Optimization

100 Examples

Simon Serovajsky

A Chapman & Hall Book

CRC Press
Taylor & Francis Group

Optimization

Optimization: 100 Examples is a book devoted to the analysis of scenarios for which the use of well-known optimization methods encounter certain difficulties. Analyzing such examples allows a deeper understanding of the features of these optimization methods, including the limits of their applicability. In this way, the book seeks to stimulate further development and understanding of the theory of optimal control. The study of the presented examples makes it possible to more effectively diagnose problems that arise in the practical solution of optimal control problems, and to find ways to overcome the difficulties that have arisen.

Features

- Vast collection of examples
- Simple accessible presentation
- Suitable as a research reference for anyone with an interest in optimization and optimal control theory, including mathematicians and engineers
- Examples differ in properties, i.e., each effect for each class of problems is illustrated by a unique example.

Simon Serovajsky is a professor of mathematics at Al-Farabi Kazakh National University in Kazakhstan. He is the author of many books published in the area of optimization and optimal control theory, mathematical physics, mathematical modelling, philosophy and history of mathematics as well as a long list of high-quality publications in learned journals.

Optimization
100 Examples

Simon Serovajsky

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

First edition published 2025
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

Library of Congress Cataloging-in-Publication Data

Names: Serovajsky, Simon, author.
Title: Optimization : 100 examples / Simon Serovajsky.
Description: First edition. | Boca Raton : C&H/CRC Press, 2024. | Includes bibliographical references and index. | Summary: "This book is devoted to the analysis of scenarios for which the use of well-known optimization methods encounter certain difficulties. Analysing such examples allows a deeper understanding of the features of these optimization methods, including the limits of their applicability. In this way, the book seeks to stimulate further development and understanding of the theory of optimal control. The study of the presented examples makes it possible to more effectively diagnose problems that arise in the practical solution of optimal control problems, and to find ways to overcome the difficulties that have arisen"-- Provided by publisher.
Identifiers: LCCN 2023059001 | ISBN 9781032500072 (hardback) | ISBN 9781032504568 (paperback) | ISBN 9781003398585 (ebook)
Subjects: LCSH: Mathematical optimization. | Control theory.
Classification: LCC QA402.5 .S4334 2024 | DDC 515/.642--dc23/eng/20240131
LC record available at https://lccn.loc.gov/2023059001

ISBN: 978-1-032-50007-2 (hbk)
ISBN: 978-1-032-50456-8 (pbk)
ISBN: 978-1-003-39858-5 (ebk)

DOI: 10.1201/9781003398585

Typeset in Latin Modern font
by KnowledgeWorks Global Ltd.

Publisher's note: This book has been prepared from camera-ready copy provided by the authors.

To school No. 25 in Almaty, which determined my future destiny, to its teachers who gave me deep knowledge, as well as to my friends classmates who were next to me all these years, no matter how many thousands of kilometers separated us.

Contents

Preface

I started solving optimal control problems during my student years. These were applied problems arising in the process of carrying out some scientific and technical projects. Gradually, my interest moved to the area of mathematical theory of optimal control. It soon became clear that I lacked knowledge in the field of functional analysis to justify optimality conditions, study the solvability of optimization problems, and prove the convergence of numerical optimization methods. I took up the study of literature. My attention was drawn to the book by B. Gelbaum and L. Olmsted ***Counterexamples in analysis*** [77].

Yes, of course, there are a significant number of excellent textbooks on functional analysis. They present basic concepts of analysis in an accessible form and provide complete proofs of the most important theorems. All this is illustrated with numerous examples showing what these concepts boil down to and how these theorems work in specific situations. However, the book by Gelbaum and Olmsted is based on completely different principles. It lacks general theoretical results with standard definitions and proofs of theorems, and the examples given are non-standard. It is not for nothing that the term *counterexamples* is used for them. They are unusual, not obvious, and seem to be on the edge, behind which the Unknown is hidden. At the same time, they do not require cumbersome and boring transformations that can "scare off" the reader, for whom mathematical analysis is not a goal, but a means for solving some other problems.

Over time, I began to teach myself. Among the many courses that I had the opportunity to teach was the optimal control theory. Over time, I had a desire to write my own version of counterexamples in relation to this area. The result was the book ***Counterexamples in the Optimal Control Theory***, published initially in Russian in 2001, and subsequently in English by Brill Academic Press in 2004 and republished in 2011 by De Gruyter [170]. It dealt with eight fairly simple optimal control problems with unusual properties.

As time went. My collection of custom examples has steadily expanded. In 2012, a new book by Gelbaum and Olmsted, ***Theorems and Counterexamples in Mathematics*** was published [78]. When the number of "bad" examples reached well over fifty, I decided that I, perhaps, should also write a new book, significantly different from the previous one in form, content, and volume. This is how the idea of ***Optimization*. 100 *examples*** came about. The missing examples had to be invented during the writing of the book. However, I got a little carried away, so in reality, there were even a few more examples. I somehow did not want to change the title, which was already officially approved at that time, especially since there was a precedent. As one knows, the Hundred Years' War lasted slightly longer than the specified number of years. In reality, in my book, there are approximately as many examples of how many years this war lasted.

The main goal of this book is to analyze examples related to the optimal control theory, for which something does not happen quite as we would like. Moreover, all examples must differ from each other either in the class of problems being solved or in the effects that arise in the process of their analysis. In addition, the examples should be extremely simple, so that, firstly, the study is not accompanied by cumbersome transformations that obscure the essence of the considered properties, and, secondly, so that the illusion does not arise that non-standard effects are characteristic only for exotic problems. In order to make it clear exactly what properties we would like to observe, we had to give the corresponding statements (we remember that Gelbaum and Olmsted also included theorems in the second book), and illustrate each of them with a "good" example. However, when presenting them, we did not strive for maximum rigor and completeness, since these issues are well covered in numerous literature. It was only important for us to draw the reader's attention to the logic of reasoning, the contribution that one or another condition of the theorem makes to the final result. Only then, when analyzing "bad" examples, it is established how this or that undesirable property is an indispensable consequence of the violation of one of the conditions of the corresponding theorem.

Naturally, this book should not be used for the first acquaintance with the optimal control theory. For these purposes, there is already a sufficient amount of relevant literature. Likewise, Gelbaum and Olmsted's book is not suitable as a textbook on analysis. It is intended for those who are already familiar with this area and would like to better understand the essence of the matter. In addition, analysis of unusual examples stimulates the desire to push the boundaries of applicability of known mathematical methods, and therefore contributes to the development of the theory itself.

