(T<t+\Delta t\,|\,T\ge t)\mathsf{P}(T\ge t)\\&=\mathsf{P}(T<t)+\mathsf{P}(T<t+\Delta t\,|\,T\ge t)(1-\mathsf{P}(T<t)).\end{aligned}$$

Hence, by (1.37), we obtain the following relation for the cumulative distribution function $\mathsf{\Phi}(t)=\mathsf{P}(T<t)$ of the random variable T:

$$\mathsf{\Phi}(t+\Delta t)-\mathsf{\Phi}(t)=\nu(1-\mathsf{\Phi}(t))\Delta t+o(\Delta t).$$

Thus, the function $\Phi(\cdot)$ is a solution to the linear differential equation

$$\dot{\mathsf{\Phi}}=\nu(1-\mathsf{\Phi})$$

with the initial condition

$$\mathsf{\Phi}(0)=0.$$

Solving this differential equation, we find that the random variable T has the following exponential cumulative distribution function and density:

$$\mathsf{\Phi}(t)=1-e^{-\nu t},\quad \phi(t)=\dot{\mathsf{\Phi}}(t)=\nu e^{-\nu t}\quad\text{for any}\quad t\ge 0.$$

As the value of the maximized goal functional on an admissible pair $(x(\cdot),u(\cdot))$, we take the mathematical expectation $E(J_T(x(\cdot),u(\cdot)))$ of the random variable

$$J_T(x(\cdot),u(\cdot))=\int_0^T e^{-\rho t}g(x(t),u(t))dt+e^{-\rho T}V(x(T)),\quad T>0.$$

where $\rho>0$ is a discount parameter; value $g(x(t),u(t))$ characterizes the instantaneous utility of the admissible pair $(x(\cdot),u(\cdot))$ at time $t\ge 0$ and quantity $V(\xi)$ characterizes the current value of capital stock $\xi\in G$ at any instant

$T > 0$. We assume that $V(\cdot)$ is an explicitly defined scalar function satisfying for any admissible trajectory $x(\cdot)$ of control system (1.27) condition

$$\int_0^\infty e^{-(\nu+\rho)t}|V(x(t))|dt < \infty.$$

According to the definition of mathematical expectation, we have

$$\mathsf{E}(J_T(x(\cdot),u(\cdot))) = \nu \int_0^\infty \Big[e^{-\nu T} \int_0^T e^{-\rho t} g(x(t),u(t))dt\Big] dT + \nu \int_0^\infty e^{-(\nu+\rho)T} V(x(T))dT.$$

Upon integrating by parts of the first integral in equality above and assuming that any admissible pair $(x(\cdot),u(\cdot))$ satisfies the condition

$$\lim_{T\to\infty} e^{-\nu T} \int_0^T e^{-\rho t} g(x(t),u(t))\,dt = 0$$

(which is true, for example, when the function $g(\cdot,\cdot)$ is bounded), we obtain

$$\mathsf{E}(J_T(x(\cdot),u(\cdot))) = \int_0^\infty e^{-(\nu+\rho)t}\,[g(x(t),u(t)) + \nu V(x(t))]\,dt.$$

Thus, the problem of maximizing the mathematical expectation of the integral $J_T(x(\cdot),u(\cdot))$ with random upper limit leads to problem of form (P). It can be easily shown that the discount parameter $\rho > 0$ is defined by the equality $\rho = 1/\mathsf{E}(T)$, where $\mathsf{E}(T)$ is the mathematical expectation of the random variable T.

To conclude this section we consider the statement of a neoclassical model of optimal economic growth. This class of optimal economic growth models can be traced back to the original problem of Ramsey [**33**].

Example 5. A neoclassical model of optimal economic growth (see [**12**], [**40**]) describes a closed aggregated economy that produces, at any time $t \geq 0$, a single homogeneous product ($\equiv$capital) with velocity $Y(t) > 0$. At every moment t, the velocity $Y(t)$ is a function of the current values of capital $K(t) > 0$ and labor resources $L(t) > 0$; the labor resources are also assumed to be homogeneous. Thus,

$$Y(t) = F(K(t),L(t)) \quad \text{for any} \quad t \geq 0. \tag{1.38}$$

The function $F(\cdot,\cdot)$ is usually called a production function. It is assumed that the production function $F(\cdot,\cdot)$ is defined and continuous on the positive quadrant

$$G = \{(K,L) \in R^2 :\ K > 0,\ L > 0\},$$

is twice continuously differentiable, and satisfies the following "neoclassical" conditions (the Inada conditions [**26**]) for any $K > 0$ and $L > 0$:

$$\frac{\partial F(K,L)}{\partial K} > 0, \qquad \frac{\partial^2 F(K,L)}{\partial K^2} < 0, \tag{1.39}$$

$$\frac{\partial F(K,L)}{\partial L} > 0, \qquad \frac{\partial^2 F(K,L)}{\partial L^2} < 0, \tag{1.40}$$

$$\lim_{K\to+0} \frac{\partial F(K,L)}{\partial K} = \infty, \quad \lim_{K\to\infty} \frac{\partial F(K,L)}{\partial K} = 0, \tag{1.41}$$

$$\lim_{L\to+0} \frac{\partial F(K,L)}{\partial L} = \infty, \quad \lim_{L\to\infty} \frac{\partial F(K,L)}{\partial L} = 0. \tag{1.42}$$

Finally, we assume that function $F(\cdot,\cdot)$ is positively homogeneous of degree one, i.e.,

$$F(\lambda K, \lambda L) = \lambda F(K,L) \quad \text{for any} \quad \lambda > 0, \quad K > 0, \quad L > 0. \tag{1.43}$$

The last condition implies that, at any unit of time, the production volume is proportional to the factors of production available at this moment. As the production function $F(\cdot,\cdot)$, one can take, for example, the standard Cobb–Douglas function (see [**27**]) of the form

$$F(K,L) = AK^{\alpha_1}L^{\alpha_2},$$

where $A > 0$, $\alpha_1 > 0$, $\alpha_2 > 0$, and $\alpha_1 + \alpha_2 = 1$.

Note that in view of (1.43), not all the conditions (1.39)–(1.42) are independent. In particular, the second condition in (1.40) follows from the second condition in (1.39) and (1.43).

In a closed economy, the product is either invested or consumed. Suppose that, at any instant $t \geq 0$, the minimum possible part of consumed product is $\varepsilon\, Y(t) > 0$, where $0 < \varepsilon < 1$ is a constant, and the part $(1-\varepsilon)Y(t)$ of product can be arbitrarily distributed between production and consumption.

Suppose that, at instant $t \geq 0$, the part

$$I(t) = u(t)Y(t), \quad 0 \leq u(t) \leq 1-\varepsilon, \tag{1.44}$$

of the product be invested in the capital stock, while the remaining part

$$C(t) = (1-u(t))Y(t) \tag{1.45}$$

be consumed. In what follows, we will consider $u(t) \in [0, 1-\varepsilon]$ as the value of control at instant t.

Therefore, in view of (1.44), the capital dynamics $K(\cdot)$ can be described by the following differential equation:

$$\dot{K} = I - \delta K = uY - \delta K, \tag{1.46}$$

where $\delta > 0$ is the capital depreciation rate. We assume that $K(0) = K_0 > 0$ at the initial moment.

Suppose that the labor resources $L(\cdot)$ satisfy the exponential growth condition, i.e.,

$$\dot{L} = \mu L, \tag{1.47}$$

where $\mu > 0$ is a constant. We will also assume that $L(0) = L_0 > 0$.

Let $\rho > 0$ be a discount parameter and, at any instant $t \geq 0$, the instantaneous utility $g(K(t), L(t), u(t))$ of the current control process $(K(\cdot), L(\cdot), u(\cdot))$ be the logarithm of the total consumption $C(t)$, i.e.,

$$g(K(t), L(t), u(t)) = \ln C(t) = \ln(1 - u(t)) + \ln F(K(t), L(t))$$

(see (1.45)).

Note that introduction of the logarithmic function $\ln C(\cdot)$ of instantaneous utility is typical for optimal economic growth problems. This is due to the fact that in this case the goal functional (1.36) in problem (P) characterizes the dynamics of the growth rate (fractional velocity) of function $C(\cdot)$ over the infinite time interval $[0, \infty)$.

Indeed, suppose for a moment that the function $C(\cdot)$ is absolutely continuous on any finite time interval along an admissible control process $(K(\cdot), L(\cdot), u(\cdot))$. Then, the total discounted value of the growth rate of the function $C(\cdot)$ along the process $(K(\cdot), L(\cdot), u(\cdot))$ is given by the integral

$$\tilde{J}(C(\cdot)) = \int_0^\infty e^{-\rho t} \frac{\dot{C}(t)}{C(t)}\, dt. \tag{1.48}$$

Then, integrating (1.48) by parts we obtain

$$\tilde{J}(C(\cdot)) = \int_0^\infty e^{-\rho t} \frac{\dot{C}(t)}{C(t)}\, dt = -\ln C(0) + \rho \int_0^\infty e^{-\rho t} \ln C(t)\, dt.$$

Thus, on the set of all admissible control processes $(K(\cdot), L(\cdot), u(\cdot))$ such that the function $C(\cdot)$ is (locally) absolutely continuous, the maximization of the goal functional $\tilde{J}(C(\cdot))$ is equivalent to the maximization of the functional

$$J(C(\cdot)) = \int_0^\infty e^{-\rho t} \ln C(t) dt$$

with the logarithmic instantaneous utility function $\ln C(\cdot)$.

The neoclassical model of optimal economic growth (with the logarithmic function of instantaneous utility) is formulated as the following optimal control problem (P_ε), $0 < \varepsilon < 1$:

$$\dot{K} = uF(K,L) - \delta K, \quad u \in U_\varepsilon = [0, 1-\varepsilon],$$

$$\dot{L} = \mu L,$$

$$K(0) = K_0, \quad L(0) = L_0,$$

$$J(K,L,u) = \int_0^\infty e^{-\rho t}\big[\ln(1-u) + \ln F(K,L)\big]\,dt \to \max. \tag{1.49}$$

It is easy to see that problem (P_ε) is a particular case of problem (P).

When studying the neoclassical optimal economic growth problem, one usually reduces the dimension of the system using the homogeneity condition (1.43) and passes to an auxiliary phase variable $x = K/L$ (capital per unit of labor force) and a one-factor production function $f(\cdot)$ of the form $f(x) = F(x,1)$, $x > 0$. In this case, in view of conditions (1.38) and (1.43), one has

$$\frac{Y(t)}{L(t)} = \frac{1}{L(t)}F(K(t),L(t)) = F\Big(\frac{K(t)}{L(t)},1\Big) = f(x(t))$$

for any $t \geq 0$. The function $f(\cdot)$ is defined and continuous on $\tilde{G} = (0,\infty)$. By conditions (1.39),

$$\frac{d}{dx}f(x) > 0, \quad \frac{d^2}{dx^2}f(x) < 0 \tag{1.50}$$

for any $x > 0$, and in view of (1.39)–(1.43),

$$\lim_{x\to+0} f(x) = 0, \quad \lim_{x\to\infty} f(x) = \infty, \quad \lim_{x\to+0}\frac{d}{dx}f(x) = \infty, \quad \lim_{x\to\infty}\frac{d}{dx}f(x) = 0. \tag{1.51}$$

Equalities (1.46) and (1.47) yield the following expression for the variable $x(t) = K(t)/L(t)$:

$$\dot{x}(t) = \frac{d}{dt}\frac{K(t)}{L(t)} = \dot{K}(t)\frac{1}{L(t)} - \frac{K(t)}{L^2(t)}\dot{L}(t) = u(t)\frac{Y(t)}{L(t)} - (\mu+\delta)\frac{K(t)}{L(t)},$$

which implies the equality

$$\dot{x}(t) = u(t)f(x(t)) - (\mu+\delta)x(t)$$

in view of the definition of the variable $x(\cdot)$ and conditions (1.38) and (1.43). The instantaneous consumption per unit of labor resource at time $t \geq 0$ is given by $c(t) = C(t)/L(t)$. According to (1.38) and (1.45), we obtain

$$c(t) = (1-u(t))\frac{Y(t)}{L(t)} = (1-u(t))f(x(t)).$$

Note that in view of (1.47) the labor $L(\cdot)$ in this model is subject to the predetermined dynamics. Therefore, the maximization of the integral functional

(1.49) is equivalent to the maximization of the functional

$$J(x(\cdot),u(\cdot)) = \int_0^\infty e^{-\rho t}\ln c(t)\,dt = \int_0^\infty e^{-\rho t}\big[\ln(1-u(t)) + \ln f(x(t))\big]\,dt,$$

which characterizes the aggregated growth rate of consumption per unit of labor force.

Thus, in terms of the phase variable $x(\cdot)$, the optimal control problem (P_ε), $0<\varepsilon<1$, reduces to the following problem $(\tilde{P}_\varepsilon)$:

$$\dot{x} = uf(x) - (\mu+\delta)x, \quad u \in U_\varepsilon = [0, 1-\varepsilon],$$

$$x(0) = x_0,$$

$$J(x,u) = \int_0^\infty e^{-\rho t}\big[\ln(1-u) + \ln f(x)\big]\,dt \to \max.$$

Here, $x_0 \in \tilde{G} = (0,\infty)$.

A qualitative description of optimal regimes in problem (P_ε) (or, equivalently, in $(\tilde{P}_\varepsilon)$) in the case of an arbitrary sufficiently small $\varepsilon > 0$ is presented in Section 12.

2. General properties of control processes

In this section we consider some general properties of control processes for the following non autonomous control system on a finite time interval $[0, T]$, $T > 0$:

$$\dot{x} = f(x, t, u), \quad u \in U \tag{2.1}$$

with a given initial state

$$x(0) = x_0. \tag{2.2}$$

Here, $t \geq 0$ is time, $x \in R^n$ is a phase variable, $u \in R^m$ is a control parameter; $x_0 \in G$, where G is a given nonempty open subset of R^n; U is a nonempty compactum in R^m; $f : G \times [0, T] \times U \mapsto R^n$ is a vector function of the phase variable $x \in G$, time $t \in [0, T]$ and the control variable $u \in U$. It is assumed that the vector function $f(\cdot, \cdot, \cdot)$ is continuous jointly in variables x, t and u, and continuously differentiable in variable x. The class of admissible controls for system (2.1) consists of all (Lebesgue) measurable functions $u : [0, T] \mapsto U$.

Substituting an admissible control $u(\cdot)$ in the right-hand side of differential equation (2.1) we arrive at the the following Cauchy problem (see also (1.2), (1.3)):

$$\dot{x} = f(x, t, u(t)), \tag{2.3}$$

$$x(0) = x_0. \tag{2.4}$$

In the general case the control $u(\cdot)$ is measurable and hence the right-hand side $f(x, \cdot, u(\cdot))$ of differential equation (2.3) is a measurable function of variable t as well.

In the case when the right-hand side $f(x, \cdot, u(\cdot))$ is a continuous function of time t the Cauchy problem (2.3), (2.4) is equivalent to integral equation

$$x(t) = x_0 + \int_0^t f(x(s), s, u(s)) ds. \tag{2.5}$$

Analogously, in the situation when the right-hand side $f(x, \cdot, u(\cdot))$ is measurable in variable t, one can consider the integral in (2.5) in the Lebesgue sense. If so, when we get a definition of solution to the Cauchy problem (2.3), (2.4) in the Carathéodory sense. In this case the Cauchy problem (2.3), (2.4) is considered as equivalent to the integral equation (2.5), i.e. a solution to the Cauchy problem (2.3), (2.4) is an absolutely continuous vector function $x : [0, \tau] \mapsto R^n$ which is defined on some interval $[0, \tau] \subset [0, T]$, $\tau > 0$, and satisfies integral equation (2.5) on $[0, \tau]$. Equivalently, the solution to the Cauchy problem

(2.3), (2.4) is an absolutely continuous vector function $x : [0, \tau] \mapsto R^n$ which defined on some interval $[0, \tau] \subset [0, T]$, $\tau > 0$, satisfies initial condition (2.4) and the following equality holds a.e. on $[0, \tau]$:

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

By the standard theorem on existence and uniqueness there is a unique (Carathéodory) solution $x(\cdot)$ to Cauchy problem (2.3), (2.4) on some interval $[0, \tau]$, $0 < \tau < T$, taking values $x(t)$ in G for all $t \in [0, \tau]$ (see [**21**]). This solution $x(\cdot)$ is called an admissible trajectory of control system (2.1) corresponding to admissible control $u(\cdot)$ on the time interval $[0, \tau]$. In this case $(x(\cdot), u(\cdot))$ is called an admissible pair or a control process of the control system (2.1) on the time interval $[0, \tau]$.

Due to the theorem on existence and uniqueness any admissible trajectory $x(\cdot)$ can be (uniquely) extended to the right on some larger time interval $[0, \tau']$, $[0, \tau] \subset [0, \tau'] \subset [0, T]$ in the open set G. Continuing this process we get that two situations are possible: either trajectory $x(\cdot)$ can be extended to the right in the open set G on the whole time interval $[0, T]$ where the corresponding admissible control $u(\cdot)$ is defined, or trajectory $x(\cdot)$ can be extended in the set G only on some "maximal" time semi-interval $[0, \tau'')$, $0 < \tau'' \leq T$.

In that follows we assume that the following condition for control system (2.1) with initial condition (2.2) holds:

(A1) *For any admissible control $u(\cdot)$ the corresponding admissible trajectory $x(\cdot)$ of the control system* (2.1) *can be extended in open set G on the whole time interval $[0, T]$ where the control $u(\cdot)$ is defined. Moreover, there is a nonempty closed subset Π of the set G such that for any admissible trajectory $x(\cdot)$ of control system* (2.1) *its values $x(t)$ belong to the set Π for all $t \in [0, T]$.*

The next condition guarantee a uniform boundedness of admissible trajectories $x(\cdot)$ of control system (2.1) on the time interval $[0, T]$:

(A2) *There exists a $C_0 \geq 0$ such that*

$$\langle x, f(x, t, u)\rangle \leq C_0(1 + \|x\|^2) \quad \textit{for any} \quad x \in \Pi,\ t \in [0, T],\ u \in U. \qquad (2.6)$$

Lemma 1. *Suppose conditions* (A1) *and* (A2) *hold. Then for any admissible trajectory $x(\cdot)$ of control system* (2.1) *the following inequality takes place:*

$$\|x(t)\| \leq \sqrt{1 + \|x(0)\|^2} e^{C_0 t} \quad \textit{for all} \quad t \in [0, T]. \qquad (2.7)$$

Proof. As far as $x(\cdot)$ is an absolutely continuous vector function on $[0, T]$ the scalar function $z : [0, T] \mapsto R^1$, $z(t) = \|x(t)\|^2$, $t \in [0, T]$, is also absolutely continuous. Hence, due to (2.1) and (2.6) a.e. on $[0, T]$ we have

$$\frac{d}{dt}\|x(t)\|^2 = 2\langle x(t), \dot{x}(t)\rangle = 2\langle x(t), f(x(t), t, u(t))\rangle \leq 2C_0(1 + \|x(t)\|^2). \quad (2.8)$$

The last condition implies inequality (2.7).

Indeed, consider the following absolutely continuous scalar function ϕ : $[0,T] \mapsto R^1$:

$$\phi(t) = e^{-2C_0 t}\|x(t)\|^2, \quad t \in [0,T]. \tag{2.9}$$

Then $\phi(0) = \|x(0)\|^2$ and a.e. on $[0,T]$ we have

$$\dot{\phi}(t) = -2C_0 e^{-2C_0 t}\|x(t)\|^2 + e^{-2C_0 t}\frac{d}{dt}\|x(t)\|^2$$

From this due to inequality (2.8) we get

$$\dot{\phi}(t) \leq -2C_0 e^{-2C_0 t}\|x(t)\|^2 + 2C_0 e^{-2C_0 t}(1 + \|x(t)\|^2) = 2C_0 e^{-2C_0 t}.$$

Hence, for all $t \in [0,T]$ the following relations hold:

$$\begin{aligned}\phi(t) = \phi(0) + \int_0^t \dot{\phi}(s)ds &\leq \|x(0)\|^2 + 2C_0\int_0^t e^{-2C_0 s}ds \\ &= 1 + \|x(0)\|^2 - e^{-2C_0 t} \leq 1 + \|x(0)\|^2.\end{aligned}$$

If so, then due to the definition of function $\phi(\cdot)$ (see (2.9)) we get

$$e^{-2C_0 t}\|x(t)\|^2 \leq 1 + \|x(0)\|^2 \quad \text{for all} \quad t \in [0.T]$$

or

$$\|x(t)\|^2 \leq (1 + \|x(0)\|^2)e^{2C_0 t} \quad \text{for all} \quad t \in [0,T],$$

or

$$\|x(t)\| \leq \sqrt{1 + \|x(0)\|^2}e^{C_0 t} \quad \text{for all} \quad t \in [0,T].$$

So, condition (2.7) is true. ■

As an immediate consequence of Lemma 1 we get that validity of both conditions (A1) and (A2) implies a uniform boundedness of admissible trajectories $x(\cdot)$ of control system (2.1) on $[0,T]$. Indeed, condition (2.7) gives a uniform estimate on possible values $x(t)$ for any admissible trajectories $x(\cdot)$ at any instant $t \in [0,T]$. So, there is a constant $M_0 \geq 0$ such that for any admissible trajectories $x(\cdot)$ the following inequality takes place:

$$\|x(t)\| \leq M_0 \quad \text{for all} \quad t \in [0,T]. \tag{2.10}$$

Another direct consequence of Lemma 1 is a uniform Lipschitzness of admissible trajectories of control system (2.1) on $[0,T]$ in the case when both conditions (A1) and (A2) are satisfied. Indeed, as far as the set U of geometric constraints is a compact subset of R^m, and due to conditions (A1) and (2.10) any admissible process $(x(\cdot),u(\cdot))$ takes values $(x(t),u(t))$, $t \in [0,T]$, in the compact set

$$\{B_{M_0}(0) \cap \Pi\} \times U \subset G \times U$$

in this case. Here

$$B_{M_0}(0) = \{x \in R^n : \|x\| \leq M_0\}$$

is a closed boll of radius $M_0 \geq 0$ with centrum 0 in R^n. Further, due to continuity of function $f(\cdot,\cdot,\cdot)$ on the compact set $\{B_{M_0}(0) \cap \Pi\} \times [0,T] \times U$ there is a $M_1 \geq 0$ such that

$$M_1 = \max_{x \in \left\{B_{M_0}(0) \cap \Pi\right\}, t \in [0,T], u \in U} \|f(t,x,u)\|.$$

Then for any admissible trajectory $x(\cdot)$ of control system (2.1) and arbitrary $\tau_1, \tau_2 \in [0,T]$ we get

$$\begin{aligned} \|x(\tau_2) - x(\tau_1)\| &= \left\| \int_{\tau_1}^{\tau_2} f(x(t),t,u(t))dt \right\| \\ &\leq \int_{\tau_1}^{\tau_2} \|f(x(t),t,u(t))\| \, dt \leq M_1 \|\tau_2 - \tau_1\|. \end{aligned}$$

So, any admissible trajectory $x(\cdot)$ of control system (2.1) satisfies Lipschitz condition

$$\|x(\tau_2) - x(\tau_1)\| \leq M_1 \|\tau_2 - \tau_1\| \quad \text{for any} \quad \tau_1, \tau_2 \in [0,T]. \tag{2.11}$$

with the same constant $M_1 \geq 0$.

In that follows $C([0,T], R^n)$, $T > 0$, denote the space of continuous vector functions $x: \ [0,T] \mapsto R^n$ with the uniform norm

$$\|x(\cdot)\|_C = \max_{t \in [0,T]} \|x(t)\|.$$

Conditions (2.10), (2.11) and the Arzelà theorem (see [**28**]) imply the following result.

Lemma 2. *Suppose conditions* (A1) *and* (A2) *hold. Then any sequence* $\{x_k(\cdot)\}$, $k = 1,2,\dots,$ *of admissible trajectories* $x_k(\cdot)$ *of control system* (2.1) *has a subsequence* $\{x_{k_i}(\cdot)\}$, $i = 1,2,\dots,$ *such that* $\{x_{k_i}(\cdot)\}$, $i = 1,2,\dots,$ *converges uniformly on* $[0,T]$ *to an absolutely continuous vector function* $\bar{x}: [0,T] \mapsto R^n$, *i.e.*

$$\lim_{i \to \infty} x_{k_i}(\cdot) = \bar{x}(\cdot) \quad in \quad C([0,T], R^n).$$

In other words the set $\{x(\cdot)\}$ *of all admissible trajectories* $x(\cdot)$ *of control system* (2.1) *is a precompact set in the space* $C([0,T], R^n)$.

Proof. Indeed, due to Lemma 1 any sequence $\{x_k(\cdot)\}$, $k = 1,2,\dots,$ of admissible trajectories $x_k(\cdot)$ of control system (2.1) is uniformly bounded on $[0,T]$ (see (2.10)). Moreover, due to (2.11) the sequence $\{x_k(\cdot)\}$, $k = 1,2,\dots,$ is equicontinuous on $[0,T]$. So, it satisfies conditions of the Arzelà theorem. Hence, there is a subsequence $\{x_{k_i}(\cdot)\}$, $i = 1,2,\dots,$ of sequence $\{x_k(\cdot)\}$, $k = 1,2,\dots,$ such that $\{x_{k_i}(\cdot)\}$ converges uniformly to continuous vector function

$\bar{x}(\cdot)$ on $[0,T]$ as $i \to \infty$. Due to (2.11) for arbitrary $i = 1, 2, \dots$ and any $\tau_1, \tau_2 \in [0,T]$ the following inequality takes place:

$$\|x_{k_i}(\tau_1) - x_{k_i}(\tau_2)\| \le M_1 \|\tau_2 - \tau_1\|.$$

Taking a limit at the right hand side of the above inequality as $i \to \infty$ we get

$$\|\bar{x}(\tau_1) - \bar{x}(\tau_2)\| \le M_1 \|\tau_2 - \tau_1\|, \quad \text{for any} \quad \tau_1, \tau_2 \in [0,T].$$

Thus, vector function $\bar{x}(\cdot)$ is Lipschitz continuous and, hence, it is absolutely continuous on $[0,T]$. ■

The following condition (if it is valid together with conditions (A1) and (A2)) guarantees that any limit point $\bar{x}(\cdot)$ (in the uniform metric of $C([0,T],R^n)$) of the set of all admissible trajectories $x(\cdot)$ of control system (2.1) on $[0,T]$ is again an admissible trajectory of control system (2.1).

(A3) *For any $x \in \Pi$ and $t \in [0,T]$ the set of all possible velocities (called a vectogram) of control system* (2.1)

$$F(x,t) = \bigcup_{u \in U} f(x,t,u) \tag{2.12}$$

is convex.

Note, that as far as U is a compact subset of R^m and the vector function $f(x,t,\cdot)$ is continuous in variable u the vectogram $F(x,t)$ (see (2.12)) is a compactum in R^n for any fixed $x \in G$ and $t \in [0,T]$. So, if assumption (A3) takes place then the vectogram $F(x,t)$ is a convex compactum in R^n for any fixed $x \in G$ and $t \in [0,T]$.

Remind the definitions (and a few related facts) of functional spaces $L^p([0,T],R^n)$, $T > 0$, $p \ge 1$, of (Lebesgue) measurable vector functions $\xi : [0,T] \mapsto R^n$ (see, for example, [**28**] and [**34**]).

We say that a measurable vector functions $\xi : [0,T] \mapsto R^n$ belongs to the space $L^p([0,T],R^n)$, $T > 0$, $1 \le p < \infty$, if

$$\int_0^T \|\xi(t)\|^p dt < \infty.$$

A measurable vector function $\xi : [0,T] \mapsto R^n$ belongs to the space $L^\infty([0,T],R^n)$, $T > 0$, if

$$\inf\{M \ge 0 : \ \|\xi(t)\| \le M \quad \text{a.e. on} \quad [0,T]\} < \infty.$$

For any $p \ge 1$ the space $L^p([0,T],R^n)$ is a normed vector space; the norm $\|\xi(\cdot)\|_{L^p}$ in $L^p([0,T],R^n)$, $p \ge 1$ is defined by equality

$$\|\xi(\cdot)\|_{L^p} = \Big(\int_0^T \|\xi(t)\|^p dt\Big)^{1/p} \quad \text{if} \quad 1 \le p < \infty$$

and

$$\|\xi(\cdot)\|_{L^\infty} = \operatorname*{ess\,supp}_{t\in[0,T]} \|\xi(t)\| \quad \text{if} \quad p = \infty.$$

Here

$$\operatorname*{ess\,supp}_{t\in[0,T]} \|\xi(t)\| = \inf\{M \geq 0 : \ \|\xi(t)\| \leq M \quad \text{a.e. on} \quad [0,T]\}$$

is the essential upper bound of the measurable vector function $\xi : [0,T] \mapsto R^n$.

Note, that $\|\xi(\cdot)\|_{L^p} = 0$, $p \geq 1$, if and only if $\xi(t) = 0$ a.e. on $[0,T]$.

It is easy to see that

$$L^\infty([0,T],R^n) \subset L^p([0,T],R^n) \subset L^r([0,T],R^n) \quad \text{for any} \quad 1 \leq r < p < \infty. \tag{2.13}$$

Here the first inclusion is obvious. Indeed, the space $L^\infty([0,T],R^n)$ consists of all essentially bounded measurable vector functions $\xi : [0,T] \mapsto R^n$. If so, then $\xi(\cdot) \in L^p([0,T],R^n)$ for any $1 \leq p < \infty$.

Let us prove the second inclusion in (2.13). If $\xi(\cdot) \in L^p([0,T],R^n)$ and $1 \leq r < p < \infty$ then

$$\|\xi(t)\|^r \leq \|\xi(t)\|^p + 1 \quad \text{for all} \quad t \in [0,T].$$

Hence,

$$\int_0^T \|\xi(t)\|^r dt \leq \int_0^T (\|\xi(t)\|^p + 1)\, dt < \infty.$$

So, inclusion $\xi(\cdot) \in L^p([0,T],R^n)$ implies the membership $\xi(\cdot) \in L^r([0,T],R^n)$ for any $1 \leq r < p < \infty$ as well.

Thus the functional space $L^1([0,T],R^n)$ is the "maximal" space among all functional spaces $L^p([0,T],R^n)$, $p \geq 1$ while the functional space $L^\infty([0,T],R^n)$ is the "minimal" one.

Recall that by the definition admissible controls $u(\cdot)$ of the control system (2.1) are measurable vector functions taking values $u(t)$ in the compact set U in R^m for all $t \in [0,T]$. Since any compact set in R^m is bounded, admissible controls $u(\cdot)$ are uniformly bounded on $[0,T]$ and hence all of them belong to the functional space $L^\infty([0,T],R^n)$. Thus, admissible controls $u(\cdot)$ can be treated as elements of any functional space $L^p([0,T],R^m)$ for arbitrary $p \geq 1$.

The convergence of a sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^p([0,T],R^n)$, $k = 1,2,\dots$, to a measurable vector function $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^p([0,T],R^n)$, in the norm of the space $L^p([0,T],R^n)$, $p \geq 1$, is called a strong convergence (in the space $L^p([0,T],R^n)$). In this case we write

$$\|\xi_k(\cdot) - \xi(\cdot)\|_{L^p} \to 0 \quad \text{as} \quad k \to \infty.$$

If $1 \leq r < p$ then condition

$$\|\xi_k(\cdot) - \xi(\cdot)\|_{L^p} \to 0 \quad \text{as} \quad k \to \infty \tag{2.14}$$

implies

$$\|\xi_k(\cdot) - \xi(\cdot)\|_{L^r} \to 0 \quad \text{as} \quad k \to \infty. \tag{2.15}$$

In the case $p = \infty$ this fact is obvious. Let us prove that (2.14) implies (2.15) in the case $1 \le r < p < \infty$ as well. In this case we have

$$\frac{p}{r} > 1 \quad \text{and} \quad \|\xi_k(\cdot) - \xi(\cdot)\|^r \in L^{\frac{p}{r}}([0,T], R^1) \quad \text{for any} \quad k = 1, 2, \ldots.$$

Further, the integral Hölder inequality (see, for example, [**34**]) says that

$$\int_0^T x(t)y(t)dt \le \Big(\int_0^T x^\alpha(t)dt\Big)^{\frac{1}{\alpha}} \Big(\int_0^T y^{\frac{\alpha}{\alpha-1}}(t)dt\Big)^{\frac{\alpha-1}{\alpha}}$$

for any $x(\cdot) \in L^\alpha([0,T], R^1)$, $y(\cdot) \in L^{\frac{\alpha}{\alpha-1}}([0,T], R^1)$ and $\alpha > 1$. From this, putting $x(t) = \|\xi_k(t) - \xi(t)\|^r$, $y(t) \equiv 1$ on $[0,T]$ and $\alpha = p/r$, we get

$$\int_0^T \|\xi_k(t) - \xi(t)\|^r dt \le T^{\frac{p-r}{p}} \left(\int_0^T \Big(\|\xi_k(t) - \xi(t)\|^r\Big)^{\frac{p}{r}} dt\right)^{\frac{r}{p}} \quad \text{for any } k = 1, 2, \ldots$$

So, condition (2.14) implies (2.15) in the case $1 \le r < p < \infty$ as well.

Note that in the general case the contrary assertion is wrong, i.e., if $1 \le r < p$ and a sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^r([0,T], R^n)$, $k = 1, 2, \ldots$, converges strongly in $L^r([0,T], R^n)$ to a measurable vector function $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^r([0,T], R^n)$, then the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, does not necessary converge to $\xi(\cdot)$ in the space $L^p([0,T], R^n)$. Even if all vector functions $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, are essentially bounded their strong limit $\xi(\cdot)$ in the space $L^r([0,T], R^n)$ can simply not belong to the space $L^p([0,T], R^n)$ in this case.

Nevertheless, if a sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, is uniformly essentially bounded on $[0,T]$, i.e. there is a constant $M \ge 0$ such that a.e. on $[0,T]$ we have

$$\|\xi_k(t)\| \le M \quad \text{for any} \quad k = 1, 2, \ldots, \tag{2.16}$$

then the strong convergence of the sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^r([0,T], R^n)$, $k = 1, 2, \ldots$, to a measurable vector function $\xi(\cdot) \in L^r([0,T], R^n)$ implies the strong convergence of the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, in all other functional space $L^p([0,T], R^n)$, for arbitrary $1 \le r < p < \infty$ excluding the case $p = \infty$. Indeed in this case due to (2.16) we have a.e. on $[0,T]$ we have

$$\|\xi_k(t) - \xi(t)\|^{p-r} \le (\|\xi_k(t)\| + \|\xi(t)\|)^{p-r} \le (2M)^{p-r} \quad \text{for any} \quad k = 1, 2, \ldots$$

Hence

$$\int_0^T \|\xi_k(t) - \xi(t)\|^p dt \le \int_0^T \|\xi_k(t) - \xi(t)\|^r \|\xi_k(t) - \xi(t)\|^{p-r}\, dt$$

$$\le (2M)^{p-r} \int_0^T \|\xi_k(t) - \xi(t)\|^r\, dt < \infty.$$

Thus, since admissible controls $u(\cdot)$ are uniformly bounded on $[0,T]$, this fact implies the equivalence of strong convergence of a sequence $\{u_k(\cdot)\}$, $k = 1,\dots,$ of admissible controls $u_k(\cdot)$ in all functional spaces $L^p([0,T],R^n)$ for any $1 \le p < \infty$.

One can prove (see, for example [**28**]) that if a sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1([0,T],R^n)$, $k = 1,2,\dots,$ converges strongly in $L^1([0,T],R^n)$ to a measurable vector function $\xi(\cdot) \in L^1([0,T],R^n)$ then there is a subsequence $\{\xi_{k_i}(\cdot)\}$, $i = 1,2,\dots$ of the sequence $\{\xi_k(\cdot)\}$, $k = 1,2,\dots,$ such that a.e. on $[0,T]$ we have

$$\xi_{k_i}(t) \to \xi(t) \quad \text{as} \quad i \to \infty.$$

Below apart from the notion of strong convergence in the functional spaces $L^p([0,T],R^n)$ we will use also a notion of weak convergence in $L^1([0,T],R^n)$ of a sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1([0,T],R^n)$, $k = 1,2,\dots,$ to a vector function $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^1([0,T],R^n)$.

We say that a sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1([0,T],R^n)$, $k = 1,2,\dots,$ converges weakly in $L^1([0,T],R^n)$ to a vector function $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^1([0,T],R^n)$, if for any vector function $h : [0,T] \mapsto R^n$, $h(\cdot) \in L^\infty([0,T],R^n)$, the following equality holds:

$$\lim_{k\to\infty} \int_0^T \langle \xi_k(t), h(t)\rangle dt = \int_0^T \langle \xi(t), h(t)\rangle \, dt. \tag{2.17}$$

Here due to the essential boundedness of function $h(\cdot)$, both integrals in (2.17) exist.

It is easy to see that the strong convergence in $L^1([0,T],R^n)$ of a sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1([0,T],R^n)$, $k = 1,2,\dots,$ to a vector function $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^1([0,T],R^n)$ implies its weak convergence to the same vector function $\xi(\cdot)$ in $L^1([0,T],R^n)$. In the general case the opposite assertion is wrong. Nevertheless, the following result takes place.

Lemma 3. *Suppose a sequence* $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1([0,T],R^n)$, $k = 1,2,\dots,$ *converges weakly in* $L^1([0,T],R^n)$ *to a vector function* $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^1([0,T],R^n)$. *Suppose also that functions* $\xi_k(\cdot)$, $k = 1,2,\dots,$ *are uniformly essentially bounded on* $[0,T]$, *i.e. there is a constant* $M \ge 0$ *such that condition* (2.16) *takes place and*

$$\|\xi_k(\cdot)\|_{L^2} \to \|\xi(\cdot)\|_{L^2} \quad \textit{as} \quad k \to \infty. \tag{2.18}$$

Then

$$\|\xi_k(\cdot) - \xi(\cdot)\|_{L^2} \to 0 \quad \textit{as} \quad k \to \infty. \tag{2.19}$$

Proof. For any $k = 1,2,\dots$ and $t \in [0,T]$ we have

$$\|\xi_k(t) - \xi(t)\|^2 = \|\xi_k(t)\|^2 - 2\langle \xi_k(t), \xi(t)\rangle + \|\xi(t)\|^2. \tag{2.20}$$

Due to (2.16) the function $\xi(\cdot)$ belongs to the space $L^\infty([0,T],R^n)$. Hence, due to the weak convergence in $L^1([0,T],R^n)$ of the sequence $\{\xi_k(\cdot)\}$, $k = 1,2,\dots,$

to the vector function $\xi(\cdot)$ we have

$$\lim_{k\to\infty}\int_0^T \langle \xi_k(t), \xi(t)\rangle dt = \int_0^T \langle \xi(t), \xi(t)\rangle dt = \int_0^T \|\xi(t)\|^2 dt.$$

Thus by (2.18), (2.19) and (2.20) we get

$$\lim_{k\to\infty}\int_0^T \|\xi_k(t) - \xi(t)\|^2\, dt = 0. \qquad \blacksquare$$

Note that since in assumptions of Lemma 3 the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, is uniformly essential bounded both conditions (2.18) and (2.19) are equivalent to analogous conditions where the norm in the space $L^2([0,T], R^n)$ is replaced by the norm in any other space $L^p([0,T], R^n)$ with $1 \le p < \infty$. In particular, if a uniformly essentially bounded sequence $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1([0,T], R^n)$, $k = 1, 2, \ldots$, converges weakly in $L^1([0,T], R^n)$ to a vector function $\xi : [0,T] \mapsto R^n$, $\xi(\cdot) \in L^1([0,T], R^n)$ and

$$\|\xi_k(\cdot)\|_{L^1} \to \|\xi(\cdot)\|_{L^1} \quad \text{as} \quad k \to \infty$$

then

$$\|\xi_k(\cdot) - \xi(\cdot)\|_{L^1} \to 0 \quad \text{as} \quad k \to \infty.$$

Although the weak convergence in $L^1[0,T], R^n)$ of a sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, $\xi_k(\cdot) \in L^1[0,T], R^n)$ to a vector function $\xi(\cdot) \in L^1([0,T], R^n)$ is weaker than the strong one sometimes, using the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, it is possible to construct another sequence $\{z_k(\cdot)\}$, $z_k(\cdot) \in L^1([0,T], R^n)$, $k = 1, 2, \ldots$, which converges to $\xi(\cdot)$ already strongly in $L^1([0,T], R^n)$ as $k \to \infty$.

The following result is a variant of the Banach–Saks theorem (see[**34**]).

Theorem 1. *Let a sequence* $\{\xi_k(\cdot)\}$, $\xi_k(\cdot) \in L^1[0,T], R^n)$, $k = 1, 2, \ldots$, *be uniformly essential bounded on* $[0,T]$, *i.e. there is a constant* $M \ge 0$ *such that condition* (2.16) *takes place. Suppose also that the sequence* $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, *converges weakly in* $L^1([0,T], R^n)$ *to a vector function* $\xi(\cdot) \in L^1([0,T], R^n)$. *Then there is a subsequence* $\{\xi_{k_i}(\cdot)\}$, $i = 1, 2, \ldots$, *of the sequence* $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, *such that the sequence* $\{z_N(\cdot)\}$, $N = 1, 2, \ldots$, *of its arithmetic means*

$$z_N(t) = \frac{1}{N}\sum_{i=1}^{N} \xi_{k_i}(t), \quad t \in [0,T], \tag{2.21}$$

converges strongly in $L^2([0,T], R^n)$ *to the vector function* $\xi(\cdot)$ *as* $N \to \infty$, *i.e.*

$$\lim_{N\to\infty}\int_0^T \|z_N(t) - \xi(t)\|^2\, dt = 0. \tag{2.22}$$

Note, that since in assumptions of Theorem 1 the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \dots$, is uniformly essential bounded the sequence of arithmetic means (see (2.21)) $\{z_N(\cdot)\}$, $N = 1, 2, \dots$, is also uniformly essential bounded. Hence, the strong convergence in $L^2([0,T], R^n)$ of the sequence $\{z_N(\cdot)\}$, $N = 1, 2, \dots$, to the vector function $\xi(\cdot)$ as $N \to \infty$ is equivalent to the strong convergence of $\{z_N(\cdot)\}$, $N = 1, 2, \dots$, to $\xi(\cdot)$ in any other space $L^p([0,T], R^n)$ for arbitrary $1 \le p < \infty$ as $N \to \infty$. In particular (2.22) implies that

$$\lim_{N\to\infty} \int_0^T \|z_N(t) - \xi(t)\| \, dt = 0.$$

Proof of Theorem 1. Without loss of generality one can assume that $\xi(t) \equiv 0$ for $t \in [0,T]$. Otherwise one can take the sequence $\{\xi_k(\cdot) - \xi(\cdot)\}$, $k = 1, 2, \dots$

Define the sequence $\{k_i\}$, $i = 1, 2, \dots$, of natural numbers k_i inductively. For $i = 1$ put $k_1 = 1$ and define k_2 as a minimal natural among all $k > k_1$ such that

$$\int_0^T \langle \xi_k(t), \xi_{k_1}(t) \rangle dt \le 1.$$

Such natural k_2 exists because the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \dots$, converges weakly in $L^1([0,T], R^n)$ to $\xi(\cdot)$.

Suppose numbers $k_1, \dots, k_i$, $i \ge 2$, are already selected. Then define k_{i+1} as a minimal natural among all $k > k_i$ such that

$$\int_0^T \langle \xi_k(t), \xi_{k_1}(t) \rangle dt \le \frac{1}{i}, \quad \dots, \quad \int_0^T \langle \xi_k(t), \xi_{k_i}(t) \rangle dt \le \frac{1}{i}.$$

Again, such number k_{i+1} exists because the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \dots$, converges weakly in $L^1([0,T], R^n)$ to $\xi(\cdot)$.

Thus the sequence $\{k_i\}$, $i = 1, 2, \dots$, is defined.

Now, consider the sequence $\{z_N\}$, $N = 1, 2, \dots$, defined by (2.21) and show that (2.22) takes place.

Indeed, the norms $\|\xi_k(t)\|$, $k = 1, 2, \dots$, are uniformly essentially bounded on $[0,T]$, i.e., there is a constant $M \ge 0$ such that (see (2.16)) a.e. on $[0,T]$

$$\|\xi_k(t)\| \le M, \quad \text{for any} \quad k = 1, 2, \dots$$

Hence,

$$\|z_N(\cdot)\|^2_{L^2} = \int_0^T \left\| \frac{\xi_{k_1}(t) + \dots + \xi_{k_N}(t)}{N} \right\|^2 dt$$
$$= \frac{\int_0^T \langle \xi_{k_1}(t) + \dots + \xi_{k_N}(t), \xi_{k_1}(t) + \dots + \xi_{k_N}(t) \rangle dt}{N^2}$$

$$\leq \frac{\left(NM^2 + 2\cdot 1 + 4\cdot \frac{1}{2} + \dots + 2(N-1)\frac{1}{N-1}\right)T}{N^2}$$
$$< \frac{(M^2+2)T}{N} \to 0 \quad \text{as} \quad N \to \infty.$$

Thus by Lemma 3 the sequence $\{z_N(\cdot)\}$, $N = 1, 2, \dots$, converges weakly in $L^2([0,T], R^n)$ to 0. The theorem is proved. ■

Lemma 4. *Suppose conditions* (A1) *and* (A2) *hold and* $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, *be a sequence of admissible trajectories* $x_k : [0,T] \mapsto R^n$, $T > 0$, *of control system* (2.1) *which is convergent uniformly on* $[0,T]$ *to an absolutely continuous vector function* $\bar{x} : [0,T] \to R^n$. *Then the sequence of their derivatives* $\{\dot{x}_k(\cdot)\}$, $k = 1, 2, \dots$, *converges weakly in* $L^1([0,T], R^n)$ *to derivative* $\dot{\bar{x}}(\cdot)$ *of the vector function* $\bar{x}(\cdot)$ *as* $k \to \infty$.

Proof. Let $h : [0,T] \mapsto R^n$ be an arbitrary essentially bounded vector function and $\varepsilon > 0$ be an arbitrary positive. Then there is a continuously differentiable vector function $h_\varepsilon : [0,T] \mapsto R^n$ such that

$$\|h_\varepsilon(\cdot) - h(\cdot)\|_{L^\infty} \leq \varepsilon. \tag{2.23}$$

As far as admissible trajectories $x_k(\cdot)$, $k = 1, 2, \dots$, are absolutely continuous vector functions the following equality holds a.e. on $[0,T]$:

$$\frac{d}{dt}\langle x_k(t), h_\varepsilon(t)\rangle = \langle \dot{x}_k(t), h_\varepsilon(t)\rangle + \langle x_k(t), \dot{h}_\varepsilon(t)\rangle.$$

Integrating this equality on $[0,T]$ for arbitrary $k = 1, 2, \dots$, we get

$$\langle x_k(T), h_\varepsilon(T)\rangle - \langle x_k(0), h_\varepsilon(0)\rangle = \int_0^T \langle \dot{x}_k(t), h_\varepsilon(t)\rangle dt + \int_0^T \langle x_k(t), \dot{h}_\varepsilon(t)\rangle\, dt. \tag{2.24}$$

As far as sequence $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, convergence uniformly to $\bar{x}(\cdot)$ on $[0,T]$ there is a limit of the left hand side of equality (2.24) and of the second integral in its right-hand side. Hence, the limit of the first integral in the right-hand side of (2.24) also exists. If so, then

$$\lim_{k\to\infty} \int_0^T \langle \dot{x}_k(t), h_\varepsilon(t)\rangle dt = \langle \bar{x}(T), h_\varepsilon(T)\rangle - \langle \bar{x}(0), h_\varepsilon(0)\rangle + \int_0^T \langle \bar{x}(t), \dot{h}_\varepsilon(t)\rangle\, dt \tag{2.25}$$

Analogously to (2.24), the following equality takes place for absolutely continuous vector function $\bar{x}(\cdot)$ a.e. on $[0,T]$:

$$\langle \bar{x}(T), h_\varepsilon(T)\rangle - \langle \bar{x}(0), h_\varepsilon(0)\rangle = \int_0^T \langle \dot{\bar{x}}(t), h_\varepsilon(t)\rangle dt + \int_0^T \langle \bar{x}(t), \dot{h}_\varepsilon(t)\rangle\, dt$$

and, hence,

$$\int_0^T \langle \dot{\bar{x}}(t), h_\varepsilon(t)\rangle dt = \langle \bar{x}(T), h_\varepsilon(T)\rangle - \langle \bar{x}(0), h_\varepsilon(0)\rangle + \int_0^T \langle \bar{x}(t), \dot{h}_\varepsilon(t)\rangle\, dt. \quad (2.26)$$

Comparing the right-hand sides of equalities (2.25) and (2.26) we arrive at equality

$$\lim_{k\to\infty} \int_0^T \langle \dot{x}_k(t), h_\varepsilon(t)\rangle dt = \int_0^T \langle \dot{\bar{x}}(t), h_\varepsilon(t)\rangle\, dt. \quad (2.27)$$

Further, by Lemma 1 there is a constant $M_1 \geq 0$ (see (2.11)) such that a.e. on $[0,T]$

$$\|\dot{x}_k(t)\| \leq M_1, \quad k = 1, 2, \ldots, \quad (2.28)$$

and (as far as $x_k(\cdot) \to \bar{x}(\cdot)$ in $C([0,T], R^n)$ as $k \to \infty$), simultaneously,

$$\|\dot{\bar{x}}(t)\| \leq M_1. \quad (2.29)$$

Due to (2.27) and (2.28) there is a $N(\varepsilon) > 0$ (depending only on ε) such that

$$\left\| \int_0^T \langle \dot{x}_k(t), h_\varepsilon(t)\rangle dt - \int_0^T \langle \dot{\bar{x}}(t), h_\varepsilon(t)\rangle dt \right\| \leq \varepsilon \quad \text{for all} \quad k \geq N(\varepsilon). \quad (2.30)$$

Further,

$$\int_0^T \langle \dot{x}_k(t), h(t)\rangle dt = \int_0^T \langle \dot{x}_k(t), h_\varepsilon(t)\rangle dt + \int_0^T \langle \dot{x}_k(t), h(t) - h_\varepsilon(t)\rangle\, dt$$

where due to (2.23) and (2.28)

$$\left\| \int_0^T \langle \dot{x}_k(t), h(t) - h_\varepsilon(t)\rangle\, dt \right\| \leq \varepsilon M_1 T.$$

Hence,

$$\left\| \int_0^T \langle \dot{x}_k(t), h(t)\rangle dt - \int_0^T \langle \dot{x}_k(t), h_\varepsilon(t)\rangle dt \right\| \leq \varepsilon M_1 T, \quad k = 1, 2, \ldots \quad (2.31)$$

Analogously, due to (2.23) and (2.29)

$$\left\| \int_0^T \langle \dot{\bar{x}}(t), h(t)\rangle dt - \int_0^T \langle \dot{\bar{x}}(t), h_\varepsilon(t)\rangle dt \right\| \leq \varepsilon M_1 T. \quad (2.32)$$

From (2.30), (2.31) and (2.32) we get

$$\left\| \int_0^T \langle \dot{x}_k(t), h(t)\rangle dt - \int_0^T \langle \dot{\bar{x}}(t), h(t)\rangle dt \right\| \leq \varepsilon(1 + 2M_1 T) \quad \text{for all} \quad k \geq N(\varepsilon). \quad (2.33)$$

Finally, as far as inequality (2.33) holds for arbitrary $\varepsilon > 0$ we get

$$\lim_{k\to\infty} \int_0^T \langle \dot{x}_k(t), h(t)\rangle dt = \int_0^T \langle \dot{\bar{x}}(t), h(t)\rangle dt. \qquad \blacksquare$$

The weak convergence in the space $L^1([0,T],R^n)$ of a uniformly integrable bounded sequence $\{\xi_k(\cdot)\}$, $k=1,2,\dots$, to a measurable vector function $\xi(\cdot)$ provides also the following pointwise characterization of the limit function $\xi(\cdot)$.

Theorem 2. *Suppose a sequence $\{\xi_k(\cdot)\}$, $k=1,2,\dots$, of measurable vector functions $\xi_k : [0,T] \mapsto R^n$, $\xi_k(\cdot) \in L^1([0,T],R^n)$, $T>0$, converges weakly in $L^1([0,T],R^n)$ to a measurable vector function $\bar{\xi} : [0,T] \mapsto R^n$, $\bar{\xi}(\cdot) \in L^1([0,T],R^N)$ as $k\to\infty$, and there is an integrable positive scalar function $\eta : [0,T] \mapsto R^1$ such that for any $k=1,2,\dots$ the following condition takes place:*

$$\|\xi_k(t)\| \le \eta(t) \quad a.e. \ on \quad [0,T]. \tag{2.34}$$

Then the following pointwise characterization of the vector function $\bar{\xi}(\cdot)$ on $[0,T]$ holds:

$$\bar{\xi}(t) \overset{a.e.}{\in} \overline{\text{conv}} \bigcup_{i\ge k} \xi_i(t) \quad for\ any \quad k=1,2,\dots$$

Proof. Denote

$$F_k(t) = \overline{\text{conv}} \bigcup_{i\ge k} \xi_i(t), \quad t\in[0,T], \quad k=1,2,\dots \tag{2.35}$$

and show that a.e. on $[0,T]$ we have

$$\bar{\xi}(t) \in F_k(t) \quad \text{for any} \quad k=1,2,\dots \tag{2.36}$$

Assume the opposite. Then there are a natural $k_* \in \{1,2,\dots\}$ and a set $\mathfrak{M} \subset [0,T]$, $\operatorname{meas}\mathfrak{M} > 0$, such that

$$\bar{\xi}(t) \notin F_{k_*}(t) \quad \text{for all} \quad t\in\mathfrak{M}. \tag{2.37}$$

Then, due to convexity and compactness of the sets $F_{k_*}(t)$, $t\in[0,T]$, (see (2.34)) and by the separation theorem (see, for example, [**35**]) for any $t\in\mathfrak{M}$ where (2.37) holds, there is a $\zeta(t)\neq 0$, $\zeta(t)\in R^n$, such that

$$\langle \bar{\xi}(t), \zeta(t)\rangle > H(F_{k_*}(t), \zeta(t)).$$

Here

$$H(A,\psi) = \sup_{a\in A}\langle a,\psi\rangle$$

is the support function of a nonempty set $A\subset R^n$ in direction $\psi\in R^n$ (see, for example, [**35**]).

Note, that the support function $H(F_k(\cdot),\psi)$ is measurable for any fixed $\psi\in R^n$ and $k=1,2,\dots$. Moreover, for any fixed $\psi\in R^n$ and $k=1,2,\dots$ we have

$$|H(F_k(t),\psi)| \le \eta(t)\|\psi\| \quad \text{a.e. on} \quad [0,T]. \tag{2.38}$$

Indeed, due to its definition (see (2.35)) for any fixed $\psi \in R^n$ and $k = 1, 2, \dots$ the support function $H(F_k(\cdot), \psi)$ is measurable as a pointwise supremum of a sequence of measurable functions $\{\langle \xi_i(\cdot), \psi \rangle\}$, $i = k, k+1, \dots$:

$$H(F_k(t), \psi) = \sup_{i \geq k} \langle \xi_i(t), \psi \rangle, \quad t \in [0, T].$$

In view of (2.34) this fact implies (2.38).

So, $H(F_k(\cdot), \psi) \in L^1([0, T], R^n)$ for any fixed $\psi \in R^n$ and $k = 1, 2, \dots$

Let $\{\zeta_i\}$, $i = 1, 2, \dots$, be a denumerable everywhere dense subset of R^n. Then due to continuity of support function $H(F_{k_*}(t), \cdot)$ for any fixed $t \in \mathfrak{M}$ there is a corresponding $i_* \in \{1, 2, \dots\}$ such that

$$\langle \bar{\xi}(t), \zeta_{i_*} \rangle > H(F_{k_*}(t), \zeta_{i_*}).$$

Consider the following denumerable set of subsets of $\mathfrak{M}$:

$$\mathfrak{M}_{\mathfrak{i}} = \{t \in \mathfrak{M} : \ \langle \bar{\xi}(t), \zeta_i \rangle > H(F_{k_*}(t), \zeta_i)\}, \quad i = 1, 2, \dots$$

Here all sets $\mathfrak{M}_{\mathfrak{i}}$, $i = 1, 2, \dots$, are measurable and

$$\mathfrak{M} = \bigcup_{i=1,2,\dots} \mathfrak{M}_i.$$

Hence, there is an index $i_* \in \{1, 2, \dots\}$ such that $\operatorname{meas} \mathfrak{M}_{i_*} > 0$ and

$$\langle \dot{\bar{x}}(t), \zeta_{i_*} \rangle > H(F_{k_*}(t), \zeta_{i_*}) \quad \text{for all} \quad t \in \mathfrak{M}_{i_*}. \tag{2.39}$$

Define a characteristic function $\chi : [0, T] \mapsto \{0, 1\}$ of measurable set $\mathfrak{M}_{i_*}$ as follows

$$\chi(t) = \begin{cases} 1 & \text{if} \quad t \in \mathfrak{M}_{i_*}, \\ 0 & \text{if} \quad t \notin \mathfrak{M}_{i_*}. \end{cases}$$

Note that $\zeta_{i_*} \chi(\cdot) \in L^\infty([0, T], R^n)$.

Due to the weak convergence of sequence $\{\xi_i(\cdot)\}$, $i = 1, 2, \dots$, to the vector function $\bar{\xi}(\cdot)$ in $L^1([0, T], R^n)$, inequality $\operatorname{meas} \mathfrak{M}_{i_*} > 0$ and condition (2.39) we get

$$\lim_{i \to \infty} \int_0^T \langle \xi_i(t), \zeta_{i_*} \chi(t) \rangle \, dt = \int_{\mathfrak{M}_{i_*}} \langle \bar{\xi}(t), \zeta_{i_*} \rangle \, dt > \int_{\mathfrak{M}_{i_*}} H(F_{k_*}(t), \zeta_{i_*}) \, dt. \tag{2.40}$$

From the other hand, due to (2.35) points $\xi_i(t)$, $i = k_*, k_* + 1, \dots$, belongs to the set $F_{k_*}(t)$ for any $t \in [0, T]$. Hence, for arbitrary $i = k_*, k_* + 1, \dots$ we have

$$\langle \xi_i(t), \zeta_{i_*} \rangle \leq H(F_{k_*}(t), \zeta_{i_*}) \quad \text{a.e. on} \quad [0, T].$$

Hence

$$\lim_{i \to \infty} \int_0^T \langle \xi_i(t), \zeta_{i_*} \chi(t) \rangle dt = \lim_{i \to \infty} \int_{\mathfrak{M}_{i_*}} \langle \xi_i(t), \zeta_{i_*} \rangle dt \leq \int_{\mathfrak{M}_{i_*}} H(F_{k_*}(t), \zeta_{i_*}) \, dt.$$

So, we have got the contradiction with (2.40). Hence, condition (2.36) holds a.e on $[0, T]$. ■

The next result is a specialized measurable selection theorem. It is well known in the optimal control literature as Filippov's lemma (see [**20**]).

Theorem 3. *Suppose $\xi : [0,T] \to G$ is a continuous vector function, $\zeta : [0,T] \mapsto R^n$ is a measurable one and the following inclusion holds a.e. on $[0,T]$:*

$$\zeta(t) \in \bigcup_{u\in U} f(\xi(t), t, u).$$

Then there is a measurable vector function $u : [0,T] \mapsto U$ such that a.e. on $[0,T]$

$$\zeta(t) = f(\xi(t), t, u(t)). \tag{2.41}$$

Proof. Consider the multivalued mapping $U(\cdot)$ defined a.e. on [0,T] by equality

$$U(t) = \{u \in U : \ \zeta(t) = f(\xi(t), t, u)\}.$$

Put $U(t) = U$ on the rest subset of $[0,T]$ of measure 0. So, in this case the multivalued mapping $U(\cdot)$ is defined on the whole time interval $[0,T]$.

As far as vector function $f(\xi, t, \cdot)$ is continuous for any fixed $t \in [0,T]$ and $\xi \in G$, and U is a compactum in R^m, the sets $U(t)$, $t \in [0,T]$, are compact. Define $u : \ [0,T] \mapsto U$, $u(t) = (u^1(t), \dots, u^m(t))$, so that the first coordinate $u^1(t)$ is the smallest possible. If there is more than one such point $u(t) \in U(t)$, it is required that $u^2(t)$ be the smallest among these, and further that $u^3(t)$ be the smallest among these, so forth, to define a unique vector $u(t) \in U$ for every $t \in [0,T]$. Obviously, this function $u : [0,T] \mapsto U$ satisfies a.e. on $[0,T]$ equality (2.41). Let us prove that $u(\cdot)$ is measurable.

Suppose coordinates $u^1(\cdot), \dots, u^{r-1}(\cdot)$, $1 \le r \le n$, are measurable on $[0,T]$ (if $r = 1$, nothing is assumed here) and prove that coordinate $u^r(\cdot)$ is measurable function as well. For arbitrary $\varepsilon > 0$ due to the Lusin theorem (see [**34**]) there is a compact set $\Delta_\varepsilon \subset [0,T]$ such that $\mathrm{meas}([0,T]\backslash\Delta_\varepsilon) \le \varepsilon$, and functions $u^1(\cdot), \dots, u^{r-1}(\cdot)$, $1 \le r \le n$, and $\zeta(\cdot)$ are continuous on Δ_ε. Let $\alpha \in R^1$ be an arbitrary real number. Show that the set

$$\Delta_{\varepsilon,\alpha} = \{\tau \in \Delta_\varepsilon : \ u^r(\tau) \le \alpha\}$$

is closed. Suppose the opposite. Then there exists a sequence $\{\tau_i\}$, $\tau_i \in \Delta_{\varepsilon,\alpha}$, $i = 1, 2, \dots$, such that $\tau_i \to \bar{\tau}$ as $i \to \infty$ and $\bar{\tau} \notin \Delta_{\varepsilon,\alpha}$, i.e. $u^r(\bar{\tau}) > \alpha$. Since U is compact there is a convergent subsequence of sequence $\{u(\tau_i)\}$, $i = 1, 2, \dots$ Without loss of generality we can assume that the sequence $\{u(\tau_i)\}$, $i = 1, 2, \dots$, itself converges to a vector $\hat{u} \in U$. By continuity we have

$$\lim_{i\to\infty} \dot{\bar{x}}(\tau_i) = \dot{\bar{x}}(\hat{\tau}), \quad \lim_{i\to\infty} u^j(\tau_i) = u^j(\hat{\tau}), \quad j = 1, \dots, r-1.$$

Thus, we have $\zeta(\hat{\tau}) = f(\xi(\hat{\tau}), \hat{\tau}, \hat{u})$. But $\hat{u}^r \le \alpha < u^r(\hat{\tau})$, which contradicts the definition of $u^r(\tau)$. Therefore, function $u^r(\cdot)$ is measurable on Δ_ε. By the Lusin theorem there exists a compact subset $\Delta_{2\varepsilon} \subset \Delta_\varepsilon$ such that meas$(\Delta_\varepsilon \backslash \Delta_{2\varepsilon}) \leq \varepsilon$ and function u^r is continuous on $\Delta_{2\varepsilon}$. Since ε is an arbitrary positive number, applying the Lusin theorem again, we get that $u^r(\cdot)$ is measurable on $[0,T]$. Hence, the vector function $u(\cdot)$ is measurable on $[0,T]$ as well. The theorem is proved. ■

Now we are ready to establish sufficient conditions for compactness of the set of admissible trajectories of control system (2.1) in the space $C([0,T],R^n)$, $T>0$.

Theorem 4. *Suppose conditions* (A1), (A2) *and* (A3) *hold and* $\{x_k(\cdot)\}$, $k=1,2,\ldots$, *be a sequence of admissible trajectories* $x_k : [0,T] \mapsto R^n$, $T>0$, *of control system* (2.1). *Then there is a subsequence* $\{x_{k_i}(\cdot)\}$, $i=1,2,\ldots$, *of sequence* $\{x_k(\cdot)\}$ $k=1,2,\ldots$, *such that* $\{x_{k_i}(\cdot)\}$, $i=1,2,\ldots$, *convergent uniformly on* $[0,T]$ *to an admissible trajectory* $\bar{x}(\cdot)$ *of control system* (2.1) *as* $i \to \infty$.

Proof. Let $\{x_k(\cdot)\}$, $k=1,2,\ldots$, be a sequence of admissible trajectories $x_k : [0,T] \mapsto R^n$, $T>0$, of control system (2.1). Then due to Lemma 2 there is a subsequence $\{x_{k_i}(\cdot)\}$, $i=1,2,\ldots$, of sequence $\{x_k(\cdot)\}$ $k=1,2,\ldots$, such that $\{x_{k_i}(\cdot)\}$, $i=1,2,\ldots$, converges uniformly on $[0,T]$ to an absolutely continuous vector function $\bar{x} : [0,T] \mapsto R^n$ as $i \to \infty$. As far as all trajectories $x_{k_i}(\cdot)$, $i=1,2,\ldots$, satisfy the initial condition (2.2) the absolutely continuous vector function $\bar{x}(\cdot)$ also satisfies (2.2). So, to complete the proof of the theorem it is sufficient to show that there is an admissible control $u : [0,T] \mapsto U$ such that a.e. on $[0,T]$ the following equality takes place:

$$\dot{\bar{x}}(t) = f(t, \bar{x}(t), u(t)). \tag{2.42}$$

To prove (2.42) we apply Filippov's lemma (see Theorem 3).

The sequence $\{x_{k_i}(\cdot)\}$, $i=1,2,\ldots$, converges uniformly on $[0,T]$ to the absolutely continuous vector function $\bar{x} : [0,T] \mapsto R^n$ as $i \to \infty$. Hence, due to Lemma 4 the sequence $\{\dot{x}_{k_i}(\cdot)\}$, $i=1,2,\ldots$, of their derivatives converges weakly in $L^1([0,T],R^n)$ to derivative $\dot{\bar{x}}(\cdot)$ of vector function $\bar{x}(\cdot)$ as $i \to \infty$. Further, by Theorem 2 on $[0,T]$ we have:

$$\dot{\bar{x}}(t) \overset{a.e.}{\in} \bigcup_{i \geq k} \overline{\text{conv}} \bigcup_{u \in U} f(x_k(t), t, u) \quad \text{for any} \quad k = 1, 2, \ldots \tag{2.43}$$

Due to continuity of vector function $f(t,\cdot,u)$ condition (2.43) implies that on $[0,T]$

$$\dot{\bar{x}}(t) \overset{a.e.}{\in} \overline{\text{conv}} \bigcup_{u \in U} f(\bar{x}(t), t, u).$$

By virtue of Theorem 3 there is an admissible control $u : [0,T] \mapsto U$ such that a.e. on $[0,T]$ equality (2.42) takes place. ■

To conclude this section we derive a few results related to the situation when the right-hand side of the control system (2.1) is affine in the control variable u. In this case the control system (2.1) has the following form:

$$\dot{x} = f_0(x,t) + \sum_{j=1}^{m} f_j(x,t)u^j, \quad u \in U. \tag{2.44}$$

Here U is a compact subset of R^m and all vector functions $f_j : G \times [0,T] \mapsto R^n$, $j = 0,\dots,m$, are continuous in variables $x \in G$ and $t \in [0,T]$, $T > 0$, and continuously differentiable in variable x. As usual we assume that initial state $x(0) = x_0 \in G$ is fixed.

The case of the affine in control system (2.44) is general enough. From one side the affine in control systems arise often in economic applications. From the other side any general nonlinear control system (2.1) (under some suitable conditions) can be approximated with arbitrary accuracy by a sequence of control systems affine in control (see [**6**] for more details).

From the technical point of view the case affine in control systems (2.44) is considerably easier for their formal analysis than the general nonlinear one.

The following result follows from Theorem 4 immediately.

Lemma 5. *Let the set U of geometric constraints in control system* (2.44) *be a nonempty convex compactum in R^m. Then any sequence $\{u_k(\cdot)\}$, $k = 1,2,\dots,$ of admissible controls $u_k : [0,T] \mapsto U$ has a subsequence $\{u_{k_i}(\cdot)\}$, $i = 1,2,\dots,$ which converges weakly in $L^1([0,T],R^N)$ to an admissible control $\bar{u} : [0,T] \mapsto U$ as $k \to \infty$.*

Proof. Indeed, consider the following control system on time interval $[0,T]$, $T > 0$:

$$\dot{x} = u, \quad u \in U \tag{2.45}$$

with initial state

$$x(0) = 0.$$

It is easy to see that conditions (A1), (A2) and (A3) are satisfied for the control system (2.45). Then by Theorem 4 the sequence $\{x_k(\cdot)\}$, $k = 1,2,\dots,$ of admissible trajectories $x_k(\cdot)$:

$$x_k(t) = \int_0^t u_k(s)\,ds, \quad t \in [0,T], \quad k = 1,2,\dots$$

of control system (2.45) has a subsequence $\{x_{k_i}(\cdot)\}$, $i = 1,2,\dots,$ which converges uniformly on $[0,T]$ to an admissible trajectory $\bar{x}(\cdot)$ as $i \to \infty$. Due to (2.45) a.e. on $[0,T]$ we have $\dot{x}_k(t) = u_k(t)$, $k = 1,2,\dots,$ and there is an admissible control $\bar{u}(\cdot)$ such that a.e. on $[0,T]$ equality $\dot{\bar{x}}(t) = \bar{u}(t)$ takes place. So, due to Lemma 4 we get that the sequence $\{u_k(\cdot)\}$, $k = 1,2\dots,$ converges weakly in $L^1([0,T],R^n)$ to the admissible control $\bar{u}(\cdot)$ as $k \to \infty$. ∎

Lemma 6. *Suppose conditions* (A1) *and* (A2) *are satisfied, and the set* U *of geometric constraints in control system* (2.44) *is a nonempty convex compactum in* R^m. *Suppose a sequence* $\{u_k(\cdot)\}$, $k = 1, 2, \dots$, *of admissible controls* $u_k : [0, T] \mapsto U$, $T > 0$, *converges weakly in* $L^1([0, T], R^n)$ *to an admissible control* $\bar{u}(\cdot)$ *as* $k \to \infty$. *Then the sequence* $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, *of admissible trajectories* $x_k : [0, T] \mapsto G$ *of the control system* (2.44) *corresponding to the sequence of controls* $\{u_k(\cdot)\}$, $k = 1, 2, \dots$, *converges uniformly to the trajectory* $\bar{x} : [0, T] \mapsto G$ *corresponding to the control* $\bar{u}(\cdot)$ *as* $k \to \infty$.

Proof. Due to Theorem 4 there is a subsequence $\{x_{k_i}(\cdot)\}$, $i = 1, 2, \dots$, of sequence $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, such that $\{x_{k_i}(\cdot)\}$, $i = 1, 2, \dots$, converges uniformly on $[0, T]$ to an admissible trajectory $\bar{x} : [0, T] \mapsto G$. Let $\{x_{k_i}(\cdot)\}$, $i = 1, 2, \dots$, be an arbitrary such convergent subsequence of sequence $\{x_k(\cdot)\}$, $k = 1, 2, \dots$

The fact that an admissible trajectory $x_{k_i}(\cdot)$, $i = 1, 2, \dots$, corresponds to an admissible control $u_{k_i}(\cdot)$ means that

$$x_{k_i}(t) = x_0 + \int_0^t f_0(x_{k_i}(s), s)\, ds + \sum_{j=1}^{m} \int_0^t f_j(x_{k_i}(s), s) u^j_{k_i}(s)\, ds \quad \text{for any } t \in [0, T]. \tag{2.46}$$

Then due to the weak convergence in $L^1([0, T], R^n)$ of the sequence $\{u_{k_i}(\cdot)\}$, $i = 1, 2, \dots$, to control $\bar{u}(\cdot)$ taking a limit as $i \to \infty$ in the equality (2.46) we get

$$\bar{x}(t) = x_0 + \int_0^t f_0(\bar{x}(s), s)\, ds + \sum_{j=1}^{m} \int_0^t f_j(\bar{x}(s), s) \bar{u}^j(s)\, ds \quad \text{for any} \quad t \in [0, T].$$

This means that admissible trajectory $\bar{x}(\cdot)$ corresponds to control $\bar{u}(\cdot)$. Since $\{x_{k_i}(\cdot)\}$, $i = 1, 2, \dots$, is an arbitrary convergent subsequence of sequence $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, and due to Theorem 4 the sequence $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, has a unique limit point $\bar{x}(\cdot)$ in $C([0, T], R^n)$. Hence, the sequence $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, converges uniformly to trajectory $\bar{x}(\cdot)$ on $[0, T]$. ■

Now we establish an upper semicontinuity property of an integral functional. This property (together with Lemmas 5 and 6) implies existence of an optimal control process in problems for systems affine in control (see Theorem 10 in Section 3).

Consider the following integral functional:

$$J(x(\cdot), \xi(\cdot)) = \int_0^T f^0(x(t), t, \xi(t))\, dt \tag{2.47}$$

where $T > 0$ and the scalar function $f^0(\cdot, \cdot, \cdot)$ is continuous jointly in variables $x \in G$, $t \in [0, T]$ and $\xi \in U$, and concave in variable ξ. Here we assume that G is a nonempty open subset of R^n and U is a nonempty convex compactum in R^m. The functional (2.47) is considered on pairs $(x(\cdot), \xi(\cdot))$ where vector function $x : [0,T] \mapsto G$ is continuous and vector function $\xi : [0,T] \mapsto U$ is measurable.

The following result establish the upper semicontinuity property of the integral functional (2.47).

Theorem 5. *Suppose a sequence* $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, *of continuous vector functions* $x_k : [0,T] \mapsto G$ *converges uniformly on time interval* $[0,T]$, $T > 0$, *to a continuous vector function* $\bar{x} : [0,T] \mapsto G$ *and a sequence* $\{\xi_k(\cdot)\}$, $k = 1, 2, \dots$, *of measurable vector functions* $\xi_k : [0,T] \mapsto U$, $T > 0$, *converges weakly in* $L^1([0,T], R^m)$ *to a measurable vector function* $\bar{\xi} : [0,T] \mapsto R^m$ *as* $k \to \infty$. *Then*

$$\overline{\lim_{k\to\infty}} \int_0^T f^0(x_k(t), t, \xi_k(t))\, dt \le \int_0^T f^0(\bar{x}(t), t, \bar{\xi}(t))\, dt, \qquad (2.48)$$

i.e., the integral functional (2.47) *is upper semicontinuous.*

Proof. Without loss of generality one can assume that

$$\lim_{k\to\infty} \int_0^T f^0(x_k(t), t, \xi_k(t))\, dt = \overline{\lim_{k\to\infty}} \int_0^T f^0(x_k(t), t, \xi_k(t))\, dt \qquad (2.49)$$

Otherwise one should pass to a subsequence.

Further, the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \dots$, and vector function $\bar{\xi}(\cdot)$ satisfy assumptions of the Banach–Saks theorem (see Theorem 1). Hence, due to Theorem 1 there is a subsequence $\{\xi_{k_i}(\cdot)\}$, $i = 1, 2, \dots$, of the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \dots$, such that the sequence $\{z_N(\cdot)\}$, $N = 1, 2, \dots$, of its arithmetic means

$$z_N(t) = \frac{1}{N} \sum_{i=1}^{N} \xi_{k_i}(t), \quad t \in [0,T], \qquad (2.50)$$

converges strongly in $L^2([0,T], R^m)$ to the vector function $\xi(\cdot)$ as $N \to \infty$. Hence, there is a subsequence $\{z_{N_j}(\cdot)\}$, $j = 1, 2, \dots$, of the sequence of measurable functions $\{z_N(\cdot)\}$, $N = 1, 2, \dots$, such that a.e. on $[0,T]$

$$z_{N_j}(t) \to \xi(t) \quad \text{as} \quad j \to \infty.$$

Since the vector function $\bar{x}(t)$ is continuous on $[0,T]$, the scalar function $f^0(\cdot,\cdot,\cdot)$ is continuous on $G \times [0,T] \times U$ and the set U is compact, the sequence of measurable scalar functions $\{\phi_j(\cdot)\}$, $j = 1, 2, \dots$, defined as $\phi_j(t) = f^0(\bar{x}(t), t, z_{N_j}(t))$ is uniformly bounded on $[0,T]$. Hence by the Lebesgue theorem (see [**34**]) we have

$$\lim_{j\to\infty} \int_0^T f^0(\bar{x}(t), t, z_{N_j}(t))\, dt = \int_0^T f^0(\bar{x}(t), t, \xi(t))\, dt. \qquad (2.51)$$

Then due to (2.50) and the concavity of function $\{f^0(\bar{x}(t), t, \cdot)\}$ in variable u at any instant $t \in [0, T]$ we get

$$\frac{1}{N_j} \sum_{i=1}^{N_j} \int_0^T f^0(\bar{x}(t), t, \xi_{k_i}(t))\, dt \leq \int_0^T f^0\Big(\bar{x}(t), t, \frac{1}{N_j} \sum_{i=1}^{N_j} \xi_{k_i}(t)\Big)\, dt$$

$$= \int_0^T f^0(\bar{x}(t), t, z_{N_j}(t))\, dt \quad \text{for any} \quad j = 1, 2, \ldots$$

As far as $\{\xi_{k_i}(\cdot)\}$, $i = 1, 2, \ldots$, is a subsequence of the sequence $\{\xi_k(\cdot)\}$, $k = 1, 2, \ldots$, and due to conditions (2.49) and (2.51) the last inequality implies

$$\lim_{i \to \infty} \int_0^T f^0(\bar{x}(t), t, \xi_{k_i}(t))\, dt \leq \int_0^T f^0(\bar{x}(t), t, \xi(t))\, dt. \tag{2.52}$$

Since the sequence $\{x_k(\cdot)\}$, $k = 1, 2, \ldots$, of continuous vector functions $x_k(\cdot)$ converges uniformly on time interval $[0, T]$ to a continuous vector function $\bar{x}(\cdot)$ inequality (2.52) implies that

$$\lim_{k \to \infty} \int_0^T f^0(x_k(t), t, \xi_k(t))\, dt \leq \int_0^T f^0(\bar{x}(t), t, \xi(t))\, dt.$$

Due to (2.49) this means that (2.48) is true. So, the theorem is proved. ■

3. Attainability domains

Here we study some properties of attainability domains of the control system (2.1) on a finite time interval $[0,T]$, $T > 0$. In that follows we assume that conditions (A1), (A2) and (A3) are satisfied.

By definition an attainability domain of the control system (2.1) from initial state $x_0 \in G$ at instant $t \in [0,T]$ is the set $X(x_0,t)$ of all points $\xi \in R^n$ which are reachable from the state x_0 at instant $t \in [0,T]$ by admissible trajectories $x(\cdot)$ of the control system (2.1) (see Figure 3). In other words, attainability domain $X(x_0,t)$ is the union of values $x(t)$ of all admissible trajectories $x(\cdot)$ of the control system (2.1) at instant $t \in [0,T]$. Briefly the definition of attainability domain $X(x_0,t)$, $t \in [0,T]$, can be written as follows:

$$X(x_0,t) = \{\xi = x(t) : \ \dot{x} = f(x,t,u), \ u \in U, \ x(0) = x_0\}.$$

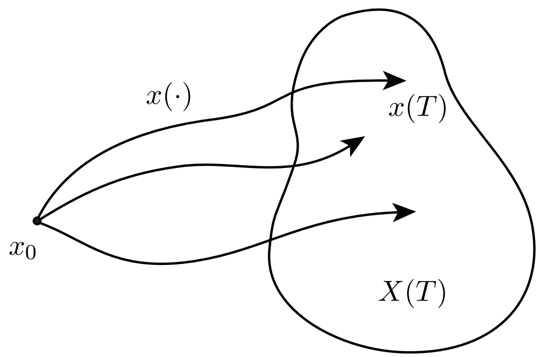

Figure 3: *Attainability domain.*

Below we consider the situation when initial state x_0 (see (2.2)) of the control system (2.1) is fixed and all attainability domains are considered from this initial state only. In this case for brevity we omit symbol x_0 and write $X(t)$ instead of $X(x_0,t)$.

By assumption (A1) all admissible trajectories $x(\cdot)$ of the control system (2.1) are defined on the whole time interval $[0,T]$ and take their values $x(t)$ in the set $\Pi \subset G$. Hence attainability domain $X(t)$ of the control system (2.1) is a nonempty subset of Π:

$$X(t) \subset \Pi \subset G \quad \text{for any} \quad t \in [0,T].$$

Note, that attainability domains $X(t)$, $t \in [0,T]$, can be considered as values of the multivalued mapping $X(\cdot)$ of time $t \in [0,T]$ to the set 2^{R^n} of all non empty subsets of R^n:

$$X : [0,T] \mapsto 2^{R^n}.$$

The notion of attainability domain plays an important role in the control theory. This is due to the fact that attainability domains $X(t)$, $t \in [0,T]$, provide a suitable geometric characterization of dynamics of the control system (2.1).

The general properties of control system (2.1) established in the previous section imply the corresponding properties of attainability domains. In particular by virtue of Lemma 1 conditions (A1) and (A2) imply a uniform boundedness of attainability domains $X(t)$, $t \in [0,T]$ (see inequality (2.10)).

As an immediate consequence of Theorem 4 we get the following result.

Theorem 6. *Suppose assumptions* (A1), (A2) *and* (A3) *hold. Then attainability domains* $X(t)$, $t \in [0,T]$, *of control system* (2.1) *are nonempty compact sets in* R^n.

Proof. Indeed, due to Theorem 4 the set of admissible trajectories $x(\cdot)$ of control system (2.1) is a nonempty compactum in the space $C([0,T], R^n)$. This means that trajectories $x(\cdot)$ of control system (2.1) are uniformly bounded on $[0,T]$ and if $\{x_k(\cdot)\}$, $k = 1, 2, \dots$, is a sequence of admissible trajectories $x_k(\cdot)$ on $[0,T]$ then it contains a subsequence $\{x_{k_i}(\cdot)\}$, $i = 1, 2, \dots$ which converges to an admissible trajectory $\bar{x}(\cdot)$ of control system (2.1) uniformly on the time interval $[0,T]$. This fact immediately implies that for any $t \in [0,T]$ the attainability domain $X(t)$ is bounded and closed subset in R^n. The theorem is proved. ■

Denote by $\Omega(R^n)$ the space of all nonempty compact subsets of R^n with the Hausdorff metric

$$h(A_1, A_2) = \inf\{r > 0 : \ A_1 \subset A_2 + B_r(0), \ A_2 \subset A_1 + B_r(0)\}, \tag{3.1}$$

algebraic sum

$$A_1 + A_2 = \{a_1 + a_2 : \ a_1 \in A_1, \ a_2 \in A_2\}$$

and multiplication on real numbers

$$\lambda A = \{\lambda a : a \in A\}.$$

Here A_1, A_2 and A are nonempty compact subsets in R^n; $\lambda \in R^1$ is a real number and $B_r(0)$ is a closed ball of radius $r \geq 0$ in R^n with the centrum in the origin.

Theorem 7. *Suppose assumptions* (A1), (A2) *and* (A3) *hold. Then attainability domain* $X : [0,T] \mapsto \Omega(R^n)$ *of the control system* (2.1) *is a Lipschitz continuous multivalued mapping.*

Proof. Due to the definition of attainability domain for arbitrary $\tau \in [0,T]$ and any $\xi \in X(\tau)$ there is an admissible trajectory $x_\xi(\cdot)$ of control system (2.1) such that $x_\xi(\tau) = \xi$.

By Lemma 1 all such admissible trajectories $x_\xi(\cdot)$ of control system (2.1) satisfy the Lipschitz condition

$$\|x_\xi(\tau_2) - x_\xi(\tau_1)\| \leq M_1\|\tau_2 - \tau_1\| \quad \text{for any} \quad \tau_1, \tau_2 \in [0,T]$$

(see (2.11)) with the same constant $M_1 \geq 0$. Hence for any $\xi_1 \in X(\tau_1)$, $\xi_2 \in X(\tau_2)$ and arbitrary $\tau_1, \tau_2 \in [0,T]$ we have

$$\|x_{\xi_1}(\tau_1) - x_{\xi_1}(\tau_2)\| \leq M_1|\tau_1 - \tau_2| \tag{3.2}$$

and

$$\|x_{\xi_2}(\tau_1) - x_{\xi_2}(\tau_2)\| \leq M_1|\tau_1 - \tau_2| \tag{3.3}$$

Since $x_{\xi_1}(\tau_1)$ is an arbitrary point from $X(\tau_1)$ and $x_{\xi_1}(\tau_2) \in X(\tau_2)$ inequality (3.2) implies inclusion

$$X(\tau_1) \subset X(\tau_2) + B_{M_1|\tau_1-\tau_2|}(0). \tag{3.4}$$

Analogously, since $x_{\xi_2}(\tau_2)$ is an arbitrary point from $X(\tau_2)$ and $x_{\xi_2}(\tau_1) \in X(\tau_1)$ inequality (3.3) implies inclusion

$$X(\tau_2) \subset X(\tau_1) + B_{M_1|\tau_1-\tau_2|}(0). \tag{3.5}$$

Together inclusions (3.4) and (3.5) imply inequality

$$h(X(\tau_1), X(\tau_2)) \leq M_1|\tau_2 - \tau_1|$$

(see (3.1)). Thus multivalued mapping $X(\cdot)$ is Lipschitz continuous. The theorem is proved. ■

It should be noted that in the case of general nonlinear control system (2.1) conditions (A1)–(A3) do not guarantee convexity of attainability domains $X(t)$, $t \in [0,T]$. Usually, the convexity property is attributed to attainability domains only in the case when the control system (2.1) is linear, i.e., it has the form

$$\dot{x} = A(t)x + u, \quad u \in U,$$

where $A(\cdot)$ is a continuous $n \times n$ matrix function on $[0,T]$ and the set U is a compactum in R^n. Some of the constructions of the optimal control theory can be considerably simplified in this case (see, for example, [**25**]). In present notes our consideration is focused mainly on the case of nonlinear control system (2.1). In this case, generally speaking, under conditions (A1)–(A3) attainability domains $X(t)$, $t \in [0,T]$, are just nonempty compact subsets of R^n.

We pass now to results on existence of an optimal admissible control.

First consider the following time-optimal problem:

$$\dot{x} = f(x,t,u), \quad u \in U, \tag{3.6}$$

$$x(0) = x_0, \quad x(T) = x_1, \tag{3.7}$$

$$T \to \min. \tag{3.8}$$

Here $x_0 \in G$, $x_1 \in G$, $x_0 \neq x_1$ and $T > 0$ is a free final time. As usual we assume that the vector function $f(\cdot,\cdot,\cdot)$ is continuous jointly in variables x, t and u, and continuously differentiable in variable x. The class of admissible controls $u(\cdot)$ of system (3.6) consists of all measurable vector functions $u : [0,T] \mapsto U$.

Theorems 6 and 7 imply the following existence result for the time-optimal problem (3.6)–(3.8).

Theorem 8. *Suppose conditions* (A1), (A2) *hold for any* $T > 0$ *and condition* (A3) *takes place. Assume also that there is at least one admissible trajectory* $\bar{x}(\cdot)$ *of control system* (3.6) *such that* $\bar{x}(0) = x_0$ *and* $\bar{x}(\bar{T}) = x_1$ *for some* $\bar{T} > 0$. *Then there is an optimal admissible pair* $(x_*(\cdot), u_*(\cdot))$ *in the time-optimal problem* (3.6)–(3.8).

Proof. Consider the set

$$\mathcal{T} = \{t \geq 0 : x_1 \in X(t)\}.$$

Since $x_1 = \bar{x}(\bar{T}) \in X(\bar{T})$ we have $\bar{T} \in \mathcal{T}$. So, the set $\mathcal{T}$ is nonempty. Denote by T_* the infimum of the set $\mathcal{T}$, i.e.,

$$T_* = \inf\{T \in \mathcal{T}\}.$$

Due to Lemma 1 admissible trajectories $x(\cdot)$ of the control system (3.6) are uniformly Lipschitz continuous on the time interval $[0,\bar{T}]$ (see (2.11)). Hence, by assumption $x_0 \neq x_1$ we have $T_* > 0$.

Let $\{T_k\}$, $k = 1,2,\ldots$, be a sequence of points $T_k \in \mathcal{T}$ such that $T_k \to T_*$ as $k \to \infty$. By the definition of an infimum of the set $\mathcal{T}$ such sequence $\{T_k\}$, $k = 1,2,\ldots$, exists. In this case $x_1 \in X(T_k)$ for any $k = 1,2,\ldots$ By Theorem 7 for arbitrary $\varepsilon > 0$ there is a natural k_1 such that for all $k \geq k_1$ we have

$$x_1 \in X(T_k) \subset X(T_*) + B_\varepsilon(0).$$

By Theorem 6 the set $X(T_*)$ is compact and due to arbitrariness of $\varepsilon > 0$ the last inclusion implies $x_1 \in X(T_*)$. Hence $T_* \in \mathcal{T}$ and there is an admissible control process $(x_*(\cdot), u_*(\cdot))$ such that $x_*(T_*) = x_1$. As far as T_* is the infimum of the set $\mathcal{T}$ it is impossible to reach state x_1 by admissible trajectory $x(\cdot)$ earlier than in time T_*. So, $(x_*(\cdot), u_*(\cdot))$ is an optimal admissible pair in time-optimal problem (3.6)–(3.8). The theorem is proved. ■

Now consider the following optimal control problem (P_T) with an integral goal functional and a free right terminal state $x(T)$ on a fixed time interval $[0,T]$, $T > 0$:

$$\dot{x} = f(x,t,u), \quad u \in U, \tag{3.9}$$

$$x(0) = x_0, \tag{3.10}$$

$$J_T(x,u) = \int_0^T f^0(x,t,u)\,dt \to \max\,. \tag{3.11}$$

Here as usual $x_0 \in G$; the vector function $f : G \times [0,T] \times U \mapsto R^n$ and the scalar function $f^0 : G \times [0,T] \times U \mapsto R^1$ are continuous. The class of admissible controls $u(\cdot)$ of system (3.9) consists of all measurable vector functions $u : [0,T] \mapsto U$.

Let us introduce an extra state coordinate $x^0 \in R^1$ and consider the following extended control system on time interval $[0,T]$:

$$\dot{\tilde{x}} = \tilde{f}(x,t,u), \quad u \in U \tag{3.12}$$

with initial condition

$$\tilde{x}(0) = (0, x_0).$$

Here $\tilde{x} = (x^0, x) \in R^1 \times R^n$ is an extended state variable, and

$$\tilde{f}(x,t,u) = (f^0(x,t,u), f(x,u)) \quad \text{for all} \quad x \in G,\ t \in [0,T] \quad \text{and} \quad u \in U.$$

In this case all admissible trajectories $\tilde{x}(\cdot)$ of the control system (3.12) are defined on $[0,T]$ and take their values $\tilde{x}(t)$ in the closed subset $\tilde{\Pi} = [0,T] \times \Pi$ of the open set $\tilde{G} = R^1 \times G$ in $R^1 \times R^n$.

Further, initial control system (3.9) satisfies conditions (A1) and (A2). Hence admissible trajectories of the control system (3.9) are uniformly bounded on the time interval $[0,T]$. Since the set U is compact, there exists a constant $\kappa_1 \geq 0$ such that, for any trajectory $x(\cdot)$ of the control system (3.9), any $u \in U$, and any $t \in [0,T]$, the following inequality holds:

$$|f^0(x(t),t,u)| \leq \kappa_1. \tag{3.13}$$

Hence there exists a $\tilde{C}_0 \geq 0$ such that

$$\langle \tilde{x}, \tilde{f}(x,t,u) \rangle \leq \tilde{C}_0(1 + \|\tilde{x}\|^2) \quad \text{for any} \quad \tilde{x} \in \tilde{G},\ t \in [0,T], u \in U.$$

Thus both conditions (A1) and (A2) are satisfied for extended control system (3.12). Due to Lemma 1 this imply that admissible trajectories $\tilde{x}(\cdot)$ of control system (3.12) are uniformly bounded on $[0,T]$. By Theorem 6 conditions (A1), (A2) and (A3) imply that attainability domains $\tilde{X}(t)$, $t \in [0,T]$, of the control system (3.12) are nonempty compact sets in $R^1 \times R^n$. Due to Theorem 7 the multivalued mapping $\tilde{X}(\cdot)$ is Lipschitz continuous on $[0,T]$.

It is easy to see that the control system (3.9) constitute a part of the control system (3.12) and there is a one-to-one correspondence between admissible trajectories $x(\cdot)$ of the control system (3.9) and admissible trajectories $\tilde{x}(\cdot)$ of the control system (3.12). Namely, if $x(\cdot)$ is an admissible trajectory of the control system (3.9) on $[0,T]$ then the absolutely continuous vector function

$\tilde{x}(\cdot) = (x^0(\cdot), x(\cdot))$ where

$$x^0(t) = \int_0^t f^0(x(s), s, u(s))\, ds, \quad t \in [0, T] \tag{3.14}$$

is an admissible trajectory of the control system (3.12) and, vice versa, if $\tilde{x}(\cdot) = (x^0(\cdot), x(\cdot))$ an admissible trajectory of the control system (3.12) then $x(\cdot)$ is an admissible trajectory of the control system (3.9).

The optimal control problem (P_T) can be reformulated in terms of the extended state variable $\tilde{x} \in R^1 \times R^n$ as follows:

$$\dot{\tilde{x}} = \tilde{f}(x, t, u), \quad u \in U, \tag{3.15}$$

$$\tilde{x}(0) = (0, x_0), \tag{3.16}$$

$$J_T(\tilde{x}(T)) = x^0(T) \to \max. \tag{3.17}$$

In terms of attainability domains the optimality in problem (P_T) of an admissible control process $(\tilde{x}_*(\cdot), u_*(\cdot))$ of the extended control system (3.15) means that coordinate $x_*^0(T)$ of the trajectory $\tilde{x}_*(\cdot)$ is maximal among all coordinates ξ^0 of points $\tilde{\xi} \in \tilde{X}(T)$ (see Figure 4). This circumstance give a possibility to treat the optimal control problem (P_T) in geometric terms of attainability domains.

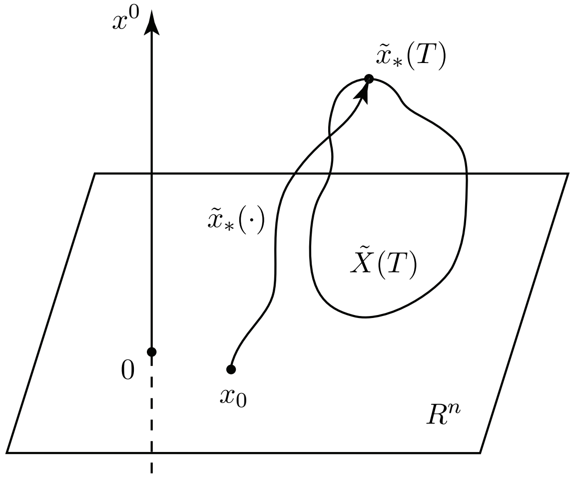

Figure 4: *Attainability domain of extended system.*

Let us introduce the following condition.

(A4) *For any $x \in G$ and $t \in [0, T]$ the set*

$$\tilde{Q}(x, t) = \big\{(z^0, z) \in R^1 \times R^n : z^0 \le f^0(t, x, u),\ z = f(x, t, u),\ u \in U\big\}$$

is convex.