In our case, there is one more important circumstance. In practice, we are often forced to solve an optimization problem without strictly justifying the optimization methods used. Under these conditions, one must allow for the possibility of the manifestation of those very undesirable effects that are described in this book. The very fact of the appearance of various difficulties in extremely simple problems indicates the widespread prevalence of the described phenomena and a fairly high probability of these effects appearing in the practical solution of applied optimization problems. Analysis of the presented examples increases the chances of correctly diagnosing the results obtained during the solving process and choosing the appropriate actions in case of any troubles.

The book consists of five parts. The first part is devoted to the problems of minimizing functions of one variable and is auxiliary. Its inclusion in the book is explained solely by the fact that almost all the features of optimal control problems described below can be found already for problems of minimizing functions. In subsequent parts, problems of optimal control of systems with a free and fixed final state are considered, in the presence of an isoperimetric condition and in the absence of initial conditions.

Each part, consisting of several chapters, begins with a brief description of known optimization methods in relation to a given class of problems, features of their practical application, as well as an example illustrating the effectiveness of the described method. Then a series of examples of similar problems is given in which the use of these methods leads to certain difficulties.

Each chapter includes four sections: *Lecture*, *Results*, *Appendix*, and *Notes*. *Lecture* includes basic theoretical material, as well as the most important examples. The reader who wants to get only a general idea of the plot being described can limit himself to reading only this paragraph.

Results invariably consist of three subsections. By turning to the *Questions* subsection, the reader can check the degree of his understanding of the lecture materials. The subsection *Conclusions* summarizes the main results of the *Lecture*. The *Problems* subsection lists directions for further development of the lecture material with links to *Appendix* or *Notes*.

If the reader wants to get a more complete and in-depth understanding of the issues discussed, he can refer to *Appendix*. Here are some additional theoretical results, as well as new examples that complement those presented in the *Lecture*.

Notes are references throughout the text of the *Lecture* and *Appendix*. It provides some explanations that may be useful to the more interested reader, as well as definitions of some concepts used in the text that goes beyond the scope of the book. Some technical calculations are also included here, which allows you to unload the main text of the chapter. In addition, there are links to literature on specific issues. Finally, it provides cross-references to material from other chapters of the book that are in some way related to the specific issues discussed in this chapter.

The preparation of the book was greatly facilitated by a special course lectured by the author for many years at the Faculty of Mechanics and Mathematics of the al-Farabi Kazakh National University.

I want to express my deep gratitude, first of all, to S. Aisagaliev, whose lectures began my acquaintance with the optimal control theory, J.L. Lions, from whose excellent books I studied and who assessed my own research in this direction, as well as B. Gelbaum and L. Olmstead, whose book inspired me to write this book. I am very grateful to S. Kabanikhin for the opportunity to publish the previous book in English and M. Ruzhansky, who connected me with Taylor and Francis Group. I also express my gratitude to A. Antipin, F. Aliev, V. Amerbaev, A. Ashimov, V. Boltyansky, A. Butkovsky, M. Dzhenaliev, A. Egorov, R. Fedorenko, A. Fursikov, Yu. Gasimov, M. Gebel, A. Iskenderov, O. Ladyzhenskaya, V. Litvinov, A. Lukyanov, K. Lurie, V. Neronov, D. Nurseitov, V. Osmolovsky, U. Raitums, T. Sirazetdinov, Sh. Smagulov, U. Sultangazin, V. Tikhomirov, N. Uraltseva, F. Vasiliev, V. Yakubovich, O. Zhautykov, with whom I had the opportunity to discuss various problems related to optimization methods over the years. I express my deep gratitude to the staff and students of the al-Farabi Kazakh National University, who to one degree or another helped me in my work. I am extremely grateful to the staff of CRC Press/Taylor and Francis Group C. Frazer, M. Kabra, and S. Kumar, who actively supported me throughout the work on the book. I am also grateful to the artists I. Saitov and B. Tasov for preparing the cover of the book. I express special gratitude to A. Teplov, without whom this book would hardly have been written. Finally, I am deeply grateful to my wife Larissa Ananyeva for her understanding and support.

I will be very grateful to readers for comments and suggestions, for any sharp criticism on individual issues and on the book as a whole. I would be especially glad to have an opportunity to add to the collection of unusual examples of optimal control problems. For those wishing to respond, I provide my address: serovajskys@mail.ru.

I dedicate this book to school No. 25 in Almaty, which determined my future destiny, to its teachers who gave me deep knowledge, as well as to my friends classmates who were next to me all these years, no matter how many thousands of kilometers separated us.

I

MINIMIZATION OF FUNCTIONS OF ONE VARIABLE

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

The main purpose of this book is to analyze unusual examples of optimal control problems, as well as to discuss the difficulties that arise in their practical study. However, similar problems appear already for the simplest problem of finding an extremum, which consists of minimizing a function of one variable in the absence of any restrictions. The first part, which is introductory, is devoted to the analysis of these problems. It consists of two chapters, the first of which contains the classical Fermat theorem characterizing the necessary condition for a local extremum of a function. The second chapter discusses some additional results related to minimization of functions, in particular, the problem of the conditional minimum of a function, minimization of a function depending on a parameter, and approximate methods for minimizing functions.

CHAPTER 1

Fermat theorem

The simplest problem in the theory of extremum is the problem of finding the unconditional minimum of a function of one variable. We confine ourselves to considering the simplest and at the same time fundamental result of the theory of extremum, according to which the derivative of a smooth function vanishes at its minimum point. In Section 1.1, we give this statement, called Fermat theorem, and consider different examples related to its practical application[1]. Additional problems arising in this case (existence and uniqueness of minimum points, Tikhonov well-posedness of the minimization problem, sufficiency of the extremum condition, minimization of non-smooth functions) are analyzed in Appendix.

1.1 LECTURE

This section deals with the problem of minimizing a differentiable function of one variable[2]. The main result here is Fermat theorem, which determines a necessary condition for a local extremum of a function. We give a proof of this result (Section 1.1.1) and describe its application to minimizing some fairly simple functions. In the process of analyzing these examples, the problems of the existence of a function minimum (Section 1.1.3), its uniqueness, the Tikhonov well-posedness of the minimization problem, and the sufficiency of the extremum condition (Section 1.1.2), as well as the applicability of the described method (Section 1.1.4) arise.

1.1.1 Fermat theorem

The simplest problem of the extremum theory is to minimize the function of one variable on the entire real axis, i.e., the problem of unconditional minimization of a function.

Problem 1.1 *Find a point of minimum for a function $f = f(x)$.*

The most natural way to analyze the behavior of a function is to study the properties of its derivative[3].

DOI: 10.1201/9781003398585-1

Theorem 1.1 (***Fermat theorem***). *In order for the differentiable function $f = f(x)$ to reach its minimum at the point x, it is necessary that it satisfies the equality*

$$f'(x) = 0. \tag{1.1}$$

Proof. Let x be a point of minimum for the function f. Therefore, the following inequality holds

$$f(y) \geq f(x) \ \ \forall y,$$

whence follows the relation[4]

$$f(x+h) \geq f(x) \ \ \forall h.$$

Using the Taylor formula, taking into account the differentiability of the considered function, we obtain the equality

$$f'(x+h) = f'(x) + f'(x)h + o(h),$$

where $o(h)/h \to 0$ as $h \to 0$. As a result, the last inequality takes the form

$$f'(x)h + o(h) \geq 0 \ \ \forall h. \tag{1.2}$$

Hence, $h > 0$ follows the relation

$$f'(x) + o(h)/h \geq 0.$$

After passage to the limit as $h \to 0$, we get

$$f'(x) \geq 0. \tag{1.3}$$

Analogically, from the formula (1.2) for $h < 0$, it follows the inequality

$$f'(x) + o(h)/h \leq 0,$$

so, after the passage to the limit as $h \to 0$, we obtain

$$f'(x) \leq 0. \tag{1.4}$$

From the inequalities (1.3) and (1.4) it follows the equality (1.1). □

Thus, the search for the minimum of a smooth function on the set of real numbers is reduced to the analysis of formula (1.1), which is an algebraic equation (as a rule, non-linear) with respect to the desired value x.

Definition 1.1 *The equality* (1.1) *is called the* ***stationary condition*** *or* ***Fermat condition***, *and its solution is called the* ***stationary point*** *or the* ***critical point***[5] *of the function f.*

By Theorem 1.1, for solving of function minimization problem, it is necessary to find the stationary points and analyze its properties[6]. Consider easy examples of using the Fermat theorem for concrete functions.

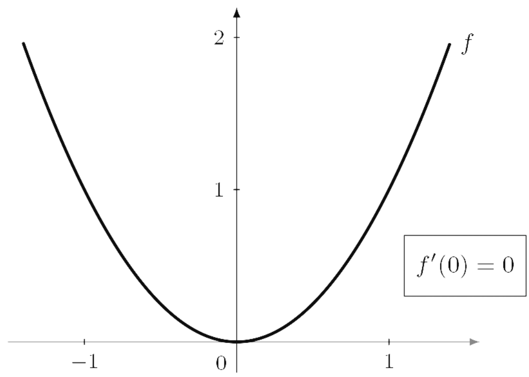

Figure 1.1 The unique stationary point is the global minimum of the function.

Example 1.1 *Find a minimum of the function* $f(x) = x^2$.

Equality (1.1) for this function is $2x = 0$. The unique solution $x = 0$ of this equation is the point of minimum for the function f; see Figure 1.1.

This example illustrates the method of solving the problem of minimizing a function using Fermat theorem. To do this, the extremal problem is reduced to the corresponding algebraic equation, which is solved directly. However, some complications are possible.

1.1.2 Non-uniqueness of the solution of the stationary condition

Consider another example.

Example 1.2 *Find a minimum of the function* $f(x) = 3x^4 – 4x^3 – 12x^2$.

Equality (1.1) is the cubic equation $x^3 – x^2 – 2x = 0$. It has three solutions: $x_1 = –1$, $x_2 = 0$, and $x_3 = 2$; see Figure 1.2. By Theorem 1.1, the minimum point of this function should be sought among these quantities. Find the values of this function at these points: $f(x_1) = –5$, $f(x_2) = 0$, and $f(x_3) = –32$. The smallest of the obtained numbers is the minimum of the considered function. Therefore, the point $x_3 = 2$ is the solution of the given minimization problem.

Because of Example 1.2, we refine the scheme for solving the function minimization problem using Fermat theorem. If several solutions are found during the analysis of the stationary condition, then the minimum point of the given function is one of them, which corresponds to the smallest of the values of the function on these solutions.

The results obtained lead to the need to clarify the concepts used, in particular, extremum points[7]; see Figure 1.3.

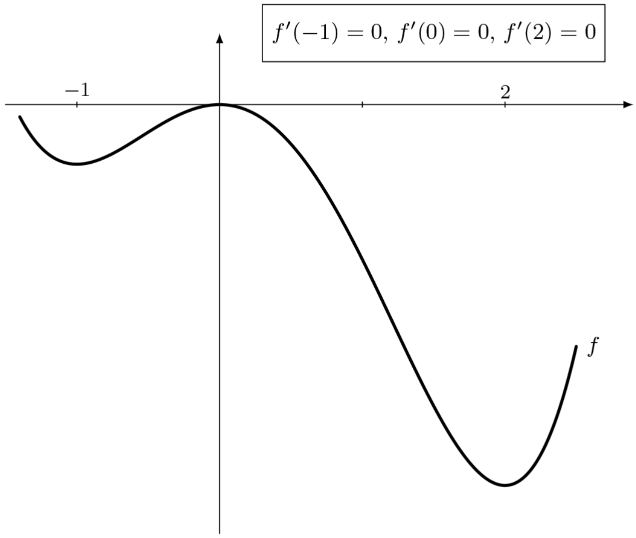

Figure 1.2 Stationary points for Example 1.2.

Definition 1.2 *A function f has the* ***local minimum*** *(****local maximum****, respectively) at a point x if there exists a neighborhood*[8] *V of this point such that the following inequality holds $f(x) \leq f(y)$ ($f(x) \geq f(y)$, respectively) for all $y \in V$. If in these relations the equal sign is possible only for $y = x$, then we have a* ***strict local minimum*** *(****maximum****). If these inequalities are valid for all values of y, then x is the point of* ***absolute*** *or* ***global minimum*** *(****absolute*** *or* ***global maximum****) of the function f.*

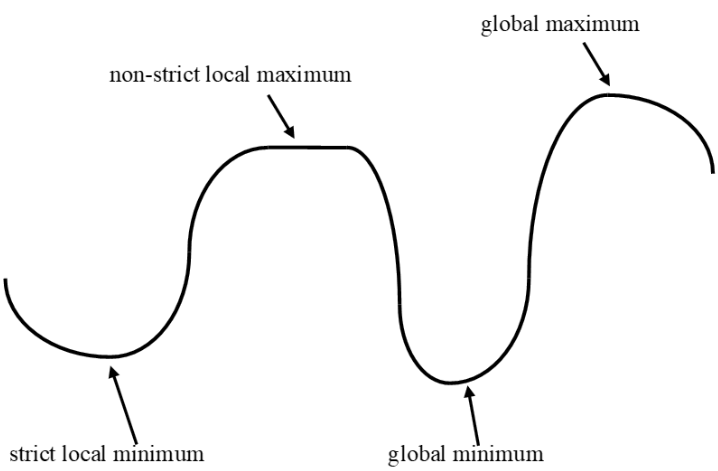

Figure 1.3 Types of function extrema.