It is easy to see that (A4) implies validity of (A3).

Theorem 9. *Suppose conditions* (A1), (A2) *and* (A4) *are valid. Then there is an optimal admissible pair* $(x_*(\cdot), u_*(\cdot))$ *in problem* (P_T).

Proof. Consider the following control system on time interval $[0,T]$:

$$\dot{x}^0 = (1-u^0)f^0(x,t,u) - u^0\kappa_1, \quad \tilde{u} = (u^0,u) \in [0,1]\times U, \tag{3.18}$$

$$\dot{x} = f(x,t,u) \tag{3.19}$$

with initial condition

$$\tilde{x}(0) = (0, x_0).$$

Here $\tilde{x} = (x^0, x) \in \tilde{G}$ and $\kappa_1 \geq 0$ is a constant satisfying (3.13). Note that the right-hand side of the control system (3.18), (3.19) does not depend on coordinate x^0. It is easy to see that the control system (3.18), (3.19) satisfies both conditions (A1) and (A2).

For any $t \in [0,T]$, $x \in G$ a vectogram of control system (3.18), (3.19) at the point $(x,t) \in G \times [0,T]$ is the set

$$W(x,t) = \big\{(z^0, z) \in R^1 \times R^n : -\kappa_1 \leq z^0 \leq f^0(x,t,u),\ z = f(x,t,u),\ u \in U\big\}.$$

Since U is a compactum and in view of conditions (A1), (A2) and (A4), the set $W(x,t)$ is convex and compact for any $x \in G$ and $t \in [0,T]$. Hence due to Theorem 6 the attainability domain $\tilde{Z}(T)$ of control system (3.18), (3.19) at time T is a compactum in $R^1 \times R^n$. Hence there is a point $\tilde{\xi}_* \in \tilde{Z}(T)$ such that its coordinate $\xi^0_*(T)$ is maximal among all coordinates ξ^0 of points $\tilde{\xi} \in \tilde{Z}(T)$. Let $\tilde{\xi}_*(\cdot) = (\xi^0(\cdot), x_*(\cdot))$ be an admissible trajectory of control system (3.18), (3.19) on $[0,T]$ such that $\tilde{\xi}_*(T) = \tilde{\xi}_*$, i.e., its coordinate $\xi^0_*(T)$ is maximal in $\tilde{Z}(T)$.

By construction any admissible trajectory $\tilde{x}(\cdot)$, of control system (3.12) is an admissible trajectory of control system (3.18), (3.19). Indeed if $\tilde{x}(\cdot)$ is an admissible trajectory of control system (3.12) corresponding to admissible control $u(\cdot)$ then $\tilde{x}(\cdot)$ is an admissible trajectory of control system (3.18), (3.19) which corresponds to admissible control $\tilde{u}(\cdot) = (0, u(\cdot))$. From the other side if an absolutely continuous vector function $\tilde{\xi}(\cdot) = (\xi^0(\cdot), \xi(\cdot))$ is an admissible trajectory of control system (3.18), (3.19) on $[0,T]$ then $\tilde{x}(\cdot) = (x^0(\cdot), \xi(\cdot))$ is an admissible trajectory of control system (3.9) where $x^0(\cdot)$ is defined according with (3.14), i.e.,

$$x^0(t) = \int_0^t f^0(\xi(s), s, u(s))\, ds, \quad t \in [0,T].$$

and, moreover, $\xi^0(t) \leq x^0(t)$ for all $t \in [0,T]$. This imply that if admissible trajectory $\tilde{\xi}_*(\cdot) = (\xi^0_*(\cdot), \xi_*(\cdot))$ of control system (3.18), (3.19) on $[0,T]$ has a maximal coordinate $\xi^0_*(T)$ in the attainability domain $\tilde{Z}(T)$ then for corresponding admissible trajectory $\tilde{x}_*(\cdot) = (x^0_*(\cdot), \xi_*(\cdot))$ of control system (3.9) we get that its coordinate $x^0_*(T)$ is maximal possible in $\tilde{X}(T)$. Hence $x_*(\cdot)$

is an optimal admissible trajectory in problem (P_T) and the corresponding admissible control $u_*(\cdot)$ is also an optimal one in (P_T). So, $(x_*(\cdot), u_*(\cdot))$ is an optimal control process in problem (P_T). The theorem is proved. ■

Lastly, using the upper semicontinuity property of an integral functional (Theorem 5) we give an alternative proof of existence of an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in the problem (P_T) in the case when the control system (2.1) is affine in control, i.e. when it has form (2.44). Note that condition (A4) holds automatically in this case.

Theorem 10. *Suppose conditions* (A1) *and* (A2) *are valid and the control system* (3.9) *is affine in control, i.e. it has a form* (2.44). *Assume also that the set U of geometric constraints is convex and compact, and the integrand $f^0(x, t, \cdot)$ is concave in variable $u \in U$ in the utility functional* (3.11). *Then there is an optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in the optimal control problem (P_T).*

Proof. Let $\{(x_k(\cdot), u_k(\cdot))\}$, $k = 1, 2, \ldots$, be a maximizing sequence of admissible pairs in problem (P_T). Then

$$\lim_{k\to\infty} \int_0^T f^0(x_k(t), t, u_k(t))\, dt = \sup_{(x(\cdot),u(\cdot))} J_T(x(\cdot), u(\cdot)), \tag{3.20}$$

where the integral functional $J_T(x(\cdot), u(\cdot))$ is defined by (3.11) and a supremum in the right-hand side of (3.20) is taken in all admissible pairs $(x(\cdot), u(\cdot))$. Then due to Lemmas 5 and 6 passing if necessary to a subsequence without loss of generality one can assume that the sequence $\{u_k(\cdot)\}$, $k = 1, 2, \ldots$, converges weakly in $L^1([0, T, R^m])$ to an admissible control $u_*(\cdot) : [0, T] \mapsto U$ and the corresponding sequence of admissible trajectories $\{x_k(\cdot)\}$, $k = 1, 2, \ldots$, converges uniformly on $[0, T]$ to the admissible trajectory $x_*(\cdot)$ corresponding to $u_*(\cdot)$. Then by Theorem 5 we have

$$\overline{\lim_{k\to\infty}} \int_0^T f^0(x_k(t), t, u_k(t))\, dt \leq \int_0^T f^0(\bar{x}_*(t), t, u_*(t))dt.$$

Due to (3.20) this inequality implies that $(x(\cdot), u(\cdot))$ is an optimal admissible pair in problem (P_T). The theorem is proved. ■

4. The Pontryagin maximum principle

The formulation of optimal control problem (P_T) (see (3.9)–(3.11)) in terms of the extended state variable $\tilde{x} \in R^1 \times R^n$ (see (3.15)–(3.17)) reduces problem (P_T) to the problem of determining a point $\tilde{\xi}_* = \tilde{x}_*(T)$ in attainability domain $\tilde{X}(T)$ such that its coordinate ξ^0_* is maximal.

If optimal control problem (P_T) satisfies conditions (A1), (A2) and (A4) then due to Theorem 9 such "maximal" point $\tilde{\xi}_* \in \tilde{X}(T)$ exits. Geometrically this means (see Figure 5) that all points $\tilde{x} \in \tilde{X}(T)$ lie below the hyperplane

$$\Gamma = \{\tilde{x} \in R^1 \times R^n : \langle \tilde{x} - \tilde{\xi}_*, \tilde{\psi}_T \rangle = 0\}$$

where

$$\tilde{\psi}_T = (\psi^0_T, \psi_T) \in R^1 \times R^n, \quad \psi^0_T = 1, \quad \psi_T = 0. \tag{4.1}$$

Thus, optimality of trajectory $\tilde{x}_*(\cdot)$ in problem (P_T) is equivalent to validity of the following condition:

$$\langle \tilde{x} - \tilde{x}_*(T), \tilde{\psi}_T \rangle \leq 0 \quad \text{for all} \quad \tilde{x} \in \tilde{X}(T). \tag{4.2}$$

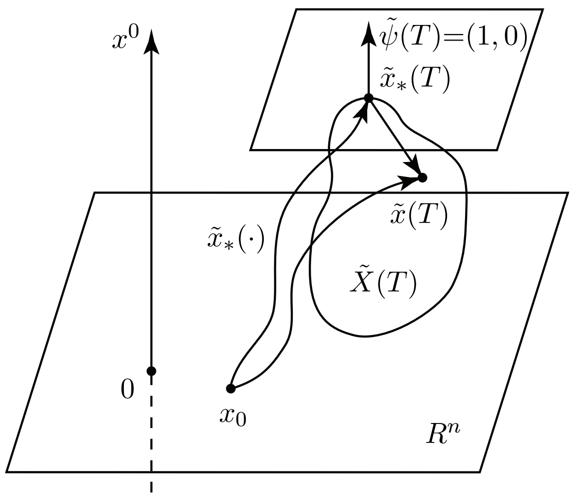

Figure 5: *Optimal trajectory of extended system.*

Let us introduce the Hamilton–Pontryagin function $\mathcal{H} : G \times [0,T] \times U \times R^1 \times R^n \mapsto R^1$ and the maximum function (or the Hamiltonian) $H : G \times [0,T] \times R^1 \times R^n \mapsto R^1$ of the problem (P_T) as follows:

$$\mathcal{H}(x,t,u,\psi^0,\psi) = \langle f(x,t,u),\psi\rangle + \psi^0 f^0(x,t,u), \tag{4.3}$$

$$H(x.t,\psi^0,\psi) = \sup_{u\in U} \mathcal{H}(x.t,u,\psi^0,\psi) \tag{4.4}$$

for any $x \in G$, $t \in [0,T]$, $u \in U$, $\psi^0 \in R^1$ and $\psi \in R^n$.

Note, that as far as the set U is compact the function $\mathcal{H}(x,t,\cdot,\psi^0,\psi)$ in the right-hand side of (4.4) always reaches its maximal value in variable $u \in U$ for any fixed $x \in G$, $t \in [0,T]$, $\psi^0 \in R^1$ and $\psi \in R^n$. Thus, the sign $\sup_{u\in U}$ can be replaced by $\max_{u\in U}$ on the right-hand side of (4.4).

The following result is the Pontryagin maximum principle for problem (P_T) (see [**32**]).

Theorem 11. *Let $(x_*(\cdot),u_*(\cdot))$ be an optimal admissible process in problem (P_T) (see* (3.9)–(3.11)*). Then there are a number $\psi^0 \in R^1$ and an absolutely continuous vector function $\psi : [0,T] \mapsto R^n$ which are called adjoint variables (or the Lagrange multipliers) such that the following conditions hold:*

1) *the vector function $\psi(\cdot)$ is a (Carathéodory) solution to the following adjoint system on $[0,T]$:*

$$\begin{aligned}\dot{\psi} &= -\frac{\partial\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi)}{\partial x}\\ &= -\left[\frac{\partial f(x_*(t),t,u_*(t))}{\partial x}\right]^*\psi - \psi^0\frac{\partial f^0(x_*(t),t,u_*(t))}{\partial x};\end{aligned} \tag{4.5}$$

2) *the following maximum condition takes place on $[0,T]$:*

$$\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t)) \stackrel{a.e.}{=} H(x_*(t),t,\psi^0,\psi(t)); \tag{4.6}$$

3) *the pair $(\psi^0,\psi(\cdot))$ is nontrivial, i.e.,*

$$\psi^0 + \|\psi(0)\| > 0; \tag{4.7}$$

4) *the following transversality conditions are valid:*

$$\psi^0 > 0, \quad \psi(T) = 0. \tag{4.8}$$

Conditions 1)–3) in the formulation of Theorem 11 are called the core relations of the Pontryagin maximum principle. These conditions are common in various formulations of the Pontryagin maximum principle for different optimal control problems. Meanwhile condition 4) is a specific one in the formulation of the Pontryagin maximum principle for problem (P_T). This condition is strongly related to the fact that (P_T) is a maximization problem with a free right endpoint.

Note also that in the context of problem (P_T) the transversality conditions (4.8) imply the nontriviality condition (4.7). Indeed, the first inequality in transversality conditions (4.8) already means that the pair of adjoint variables $(\psi^0,\psi(\cdot))$ is nontrivial, i.e, that (4.7) holds. Moreover, the first inequality in

(4.8) means that optimal control problem (P_T) is normal, i.e., the Lagrange multiplier ψ^0 (which corresponds to the maximized functional (3.11)) does not vanish.

It is easy to see that conditions 1)–4) in Theorem 11 are invariant under multiplication of adjoint variables ψ^0, $\psi(\cdot)$ on positive numbers. If conditions 1)–4) are valid with adjoint variables ψ^0, $\psi(\cdot)$ then for any $\lambda > 0$ they are satisfied also with adjoint variables

$$\psi^0_\lambda = \lambda\psi^0 \quad \text{and} \quad \psi_\lambda(t) = \lambda\psi(t), \quad t \in [0,T].$$

Thus, multiplying both variables ψ^0 and $\psi(\cdot)$ in $1) - 4)$ by $\lambda = 1/\psi^0$ one can always assume that transversality conditions (4.8) have the form

$$\psi^0 = 1, \qquad \psi(T) = 0.$$

In the last case we say that relations of the maximum principle for problem (P_T) are in the normal form.

To prove Theorem 11 we will use the classical method of needle variations (see [**32**]). Let us describe this method.

Consider an admissible control $u(\cdot)$ of system (3.15). This means that $u(\cdot)$ is a bounded measurable function on $[0,T]$, $T > 0$, and $u(t) \in U$ for all $t \in [0,T]$. If so, then almost all points $\tau \in (0,T)$ are the points of approximate continuity of control $u(\cdot)$ (see [**31**]). For each such point τ there is a measurable set $\mathfrak{M} \subset [0,T]$ such that $\tau \in \mathfrak{M}$, $u(\cdot)$ is continuous at point τ along $\mathfrak{M}$, and

$$\lim_{h_1\to 0+, h_2\to 0+} \frac{\operatorname{meas}\{\mathfrak{M}\bigcap[\tau - h_1, \tau + h_2]\}}{h_1 + h_2} = 1.$$

Here the limit is taken with $h_1 \to 0+$ and $h_2 \to 0+$ in arbitrary ways.

It is clear that any continuity point of an admissible control $u(\cdot)$ is a fortiori a point of its approximate continuity.

If $\phi : [0,T] \times U \mapsto R^n$ is a continuous vector function, $u(\cdot)$ is an admissible control, $\tau \in (0,T)$ is a point of approximate continuity of $u(\cdot)$, $\varepsilon > 0$ is small and $\sigma \geq 0$ then the following condition holds:

$$\int_{\tau-\varepsilon\sigma}^{\tau} \phi(t, u(t))\, dt = \varepsilon\sigma\phi(\tau, u(\tau)) + \mathrm{o}(\varepsilon) \tag{4.9}$$

where

$$\frac{\|\mathrm{o}(\varepsilon)\|}{\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0+\,.$$

Let $\tau \in (0,T)$ be a point of approximate continuity of control $u(\cdot)$. Then take an arbitrary vector $v \in U$, a $\sigma \geq 0$ and a small $\varepsilon > 0$ such that $(\tau - \varepsilon\sigma, \tau] \subset [0,T]$. The vector function $u^v_{\varepsilon,\sigma,\tau}(\cdot)$ defined by equality

$$u^v_{\varepsilon,\sigma,\tau}(t) = \begin{cases} u(t), & t \notin (\tau - \varepsilon\sigma, \tau], \\ v, & t \in (\tau - \varepsilon\sigma, \tau]. \end{cases}$$

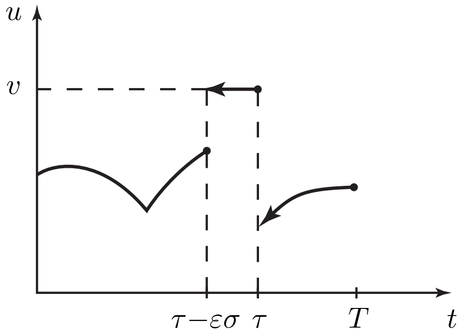

Figure 6: *Elementary needle variation.*

is called an elementary needle variation of the control $u(\cdot)$ (see Figure 6).

It is easy to see that as far as the control $u(\cdot)$ is admissible the vector function $u^v_{\varepsilon,\sigma,\tau}(\cdot)$ is an admissible control of system (3.15) as well; if control $u(\cdot)$ is a piecewise continuous function which is continuous from the left then $u^v_{\varepsilon,\sigma,\tau}(\varepsilon,\cdot)$ is also a piecewise continuous function which is continuous from the left.

The next result describes the main property of the elementary needle variation $u^v_{\varepsilon,\sigma,\tau}(\varepsilon,\cdot)$ in the context of control system (3.15).

Lemma 7. *Let $u^v_{\varepsilon,\sigma,\tau}(\cdot)$, $\tau \in (0,T)$, $v \in U$, $\sigma \geq 0$ and $\varepsilon > 0$ be an elementary needle variation of an admissible control $u(\cdot)$, and $\tilde{x}(\cdot) = (x^0(\cdot), x(\cdot))$ be the corresponding to $u(\cdot)$ admissible trajectory of control system* (3.15)*. Then there is a sufficiently small $\varepsilon_0 > 0$ such that for any $0 < \varepsilon \leq \varepsilon_0$ the corresponding to $u^v_{\varepsilon,\sigma,\tau}(\cdot)$ admissible trajectory $\tilde{x}^v_{\varepsilon,\sigma,\tau}(\cdot) = (x^{v,0}_{\varepsilon,\sigma,\tau}(\cdot), x^v_{\varepsilon,\sigma,\tau}(\cdot))$ of control system* (3.15) *is defined on the whole time interval* $[0,T]$*, the values $\tilde{x}^v_{\varepsilon,\sigma,\tau}(t)$, $t \in [0,T]$, belongs to the set $\tilde{G}$ and*

$$\tilde{x}^v_{\varepsilon,\sigma,\tau}(t) = \tilde{x}(t) + \varepsilon \tilde{y}^v_{\sigma,\tau}(t) + \tilde{o}(\varepsilon, t) \tag{4.10}$$

where $\tilde{y}^v_{\sigma,\tau} : [0,T] \mapsto R^1 \times R^n$ is (a unique) solution of the linear differential system

$$\dot{\tilde{y}} = \frac{\partial \tilde{f}(x(t), t, u(t))}{\partial \tilde{x}} \tilde{y} \tag{4.11}$$

satisfying the initial condition

$$\tilde{y}(\tau) = \sigma \left[\tilde{f}(x(\tau), \tau, v) - \tilde{f}(x(\tau), \tau, u(\tau)) \right] \tag{4.12}$$

and $\tilde{o}(\varepsilon,\cdot)$ is an absolutely continuous vector function on $[0,T]$ *such that*

$$\frac{\|\tilde{o}(\varepsilon, t)\|}{\varepsilon} \to 0 \quad \textit{as} \quad \varepsilon \to 0+$$

uniformly for $t \in [0,T]$.

Proof. First let us calculate the increment that the admissible trajectory $\tilde{x}(\cdot)$ will get at instant $\tau \in (0,T)$ from result of application of elementary needle variation $u^v_{\varepsilon,\sigma,\tau}(\cdot)$, i.e., the difference

$$\Delta\tilde{x}(\tau) = \tilde{x}^v_{\varepsilon,\sigma,\tau}(\tau) - \tilde{x}(\tau).$$

As far as admissible controls $u(\cdot)$ and $u^v_{\varepsilon,\sigma,\tau}(\cdot)$ coincide on the time interval $[0,\tau-\varepsilon\sigma]$ the corresponding admissible trajectories $\tilde{x}(\cdot)$ and $\tilde{x}^v_{\varepsilon,\sigma,\tau}(\cdot)$ are also coincide on $[0,\tau-\varepsilon\sigma]$. Hence

$$\tilde{x}^v_{\varepsilon,\sigma,\tau}(\tau) - \tilde{x}(\tau) = \int_{\tau-\varepsilon\sigma}^{\tau} \left[\tilde{f}(x^v_{\varepsilon,\sigma,\tau}(t), t, u^v_{\varepsilon,\sigma,\tau}(t)) - \tilde{f}(x(t), t, u(t)))\right] dt.$$

Since τ is a point of approximate continuity of $u(\cdot)$ and due to the fact that $u^v_{\varepsilon,\sigma,\tau}(t) \equiv v$ on $(\tau-\varepsilon\sigma,\tau]$ we get (see (4.9))

$$\tilde{x}^v_{\varepsilon,\sigma,\tau}(\tau) - \tilde{x}(\tau) = \int_{\tau-\varepsilon\sigma}^{\tau} \tilde{f}(\tilde{x}^v_{\varepsilon,\sigma,\tau}(t), t, v)\, dt - \varepsilon\sigma\tilde{f}(\tilde{x}(\tau), \tau, u(\tau)) + \tilde{o}(\varepsilon) \quad (4.13)$$

where

$$\frac{\|\tilde{o}(\varepsilon)\|}{\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0+\,.$$

Further, due to the theorem on continuous dependence of solution to the differential equation on parameters (see, for example, [**3**]) we have

$$\tilde{x}^v_{\varepsilon,\sigma,\tau}(t) = \tilde{x}(t) + \tilde{O}(\varepsilon, t) = \tilde{x}(\tau) + (\tilde{x}(t) - \tilde{x}(\tau)) + \tilde{O}(\varepsilon, t) \quad (4.14)$$

where $\tilde{O}(\varepsilon,\cdot)$ is an absolutely continuous function on $[0,T]$ such that

$$\tilde{O}(\varepsilon, t) \to 0 \quad \text{as} \quad \varepsilon \to 0+\,.$$

uniformly for $t \in [0,T]$.

Substituting (4.14) in (4.13) we get

$$\tilde{x}^v_{\varepsilon,\sigma,\tau}(\tau) = \tilde{x}(\tau) + \varepsilon\sigma\left[\tilde{f}(\tilde{x}(\tau), \tau, v) - \tilde{f}(\tilde{x}(\tau), \tau, u(\tau))\right] + \tilde{o}(\varepsilon) \quad (4.15)$$

where

$$\frac{\|\tilde{o}(\varepsilon)\|}{\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0+\,.$$

Let $\tilde{y}^v_{\sigma,\tau} : [0,T] \mapsto R^1 \times R^n$ be (a unique) solution of the linear differential system (4.11) satisfying the initial condition (4.12). Then (see (4.15)) due to the theorem on differentiability of solution to the differential equation on initial conditions (see, for example [**3**]) the equality (4.10) is valid on $[0,T]$. ■

Now we are ready to prove the Pontryagin maximum principle for problem (P_T).

Proof of Theorem 11. Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible process in problem (P_T) (see (3.9)–(3.11)) and $\tilde{x}_*(\cdot) = (x_*^0(\cdot), x_*(\cdot))$ be the corresponding to $u_*(\cdot)$ admissible trajectory of the extended control system (3.15). Consider an arbitrary elementary needle variation $u^v_{\varepsilon,\sigma,\tau}(\cdot)$, $\tau \in (0, T)$, $v \in U$, $\sigma \geq 0$, $\varepsilon > 0$ of the optimal control $u_*(\cdot)$ and let $\tilde{x}^v_{\varepsilon,\sigma,\tau}(\cdot)$ be the corresponding to $u^v_{\varepsilon,\sigma,\tau}(\cdot)$ admissible trajectory of the control system (3.15). Then due to Lemma 7 the following equality takes place:

$$\tilde{x}^v_{\varepsilon,\sigma,\tau}(t) = \tilde{x}_*(t) + \varepsilon \tilde{y}^v_{\sigma,\tau}(t) + \tilde{o}(\varepsilon, t), \quad t \in [0, T] \tag{4.16}$$

where $\tilde{y}^v_{\sigma,\tau} : [0, T] \mapsto R^1 \times R^n$ is (a unique) solution of the linear differential system

$$\dot{\tilde{y}} = \frac{\partial \tilde{f}(x_*(t), t, u_*(t))}{\partial \tilde{x}} \tilde{y} \tag{4.17}$$

satisfying the initial condition

$$\tilde{y}(\tau) = \sigma \left[\tilde{f}(x_*(\tau), \tau, v) - \tilde{f}(x_*(\tau), \tau, u_*(\tau)) \right] \tag{4.18}$$

and $\tilde{o}(\varepsilon, \cdot) : [0, T] \mapsto R^1 \times R^n$ is an absolutely continuous vector function such that

$$\frac{\|\tilde{o}(\varepsilon, t)\|}{\varepsilon} \to 0 \quad \text{as} \quad \varepsilon \to 0+$$

uniformly for $t \in [0, T]$.

Due to (4.2) the following inequality takes place:

$$\langle \tilde{x}^v_{\varepsilon,\sigma,\tau}(T) - \tilde{x}_*(T), \tilde{\psi}_T \rangle \leq 0 \tag{4.19}$$

where $\tilde{\psi}_T = (1, 0) \in R^1 \times R^n$ (see (4.1)). If so, then due to (4.16) we have

$$\langle \tilde{y}^v_{\sigma,\tau}(T), \tilde{\psi}_T \rangle + \langle \frac{\tilde{o}(\varepsilon, T)}{\varepsilon}, \tilde{\psi}_T \rangle \leq 0.$$

By virtue of (4.19) passing to a limit in the last inequality as $\varepsilon \to 0+$ we get

$$\langle \tilde{y}^v_{\sigma,\tau}(T), \tilde{\psi}_T \rangle \leq 0. \tag{4.20}$$

Let $\tilde{\psi} : [0, T] \mapsto R^1 \times R^n$, $\tilde{\psi}(\cdot) = (\psi^0(\cdot), \psi(\cdot))$, be (a unique) solution to the following linear differential system

$$\dot{\tilde{z}} = -\left[\frac{\partial \tilde{f}(x_*(t), t, u_*(t))}{\partial \tilde{x}} \right]^* \tilde{z} \tag{4.21}$$

satisfying the boundary condition

$$\tilde{z}(T) = \tilde{\psi}_T. \tag{4.22}$$

Note, that such defined absolutely continuous vector function $\tilde{\psi}(\cdot)$ does not dependent on parameters $\tau \in (0, T)$, $\sigma > 0$, $v \in U$ and $\varepsilon > 0$ of the elementary needle variation $u^v_{\varepsilon,\sigma,\tau}(\cdot)$.

As far the right-hand side of differential system (4.21) does not depend on the coordinate x^0 of the state variable $\tilde{x}$ the coordinate $\psi^0(\cdot)$ of the solution $\tilde{\psi}(\cdot)$ must be a constant, i.e., $\psi^0(t) \equiv \psi^0$ for all $t \in [0,T]$.

Due to the construction of function $\tilde{\psi}(\cdot) = (\psi^0, \psi(\cdot))$ the absolutely continuous function $\psi : [0,T] \mapsto R^n$ is a (Carathéodory) solution of the following differential system on $[0,T]$ (see (4.21)):

$$\dot{\psi} = -\left[\frac{\partial f(x_*(t),t,u_*(t))}{\partial x}\right]^* \psi + \psi^0 \frac{\partial f^0(x_*(t),t,u_*(t))}{\partial x},$$

i.e. $\psi(\cdot)$ is a solution to the adjoint system (4.5).

Further due to the boundary condition (4.22) we have (see (4.1))

$$\psi^0 = 1, \quad \psi(T) = 0.$$

Hence ψ^0 and absolutely continuous function $\psi(\cdot)$ satisfy the transversality conditions (4.8). If so, then the nontriviality condition (4.7) is valid as well.

By virtue of (4.17) and (4.21) the following equality holds a.e. on $[0,T]$:

$$\frac{d}{dt}\langle \tilde{y}^v_{\sigma,\tau}(t), \tilde{\psi}(t)\rangle = \left\langle \frac{\partial \tilde{f}(x_*(t),t,u_*(t))}{\partial \tilde{x}} \tilde{y}^v_{\sigma,\tau}(t), \tilde{\psi}(t) \right\rangle - \left\langle \tilde{y}^v_{\sigma,\tau}(t), \left[\frac{\partial f(x_*(t),t,u_*(t))}{\partial \tilde{x}}\right]^* \tilde{\psi}(t) \right\rangle = 0.$$

Hence, $\langle y^v_{\sigma,\tau}(t), \tilde{\psi}(t)\rangle \equiv \mathrm{const}$ and we have (see (4.18), (4.20))

$$\begin{aligned}\langle \tilde{y}^v_{\sigma,\tau}(T), \tilde{\psi}(T)\rangle &= \langle \tilde{y}^v_{\sigma,\tau}(\tau), \tilde{\psi}(\tau)\rangle \\ &= \sigma\langle \tilde{f}(x_*(\tau),\tau,v) - \tilde{f}(x_*(\tau),\tau,u_*(\tau)), \tilde{\psi}(\tau)\rangle \le 0.\end{aligned}$$

As far as $v \in U$, $\sigma \ge 0$ are arbitrary, the instant $\tau \in (0,T)$ is an arbitrary point of approximate continuity of the optimal control $u_*(\cdot)$ and the vector function $\tilde{\psi}(\cdot)$ does not depend on parameters τ, σ, v and ε of the elementary variation $u^v_{\varepsilon,\sigma,\tau}(\cdot)$ the last inequality implies that a.e. on $[0,T]$ we have

$$\sup_{v\in U}\langle \tilde{f}(x_*(\tau),\tau,v), \tilde{\psi}(\tau)\rangle \le \langle \tilde{f}(x_*(\tau),\tau,u_*(\tau)), \tilde{\psi}(\tau)\rangle$$

or, equivalently (see (4.3), (4.4)),

$$\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t)) \overset{a.e.}{=} H(x_*(t),t,\psi^0,\psi(t)).$$

Thus the maximum condition (4.6) is true a.e. on $[0,T]$. The theorem is proved. ∎

Lemma 8. *Assume that the optimal control problem* (P_T) *(see* (3.9)–(3.11)*) is autonomous, i.e., the right-hand side* $f(\cdot,\cdot,\cdot)$ *of control system* (3.9) *and the integrand* $f^0(\cdot,\cdot,\cdot)$ *in the functional* (3.11) *do not depend on the variable* t*, i.e.,* $f(x,t,u) \equiv f(x,u)$ *and* $f^0(x,t,u) \equiv f^0(x,u)$ *for all* $x \in G$ *and* $u \in U$*. Assume also, that an admissible control process* $(x_*(\cdot),u_*(\cdot))$ *satisfies* *the core relations of the maximum principle* (4.5) *and* (4.6) *together with a pair of adjoint variables* $(\psi^0, \psi(\cdot))$ *on a time interval* $[0, T]$, $T > 0$. *Then there the following stationarity condition for the Hamiltonian takes place a.e. on* $[0, T]$*:*

$$\frac{d}{dt} H(x_*(t), \psi^0, \psi(t)) \overset{a.e.}{=} 0. \tag{4.23}$$

Here $H(\cdot, \cdot, \cdot)$ *is the (autonomous) Hamiltonian (see* (4.4)*) of problem* (P_T).

Proof. Since admissible trajectory $x_*(\cdot)$ and adjoint variable $\psi(\cdot)$ are Lipschitz continuous, and ψ^0 is a constant the function $h(\cdot) : [0, T] \mapsto R^1$ defined by the equality

$$h(t) = H(x_*(t), \psi^0, \psi(t)) = \max_{u \in U} \left[\langle f(x_*(t), u), \psi(t) \rangle + \psi^0 f^0(x_*(t), u) \right],$$

(for all $t \in [0, T]$) is Lipschitz continuous on $[0, T]$ as well. Hence $h(\cdot)$ is an absolutely continuous function and its derivative $\dot{h}(\cdot)$ exists a.e. on $[0, T]$.

Denote by $\mathfrak{M}$ the set of all points $\tau \in [0, T]$ at which all functions $h(\cdot)$, $x_*(\cdot)$ and $\psi(\cdot)$ are differentiable, the control $u_*(\cdot)$ is approximative continuous, and both equality (4.5) and the maximum condition (4.6) hold. It is clear that set $\mathfrak{M}$ is of full measure on $[0, T]$, i.e. $\text{meas}\,\{[0, T] \setminus \mathfrak{M}\} = 0$.

Let $\tau \in \mathfrak{M}$ be arbitrary. Now we will demonstrate that (4.23) takes place at τ.

It follows from the definition of the approximate continuity of $u_*(\cdot)$ at τ that there are two sequences $\{t_i^-\}$, $t_i^- \in \mathfrak{M}$, $i = 1, 2, \dots$, and $\{t_i^+\}$, $t_i^+ \in \mathfrak{M}$, $i = 1, 2, \dots$, such that

$$t_i^- \to \tau - 0, \quad \text{as} \quad i \to \infty,$$

$$t_i^+ \to \tau + 0, \quad \text{as} \quad i \to \infty.$$

Consider behavior of function $h(\cdot)$ at points $t_i = t_i^+$, $i = 1, 2, \dots$ Due to the maximum condition (4.6) and equality (4.5) for any $i = 1, 2, \dots$ we get

$$h(t_i) - h(\tau) = H(x_*(t_i), \psi^0, \psi(t_i)) - H(x_*(\tau), \psi^0, \psi(\tau))$$

$$= \mathcal{H}(x_*(t_i), u_*(t_i), \psi^0, \psi(t_i)) - \mathcal{H}(x_*(\tau), u_*(\tau), \psi^0, \psi(\tau))$$

$$= \left\langle \frac{\partial}{\partial x} \mathcal{H}\big(x_*(\tau) + \theta_i(x_*(t_i) - x_*(\tau)), u_*(t_i), \psi^0, \psi(\tau) + \theta_i(\psi(t_i) - \psi(\tau))\big), x_*(t_i) - x_*(\tau) \right\rangle$$

$$+ \left\langle \frac{\partial}{\partial \psi} \mathcal{H}\big(x_*(\tau) + \theta_i(x_*(t_i) - x_*(\tau)), u_*(t_i), \psi^0, \psi(\tau) + \theta_i(\psi(t_i) - \psi(\tau))\big), \psi(t_i) - \psi(\tau) \right\rangle$$

$$= \Big[\Big\langle \frac{\partial}{\partial x}\mathcal{H}(x_*(\tau),u_*(\tau),\psi^0,\psi(\tau)),\dot{x}_*(\tau)\Big\rangle$$

$$+\Big\langle \frac{\partial}{\partial \psi}\mathcal{H}(x_*(\tau),u_*(\tau),\psi^0,\psi(\tau)),\dot{\psi}(\tau)\Big\rangle\Big](t_i-\tau)+\mathrm{o}(t_i-\tau)$$

$$= \Big[-\langle\dot{\psi}(\tau),\dot{x}_*(\tau)\rangle+\langle\dot{x}_*(\tau),\dot{\psi}(\tau)\rangle\Big]\,(t_i-\tau)+\mathrm{o}(t_i-\tau)=\mathrm{o}(t_i-\tau)$$

where $\theta_i \in [\tau, t_i]$ and

$$\frac{\mathrm{o}(t_i-\tau)}{t_i-\tau}\to 0 \quad \text{as} \quad i\to\infty.$$

Dividing the final inequality by $t_i - \tau$ and passing to a limit as $i \to \infty$ we get

$$\dot{h}(\tau) \geq 0.$$

Analogously, considering behavior of function $h(\cdot)$ at points $t_i = t_i^-$, $i = 1, 2, \ldots$, one can prove the opposite inequality

$$\dot{h}(\tau) \leq 0.$$

Thus the following equality holds at arbitrary $\tau \in \mathfrak{M}$:

$$\dot{h}(\tau) = 0.$$ ■

By virtue of Theorem 11 and Lemma 8 the stationarity condition

$$\frac{d}{dt}H(x_*(t),\psi^0,\psi(t)) \stackrel{a.e.}{=} 0 \tag{4.24}$$

holds along any optimal (in autonomous problem (P_T)) admissible trajectory $x_*(\cdot)$ taken together with the corresponding due to the Pontryagin maximum principle (Theorem 11) adjoint variables ψ^0 and $\psi(\cdot)$. Nevertheless it should be noted that Lemma 8 does not assume that the pair $(x_*(\cdot), u_*(\cdot))$ is optimal. Any admissible pair $(x_*(\cdot), u_*(\cdot)$ satisfying together with a pair of adjoint variables $(\psi^0, \psi(\cdot))$ conditions (4.5) and (4.6) of the maximum principle on $[0, T]$, $T > 0$, must satisfy also condition (4.24).

Now consider the situation when the final instant of time $T > 0$ is free in problem (P_T). In this case the following version of the Pontryagin maximum principle is true.

Theorem 12. *Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible process in problem (P_T) (see* (3.9)–(3.11)*) with a free final time $T > 0$ and $T_* > 0$ be the corresponding optimal final instant of time. Then there are a number $\psi^0 \in R^1$ and an absolutely continuous vector function $\psi : [0, T] \mapsto R^n$ (adjoint variables) such that the optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ satisfies together with pair of adjoint variables $(\psi^0, \psi(\cdot))$ to all conditions* $1) - 4)$ *of Theorem* 11 *on $[0, T_*]$. Moreover, the following transversality in time condition holds at the final instant T_*:*

$$H(x_*(T_*),T_*,\psi^0,\psi(T_*)) = 0. \tag{4.25}$$

Proof. The optimality of admissible pair $(x_*(\cdot), u_*(\cdot))$ on time interval $[0, T_*]$, $T_* > 0$, in problem (P_T) with a free final time $T > 0$ is equivalent to validity of the following condition at instant T_* (see (4.2)):

$$\langle \tilde{x} - \tilde{x}_*(T_*), \tilde{\psi}_{T_*} \rangle \leq 0 \quad \text{for any} \quad \tilde{x} \in \tilde{X}(T) \quad \text{and} \quad T > 0 \tag{4.26}$$

where $\tilde{\psi}_{T_*} = (1, 0) \in R^1 \times R^n$.

In particular inequality (4.26) must hold for all $\tilde{x} \in \tilde{X}(T_*)$. In this case we are in the situation described in Theorem 11. Indeed, let $\tilde{\psi} : [0, T] \mapsto R^1 \times R^n$, $\tilde{\psi}(\cdot) = (\psi^0(\cdot), \psi(\cdot))$, be (a unique) solution to the linear differential equation (4.21) with boundary conditions (4.22). Then as it is demonstrated in the proof of Theorem 11 the optimal control process $(x_*(\cdot), u_*(\cdot))$ satisfies together with the pair $(\psi^0, \psi(\cdot))$ to conditions 1) $-$ 4) of Theorem 11 on $[0, T_*]$.

Let us show that condition (4.25) is also satisfied with this pair of adjoint variables $(\psi^0, \psi(\cdot))$.

For arbitrary $\alpha > 0$ and arbitrary $v \in U$ define an admissible control $u_\alpha^v(\cdot)$ on time interval $[0, T_* + \alpha]$ as follows:

$$u_\alpha^v(t) = \begin{cases} u_*(t), & t \in [0, T_*], \\ v, & t \in (T_*, T_* + \alpha]. \end{cases}$$

So, $u_\alpha^v(\cdot)$ coincide with $u_*(\cdot)$ on the time interval $[0, T_*]$ and $u_\alpha^v(t) \equiv v$ on the rest semi-interval $(T_*, T_* + \alpha]$.

Then it is easy to see that the corresponding admissible trajectory $\hat{x}_\alpha^v(\cdot)$ of control system (3.15) takes values

$$\hat{x}_\alpha^v(t) = \tilde{x}_*(T_*) + (t - T_*)\tilde{f}(x_*(T_*), T_*, v) + \tilde{o}(t - T_*)$$

on the time interval $[T_*, \alpha]$. Here

$$\frac{\|\tilde{o}(t - T_*)\|}{t - T_*} \to 0 \quad \text{as} \quad t - T_* \to 0+ .$$

Due to (4.26) we have

$$\langle \hat{x}_\alpha^v(t) - \tilde{x}_*(T_*), \tilde{\psi}_{T_*} \rangle \leq 0 \quad \text{for all} \quad t \in (T_*, T_* + \alpha].$$

Since $\tilde{\psi}_{T_*} = (1, 0) \in R^1 \times R^n$ we have

$$\langle \tilde{f}(x_*(T_*), T_*, v), \tilde{\psi}_{T_*} \rangle = f^0(x_*(T_*), T_*, v) \leq 0.$$

As far as $v \in U$ is arbitrary the last inequality implies

$$H(x_*(T_*), T_*, \psi^0, \psi(T_*)) = \max_{u \in U} f^0(x_*(T_*), T_*, u) \leq 0. \tag{4.27}$$

Let us demonstrate that in fact the last condition holds as an equality. Indeed, assume the contrary, i.e., that

$$H(x_*(T_*), T_*, \psi^0, \psi(T_*)) = \max_{u \in U} f^0(x_*(T_*), T_*, u) < 0.$$

Then due to continuity of the function $H(\cdot,\psi^0,\psi(T_*))$ and continuity of the optimal trajectory $x_*(\cdot)$ we have that there is an $\alpha>0$ such that the following inequality holds on some left neighborhood $(T_*-\varepsilon,T_*]$: $\varepsilon>0$, of point T_*:

$$\max_{u\in U} f^0(x_*(t),t,u)\le -\alpha<0,\quad \text{for all}\quad t\in(T_*-\varepsilon,T_*].$$

Since $(\psi^0,\psi(T_*))=(1,0)\in R^1\times R^n$ and $u_*(\cdot)$ takes values in the set U we get

$$f^0(x_*(t),t,u_*(t))\le -\alpha<0,\quad \text{for all}\quad t\in(T_*-\varepsilon,T_*].$$

But in this case

$$\begin{aligned} x_*^0(t)&=x_*^0(T_*)+\int_{T_*}^{t} f^0(x_*(s),s,u_*(s))ds\\ &\ge x_*^0(T_*)-\alpha(t-T_*)>x_*^0(T_*)\quad \text{for all}\quad t\in(T_*-\varepsilon,T_*) \end{aligned}$$

that contradicts to assumption about optimality of control process $(x_*(\cdot),u_*(\cdot))$ on time interval $[0,T_*]$, $T_*>0$, in problem (P_T) with a free final time $T>0$. Hence, condition (4.27) holds as an equality. Thus, the transversality in time condition (4.25) is proved. ■

Consider now the situation then the optimal control problem (P_T) is non autonomous with fixed final time $T>0$ and both the right-hand side $f(\cdot,\cdot,\cdot)$ of control system (3.9), and the integrand $f^0(\cdot,\cdot,\cdot)$ in the functional (3.11) are continuously differentiable in variables $(x.t)$. In this case (see Section 1) the non autonomous optimal control problem (P_T) can be reduced to an autonomous one by introducing of an auxiliary phase variable $x^{n+1}\in R^1$ such that

$$\dot{x}^{n+1}=1,\quad x^{n+1}(0)=0.$$

In that case $x^{n+1}(t)\equiv t$ for all $t\ge 0$.

Denote the corresponding new Hamilton–Pontryagin function by

$$\tilde{\mathcal{H}}(x,x^{n+1},u,\psi^0,\psi,\psi^{n+1})=\mathcal{H}(x,x^{n+1},u,\psi^{n+1},\psi^0,\psi)+\psi^{n+1}$$

preserving notation

$$\mathcal{H}(x,x^{n+1},u,\psi^0,\psi)=\langle f(x,x^{n+1},u),\psi\rangle+\psi^0 f^0(x,x^{n+1},u)$$

for the old one. Here $x\in G$, $x^{n+1}\in[0,T]$, $u\in U$, $\psi\in R^n$, $\psi^0\in R^1$ and $\psi^{n+1}\in R^1$.

Let $(\tilde{x}_*(\cdot),u_*(\cdot))$ be an optimal process in the nonautonomous optimal control problem (P_T) with a fixed final time $T>0$. Then treating (P_T) as autonomous optimal control problem as described above (i.e. with additional state coordinate $\dot{x}^{n+1}(t)\equiv 1$ on $[0,T]$, $x^{n+1}(0)=0$) and applying Theorem 11 we get that there are adjoint variables ψ^0, $\psi(\cdot)$, $\psi^{n+1}(\cdot)$ such that $\psi^0\in R^1$ is a constant, $\psi:[0,T]\mapsto R^n$ and $\psi^{n+1}:[0,T]\mapsto R^1$ are absolutely continuous functions, and the following conditions hold:

1) the vector function $\psi(\cdot)$ is a (Carathéodory) solution to the following adjoint system on $[0, T]$:

$$\begin{aligned}\dot{\psi} &= -\frac{\partial \mathcal{H}(x_*(t), t, u_*(t), \psi^0, \psi(t))}{\partial x} \\ &= -\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x}\right]^* \psi(t) - \psi^0 \frac{\partial f^0(x_*(t), t, u_*(t))}{\partial x},\end{aligned}$$

while the adjoint variable $\psi^{n+1}(\cdot)$ satisfies the following equality a.e.on $[0, T]$;

$$\begin{aligned}\dot{\psi}^{n+1}(t) &= -\frac{\partial \mathcal{H}(x_*(t), t, u_*(t), \psi^0, \psi(t))}{\partial t} \\ &= -\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial t}\right]^* \psi(t) - \psi^0 \frac{\partial f^0(x_*(t), t, u_*(t))}{\partial t};\end{aligned} \tag{4.28}$$

2) the following maximum condition takes place on $[0, T]$:

$$\mathcal{H}(x_*(t), t, u_*(t), \psi^0, \psi(t)) \overset{a.e.}{=} H(x_*(t), t, \psi^0, \psi(t));$$

3) The pair $(\psi^0, \psi(\cdot))$ is nontrivial, i.e.,

$$\psi^0 + \|\psi(0)\| > 0; \tag{4.29}$$

4) the following transversality conditions are valid:

$$\psi^0 > 0, \quad \psi(T) = 0, \quad \psi^{n+1}(T) = 0. \tag{4.30}$$

As far as $\psi^0 > 0$, similarly to Theorem 11, the nontriviality condition (4.29) follows from the transversality condition (4.30) in this case.

Note also that due to the transversality condition $\psi^{n+1}(T) = 0$ (see (4.30)) and equality (4.28) the following explicit characterization of the adjoint variable $\psi^{n+1}(\cdot)$ holds:

$$\psi(t) = \int_t^T \frac{\partial}{\partial s} \mathcal{H}(x_*(s), s, u_*(s), \psi^0, \psi(s))\, ds \quad \text{for all} \quad t \in [0, T].$$

Thus the adjoint variable $\psi^{n+1}(\cdot)$ can be completely excluded from the relations of the maximum principle in this case.

In particular, due to Lemma 8 the stationarity condition for the Hamiltonian on $[0, T]$ (see (4.24))

$$\frac{d}{dt} H(x_*(t), t, \psi^0, \psi(t)) + \dot{\psi}^{n+1}(t) \overset{a.e.}{=} 0$$

can be rewritten equivalently in this case as the follows:

$$\frac{d}{dt} H(x_*(t), t, \psi^0, \psi(t)) \overset{a.e.}{=} \frac{\partial}{\partial t} \mathcal{H}(x_*(t), t, u_*(t), \psi^0, \psi(t)).$$

Thus we have proved the following result.