Obviously, of the three stationary points in Example 1.2, the first corresponds to the local minimum of the considered function, the second to the local maximum of the function, and the third to its global minimum; see Figure 1.2.

Let some relation Q be given, which may or may not be satisfied by some objects from the set on which the extremal problem P is carried out[9].

Definition 1.3 *The relation Q is called a **necessary extremum condition** for problem P if any solution to this problem satisfies relation Q. The relation Q is called a **sufficient extremum condition** for problem P if any object satisfying it turns out to be a solution to problem P.*

If the extremum conditions are necessary and sufficient[10], then its solutions and only they turn out to be solutions of the extremum problem under study, i.e., extremum problem and extremum conditions conditions are equivalent; see Figure 1.4. In the general case, the set of solutions to the stationary condition turns out to be wider than the set of minimum points of the considered function[11], i.e., the stationary condition is a necessary condition for the local function minimum[12]. Particularly, the stationary condition for Example 1.1 is a necessary and sufficient condition for the function minimum, but for Example 1.2 it is necessary, but not sufficient[13].

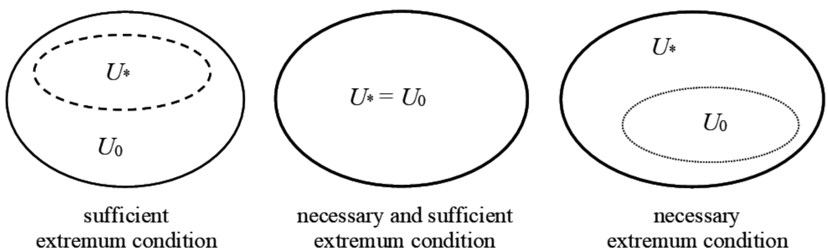

Figure 1.4 Relations between the set U_0 of extremum problem solution and the set U_* of extremum condition solution.

Obviously, the maximum of the function f corresponds to the minimum of the function g, characterized by the equality $g(x) = -f(x)$. As a result, maximization problems can always be reduced to minimization problems, and therefore do not require the development of a special theory. In particular, the stationary condition can also be used to find the maxima of a function.

The following example illustrates another important property of the problem of finding extrema.

Example 1.3 *Find a minimum of the function $f(x) = x^4 - 2x^2$.*

From the stationary condition, we find three solutions of the corresponding algebraic equation (1.1) $x_1 = -1$, $x_2 = 0$, and $x_3 = 1$; see Figure 1.5. The second of them corresponds to the local maximum of the function under consideration, and the other two are solutions to the problem of its minimization, since this function takes the same value on them[14]. In this case, we have simultaneously the absence of both the uniqueness of the solution of the problem and the sufficiency of the stationary condition[15].

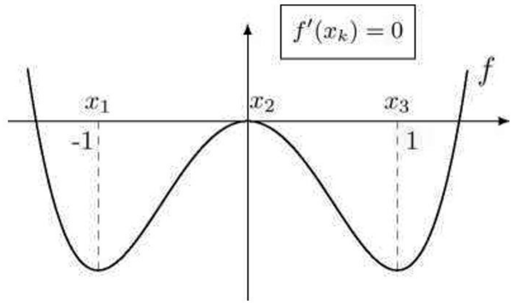

Figure 1.5 The function has two points of minimum.

The non-uniqueness of the minimum points of the function in the example considered is due to the fact that this function is even, i.e., invariant with respect to changes in the sign of the argument[16]. Naturally, the ambiguity of the solution to the problem of finding the function extremum can be observed in the absence of this property.

Example 1.4 *Find a maximum of the function* $f(x) = x^6 - 6x^5 + 13x^4 - 12x^3 + 4x^2$.

The corresponding stationarity condition is a fifth-order algebraic equation

$$3x^5 - 15x^4 + 26x^3 - 18x^2 + 4x = 0.$$

It has the following solutions

$$x_1 = 0,\ x_2 = 1 - 1/\sqrt{3},\ x_3 = 1,\ x_4 = 1 + 1/\sqrt{3},\ x_5 = 2.$$

Obviously, the first, third, and fifth solutions correspond to the absolute maxima of the function under consideration, and the second and fourth are the points of its local minima; see Figure 1.6. Thus, the problem under consideration has three

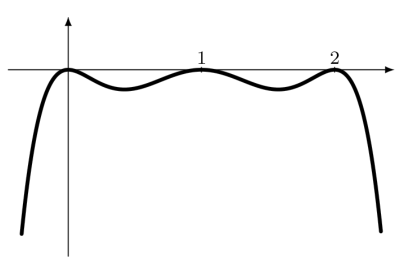

Figure 1.6 The function has three maximum points.

solutions, and the stationary condition is a necessary but not sufficient condition for the maximum[17].

As follows from the next examples, the set of global minimum points of the function can be quite large.

Example 1.5 *Find a minimum of the function* $f(x) = \sin x$.

The corresponding stationary condition $\cos x = 0$ has an infinite set of solutions $x_k = \pi/2 + k\pi$, where k is an arbitrary integer; see Figure 1.7. In this case, even values of k correspond to absolute minimum points, and odd values correspond to maximum points. Thus, the stationary condition for the considered function is not a sufficient minimum condition, and the set of solutions of the corresponding minimization problem turns out to be infinite[18].

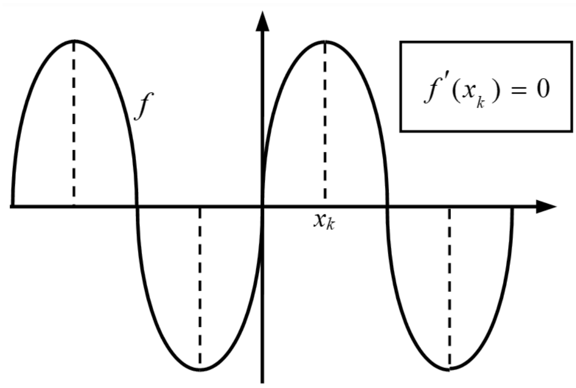

Figure 1.7 The function has an infinite set of minimum points.

A situation is possible when the set of minimum points of a differentiable function turns out to be even wider.

Example 1.6 *Find a minimum of the function* $f = f(x)$ *that is equal to* $(x + 1)^2$ *for* $x < -1$, *to zero for* $-1 \le x \le 1$ *and to* $(x - 1)^2$ *for* $x > 1$.

The derivative of this function vanishes at all points from the segment $[-1, 1]$, which are the points of its minimum; see Figure 1.8. Thus, in this case, the stationary condition is a necessary and sufficient condition for a minimum, and the set of solutions to the problem of minimizing the function f is not only infinite, but not even countable[19].

Consider another property of extremum problems[20].

Example 1.7 *Find a minimum of the function* $f(x) = x^2/(1 + x^4)$.

Find the derivative

$$f(x) = \frac{2x(1 - x^4)}{(1 + x^4)^2}.$$

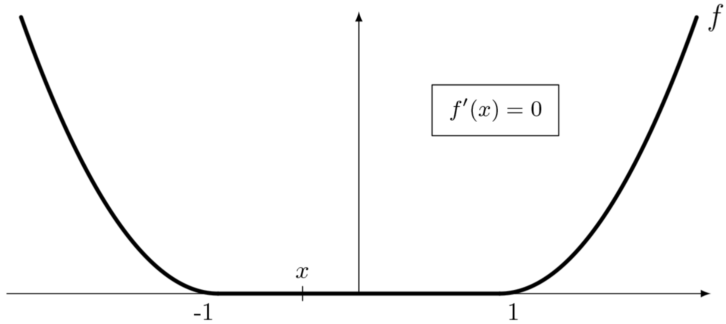

Figure 1.8 The set of points of minimum is uncountable.

The corresponding stationary condition has three solutions −1, 0, and 1. Determine the value of the function at these points: $f(0) = 0$, $f(-1) = f(1) = 1/2$. Obviously, 0 is the absolute minimum point of this function, and the values −1, 1 are its absolute maximum points; see Figure 1.9. Thus, the problem of minimizing the function f has a unique solution, and the corresponding stationary condition gives a necessary but not sufficient minimum condition. We have already encountered a similar situation in Example 1.2. However, this example has another surprising property.

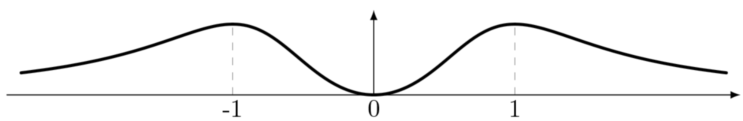

Figure 1.9 The function from Example 1.2.

Consider the sequence $\{x_k\}$ characterized by the equalities[21] $x_k = k$, $k = 1, 2, \ldots$. Obviously, the corresponding sequence of values $\{f(x_k)\}$ converges to zero, i.e., to the minimum of the considered function. Sequences that have this property are called ***minimizing sequences***. At the same time, the sequence $\{x_k\}$ itself diverges and does not have a minimum point (problem solution) as its limit. The result obtained leads to the following concept[22].

Definition 1.4 *A function minimization problem is called* ***Tikhonov well-posed*** *if any minimizing sequence for it converges to the point of minimum for this function.*

Thus, in Example 1.7, we consider the problem of minimizing a function that is ***Tikhonov ill-posed***[23].

Many methods for the approximate solution of problems in the theory of extremum are based on iterative processes in which the subsequent approximation is selected in such a way that the corresponding value of the minimized quantity turns out to be less than its previous value[24]. For ill-posed problems, such methods do not guarantee finding the minimum point of the function with the desired exactness, even if they converge[25].

In this section, functions were considered for which the corresponding stationary conditions had a non-unique solution. However, the opposite situation is possible, when the stationary condition has no solution at all.

1.1.3 Absence of function minimum points

We continue to apply Fermat theorem to find the minima of functions.

Example 1.8 *Find a minimum of the function $f(x) = x$.*

The stationary condition in this case is reduced to the equality $1 = f'(x) = 0$, which has no solution. It is clear that the problem of minimizing the considered function also has no solution; see Figure 1.10. Nevertheless, since the sets of solutions to this problem and relation (1.1) coincide (both are empty), the extremum condition used is necessary and sufficient.

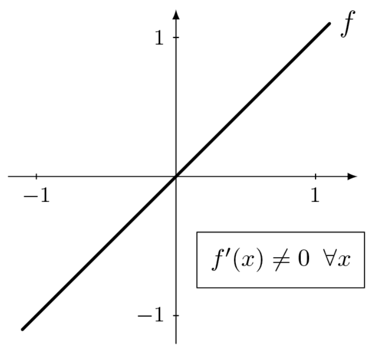

Figure 1.10 There are no stationary points for the unsolvable extremal problem.

The absence of stationary points allows us to conclude that the problem of minimizing the function under consideration turns out to be unsolvable[26]. Note that the insufficiency of the extremum condition and the absence of a minimum of the function are independent properties. In particular, Example 1.2 has a minimum but no sufficiency, and Example 1.7 has sufficiency but no minimum. The following example provides additional information about the relationship between these properties.

Example 1.9 *Find a minimum of the function $f(x) = x^3$.*

The necessary extremum condition has a unique solution $x = 0$ that does not minimize the function f; see Figure 1.11. The problem of its minimization has no solution, so relation (1.1) is a necessary but not sufficient condition for an extremum.

It is characteristic that in the last example the only point of stationary is not even a point of local extremum of the considered function[27]. In this case, we are dealing with both the absence of minimum points and the insufficiency of the extremum condition[28].

We note one more possible situation.

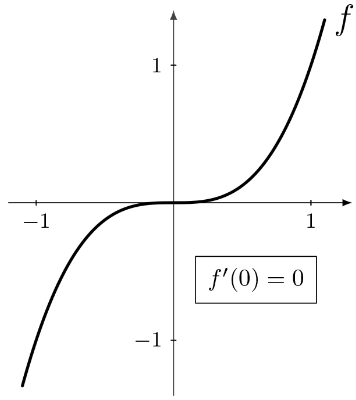

Figure 1.11 There are no stationary points for the unsolvable extremal problem.