Theorem 13. *Assume that both the right-hand side $f(\cdot,\cdot,\cdot)$ of control system* (3.9)*, and the integrand $f^0(\cdot,\cdot,\cdot)$ in the functional* (3.11) *are continuously differentiable in (x,t). Let $(x_*(\cdot),u_*(\cdot))$ be an optimal admissible process in the optimal control problem (P_T). Then there are an absolutely continuous vector function $\psi:[0,T]\mapsto R^n$ and a number $\psi^0\in R^1$ (adjoint variables) such that the following conditions hold:*

1) *the vector function $\psi(\cdot)$ is a (Carathéodory) solution to the following adjoint system on $[0,T]$:*

$$\dot\psi = -\frac{\partial \mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi)}{\partial x} = -\left[\frac{\partial f(x_*(t),t,u_*(t))}{\partial x}\right]^*\psi - \psi^0\frac{\partial f^0(x_*(t),t,u_*(t))}{\partial x}; \tag{4.31}$$

2) *the following maximum condition takes place on $[0,T]$:*

$$\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t)) \stackrel{a.e.}{=} H(x_*(t),t,\psi^0,\psi(t)); \tag{4.32}$$

3) *the pair $(\psi^0,\psi(\cdot))$ is nontrivial, i.e.,*

$$\psi^0 + \|\psi(0)\| > 0; \tag{4.33}$$

4) *the following transversality conditions are valid:*

$$\psi^0 > 0, \qquad \psi(T) = 0. \tag{4.34}$$

5) *the Hamiltonian $H(x_*(\cdot),\psi^0,\psi(\cdot))$ is an absolutely continuous function of time t on $[0,T]$ and the following stationarity condition takes place:*

$$\frac{d}{dt}H(x_*(t),t,\psi^0,\psi(t)) \stackrel{a.e.}{=} \frac{\partial}{\partial t}\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t)). \tag{4.35}$$

It is easy to see that if the final instant $T>0$ in problem (P_T) is free in addition to other assumptions of Theorem 13 then the transversality in time condition

$$H(x_*(T),T,\psi^0,\psi(T)) = 0 \tag{4.36}$$

is valid as well (see (4.25)).

Lastly we consider the following slightly more general optimal control problem $(\hat P_T)$ with free right terminal state on a fixed time interval $[0,T]$, $T>0$:

$$\dot x = f(x,t,u), \quad u\in U, \tag{4.37}$$

$$x(0) = x_0,$$

$$J(T,x(\cdot),u(\cdot)) = \int_0^T f^0(x(t),u(t))\,dt + V(x(T)) \to \max. \tag{4.38}$$

Here all data in problem $(\hat P_T)$ are the same as in problem (P_T). We assume also that both the right-hand side $f(\cdot,\cdot,\cdot)$ of control system (4.37) and the integrand $f^0(\cdot,\cdot,\cdot)$ in the functional (4.38) are continuously differentiable in variables (x,t), and the scalar function $V(\cdot):G\mapsto R^1$ in (4.38) is twice continuously differentiable. As usual the class of admissible controls $u(\cdot)$ of system (4.37) consists of all measurable vector functions $u : [0,T] \mapsto U$.

It is easy to see that in the context of problem $(\hat{P}_T)$ maximization of functional (4.38) is equivalent to maximization of the following integral functional (see Section 1):

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

where function

$$\hat{f}^0(x,t,u) = f^0(x,t,u) + \left\langle \frac{\partial V(x)}{\partial x}, f(x,t,u) \right\rangle \quad \text{for all} \quad x \in G, \; u \in U$$

is continuous in x, t, u and continuously differentiable function in (x,t).

Denote the corresponding new Hamilton–Pontryagin function by

$$\hat{\mathcal{H}}(x,t,u,\psi^0,\psi) = \mathcal{H}(x,t,u,\psi^0,\psi) + \psi^0 \left\langle \frac{\partial V(x)}{\partial x}, f(x,t,u) \right\rangle$$

preserving notation

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

for the old one. Analogously, define the Hamiltonian of problem $(\hat{P}_T)$

$$\hat{H}(x,t,\psi^0,\psi) = \sup_{u \in U} \hat{\mathcal{H}}(x,t,u,\psi^0,\psi)$$

preserving notation

$$H(x,t,\psi^0,\psi) = \sup_{u \in U} \mathcal{H}(x,t,u,\psi^0,\psi) = \sup_{u \in U} \left[\langle f(x,t,u), \psi \rangle + \psi^0 f^0(x,t,u) \right]$$

for the Hamiltonian of problem (P_T). Here $x \in G$, $t \in [0,T]$, $u \in U$, $\psi^0 \in R^1$ and $\psi \in R^n$.

Now applying Theorem 11 we get that if $(x_*(\cdot), u_*(\cdot))$ is an optimal admissible process in the optimal control problem $(\hat{P}_T)$ then there are an absolutely continuous vector function $\hat{\psi} : [0,T] \mapsto R^n$ and a number $\psi^0 \in R^1$ (adjoint variables) such that the following conditions hold:

1) the vector function $\hat{\psi}(\cdot)$ is a (Carathéodory) solution to the following adjoint system on $[0,T]$:

$$\dot{\hat{\psi}} = -\frac{\partial \hat{\mathcal{H}}(x_*(t), t, u_*(t), \psi^0, \hat{\psi}(t))}{\partial x} = -\frac{\partial \mathcal{H}(x_*(t), t, u_*(t), \psi^0, \hat{\psi}(t))}{\partial x}$$
$$-\psi^0 \left[\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x} \right]^* \frac{\partial V(x_*(t))}{\partial x} - \left[\frac{\partial^2 V(x_*(t))}{\partial x^2} \right]^* f(x_*(t), t, u_*(t)) \right]; \tag{4.39}$$

2) the following maximum condition takes place on $[0,T]$:

$$\hat{\mathcal{H}}(x_*(t), t, u_*(t), \psi^0, \hat{\psi}(t)) \stackrel{a.e.}{=} \hat{H}(x_*(t), t, \psi^0, \hat{\psi}(t)); \tag{4.40}$$

3) the pair $(\psi^0, \hat{\psi}(\cdot))$ is nontrivial, i.e.,

$$\psi^0 + \|\hat{\psi}(0)\| > 0; \tag{4.41}$$

4) the following transversality conditions are valid:

$$\psi^0 > 0, \quad \hat{\psi}(T) = 0. \tag{4.42}$$

Note, that here as above the transversality condition (4.42) implies the nontriviality condition (4.41).

Let us introduce a new adjoint variable $\psi : [0, T] \mapsto R^n$:

$$\psi(t) = \hat{\psi}(t) + \psi^0 \frac{\partial V(x_*(t))}{\partial x}, \quad t \in [0, T] \tag{4.43}$$

and rewrite conditions (4.39)-(4.42) in terms of the adjoint variables ψ^0 and $\psi(\cdot)$.

First, due to (4.39) and (4.43) we have a.e. on $[0, T]$

$$\dot{\psi}(t) = -\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x}\right]^* \left(\psi(t) - \psi^0 \frac{\partial V(x_*(t))}{\partial x}\right) - \psi^0 \frac{\partial f^0(x_*(t), t, u_*(t)}{\partial x}$$

$$-\psi^0 \left[\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x}\right]^* \frac{\partial V(x_*(t))}{\partial x} - \left[\frac{\partial^2 V(x_*(t))}{\partial x^2}\right]^* f(x_*(t), t, u_*(t))\right]$$

$$+\psi^0 \left[\frac{\partial^2 V(x_*(t))}{\partial x^2}\right]^* f(x_*(t), t, u_*(t))$$

or

$$\dot{\psi}(t) = -\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x}\right]^* \psi(t) - \psi^0 \frac{\partial f^0(x_*(t), t, u_*(t))}{\partial x}. \tag{4.44}$$

Second, due to (4.43) the maximum condition (4.40) takes the form

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

Third, it is easy to see that due to (4.34) the nontriviality condition (4.41) is equivalent to condition

$$\psi^0 + \|\psi(0)\| > 0; \tag{4.45}$$

Finally, due to (4.34) the transversality condition (4.42) transforms in the following one:

$$\psi^0 > 0, \quad \psi(T) = \psi^0 \frac{\partial V(x_*(T))}{\partial x}. \tag{4.46}$$

Due to (4.44)-(4.46) the following result is the Pontryagin maximum principle for problem $(\hat{P}_T)$.

Theorem 14. *Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible process in the optimal control problem $(\hat{P}_T)$. Then there are a number $\psi^0 \in R^1$ and an absolutely continuous vector function $\psi : [0, T] \mapsto R^n$ (adjoint variables) such that the following conditions hold:*

1) *the vector function $\psi(\cdot)$ is a (Carathéodory) solution to the following adjoint system on* $[0,T]$*:*

$$\dot{\psi}(t) = -\frac{\partial \mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t))}{\partial x}$$
$$= -\left[\frac{\partial f(x_*(t),t,u_*(t))}{\partial x}\right]^* \psi(t) - \psi^0 \frac{\partial f^0(x_*(t),t,u_*(t))}{\partial x};$$

2) *the following maximum condition takes place on* $[0,T]$*:*

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

3) *the pair $(\psi^0,\psi(\cdot))$ is nontrivial, i.e.,*

$$\psi^0 + \|\psi(0)\| > 0;$$

4) *the following transversality conditions are valid:*

$$\psi^0 > 0, \quad \psi(T) = \psi^0 \frac{\partial V(x_*(T))}{\partial x}.$$

Note that if both the right-hand side $f(\cdot,\cdot,\cdot)$ of control system (4.37), and the integrand $f^0(\cdot,\cdot,\cdot)$ in the functional (4.38) are continuously differentiable in (x,t) then due to Lemma 8 (similarly to Theorem 13) we get that the following stationarity condition (see (4.35)) holds along the optimal process $(x_*(\cdot),u_*(\cdot))$ on $[0,T]$:

$$\frac{d}{dt}H(x_*(t),t,\psi^0,\psi(t)) \stackrel{a.e.}{=} \frac{\partial}{\partial t}\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t)).$$

Moreover, it is easy to see that if the final time T is free in problem $(\hat{P}_T)$ in addition to other assumptions of Theorem 14 then the transversality in time condition

$$H(x_*(T),T,\psi^0,\psi(T)) = 0$$

is valid as well (see (4.25)).

5. Optimal economic growth problem

Starting from this section we focus consideration on the infinite-horizon optimal control problem (P) (see (1.34)–(1.36))):

$$\dot{x} = f(x, u), \quad u \in U, \tag{5.1}$$

$$x(0) = x_0, \tag{5.2}$$

$$J(x, u) = \int_0^\infty e^{-\rho t} g(x, u)\, dt \to \max. \tag{5.3}$$

Here, $x = (x^1, \dots, x^n) \in R^n$ and $u = (u^1, \dots, u^m) \in R^m$ are the phase vector of the system and the vector of control parameters (control), respectively; $x_0 \in G$ where G is a given open subset of R^n and U is a nonempty compactum in R^m; $\rho \geq 0$ is a discount parameter. Further, $f : G \times U \mapsto R^n$ is a vector function of two vector variables $x \in G$ and $u \in U$. The instantaneous utility $g : G \times U \mapsto R^1$ is a scalar function of two vector variables $x \in G$ and $u \in U$. We assume that both vector function $f(\cdot, \cdot)$ and scalar function $g(\cdot, \cdot)$ are continuous jointly in variables x and u, and continuously differentiable in variable x. The class of admissible controls in problem (P) consists of all measurable vector functions $u : [0, \infty) \mapsto U$.

Let $u(\cdot)$ be an admissible control in problem (P). Then substituting $u(\cdot)$ in the right-hand side of the control system (5.1) we get the following Cauchy problem on infinite time interval $[0, \infty)$:

$$\dot{x} = f(x, u(t)), \tag{5.4}$$

$$x(0) = x_0. \tag{5.5}$$

We assume that for arbitrary $T > 0$ the (Carathéodory) solution $x(\cdot)$ to the Cauchy problem (5.4), (5.5) exists on $[0, T]$ and takes values in G. In this case the solution $x(\cdot)$ is defined on $[0, \infty)$ and it is locally absolutely continuous, i.e., the vector function $x : [0, T] \mapsto R^n$ is absolutely continuous on any finite time interval $[0, T]$, $T > 0$. The vector function $x(\cdot)$ is called an admissible trajectory corresponding to admissible control $u(\cdot)$. The pair $(x(\cdot), u(\cdot))$ is called an admissible pair or a control process. So, all admissible pairs are defined on the whole infinite time interval $[0, \infty)$ and we assume that for any admissible pair $(x(\cdot), u(\cdot))$ the corresponding integral in the utility functional (5.3) converges absolutely. An admissible pair $(x_*(\cdot), u_*(\cdot))$ is called optimal in problem (P) if it provides the maximal value of the functional (5.3).

Problem (P) naturally arises in various studies of economic growth processes. So, we consider (P) as a generic optimal economic growth problem.

Now let us specify assumptions about the problem (P).

In that follows we assume that for arbitrary $T > 0$ the control system (5.1) satisfies conditions (A1) and (A2) (see Section 2), i.e., we assume that

1) for arbitrary $T > 0$ there is a nonempty closed subset Π of the set G such that for any admissible trajectory $x(\cdot)$ of control system (5.1) its values $x(t)$ belong to the set Π for all $t \in [0, T]$;

2) for arbitrary $T > 0$ there exists a $C_0 \geq 0$ such that

$$\langle x, f(x,u)\rangle \leq C_0(1 + \|x\|^2) \quad \text{for any} \quad x \in \Pi,\ u \in U.$$

In this case by Lemma 1 admissible trajectories of control system (5.1) are uniformly bounded on any final time interval $[0, T]$, $T > 0$. Moreover, for any $T > 0$ due to Lemma 2 the set of admissible trajectories of control system (5.1) is a precompactum in the space $C([0,T], R^n)$.

In that follows we assume also that the following condition about the control system (5.1) and the integrand $g(\cdot,\cdot)$ in the utility functional (5.3) holds.

(A5) *The control system* (5.1) *is affine in control variable u, i.e., it has the following form (see* (2.44)*):*

$$\dot{x} = f_0(x) + \sum_{j=1}^{m} f_j(x)u^j, \quad u \in U. \tag{5.6}$$

Here all vector functions $f_j(\cdot) : G \mapsto R^n$, $j = 0, \dots, m$, are continuously differentiable on G. Moreover, the set U is a convex compactum in R^m and the instantaneous utility $g(x, \cdot)$ is a concave function of variable u for any fixed $x \in G$.

It is easy to see that validity of condition (A5) implies that for any $x \in G$ the vectogram of the control system (5.1), i.e., the set

$$F(x) = \bigcup_{u \in U} f(x,u)$$

and the set

$$\tilde{Q}(x) = \big\{(z^0, z) \in R^1 \times R^n :\, z^0 \leq g(x,u),\ z = f(x,u),\ u \in U\big\}$$

are convex. So, both conditions (A3) and (A4) (see Section 2) are satisfied in this case.

Note, that assumption about the affine in control structure (5.6) of the right-hand side of control system (5.1) (see (A5)) is of technical character mainly. In fact, all results presented below can be proved (under some suitable conditions) in the general nonlinear case as well. Since the general proofs are cumbersome (see the monograph [**6**] for more details) for the sake of simplicity of presentation we restrict our analysis here by the case of affine in control system (5.6).

Finally, we assume that the following condition on the problem (P) takes place.

(A6) *There exist positive decreasing functions* $\mu(\cdot)$ *and* $\omega(\cdot)$ *on* $[0,\infty)$ *such that* $\mu(t)\to\ 0+$ *and* $\omega(t)\to\ 0+$ *as* $t\to\infty$, *and for every admissible pair* $(x(\cdot),u(\cdot))$ *the following inequalities hold:*

$$e^{-\rho t}\max_{u\in U}|g(x(t),u)|\leq\mu(t)\quad\textit{for any}\quad t\geq 0,\tag{5.7}$$

$$\int_T^{\infty}e^{-\rho t}|g(x(t),u(t))|\,dt\leq\omega(T)\quad\textit{for any}\quad T\geq 0.\tag{5.8}$$

Now we are ready to establish existence of an optimal admissible control in problem (P).

Theorem 15. *Suppose that for arbitrary* $T>0$ *conditions* (A1), (A2) *take place, and conditions* (A5), (A6) *hold as well. Then, there exists an optimal admissible control* $u_*(\cdot)$ *in problem* (P).

Proof. Let $\{T_k\}$, $k=1,2,\dots$, be an arbitrary increasing sequence of positive numbers such that $\lim_{k\to\infty}T_k=\infty$.

Consider the following auxiliary sequence $\{(Q_k)\}$, $k=1,2,\dots$, of optimal control problems (Q_k), $k=1,2,\dots$, each of which is defined on its own finite time interval $[0,T_k]$:

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

$$x(0)=x_0,$$

$$J_k(x,u)=\int_0^{T_k}e^{-\rho t}g(x,u)\,dt\to\max.$$

Here, for any $k=1,2,\dots$, the functions $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$, the sets U and G, and the initial state x_0 in problem (Q_k) are the same as in the original problem (P). The only difference of (Q_k) from (P) is that it is considered on the finite time interval $[0,T_k]$. As in the problem (P), the admissible controls in (Q_k) are measurable vector functions $u:[0,T_k]\mapsto U$.

By conditions (A1), (A2) and (A5) and Theorem 10 (see Section 3), for any $k=1,2,\dots$ there exists an optimal admissible pair $(x_k(\cdot),u_k(\cdot))$ in problem (Q_k). The pair $(x_k(\cdot),u_k(\cdot))$ is defined on the finite time interval $[0,T_k]$. Assume that this pair is extended in an arbitrary admissible way to the whole time interval $[0,\infty)$, thus giving an admissible pair for problem (P).

Consider the sequence $\{(x_k(\cdot),u_k(\cdot)\}$, $k=1,2,\dots$, of admissible control processes on the finite time interval $[0,T_1]$.

By conditions (A1), (A2) and (A5) (see Lemmas 5 and 6 in Section 2) there exist a subsequence $\{(x_{1,k}(\cdot), u_{1,k}(\cdot))\}$, $k = 1, 2, \ldots$, of the sequence $\{(x_k(\cdot), u_k(\cdot))\}$, $k = 1, 2, \ldots$, and an admissible pair $(x_*(\cdot), u_*(\cdot))$ of the control system (5.1) on the time interval $[0, T_1]$ such that the sequence $\{u_{1,k}\}$, $k = 1, 2, \ldots$, converges to $u_*(\cdot)$ weakly in $L^1([0, T_1], R^m)$ and the sequence $\{x_{1,k}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $x_*(\cdot)$ uniformly on the interval $[0, T_1]$ (i.e., in the space $C([0, T_1], R^n)$) as $k \to \infty$.

Now, consider the sequence of admissible pairs $\{(x_{1,k}(\cdot), u_{1,k}(\cdot))\}$, $k = 1, 2, \ldots$, on the time interval $[0, T_2]$.

By analogy with the previous case, conditions (A1), (A2), and (A5) (see Lemmas 5 and 6) imply that there exists a subsequence $\{(x_{2,k}(\cdot), u_{2,k}(\cdot))\}$, $k = 1, 2, \ldots$, of the sequence $\{(x_{1,k}(\cdot), u_{1,k})(\cdot)\}$, $k = 1, 2, \ldots$, and an an admissible pair $(x_*(\cdot), u_*(\cdot))$ of the control system (5.1) on the time interval $[0, T_2]$ such that the sequence $\{u_{2,k}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $u_*(\cdot)$ weakly in $L^1([0, T_2], R^m)$ and the sequence $\{x_{2,k}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $x_*(\cdot)$ uniformly on the interval $[0, T_2]$ as $k \to \infty$. Here we preserve notation for the control process $(x_*(\cdot), u_*(\cdot))$ on the time interval $[0, T_2]$ because it is an extension of the process previously constructed on $[0, T_1]$.

Repeating this procedure infinitely many times, we successively construct an admissible pair $(x_*(\cdot), u_*(\cdot))$ on the whole time interval $[0, \infty)$. Simultaneously, we obtain a family of auxiliary optimal control problems $\{(Q_{i,k})\}$, $i = 1, 2, \ldots$, $k = 1, 2, \ldots$, on finite time intervals $[0, T_{i,k}]$, $T_i \leq T_{i,k}$, and a family of optimal admissible pairs $\{(x_{i,k}(\cdot), u_{i,k}(\cdot))\}$ in problems $(Q_{i,k})$ such that for any $i = 1, 2, \ldots$ the sequence $\{u_{i,k}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $u_*(\cdot)$ weakly in $L^1([0, T_i], R^m)$ and the sequence $\{x_{i,k}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $x_*(\cdot)$ uniformly on the interval $[0, T_i]$ as $k \to \infty$.

Take a diagonal subsequence $\{(Q_{k,k})\}$, $k = 1, 2, \ldots$, in the family of problems $\{(Q_{i,k})\}$, $i = 1, 2, \ldots$, $k = 1, 2, \ldots$, and denote it by $\{(Q_{m(k)})\}$, $k = 1, 2, \ldots$. Accordingly, we obtain a sequence of optimal admissible pairs $\{(x_{m(k)}(\cdot), u_{m(k)}(\cdot))\}$, $k = 1, 2, \ldots$, in problems $(Q_{m(k)})$ on the finite time intervals $[0, T_{m(k)}]$, $m(k) \geq k$. By construction for any $T > 0$ the sequence $\{u_{m(k)}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $u_*(\cdot)$ weakly in $L^1([0, T], R^m)$ and the sequence $\{x_{m(k)}(\cdot)\}$, $k = 1, 2, \ldots$, converges to $x_*(\cdot)$ uniformly on the interval $[0, T]$ as $k \to \infty$. Note that by construction $m(k) \geq k$ for $k = 1, 2, \ldots$

Let us prove that the admissible pair $(x_*(\cdot), u_*(\cdot))$ thus constructed is optimal in problem (P).

Suppose the contrary; i.e., let there exist an admissible pair $(\hat{x}(\cdot), \hat{u}(\cdot))$ and a number $\varepsilon > 0$ such that

$$J(x_*(\cdot), u_*(\cdot)) < J(\hat{x}(\cdot), \hat{u}(\cdot)) - \varepsilon. \tag{5.9}$$

By virtue of condition (A6) (see (5.8)), for the chosen number ε there exists a natural k_1 such that

$$\omega(T_{m(k)}) < \frac{\varepsilon}{4} \tag{5.10}$$

for all $k \geq k_1$.

Consider the sequence of pairs $\{(x_{m(k)}(\cdot), u_{m(k)}(\cdot))\}$, $k = 1, 2, \ldots$, on the time interval $[0, T_{m(k_1)}]$.

Due to Theorem 5 (see Section 2) the integral functional

$$J_{m(k_1)}(x(\cdot), u(\cdot)) = \int_0^{T_{m(k_1)}} e^{-\rho t} g(x(t), u(t))\, dt$$

is upper semicontinuous. Hence there is a $k_2 \geq k_1$ such that for any $k \geq k_2$ we have

$$\int_0^{T_{m(k_1)}} e^{-\rho t} g(x_*(t), u_*(t))\, dt \geq \int_0^{T_{m(k_1)}} e^{-\rho t} g(x_{m(k)}(t), u_{m(k)}(t))\, dt - \frac{\varepsilon}{4}. \tag{5.11}$$

Since the pair $\left(x_{m(k_2)}(\cdot), u_{m(k_2)}(\cdot)\right)$ is optimal in problem $(Q_{m(k_2)})$, we have

$$\begin{aligned} J_{m(k_2)}\left(x_{m(k_2)}(\cdot), u_{m(k_2)}(\cdot)\right) &= \int_0^{T_{m(k_2)}} e^{-\rho t} g\left(x_{m(k_2)}(t), u_{m(k_2)}(t)\right)\, dt \\ &\geq \int_0^{T_{m(k_2)}} e^{-\rho t} g(\hat{x}(t), \hat{u}(t))\, dt. \end{aligned} \tag{5.12}$$

Hence, by conditions (5.10) and (5.8) we obtain

$$\begin{aligned} \int_0^{T_{m(k_2)}} e^{-\rho t} g\left(x_{m(k_2)}(t), u_{m(k_2)}(t)\right)\, dt &\geq \int_0^{\infty} e^{-\rho t} g(\hat{x}(t), \hat{u}(t))\, dt - \frac{\varepsilon}{4} \\ &= J(\hat{x}, \hat{u}) - \frac{\varepsilon}{4}. \end{aligned} \tag{5.13}$$

Therefore, in view of conditions (5.12) and (5.13), the following inequalities hold:

$$\begin{aligned} J(\hat{x}(\cdot), \hat{u}(\cdot)) &\leq \int_0^{T_{k_2}} e^{-\rho t} g\big(x_{m(k_2)}(t), u_{m(k_2)}(t)\big)\, dt + \frac{\varepsilon}{4} \\ &\leq \int_0^{T_{k_1}} e^{-\rho t} g\big(x_{m(k_2)}(t), u_{m(k_2)}(t)\big)\, dt + \frac{\varepsilon}{2}, \end{aligned}$$

which, by (5.11), imply the inequality

$$J(\hat{x}(\cdot), \hat{u}(\cdot)) \leq \int_0^{T_{k_1}} e^{-\rho t} g(x_*(t), u_*(t))\, dt + \frac{3\varepsilon}{4}.$$

Thus, (5.8) yields

$$J(\hat{x}(\cdot), \hat{u}(\cdot)) \leq \int_0^{\infty} e^{-\rho t} g(x_*(t), u_*(t))\, dt + \varepsilon = J(x_*, u_*) + \varepsilon,$$

which contradicts assumption (5.9). The theorem is proved. ■

6. Finite-horizon approximations

In this section we describe a finite-horizon approximation approach (see [**4**]–[**6**]) for deriving different versions of the maximum principle for the infinite-horizon optimal control problem (P) (see (5.1)–(5.3)) from the classical Pontryagin maximum principle for finite-horizon problem (P_T) (see Theorem 11 in Section 4). As it was already mentioned in Section 5 for the sake of simplicity of presentation we consider only the case when the control system (5.1) in problem (P) can be represented as (5.6), the set U is convex and compact, and the scalar function $g(x, \cdot)$ is concave in variable u for any fixed $x \in G$. So, we assume that conditions (A5) and (A6) for problem (P) are satisfied together with conditions (A1), (A2) (on any finite time interval $[0, T]$, $T > 0$). Regarding the general nonlinear case we refer to [**6**].

Note also, that approximation approach presented here can be used not only for developing general results but it can be applied to particular problems directly, i.e., for deriving specialized versions of the Pontryagin maximum principle in specific nonstandard situations.

Now, let us pass to the construction of a sequence of finite-horizon approximating problems (P_k), $k = 1, 2, \dots$.

Let $u_*(\cdot)$ be a given optimal control in problem (P). Choose a sequence of continuously differentiable vector functions $\{z_k(\cdot)\}$, $k = 1, 2, \dots$, $z_k : [0, \infty) \mapsto R^m$, and a sequence of positive numbers $\{\sigma_k\}$, $k = 1, 2, \dots$, such that

$$\sup_{t \in [0,\infty)} \|z_k(t)\| \leq \max_{u \in U} \|u\| + 1, \tag{6.1}$$

$$\int_0^\infty e^{-(\rho+1)t} \|z_k(t) - u_*(t)\|^2 \, dt \leq \frac{1}{k}, \tag{6.2}$$

$$\sup_{t \in [0,\infty)} \|\dot{z}_k(t)\| \leq \sigma_k < \infty, \tag{6.3}$$

$$\sigma_k \to \infty \quad \text{as} \quad k \to \infty.$$

Obviously, such sequences $\{z_k(\cdot)\}$, $k = 1, 2, \dots$, and $\{\sigma_k\}$, $k = 1, 2, \dots$, exist.

Next, we choose a monotonically increasing sequence of positive numbers $\{T_k\}$, $k = 1, 2, \dots$, such that

$$T_k \to \infty \quad \text{as} \quad k \to \infty$$

and

$$\omega(T_k) \leq \frac{1}{k(1+\sigma_k)} \quad \text{for any} \quad k = 1, 2, \ldots. \tag{6.4}$$

Recall that, according to condition (A6), the function $\omega(\cdot)$ satisfies inequality (5.8).

For an arbitrary $k = 1, 2, \ldots$, define a problem (P_k) as follows:

$$\dot{x} = f(x, u), \quad u \in U, \tag{6.5}$$

$$x(0) = x_0, \tag{6.6}$$

$$J_k(x, u) = \int_0^{T_k} e^{-\rho t}\Big[g(x, u) - \frac{1}{1+\sigma_k} e^{-t}\|u - z_k(t)\|^2\Big]\, dt \to \max. \tag{6.7}$$

Here, $x \in R^n$, $u \in R^m$, $x_0 \in G$, the vector function $f(\cdot,\cdot)$, the scalar function $g(\cdot,\cdot)$, the discount parameter $\rho \geq 0$, and the sets U and G are the same as in the original problem (P). As usual, the class of admissible controls in problem (P_k), $k = 1, 2, \ldots$, consists of all measurable vector functions $u : [0, T_k] \mapsto U$.

We say that (P_k), $k = 1, 2, \ldots$, is an approximating sequence of problems corresponding to the optimal control $u_*(\cdot)$.

Since problem (P) satisfies conditions (A1), (A2)(on any finite time interval $[0, T]$, $T > 0$) and (A5), problem (P_k) also satisfies conditions (A1), (A2) and (A5) for any $k = 1, 2, \ldots$. Hence, by Theorem 10 (see Section 3), for any $k = 1, 2, \ldots$, there exists an optimal admissible pair $(x_k(\cdot), u_k(\cdot))$ in problem (P_k) (on the time interval $[0, T_k]$). Further, we will assume that any pair $(x_k(\cdot), u_k(\cdot))$ that is admissible in problem (P_k), $k = 1, 2, \ldots$, is extended from the time interval $[0, T_k]$ to the entire infinite time interval $[0, \infty)$ in an arbitrary way admissible for the control system (5.1).

In the case of systems that are affine in control, the following result provides the main instrument for analyzing problem (P) by means of the approximating sequence of problems (P_k), $k = 1, 2, \ldots$.

Lemma 9. *Suppose that for arbitrary $T > 0$ conditions* (A1), (A2) *take place, and conditions* (A5), (A6) *hold as well. Further, let $u_*(\cdot)$ be a given optimal admissible control in problem (P) and $\{(P_k)\}$, $k = 1, 2, \ldots$, be an approximating sequence of optimal control problems that corresponds to $u_*(\cdot)$. Finally, let $u_k(\cdot)$ be an optimal control in problem (P_k), $k = 1, 2, \ldots$. Then, for any $T > 0$, we have*

$$u_k(\cdot) \to u_*(\cdot) \quad \textit{in} \quad L^2([0, T], R^m) \quad \textit{as} \quad k \to \infty.$$

Proof. Take an arbitrary $T > 0$. Let k_1 be a natural number such that

$T_{k_1} \geq T$. Then, for any $k \geq k_1$, we have

$$J_k(x_k(\cdot), u_k(\cdot)) = \int_0^{T_k} e^{-\rho t} \left[g(x_k(t), u_k(t)) - e^{-t} \frac{\|u_k(t) - z_k(t)\|^2}{1+\sigma_k} \right] dt$$
$$\leq \int_0^{T_k} e^{-\rho t} g(x_k(t), u_k(t))\, dt - \frac{e^{-(\rho+1)T}}{1+\sigma_k} \int_0^T \|u_k(t) - z_k(t)\|^2\, dt,$$

where $u_k(\cdot)$ is an optimal control in problem (P_k) and $x_k(\cdot)$ is the corresponding trajectory of the control system (5.1) (on the time interval $[0, T_k]$). Recall that the optimal pair $(x_k(\cdot), u_k(\cdot))$ is extended from the time interval $[0, T_k]$ to the entire infinite time interval $[0, \infty)$ in a way that is admissible for the control system (5.1).

Let $x_*(\cdot)$ be the trajectory corresponding to the control $u_*(\cdot)$. Since the control $u_k(\cdot)$ is optimal in problem (P_k), $k = 1, 2, \ldots$, the control $u_*(\cdot)$ is optimal in problem (P), the pair $(x_*(\cdot), u_*(\cdot))$ is admissible in problem (P_k), and the pair $(x_k(\cdot), u_k(\cdot))$ is admissible in problem (P), it follows from conditions (5.8), (6.2), and (6.4) that

$$\frac{e^{-(\rho+1)T}}{1+\sigma_k} \int_0^T \|u_k(t) - z_k(t)\|^2\, dt \leq \int_0^{T_k} e^{-\rho t} g(x_k(t), u_k(t))\, dt - J_k(x_*(\cdot), u_*(\cdot))$$
$$\leq \int_0^{T_k} e^{-\rho t} g(x_k(t), u_k(t))\, dt - J(x_*(\cdot), u_*(\cdot)) + \omega(T_k)$$
$$+ \int_0^{\infty} \frac{e^{-(\rho+1)t}}{1+\sigma_k} \|u_*(t) - z_k(t)\|^2\, dt$$
$$\leq \int_0^{T_k} e^{-\rho t} g(x_k(t), u_k(t))\, dt - J(x_*(\cdot), u_*(\cdot)) + \frac{2}{k(1+\sigma_k)}$$
$$\leq J(x_k(\cdot), u_k(\cdot)) - J(x_*(\cdot), u_*(\cdot)) + \frac{3}{k(1+\sigma_k)} \leq \frac{3}{k(1+\sigma_k)}$$

for all $k \geq k_1$. Therefore,

$$\int_0^T \|u_k(t) - z_k(t)\|^2\, dt \leq \frac{3e^{(\rho+1)T}}{k},$$

and so, by condition (6.2), we obtain

$$\left(\int_0^T \|u_k(t) - u_*(t)\|^2\, dt \right)^{1/2}$$
$$\leq \left(\int_0^T \|u_*(t) - z_k(t)\|^2\, dt \right)^{1/2} + \left(\int_0^T \|u_k(t) - z_k(t)\|^2\, dt \right)^{1/2}$$
$$\leq \left(\frac{e^{(\rho+1)T}}{k} \right)^{1/2} + \left(\frac{3e^{(\rho+1)T}}{k} \right)^{1/2} = \left(1 + \sqrt{3}\right) \left(\frac{e^{(\rho+1)T}}{k} \right)^{1/2}.$$

Thus, for arbitrary $\varepsilon > 0$, there exists a $k_2 \geq k_1$ such that

$$\left(\int_0^T \|u_k(t) - u_*(t)\|^2\, dt\right)^{1/2} \leq \varepsilon \quad \text{for any} \quad k \geq k_2.$$

The lemma is proved. ∎

Note that Lemma 9 implies that passing to a subsequence if necessary, we can always assume that for any $T > 0$

$$u_k(t) \to u_*(t) \quad \text{for a.e.} \quad t \in [0,T] \quad \text{as} \quad k \to \infty \tag{6.8}$$

(see, for example, [**28**]). Then, by Lemma 6 for the sequence $\{x_k(\cdot)\}$, $k = 1,2,\ldots$, of the corresponding trajectories, we have

$$x_k(\cdot) \to x_*(\cdot) \quad \text{in} \quad C([0,T], R^n) \quad \text{as} \quad k \to \infty. \tag{6.9}$$

Now using Lemma 9, we prove an approximation-based version of the Pontryagin maximum principle for problem (P) in the case when system (5.1) is affine in control.

For each $k = 1,2,\ldots$ the problem (P_k) is an optimal control problem (P_{T_k}) (see (3.9)-(3.11)) with a free right end-point on the fixed time interval $[0,T_k]$ satisfying conditions of Theorem 13 (see Section 4). So, if an admissible pair $(x_k(\cdot), u_k(\cdot))$ is optimal in problem (P_k) for some $k = 1,2,\ldots$, then, due to Theorem 13, there exists a pair $(\psi_k^0, \psi_k(\cdot))$ of adjoint variables corresponding to the pair $(x_k(\cdot), u_k(\cdot))$ such that the pair $(x_k(\cdot), u_k(\cdot))$, together with $(\psi_k^0, \psi_k(\cdot))$, satisfies the core relations (4.31)–(4.33) of the Pontryagin maximum principle (for problem (P_k)), the transversality conditions (4.34) and the stationarity condition (4.35).

Recall that the adjoint variable $\psi_k(\cdot)$ is a solution, for given $(x_k(\cdot), u_k(\cdot))$, to the adjoint system for problem (P_k); i.e., on the time interval $[0,T_k]$, we have

$$\dot{\psi}_k(t) \stackrel{\text{a.e.}}{=} -\left[\frac{\partial f(x_k(t), u_k(t))}{\partial x}\right]^* \psi_k(t) - \psi^0 e^{-\rho t} \frac{\partial g(x_k(t), u_k(t))}{\partial x}; \tag{6.10}$$

the maximum condition

$$\mathcal{H}_k\big(x_k(t), t, u_k(t), \psi_k^0, \psi_k(t)\big) \stackrel{\text{a.e.}}{=} H_k\big(x_k(t), t, \psi_k^0, \psi_k(t)\big) \tag{6.11}$$

holds on $[0,T_k]$; the transversality conditions for problem (P_k) are the following:

$$\psi_k^0 > 0, \quad \psi_k(T_k) = 0 \tag{6.12}$$

and stationarity condition for problem (P_k) is

$$\frac{d}{dt} H_k\big(x_k(t), t, \psi_k^0, \psi_k(t)\big) \stackrel{\text{a.e.}}{=} \frac{\partial}{\partial t} \mathcal{H}_k\big(x_k(t), t, u_k(t), \psi_k^0, \psi_k(t)\big). \tag{6.13}$$

Here, the Hamilton–Pontryagin function $\mathcal{H}_k : G \times [0,\infty) \times U \times R^1 \times R^n \mapsto R^1$ and the Hamiltonian $H_k : G \times [0,\infty) \times R^1 \times R^n \mapsto R^1$ for problem (P_k), $k = 1, 2, \ldots$, are defined by the equalities (see Section 4)

$$\mathcal{H}_k(x,t,u,\psi^0,\psi) = \langle f(x,u),\psi\rangle + \psi^0 e^{-\rho t} g(x,u) - \psi^0 e^{-(\rho+1)t}\frac{\|u - z_k(t)\|^2}{1+\sigma_k},$$
$$H_k(x,t,\psi^0,\psi) = \sup_{u\in U} \mathcal{H}_k(x,t,u,\psi^0,\psi) \tag{6.14}$$

for any $x \in G$, $t \in [0,\infty)$, $u \in U$, $\psi^0 \in R^1$ and $\psi \in R^n$.

In what follows, we assume without loss of generality that the vector function $\psi_k(\cdot)$ is continuously extended to the infinite interval $[0,\infty)$ by zero (see (6.12)) for any $k = 1, 2, \ldots$; i.e., we set $\psi_k(t) = \psi_k(T_k) = 0$ for any $t \geq T_k$.

The following version of the Pontryagin maximum principle for problem (P) is formulated in terms of the approximating sequence $\{(P_k)\}$, $k = 1, 2, \ldots$; this gives us grounds to call it an approximation-based maximum principle.

Theorem 16. *Suppose that for arbitrary $T > 0$ conditions* (A1), (A2) *take place, and conditions* (A5), (A6) *hold as well. Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible pair in problem (P), and let $\{(P_k)\}$, $k = 1, 2, \ldots$, be an approximating sequence of optimal control problems that corresponds to the control $u_*(\cdot)$. In addition, for any $k = 1, 2, \ldots$, let $(x_k(\cdot), u_k(\cdot))$ be an optimal admissible pair in problem (P_k) and $(\psi_k^0, \psi_k(\cdot))$ be a pair of adjoint variables corresponding to the pair $(x_k(\cdot), u_k(\cdot))$ in problem (P_k); i.e., the pair $(x_k(\cdot), u_k(\cdot))$, together with the adjoint variables $(\psi_k^0, \psi_k(\cdot))$, satisfies the relations* (6.10)–(6.13) *of the Pontryagin maximum principle for problem (P_k) for any $k = 1, 2, \ldots$. Finally, let the sequences $\{\psi_k(0)\}$, $k = 1, 2, \ldots$, and $\{\psi_k^0\}$, $k = 1, 2, \ldots$, be bounded and the inequalities*

$$\psi_k^0 + \|\psi_k(0)\| \geq a, \quad k = 1, 2, \ldots, \tag{6.15}$$

hold for some $a > 0$.

Then, there exist a subsequence of the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k^0, \psi_k(\cdot))\}$, $k = 1, 2, \ldots$ (which is again denoted by $\{(x_k(\cdot), u_k(\cdot), \psi_k^0, \psi_k(\cdot))\}$ in what follows, a real number $\psi^0 \geq 0$ and a locally absolutely continuous vector function $\psi : [0,\infty) \mapsto R^n$ (adjoint variables) such that for any $T > 0$, we have:

1)

$$u_k(t) \to u_*(t) \quad \textit{for a.e.} \quad t \in [0,T] \quad \textit{as } k \to \infty, \tag{6.16}$$
$$x_k(\cdot) \to x_*(\cdot) \quad \textit{in } C([0,T],R^n) \quad \textit{as } k \to \infty, \tag{6.17}$$
$$\psi_k^0 \to \psi^0 \quad \textit{as } k \to \infty, \tag{6.18}$$
$$\psi_k(\cdot) \to \psi(\cdot) \quad \textit{in } C([0,T],R^n) \quad \textit{as } k \to \infty; \tag{6.19}$$

2) *the vector function* $\psi(\cdot)$ *is a (Carathéodory) solution to the following adjoint system on* $[0,\infty)$*:*

$$\dot{\psi} = -\frac{\partial \mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi)}{\partial x} = -\left[\frac{\partial f(x_*(t),t,u_*(t))}{\partial x}\right]^* \psi - \psi^0 \frac{\partial f^0(x_*(t),t,u_*(t))}{\partial x}; \qquad (6.20)$$

3) *The following maximum condition takes place on* $[0,\infty)$*:*

$$\mathcal{H}(x_*(t),t,u_*(t),\psi^0,\psi(t)) \stackrel{a.e.}{=} H(x_*(t),t,\psi^0,\psi(t)); \qquad (6.21)$$

4) *the pair* $(\psi^0,\psi(\cdot))$ *is nontrivial, i.e.,*

$$\psi^0 + \|\psi(0)\| > 0. \qquad (6.22)$$

5) *the following stationarity condition holds:*

$$H(x_*(t),t,\psi^0,\psi(t)) = \psi^0 \rho \int_t^\infty e^{-\rho s} g(x_*(s),u_*(s))\,ds \quad \text{for any} \quad t \geq 0. \qquad (6.23)$$

Here, the Hamilton–Pontryagin function $\mathcal{H} : G\times[0,\infty)\times U\times R^1\times R^n \mapsto R^1$ and the Hamiltonian $H : G \times [0,\infty) \times R^1 \times R^n \mapsto R^1$ for problem (P) are defined by the equalities (see Section 4)

$$\mathcal{H}(x,t,u,\psi^0,\psi) = \langle f(x,u),\psi\rangle + \psi^0 e^{-\rho t} g(x,u),$$
$$H(x,t,\psi^0,\psi) = \sup_{u\in U} \mathcal{H}(x,t,u,\psi^0,\psi)$$

for any $x \in G$, $t \in [0,\infty)$, $u \in U$, $\psi \in R^n$, and $\psi^0 \in R^1$.

Proof. By Lemma 9, passing to a subsequence if necessary, we obtain conditions (6.16) and (6.17) for an arbitrary $T > 0$ (see (6.8) and (6.9)).

Next, by the hypotheses of the theorem, the sequence $\{\psi_k^0\}$, $k = 1,2,\dots$, is bounded. Hence, passing to a subsequence if necessary, we can assume without loss of generality that condition (6.18) holds for some $\psi^0 \geq 0$.

Now, we extract a subsequence from the sequence $\{\psi_k(\cdot)\}$, $k = 1,2,\dots$, such that condition (6.19) holds for any $T > 0$, where $(\psi^0,\psi(\cdot))$ is a nonzero pair of adjoint variables corresponding to the pair $(x_*(\cdot),u_*(\cdot))$.

Consider the sequence of absolutely continuous functions $\{\psi_k(\cdot)\}$, $k = 1,2,\dots$, on the time interval $[0,T_1]$. By virtue of the adjoint system (6.10), the boundedness of the sequence $\{\psi_k(0)\}$, $k = 1,2,\dots$ (see the hypotheses of the theorem), and Lemma 1, this sequence is uniformly bounded on $[0,T_1]$. Moreover, the sequence of derivatives $\{\dot{\psi}_k(\cdot)\}$, $k = 1,2,\dots$, is also uniformly bounded on $[0,T_1]$. Therefore, by the Arzelà theorem, passing to a subsequence (which is denoted below by $\{\psi_{1,k}(\cdot)\}$, $k = 1,2,\dots$) if necessary, we obtain

$$\psi_{1,k}(\cdot) \to \psi(\cdot) \quad \text{in} \quad C([0,T_1],R^n) \quad \text{as} \quad k \to \infty,$$

where $\psi : [0, T_1] \mapsto R^n$ is an absolutely continuous function.

Now, consider the sequence $\{\psi_{1,k}(\cdot)\}$, $k = 1, 2, \dots$, on the time interval $[0, T_2]$. By analogy with the previous case, passing, if necessary, to a subsequence $\{\psi_{2,k}(\cdot)\}$, $k = 1, 2, \dots$, in the sequence $\{\psi_{1,k}(\cdot)\}$, $k = 1, 2, \dots$, we obtain

$$\psi_{2,k}(\cdot) \to \tilde{\psi}(\cdot) \quad \text{in} \quad C([0, T_2], R^n) \quad \text{as} \quad k \to \infty,$$

where $\tilde{\psi} : [0, T_2] \mapsto R^n$ is an absolutely continuous function, which obviously coincides with the function $\psi(\cdot)$ on $[0, T_1]$. For short, we will again denote it by $\psi(\cdot)$.

Successively applying this procedure on the time intervals $[0, T_i]$, $i = 3, 4, \dots$, we find that there exist a locally absolutely continuous function $\psi : [0, \infty) \mapsto R^n$ and subsequences of absolutely continuous functions $\psi_{i,k} : [0, T_i] \mapsto R^n$, $i, k = 1, 2, \dots$, such that the sequence $\{\psi_{i+1,k}(\cdot)\}$, $k = 1, 2, \dots$, is a subsequence of the sequence $\{\psi_{i,k}(\cdot)\}$, $k = 1, 2, \dots$, for any $i = 1, 2, \dots$, and

$$\psi_{i,k}(\cdot) \to \psi(\cdot) \quad \text{in} \quad C([0, T_i], R^n) \quad \text{as} \quad k \to \infty$$

for any $i = 1, 2, \dots$.

Since

$$T_i \to \infty \quad \text{as} \quad i \to \infty,$$

we obtain

$$\psi_{k,k}(\cdot) \to \psi(\cdot) \quad \text{in} \quad C([0, T], R^n) \quad \text{as} \quad k \to \infty$$

for any $T > 0$. For simplicity, in what follows we will again denote the sequence $\{\psi_{k,k}(\cdot)\}$, $k = 1, 2, \dots$, by $\{\psi_k(\cdot)\}$, $k = 1, 2, \dots$.

It follows immediately from the construction of the sequence $\{\psi_k(\cdot)\}$, $k = 1, 2, \dots$, that condition (6.19) holds for any $T > 0$.

Next, for any $t \geq 0$ and all sufficiently large numbers k, by virtue of the adjoint system (6.10) we have

$$\psi_k(t) = \psi_k(0) - \int_0^t \left(\left[\frac{\partial f(x_k(s), u_k(s))}{\partial x} \right]^* \psi_k(s) + \psi_k^0 e^{-\rho s} \frac{\partial g(x_k(s), u_k(s))}{\partial x} \right) ds. \tag{6.24}$$

The integrands in (6.24) are uniformly bounded on $[0, t]$, and in view of conditions (6.16)–(6.19) proved above, we have

$$\lim_{k \to \infty} \left(\left[\frac{\partial f(x_k(s), u_k(s))}{\partial x} \right]^* \psi_k(s) + \psi_k^0 e^{-\rho s} \frac{\partial g(x_k(s), u_k(s))}{\partial x} \right)$$
$$= \left[\frac{\partial f(x_*(s), u_*(s))}{\partial x} \right]^* \psi(s) + \psi_k^0 e^{-\rho s} \frac{\partial g(x_*(s), u_*(s))}{\partial x}$$

for a.e. $s \in [0, t]$. Since $\psi_k(0) \to \psi(0)$ as $k \to \infty$, we can apply the Lebesgue theorem (see [**34**]) and pass to the limit as $k \to \infty$ under the sign of integral in (6.24). Thus, we obtain

$$\psi(t) = \psi(0) - \int_0^t \left(\left[\frac{\partial f(x_*(s), u_*(s))}{\partial x} \right]^* \psi_*(s) + \psi_*^0 e^{-\rho s} \frac{\partial g(x_*(s), u_*(s))}{\partial x} \right) ds.$$

Therefore, the locally absolutely continuous vector function $\psi(\cdot)$ satisfies the adjoint system (6.20) on the infinite time interval $[0, \infty)$.