Example 1.10 *Find a minimum of the function* $f(x) = x^3 - 3x$.

The stationary condition here has two solutions: $x_1 = -1$ and $x_2 = 1$. The first of these corresponds to a local maximum, and the second to a local minimum of the function f, while there is no global minimum; see Figure 1.12. We again face the lack of sufficiency of the extremum condition in the case of insolvability of the minimization problem, however, with a different combination of properties than in the previous example.

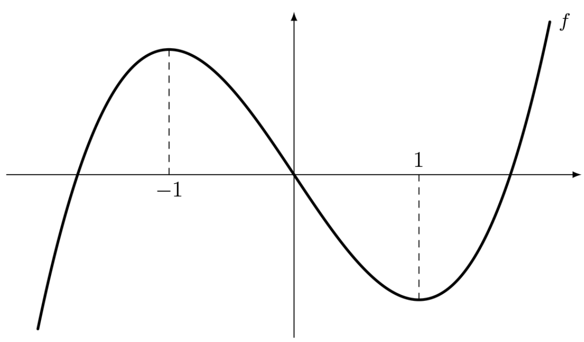

Figure 1.12 Two stationary points correspond to the local maximum and minimum.

It remains to clarify one more question related to the applicability of Fermat theorem.

1.1.4 Inapplicability of Fermat theorem

The following examples correspond to cases where the stationary condition is inapplicable for finding the minimum of a given function.

Example 1.11 *Find a minimum of the function* $f(x) = |x|$.

Due to the fact that f is not differentiable, Fermat theorem is not applicable here[29]. Thus, the search for a real-life minimum point $x = 0$ (see Figure 1.13) requires the involvement of another mathematical apparatus.

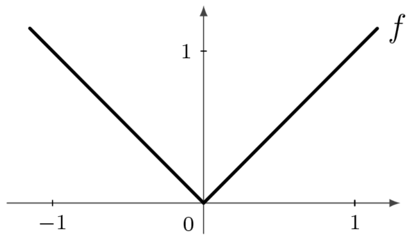

Figure 1.13 The stationarity condition is not applicable because of the non-smoothness of the function.

Example 1.12 *Find a minimum of the function* $f(x) = (x+1)^2$ *on the interval* $[0, 1]$.

The stationary condition has a unique solution $x = -1$; see Figure 1.14. However, this value is not a solution to the problem under consideration, since it lies outside the specified interval. Thus, the application of the stationary condition in this case does not lead to the desired results[30], because in Theorem 1.1, the problem of the unconditional extremum of a function was considered. It is clear that in the presence of restrictions, the solution of the problem may not satisfy the stationary condition[31].

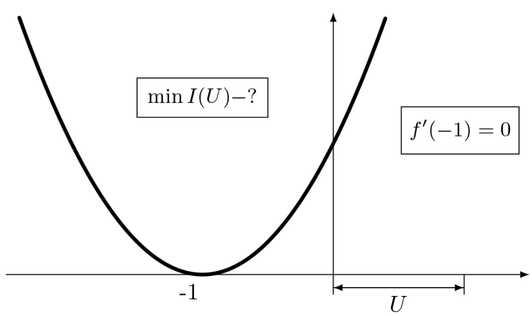

Figure 1.14 In the conditional extremum problem, the solution of the stationary condition is not a minimum point.

RESULTS

Here is a list of questions in the field of problems of minimizing the function and stationary conditions, the main conclusions on this topic, as well as the problems that arise in this case, partially solved in Appendix, partially taken out in Notes.

Questions

It is required to answer questions about the properties of the problem of minimizing a function of one variable and the stationary condition.

1. What are the applicability conditions of Fermat theorem?
2. What class of mathematical problems does the stationary condition belong to?
3. Applying the stationary condition, we do not solve the minimization problem, but reduce it to a different problem. What justifies such an approach?
4. What is the difference between local and absolute extremum?
5. What is the difference between strict and non-strict extremum?
6. What is the difference between a necessary and sufficient condition for an extremum?
7. Why is the stationary condition satisfied not only for the minimum points of the function, but also for the maximum points?
8. Why is the stationary condition satisfied not only for the global extremum points of the function, but also for its local extremum points?
9. What can be the relationship between the set of solutions to the problem of minimizing a function and the set of stationary points?
10. Can the number of stationary points be less than the number of local extrema of the function?
11. What should be done if the solution of the stationary condition is not unique?
12. How can one distinguish between the minimum and maximum points of a function that equally satisfy the stationary condition?
13. Suppose that two solutions to the stationary condition have been found. How can we know if we have the insufficiency of the extremum condition or with the non-uniqueness of the minimum points?
14. Assume that a continuous function has several local extrema. How can the various minimum and maximum points be located relative to each other?
15. How many solutions can the stationary condition have?
16. How many solutions can the problem of minimizing a continuous function have?
17. What properties can solutions of the stationary condition for a function of one variable have in the general case?
18. What is the difference between the properties of Examples 1.2 and 1.6?

19. Can a Tikhonov well-posed problem not have a solution?
20. Why are minimization problems with non-unique solutions not Tikhonov well-posed?
21. What difficulties may arise in the practical implementation of problems that are Tikhonov ill-posed?
22. Why the stationary condition makes sense even in the absence of a solution to the function minimization problem?
23. What properties can the stationary condition have if the problem is unsolvable?
24. What conclusion can be drawn in the absence of a solution to the stationary condition?
25. In what cases can the stationary condition be inapplicable?
26. Why is the stationary condition inapplicable for non-smooth functions?
27. Why is the stationary condition inapplicable for the conditional extremum problem?
28. Can the stationary condition be effective in the problem on the conditional extremum of a function?
29. What can be the solution to the problem of minimizing a continuous function on an interval?

Conclusions

Based on the study of the minimization problem for functions using Fermat theorem and the results of the analysis of considered examples, we can have the following conclusions.

- The unconditional function minimization problem can have one, several, an infinite set of minimum points, or not have them at all.
- To solve the minimization problem for a differentiable function of one variable on the set of real numbers, one can use Fermat theorem.
- Fermat theorem reduces the function minimization problem to the stationary condition.
- The stationary condition is an algebraic equation.
- The stationary condition is satisfied both for the minimum points of the function and for its maximum points.
- The stationary condition is fulfilled not only for the absolute extremum points of the function, but also for the local extremum points.