Now, consider the maximum condition (6.11) on the interval $[0, T_k]$ for $k = 1, 2, \dots$:

$$\langle f(x_k(t), u_k(t)), \psi_k(t) \rangle + \psi_k^0 e^{-\rho t} g(x_k(t), u_k(t)) - \psi_k^0 e^{-(\rho+1)t} \frac{\|u_k(t) - z_k(t)\|^2}{1 + \sigma_k}$$

$$\stackrel{\text{a.e.}}{=} \max_{u \in U} \left[\langle f(x_k(t), u), \psi_k(t) \rangle + \psi_k^0 e^{-\rho t} g(x_k(t), u) - \psi_k^0 e^{-(\rho+1)t} \frac{\|u - z_k(t)\|^2}{1 + \sigma_k} \right].$$

Since $T_k \to \infty$ and $\sigma_k \to \infty$ as $k \to \infty$, it follows from conditions (5.7), (6.16)-(6.19) (which hold for any $T > 0$) that condition (6.11) yields the maximum condition (6.21) at the limit.

The pair $(\psi^0, \psi(\cdot))$ is nontrivial (see condition (6.22)) by condition (6.15).

Thus, the optimal pair $(x_*(\cdot), u_*(\cdot))$, together with the pair of Lagrange multipliers $(\psi^0, \psi(\cdot))$, satisfies the core relations of the Pontryagin maximum principle (6.20)–(6.22); i.e., the pair $(\psi^0, \psi(\cdot))$ is a pair of adjoint variables that corresponds to the optimal pair $(x_*(\cdot), u_*(\cdot))$.

Consider condition (6.13). The definition of the function $\mathcal{H}_k(\cdot, \cdot, \cdot, \cdot, \cdot)$ (see (6.14)) on the time interval $[0, T_k]$ implies the equality

$$\frac{d}{dt} H_k\big(x_k(t), t, \psi_k(t), \psi_k^0\big) \stackrel{\text{a.e.}}{=} \frac{\partial}{\partial t} \mathcal{H}_k\big(x_k(t), t, u_k(t), \psi_k(t), \psi_k^0\big)$$

$$= -\psi_k^0 \rho e^{-\rho t} g(x_k(t), u_k(t)) + \psi_k^0 (\rho + 1) e^{-(\rho+1)t} \frac{\|u_k(t) - z_k(t)\|^2}{1 + \sigma_k}$$

$$+ 2\psi_k^0 e^{-(\rho+1)t} \frac{\langle u_k(t) - z_k(t), \dot{z}_k(t) \rangle}{1 + \sigma_k}.$$

Take arbitrary $t > 0$ and k such that $T_k > t$, and integrate this equality over the time interval $[t, T_k]$. Then, taking account of the boundary condition (6.12), we obtain

$$H_k\big(x_k(t), t, \psi_k(t), \psi_k^0\big) = \psi_k^0 e^{-\rho T_k} \max_{u \in U} \left[g(x_k(T_k), u) - e^{-T_k} \frac{\|u - z_k(T_k)\|^2}{1 + \sigma_k} \right]$$

$$-\psi_k^0 \rho \int_t^{T_k} e^{-\rho s} g(x_k(s), u_k(s))\, ds + \psi_k^0 (\rho + 1) \int_t^{T_k} e^{-(\rho+1)s} \frac{\|u_k(s) - z_k(s)\|^2}{1 + \sigma_k}\, ds$$

$$+ 2\psi_k^0 \int_t^{T_k} e^{-(\rho+1)s} \frac{\langle u_k(s) - z_k(s), \dot{z}_k(s) \rangle}{1 + \sigma_k}\, ds.$$

Using relations (6.3) and (6.4), the already proved conditions (6.16)-(6.19) (which hold for any $T > 0$), and inequalities (6.1)-(6.3), we can pass to the limit in this equality as $k \to \infty$. At the limit, we obtain condition (6.23). The proof of the theorem is complete. ∎

To conclude this section we formulate a variant of Theorem 16 in the case when relations of the approximation-based maximum principle for problems (P_k) have the normal form, i.e., $\psi_k^0 = 1$ for any $k = 1, 2, \ldots$. We will use this formulation later in Sections 9 and 10 for deriving different normal-form versions of the maximum principle for problem (P) when the control system (5.1) is affine in control.

We will use the following formulation of the normal-form maximum principle for problem (P_k).

Let $(x_k(\cdot), u_k(\cdot))$ be an optimal admissible pair in problem (P_k) for some natural $k \in \{1, 2, \ldots\}$. Then, there exists an adjoint variable $\psi_k(\cdot)$ corresponding to the pair $(x_k(\cdot), u_k(\cdot))$, such that the pair $(x_k(\cdot), u_k(\cdot))$, together with the adjoint variable $\psi_k(\cdot)$, satisfies the normal-form maximum principle (for problem (P_k)) and the transversality condition (6.12) holds.

Recall that, for given $(x_k(\cdot), u_k(\cdot))$, $\psi_k(\cdot)$ is a solution to the normal-form adjoint system; i.e., on the time interval $[0, T_k]$, we have

$$\dot{\psi}_k(t) \stackrel{\text{a.e.}}{=} -\left[\frac{\partial f(x_k(t), u_k(t))}{\partial x}\right]^* \psi_k(t) - e^{-\rho t}\frac{\partial g(x_k(t), u_k(t))}{\partial x}. \qquad (6.25)$$

Further, the pair $(x_k(\cdot), u_k(\cdot))$, together with the adjoint variable $\psi_k(\cdot)$, satisfies the maximum condition in the normal form on the interval $[0, T_k]$; i.e.,

$$\mathcal{H}_k(x_k(t), t, u_k(t), \psi(t)) \stackrel{\text{a.e.}}{=} H_k(x_k(t), t, \psi_k(t)). \qquad (6.26)$$

Here, the Hamilton–Pontryagin function $\mathcal{H}_k : G \times [0, \infty) \times U \times R^n \mapsto R^1$ and the Hamiltonian $H_k : G \times [0, \infty) \times R^n \mapsto R^1$ in the normal form for problem (P_k) are defined in the standard way (see Section 4):

$$\mathcal{H}_k(x, t, u, \psi) = \langle f(x, u), \psi\rangle + e^{-\rho t} g(x, u) - e^{-(\rho+1)t}\frac{\|u - z_k(t)\|^2}{1 + \sigma_k},$$

$$H_k(x, t, \psi) = \sup_{u \in U} \mathcal{H}_k(x, t, u, \psi)$$

for any $x \in G$, $t \in [0, \infty)$, $u \in U$, and $\psi \in R^n$. Moreover, the following transversality condition

$$\psi_k(T_k) = 0 \qquad (6.27)$$

takes place.

As above, we assume that the vector function $\psi_k(\cdot)$ is continuously extended to the infinite interval $[0, \infty)$ by zero (see (6.27)) for any $k = 1, 2, \ldots$; i.e., we set $\psi_k(t) = \psi_k(T_k) = 0$ for any $t \geq T_k$.

It is easy to see that the following normal-form approximation-based version of the Pontryagin maximum principle for problem (P) is a corollary of Theorem 16.

Theorem 17. *Suppose that for arbitrary $T > 0$ conditions* (A1), (A2) *take place, and conditions* (A5), (A6) *hold as well. Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible pair in problem (P), let $\{(P_k)\}$, $k = 1, 2, \dots$, be a sequence of approximating problems that corresponds to the optimal control $u_*(\cdot)$, and let $(x_k(\cdot), u_k(\cdot))$ be an optimal pair in problem (P_k). Moreover, for any $k = 1, 2, \dots$, let $\psi_k(\cdot)$ be an adjoint variable corresponding to the pair $(x_k(\cdot), u_k(\cdot))$ in problem (P_k); i.e., $\psi_k(\cdot)$ and $(x_k(\cdot), u_k(\cdot))$ satisfy conditions* (6.25) *and* (6.26) *of the normal-form maximum principle for problem (P_k) on the time interval $[0, T_k]$ and the transversality condition* (6.27) *holds. Finally, let the sequence $\{\psi_k(0)\}$, $k = 1, 2, \dots$, be bounded. Then, there exists a subsequence of $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \dots$ (which we again denote by $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$), such that for any $T > 0$*

1)

$$u_k(t) \to u_*(t) \quad \text{for a.e.} \quad t \in [0, T] \quad \text{as } k \to \infty, \tag{6.28}$$

$$x_k(\cdot) \to x_*(\cdot) \quad \text{in} \quad C([0, T], R^n) \quad \text{as} \quad k \to \infty, \tag{6.29}$$

$$\psi_k(\cdot) \to \psi(\cdot) \quad \text{in} \quad C([0, T], R^n) \quad \text{as} \quad k \to \infty; \tag{6.30}$$

2) *the vector function $\psi(\cdot)$ is a (Carathéodory) solution to the following normal form adjoint system on $[0, \infty)$:*

$$\begin{aligned} \dot{\psi} &= -\frac{\partial \mathcal{H}(x_*(t), t, u_*(t), \psi, \psi^0)}{\partial x} \\ &= -\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x}\right]^* \psi - \frac{\partial f^0(x_*(t), t, u_*(t))}{\partial x}; \end{aligned} \tag{6.31}$$

3) *the following normal-form maximum condition takes place on $[0, \infty)$:*

$$\mathcal{H}(x_*(t), t, u_*(t), \psi(t)) \stackrel{a.e.}{=} H(x_*(t), t, \psi(t)); \tag{6.32}$$

4) *the following normal-form stationarity condition holds*:

$$H(x_*(t), t, \psi(t)) = \rho \int_t^\infty e^{-\rho s} g(x_*(s), u_*(s))\, ds \quad \text{for any} \quad t \geq 0. \tag{6.33}$$

Here, the normal form Hamilton–Pontryagin function $\mathcal{H} : G \times [0, \infty) \times U \times R^n \mapsto R^1$ and the normal form Hamiltonian $H : G \times [0, \infty) \times R^n \mapsto R^1$ for problem (P) are defined in the standard way (see Section 4):

$$\mathcal{H}(x, t, u, \psi) = \langle f(x, u), \psi \rangle + e^{-\rho t} g(x, u),$$

$$H(x, t, \psi) = \sup_{u \in U} \mathcal{H}(x, t, u, \psi)$$

for any $x \in G$, $t \in [0, \infty)$, $u \in U$, and $\psi \in R^n$.

7. Maximum principle and transversality conditions for optimal economic growth problems

In this section, using the approximation-based maximum principle (see Theorem 16 in Section 6), we develop a general version of the Pontryagin maximum principle for problem (P) in the case when system (5.1) is affine in control. Then we consider relations of the maximum principle in more details. In particular, we discuss their compatibility with transversality conditions at infinity often used in economic applications.

Now we derive a general version of the Pontryagin maximum principle for problem (P) as a corollary of Theorem 16.

Theorem 18. *Suppose that for arbitrary $T > 0$ conditions* (A1), (A2) *take place, and conditions* (A5), (A6) *hold as well. Let $(x_*(\cdot), u_*(\cdot))$ be an optimal pair in problem (P). Then there are a real number $\psi^0 \geq 0$ and a locally absolutely continuous vector function $\psi : [0,\infty) \mapsto R^n$ (adjoint variables) such that we have:*

1) *the vector function $\psi(\cdot)$ is a (Carathéodory) solution to the following adjoint system on $[0,\infty)$:*

$$\begin{aligned}\dot{\psi} &= -\frac{\partial \mathcal{H}(x_*(t), t, u_*(t), \psi^0, \psi)}{\partial x} \\ &= -\left[\frac{\partial f(x_*(t), t, u_*(t))}{\partial x}\right]^* \psi - \psi^0 \frac{\partial f^0(x_*(t), t, u_*(t))}{\partial x};\end{aligned} \tag{7.1}$$

2) *the following maximum condition takes place on $[0,\infty)$:*

$$\mathcal{H}(x_*(t), t, u_*(t), \psi^0, \psi(t)) \stackrel{a.e.}{=} H(x_*(t), t, \psi^0, \psi(t)); \tag{7.2}$$

3) *the pair $(\psi^0, \psi(\cdot))$ is nontrivial, i.e.,*

$$\psi^0 + \|\psi(0)\| > 0; \tag{7.3}$$

4) *the following stationarity condition holds:*

$$H(x_*(t), t, \psi^0, \psi(t)) = \psi^0 \rho \int_t^\infty e^{-\rho s} g(x_*(s), u_*(s))\, ds \quad \textit{for any} \quad t \geq 0. \tag{7.4}$$

Proof. Let $\{(P_k)\}$, $k = 1, 2, \ldots$, be a sequence of approximating problems that correspond to the optimal control $u_*(\cdot)$ (see Section 6). Next, let $(x_k(\cdot), u_k(\cdot))$ be an optimal admissible pair in problem (P_k) (see (6.5)–(6.7)). For any $k = 1, 2, \ldots$, according to the Pontryagin maximum principle for non autonomous problems (see Theorem 13 in Section 4), there exists a pair of adjoint variables $(\psi_k^0, \psi_k(\cdot))$ that correspond to the optimal pair $(x_k(\cdot), u_k(\cdot))$ in problem (P_k), i.e., such that the pair $(x_k(\cdot), u_k(\cdot))$, together with adjoint variables $(\psi_k^0, \psi_k(\cdot))$, satisfies the relations (6.10)–(6.13) of the maximum principle for any $k = 1, 2, \dots$.

Since $\psi_k^0 > 0$, the quantity $c_k = \|\psi_k(0)\| + \psi_k^0$ is positive. Let us pass to the normalized adjoint variables $\psi_k(\cdot)/c_k$ and ψ_k^0/c_k in the relations of the maximum principle for problem (P_k) and retain the previous notations $\psi_k(\cdot)$ and ψ_k^0 for them. Then, the normalized adjoint variables satisfy the nontriviality condition

$$\psi_k^0 + \|\psi_k(0)\| = 1 \tag{7.5}$$

and the transversality condition (6.12); moreover, it is obvious that the pair $(x_k(\cdot), u_k(\cdot))$, together with the normalized adjoint variables $(\psi_k^0, \psi_k(\cdot))$, satisfies the core relations of the Pontryagin maximum principle (for problem (P_k)), $k = 1, 2, \dots$. By the nontriviality condition (7.5), the sequences $\{\psi_k^0\}$, $k = 1, 2, \dots$, and $\{\psi_k(0)\}$, $k = 1, 2, \dots$, are bounded and the condition (6.15) holds with $a = 1$. Hence, the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k^0, \psi_k(\cdot)\}$, $k = 1, 2, \dots$, satisfies all the hypotheses of Theorem 16. By Theorem 16, there exists a subsequence of the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k^0, \psi_k(\cdot))\}$, $k = 1, 2, \dots$ (which we again denote by $\{(x_k(\cdot), u_k(\cdot), \psi_k^0, \psi_k(\cdot))\}$, $k = 1, 2, \dots$), such that conditions (6.16), (6.17) are valid and the sequence of adjoint variables $\{(\psi_k^0, \psi_k(\cdot))\}$, $k = 1, 2, \dots$, satisfy conditions (6.18), (6.19) for any $T > 0$. The limit of this sequence is a nontrivial pair of adjoint variables $(\psi^0, \psi(\cdot))$ that correspond to the optimal pair $(x_*(\cdot), u_*(\cdot))$ in problem (P). Thus, the pair $(x_*(\cdot), u_*(\cdot))$, together with the pair of adjoint variables $(\psi^0, \psi(\cdot))$, satisfies the core relations (6.20)–(6.22) of the Pontryagin maximum principle for problem(P); finally, the pair $(x_*(\cdot), u_*(\cdot))$, together with the adjoint variables $(\psi^0, \psi(\cdot))$, satisfies the stationarity condition (6.23). The theorem is proved. ■

Note that the formulation of Theorem 18 is weaker than that of the corresponding classical result for a finite-horizon optimal control problem with a free right endpoint considered above in Section 4 (see Theorem 13).

Indeed, due to Theorem 13 the adjoint variables $\psi(\cdot)$ and ψ^0 can be chosen such that they satisfy the following transversality conditions at the right endpoint (see (4.34)):

$$\psi^0 = 1, \quad \psi(T) = 0. \tag{7.6}$$

We can see that Theorem 18 does not contain any analogs of the transversality conditions (7.6).

The transversality conditions (7.6) contain important information about the pair of adjoint variables $(\psi^0, \psi(\cdot))$ corresponding, by virtue of the Pontryagin maximum principle, to the optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem (P_T). Without the transversality conditions (7.6), the relations of the maximum principle (Theorem 13) (4.31) and (4.32) (together with the nontriviality condition (4.33), and stationarity condition (4.35) are incomplete: they define a very large set of extremals.

The transversality conditions (7.6) are important because, first, they guarantee the normality of problem (P_T) and, second, they uniquely define the adjoint variable $\psi(\cdot)$ as a solution to the system of linear differential equations (4.31) on $[0,T]$ with the boundary condition $\psi(T)=0$.

Let us explain the last fact in more details.

On a finite time interval $[0,T]$, $T>0$, consider the inhomogeneous system of linear differential equations

$$\dot{y} = A(t)y + b(t) \tag{7.7}$$

with initial condition

$$y(\xi) = y_\xi. \tag{7.8}$$

Here, $y \in R^n$; $A(\cdot)$ and $b(\cdot)$ are an $n\times n$ matrix function and an n-dimensional vector function, respectively, defined on $[0,T]$; and y_ξ is a given state of the system at moment $\xi \in [0,T]$. Suppose that all the components of the matrix function $A(\cdot)$ and all the coordinates of the vector function $b(\cdot)$ are measurable and bounded. Then, there exists a unique solution $y(\cdot)$ (in the sense of Carathéodory) of system (7.7) on $[0,T]$ with the initial condition (7.8); moreover, the following Cauchy formula holds for any $t \in [0,T]$ (see, for example, [**17**], [**24**]):

$$y(t) = Y(t)\Big[[Y(\xi)]^{-1}y_\xi + \int_\xi^t [Y(s)]^{-1}b(s)\,ds\Big]. \tag{7.9}$$

Here, $Y(\cdot)$ is a normalized fundamental matrix of the linear homogeneous system

$$\dot{y} = A(t)y. \tag{7.10}$$

More precisely, $Y(\cdot)$ is an $n\times n$ matrix function defined on $[0,T]$ whose columns $y_i(\cdot)$, $i=1,\dots,n$, are (linearly independent) solutions of system (7.10) on $[0,T]$ that satisfy the initial conditions

$$y_i^j(0) = \delta_{i,j}, \quad i,j=1,\dots,n,$$

where

$$\delta_{i,i}=1, \quad i=1,\dots,n, \quad \text{and} \quad \delta_{i,j}=0 \quad \text{for} \quad i\neq j, \;\; i,j=1,\dots,n.$$

Now, consider the adjoint system (4.31), which is included in the relations of the Pontryagin maximum principle for problem (P_T) (Theorem 13). This is a linear inhomogeneous system of differential equations on the finite time interval $[0,T]$. All the components of the matrix function

$$A(t) = -\left[\frac{\partial f(x_*(t),u_*(t))}{\partial x}\right]^*, \quad t\in[0,T],$$

and the coordinates of the vector function

$$b(t) = -e^{-\rho t}\frac{\partial g(x_*(t), u_*(t))}{\partial x}, \quad t \in [0, T],$$

are measurable and bounded (here, we took into account that $\psi^0 = 1$ in view of conditions (7.6)). Moreover, the vector function $\psi(\cdot)$, as an adjoint variable corresponding to the optimal pair $(x_*(\cdot), u_*(\cdot))$ by virtue of the Pontryagin maximum principle, satisfies the equality $\psi(T) = 0$ (see (7.6)).

Let $Z_*(\cdot)$ be the normalized fundamental matrix of the linear homogeneous system

$$\dot{z}(t) = -\left[\frac{\partial f(x_*(t), u_*(t))}{\partial x}\right]^* z(t). \tag{7.11}$$

Thus, the columns z_i, $i = 1, \dots, n$, of the $n \times n$ matrix function Z_* are (linearly independent) solutions of system (7.11) on $[0, T]$ that satisfy the initial conditions

$$z_i^j(0) = \delta_{i,j}, \quad i, j = 1, \dots, n.$$

Denote by $I_*^T(\cdot)$ the following vector function defined on the interval $[0, T]$ and taking values in R^n:

$$I_*^T(t) = \int_t^T e^{-\rho s}[Z_*(s)]^{-1}\frac{\partial g(x_*(s), u_*(s))}{\partial x}\, ds \quad \text{for any} \quad t \in [0, T].$$

In terms of the matrix function $Z_*(\cdot)$ and the vector function $I_*^T(\cdot)$ (for $\xi = T$ and in view of the equality $\psi(T) = 0$), the Cauchy formula (7.9) is expressed as

$$\psi(t) = Z_*(t) I_*^T(t) \quad \text{for any} \quad t \in [0, T]. \tag{7.12}$$

So, due to conditions of the Pontryagin maximum principle for finite-horizon optimal control problem (P_T)(see Theorem 13) formula (7.12) uniquely defines the adjoint variable $\psi(\cdot)$ on the corresponding time-interval $[0, T]$.

For the optimal control problem (P) on the infinite time interval $[0, \infty)$, the question of deriving additional conditions that characterize the behavior of the adjoint variable at infinity has been considered by many authors (see, for example, [**7**], [**10**], [**18**], [**30**], [**36**] and [**39**]). A comprehensive survey of results in this direction is given in [**6**].

As a rule, one suggests using the following "natural" generalizations of conditions (7.6) as the transversality conditions at infinity:

$$\psi^0 = 1, \quad \lim_{t\to\infty} \psi(t) = 0 \tag{7.13}$$

or

$$\psi^0 = 1, \quad \lim_{t\to\infty} \langle x_*(t), \psi(t)\rangle = 0. \tag{7.14}$$

Note that using conditions (7.14) as necessary optimality conditions is sometimes motivated by the fact that these conditions are contained in the well-known (based on the relations of the maximum principle) sufficient optimality conditions for problem (P) (for more detail on the sufficient optimality conditions for problem (P), see Section 11).

In the general case, the standard transversality conditions (7.13) and (7.14) do not hold for the optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ and a pair of adjoint variables $(\psi^0, \psi(\cdot))$ corresponding to $(x_*(\cdot), u_*(\cdot))$ in problem (P). Moreover, both conditions in (7.13), as well as both conditions in (7.14), may be violated simultaneously and in various combinations.

For the case $\rho = 0$, examples of such "pathological" behavior of adjoint variables were constructed in [**23**], [**38**]. For the case $\rho > 0$, an example of an abnormal optimal admissible pair (i.e., such that necessarily $\psi^0 = 0$) in problem (P) was constructed in [**30**]. Further examples of pathological behavior of adjoint variables at infinity in the case $\rho > 0$ are given in [**6**].

Let us consider two examples illustrating incompatibility (in the general case) of relations of the maximum principle (7.1)–(7.4) with transversality conditions (7.13) and (7.14) in the situation $\rho > 0$.

First consider a little bit modified version of example presented in [**30**].

Example 6. Consider the following optimal control problem:

$$\dot{x} = (2x + u))\phi(x), \quad u \in U = [-1, 0], \tag{7.15}$$

$$x(0) = 0, \tag{7.16}$$

$$J(x, u) = \int_0^\infty e^{-t}(2x + u)\,dt \to \max. \tag{7.17}$$

Here, $\phi : R^1 \mapsto R^1$ is a smooth nonnegative bounded function such that $\phi(x) = 1$ for $|x| \leq 1$ and $\phi(x) = 0$ for $|x| \geq 2$.

Let $G = R^1$. It is clear that problem (7.15)–(7.17) satisfies conditions (A1), (A2) (on any finite time interval $[0, T]$) and conditions (A5), (A5) as well; i.e., it is a particular case of problem (P). The only optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem (7.15)–(7.17) has the form

$$x_*(t) = 0 \quad \text{for all } \ t \geq 0, \quad u_*(t) \overset{\text{a.e.}}{=} 0.$$

Indeed, any admissible control u different from $u_*(\cdot)$ takes negative values on a set of positive measure. Since the corresponding admissible trajectory $x(\cdot)$ satisfies the inequality $x(t) \leq 0$ for any $t \geq 0$ (see (7.15) and (7.16)), this immediately gives a negative value of the goal functional $J(x(\cdot), u(\cdot))$ (see (7.17)) at the pair $(x(\cdot), u(\cdot))$, whereas $J(x_*(\cdot), u_*(\cdot)) = 0$.

By Theorem 18, the optimal pair $(x_*(\cdot), u_*(\cdot))$, together with a pair of adjoint variables $(\psi^0, \psi(\cdot))$ corresponding to $(x_*(\cdot), u_*(\cdot))$, satisfies the relations (7.1)–(7.4) of the Pontryagin maximum principle.

Let us show that for any pair of Lagrange multipliers $(\psi^0, \psi(\cdot))$ corresponding to the optimal pair $(x_*(\cdot), u_*(\cdot))$, the equality $\psi^0 = 0$ follows from the Pontryagin maximum principle (Theorem 18).

The Hamilton–Pontryagin function for problem (7.15)–(7.17) has the form
$\mathcal{H}(x,t,u,\psi,\psi^0) = \psi(2x+u)\phi(x) + \psi^0 e^{-t}(2x+u) = \big(\psi\phi(x) + \psi^0 e^{-t}\big)(2x+u)$.
Hence, the variable $\psi(\cdot)$ is a solution to the following adjoint system (see (7.1)):

$$\dot{\psi} = -2\big(\psi + \psi^0 e^{-t}\big). \tag{7.18}$$

Here, we took into account that

$$\phi(x_*(t)) = 1 \quad \text{and} \quad \frac{d}{dx}\phi(x_*(t)) = 0 \quad \text{for any} \quad t \geq 0$$

according to the definition of the function $\phi(\cdot)$. It is easy to see that the maximum condition (7.2) in this case implies the inequality

$$\psi(t) + \psi^0 e^{-t} \geq 0 \quad \text{for any} \quad t \geq 0. \tag{7.19}$$

Solving the adjoint equation (7.18), we obtain

$$\psi(t) = -2\psi^0 e^{-t} + (\psi(0) + 2\psi^0)e^{-2t} \quad \text{for any} \quad t \geq 0.$$

This implies that if $\psi^0 > 0$, then

$$\psi(t) + \psi^0 e^{-t} = -\psi^0 e^{-t} + (\psi(0) + 2\psi^0)e^{-2t} < 0$$

for all sufficiently large t, which contradicts condition (7.19) and, hence, the maximum condition (7.2).

Thus, the equality $\psi^0 = 0$ necessarily holds in the relations of the Pontryagin maximum principle for problem (7.15)–(7.17).

Now, consider the limit relations at infinity in the transversality conditions (7.13) and (7.14). In the general case, these relations are different; they do not follow from each other or from the relations of the maximum principle (7.1)–(7.4).

In the following example (see [**6**]), irrespective of the value of the adjoint variable ψ^0, none of the limit relations at infinity in (7.13) and (7.14) can hold together with the conditions of the maximum principle (7.1)–(7.4) for problem (P).

Example 7. Consider the following optimal control problem:

$$\dot{x} = u - x, \qquad u \in U = [1/2, 1], \tag{7.20}$$

$$x(0) = 1, \tag{7.21}$$

$$J(x,u) = \int_0^\infty e^{-t} \ln \frac{1}{x - 1/2}\, dt \to \max. \tag{7.22}$$

Set $G = (1/2, \infty)$. Obviously, conditions (A1), (A2) (on any finite time interval $[0,T]$, $T > 0$) are satisfied and (A5), (A6) hold in this case. The only optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem (7.20)–(7.22) has the form

$$x_*(t) = \frac{1+e^{-t}}{2} \quad \text{for any} \quad t \geq 0, \qquad u_*(t) \overset{\text{a.e.}}{=} \frac{1}{2}.$$

Let $(\psi^0, \psi(\cdot))$ be a pair of adjoint variables that satisfy, together with the optimal pair $(x_*(\cdot), u_*(\cdot))$, the core relations of the Pontryagin maximum principle by virtue of Theorem 18. The Hamilton–Pontryagin function for problem (7.20)–(7.22) has the form

$$\mathcal{H}(x, t, u, \psi, \psi^0) = \psi(u - x) - \psi^0 e^{-t} \ln(x - 1/2).$$

Hence, the variable $\psi(\cdot)$ is a solution to the following adjoint system (see (7.1)):

$$\dot{\psi} = \psi + 2\psi^0. \tag{7.23}$$

Here we used the fact that

$$\frac{d}{dx} \ln(x_*(t) - 1/2) = \frac{1}{x_*(t) - 1/2} = 2e^t.$$

It is obvious that the maximum condition (7.2) in this case implies the inequality

$$\psi(t) \leq 0 \quad \text{for any} \quad t \geq 0.$$

Let us check whether the limit relations at infinity in the transversality conditions (7.13) and (7.14) can hold.

Solving the adjoint equation (7.23), we obtain

$$\psi(t) = (\psi(0) + 2\psi^0)e^t - 2\psi^0.$$

If $\psi^0 = 0$, then $\psi(t) = \psi(0)e^t$. In this case, by the maximum condition (7.2) and the nontriviality condition (7.3), we have $\psi(0) < 0$. Hence,

$$\psi(t) \to -\infty \quad \text{and} \quad x_*(t)\psi(t) = \psi(t)\frac{1+e^{-t}}{2} \to -\infty \quad \text{as} \quad t \to \infty,$$

which contradicts both the limit condition in (7.13) and that in (7.14).

Let $\psi^0 > 0$. In this case, without loss of generality (or multiplying both adjoint variables ψ and ψ^0 by $1/\psi^0$), we may assume that $\psi^0 = 1$. Then,

$$\psi(t) = (\psi(0) + 2)e^t - 2.$$

If $\psi(0) \neq -2$, then $\lim_{t\to\infty} \psi(t) = -\infty$, while if $\psi(0) = -2$, then $\psi(t) = -2$ for any $t \geq 0$. In both cases, the limit relation in (7.13) is violated. Since $\lim_{t\to\infty} x_*(t) = 1/2$, it follows that $\lim_{t\to\infty} x_*(t)\psi(t) \neq 0$ in both cases. Thus, the limit condition in (7.14) is also violated.

The examples considered demonstrate a qualitative difference between the infinite-horizon optimal control problem (P) and an analogous problem (P_T) on a finite time interval (see Section 4). In particular, in the case of problem (P) the natural transversality conditions (7.13) and (7.14) may not be compatible with the relations (7.1)–(7.4) of the Pontryagin maximum principle.

The reason of this phenomenon is related to the fact that the control process of system (5.1) in problem (P) is considered on an infinite time interval. The presence of an infinite horizon gives rise to a singularity in problem (P).

To illustrate this fact, we assume that $\rho > 0$ and change time in problem (P) as follows:

$$\tau(t) = 1 - e^{-\rho t}, \qquad t \in [0, \infty). \tag{7.24}$$

The continuously differentiable function $\tau : [0, \infty) \mapsto [0, 1)$ monotonically increases. Hence, the change (7.24) leads to the following optimal control problem (Q) on the finite time interval $[0, 1)$:

$$\dot{x} = \frac{1}{\rho(1-\tau)} f(x, u), \qquad u \in U, \tag{7.25}$$

$$x(0) = x_0,$$

$$J(x, u) = \int_0^1 g(x, u)\, d\tau \to \max.$$

Here, as usual, the class of admissible controls consists of all measurable functions $u : [0, 1) \mapsto U$.

Problem (Q) is equivalent to the original problem (P). The control system (7.25) in problem (Q) has a singularity at the final moment $\tau_1 = 1$. In view of this singularity, the right-hand side of the control system (7.25) may tend to infinity when time τ approaches 1 from the left. The constructions of the classical optimal control theory [**3**], [**32**], which are based on standard assumptions about the dynamics of the system and the set of admissible controls, suggest that the right-hand side of the control system is bounded along admissible trajectories over the entire finite time interval under consideration. This fact plays an important role when applying the needle variation method. Thus, in spite of the fact that problem (Q) is an optimal control problem on the bounded time interval $[0, 1)$, it cannot be placed within the framework of the standard optimal control theory because of its singularity.

When studying the infinite-horizon optimal control problem (P), it is often convenient to pass to the so-called current adjoint variable, current Hamilton–Pontryagin function, and current Hamiltonian in the relations of the Pontryagin maximum principle; at every moment $t \geq 0$, the values of these current quantities differ from the values of the standard adjoint variable, Hamilton–Pontryagin function, and Hamiltonian by the exponential factor $e^{\rho t}$. The term "current" is associated with the fact that, in many economic problems, the discounting factor $e^{-\rho t}$ characterizes the preference for the current utility $g(x(t), u(t))$ at moment $t > 0$ over the utility at the initial moment $t_0 = 0$. Let us give necessary explanations.

Let $\psi(\cdot)$ be the adjoint variable that appears in relations (7.1)–(7.4) of the Pontryagin maximum principle for problem (P). Define a current adjoint variable $p : [0, \infty) \mapsto R^n$ by

$$p(t) = e^{\rho t}\psi(t) \quad \text{for any} \quad t \geq 0. \tag{7.26}$$

Accordingly, for any $x \in G$, $u \in U$, $p \in R^n$, and $\psi^0 \geq 0$, we define a current Hamilton–Pontryagin function $\mathcal{M}(x, u, p, \psi^0)$ and a current Hamiltonian $M(x, p, \psi^0)$ by

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

$$M(x, p, \psi^0) = e^{\rho t}H(x, t, \psi, \psi^0) = \sup_{u\in U}\big(\langle f(x, u), p\rangle + \psi^0 g(x, u)\big).$$

In terms of the current value adjoint variable $p(\cdot)$ and the functions $\mathcal{M}(\cdot, \cdot, \cdot)$ and $M(\cdot, \cdot)$, relations (7.1)–(7.2) of the Pontryagin maximum principle for problem (P) can be rewritten in the following equivalent form:

$$\dot{p}(t) \overset{\text{a.e.}}{=} \rho p(t) - \frac{\partial}{\partial x}\mathcal{M}(x_*(t), u_*(t), p(t), \psi^0) \tag{7.27}$$

$$= \rho p(t) - \left[\frac{\partial f(x_*(t), u_*(t))}{\partial x}\right]^* p(t) - \psi^0 \frac{\partial g(x_*(t), u_*(t))}{\partial x},$$

$$(x_*(t), u_*(t), p(t), \psi^0) \overset{\text{a.e.}}{=} M(x_*(t), p(t), \psi^0),$$

$$\|p(0)\| + \psi^0 > 0. \tag{7.28}$$

The transversality conditions (7.13) and (7.14) in terms of the variable $p(\cdot)$ are given by

$$\psi^0 = 1, \quad \lim_{t\to\infty} e^{-\rho t}p(t) = 0 \tag{7.29}$$

and

$$\psi^0 = 1, \quad \lim_{t\to\infty} e^{-\rho t}\langle x_*(t), p(t)\rangle = 0, \tag{7.30}$$

and the stationarity of the Hamiltonian (7.4) has the form

$$M(x_*(t), p(t), \psi^0) = \psi^0 \rho e^{\rho t}\int_t^\infty e^{-\rho s} g(x_*(s), u_*(s))\, ds \quad \text{for any} \quad t \geq 0. \tag{7.31}$$

Finally, the Pontryagin maximum principle in the normal form (with $\psi^0 = 1$) for problem (P) is expressed in terms of the current variable $p(\cdot)$ as

$$\begin{aligned}\dot{p}(t) &\overset{\text{a.e.}}{=} \rho p(t) - \frac{\partial}{\partial x}\mathcal{M}(x_*(t), u_*(t), p(t)) \\ &= \rho p(t) - \left[\frac{\partial f(x_*(t), u_*(t))}{\partial x}\right]^* p(t) - \frac{\partial g(x_*(t), u_*(t))}{\partial x},\end{aligned} \tag{7.32}$$

$$\mathcal{M}(x_*(t), u_*(t), p(t)) \overset{\text{a.e.}}{=} M(x_*(t), p(t)). \tag{7.33}$$

Here, for any $x \in G$, $u \in U$, and $p \in R^n$,

$$\mathcal{M}(x,u,p) = e^{\rho t}\mathcal{H}(x,u,\psi) = \langle f(x,u), p\rangle + g(x,u), \tag{7.34}$$

$$M(x,p) = e^{\rho t}H(x,t,\psi) = \sup_{u\in U}\big(\langle f(x,u),p\rangle + g(x,u)\big). \tag{7.35}$$

Note that the nontriviality condition (7.28) holds automatically because $\psi^0 = 1$ in this case, and the stationarity condition (7.31) in the normal form is expressed in terms of the current value adjoint variable $p(\cdot)$ as

$$M(x_*(t),p(t)) = \rho e^{\rho t}\int_t^\infty e^{-\rho s} g(x_*(s),u_*(s))\,ds \quad \text{for any} \quad t \geq 0. \tag{7.36}$$

8. Economic interpretation of the maximum principle

Here we discuss an economic interpretation of the Pontryagin maximum principle. On whole we follow the standard view on the maximum principle that has been established in the economic literature (see [**16**], [**27**] and [**41**]).

The economic interpretation of the maximum principle is based on its relation to Bellman's dynamic programming method (see [**9**] and [**32**]) and interpretation of value $V(x)$ of an optimal value function $V(\cdot)$ as the net value of the corresponding capital vector (capital stock) $x \in G$.

Assuming that the optimal value function $V(\cdot)$ is sufficiently smooth and applying the dynamic programming technique, we obtain a version of the Pontryagin maximum principle for problem (P) in the normal form (i.e. with $\psi^0 = 1$) in terms of the current value adjoint variable $p(\cdot)$ (see (7.26)), the current value Hamilton–Pontryagin function $\mathcal{M}(\cdot,\cdot,\cdot)$, and the current value Hamiltonian $M(\cdot,\cdot)$ (see formulas (7.34) and (7.35)). Together with the core relations (7.32) and (7.33) this version of the maximum principle contains also additional condition of stationarity of the Hamiltonian (see (7.36)). The current value adjoint variable $p(\cdot)$ in the maximum principle such obtained is defined as the gradient of the optimal value function $V(\cdot)$ along the optimal trajectory $x_*(\cdot)$ under consideration. The latter fact allows us to interpret the coordinates of the vector $p(t)$ as "marginal" or "shadow" prices of units of coordinates of the capital stock $x_*(t)$ at moment $t \geq 0$, thus rendering the relations of the maximum principle economically meaningful.

Suppose that the optimal control problem (P) (see (5.1)–(5.3)) is a model of optimal capital accumulation of an enterprise. Suppose also that at any moment $t \geq 0$, the state of an enterprise is completely characterized by a capital stock $x(t) = (x^1(t), \ldots, x^n(t)) \in R^n$ whose coordinates are identified with different production assets. Further, we will assume that the coordinates of the control vector $u(t) = (u^1(t), \ldots, u^m(t)) \in R^m$ characterize amounts of capital installed in the unit of time following the moment $t \geq 0$. In this case, the control $u(\cdot)$ can be identified with the investment policy of the enterprise.

Suppose that the vector $u(t)$ is constrained at any moment $t \geq 0$; i.e., $u(t) \in U$ for any $t \geq 0$, where U is a nonempty compact set in R^m.

Thus, we assume that the control system

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

(see (5.1)) describes the dynamics of the capital stock $x(\cdot)$ depending on the investment policy $u(\cdot)$ chosen.

For example, when the production assets are homogeneous ($x \in R^1$), equipment does not depreciate or become obsolete, and the newly bought equipment cannot be sold or somehow withdrawn from the production process (for example, mothballed), the dynamics of the capital of the enterprise is described by the following control system:

$$\dot{x}(t) = u(t), \qquad u(t) \in U = \{u \in R^1 : 0 \le u \le u_{\max}\}.$$

Here, $u_{\max} > 0$ is the maximum possible installation velocity of production assets.

If we take into account the depreciation of capital, then the control system describing the dynamics of capital takes the form

$$\dot{x}(t) = u(t) - \delta x(t), \quad u(t) \in U = \{u \in R^1 : 0 \le u \le u_{\max}\}.$$

Here, $\delta > 0$ is the depreciation rate of capital.

We assume that at the moment $t_0 = 0$ the initial state of the system is known:

$$x(0) = \xi \in G.$$

Here G is a given nonempty open subset of R^n.

We assume also that the technologies used in the enterprise do not change and the production assets $x(t)$ available at any moment $t \ge 0$ are fully utilized. In this case, the production output and expenses on various factors of production, such as labor force, materials, energy, etc., are uniquely defined at any moment t by the vector $x(t)$.

Suppose that the instantaneous profit $g(x(t), u(t))$ of the enterprise at any moment $t \ge 0$ (in prices at the current time t) is completely determined by the values of the production assets $x(t)$ available at this moment t and by the control vector $u(t)$. In this case, the quantity $g(x(t), u(t))$ takes into account the expenses for the current production (which are fully determined by the capital $x(t)$), the expenses for acquiring new production facilities and setting them up, as well as the revenue from selling finished goods.

Consider the situation when the enterprise operates for an infinitely long time. Let $\rho > 0$ be a subjective discount rate (for example, in the case of enterprise the discount rate can be associated with the rate of inflation). Then, the aggregated profit allowing for inflation (in prices at the initial moment $t_0 = 0$) of the enterprise over the entire infinite time interval of its operation is defined by a goal functional of the form

$$J(x(\cdot), u(\cdot)) = \int_0^\infty e^{-\rho t} g(x(t), u(t))\, dt.$$

Thus, our mathematical model of optimal control of the capital accumulation of an enterprise is given by the following optimal control problem $P(\xi, 0)$:

$$\dot{x} = f(x, u), \quad u \in U, \tag{8.1}$$

$$x(0) = \xi, \tag{8.2}$$

$$J(x, u) = \int_0^\infty e^{-\rho t} g(x, u)\, dt \to \max. \tag{8.3}$$

We assume that for any initial state $\xi \in G$ and arbitrary $T > 0$, conditions (A1), (A2) take place, and conditions (A5), (A6) are satisfied as well; i.e., for any $\xi \in G$ problem $P(\xi, 0)$ is a problem (P) with $x_0 = \xi$.

Note, that in many economic optimal control problems of the form (8.1)–(8.3), the vector function $f(\cdot, \cdot)$, the scalar function $g(\cdot, \cdot)$, and the set G may satisfy some additional general conditions that are typical of various classes of economic models. For example, it is often natural to consider the situation when all coordinates of the phase vector (capital) $x(t)$ are positive at any moment $t \geq 0$. In this case,

$$G = \{x \in R^n : x^i > 0,\ i = 1, 2, \dots, n\}.$$

As regards the function $g(\cdot, \cdot)$, it is sometimes natural to assume that function $g(\cdot, u)$ monotonically increases in each coordinate of the vector x for any $u \in U$, i.e., $\partial g(x, u)/\partial x^i > 0$, $i = 1, \dots, n$, for any $x \in G$ and $u \in U$. This condition means that the instantaneous profit increases with the increase of the amount of equipment available. In this case, the instantaneous profit $g(\cdot, u)$ often turns out to be a concave function of the coordinates of the variable x for any fixed $u \in U$. In many economic problems, the function $f(\cdot, \cdot)$ is such that the control system (8.1) is a system with saturation. In this case, all trajectories of the control system (8.1) with the given initial condition (8.2) take values in a certain bounded subset of the phase space R^n. Distinguishing special subclasses of the general problem (P) sometimes allows one to strengthen the general formulation of the maximum principle (Theorem 18) and to obtain additional conditions characterizing the adjoint variables.

By Theorem 15, problem $P(\xi, 0)$ is solvable for any initial state $x(0) = \xi \in G$. Denote the corresponding optimal value of the utility functional by $V(\xi)$. Thus, value $V(\xi)$ is defined for any $\xi \in G$. The functional $V(\cdot)$ is usually called the optimal value function or the value function in problem (8.1)–(8.3). In our case, $V(\xi)$ is, in essence, the maximum possible total discounted profit of an enterprise that has capital stock $x(0) = \xi$ at the initial moment $t_0 = 0$. This is the profit that can be earned provided that the control of the enterprise on the entire infinite time interval $[0, \infty)$ is optimal.

Define the net value (or simply value) of capital vector $\xi \in G$ as the corresponding value of $V(\xi)$. We stress that such defined value $V(\xi)$ is the cost of capital $\xi \in G$ in prices at the initial moment $t_0 = 0$. This cost of the capital vector ξ is in no way related to its "market" price at this moment, i.e., to the price of production assets ξ in the relevant production asserts markets at the moment $t_0 = 0$, and expresses exclusively the potential profit that can be made from given capital vector ξ over the infinite time interval $[0,\infty)$ provided that the enterprise implements an optimal investment policy. In this sense, the net value $V(\xi)$ of the capital $\xi \in G$ is its "true" value from the viewpoint of the enterprise. It is the net value of the capital vector ξ that one should take into account when determining the investment policy of the enterprise.

Note, that in contrast to the net value $V(\xi)$ of capital stock $\xi \in G$, its market price at any moment $t \geq 0$ can be observed directly. However, it is determined by subjective expectations of agents operating in appropriate markets (sellers and buyers) rather than by the potential aggregated profit of the enterprise. As a result, in real markets, the production assets of the enterprise may be underestimated or overestimated with respect to their net values.

Apart from the net value $V(\xi)$ of capital stock $\xi \in G$ at the initial moment $t_0 = 0$ (in prices at the moment $t_0 = 0$) define its current net value $V(\xi,\tau)$ at instant $\tau \geq 0$ as the optimal value of the functional in the following optimal control problem $P(\xi,\tau)$:

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

$$x(\tau) = \xi,$$

$$J_\tau(x,u) = \int_\tau^\infty e^{-\rho(t-\tau)} g(x,u)\, ds \to \max$$

Hear $\xi \in G$ and $\tau \geq 0$. All other data in problem $P(\xi,\tau)$ are the same as in initial problem (P). It is easy to see that problem $P(\xi,0)$ is a particular case of problem $P(\xi,\tau)$ corresponding to the case $\tau = 0$.

Let $(x(\cdot),u(\cdot))$ be an admissible trajectory in problem $P(\xi,0)$, $\xi \in G$. In this case the pair $(x(\cdot),u(\cdot))$ is defined on the infinite time interval $[0,\infty)$ and $x(0) = \xi$. For any $\tau \geq 0$, consider a pair $(x_\tau(\cdot),u_\tau(\cdot))$, where $x_\tau(t) = x(t-\tau)$ and $u_\tau(t) = u(t-\tau)$ for $t \geq \tau$. It is easy to see that $(x_\tau(\cdot),u_\tau(\cdot))$ is an admissible pair in the problem $P(\xi,\tau)$. Vice versa if $(x_\tau(\cdot),u_\tau(\cdot))$ is an admissible pair in the problem $P(\xi,\tau)$ then the pair $(x(\cdot),u(\cdot))$, $x(s) = x_\tau(s+\tau)$, $u(s) = u_\tau(s+\tau)$, $s \geq 0$, is an admissible pair in the problem $P(\xi,0)$. So, the change of time $t = s+\tau$, $s \geq 0$, provide one-to-one correspondence between admissible pairs in problems $P(\xi,0)$ and $P(\xi,\tau)$ for any $\xi \in G$ and $\tau \geq 0$.

Moreover, it is easy to see that the following equality takes place for corresponding to each other pairs $(x_\tau(\cdot),u_\tau(\cdot))$ and $(x(\cdot),u(\cdot))$:

$$\int_\tau^\infty e^{-\rho(t-\tau)} g(x_\tau(t),u_\tau(t))\, dt = \int_0^\infty e^{-\rho s} g(x(s),u(s))\, ds.$$

This equality implies that the change of time $t = s + \tau$, $s \geq 0$, provide also one-to-one correspondence between optimal admissible pairs in problems $P(\xi, 0)$ and $(P(\xi, \tau))$ for any $\xi \in G$ and $\tau > 0$. So, if $x_*(\cdot), u_*(\cdot)$ is an optimal admissible pair in problem $P(\xi, 0)$ then the corresponding pair $x_{\tau,*}(\cdot), u_{\tau,*}(\cdot)$, $x_{\tau,*}(t) = x_*(t - \tau)$ and $u_{\tau,*}(t) = u_*(t - \tau)$ for $t \geq \tau$, is an optimal admissible pair in problem $P(\xi, \tau)$ and we have

$$V(\xi, \tau) = V(\xi) \quad \text{for any} \quad \xi \in G \quad \text{and} \quad \tau \geq 0. \tag{8.4}$$

Thus, in problem $P(\xi, \tau)$, the current net value $V(\xi, \tau)$ of capital stock ξ does not depend on time $\tau \geq 0$. If the enterprise has initial capital $\xi \in G$ at moment τ and implements an optimal investment policy, it earns aggregated profit equal to $V(\xi)$ in prices at time τ, allowing for inflation. Thus, $V(\xi)$ is the net value of the capital vector $x(t)$ in current prices at any moment $t \geq 0$.

Suppose that the value function $V(\cdot)$ is twice continuously differentiable on the set G.

Let $\xi \in G$ be arbitrary and $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible pair in the problem $P(\xi, 0)$. Then it is easy to see that the tale of the pair $(x_*(\cdot), u_*(\cdot))$, considered on the time interval $[\tau, \infty)$, i.e. the pair $(x_{\tau,*}(\cdot), u_{\tau,*}(\cdot))$, $x_{\tau,*}(t) = x_*(t)$, $u_{\tau,*}(t) = u_*(t)$, $t \geq \tau$, is an optimal pair in the problem $P(x_*(\tau), \tau)$ (note that this fact is usually called the Bellman optimality principle). Hence, as it is shown above, for any $\tau \geq 0$, the pair $(\bar{x}_*(\cdot), \bar{u}_*(\cdot))$, $\bar{x}_*(s) = x_{\tau,*}(s+\tau) = x_*(s+\tau)$, $\bar{u}_*(s) = u_{\tau,*}(s+\tau) = u_*(s+\tau)$, $s \geq 0$, is an optimal admissible pair in the problem $P(x_*(\tau), 0)$ and due to (8.4) we have

$$V(x_*(\tau)) = e^{\rho\tau} \int_\tau^\infty e^{-\rho t} g(x_*(t), u_*(t))\, dt \tag{8.5}$$

for any $\tau \geq 0$.

Consider the motion of the system along the optimal (in the problem $P(\xi, 0)$) trajectory $x_*(\cdot)$.

At time $t \geq 0$, the system is in the state $x_*(t)$. On a small time interval $[t, t + \Delta t]$, $\Delta t > 0$, the system goes over to the state $x_*(t + \Delta t)$. By condition (8.5), we have

$$e^{-\rho(t+\Delta t)} V(x_*(t + \Delta t)) - e^{-\rho t} V(x_*(t)) = -\int_t^{t+\Delta t} e^{-\rho s} g(x_*(s), u_*(s))\, ds.$$

Let us divide this equality by $\Delta t > 0$, and let Δt go to zero. Since the value function $V(\cdot)$ is continuously differentiable, we find that the optimal pair $(x_*(\cdot), u_*(\cdot))$ satisfies the following equality for a.e. $t \geq 0$:

$$\left\langle \frac{\partial}{\partial x} V(x_*(t)), f(x_*(t), u_*(t)) \right\rangle - \rho V(x_*(t)) + g(x_*(t), u_*(t)) = 0. \tag{8.6}$$

Note that equality (8.6) holds a.e. on $[0, \infty)$ along any optimal admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem $P(\xi, 0)$ for arbitrary $\xi \in G$.

Again, let $\xi \in G$ be arbitrary and define an admissible pair $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ with the initial condition $x(0) = \xi$ as follows. Take a small time interval $\Delta = [0, \Delta t]$, $\Delta t > 0$, and define, on this interval, a trajectory $\tilde{x}(\cdot)$ that starts at the point ξ under constant control $\tilde{u}(t) = \bar{u}$, $t \in \Delta$, where $\bar{u} \in U$ is arbitrary. Such motion on the time interval Δ brings the capital vector $x(0) = \xi$ to the state

$$\tilde{x}(\Delta t) = \xi + \Delta t f(\xi, \bar{u}) + \mathrm{o}(\Delta t).$$

We will assume that the further motion of the system from the point $\tilde{x}(\Delta t)$ on the infinite time interval $[\Delta t, \infty)$ is optimal (by Theorem 15, an optimal admissible control in problem $(P(\tilde{x}(\Delta t), \Delta t))$ exists). In this case, the value of the goal functional on the admissible pair $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ is given by (see (8.5))

$$J(\tilde{x}(\cdot), \tilde{u}(\cdot)) = \int_0^{\Delta t} e^{-\rho t} g(\tilde{x}(t), \bar{u})\, dt + e^{-\rho \Delta t} V(\tilde{x}(\Delta t)).$$

Since the admissible pair $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ is not necessarily optimal in problem $P(\xi, 0)$, we have $J(\tilde{x}(\cdot), \tilde{u}(\cdot)) \le V(\xi)$ and, hence, the following inequality holds:

$$e^{-\rho \Delta t} V(\tilde{x}(\Delta t)) - V(\xi) + \int_0^{\Delta t} e^{-\rho t} g(\tilde{x}(t), \bar{u})\, dt \le 0.$$

Dividing the last inequality by $\Delta t > 0$ and letting Δt go to zero, we obtain (since the function $V(\cdot)$ is continuously differentiable)

$$\Big\langle \frac{\partial}{\partial x} V(\xi), f(\xi, \bar{u}) \Big\rangle - \rho V(\xi) + g(\xi, \bar{u}) \le 0.$$

Hence, due to the arbitrariness of $\xi \in G$ and $\bar{u} \in U$, we find that

$$\Big\langle \frac{\partial}{\partial x} V(x), f(x, u) \Big\rangle - \rho V(x) + g(x, u) \le 0 \quad \text{for any} \quad x \in G \quad \text{and} \quad u \in U. \tag{8.7}$$

Let $t \ge 0$ be such that equality (8.6) holds at this moment. Then, in view of (8.7), the following inequality holds for any $x \in G$ at this moment:

$$\Big\langle \frac{\partial}{\partial x} V(x), f(x, u_*(t)) \Big\rangle - \rho V(x) + g(x, u_*(t)) \le 0,$$

while for $x = x_*(t)$ we have the equality:

$$\Big\langle \frac{\partial V(x_*(t))}{\partial x}, f(x_*(t), u_*(t)) \Big\rangle - \rho V(x_*(t)) + g(x_*(t), u_*(t)) = 0.$$

Thus, the function $\phi : G \mapsto R^1$ defined by

$$\phi(x) = \Big\langle \frac{\partial V(x)}{\partial x}, f(x, u_*(t)) \Big\rangle - \rho V(x) + g(x, u_*(t)) \quad \text{for any} \quad x \in G$$

attains its maximum at the point $x_*(t) \in G$. As the function $V(\cdot)$ is continuously differentiable on G, we have

$$\frac{\partial}{\partial x} \phi(x_*(t)) = 0$$

for a.e. $t \geq 0$; i.e.,

$$\frac{\partial^2}{\partial x^2} V(x_*(t)) f(x_*(t), u_*(t)) + \left[\frac{\partial}{\partial x} f(x_*(t), u_*(t)) \right]^* \frac{\partial}{\partial x} V(x_*(t)) \\ - \rho \frac{\partial}{\partial x} V(x_*(t)) + \frac{\partial}{\partial x} g(x_*(t), u_*(t)) \stackrel{\text{a.e.}}{=} 0. \quad (8.8)$$

Next, in view of (8.7), the inequality

$$\left\langle \frac{\partial}{\partial x} V(x_*(t)), f(x_*(t), u) \right\rangle + g(x_*(t), u) \leq \rho V(x_*(t)) \quad (8.9)$$

holds for any $u \in U$ and any $t \geq 0$; in view of (8.6), for $u = u_*(t)$ this inequality turns into equality for a.e. $t \geq 0$. Hence,

$$\rho V(x_*(t)) \stackrel{\text{a.e.}}{=} \max_{u \in U} \left\{ \left\langle \frac{\partial}{\partial x} V(x_*(t)), f(x_*(t), u) \right\rangle + g(x_*(t), u) \right\}. \quad (8.10)$$

Thus, conditions (8.6), (8.9), and (8.10) imply that the following maximum condition holds on the entire infinite time interval $[0, \infty)$:

$$\left\langle \frac{\partial}{\partial x} V(x_*(t)), f(x_*(t), u_*(t)) \right\rangle + g(x_*(t), u_*(t)) \\ \stackrel{\text{a.e.}}{=} \max_{u \in U} \left\{ \left\langle \frac{\partial}{\partial x} V(x_*(t)), f(x_*(t), u) \right\rangle + g(x_*(t), u) \right\}. \quad (8.11)$$

Due to assumptions made for any $x \in G$ there is an optimal admissible control $u_*(\cdot)$ in problem $P(x, 0)$. Since the value function $V(\cdot)$ is assumed to be continuously differentiable on G this fact and condition (8.10) (which is valid along any optimal admissible trajectory $x_*(\cdot)$ in any problem $P(x, 0)$ with arbitrary $x \in G$) imply that the value function $V(\cdot)$ is a solution to the following first order partial differential equation on G in this case:

$$\max_{u \in U} \left\{ \left\langle \frac{\partial}{\partial x} V(x), f(x, u) \right\rangle + g(x, u) \right\} - \rho V(x) = 0. \quad (8.12)$$

Equation (8.12) is called the Bellman equation (for problem (P)). Consideration of the Bellman equation constitutes the essence of the Bellman dynamic programming method. If the maximizer $u_*(x) \in U$ in (8.12) is unique for any $x \in G$ then due to (8.11) the function $u_* : G \mapsto U$ is an optimal synthesis (see [**32**]) in problem (P).

Define a vector function $p : [0, \infty) \mapsto R^n$ by

$$p(t) = \frac{\partial}{\partial x} V(x_*(t)), \qquad t \in [0, \infty). \quad (8.13)$$

Since the function $V(\cdot)$ is twice continuously differentiable on G and the trajectory $x_*(\cdot)$ is absolutely continuous on any finite time interval $[0, T]$, $T > 0$, the vector function $p(\cdot)$ is also absolutely continuous on any finite time interval

$[0,T]$, $T>0$; i.e., it is locally absolutely continuous. Hence, by virtue of (8.8), for a.e. $t\in[0,\infty)$ we have

$$\begin{aligned}\dot p(t) &= \frac{\partial^2}{\partial x^2}V(x_*(t))f(x_*(t),u_*(t))\\ &= -\left[\frac{\partial}{\partial x}f(x_*(t),u_*(t))\right]^*\frac{\partial}{\partial x}V(x_*(t)) + \rho\frac{\partial}{\partial x}V(x_*(t)) - \frac{\partial}{\partial x}g(x_*(t),u_*(t)),\end{aligned}$$

or

$$\dot p(t) = \rho p(t) - \left[\frac{\partial}{\partial x}f(x_*(t),u_*(t))\right]^* p(t) - \frac{\partial}{\partial x}g(x_*(t),u_*(t)). \tag{8.14}$$

In terms of the function $p(\cdot)$ (see (8.13)), the maximum condition (8.11) is expressed as

$$\big\langle f(x_*(t),u_*(t)),p(t)\big\rangle + g(x_*(t),u_*(t)) \overset{\text{a.e.}}{=} \max_{u\in U}\big\{\big\langle f(x_*(t),u),p(t)\big\rangle + g(x_*(t),u)\big\}. \tag{8.15}$$

Moreover, by virtue of (8.5) and (8.6), we obtain

$$\big\langle f(x_*(t),u_*(t)),p(t)\big\rangle + g(x_*(t),u_*(t)) = \rho e^{\rho t}\int_t^\infty e^{-\rho s}g(x_*(s),u_*(s))\,ds \tag{8.16}$$

for a.e. $t\ge 0$.

According to (7.34) and (7.35), we define the current value Hamilton–Pontryagin function $\mathcal{M}(\cdot,\cdot,\cdot)$ and the current value Hamiltonian $M(\cdot,\cdot)$ in the normal form by the equalities

$$\mathcal{M}(x,u,p) = \langle f(x,u),p\rangle + g(x,u), \quad M(x,p) = \sup_{u\in U}\big(\langle f(x,u),p\rangle + g(x,u)\big)$$

for any $x\in G$, $u\in U$, and $p\in R^n$; i.e., we set

$$\psi^0 = 1. \tag{8.17}$$

Then, in terms of the functions $\mathcal{M}(\cdot,\cdot,\cdot)$ and $M(\cdot,\cdot)$, formula (8.14) can be rewritten as (see (7.32))

$$\dot p(t) \overset{\text{a.e.}}{=} \rho p(t) - \frac{\partial}{\partial x}\mathcal{M}(x_*(t),u_*(t),p(t)), \tag{8.18}$$

and the maximum condition (8.15) is expressed as (see (7.33))

$$\mathcal{M}(x_*(t),u_*(t),p(t)) \overset{\text{a.e.}}{=} M(x_*(t),p(t)). \tag{8.19}$$

The stationarity condition (8.16) in terms of the function $M(\cdot,\cdot)$ has the form (see (7.36))

$$M(x_*(t),p(t)) = \rho e^{\rho t}\int_t^\infty e^{-\rho s}g(x_*(s),u_*(s))\,ds \quad \text{for any} \quad t\ge 0. \tag{8.20}$$

By virtue of equality (8.17), the nondegeneracy condition (7.28) holds automatically.

Thus, we have established the following version of the Pontryagin maximum principle for problem (P).

Theorem 19. *Suppose that for any initial state $x_0 \in G$ conditions* (A1), (A2) *(on any finite time interval $[0, T]$, $T > 0$) take place, and condition* (A6) *is satisfied as well. Assume that the value function $V(\cdot)$ is twice continuously differentiable on the set G. Let an admissible pair $(x_*(\cdot), u_*(\cdot))$ be optimal in problem (P). Then, the pair $(x_*(\cdot), u_*(\cdot))$, together with the current value adjoint variable $p(\cdot)$ where*

$$p(t) = \frac{\partial}{\partial x} V(x_*(t)) \quad \text{for any} \quad t \geq 0,$$

satisfies relations (8.18) *and* (8.19) *of the Pontryagin maximum principle. Moreover, the optimal pair $(x_*(\cdot), u_*(\cdot))$ and the adjoint variable $p(\cdot)$ satisfy the stationarity condition* (8.20).

Note that Theorem 19 is formulated in terms of the current value adjoint variable $p(\cdot)$ (see (7.26)). This result contains some additional information about the pair of adjoint variables $(\psi^0, p(\cdot))$ (however, the additional conditions are obtained here under the assumption that the value function $V(\cdot)$ is twice continuously differentiable). In particular, Theorem 19 is a version of the Pontryagin maximum principle in the normal form ($\psi^0 = 1$), which is not generally true (see Example 6 in previous Section 7, in which the equality $\psi^0 = 0$ necessarily holds). Theorem 19 provides also a relationship between the current value adjoint variable $p(\cdot)$ and the value function $V(\cdot)$ (the current value adjoint variable $p(\cdot)$ is represented as a gradient of $V(\cdot)$ along an optimal trajectory $x_*(\cdot)$).

Now, we pass directly to the economic interpretation of the Pontryagin maximum principle.

First of all, recall that the net value $V(x)$ of the capital vector $x \in G$ was defined above as its value at the initial moment $t_0 = 0$ (in prices at the moment $t_0 = 0$). In fact, this value $V(x)$ is always equal to the net present value $V(x, t)$ of the capital vector x at any subsequent moment $t \geq 0$, i.e., to the net value of capital x in current prices at any moment $t \geq 0$ (see (8.4)). Thus, in problem (P), the net value of capital stock in current prices does not depend on time. If the enterprise has initial capital $x \in G$ at some moment $\tau \geq 0$ and implements an optimal investment policy, then it earns aggregated profit equal to $V(x)$ in prices at time τ, allowing for inflation. Thus, $V(x)$ is the net value of the capital vector x in current prices at any moment $t \geq 0$.

Consider the Hamilton–Pontryagin function $\mathcal{M}(\cdot, \cdot, \cdot)$ and the maximum condition (8.19).

Since the function $V(\cdot)$ of the net value of the capital vector is continuously differentiable, the increment of the net value of capital when the system passes from state $x \in G$ to state $x + \Delta x \in G$ is given by

$$V(x + \Delta x) - V(x) = \Big\langle \frac{\partial}{\partial x} V(x), \Delta x \Big\rangle + o(\|\Delta x\|).$$

Thus, the coordinates of the vector $\partial V(x)/\partial x$ characterize the increment of the net value of the capital vector when the system passes from point x to a nearby point $x + \Delta x$. When the capital increment Δx is small, the resulting increment of the net value of capital is a weighted sum of the coordinates of Δx with weights equal to the coordinates of the vector $\partial V(x)/\partial x$. Hence, the ith ($i = 1, 2, \dots, n$) coordinate of the vector $\partial V(x)/\partial x$ can be interpreted as the "marginal" price per unit of the ith coordinate of the capital vector x. Since $V(x)$ is the net value of the capital vector x in current prices at an arbitrary moment $t \geq 0$, the marginal price $\partial V(x)/\partial x$ is also the current value marginal price at arbitrary moment $t \geq 0$.

Since, by Theorem 19, the current value adjoint variable $p(\cdot)$ is defined by the equality

$$p(t) = \frac{\partial}{\partial x} V(x_*(t)) \quad \text{for any} \quad t \geq 0,$$

the coordinates of the vector $p(t)$ are the current value marginal prices of the corresponding units of the capital vector $x_*(t)$ (at an arbitrary current moment $t \geq 0$). These marginal prices are also called shadow prices because they generally differ from the market prices of units of the coordinates of the capital vector $x_*(t)$ that are observed in appropriate markets at moment t.

Note that the adjoint variable $\psi(\cdot)$ in Theorem 18 differs from the current adjoint variable $p(\cdot)$ in Theorem 19 by the exponential factor (see (7.26)):

$$\psi(t) = e^{-\rho t} p(t) \quad \text{for any} \quad t \geq 0.$$

Taking into account that $e^{-\rho t}$ is the value of the currency at moment t allowing for inflation, it is natural to interpret the coordinates of the adjoint variable $\psi(t)$ as "reduced" marginal prices per units of the coordinates of the capital vector at point $x_*(t)$; by this we mean that they are reduced to the initial moment $t_0 = 0$ (they express the marginal cost of units of capital at point $x_*(t)$ in prices of the initial moment t_0, i.e., allowing for inflation).

Consider the values $\mathcal{M}(x_*(t), u_*(t), p(t))$ of the current value Hamilton–Pontryagin function in the normal form $\mathcal{M}(\cdot, \cdot, \cdot)$ at moment t. According to the definition of the Hamilton–Pontryagin function, we have

$$\mathcal{M}(x_*(t), u_*(t), p(t)) = \big\langle f(x_*(t), u_*(t)), p(t) \big\rangle + g(x_*(t), u_*(t)). \tag{8.21}$$

Here, as above,

$$p(t) = \frac{\partial}{\partial x} V(x_*(t)) \quad \text{for any} \quad t \geq 0$$

is the vector of marginal prices of units of capital $x_*(t)$. Hence, for almost every t, the equality

$$V(x_*(t+\Delta t)) - V(x_*(t)) = \Big\langle \frac{\partial}{\partial x} V(x_*(t)), f(x_*(t), u_*(t)) \Big\rangle \Delta t + o(\Delta t)$$

implies that the first term $\langle f(x_*(t), u_*(t)), p(t)\rangle$ on the right-hand side of (8.21) is the increment of the net value of capital $x_*(t)$ per unit of time at moment t under the optimal investment policy $u_*(t)$. This term reflects the quality of the control process $u_*(t)$ at moment t from the viewpoint of the variation of the net value of capital, or, which is essentially the same, from the viewpoint of the total subsequent profit. The second term $g(x_*(t), u_*(t))$ is the current actual profit of the enterprise per unit of time at moment t (instantaneous current actual profit). Thus, the quantity $\mathcal{M}(x_*(t), u_*(t), p(t))$ consists of the instantaneous net current profit $\langle f(x_*(t), u_*(t)), p(t)\rangle$ and the instantaneous current actual profit $g(x_*(t), u_*(t))$ of the enterprise; i.e., this quantity can be interpreted as an aggregated expression for the instantaneous total current profit of the enterprise at moment t; it is expressed in current prices at the moment t. It is the quantity that most comprehensively characterizes the chosen investment policy of the enterprise. By the maximum condition (see (8.19)), the optimal investment policy $u_*(t)$ must maximize the instantaneous total current profit for a.e. $t \geq 0$. At any moment t, the coordinates of the vector of marginal prices $p(t)$ play the role of weight coefficients corresponding to the variation of the net value of the capital vector $x_*(t)$. This is the economic interpretation of the Hamilton–Pontryagin function and the maximum condition (8.19).

Now, we pass on to the economic interpretation of the adjoint system (8.18).

Suppose that the economy within which the enterprise operates is in dynamic equilibrium. This, in particular, implies that the market prices of units of basic production assets $x_*^i(t)$, $i = 1, \dots, n$, at any moment $t \geq 0$ coincide with the marginal prices $p^i(t)$, $i = 1, \dots, n$.

The adjoint system gives an expression for the growth velocities of current value marginal prices of capital units per unit of time along an optimal trajectory $x_*(\cdot)$.

Let us represent the adjoint system as

$$\dot{p}(t) + \left[\frac{\partial}{\partial x} f(x_*(t), u_*(t))\right]^* p(t) + \frac{\partial}{\partial x} g(x_*(t)) \overset{\text{a.e.}}{=} \rho p(t). \qquad (8.22)$$

Taking the scalar product of both sides of (8.22) with the vector $e = (1, \dots, 1) \in R^n$, all of whose coordinates are ones, we obtain

$$\langle \dot{p}(t), e\rangle + \Big\langle \frac{\partial}{\partial x} f(x_*(t), u_*(t)) e, p(t) \Big\rangle + \Big\langle \frac{\partial}{\partial x} g(x_*(t)), e \Big\rangle \overset{\text{a.e.}}{=} \rho \langle p(t), e\rangle.$$

The left-hand side of this equality contains the increment of the total current profit of the enterprise per unit of time after purchasing, at moment t, the unit capital vector e along an optimal trajectory. Indeed, the first term $\langle \dot{p}(t), e \rangle$ on the left-hand side is the increment of current profit, per unit of time following the moment t, associated with the variation of the vector of marginal prices $p(t)$. The second term gives an expression for the increment of current profit, per unit of time following the moment t, associated with the variation of the net value of capital. The third term on the left-hand side gives an expression for the increment of current profit, per unit of time following the moment t, associated with the variation of the flow of the current profit as a result of acquiring unit capital vector e. The quantity $\rho\langle p(t), e \rangle$ on the right-hand side of the above equality is the value by which the funds $\langle p(t), e \rangle$ spent on the acquisition of unit capital vector e are depreciated per unit of time as a result of inflation.

Thus, the adjoint system (8.18) is a balance relation stating that, along an optimal trajectory, after purchasing a unit capital vector, the increment of the total current profit of the enterprise per unit of time is equal to the value by which the funds spent on acquiring this unit capital vector are depreciated per this unit of time.

Now suppose that, in addition to the investment of funds in the acquisition of production assets, the enterprise can use a certain financial instrument available in the economy, for example, it can put this money into a bank account at interest rate ν. In this case, the value $\langle p(t), e \rangle$ of funds will yield a profit of $\nu\langle p(t), e \rangle$ on the unit time interval following the moment t. If the inequality

$$\rho < \nu$$

holds, then the investment of funds in the acquisition of basic production assets turns out to be less profitable than putting this money into a bank account (the situation of high interest rate). Conversely, if the inequality

$$\rho > \nu$$

holds, then the investment of funds in the acquisition of basic production assets is more profitable than putting them into a bank account (the situation of low interest rate). However, both these situations contradict the assumption that the economy is in dynamic equilibrium, when any deviation from the optimal regime is inexpedient. Therefore, the equality

$$\rho = \nu \tag{8.23}$$

must hold in the case under consideration. This equality implies that, in dynamic equilibrium, the profitability of production coincides with the profitability of the alternative financial instrument available in the economy. Thus condition (8.23) relates the discount parameter ρ to the interest rate ν.

Now, we pass to an interpretation of the transversality conditions at infinity.

The boundary condition (see (7.29)),

$$\lim_{t\to\infty} e^{-\rho t} p(t) = 0,$$

implies that under optimal investment policy, the current value marginal prices of the units of basic production assets increase along an optimal trajectory slower than inflation; i.e., the reduced (to the initial moment) marginal prices per unit of capital, $\psi(t) = e^{-\rho t}p(t)$, tend to zero at infinity. Thus, under optimal investment policy, the costs of units of capital (in prices of the initial moment) are completely exhausted in the limit.

Consider the alternative boundary condition at infinity (see (7.30)):

$$\lim_{t\to\infty} e^{-\rho t} \langle x_*(t), p(t)\rangle = 0. \tag{8.24}$$

Suppose that this condition does not hold along the optimal trajectory $x_*(\cdot)$. Suppose also that $x_*(t) > 0$ and $p(t) > 0$ for any $t \geq 0$. Then, there exist an increasing sequence of time points T_i, $i = 1, 2, \dots$, $\lim_{i\to\infty} T_i = \infty$, and a number $A > 0$ such that

$$\lim_{i\to\infty} e^{-\rho T_i} \langle x_*(T_i), p(T_i)\rangle \geq A.$$

Hence, there exists a natural N such that, for any $i \geq N$, we have

$$e^{-\rho T_i} \langle x_*(T_i), p(T_i)\rangle > \frac{A}{2}.$$

By condition (A6), the integral

$$\int_0^\infty e^{-\rho t} g(x_*(t), u_*(t))\, dt$$

converges absolutely. Therefore, without loss of generality, we may assume that the following inequality holds for any $i \geq N$:

$$\int_0^{T_i} e^{-\rho t} g(x_*(t), u_*(t)\, dt > \int_0^\infty e^{-\rho t} g(x_*(t), u_*(t))\, dt - \frac{A}{2},$$

which implies that

$$\begin{aligned} J(x_*(\cdot), u_*(\cdot)) &= \int_0^\infty e^{-\rho t} g(x_*(t), u_*(t))\, dt \\ &< \int_0^{T_i} e^{-\rho t} g(x_*(t), u_*(t))\, dt + e^{-\rho T_i} \langle x_*(T_i), p(T_i)\rangle \end{aligned}$$

for any $i \geq N$. However, it follows from this inequality that if the marginal prices $p(T_i)$ coincide with the market prices at the moments T_i, selling all available production assets $x_*(T_i)$ at these moments will yield a profit (allowing for inflation) greater than the profit one would obtain by continuing to exploit these assets over the whole infinite time interval. In this case, the production assets $x_*(T_i)$ are objectively overestimated at the moments T_i. However, this fact contradicts our assumption that the economy is in dynamic equilibrium.

Thus, the boundary condition (8.24) implies that capital cannot be overestimated at infinity; i.e., the economy at infinity is also in equilibrium.

Finally, consider the stationarity condition (8.20). By (8.5), this condition implies that the following equality holds along the optimal trajectory

$$\rho V(x_*(t)) = \big\langle f(x_*(t), u_*(t)), p(t)\big\rangle + g(x_*(t), u_*(t)) \quad \text{for a.e.} \quad t \geq 0. \quad (8.25)$$

The quantity

$$\big\langle f(x_*(t), u_*(t)), p(t)\big\rangle + g(x_*(t), u_*(t))$$

on the right-hand side of (8.25) is the total profit earned by capital $x_*(t)$ per unit of time following the moment t (the increment of the net value of capital plus the actual profit in terms of money). The quantity $V(x_*(t))$ on the left-hand side is the value of capital $x_*(t)$ in prices of the current moment t. If one puts this money into a bank account at interest rate ρ, this will yield a profit

$$\rho V(x_*(t))$$

over the time unit following the moment t.

Thus, equality (8.25) provides a rule for determining the net value $V(x_*(t))$ of capital $x_*(t)$: the value $V(x_*(t))$ of capital $x_*(t)$ at every moment $t \geq 0$ is equal to the amount of money that, being put into a bank account at interest rate ρ, yields, in the following time unit, a profit equal to the total current profit of the enterprise over this time unit.

This is the economic interpretation of the stationarity condition (8.20).

9. Case of dominating discount

In this section, we derive a normal-form version of the Pontryagin maximum principle in the situation when the discount factor $e^{-\rho t}$ in the functional (5.3) suppresses the growth of the gradient $\partial g(\cdot,u)/\partial x$ of function $g(\cdot,u)$ for any $u \in U$, and the growth of admissible trajectories $x(\cdot)$ of the control system (5.1) and of the trajectories of the corresponding linearized system of differential equations. We call this situation a case of dominating discount.

For linear control system (5.3) the case of dominating discount was considered first in [**7**]. In general nonlinear situation this case was investigated in [**6**].

The following condition characterizes the growth of the gradient $\partial g(\cdot,u)/\partial x$ of function $g(\cdot,u)$ for any $u \in U$.

(A7) *There exist constants $\kappa \geq 0$ and $r \geq 0$ such that*

$$\left\| \frac{\partial g(x,u)}{\partial x} \right\| \leq \kappa(1 + \|x\|^r) \quad \textit{for any} \quad x \in G \quad \textit{and} \quad u \in U.$$

For a given admissible pair $(x(\cdot),u(\cdot))$ of the control system (5.1), we denote by $Y_{(x(\cdot),u(\cdot))}(\cdot)$ the normalized fundamental matrix of the linear system of differential equations (linearized system)

$$\dot{y} = \frac{\partial f(x(t),u(t))}{\partial x} y. \tag{9.1}$$

Recall that an $n \times n$ matrix function $Y_{(x(\cdot),u(\cdot))}(\cdot)$ is called the normalized fundamental matrix of the homogeneous system of linear differential equations (9.1) if the columns $y_i(\cdot)$, $i = 1,\dots,n$, of this matrix are (Carathéodory) solutions to system (9.1) with the initial conditions $y_i^j(0) = \delta_{i,j}$, $i,j = 1,\dots,n$, where $\delta_{i,i} = 1$ for $i = 1,\dots,n$ and $\delta_{i,j} = 0$ for any $i \neq j$.

Analogously denote by $Z_{(x(\cdot),u(\cdot))}(\cdot)$ the normalized fundamental matrix of the corresponding adjoint system

$$\dot{z} = -\left[\frac{\partial f(x(t),u(t))}{\partial x}\right]^* z.$$

Here, we assume that the matrix functions $Y_{(x(\cdot),u(\cdot))}(\cdot)$ and $Z_{(x(\cdot),u(\cdot))}(\cdot)$ are considered on the whole infinite time interval $[0,\infty)$.

The matrix functions $Y_{(x(\cdot),u(\cdot))}(\cdot)$ and $Z_{(x(\cdot),u(\cdot))}(\cdot)$ are related. In particular, it follows directly from definitions of the matrix functions $Y_{(x(\cdot),u(\cdot))}(\cdot)$

and $Z_{(x(\cdot),u(\cdot))}(\cdot)$ that if $y_i(\cdot)$, $i \in \{1,\dots,n\}$, is an arbitrary ith column of the matrix $Y_{(x(\cdot),u(\cdot))}(\cdot)$ and $z_j(\cdot)$, $j \in \{1,\dots,n\}$, is an arbitrary jth column of the matrix $Z_{(x(\cdot),u(\cdot))}(\cdot)$, then

$$\langle y_i(t), z_j(t)\rangle \equiv \delta_{i,j}$$

on the interval $[0,\infty)$. This immediately implies the equality

$$[Z_{(x(\cdot),u(\cdot))}(t)]^{-1} = [Y_{(x(\cdot),u(\cdot))}(t)]^* \quad \text{for any} \quad t \in [0,T]. \tag{9.2}$$

Recall that the normal-form Hamilton–Pontryagin function $\mathcal{H}\colon G\times[0,\infty)\times U \times R^n \mapsto R^1$ and the normal-form Hamiltonian $H : G \times [0,\infty) \times R^n \mapsto R^1$ were defined in Section 6 as follows:

$$\mathcal{H}(x,t,u,\psi) = \langle f(x,u),\psi\rangle + e^{-\rho t} g(x,u),$$

$$H(x,t,\psi) = \sup_{u\in U} \mathcal{H}(x,t,u,\psi)$$

for any $x \in G$, $t \in [0,\infty)$, $u \in U$, and $\psi \in R^n$.

Next, denote by $\|A\|$ the standard norm of an $n \times n$ matrix A considered as a linear operator in R^n; i.e., we set

$$\|A\| = \max_{x\in R^n,\ x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

The following condition characterize the growth of trajectories of systems (5.1) and (9.1):

(A8) *There exist numbers $\lambda \in R^1$, $C_1 \geq 0$, $C_2 \geq 0$, and $C_3 \geq 0$ such that any admissible pair $(x(\cdot),u(\cdot))$ of system (5.1) satisfies the conditions*

$$\|x(t)\| \leq C_1 + C_2 e^{\lambda t} \quad \textit{for any} \quad t \geq 0$$

and

$$\left\|Y_{(x(\cdot),u(\cdot))}(t)\right\| \leq C_3 e^{\lambda t} \quad \textit{for any} \quad t \geq 0.$$

Note that in the case of a convex set G, condition (A7) implies that there exist constants $C_4 \geq 0$ and $C_5 \geq 0$ such that any admissible pair $(x(\cdot),u(\cdot))$ of the control system (5.1) satisfies the inequality

$$|g(x(t),u(t))| \leq C_4 + C_5\|x(t)\|^{r+1} \quad \text{for any} \quad t \geq 0. \tag{9.3}$$

Indeed, in this case, we have

$$g(x(t),u(t)) - g(x_0,u(t)) = \int_0^1 \Big\langle \frac{\partial}{\partial x} g\big(x_0 + s(x(t)-x_0), u(t)\big),\ x(t)-x_0 \Big\rangle\, ds$$

for any $t \geq 0$. Since the set U is compact, this implies inequality (9.3) provided that condition (A7) holds.

Next, condition (9.3) and the first inequality in (A8) imply the inequality

$$e^{-\rho t}|g(x(t),u(t))| \leq C_6 e^{-\rho t} + C_7 e^{-(\rho-(r+1)\lambda)t} \quad \text{for any} \quad t \geq 0$$

along any admissible pair $(x(\cdot), u(\cdot))$ of system (5.1) (with constants $C_6 \geq 0$ and $C_7 \geq 0$ independent of the pair $(x(\cdot), u(\cdot))$).

So, if the set G is convex and $\rho > 0$, conditions (A7), (A8) and the inequality $\rho > (r+1)\lambda$ imply condition (A6).

The following condition is called a dominating discount condition:

(A9) *The following inequality holds:*

$$\rho > (r+1)\lambda.$$

Condition (A9) means that the discount parameter ρ in the goal functional (5.3) dominates the growth parameters r and λ (see conditions (A7) and (A8)).

The proof of the following version of the Pontryagin maximum principle for problem (P) is based on Theorem 17 obtained in Section 6 above.

Theorem 20. *Suppose that for any $T > 0$ conditions* (A1), (A2) *take place, and conditions* (A5)–(A8) *are valid as well. Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible pair in problem (P). Then, there exists an adjoint variable $\psi(\cdot)$ (corresponding to the pair $(x_*(\cdot), u_*(\cdot))$) such that the following conditions hold*:

1) *the optimal admissible pair $(x_*(\cdot), u_*(\cdot))$, together with the adjoint variable $\psi(\cdot)$, satisfies the core relations of the Pontryagin maximum principle in the normal form on the infinite time interval $[0, \infty)$:*

$$\dot{\psi}(t) \stackrel{\text{a.e.}}{=} -\left[\frac{\partial f(x_*(t), u_*(t))}{\partial x}\right]^* \psi(t) - e^{-\rho t}\frac{\partial g(x_*(t), u_*(t))}{\partial x}, \tag{9.4}$$

$$\mathcal{H}(x_*(t), t, u_*(t), \psi(t)) \stackrel{\text{a.e.}}{=} H(x_*(t), t, \psi(t)); \tag{9.5}$$

2) *the optimal admissible pair $(x_*(\cdot), u_*(\cdot))$, together with the adjoint variable $\psi(\cdot)$, satisfies the normal-form stationarity condition*

$$H(x_*(t), t, \psi(t)) = \rho \int_t^\infty e^{-\rho s} g(x_*(x), u_*(s))\, ds \quad \textit{for any} \quad t \geq 0; \tag{9.6}$$

3) *for any $t \geq 0$, the integral*

$$I_*(t) = \int_t^\infty e^{-\rho s}[Z_*(s)]^{-1}\frac{\partial g(x_*(s), u_*(s))}{\partial x}\, ds, \tag{9.7}$$

where $Z_*(\cdot) = Z_{(x_*(\cdot), u_*(\cdot))}(\cdot)$, converges absolutely and the following equality holds:

$$\psi(t) = Z_*(t) I_*(t). \tag{9.8}$$

Proof. Let $\{(P_k)\}$, $k = 1, 2, \ldots$, be a sequence of approximating optimal control problems that corresponds to the control $u_*(\cdot)$ (see Section 6), and let $(x_k(\cdot), v_k(\cdot))$ be an optimal admissible pair in problem (P_k) (on the time interval $[0, T_k]$). According to the classical Pontryagin maximum principle, for any $k = 1, 2, \dots$ there exists an adjoint variable $\psi_k(\cdot)$ (corresponding to the pair $(x_k(\cdot), v_k(\cdot))$) such that the pair $(x_k(\cdot), v_k(\cdot))$, together with the adjoint variable $\psi_k(\cdot)$, satisfies the core relations of the Pontryagin maximum principle in the normal form for problem (P_k) (see (6.25), (6.26)):

$$\dot{\psi}_k(t) \overset{\text{a.e.}}{=} -\left[\frac{\partial f(x_k(t), u_k(t))}{\partial x}\right]^* \psi_k(t) - e^{-\rho t}\frac{\partial g(x_k(t), u_k(t))}{\partial x}, \tag{9.9}$$

$$\mathcal{H}_k(x_k(t), t, u_k(t), \psi_k(t)) \overset{\text{a.e.}}{=} H_k(x_k(t), t, \psi_k(t)).$$

Moreover, the following transversality condition holds (see (6.27)):

$$\psi_k(T_k) = 0. \tag{9.10}$$

Henceforth, we assume that the adjoint variable $\psi_k(\cdot)$ is continuously extended by zero to the entire infinite interval $[0, \infty)$ for any $k = 1, 2, \dots$; i.e., $\psi_k(t) = 0$ for any $t \geq T_k$.

Denote by $Z_k(\cdot)$ and $Y_k(\cdot)$, $k = 1, 2, \dots$, the normalized fundamental matrices of the homogeneous systems of linear differential equations

$$\dot{z} = -\left[\frac{\partial f(x_k(t), u_k(t))}{\partial x}\right]^* z$$

and

$$\dot{y} = \left[\frac{\partial f(x_k(t), u_k(t))}{\partial x}\right] y$$

on the interval $[0, T_k]$. By virtue of (A8), we have

$$\|Y_k(t)\| \leq C_3 e^{\lambda t} \quad \text{for any} \quad t \geq 0.$$

Recall that the Hamilton–Pontryagin function $\mathcal{H}_k(\cdot, \cdot, \cdot, \cdot)$ and the Hamiltonian $H_k(\cdot, \cdot, \cdot)$ for problem (P_k), $k = 1, 2, \dots$, are defined by the equalities (see Section 6)

$$\mathcal{H}_k(x, t, u, \psi) = \langle f(x, u), \psi \rangle + e^{-\rho t}\left[g(x, u) - e^{-t}\frac{\|u - z_k(t)\|^2}{1 + \sigma_k}\right],$$

$$H_k(x, t, \psi) = \sup_{u \in U} \mathcal{H}_k(x, t, u, \psi)$$

for any $x \in G$, $t \in [0, \infty)$, $u \in U$, and $\psi \in R^n$. Moreover, for any $k = 1, 2, \dots$ (see (4.35)), we have the following relation on the interval $[0, T_k]$:

$$\frac{d}{dt} H_k(x_k(t), t, \psi_k(t)) \overset{\text{a.e.}}{=} \frac{\partial}{\partial t}\mathcal{H}_k(x_k(t), t, u_k(t), \psi_k(t)).$$

Let us show that the sequence $\{\psi_k(0)\}$, $k = 1, 2, \dots$, is bounded.

Using the standard representation of the solution $\psi_k(\cdot)$ to the adjoint system (9.9) with the boundary condition (9.10) in terms of the matrix $Z_k(\cdot) = Z_{(x_k,v_k)}(\cdot)$ by the Cauchy formula (7.9) , we obtain

$$\psi_k(0) = \int_0^{T_k} e^{-\rho s}[Z_k(s)]^{-1}\left[\frac{\partial g(x_k(t), u_k(t))}{\partial x}\right] ds.$$

Hence, by (9.2) we have

$$\|\psi_k(0)\| \le \int_0^{T_k} e^{-\rho s}\|Y_k(s)\|\left\|\frac{\partial g(x_k(t), u_k(t))}{\partial x}\right\| ds$$

and, by conditions (A7) and (A6), we obtain

$$\|\psi_k(0)\| \le \int_0^{T_k} \left(C_8 e^{-(\rho-\lambda)s} + C_9 e^{-(\rho-(r+1)\lambda)s}\right) ds,$$

where $C_8 \ge 0$ and $C_9 \ge 0$ are independent of $k = 1, 2, \ldots$. This and condition (A9) imply that the sequence $\{\psi_k(0)\}$, $k = 1, 2, \ldots$, is bounded.

Thus, the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \ldots$, satisfies the conditions of Theorem 17. By Theorem 17, there exists a subsequence of the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \ldots$ (henceforth, we will denote it again by $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \ldots$), such that, for any $T > 0$, conditions (6.28), (6.29) hold for the admissible pairs $(x_k(\cdot), u_k(\cdot))$, $k = 1, 2, \ldots$, in problems (P_k) and condition (6.30) holds for the adjoint variables $\psi_k(\cdot)$, $k = 1, 2, \ldots$. Moreover, in the latter condition, the limit function $\psi(\cdot)$ is an adjoint variable corresponding to the pair $(x_*(\cdot), u_*(\cdot))$ in problem (P) (see (6.31)–(6.33)). Thus, the optimal pair $(x_*(\cdot), u_*(\cdot))$, together with the adjoint variable $\psi(\cdot)$, satisfies the core relations (9.4) and (9.5) of the Pontryagin maximum principle in the normal form for problem (P). Finally, the pair $(x_*(\cdot), u_*(\cdot))$ and the variable $\psi(\cdot)$ satisfy the stationarity condition (9.6).

Thus, assertions 1) and 2) are proved.

Let us prove assertion 3). Consider the integral $I_*(t)$ (see (9.7)) for arbitrary $t \ge 0$.

First, notice that conditions (6.28) and (6.29) imply the convergence

$$Z_k(s) \to Z_*(s) \quad \text{for any} \quad s \ge 0. \tag{9.11}$$

Hence,

$$\begin{aligned} I_*(t) &= \lim_{T\to\infty} \int_t^T e^{-\rho s}[Z_*(s)]^{-1}\frac{\partial g(x_*(s), u_*(s))}{\partial x} ds \\ &= \lim_{T\to\infty} \lim_{k\to\infty} \int_t^T e^{-\rho s}[Z_k(s)]^{-1}\left[\frac{\partial g(x_k(s), u_k(s))}{\partial x}\right] ds \end{aligned}$$

if all the limits in this formula exist. Let us show that this is indeed the case. Conditions (9.2), (A7), and (A8) imply that the following inequality holds for any $s \geq 0$:

$$e^{-\rho s} \left\| [Z_k(s)]^{-1} \right\| \left\| \left[\frac{\partial g(x_k(s), u_k(s))}{\partial x} \right] \right\| \leq C_{10} e^{-(\rho-\lambda)s} + C_{11} e^{-(\rho-(r+1)\lambda)s}$$

with some constants $C_{10} \geq 0$ and $C_{11} \geq 0$ that are independent of $k = 1, 2, \dots$. Hence, by the domination condition (A9) and the Lebesgue theorem, we have

$$\int_t^T e^{-\rho s} [Z_*(s)]^{-1} \frac{\partial g}{\partial x}(x_*(s), u_*(s))\, ds = \lim_{k\to\infty} \int_t^T e^{-\rho s} [Z_k(s)]^{-1} \left[\frac{\partial g(x_k(s), u_k(s))}{\partial x} \right] ds.$$

Thus, we see that the integral

$$I_*(t) = \int_t^\infty e^{-\rho s} [Z_*(s)]^{-1} \frac{\partial g(x_*(s), u_*(s))}{\partial x}\, ds$$

converges absolutely.

Now, let us prove (9.8). To this end, we integrate the adjoint equation for $\psi_k(\cdot)$ (see (9.9)) on the interval $[t, T_k]$ for sufficiently large k (i.e., for $T_k \geq t$). Taking into account the transversality condition (9.10), we obtain

$$\psi_k(t) = Z_k(t) \int_t^{T_k} e^{-\rho s} Z_k^{-1}(s) \left[\frac{\partial g(x_k(t), u_k(s))}{\partial x} \right] ds. \tag{9.12}$$

By conditions (6.28) and (6.29) (which hold for any fixed $T > 0$), we have

$$\frac{\partial g(x_k(s), u_k(s))}{\partial x} \to \frac{\partial g(x_*(s), u_*(s))}{\partial x} \quad \text{for a.e.} \quad s \geq 0.$$

These convergence, conditions (9.11) and (6.30), and the absolute convergence of the integral $I_*(t)$ imply the required equality (9.8) as the limit of equality (9.12) as $k \to \infty$. Assertion 3) is proved. The theorem is proved. ■

Recall a few concepts and results of stability theory (see [**14**], [**17**] for more details).

Consider a system of linear differential equations

$$\dot{x} = A(t)x. \tag{9.13}$$

Here, $t \in [0, \infty)$, $x \in R^n$, and all entries of the real $n \times n$ matrix A are bounded measurable functions.

Let $x(\cdot)$ be a nonzero solution to the differential system (9.13). Then, the number

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

is called the characteristic Lyapunov exponent or, briefly, the characteristic exponent of the solution $x(\cdot)$. Note that the characteristic exponent λ differs by the sign from the characteristic number introduced by A.M. Lyapunov (see [**29**]).

It is well known that the characteristic exponent λ of any nonzero solution x of system (9.13) is finite. The set of characteristic exponents corresponding to all nonzero solutions of system (9.13) is called the *spectrum* of system (9.13). The spectrum of system (9.13) always consists of at most n different numbers.

The solutions of the system of linear differential equations (9.13) form a finite-dimensional linear space of dimension n. Any basis of this space, i.e., any n linearly independent solutions $x_1, \dots, x_n$, is called a fundamental system of solutions of system (9.13). A fundamental system of solutions $x_1, \dots, x_n$ is said to be normal if the sum of characteristic exponents of these solutions $x_1, \dots, x_n$ is minimal among all fundamental systems of solutions of the differential system (9.13).