- Fermat theorem gives a necessary condition for a local extremum of a function, i.e., any local extremum point satisfies the stationary condition, but solutions to the stationary condition may not be the local extremum points of the function.
- If the stationary condition has several solutions, then the optimal one is the one that corresponds to the smallest of the values of the minimized function on these solutions.
- The function minimization problem can to be Tikhonov ill-posed such that minimizing sequences are not required to converge to the minimum point of this function.
- If the stationary condition has no solutions, then this means that there are no minimum points for this function.
- The stationary condition is not applicable if a non-smooth function is considered or this is a conditional function extremum.

Problems

On the basis of the results obtained above, we have the following problems of function minimization, which require further analysis.

1. **Solvability**. In Examples 1.8, 1.9, and 1.10, there is no solution to the problem of minimizing the corresponding functions, while in all other cases the problem has a solution. We would like to know what properties of the function guarantee the existence of a solution to the problem. This issue is explored in Appendix.
2. **Uniqueness of solution**. In Examples 1.3, 1.5, and 1.6, the solution to the minimization problem is not unique, while in all other cases, the problem has a unique solution if, of course, it exists. It is desirable to establish the properties of the function that guarantee the uniqueness of the problem solution. This problem is analyzed in Appendix.
3. **Tikhonov well-posedness**. In Example 1.7, we considered the Tikhonov ill-posed problem. It would be interesting to establish under what conditions such a problem turns out to be well-posed. This issue is considered in Appendix.
4. **Concept of approximate solution**. The Tikhonov ill-posedness of the problem actualized the question of what is meant by an approximate solution to the problem. The corresponding definitions are given in the next chapter.
5. **Sufficiency of extremum condition**. In Example 1.2, we discovered the insufficiency of the stationary condition by calculating the values of the function in the found solutions. One would like to check whether the solutions of the stationary condition are minimum points without calculating the function in them. This seems to be relevant, at least in the case when only one solution of the stationary condition is somehow found so that there is nothing to compare

the value of the function at this point with. This problem is investigated in Appendix.

6. **Algorithm for solving**. In all the examples considered, extremely simple functions were investigated. As a result, finding solutions to the stationary condition did not cause any serious difficulties. However, in the general case, we are dealing with a non-linear algebraic equation, the solution of which is difficult to find explicitly. It is necessary to be able to find stationary points for fairly complex functions. This problem is explored in the next chapter.

7. **Non-smooth problems**. In Example 1.11, we could not use the stationary condition because of non-differentiability of the minimized function. It would be desirable to establish some generalization of the stationary condition in the case of non-smooth functions. This result is given in Appendix.

8. **Global extremum**. Fermat theorem gives a necessary condition for a local extremum. The problem of finding the global extremum is much more difficult. About methods for finding the global extremum see Notes[32].

9. **Conditional extremum**. In Example 1.12, the stationary condition was ineffective due to the presence of additional restrictions. The development of methods for solving the problem for the conditional extremum of a function seems to be very important. In the next chapter, we consider the problem of minimizing a function on an interval and the problem of minimizing a function of two variables related by some equality. More difficult conditional extremum problems related to optimal control theory are explored in later parts of the book.

10. **Alternative methods**. To solve the problem of minimizing functions, the stationarity condition was used. However, there are other methods for solving such problems; see Notes[33].

11. **Dependence on parameters**. When solving applied problems for an extremum, the value to be minimized often depends on some parameters. In practice, these parameters are known with some errors. In this regard, the question arises, how small an error in determining the parameters will affect the accuracy of solving the minimization problem? Let us also note such a circumstance. The minimized functions in Examples 1.2 and 1.3 and, respectively, in Examples 1.9 and 1.10 are quite close. In particular, they are polynomials of the same order, differing only in lower terms. At the same time, the properties of the corresponding problems differ significantly. Particularly, we considered the function $f(x) = x^3 – ax$, where the parameter a is zero in Example 1.9 and three in Example 1.10. However, in the first case, there is a unique solution of the stationary condition, which is an inflection point, and in the second, there are two solutions, which are points of local extrema. Obviously, for arbitrarily small positive values of this parameter, the function will have the same properties as in Example 1.10. However, with $a = 0$, we get Example 1.9 with qualitatively

different properties. Consequently, with a small change in the parameter, the problem of minimizing a function qualitatively changes its properties. Such effects are discussed in the next chapter.

12. **Generalizations**. We have considered the simplest problem of minimizing a function of one variable. In Appendix, the established results are naturally extended to functions of many variables. In the subsequent parts of the book, optimal control problems will be considered in which functionals defined on a certain class of functions are minimized under additional constraints.

1.2 APPENDIX

When analyzing the examples discussed earlier in the Lecture, we had some problems that need additional research. There are the problems of existence and uniqueness of a function minimum, a sufficient condition for an extremum, and the minimization of a non-smooth function. Below, we present results that clarify the situation, and the examples given earlier will be considered as applications. The final section provides information about the problem of minimizing functions of many variables.

1.2.1 Existence of function minimum

In most of the examples considered earlier, the minimum of the function exists, but in some cases the minimization problem was be unsolvable. All these examples are quite simple, so that it was quite easy to find a solution to the problem or to establish the absence of its solution. However, it would be extremely interesting to know why some functions have minima and others do not. The classical result in the field of solvability of extremum theory problems is the following assertion[34].

Theorem 1.2 (***Weierstrass theorem***). *The continuous function on a closed*[35] *bounded*[36] *set has its minimum and maximum.*

Proof. Let a continuous function f defined in a closed bounded domain U of the real line be given. Then the set of values $f(U)$ is bounded. Therefore, there exists $\inf f(U)$, and hence a sequence $\{x_k\}$ of elements of the set U such that $f(x_k) \to \inf f(U)$. Since the set U is bounded, this sequence is bounded too. Using the Bolzano–Weierstrass theorem[37], we extract from it a subsequence $\{x_s\}$ such that $x_s \to x$. Since the set U is closed[38], the inclusion $x \in U$ is true. Taking into account the continuity of the function f, we obtain the convergence $f(x_s) \to f(x)$. However, the entire sequence $\{f(x_k)\}$, and hence any of its subsequences, has the number $\inf f(U)$ as its limit. This implies that there exists a point x from the set U for which the equality $f(x) = \inf f(U)$ holds. Thus, the considered function reaches its minimum on the given set. The existence of its maximum is proved similarly. □

The requirement that the set on which the function is minimized be closed is very essential. Thus, in the problem of minimizing the function $f(x) = x$ on an open interval $(0, 1)$, we have the function that is continuous in a bounded set. However, this problem has no solution due to the absence of a minimum positive number. The same function on the closed unbounded set $(-\infty, 0]$, also has no minimum. On the other hand, the function $f(x) = 1/x$ has no extremum on the closed bounded set $[-1, 1]$, since it is not there continuous.