It turns out that a normal fundamental system of solutions for (9.13) always exists. If $x_1(\cdot), \dots, x_n(\cdot)$ is such a fundamental system of solutions, then the characteristic exponents of the solutions $x_1, \dots, x_n$ cover the entire spectrum of the system (9.13). This means that for any number λ in the spectrum of system (9.13), there necessarily exists a solution of this system from the set $x_1(\cdot), \dots, x_n(\cdot)$ that has this number as its characteristic exponent. Note that different solutions $x_j(\cdot)$ and x_k in a fundamental system $x_1(\cdot), \dots, x_n(\cdot)$ may have the same characteristic exponent.

Let

$$\sigma = \sum_{s=1}^{l} n_s \lambda_s$$

be the sum of all numbers $\lambda_1, \dots, \lambda_l$ of the spectrum of the differential system (9.13), counted with their multiplicities n_s, $s = 1, \dots, l$.

The linear system of differential equations (9.13) 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 trace of the matrix $A(s)$ (i.e., the sum of all its elements that lie on the principal diagonal).

Let $X(\cdot)$ be a fundamental matrix of the linear system of differential equations (9.13); i.e., the columns of the matrix $X(\cdot)$ form a fundamental system of solutions. Then, the Cauchy matrix of system (9.13) is an $n \times n$ matrix function $K(\cdot, \cdot)$ given by

$$K(s,t) = X(s)[X(t)]^{-1} \quad \text{for any} \quad s \geq 0, \quad t \geq 0.$$

The Cauchy matrix $K(\cdot, \cdot)$ of system (9.13) is defined uniquely; it does not depend on what fundamental matrix $X(\cdot)$ is used in its definition.

If the differential system (9.13) is regular, then, for any $\varepsilon > 0$, its Cauchy matrix $K(\cdot,\cdot)$ satisfies the following inequality:

$$\|K(s,t)\| \le \kappa_1 e^{\bar{\lambda}(s-t)+\varepsilon s} \quad \text{for any} \quad t \ge 0 \quad \text{and any} \quad s \ge t, \tag{9.14}$$

where $\bar{\lambda}$ is the maximal element of the spectrum of system (9.13) and the constant $\kappa_1 \ge 0$ depends only on ε. The proofs of this and other results presented above are given in [**14**], [**17**].

Now, we deduce two corollaries from Theorem 20 proved above.

Corollary 1. *Suppose that the conditions of Theorem* 20 *hold and the system of linear differential equations*

$$\dot{y}(t) = \frac{\partial f(x_*(t), u_*(t))}{\partial x} y(t) \tag{9.15}$$

is regular. Let, in addition,

$$\lambda \ge \bar{\lambda},$$

where $\bar{\lambda}$ is the maximal element of the spectrum of system (9.15). *Then, for any $\varepsilon > 0$, the adjoint variable ψ involved in the formulation of Theorem* 20 *satisfies the following inequality*:

$$\|\psi(t)\| \le \kappa_2 \big(e^{-\rho t+\varepsilon t} + e^{-(\rho - r\lambda)t+\varepsilon t}\big) \quad \textit{for any} \quad t \ge 0. \tag{9.16}$$

Here, the constant $\kappa_2 \ge 0$ depends only on ε.

Proof. By conditions (9.2) and (9.8), for any $t \ge 0$ we have

$$\begin{aligned}
\psi(t) &= \int_t^{\infty} e^{-\rho s} \big[[Y_*(t)]^*\big]^{-1} [Y_*(s)]^* \frac{\partial g(x_*(s), u_*(s))}{\partial x}\, ds \\
&= \int_t^{\infty} e^{-\rho s} \big[Y_*(s)[Y_*(t)]^{-1}\big]^* \frac{\partial g(x_*(s), u_*(s))}{\partial x}\, ds \\
&= \int_t^{\infty} e^{-\rho s} [K(s,t)]^* \frac{\partial g(x_*(s), u_*(s))}{\partial x}\, ds,
\end{aligned}$$

where $Y_*(\cdot) = Y_{(x_*(\cdot), u_*(\cdot))}(\cdot)$ is the normalized fundamental matrix of the linear system (9.15) and $K(s,t) = Y_*(s)[Y_*(t)]^{-1}$, $s,t \ge 0$, is the Cauchy matrix of system (9.15).

Hence, by conditions (A7) and (A8) and estimate (9.14), the following inequality holds for any $0 < \varepsilon < \min\{\rho - \lambda, \rho - (r+1)\lambda\}$:

$$\begin{aligned}
\|\psi(t)\| &\le \int_t^{\infty} e^{-\rho s} \big\|[K(s,t)]^*\big\| \left\| \left[\frac{\partial g(x_*(s), u_*(s))}{\partial x} \right] \right\| ds \\
&\le C_{12} \int_t^{\infty} e^{-\rho s} e^{\bar{\lambda}(s-t)} e^{\varepsilon s} \big(1 + e^{r\lambda s}\big)\, ds \\
&\le \kappa_2 \big(e^{-\rho t+\varepsilon t} + e^{-(\rho - r\lambda)t+\varepsilon t}\big).
\end{aligned}$$

Here, $C_{12} \geq 0$ and $\kappa_2 \geq 0$ are constants that depend only on ε. Hence, estimate (9.16) holds for any $\varepsilon > 0$. The corollary is proved. ■

It is easily seen that when $\rho > 0$, the the estimate (9.16) implies that both transversality conditions (see (7.13) and (7.14))

$$\lim_{t\to\infty} \psi(t) = 0 \quad \text{and} \quad \lim_{t\to\infty} \langle x_*(t), \psi(t)\rangle = 0$$

hold.

To conclude this section, consider the situation when the control system (5.1) is linear and autonomous. In this case, the optimal control problem in question is formulated as follows.

Problem $(P2)$:

$$\dot{x} = Ax + u, \quad u \in U, \tag{9.17}$$

$$x(0) = x_0,$$

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

Here, A is an $n \times n$ real constant matrix. The other data in problem $(P2)$ are the same as in the original problem (P).

Let λ_A stand for the maximal real part of the eigenvalues of the matrix A. Then, the number λ_A is the maximal element of the spectrum of the linear homogeneous system

$$\dot{x} = Ax$$

corresponding to the control system (9.17). In addition, the following inequality holds for any $\varepsilon > 0$:

$$\left\|e^{At}\right\| \leq C_{13} e^{(\lambda_A+\varepsilon)t} \quad \text{for any} \quad t \geq 0.$$

Here, e^{At} is the exponential of the matrix A and $C_{13} \geq 0$ is a constant depending only on ε (see, for example, [**17**] for the proof of this fact).

Next, using the standard representation of the solution to the linear system (9.17) in terms of the exponential e^{At} (see, for example, [**24**]), we obtain the following inequality for any $\varepsilon > 0$ and any admissible trajectory x of the control system (9.17):

$$\|x(t)\| \leq C_{14} + C_{15} e^{(\lambda_A+\varepsilon)t} \quad \text{for any} \quad t \geq 0.$$

Here, the constants $C_{14} \geq 0$ and $C_{15} \geq 0$ depend only on ε. Thus, condition (A8) obviously holds for any $\lambda > \lambda_A$.

This observation and the fact that any homogeneous system of linear differential equations with constant coefficients is regular (see [**17**]) yield the following version of Corollary 1 above for problem $(P2)$.

Corollary 2. *Suppose that for any $T > 0$ conditions* (A1), (A2) *take place, conditions* (A5), (A6) *are also valid, and, in addition, $\rho > (r+1)\lambda_A$. Let $(x_*(\cdot), u_*(\cdot))$ be an optimal admissible pair in problem $(P2)$. Then, all the conditions of Theorem* 20 *are satisfied, and the pair $(x_*(\cdot), u_*(\cdot))$ satisfies the conditions of Theorem* 20 *with some adjoint variable $\psi(\cdot)$. Moreover, for any $\varepsilon > 0$, the adjoint variable $\psi(\cdot)$ satisfies the inequality*

$$\|\psi(t)\| \leq \kappa_3\big(e^{-\rho t+\varepsilon t} + e^{-(\rho - r\lambda_A)t+\varepsilon t}\big) \quad \textit{for any} \quad t \geq 0, \tag{9.18}$$

where the constant $\kappa_3 \geq 0$ depends only on ε.

Note that when $\rho > 0$, estimate (9.18) implies that both transversality conditions (see (7.13) and (7.14))

$$\lim_{t\to\infty} \psi(t) = 0 \quad \text{and} \quad \lim_{t\to\infty} \langle x_*(t), \psi(t)\rangle = 0$$

hold.

10. Monotonous case

In this section we prove another normal-form version of the Pontryagin maximum principle in the situation when a monotonicity condition holds (see [**4**]–[**6**]). In this case we show that all coordinates of the adjoint variable $\psi(\cdot)$ are positive and the stationarity condition for the Hamiltonian takes place together with the core relations of the maximum principle. This version of the maximum principle turn out to be especially effective in applications to optimal economic growth problems.

In what follows, we write $a > 0$ (respectively, $a \geq 0$) if all coordinates of the vector $a \in R^n$ are positive (respectively, nonnegative). Similarly, the relation $A > 0$ (respectively, $A \geq 0$) for an $n \times n$ matrix A means that all components of A are positive (respectively, nonnegative).

The following condition is called a monotonicity condition.

(**A10**) *The following inequalities hold for any admissible trajectory* $x(\cdot)$ *of the control system* (5.1) *and any vector* $u \in U$*:*

$$\frac{\partial g(x(t),u)}{\partial x} > 0, \quad \frac{\partial f^i(x(t),u)}{\partial x^j} \geq 0, \quad i,j = 1,\dots,n, \quad i \neq j, \quad \text{for any } t \geq 0.$$

Note that in the one-dimensional case ($n = 1$), the second inequality in condition (A10) holds automatically.

The monotonicity condition (A10) and a number of other positivity conditions are characteristic of many optimal economic growth problems. Indeed, in optimal economic growth problems, the coordinates of the phase vector x represent, as a rule, various (positive) factors of production. Moreover, it is usually assumed that the instantaneous utility in such problems and, sometimes, the growth velocities of all factors of production increase as the factors of production themselves increase. Within the framework of problem (P), assumptions of this kind imply that the integrand $g(\cdot,u)$ in the goal functional (5.3) monotonically increases, while the coordinates of the vector function $f(\cdot,u)$ in the control system (5.1) do not decrease as the coordinates of the phase vector x increase for any $u \in U$. These assumptions imply that condition (A10) holds.

The instantaneous utility function $g(\cdot,\cdot)$ and the growth velocity of the factors of production often turn out to be positive along an optimal pair $(x_*(\cdot),u_*(\cdot))$, which implies that the conditions $g(x_*(t),u_*(t)) > 0$ and

$f(x_*(t), u_*(t)) > 0$ hold for any $t \geq 0$. These conditions, just as the above-mentioned positivity of the vector x, appear (in various combinations) in the statements of many optimal economic growth problems.

Recall that the normal-form Hamilton–Pontryagin function $\mathcal{H} : G \times [0, \infty) \times U \times R^n \mapsto R^1$ and the normal-form Hamiltonian $H : G \times [0, \infty) \times R^n \mapsto R^1$ were defined in Section 6 as follows:

$$\mathcal{H}(x, t, u, \psi) = \langle f(x, u), \psi \rangle + e^{-\rho t} g(x, u),$$

$$H(x, t, \psi) = \sup_{u \in U} \mathcal{H}(x, t, u, \psi)$$

for any $x \in G$, $t \in [0, \infty)$, $u \in U$, and $\psi \in R^n$.

The following result strengthens the general versions of the Pontryagin maximum principle (Theorems 18) for problem (P) in the case when the additional monotonicity condition (A10) holds.

Theorem 21. *Suppose that for any* $T > 0$ *conditions* (A1), (A2) *take place, and conditions* (A5), (A10) *are valid as well. Moreover, suppose that there exists a vector* $u_0 \in U$ *such that* $f(x_0, u_0) > 0$. *Let* $(x_*(\cdot), u_*(\cdot))$ *be an optimal admissible pair in problem* (P). *Then, there exists an adjoint variable* $\psi(\cdot)$ *(corresponding to the pair* $(x_*(\cdot), u_*(\cdot))$*) such that the following conditions hold:*

1) *the admissible pair* $(x_*(\cdot), u_*(\cdot))$, *together with the adjoint variable* $\psi(\cdot)$ *satisfies the core relations of the Pontryagin maximum principle in the normal form on the infinite time interval* $[0, \infty)$

$$\dot{\psi}(t) \stackrel{\text{a.e.}}{=} -\left[\frac{\partial f(x_*(t), u_*(t))}{\partial x}\right]^* \psi(t) - e^{-\rho t} \frac{\partial g(x_*(t), u_*(t))}{\partial x}, \tag{10.1}$$

$$\mathcal{H}(x_*(t), t, u_*(t), \psi(t)) \stackrel{\text{a.e.}}{=} H(x_*(t), t, \psi(t)); \tag{10.2}$$

2) *for any* $t \geq 0$, *we have*

$$\psi(t) > 0; \tag{10.3}$$

3) *the admissible pair* $(x_*(\cdot), u_*(\cdot))$, *together with the adjoint variable* $\psi(\cdot)$, *satisfies the normal-form stationarity condition:*

$$H(x_*(t), t, \psi(t)) = \rho \int_t^\infty e^{-\rho s} g(x_*(x), u_*(s))\, ds \quad \textit{for any} \quad t \geq 0. \tag{10.4}$$

Proof. Let $\{(P_k)\}$, $k = 1, 2, \ldots$, be an approximating sequence of optimal control problems corresponding to the control $u_*(\cdot)$ (see Section 6) and $(x_k(\cdot), u_k(\cdot))$ be an optimal admissible pair in problem (P_k). According to the standard formulation of the Pontryagin maximum principle in the normal form (i.e., for $\psi_k^0 = 1$), for any $k = 1, 2, \ldots$ there exists an adjoint variable $\psi_k(\cdot)$ corresponding to the optimal pair $(x_k(\cdot), u_k(\cdot))$ in problem (P_k); i.e., the admissible pair $(x_k(\cdot), u_k(\cdot))$, together with the adjoint variable $\psi_k(\cdot)$, satisfies the Pontryagin maximum principle in the normal form (for problem (P_k)) on the time interval $[0, T_k]$ (see (6.25), (6.26)):

$$\dot{\psi}_k(t) \stackrel{\text{a.e.}}{=} -\left[\frac{\partial f(x_k(t), u_k(t))}{\partial x}\right]^* \psi_k(t) - e^{-\rho t}\frac{\partial g(x_k(t), u_k(t))}{\partial x}, \tag{10.5}$$

$$\mathcal{H}_k(x_k(t), t, v_k(t), \psi_k(t)) \stackrel{\text{a.e.}}{=} H_k(x_k(t), t, \psi_k(t)).$$

In addition, the transversality condition (see (6.27))

$$\psi_k(T_k) = 0 \tag{10.6}$$

holds. Here,

$$\mathcal{H}_k(x, t, u, \psi) = \langle f(x, u), \psi\rangle + e^{-\rho t}\Big[g(x, u) - e^{-t}\frac{\|u_i - z_k(t)\|^2}{1 + \sigma_k}\Big],$$

$$H_k(x, t, \psi) = \sup_{uv\in U} \mathcal{H}_k(x, t, u, \psi)$$

for any $x \in G$, $t \in [0, \infty)$, $u \in U$, and $\psi \in R^n$. Moreover, for any $k = 1, 2, \ldots$ (see (6.13)), we have the following relation on the interval $[0, T_k]$:

$$\frac{d}{dt}H_k(x_k(t), t, \psi_k(t)) \stackrel{\text{a.e.}}{=} \frac{\partial}{\partial t}\mathcal{H}_k(x_k(t), t, u_k(t), \psi_k(t)). \tag{10.7}$$

It follows from condition (A10), the adjoint system (10.5), and the transversality condition (10.6) that for any $k = 1, 2, \ldots$ the adjoint variable satisfies the inequality $\psi_k(t) > 0$ for all t in some interval adjoining T_k from the left.

Let us show that, in fact,

$$\psi_k(t) > 0 \quad \text{for any} \quad t \in [0, T_k). \tag{10.8}$$

Suppose the contrary. Then, for some natural k, there exists a $\tau \in [0, T_k)$ such that at least one coordinate of the vector $\psi_k(\tau)$ vanishes. Let ξ be the maximal among all such numbers τ. Then, it is obvious that $\xi < T_k$. Next, let $i \in \{1, 2, \ldots, n\}$ be an index such that $\psi_k^i(\xi) = 0$. Then,

$$\psi_k(t) > 0 \quad \text{for any} \quad t \in (\xi, T_k), \tag{10.9}$$

and by virtue of the adjoint system (10.5), we have

$$\begin{aligned}\dot{\psi}_k^i(t) &\stackrel{\text{a.e.}}{=} -\sum_{j=1}^{n}\frac{\partial f^j(x_k(t), u_k(t))}{\partial x^i}\psi_k^j(t) - e^{-\rho t}\frac{\partial g(x_k(t), u_k(t))}{\partial x^i}\\ &= -\frac{\partial f^i(x_k(t), u_k(t))}{\partial x^i}\psi_k^i(t) - \sum_{j\neq i}\frac{\partial f^j(x_k(t), u_k(t))}{\partial x^i}\psi_k^j(t) - e^{-\rho t}\frac{\partial g(x_k(t), u_k(t))}{\partial x^i}\end{aligned} \tag{10.10}$$

for all $t \in [\xi, T_k]$.

On the finite time interval $[0, T_k]$, all trajectories of the control system (5.1) with the initial condition (5.2) are uniformly bounded (see (2.10)), and the control $u(\cdot)$ takes values in the compact set U. Therefore, since the partial derivative $\partial f^i(\cdot,\cdot)/\partial x^i$, $i=1,2,\dots,n$, is continuous on the set $G\times U$, there exists a constant $C_1\geq 0$ such that

$$\left|\frac{\partial f^i(x_k(t),u)}{\partial x^i}\right|\leq C_1\quad\text{for any}\quad t\in[\xi,T_k],\quad u\in U,\quad i=1,\dots,n.$$

Condition (A10) implies that the measurable function $w:[\xi,T_k]\mapsto R^n$ defined by

$$w(t)=-\sum_{j\neq i}\frac{\partial f^j(x_k(t),u_k(t))}{\partial x^i}\psi_k^j(t)-e^{-\rho t}\frac{\partial g(x_k(t),u_k(t))}{\partial x^i},\quad t\in[\xi,T_k],$$

is negative for a.e. $t\in[\xi,T_k]$. By virtue (10.10), for arbitrary $i=1,2,\dots,n$ the scalar function $\psi_k^i(\cdot)$ satisfies on $[\xi,T_k]$ to the differential inequality

$$\dot{\psi}_k^i\leq C_1\psi_k^i+w(t)$$

with the boundary condition

$$\psi_k^i(\xi)=0.$$

Hence, by the Cauchy formula (7.9) for linear differential equations, we obtain

$$\psi_k^i(t)\leq e^{C_1t}\int_\xi^t e^{-C_1s}w(s)\,ds<0\quad\text{for any}\quad t\in(\xi,T_k),\quad i=1,2,\dots,n,$$

which contradicts inequality (10.9). Hence, condition (10.8) holds.

Now, let us show that the sequence $\{\psi_k(0)\}$, $k=1,2,\dots$, is bounded.

Indeed, by condition (10.7), we have the following relation on the interval $[0,T_k]$:

$$\begin{aligned}\frac{d}{dt}H_k(x_k(t),&t,\psi_k(t))\overset{\text{a.e.}}{=}\frac{\partial}{\partial t}\mathcal{H}_k(x_k(t),t,v_k(t),\psi_k(t))\\&=-\rho e^{-\rho t}g(x_k(t),u_k(t))+(\rho+1)e^{-(\rho+1)t}\frac{\|u_k(t)-z_k(t)\|^2}{1+\sigma_k}\\&\qquad+2e^{-(\rho+1)t}\frac{\langle u_k(t)-z_k(t),\dot{z}_k(t)\rangle}{1+\sigma_k}.\end{aligned}$$

Integrating this equality over the time interval $[0,T_k]$ and taking into account the transversality condition at the right endpoint (10.6), we obtain

$$\begin{aligned}H_k(x_0,&0,\psi_k(0))=e^{-\rho T_k}\max_{u\in U}\Big[g(x_k(T_k),u)-e^{-T_k}\frac{\|u-z_k(T_k)\|^2}{1+\sigma_k}\Big]\\&+\rho\int_0^{T_k}e^{-\rho t}g(x_k(t),u_k(t))\,dt-(\rho+1)\int_0^{T_k}e^{-(\rho+1)t}\frac{\|u_k(t)-z_k(t)\|^2}{1+\sigma_k}\,dt\\&-2\int_0^{T_k}e^{-(\rho+1)t}\frac{\langle u_k(t)-z_k(t),\dot{z}_k(t)\rangle}{1+\sigma_k}\,dt.\end{aligned}$$

This equality, combined with conditions (5.7), (5.8), (6.2), and (6.3), implies that there exists a $C_2 > 0$ such that

$$H_k(x_0, 0, \psi_k(0)) \le C_2 \quad \text{for any} \quad k = 1, 2, \dots .$$

Further, inclusion $u_0 \in U$ implies the inequality

$$\langle f(x_0, u_0), \psi_k(0)\rangle + g(x_0, u_0) - \frac{\|u_0 - z_k(0)\|^2}{1 + \sigma_k} \le H_k(x_0, 0, \psi_k(0)),$$

and so

$$\langle f(x_0, u_0), \psi_k(0)\rangle \le C_2 + |g(x_0, u_0)| + (2|U| + 1)^2,$$

where $|U| = \max_{u \in U} \|u\|$. This inequality, combined with the condition $f(x_0, u_0) > 0$ and condition (10.8), implies that the sequence $\{\psi_k(0)\}$, $k = 1, 2, \dots$, is bounded.

Thus, the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \dots$, satisfies all the conditions of Theorem 17. Hence, there exists a subsequence of the sequence $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \dots$ (which we again denote by $\{(x_k(\cdot), u_k(\cdot), \psi_k(\cdot))\}$, $k = 1, 2, \dots$), such that, for any $T > 0$, the sequence of adjoint variables $\psi_k(\cdot)$, $k = 1, 2, \dots$, converges uniformly on the time interval $[0, T]$ to an adjoint variable $\psi(\cdot)$ corresponding to the optimal pair $(x_*(\cdot), u_*(\cdot))$ in problem (P); i.e., the pair $(x_*(\cdot), u_*(\cdot))$, together with the adjoint variable $\psi(\cdot)$, satisfies the core relations (10.1) and (10.2) of the normal-form Pontryagin maximum principle for problem (P) (see (6.31), (6.32)). Finally, the pair $(x_*(\cdot), u_*(\cdot))$ and the adjoint variable $\psi(\cdot)$ satisfy the stationarity condition (10.4) (see (6.33)). Thus, assertions 1) and 3) of the theorem are proved for the optimal pair $(x_*(\cdot), u_*(\cdot))$ and the adjoint variable $\psi(\cdot)$.

Since the adjoint variables $\psi_k(\cdot)$ are positive on the time intervals $[0, T_k)$, $k = 1, \dots$, the adjoint variable $\psi(\cdot)$ is nonnegative on the whole infinite interval $[0, \infty)$. Let us show that none of the coordinates of the vector function $\psi(\cdot)$ can vanish at any moment $t \ge 0$. Indeed, suppose that such a moment exists. Then, reasoning as above when proving the positivity of the vector functions $\psi_k(\cdot)$ on the intervals $[0, T_k)$, $k = 1, \dots$, we combine condition (A10) and the fact that the adjoint variable $\psi(\cdot)$ is a solution to the adjoint system (10.1) and arrive at a contradiction. Thus, $\psi(t) > 0$ for any $t \ge 0$. Hence, inequality (10.3) takes place. So, assertion 2) is proved. The theorem is proved. ■

Now, let us draw several corollaries to Theorem 21.

Corollary 3. *By virtue of the stationarity condition* 3) *in Theorem* 21*, at the initial moment we have*

$$\langle f(x_0, u_0), \psi(0)\rangle + g(x_0, u_0) \le H(x_0, 0, \psi(0)) = \rho \int_0^\infty e^{-\rho t} g(x_*(t), u_*(t))\, dt.$$

Therefore, it follows from inequality (5.8) *that*

$$\langle f(x_0,u_0),\psi(0)\rangle \le \rho\int_0^\infty e^{-\rho t} g(x_*(t),u_*(t))\,dt - g(x_0,u_0) \le \rho\omega(0) - g(x_0,u_0).$$

Since the coordinates of the vector $f(x_0,u_0)$ *are positive (see the conditions of Theorem* 21*) and the coordinates of the vector* $\psi(0)$ *are also positive* (*condition* 2) *of Theorem* 21), *the last inequality yields an upper bound for the initial state* $\psi(0)$ *of the adjoint variable* $\psi(\cdot)$*.*

Corollary 4. *Suppose that the conditions of Theorem* 21 *hold and, moreover, there exists a vector* $a \in R^n$ *such that all its coordinates are positive* ($a > 0$) *and the following inequality holds starting from some moment* $\tau \ge 0$:

$$f(x_*(t),u_*(t)) \ge a \quad \text{for a.e.} \quad t \ge \tau. \tag{10.11}$$

Then, the adjoint variable ψ *corresponding to the optimal pair* $(x_*(\cdot),u_*(\cdot))$ *by virtue of Theorem* 21 (*i.e., conditions* 1), 2), *and* 3) *hold*) *satisfies the transversality condition*

$$\lim_{t\to\infty} \psi(t) = 0. \tag{10.12}$$

Proof. Let us prove that the variable $\psi(\cdot)$ satisfies condition (10.12). Indeed, conditions (10.4) and (10.11) yield

$$\limsup_{t\to\infty}\langle a,\psi(t)\rangle \le \lim_{t\to\infty}\max_{u\in U}\langle f(x_*(t),u),\psi(t)\rangle = \lim_{t\to\infty} H(x_*(t),t,\psi(t)) = 0;$$

this condition, combined with (10.3), implies that equality (10.12) holds. The corollary is proved. ■

Corollary 5. *Suppose that the conditions of Theorem* 21 *hold. Suppose, in addition, that there exists a* $\theta > 0$ *such that, starting from some moment* $\tau \ge 0$, *we have*

$$x_*(t) \ge 0 \quad \text{for any} \quad t \ge \tau \tag{10.13}$$

and

$$\frac{\partial f(x_*(t),u_*(t))}{\partial x} - \theta I \ge 0 \quad \text{for a.e.} \quad t \ge \tau, \tag{10.14}$$

where I *is the* $n \times n$ *identity matrix. Then, the adjoint variable* $\psi(\cdot)$ *corresponding to the optimal pair* $(x_*(\cdot),u_*(\cdot))$ *by virtue of Theorem 21* (*i.e., conditions* 1), 2), *and* 3) *hold*) *satisfies the transversality condition*

$$\lim_{t\to\infty}\langle x_*(t),\psi(t)\rangle = 0. \tag{10.15}$$

Proof. By virtue of the control system (5.1) and the adjoint system in the normal form (10.5), we have

$$\frac{d}{dt}\langle x_*(t),\psi(t)\rangle \overset{\text{a.e.}}{=} \langle f(x_*(t),u_*(t)),\psi(t)\rangle - \left\langle x_*(t), \left[\frac{\partial f(x_*(t),u_*(t))}{\partial x}\right]^* \psi(t)\right\rangle$$
$$- e^{-\rho t}\left\langle x_*(t), \frac{\partial g(x_*(t),u_*(t))}{\partial x}\right\rangle$$

on the infinite time interval $[\tau,\infty)$.

By conditions (10.13) and (A10), on the interval $[\tau,\infty)$ we have

$$e^{-\rho t}\left\langle x_*(t), \frac{\partial g(x_*(t),u_*(t))}{\partial x}\right\rangle \overset{\text{a.e.}}{\geq} 0.$$

Hence, the following inequality holds on $[\tau,\infty)$:

$$\frac{d}{dt}\langle x_*(t),\psi(t)\rangle \overset{\text{a.e.}}{\leq} -\left\langle x_*(t), \left[\frac{\partial f(x_*(t),u_*(t))}{\partial x}\right]^* \psi(t)\right\rangle$$
$$+ \mathcal{H}(x_*(t),t,u_*(t),\psi(t)) - e^{-\rho t} g(x_*(t),u_*(t)).$$

Taking into account condition (5.7) from (A3) and condition (10.14), we obtain

$$\frac{d}{dt}\langle x_*(t),\psi(t)\rangle \leq -\theta\langle x_*(t),\psi(t)\rangle + \mathcal{H}(x_*(t),t,u_*(t),\psi(t)) + \mu(t)$$

for almost every $t \geq \tau$. Thus, the following inequality holds on $[\tau,\infty)$:

$$\frac{d}{dt}\langle x_*(t),\psi(t)\rangle \overset{\text{a.e.}}{\leq} -\theta\langle x_*(t),\psi(t)\rangle + \alpha(t),$$

where α is a positive function on $[\tau,\infty)$ such that

$$\alpha(t) \geq H(x_*(t),t,\psi(t)) + \mu(t)$$

and

$$\alpha(t) \to 0 \quad \text{as} \quad t \to \infty.$$

Obviously, in view of conditions (10.4) and (5.7), such a function α exists. Then, by (10.13), (10.3), and the Cauchy formula (7.9) for linear systems, we obtain

$$0 \leq \langle x_*(t),\psi(t)\rangle \leq e^{-\theta(t-\tau)}\langle x_*(\tau),\psi(\tau)\rangle + e^{-\theta t}\int_\tau^t e^{\theta s}\alpha(s)\,ds. \tag{10.16}$$

Since

$$e^{-\theta(t-\tau)}\langle x_*(\tau),\psi(\tau)\rangle \to 0 \quad \text{as} \quad t \to \infty,$$

we prove the transversality condition (10.15) if we show that

$$e^{-\theta t}\int_\tau^t e^{\theta s}\alpha(s)\,ds \to 0 \quad \text{as} \quad t \to \infty.$$

Let us prove this fact. Since the function α is positive, it suffices to consider the situation when

$$\int_{\tau}^{t} e^{\theta s}\alpha(s)\,ds \to \infty \quad \text{as} \quad t \to \infty.$$

In this case, by l'Hôpital's rule, we obtain

$$\lim_{t\to\infty} e^{-\theta t}\int_{\tau}^{t} e^{\theta s}\alpha(s)\,ds = \lim_{t\to\infty}\frac{\frac{d}{dt}\int_{\tau}^{t} e^{\theta s}\alpha(s)\,ds}{\frac{d}{dt}e^{\theta t}} = \lim_{t\to\infty}\frac{\alpha(t)}{\theta} = 0.$$

Thus, the transversality condition (10.15) is proved. The corollary is proved. ■

Note that the transversality condition (10.15) is frequently used in sufficient optimality conditions based on the Pontryagin maximum principle (see Section 11).

11. Maximum principle and sufficient conditions of optimality

Here we present a variant of Arrow's theorem (see [**2**] and [**37**]) on sufficient optimality conditions in the form of the Pontryagin maximum principle for the optimal control problem (P). Results of this type (sometimes they are formulated under somewhat different assumptions) are often used when substantiating the optimality of extremals of the Pontryagin maximum principle for optimal economic growth problems.

When proving this variant of Arrow's theorem for problem (P), we mainly follow the study [**37**].

Recall that, the normal-form Hamilton–Pontryagin function $\mathcal{H} : G \times [0,\infty) \times U \times R^n \mapsto R^1$ and the normal form Hamiltonian $H : G \times [0,\infty) \times R^n \mapsto R^1$ are defined in Section 6 as follows:

$$\mathcal{H}(x,t,u,\psi) = \langle f(x,u),\psi\rangle + e^{-\rho t} g(x,u),$$

$$H(x,t,\psi) = \sup_{u\in U} \mathcal{H}(x,t,u,\psi)$$

for any $x \in G$, $t \in [0,\infty)$, $u \in U$ and $\psi \in R^n$.

Theorem 22. *Suppose that for any $T > 0$ conditions* (A1), (A2) *take place, and condition* (A6) *is satisfied as well. Assume that an admissible pair $(x_*(\cdot),u_*(\cdot))$, together with an adjoint variable $\psi(\cdot)$ satisfies the core relations* (6.31) *and* (6.32) *of the Pontryagin maximum principle in the normal form for problem (P). In addition, suppose that the set G is convex and the normal-form Hamiltonian $H(\cdot,t,\psi(t))$ in problem (P) (see Section 6) is a concave function of the variable x on G for any $t \geq 0$. Finally, suppose that the following boundary condition at infinity holds for any admissible trajectory $x(\cdot)$:*

$$\liminf_{t\to\infty} \big\langle \psi(t), x(t) - x_*(t) \big\rangle \geq 0. \tag{11.1}$$

Then, the admissible pair $(x_*(\cdot),u_*(\cdot))$ is optimal in problem (P).

Proof. Suppose that the maximum condition in the normal form (see (6.32))

$$\mathcal{H}(x_*(t),t,u_*(t),\psi(t)) = H(x_*(t),t,\psi(t)) \tag{11.2}$$

holds at a moment $t \geq 0$. For this moment t, consider the normal-form Hamiltonian $H(\cdot,t,\psi(t))$ of problem (P) as a function of the variable x on the open convex set G.

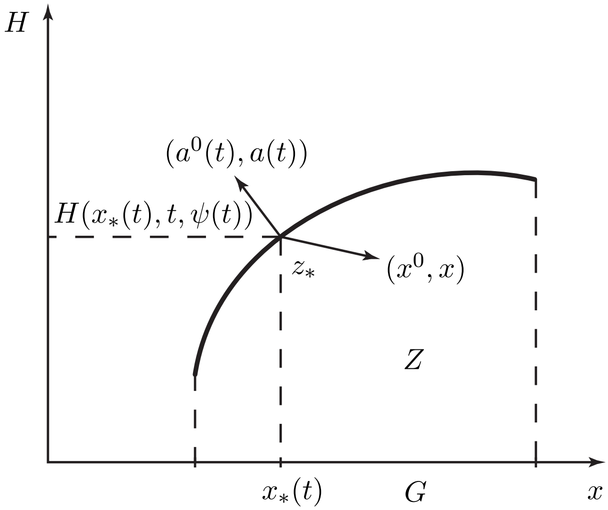

Figure 7: *Point z_* at the boundary of set Z.*

It is easy to see that the point

$$z_* = \big(H(x_*(t), t, \psi(t)), x_*(t)\big) \in R^1 \times R^n$$

belongs to the boundary of the set

$$Z = \big\{(x^0, x) \in R^1 \times R^n : \ x^0 \le H(x, t, \psi(t)),\, x \in G\big\};$$

moreover, since the function $H(\cdot, t, \psi(t))$ is concave in the variable x on the set G, the set Z is convex. Hence, by the separation theorem (see, for example, [**35**]), there exists a nonzero vector $(a^0(t), a(t)) \in R^1 \times R^n$ such that (see Figure 7)

$$a^0(t)x^0 + \langle a(t), x\rangle \le a^0(t)H(x_*(t), t, \psi(t)) + \langle a(t), x_*(t)\rangle \quad \text{for any } (x^0, x) \in Z. \tag{11.3}$$

Since the set G is open, $x_*(t)$ is an interior point of G, and the function $H(\cdot, t, \psi(t))$ is continuous on G, there exists a sufficiently small number $\delta > 0$ such that

$$x_*(t) + \xi \in G \quad \text{and} \quad z_* + (-1, \xi) = \big(H(x_*(t), t, \psi(t)) - 1, x_*(t) + \xi\big) \in Z$$

for any $\|\xi\| \le \delta$. Then, condition (11.3) implies the inequality

$$-a^0(t) + \langle a(t), \xi\rangle \le 0 \quad \text{for any} \quad \|\xi\| \le \delta.$$

This shows that $a^0(t) > 0$. Normalizing the vector $(a^0(t), a(t))$ if necessary, we may assume without loss of generality that $a^0(t) = 1$. Then, by condition (11.3), we obtain

$$\langle a(t), x - x_*(t)\rangle \le H(x_*(t), t, \psi(t)) - H(x, t, \psi(t)) \quad \text{for any} \quad x \in G. \tag{11.4}$$

Since the maximum condition (11.2) holds at the moment t, it follows from (11.4) that

$$\big\langle f(x,u_*(t)) - f(x_*(t),u_*(t)), \psi(t)\big\rangle + e^{-\rho t}g(x,u_*(t)) - e^{-\rho t}g(x_*(t),u_*(t))$$
$$\leq H(x,t,\psi(t)) - H(x_*(t),t,\psi(t)) \leq -\langle a(t), x - x_*(t)\rangle$$

for any $x \in G$. Thus, for the function

$$\phi(x) = \big\langle f(x,u_*(t)) - f(x_*(t),u_*(t)), \psi(t)\big\rangle$$
$$+ e^{-\rho t}g(x,u_*(t)) - e^{-\rho t}g(x_*(t),u_*(t)) + \langle a(t), x - x_*(t)\rangle,$$

which is continuously differentiable on the open set G, we have

$$\phi(x) \leq 0 \quad \text{for any} \quad x \in G \quad \text{and} \quad \phi(x_*(t)) = 0.$$

Therefore, the function $\phi(\cdot)$ attains its absolute maximum at the point $x_*(t)$. Thus, the equality

$$\left[\frac{\partial f(x_*(t),u_*(t))}{\partial x}\right]^* \psi(t) + e^{-\rho t}\frac{\partial g(x_*(t),u_*(t))}{\partial x} + a(t) = 0$$

holds, and so

$$a(t) = -\left[\frac{\partial f(x_*(t),u_*(t))}{\partial x}\right]^* \psi(t) - e^{-\rho t}\frac{\partial g(x_*(t),u_*(t))}{\partial x}. \tag{11.5}$$

Equality (11.5) is obtained under the assumption that the maximum condition (11.2) holds at the moment t. Since the maximum condition (11.2) holds for a.e. $t \geq 0$, equality (11.5) also holds for a.e. $t \geq 0$. By the hypothesis of the theorem, the admissible pair (x_*, u_*), together with the adjoint variable ψ, satisfies the adjoint system in the normal form (6.31). Since (11.5) holds for a.e. $t \geq 0$, it follows that $a(t) = \dot\psi(t)$ for a.e. $t \geq 0$. Hence, condition (11.4) implies that

$$\langle \dot\psi(t), x - x_*(t)\rangle \leq H(x_*(t),t,\psi(t)) - H(x,t,\psi(t)) \quad \text{for any} \quad x \in G$$

for a.e. $t \geq 0$.

Now, let $(x(\cdot), u(\cdot))$ be an arbitrary admissible pair. Then, the above inequality and the maximum condition (11.2) yield

$$\langle \dot\psi(t), x(t) - x_*(t)\rangle \leq H(x_*(t),t,\psi(t)) - H(x(t),t,\psi(t))$$
$$\leq \langle f(x_*(t),u_*(t)), \psi(t)\rangle + e^{-\rho t}g(x_*(t),u_*(t))$$
$$- \langle f(x(t),u(t)), \psi(t)\rangle - e^{-\rho t}g(x(t),u(t))$$

for a.e. $t \geq 0$. Consequently,

$$\frac{d}{dt}\langle \psi(t), x(t) - x_*(t)\rangle + e^{-\rho t}g(x(t),u(t)) \leq e^{-\rho t}g(x_*(t),u_*(t))$$

for a.e. $t \geq 0$. Integrating this inequality over an arbitrary time interval $[0, T]$, $T > 0$, and taking into account the initial condition (5.2), we obtain

$$\int_0^T e^{-\rho t} g(x(t), u(t))\, dt + \langle \psi(T), x(T) - x_*(T) \rangle \leq \int_0^T e^{-\rho t} g(x_*(t), u_*(t))\, dt.$$

Passing to the limit as $T \to \infty$ and applying (11.1), we infer that

$$\int_0^\infty e^{-\rho t} g(x(t), u(t))\, dt \leq \int_0^\infty e^{-\rho t} g(x_*(t), u_*(t))\, dt.$$

The theorem is proved. ■

Note that in some cases the boundary condition (11.1) can be deduced from the transversality conditions at infinity (7.13) or (7.14). In particular, it is obvious that the boundary condition (11.1) certainly holds if the transversality conditions (7.14) hold and, starting from some moment $\tau \geq 0$, all coordinates of the adjoint variable $\psi(t)$ and of any admissible trajectory $x(t)$ are nonnegative for any $t \geq \tau$. Or alternatively, if all admissible trajectories $x(\cdot)$ are bounded (i.e., if any admissible trajectory $x(\cdot)$ satisfies the inequality $\sup_{t \geq 0} \|x(t)\| < \infty$), then condition (11.1) follows from the transversality conditions (7.13).

Theorem 22 and Corollary 5 to Theorem 21 imply the following criterion of optimality of an admissible pair $(x_*(\cdot), u_*(\cdot))$ in problem (P).

Corollary 6. *Suppose that the set G is convex and the function $H(\cdot, t, \psi) : G \mapsto R^1$ is concave in the variable x for any fixed $t \geq 0$ and $\psi > 0$. Moreover, suppose that for any admissible pair $(x(\cdot), u(\cdot))$ of system* (5.1), *there exists a number $\theta > 0$ such that the following conditions hold starting from some moment $\tau \geq 0$:*

$$x(t) \geq 0 \quad \text{for any} \quad t \geq \tau$$

and

$$\frac{\partial f(x(t), u(t))}{\partial x} - \theta I \geq 0 \quad \text{for a.e.} \quad t \geq \tau$$

(see (10.13) *and* (10.14)*). Then, the conditions of Theorem* 21 *are necessary and sufficient optimality conditions in problem (P).*

12. Example

Here we consider the neoclassical model of optimal economic growth with the logarithmic instantaneous utility function. The presented study of this model is based on application of the variant of the Pontryagin maximum principle for infinite-horizon optimal controls developed in Section 10 in the situation when the monotonicity condition (A10) takes place.

The neoclassical model of optimal economic growth with logarithmic instantaneous utility function is formulated as the following optimal control problem (P_ε), $0 < \varepsilon < 1$ (see Section 1 for more details):

$$\dot{K} = uF(K, L) - \delta K, \quad u \in U_\varepsilon = [0, 1 - \varepsilon],$$

$$\dot{L} = \mu L,$$

$$K(0) = K_0, \quad L(0) = L_0,$$

$$J(K, L, u) = \int_0^\infty e^{-\rho t} \big[\ln(1 - u) + \ln F(K, L)\big]\, dt \to \max.$$

Here, $K \in R^1$ and $L \in R^1$ are phase variables (capital and labor); $\mu > 0$; $\delta \geq 0$, $\rho > 0$; $K_0 > 0$ and $L_0 > 0$ are given initial states of the system; and $F(\cdot, \cdot)$ is a production function satisfying the neoclassical conditions (1.39)–(1.42) and condition (1.43). The maximum in problem (P_ε) is sought in the class of measurable functions $u : [0, \infty) \mapsto U_\varepsilon$.

It is easy to see that introducing new state variables

$$\tilde{K}(t) = e^{\delta t} K(t), \quad \tilde{L}(t) = e^{\delta t} L(t), \quad t \geq 0$$

due to condition (1.43) one can rewrite the problem (P_ε), $0 < \varepsilon < 1$ in terms of variables $(\tilde{K}(\cdot), \tilde{L}(\cdot))$ in the following equivalent form:

$$\dot{\tilde{K}} = uF(\tilde{K}, \tilde{L}), \quad u \in U_\varepsilon = [0, 1 - \varepsilon], \tag{12.1}$$

$$\dot{\tilde{L}} = (\mu + \delta)\tilde{L}, \tag{12.2}$$

$$\tilde{K}(0) = K_0, \quad \tilde{L}(0) = L_0, \tag{12.3}$$

$$J(\tilde{K}, \tilde{L}, u) = \int_0^\infty e^{-\rho t} \big[\ln(1 - u) + \ln F(\tilde{K}, \tilde{L})\big]\, dt \to \max. \tag{12.4}$$

So, below we identify the problem (P_ε), $0 < \varepsilon < 1$, with optimal control problem (12.1)–(12.4).

Conditions (1.39)–(1.42) and the positivity of the production function $F(\cdot,\cdot)$ imply that for arbitrary admissible control $u(\cdot)$, the corresponding trajectory $(\tilde{K}(\cdot),\tilde{L}(\cdot))$ of the control system (12.1), (12.2) with the initial conditions (12.3) is defined on $[0,\infty)$ and lies in the set

$$G=\big\{(\tilde{K},\tilde{L})\in R^2:\ \tilde{K}>0,\ \tilde{L}>0\big\}.$$

It is easily seen that problem (P_ε) (see (12.1)–(12.4)) satisfies conditions (A1), (A2) and (A5), (A6). Indeed, condition (A1) is obviously satisfied and the boundedness condition (A2) holds for problem (P_ε) by conditions (1.39)–(1.42). The control system (12.1), (12.2) is affine in control, the set U_ε is convex, and the instantaneous utility function

$$g(\tilde{K},\tilde{L},u)=\ln(1-u)+\ln F(\tilde{K},\tilde{L}),\quad \tilde{K}>0,\quad \tilde{L}>0,\quad u\in U_\varepsilon,$$

is concave with respect to the variable u. Hence, problem (P_ε) satisfies condition (A5). Next, in view of the neoclassical conditions (1.39)–(1.42), the production function $F(\tilde{K}(\cdot),\tilde{L}(\cdot))$ does not decrease on $[0,\infty)$ along any admissible trajectory $(\tilde{K}(\cdot),\tilde{L}(\cdot))$, and its growth can be uniformly estimated by an exponential function. Therefore, the function $g(\cdot,\cdot)$ is of at most polynomial growth. Hence, there exists a positive integrable function $\tilde{\mu}(\cdot)$ on $[0,\infty)$ such that $\tilde{\mu}(t)\to 0+$ as $t\to\infty$ and the following condition holds for any admissible trajectory $(\tilde{K}(\cdot),\tilde{L}(\cdot))$:

$$e^{-\rho t}\max_{u\in U_\varepsilon}\big|g\big(\tilde{K}(t),\tilde{L}(t),u\big)\big|\le e^{-\rho t}\big[|\ln\varepsilon|+|\ln F(\tilde{K}(t),\tilde{L}(t))|\big]\le\tilde{\mu}(t)$$

for any $t\ge 0$. Thus, condition (A6) also holds.

By Theorem 15, there exists an optimal admissible control $u_*(\cdot)$ in problem (P_ε). Let $(\tilde{K}_*(\cdot),\tilde{L}_*(\cdot))$ be the admissible trajectory corresponding to the control $u_*(\cdot)$.

The neoclassical conditions (1.39)–(1.42) and the fact that the production function $F(\cdot,\cdot)$ is positive imply

$$\frac{\partial\ln F(\tilde{K},\tilde{L})}{\partial K}=\frac{1}{F(\tilde{K},\tilde{L})}\frac{\partial F(\tilde{K},\tilde{L})}{\partial K}>0\quad\text{for any}\quad\tilde{K}>0,\quad\tilde{L}>0,$$

$$\frac{\partial\ln F(\tilde{K},\tilde{L})}{\partial L}=\frac{1}{F(\tilde{K},\tilde{L})}\frac{\partial F(\tilde{K},\tilde{L})}{\partial L}>0\quad\text{for any}\quad\tilde{K}>0,\quad\tilde{L}>0.$$

For the vector $u_0=1-\varepsilon\in U_\varepsilon$, the right-hand side of the control system (12.1), (12.2) is positive at the initial point (K_0,L_0).

Thus, problem (P_ε), $0<\varepsilon<1$, satisfies the conditions of Theorem 21.

Further, it is easy to see that the neoclassical conditions (1.39)–(1.42) and condition (1.43) imply that for any admissible trajectory $(\tilde{K}(\cdot),\tilde{L}(\cdot))$ there is a constant $C_1 > 0$ such that

$$\frac{\tilde{K}(t)}{\tilde{L}(t)} \le C_1 \quad \text{for all} \quad t \ge 0.$$

Hence for any admissible trajectory $(\tilde{K}(\cdot), \tilde{L}(\cdot))$ there is a constant $C_2 > 0$ such that

$$\frac{\partial}{\partial K} F\left(\tilde{K}(t), \tilde{L}(t)\right) = \frac{\partial}{\partial K} F\left(\frac{\tilde{K}(t)}{\tilde{L}(t)}, 1\right) \ge C_2 \quad \text{for all} \quad t \ge 0.$$

Thus, if the inequality $u_*(t) \overset{\text{a.e.}}{\ge} \theta$ with a positive number θ holds for the optimal control $u_*(\cdot)$ starting from some moment $\tau \ge 0$, then this control $u_*(\cdot)$ satisfies the conditions of Corollary 5 to Theorem 21. So, in this case the adjoint variable $\psi(\cdot) = (\psi^1(\cdot), \psi^2(\cdot))$ corresponding to the optimal process $(\tilde{K}_*(\cdot), \tilde{L}_*(\cdot), u_*(\cdot))$ by virtue of the maximum principle in the normal form (Theorem 21) satisfies the transversality condition (10.15).

Applying Theorem 21 and Corollary 5, we obtain the following version of the Pontryagin maximum principle for problem (P_ε), $0 < \varepsilon < 1$.

Theorem 23. *Let $u_*(\cdot)$ be an optimal control in problem (P_ε), $0 < \varepsilon < 1$, and $(\tilde{K}_*(\cdot), \tilde{L}_*(\cdot))$ be the corresponding optimal trajectory. Then, there exists an adjoint variable $\psi(\cdot) = (\psi^1(\cdot), \psi^2(\cdot))$ (corresponding to the process $(\tilde{K}_*(\cdot), \tilde{L}_*(\cdot), u_*(\cdot))$) such that the following conditions hold:*

1) *the optimal process $(\tilde{K}_*(\cdot), \tilde{L}_*(\cdot), u_*(\cdot))$, together with the adjoint variable $\psi(\cdot) = (\psi^1(\cdot), \psi^2(\cdot))$, satisfies the core relations of the Pontryagin maximum principle in the normal form on the infinite time interval $[0, \infty)$ (see (10.1) and (10.2)):*

$$\begin{aligned} \dot{\psi}^1(t) &\overset{\text{a.e.}}{=} -u_*(t) \frac{\partial F(\tilde{K}_*(t), \tilde{L}_*(t))}{\partial K} \psi^1(t) \\ &\qquad - \frac{e^{-\rho t}}{F(\tilde{K}_*(t), \tilde{L}_*(t))} \frac{\partial F(\tilde{K}_*(t), \tilde{L}_*(t))}{\partial K}, \\ \dot{\psi}^2(t) &\overset{\text{a.e.}}{=} -u_*(t) \frac{\partial F(\tilde{K}_*(t), \tilde{L}_*(t))}{\partial L} \psi^1(t) - (\mu + \delta) \psi^2(t) \\ &\qquad - \frac{e^{-\rho t}}{F(\tilde{K}_*(t), \tilde{L}_*(t))} \frac{\partial F(\tilde{K}_*(t), \tilde{L}_*(t))}{\partial L}, \end{aligned} \tag{12.5}$$

$$\mathcal{H}(\tilde{K}_*(t), \tilde{L}_*(t), t, u_*(t), \psi(t)) \overset{\text{a.e.}}{=} H(\tilde{K}_*(t), \tilde{L}_*(t), t, \psi(t));$$

2) *for any $t \ge 0$, we have (see (10.3))*

$$\psi^1(t) > 0, \quad \psi^2(t) > 0; \tag{12.6}$$

3) *if there exists a number $\theta > 0$ such that*

$$u_*(t) \overset{\text{a.e.}}{\ge} \theta$$

starting from some moment $\tau \geq 0$, *then the transversality condition*

$$\lim_{t\to\infty}\Big(\psi^1(t)\tilde{K}_*(t)+\psi^2(t)\tilde{L}_*(t)\Big)=0, \tag{12.7}$$

holds.

Here,

$$\mathcal{H}(\tilde{K},\tilde{L},t,u,\psi)=uF(\tilde{K},\tilde{L})\psi^1+(\mu+\delta)\tilde{L}\psi^2+e^{-\rho t}\big[\ln(1-u)+\ln F(\tilde{K},\tilde{L})\big],$$

$$H(\tilde{K},\tilde{L},t,\psi)=\max_{u\in U_\varepsilon}\mathcal{H}(\tilde{K},\tilde{L},t,u,\psi)$$

is the Hamilton–Pontryagin function and the Hamiltonian in the normal form for problem (P_ε).

Note that since both phase variables $\tilde{K}(\cdot)$, $\tilde{L}(\cdot)$ and coordinates $\psi^1(\cdot)$, $\psi^2(\cdot)$ of adjoint variable $\psi(\cdot)$ take positive values on $[0,\infty)$ (see (12.6)) transversality condition (12.7) implies

$$\lim_{t\to\infty}\psi^1(t)\tilde{K}_*(t)=0 \quad \text{and} \quad \lim_{t\to\infty}\psi^2(t)\tilde{L}_*(t)=0. \tag{12.8}$$

Below we consider problem (P_ε) with an arbitrary small parameter $\varepsilon>0$. We will show that for any initial state $(K_0,L_0)\in G$, the parameter ε has no effect on the optimal control process (K_*,L_*,u_*) provided that this parameter is sufficiently small.

Let us introduce the phase variable $x(\cdot)$, $x(t)=\tilde{K}(t)/\tilde{L}(t)$, $t\geq 0$, in problem (P_ε). In terms of the phase variable $x(\cdot)$, the optimal control problem (P_ε) can be rewritten as the following equivalent problem $(\tilde{P}_\varepsilon)$, $0<\varepsilon<1$ (see Section 1 for details):

$$\dot{x}=uf(x)-(\mu+\delta)x, \quad u\in U_\varepsilon=[0,1-\varepsilon], \tag{12.9}$$

$$x(0)=x_0,$$

$$J(x,u)=\int_0^\infty e^{-\rho t}\big[\ln(1-u)+\ln f(x)\big]\,dt\to\max.$$

Here, $x_0\in\tilde{G}=\{x\in R^1: x>0\}$ and $f(x)=F(x,1)$ for any $x\in\tilde{G}$. The other data in problem $(\tilde{P}_\varepsilon)$ are the same as in problem (P_ε).

Note that function $f(\cdot)$ satisfies the neoclassical conditions (1.50), (1.51).

Let $\psi(\cdot)=(\psi^1(\cdot),\psi^2(\cdot))$ be an adjoint variable corresponding to the optimal control process $(\tilde{K}_*(\cdot),\tilde{L}_*(\cdot),u_*(\cdot)))$ in problem (P_ε) by virtue of Theorem 23. Then, $u_*(\cdot)$ is an optimal control in problem $(\tilde{P}_\varepsilon)$ and $x_*(\cdot)$, $x_*(t)=\tilde{K}_*(t)/\tilde{L}_*(t)$, $t\geq 0$, is the corresponding optimal admissible trajectory in this problem.

Set

$$p(t)=e^{\rho t}\psi^1(t)\tilde{L}_*(t), \quad t\geq 0.$$

Then, by conditions (12.2) and (12.5), the adjoint variable $p(\cdot)$ takes positive values and satisfies the following differential equation on the time interval $[0, \infty)$:

$$\dot{p} = \left[\rho + \mu + \delta - u_*(t)\frac{d}{dx}f(x_*(t))\right]p - \frac{1}{f(x_*(t))}\frac{d}{dx}f(x_*(t)). \tag{12.10}$$

Here, we used the homogeneity condition (1.43) and the fact that

$$\frac{d}{dx}f(x) = \frac{\partial}{\partial x}F(x,1) \quad \text{for any} \quad x > 0.$$

Note that (12.10) is exactly the normal form adjoint system for the current value adjoint variable $p(\cdot)$ in problem $(\tilde{P}_\varepsilon)$ as it was introduced in Section 7 (see (7.27)).

Thus Theorem 23 implies the following normal-form version of the Pontryagin maximum principle for problem $(\tilde{P}_\varepsilon)$ in terms of the current value adjoint variable $p(\cdot)$, the current value Hamilton–Pontryagin function $\mathcal{M}(\cdot,\cdot,\cdot)$, and the current value Hamiltonian $M(\cdot,\cdot)$ (see Section 7).

Theorem 24. *Let $u_*(\cdot)$ be an optimal control in problem $(\tilde{P}_\varepsilon)$, $0 < \varepsilon < 1$, and $x_*(\cdot)$ be the corresponding optimal trajectory. Then, there exists a current value adjoint variable $p(\cdot)$ (corresponding to the pair $(x_*(\cdot), u_*(\cdot))$) such that the following conditions hold:*

1) *the optimal pair $(x_*(\cdot), u_*(\cdot))$, together with the current value adjoint variable $p(\cdot)$ satisfies the core relations of the Pontryagin maximum principle in the normal form on the infinite time interval $[0, \infty)$ (see (7.32) and (7.33)):*

$$\dot{p}(t) \overset{\text{a.e.}}{=} \left[\rho + \mu + \delta - u_*(t)\frac{d}{dx}f(x_*(t))\right]p - \frac{1}{f(x_*(t))}\frac{d}{dx}f(x_*(t)),$$

$$\mathcal{M}(x_*(t), u_*(t), p(t)) \overset{\text{a.e.}}{=} M(x_*(t), p(t));$$

2) *for any $t \geq 0$, we have (see (12.6))*

$$p(t) > 0; \tag{12.11}$$

3) *if there exists a number $\theta > 0$ such that*

$$u_*(t) \overset{\text{a.e.}}{\geq} \theta \tag{12.12}$$

starting from some moment $\tau \geq 0$, then the following transversality condition holds (see (12.8)):

$$\lim_{t\to\infty} e^{-\rho t}p(t)x_*(t) = 0. \tag{12.13}$$

Here, the current value Hamilton–Pontryagin function $\mathcal{M}(\cdot,\cdot,\cdot)$ and the current value Hamiltonian $M(\cdot,\cdot)$ in the normal form for problem $(\tilde{P}_\varepsilon)$ are defined *in a standard way (see* (7.34) *and* (7.35)*):*

$$\mathcal{M}(x,u,p) = (uf(x) - (\mu+\delta)x)p + \ln(1-u) + \ln f(x),$$
$$M(x,p) = \max_{u\in U_\varepsilon} \mathcal{M}(x,u,p) \tag{12.14}$$

for any $x > 0$*,* $u \in U_\varepsilon$ *and* $p > 0$*.*

Note that reduced problem $(\tilde{P}_\varepsilon)$ satisfies the conditions of Theorem 21 not for all values of parameters μ, δ and ε, and not for all initial states x_0. Indeed, in view of the neoclassical conditions, inequality $(1-\varepsilon)f(x_0) < (\mu+\delta)x_0$ takes place for any sufficiently large x_0, that contradicts the conditions of Theorem 21. However, the problem (P_ε) (see (12.1)–(12.4)) satisfies all conditions of Theorem 21. Therefore, we first apply Theorem 21 to problem (P_ε) and then, introducing the new phase variable $x(\cdot)$ and the current value adjoint variable $p(\cdot)$, we derive from the result thus obtained (Theorem 23) the required version of the Pontryagin maximum principle (Theorem 24) for problem $(\tilde{P}_\varepsilon)$.

Now let us construct the Hamiltonian system of the Pontryagin maximum principle for problem $(\tilde{P}_\varepsilon)$ (see [**6**] for more details about the methodology of construction of the Hamiltonian system of the Pontryagin maximum principle).

By Theorem 24, we have $p(t) > 0$ for any $t \geq 0$. Since all admissible trajectories $x(\cdot)$ of the control system (12.9) are positive, any pair $(x_*(\cdot), p(\cdot))$, where $x_*(\cdot)$ is an optimal trajectory in problem $(\tilde{P}_\varepsilon)$ and $p(\cdot)$ is a current adjoint variable corresponding to this trajectory, lies in the positive quadrant

$$\Gamma = \big\{(x,p) \in R^2 : \, x > 0, \ p > 0\big\}$$

for all $t \geq 0$.

Thus, we will construct the Hamiltonian system of the maximum principle for problem $(\tilde{P}_\varepsilon)$ in the open set Γ. In what follows, for brevity we will refer to this system as the Hamiltonian system.

Resolving the maximum condition (12.14) for the control u in the set Γ, we find that the maximum in (12.14) is attained at the point

$$u(x,p) = \begin{cases} 0 & \text{if} \quad 0 < p < \frac{1}{f(x)}, \\ 1 - \frac{1}{f(x)p} & \text{if} \quad \frac{1}{f(x)} \leq p \leq \frac{1}{\varepsilon f(x)}, \\ 1-\varepsilon & \text{if} \quad p > \frac{1}{\varepsilon f(x)}; \end{cases} \tag{12.15}$$

moreover,

$$M(x,p) = (u(x,p)f(x) - (\mu+\delta)x)\,p + \ln(1 - u(x,p)) + \ln f(x)$$

everywhere in the set Γ.

According to condition (12.15), we introduce the following sets:

$$\Gamma_1 = \left\{(x,p) \in \Gamma :\, 0 < p < \frac{1}{f(x)}\right\}, \quad \Gamma_2 = \left\{(x,p) \in \Gamma : \frac{1}{f(x)} \leq p \leq \frac{1}{\varepsilon f(x)}\right\},$$

$$\Gamma_3 = \left\{(x,p) \in \Gamma : p > \frac{1}{\varepsilon f(x)}\right\}.$$

Then it is obvious that

$$\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3.$$

In the set Γ_1, the Hamiltonian system has the form (see (12.15))

$$\dot{x} = -(\mu + \delta)x,$$

$$\dot{p} = (\rho + \mu + \delta)p - \frac{1}{f(x)}\frac{d}{dx}f(x).$$

Thus, the Hamiltonian system has no equilibrium states in the set Γ_1. All of its trajectories have exponentially decaying x-coordinate in this set.

Define a function $y_1 : (0, \infty) \mapsto R^1$ as

$$y_1(x) = \frac{1}{(\rho + \mu + \delta) f(x)} \frac{d}{dx} f(x) \quad \text{for any} \quad x > 0.$$

In view of the neoclassical conditions, the function $y_1(\cdot)$ monotonically decreases,

$$y_1(x) \to \infty \quad \text{as} \;\; x \to 0 \quad \text{and} \quad y_1(x) \to 0 \quad \text{as} \;\; x \to \infty,$$

and the equality $y_1(x) = 1/f(x)$ is satisfied at a unique point $\hat{x}$, which is the root of the equation

$$\frac{d}{dx} f(x) = \rho + \mu + \delta. \tag{12.16}$$

Define a curve V_{12}^0 in the set Γ_1 as

$$V_{12}^0 = \big\{(x,p) \in \Gamma_1 :\, x \geq \hat{x},\; p = y_1(x)\big\}.$$

The curve V_{12}^0 is the locus of points in the set Γ_1 at which the p-coordinates of the trajectories of the Hamiltonian system have zero derivatives.

The curve V_{12}^0 divides the set Γ_1 into two subsets

$$V_1^{--} = \big\{(x,p) \in \Gamma_1 :\, x > 0,\; 0 < p < y_1(x)\big\},$$

$$V_1^{-+} = \left\{(x,p) \in \Gamma_1 :\, \hat{x} < x < \infty,\; y_1(x) < p < \frac{1}{f(x)}\right\}.$$

Here, V_1^{--} is the set of points in Γ_1 at which both coordinates $x(\cdot)$ and $p(\cdot)$ of the trajectories of the Hamiltonian system have negative derivatives, and V_1^{-+} is the set of points in Γ_1 at which the x-coordinates of the trajectories of the Hamiltonian system have negative derivatives while the derivative of the p-coordinate is positive.

The disposition of the sets V_1^{--} and V_1^{-+} and the curve V_{12}^0 is shown in Figure 8.

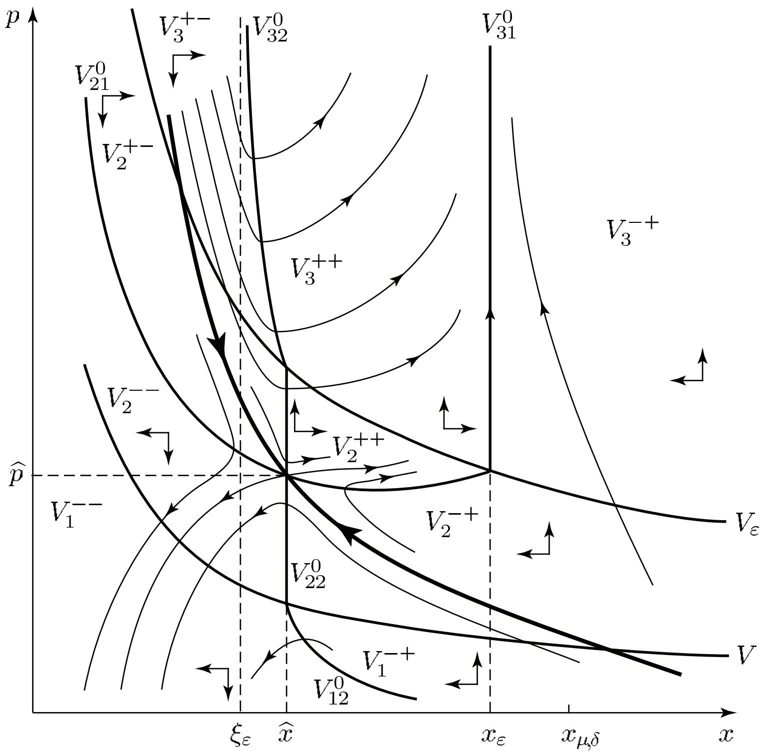

Figure 8: *Trajectories of the Hamiltonian system for problem* $(\tilde{P}_\varepsilon)$.

It is obvious that

$$\Gamma_1 = V_{12}^0 \cup V_1^{--} \cup V_1^{-+}.$$

In the set V_1^{--}, both coordinates $x(\cdot)$ and $p(\cdot)$ of the solutions to the Hamiltonian system decrease. In the set V_1^{-+}, the x-coordinate decreases while the p-coordinate increases. The derivative of the p-coordinate of the solutions to the Hamiltonian system changes its sign on the curve V_{12}^0.

If some trajectory of the Hamiltonian system starts from the set V_1^{--} or has the initial state on the curve V_{12}^0, then this trajectory belongs to the set V_1^{--} at any subsequent moment. It is easily seen that the p-coordinate of this trajectory vanishes in finite time, which contradicts condition (12.11) of the maximum principle (Theorem 24). Thus, the trajectories of the Hamiltonian system that enter the set V_1^{--} (or hit the curve V_{12}^0) at some moment cannot correspond to any optimal pairs $(x_*(\cdot), u_*(\cdot))$ in problem $(\tilde{P}_\varepsilon)$.

If a trajectory of the Hamiltonian system starts from the set V_1^{-+}, then its x-coordinate decays exponentially, and it either hits the curve V_{12}^0 at some moment and then enters the set V_1^{--} (such trajectories cannot correspond to optimal pairs in problem $(\tilde{P}_\varepsilon)$), or intersects the curve

$$V = \left\{ (x,p) \in \Gamma : \, p = \frac{1}{f(x)} \right\}$$

in finite time (to the right of the point $\hat{x}$), leaves the set Γ_1, and enters the set Γ_2.

The behavior of the trajectories of the Hamiltonian system in the set Γ_1 is illustrated in Figure 8. Note that the parameter ε has no effect on the trajectories of the Hamiltonian system in the set Γ_1.

In the set $\text{int}\Gamma_2$, the Hamiltonian system of the maximum principle has the form (see (12.15))

$$\dot{x} = f(x) - (\mu + \delta)x - \frac{1}{p}, \tag{12.17}$$

$$\dot{p} = \Big(\rho + \mu + \delta - \frac{d}{dx} f(x) \Big) p. \tag{12.18}$$

Define a function $y_2(\cdot)$ on the interval $(0, x_{\mu,\delta})$, where $x_{\mu,\delta}$ is a (unique) root of the equation

$$f(x) = (\mu + \delta)x, \tag{12.19}$$

by the equality

$$y_2(x) = \frac{1}{f(x) - (\mu + \delta)x} \quad \text{for any} \quad 0 < x < x_{\mu,\delta}$$

and curves V_{21}^0 and V_{22}^0 in the set Γ_2 as follows:

$$V_{21}^0 = \big\{ (x,p) \in \Gamma_2 : \, 0 < x < x_{\mu,\delta}, \; p = y_2(x) \big\}, \quad V_{22}^0 = \big\{ (x,p) \in \Gamma_2 : \, x = \hat{x} \big\}.$$

Recall that the point $\hat{x}$ is defined as the root of equation (12.16).

The curve V_{21}^0 is the locus of points in Γ_2 at which the x-coordinates of the trajectories of the Hamiltonian system have zero derivatives. The curve (segment) V_{22}^0 has a similar meaning. On this segment, the p-coordinates of the trajectories of the Hamiltonian system have zero derivatives. Since

$$\frac{1}{f(x) - (\mu + \delta)x} > \frac{1}{f(x)}$$

for any $x \in (0, x_{\mu,\delta})$, the curve V_{21}^0 lies above the curve V. By the neoclassical conditions, the curve V_{21}^0 has a unique point of intersection with the curve

$$V_\varepsilon = \left\{ (x,p) \in \Gamma : \, x > 0, \; p = \frac{1}{\varepsilon f(x)} \right\}.$$

This intersection point $(x_\varepsilon, p_\varepsilon)$ (it depends on the parameter ε) is specified by the equalities

$$f(x_\varepsilon) = \frac{(\mu+\delta)x_\varepsilon}{1-\varepsilon}, \quad p_\varepsilon = \frac{1}{\varepsilon f(x_\varepsilon)}.$$

To the left of the point x_ε, the curve V_{21}^0 lies below the curve V_ε. Thus,

$$V_{21}^0 = \big\{(x,p) \in \Gamma : 0 < x \le x_\varepsilon, \ p = y_2(x)\big\}.$$

Note that, in view of neoclassical conditions and definitions of points $\hat{x}$ and $x_{\mu,\delta}$ as roots of equations (12.16) and (12.19) respectively, for any sufficiently small values of ε, we have $\hat{x} < x_\varepsilon < x_{\mu,\delta}$ and $x_\varepsilon \to x_{\mu,\delta}$ as $\varepsilon \to 0$. Henceforth, we consider the situation when the parameter ε is sufficiently small.

The curve V_{21}^0 and the segment V_{22}^0 have a unique intersection point $(\hat{x}, \hat{p}) \in \mathrm{int}\Gamma_2$, where

$$\hat{p} = \frac{1}{f(\hat{x}) - (\mu+\delta)\hat{x}}.$$

This point $(\hat{x}, \hat{p})$ is a unique equilibrium state of the Hamiltonian system in the set Γ_2.

The curve V_{21}^0 and the segment V_{22}^0 divide the set $\mathrm{int}\Gamma_2$ into four subsets:

$$V_2^{--} = \left\{(x,p) \in \Gamma_2 : 0 < x < \hat{x}, \ \frac{1}{f(x)} < p < y_2(x)\right\},$$

$$V_2^{+-} = \left\{(x,p) \in \Gamma_2 : 0 < x < \hat{x}, y_2(x) < p < \frac{1}{\varepsilon f(x)}\right\},$$

$$V_2^{++} = \left\{(x,p) \in \Gamma_2 : \hat{x} < x < x_\varepsilon, y_2(x) < p < \frac{1}{\varepsilon f(x)}\right\},$$

$$V_2^{-+} = \left\{(x,p) \in \Gamma_2 : \hat{x} < x < x_\varepsilon, \ \frac{1}{f(x)} < p < y_2(x)\right\} \cup \left\{(x,p) \in \Gamma_2 : x \ge x_\varepsilon, \frac{1}{f(x)} < p < \frac{1}{\varepsilon f(x)}\right\}.$$

Here, the subscript indicates that the relevant subset belongs to the set Γ_2, and the superscripts indicate the signs of the derivatives of the coordinates $x(\cdot)$ and $p(\cdot)$ for the solutions to the Hamiltonian system in these sets. The disposition of the sets V_2^{--}, V_2^{+-}, V_2^{-+}, and V_2^{++}, the curve V_{21}^0, and the segment V_{22}^0 is shown in Figure 8.

The following equality obviously holds:

$$\Gamma_2 = V_{21}^0 \cup V_{22}^0 \cup V \cup V_\varepsilon \cup V_2^{--} \cup V_2^{+-} \cup V_2^{++} \cup V_2^{-+}.$$

In the set V_2^{--}, both coordinates $x(\cdot)$ and $p(\cdot)$ of the solutions to the Hamiltonian system decrease. Any solution that starts from a point of the set V_2^{--} certainly intersects the curve V in finite time and enters the set V_1^{--}.

Therefore, such solutions cannot correspond to any optimal pairs in problem $(\tilde{P}_\varepsilon)$ (see the analysis performed above when considering the behavior of the trajectories of the Hamiltonian system in the set V_1^{--}).

In the set V_2^{+-}, the x-coordinate of any solution to the Hamiltonian system increases (its derivative changes the sign on the curve V_{21}^0), while the p-coordinate decreases. If the trajectory does not intersect the curve V_{21}^0 on a finite time interval, then it either tends to the equilibrium state $(\hat{x}, \hat{p})$ as $t \to \infty$ or intersects, in finite time, either the segment V_{22}^0 and enters the set V_2^{++} or the curve V_ε on the interval $[\zeta_\varepsilon, \hat{x}]$, $\zeta_\varepsilon < \hat{x}$, where ζ_ε is the root of the equation

$$\frac{d}{dx} f(x) = \frac{\rho + \mu + \delta}{\varepsilon},$$

and enters the set Γ_3.

In the set V_2^{++}, both coordinates $x(\cdot)$ and $p(\cdot)$ of any solution to the Hamiltonian system increase. In this set, any solution to the Hamiltonian system intersects the curve V_ε in finite time (on the interval between the points $\hat{x}$ and x_ε) and enters the set Γ_3.

In the set V_2^{-+}, the x-coordinate of any solution to the Hamiltonian system of the maximum principle decreases (its derivative changes the sign on the curve V_{21}^0), while the p-coordinate increases. Thus, in this domain, an arbitrary trajectory of the Hamiltonian system either tends to the equilibrium state $(\hat{x}, \hat{p})$ as $t \to \infty$ on the infinite time interval or intersects, in finite time, either the curve V_{21}^0 and enters the set V_2^{++}, or the segment V_{22}^0 and enters the set V_2^{--}, or the curve V_ε and enters the set Γ_3. Note that if a trajectory intersects the segment V_{22}^0 and enters the set V_2^{--}, then it cannot correspond to any optimal regime in problem $(\tilde{P}_\varepsilon)$ because in this case it will later enter the set V_1^{--}.

Consider the equilibrium state $(\hat{x}, \hat{p})$ of the Hamiltonian system in the set Γ_2. This equilibrium state corresponds a stationary trajectory

$$x_*(t) \equiv \hat{x}, \quad u_*(t) \stackrel{\text{a.e.}}{=} 1 - \frac{1}{f(\hat{x})\hat{p}}$$

of the control system (12.9) in problem $(\tilde{P}_\varepsilon)$ with the initial condition $x_0 = \hat{x}$.

Linearizing the Hamiltonian system (12.17), (12.18) in the neighborhood of the equilibrium state $(\hat{x}, \hat{p})$, we obtain the following linear system with constant coefficients:

$$\dot{x} = \left[\frac{d}{dx} f(\hat{x}) - (\mu + \delta)\right](x - \hat{x}) + \frac{1}{\hat{p}^2}(p - \hat{p}), \tag{12.20}$$

$$\dot{p} = -\frac{d^2}{dx^2} f(\hat{x})\hat{p}(x - \hat{x}). \tag{12.21}$$

Here we took into account that the point $\hat{x}$ is the root of equation (12.16).

The matrix A of the linear differential system (12.20), (12.21) has the form

$$A = \begin{pmatrix} \frac{d}{dx} f(\hat{x}) - (\mu + \delta) & \frac{1}{\hat{p}^2} \\ -\frac{d^2}{dx^2} f(\hat{x})\hat{p} & 0 \end{pmatrix},$$

and its eigenvalues $\lambda_{1,2}$ are given by

$$\lambda_{1,2} = \frac{1}{2}\left[\frac{d}{dx} f(\hat{x}) - (\mu + \delta) \pm \sqrt{\left(\mu + \delta - \frac{d}{dx} f(\hat{x})\right)^2 - \frac{4}{\hat{p}}\frac{d^2}{dx^2} f(\hat{x})}\,\right].$$

By virtue of the neoclassical conditions, these eigenvalues are real numbers of different signs. Hence, in a neighborhood Ω of the equilibrium state $(\hat{x}, \hat{p})$, the Hamiltonian system (12.17), (12.18) is of saddle type (see [**24**]). In this case, there are only two phase trajectories of the Hamiltonian system that approach the equilibrium state $(\hat{x}, \hat{p})$ asymptotically as $t \to \infty$. One such trajectory $(x_1(\cdot), p_1(\cdot))$ approaches the point $(\hat{x}, \hat{p})$ from the left from the set V_2^{+-} (we call this trajectory $(x_1(\cdot), p_1(\cdot))$ the left equilibrium trajectory), while the second trajectory $(x_2(\cdot), p_2(\cdot))$ approaches the point $(\hat{x}, \hat{p})$ from the right from the set V_2^{-+} (we call this trajectory $(x_2(\cdot), p_2(\cdot))$ the right equilibrium trajectory).

Indeed, the set V_2^{--} lies to the left of the point $(\hat{x}, \hat{p})$ (of the segment V_{22}^0), and the x-coordinate decreases in this set (see Figure 8). Therefore, it is impossible to reach the point $(\hat{x}, \hat{p})$ from the set V_2^{--}. Similarly, the set V_2^{++} lies to the right of the point $(\hat{x}, \hat{p})$, and the x-coordinate increases in this set (see Figure 8). Therefore, it is impossible to reach the point $(\hat{x}, \hat{p})$ from the set V_2^{++}. In the neighborhood Ω, the equilibrium state $(\hat{x}, \hat{p})$ is of saddle type; therefore, only the two trajectories $(x_1(\cdot), p_1(\cdot))$ and $(x_2(\cdot), p_2(\cdot))$ tend to this point asymptotically as $t \to \infty$. Since the equilibrium state $(\hat{x}, \hat{p})$ is of saddle type, these trajectories cannot approach the point $(\hat{x}, \hat{p})$ from the same set (either from V_2^{+-} or from V_2^{-+}), so one of these trajectories approaches $(\hat{x}, \hat{p})$ from the left from V_2^{+-} and the other from the right from V_2^{-+} (see Figure 8).

In the set Γ_3, the Hamiltonian system of the maximum principle has the form (see (12.15))

$$\dot{x} = (1 - \varepsilon) f(x) - (\mu + \delta)x,$$

$$\dot{p} = \left(\rho + \mu + \delta - (1 - \varepsilon)\frac{d}{dx} f(x)\right) p - \frac{1}{f(x)}\frac{d}{dx} f(x).$$

Denote by $\xi_\varepsilon > 0$ the (unique) root of the equation

$$\frac{d}{dx} f(x) = \frac{\rho + \mu + \delta}{1 - \varepsilon}.$$

By the neoclassical conditions, we have (see (12.16))

$$\xi_\varepsilon < \hat{x} \quad \text{and} \quad \xi_\varepsilon \to \hat{x} - 0 \quad \text{as} \;\; \varepsilon \to 0.$$

Therefore,

$$\rho + \mu + \delta < \frac{d}{dx} f(x) < \frac{\rho + \mu + \delta}{1 - \varepsilon} \quad \text{for} \quad x \in (\xi_\varepsilon, \hat{x}). \tag{12.22}$$

Moreover,

$$\frac{d}{dx} f(x) < \rho + \mu + \delta \quad \text{for} \quad x \in (\hat{x}, \infty). \tag{12.23}$$

Define a function $y_3 : (\xi_\varepsilon, \infty) \mapsto R^1$, a line V_{31}^0, and a curve V_{32}^0 as follows:

$$y_3(x) = \frac{\frac{d}{dx} f(x)}{f(x)\big[\rho + \mu + \delta - (1 - \varepsilon)\frac{d}{dx} f(x)\big]}, \qquad x \in (\xi_\varepsilon, \infty),$$

$$V_{31}^0 = \big\{(x, p) \in \Gamma_3 : \, x = x_\varepsilon\big\}, \quad V_{32}^0 = \big\{(x, p) \in \Gamma_3 : \, x \in (\xi_\varepsilon, \infty), \; p = y_3(x)\big\}.$$

The vertical line V_{31}^0 is the locus of points in Γ_3 at which the x-coordinates of the trajectories of the Hamiltonian system have zero derivatives, while the curve V_{32}^0 is the locus of points in the set Γ_3 at which the p-coordinates of the trajectories of the Hamiltonian system have zero derivatives.

The following equality holds at the point $\hat{x}$:

$$y_3(\hat{x}) = \frac{1}{\varepsilon f(\hat{x})}.$$

Next, by condition (12.22), for any $x \in (\xi_\varepsilon, \hat{x})$ we have

$$y_3(x) > \frac{\rho + \mu + \delta}{f(x)[\rho + \mu + \delta - (1 - \varepsilon)(\rho + \mu + \delta)]} = \frac{1}{\varepsilon f(x)}.$$

By virtue of inequality (12.23), for any $x \in (\hat{x}, \infty)$ we obtain

$$y_3(x) < \frac{\rho + \mu + \delta}{f(x)[\rho + \mu + \delta - (1 - \varepsilon)(\rho + \mu + \delta)]} = \frac{1}{\varepsilon f(x)}.$$

Thus, the line V_{31}^0 and the curve V_{32}^0 do not intersect. Hence, the Hamiltonian system has no equilibrium states in the set Γ_3. The disposition of the line V_{31}^0 and the curve V_{32}^0 in the set Γ_3 is shown in Figure 8.

The line V_{31}^0 and the curve V_{32}^0 divide the set Γ_3 into three subsets (see Figure 8):

$$V_3^{+-} = \{(x,p) \in \Gamma_3 : x \le \xi_\varepsilon\} \cup \left\{(x,p) \in \Gamma_3 : \xi_\varepsilon < x < \hat{x},\ \frac{1}{\varepsilon f(x)} < p < y_3(x)\right\},$$

$$V_3^{++} = \{(x,p) \in \Gamma_3 : \xi_\varepsilon < x < \hat{x},\ p > y_3(x)\} \cup \left\{(x,p) \in \Gamma_3 : \hat{x} \le x < x_\varepsilon,\ p > \frac{1}{\varepsilon f(x)}\right\},$$

$$V_3^{-+} = \left\{(x,p) \in \Gamma_3 : x > x_\varepsilon,\ p > \frac{1}{\varepsilon f(x)}\right\}.$$

In the subset V_3^{+-}, the x-coordinate increases while the p-coordinate decreases. In the subset V_3^{++}, both the x-coordinate and the p-coordinate increase. In the subset V_3^{-+}, the x-coordinate decreases while the p-coordinate increases.

Any trajectory of the Hamiltonian system starts from the set V_3^{+-} either enters the set V_2^{+-} in finite time, or enters the set V_3^{++} also in finite time.

If a trajectory of the Hamiltonian system starts from the set V_3^{-+} or from the set V_3^{++}, then it remains in the set V_3^{-+} or, respectively, in the set V_3^{++} for all $t \ge 0$ and asymptotically approaches the vertical line V_{31}^0 as $t \to \infty$ (see Figure 8).

The line V_{31}^0 itself is a phase trajectory of the Hamiltonian system, with the upward vertical motion along this trajectory. Such trajectories cannot correspond to any optimal pair in problem $(\tilde{P}_\varepsilon)$ for any initial condition $x_0 > 0$. Indeed, let us show that when a trajectory $x(\cdot)$ of the Hamiltonian system asymptotically approaches the line V_{31}^0 as $t \to \infty$ (in the set V_3^{-+} or V_3^{++}) or coincides with this line, the transversality condition at infinity (12.13) is violated. Since $\lim_{t\to\infty} x(t) = x_\varepsilon$ in this case, condition (12.13) is equivalent to the condition

$$\lim_{t\to\infty} e^{-\rho t} p(t) = 0. \qquad (12.24)$$

Since $f(x_{\mu,\delta}) = (\mu + \delta)x$ and $x_\varepsilon \to x_\mu$ as $\varepsilon \to 0$, we have

$$\frac{d}{dx} f(x) < \mu + \delta - \varepsilon$$

everywhere in some neighborhood of the point x_ε for sufficiently small $\varepsilon > 0$. Moreover, the following inequality holds everywhere in the set Γ_3:

$$\frac{1}{f(x)} < \varepsilon p.$$

Taking into account the above two inequalities, which hold when a trajectory $(x(\cdot), p(\cdot))$ approaches the line V_{31}^0 or coincides with this line, we obtain

$$\frac{d}{dt}\big(e^{-\rho t}p(t)\big) = e^{-\rho t}\left(\mu + \delta - (1-\varepsilon)\frac{d}{dx}f(x(t))\right)p(t) - e^{-\rho t}\frac{1}{f(x(t))}\frac{d}{dx}f(x(t))$$
$$> e^{-\rho t}\left(\mu + \delta - \frac{d}{dx}f(x(t))\right)p(t) > \varepsilon e^{-\rho t}p(t)$$

for sufficiently large $t > 0$. Hence,

$$\lim_{t\to\infty} e^{-\rho t}p(t) = \infty,$$

which contradicts condition (12.24). Thus, the trajectories of the Hamiltonian system that enter the sets V_3^{-+} and V_3^{++}, as well as the trajectory going along the line V_{31}^0, cannot correspond to any optimal pair in problem $(\tilde{P}_\varepsilon)$ for any initial condition $x_0 > 0$.

Let us summarize the analysis of the Hamiltonian system of the maximum principle for problem $(\tilde{P}_\varepsilon)$ for sufficiently small values of $\varepsilon > 0$.

According to this analysis, the Hamiltonian system of the maximum principle has a unique equilibrium state $(\hat{x}, \hat{p})$ that does not depend on the parameter ε. The Hamiltonian system of the maximum principle has only two trajectories that asymptotically tend to the equilibrium state $(\hat{x}, \hat{p})$ as $t \to \infty$: these are the left equilibrium trajectory $(x_1(\cdot), p_1(\cdot))$ and the right equilibrium trajectory $(x_2(\cdot), p_2(\cdot))$.

All trajectories $(x(\cdot), p(\cdot))$ of the Hamiltonian system of the maximum principle can be classified in tree classes according to their asymptotic behaviors at infinity:

1) trajectory $(x(\cdot), p(\cdot))$ tends to the equilibrium state $(\hat{x}, \hat{p})$ as t goes to infinity (see Figure 8). In this case $(x(\cdot), p(\cdot))$ is either the left equilibrium trajectory $(x_1(\cdot), p_1(\cdot))$, or the right equilibrium trajectory $(x_2(\cdot), p_2(\cdot))$, or the stationary equilibrium trajectory $(x(t), p(t)) \equiv (\hat{x}, \hat{p})$ for all $t \geq 0$. It is easy to see that the transversality condition (12.13) is satisfied in this case. Hence any trajectory $(x(\cdot), p(\cdot))$ of this type satisfies conditions of Theorem 24;

2) starting from some instant $\tau > 0$ and for all $t \geq \tau$ trajectory $(x(\cdot), p(\cdot))$ lies in the set V_1^{--}(see Figure 8). It is easily seen that the p-coordinate of this trajectory vanishes in finite time, which contradicts condition (12.11) of the maximum principle (Theorem 24). So, such trajectories cannot correspond to any optimal pair in problem $(\tilde{P}_\varepsilon)$ for any initial condition $x_0 > 0$.

3) trajectory $(x(\cdot), p(\cdot))$ asymptotically approaches the vertical line V_{31}^0 as $t \to \infty$ (see Figure 8). Such trajectories cannot correspond to any optimal pair in problem $(\tilde{P}_\varepsilon)$ for any initial condition $x_0 > 0$ because the transversality condition at infinity (12.13) is violated in this case, while $u(t) = 1 - \varepsilon >$

0 starting from some instant $\tau > 0$ (see (12.12)). So, they do not satisfy Theorem 24.

Thus, only trajectories $(x(\cdot), p(\cdot))$ of type 1) which tend to the equilibrium state $(\hat{x}, \hat{p})$ as t goes to infinity (or the stationary equilibrium trajectory $(x(t), p(t)) \equiv (\hat{x}, \hat{p})$ for all $t \geq 0$) can correspond to optimal processes in problem $(\tilde{P}_\varepsilon)$.

According to the analysis performed, for any initial state x_0 the Hamiltonian system of the maximum principle has a trajectory that satisfies the necessary optimality conditions (Theorem 24). Depending on the initial condition x_0, this is either the right equilibrium trajectory, the left equilibrium trajectory, or the stationary trajectory $(\hat{x}, \hat{p})$. All other trajectories of the Hamiltonian system do not satisfy the conditions of Theorem 24, because either the p-coordinate of these trajectories vanishes in finite time and they leave the set Γ, or these trajectories asymptotically approach the line V_{31}^0 as $t \to \infty$ (or go along this line), thereby violating the transversality condition (12.13). Thus for any $x < \hat{x}$ there is a unique value $p_*(x)$ such that point $(x, p_*(x))$ lies on the graph of the left equilibrium trajectory $(x_1(\cdot), p_1(\cdot))$. Analogously, for any $x > \hat{x}$ there is a unique value $p_*(x)$ such that point $(x, p_*(x))$ lies on the graph of the right equilibrium trajectory $(x_2(\cdot), p_2(\cdot))$. We put $p_*(\hat{x}) = \hat{p}$.

Thus, the necessary optimality conditions (Theorem 24) describe unambiguously the optimal control in problem $(\tilde{P}_\varepsilon)$ (in the form of an optimal synthesis $u_*(x)$ (see [**32**])) in terms of function $p_* : (0, \infty) \mapsto (0, \infty)$ such defined. It follows from our analysis that

$$\operatorname{graph} p_*(\cdot) \subset V_3^{+-} \cup V_2^{+-} \cup V_2^{-+} \cup V_1^{-+}.$$

Finally we have

$$u_*(x) = \begin{cases} 1 - \varepsilon & \text{if} \quad (x, p_*(x)) \in V_3^{+-}, \\ 1 - \frac{1}{p_*(x) f(x)} & \text{if} \quad (x, p_*(x)) \in V_2^{+-} \cup V_2^{-+}, \\ 0 & \text{if} \quad (x, p_*(x)) \in V_1^{-+}. \end{cases}$$

For an arbitrary initial state $x_0 > 0$, the optimal synthesis $u_*(x)$, $x \in (0, \infty)$, uniquely defines the optimal trajectory $x_*(\cdot)$ in problem $(\tilde{P}_\varepsilon)$ as the solution to the Cauchy problem

$$\dot{x} = u_*(x) f(x) - (\mu + \delta) x,$$
$$x(0) = x_0$$

and the corresponding optimal control $u_*(\cdot)$ as a function

$$u_*(t) = u_*(x_*(t)), \quad t \in [0, \infty).$$

For any given initial state x_0, the optimal pair $(x_*(\cdot), u_*(\cdot))$ generated by this optimal synthesis in problem $(\tilde{P}_\varepsilon)$ does not depend on the parameter ε provided that ε is sufficiently small.

Bibliography

[1] P. Aghion and P. Howitt, Endogenous Growth Theory, Cambridge, MA, MIT Press, 1998.

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

[3] V.M. Alekseev, V.M. Tikhomirov and S.V. Fomin, Optimal control, New York, NY, Plenum Press, 1987.

[4] 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), 89–91.

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

[6] S.M. Aseev, A.V. Kryazhimskii, The Pontryagin maximum principle and optimal economic growth problems, Proceedings of the Steklov Institute of Mathematics, 257 (2007), 1–255.

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

[8] R.J. Barro, X. Sala-i-Martin, Economic Growth, New York, McGraw Hill, 1995.

[9] R. Bellman, Dynamic Programming, Princeton, NJ, Princeton Univ. Press, 1957.

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

[11] D. W. Bushaw, Differential Equations with a Discontinuous Forcing Term, PhD thesis, Princeton University, 1952.

[12] D. Cass, Optimum growth in an aggregative model of capital accumulation, Rev. Econ. Stud., 32 (1965), 233–240.

[13] L. Cesari, Optimization – Theory and Applications. Problems with Ordinary Differential Equations, New York, Springer, 1983.

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

[15] F.H. Clarke, Optimization and Nonsmooth Analysis, New York, J. Wiley, 1983.

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

[17] B.P. Demidovich, Lectures on the Mathematical Theory of Stability, Moscow, Nauka, 1967.

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

[19] A. A. Fel'dbaum, On the synthesis of optimal systems of automatic control, Transactions of the Second National Conference on the Theory of Automatic Control 2. Izdatel'stvo Akademii Nauk SSSR, Moscow, (1955), 325–360 (In Russian).

[20] A.F. Filippov, On some problems in optimal control theory, Vestn. Mosk. Univ., Ser. 1: Mat., Mekh., Astron., Fiz., Khim., 2 (1959), 25–32.

[21] A.F. Filippov, Differential equations with discontinuous right-hand sides. Dordrecht, Kluwer, 1988.
[22] B. V. Gnedenko, *Theory of Probability*, Newark, NJ, Gordon and Breach, 1997.
[23] H. Halkin, Necessary conditions for optimal control problems with infinite horizons, Econometrica, 42 (1974), 267–272.
[24] P. Hartman, Ordinary Differential Equations, New York, J. Wiley & Sons, 1964.
[25] H. Hermes, J.P. La Salle, Functional analysis and time optimal control. New York, Acad. Press, 1969.
[26] K. Inada, On a Two-Sector Model of Economic Growth: Comments and a Generalization, Rev. Econ. Stud., 30 [2] (1963), 119–127.
[27] M.D. Intriligator, Mathematical Optimization and Economic Theory, Englewood Cliffs, NJ, Prentice-Hall, 1971.
[28] A.N. Kolmogorov, S.V. Fomin, Elements of the Theory of Functions and Functional Analysis, Mineola, NY, Dover Publications, 1999.
[29] A. Liapounoff, Problème général de la stabilité du mouvement, Princeton, NJ, Princeton Univ. Press, 1947.
[30] P. Michel, On the transversality condition in infinite horizon optimal problems, Econometrica, 50 (1982), 975–985.
[31] I.P. Natanson, Theory of Functions of a Real Variable, New York, Frederick Ungar Publishing Co., 1955.
[32] L.S. Pontryagin, V.G. Boltyanskij, R.V. Gamkrelidze, E.F. Mishchenko, The Mathematical Theory of Optimal Processes, Oxford, Pergamon, 1964.
[33] F.P. Ramsey, A Mathematical Theory of Saving, Econ. J. 38 (1928), 543–559.
[34] F. Riesz and B. Sz.-Nagy, Functional Analysis, Dover, 1990
[35] R.T. Rockafellar, Convex Analysis, Princeton, NJ, Princeton University Press, 1970.
[36] A. Seierstad, Necessary conditions for nonsmooth, infinite-horizon optimal control problems, J. Optim. Theory Appl., 103 (1) (1999), 201–229.
[37] A. Seierstad, K. Sydsæter, Optimal Control Theory with Economic Applications, Amsterdam, North-Holland, 1987.
[38] K. Shell, Applications of Pontryagin's maximum principle to economics, Mathematical Systems Theory and Economics 1, Springer, Berlin (Lect. Notes Oper. Res. Math. Econ., 11) (1969), 241–292.
[39] G.V. Smirnov, Transversality condition for infinite-horizon problems, J. Optim. Theory Appl., 88 (3), (1996), 671–688.
[40] R.M. Solow, Growth Theory: An Exposition, New York, Oxford Univ. Press, 1970.
[41] M.L. Weitzman, Income, Wealth, and the Maximum Principle, Cambridge, MA, Harvard Univ. Press, 2003.