We considered the minimization problem for the function $f(x) = (x+1)^2$ on the interval $[0, 1]$; see Example 1.11. We are indeed dealing with a continuous function defined on a closed bounded set. Thus, this problem has a solution due to Theorem 1.2. However, for all the other examples considered earlier, we are unable to use this result, since they consider problems for an unconditional extremum, which means that we are talking about minimizing a function on the real line, which is not a bounded set. Nevertheless, in most of the considered examples, the minimum of the function exists. In this regard, we would like to have a statement that guarantees the solvability of the minimization problem for functions in unbounded domains. Such a result can be established if the function satisfies some additional property.

Definition 1.5 *A function f is called* ***coercive*** *if for $|x_k| \to \infty$ there is a convergence $f(x_k) \to +\infty$.*

The following assertion is true[39].

Theorem 1.3 *The continuous coercive function lower bounded on a closed subset of the real line reaches its minimum.*

Proof. Since the function f is lower bounded on the closed set U, its infimum $\inf I(U)$ exists. Thus, there is a sequence $\{x_k\}$ of elements of the set U such that $f(x_k) \to \inf f(U)$. Suppose that this sequence is not bounded, i.e., $|x_k| \to \infty$. Then, due to the coercivity of the function f, we obtain $f(x_k) \to +\infty$. However, it was previously established that the sequence $\{f(x_k)\}$ has the lower bound of this function as its limit. Therefore, the assumption that the sequence $\{x_k\}$ is unbounded is not true. The proof ends in the same way as in Theorem 1.2. □

Let us now turn to the examples considered earlier. Everywhere, with the exception of Example 1.11, the whole set of real numbers, which is closed, was used as the domain of definition. All considered functions were continuous. In Example 1.1, the function $f(x) = x^2$ was minimized, which is obviously bounded below and coercive. Thus, the existence of its minimum follows from Theorem 1.3. The functions $f(x) = 3x^4 – 4x^3 – 12x^2$, $f(x) = x^4 – 2x^2$ and $f(x) = |x|$ considered in Examples 1.2, 1.3 and 1.10, as well as the function from Example 1.6, have similar properties. In all these cases, the solvability of the minimization problem is also established using Theorem 1.3. On the other hand, this statement is not applicable to the functions $f(x) = x$, $f(x) = x^3$ and $f(x) = x^3 – 3x$ considered in Examples 1.8, 1.9 and 1.10 because they are unbounded from below. All of them do not have minimums. There remains the function $f(x) = \sin x$ from Example 1.5, which, being bounded, is not coercive. However, it reaches its minimum. Note that the conditions of Theorems 1.2 and 1.3 are sufficient conditions for the existence of a minimum of the function, i.e., violation of these properties does not mean that the corresponding minimization problem turns out to be unsolvable[40].

1.2.2 Uniqueness of function minimum

In many of the examples considered earlier, the minimum point of the function was unique, but in some cases, the solution was not unique. In this regard, we would like to reveal the properties of the function, under which the uniqueness of the solution of the minimization problem would be guaranteed.

Definition 1.6 *A function f determined on the interval $[a, b]$ is* ***convex***[41] *if the following inequality holds*

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \;\; \forall x, y \in [a, b], \alpha \in (0, 1).$$

If in this relation the equal sign is realized only when $x = y$, then the function is called ***strictly convex***.

The convexity of a function defined on the whole number axis is defined similarly. Geometrically, the convexity of a function means that the part of the curve $f = f(x)$, connecting the points with coordinates $(x, f(x))$ and $(y, f(y))$, is located no higher (for a strictly convex function – below) the line segment connecting these points; see Figure 1.15.

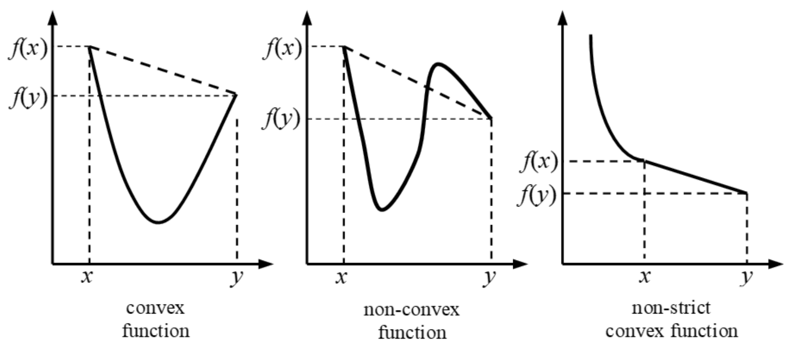

Figure 1.15 Convexity of functions.

The sufficient condition for the uniqueness of the minimum of a function gives the following statement[42].

Theorem 1.4 *The strictly convex function on an interval or on a real line has at most one minimum point*[43].

Proof. Suppose there are two distinct points x and y of the minimum of a strictly convex function f on the set U, which is a segment or the entire set of real numbers. Then the point $\alpha x + (1-\alpha)y$ for any number $\alpha \in (0, 1)$ belongs to the considered set[44]. Taking into account the strict convexity of the function, we establish the inequality

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y) = \alpha \min f(U) + (1-\alpha) \min f(U) = \min f(U).$$

Thus, there is such an element of the set U, on which the value of the functional is less than the minimum possible value. Thus, the assumption of the existence of two different minimum points led to a contradiction. □

Let us make sure that the function $f(x) = x^2$ considered in Example 1.1 is strictly convex. Indeed, we have

$$f((\alpha x + (1–\alpha)y)–\alpha f(x)–(1–\alpha)f(y) = [\alpha x + (1–\alpha)y]^2–\alpha x^2–(1–\alpha)y^2$$
$$= \alpha(\alpha - 1)(x^2–2xy + y^2) = \alpha(\alpha - 1)(x–y)^2.$$

Obviously, the value on the right side of this equality is not positive because of the inclusion $\alpha \in (0, 1)$, and the equality to zero here is possible only for $x = y$. Thus, the function under consideration turns out to be strictly convex, and the uniqueness of the previously established solution to the problem of its minimization follows from Theorem 1.4. Similarly, the uniqueness of the minimum point of the function $f(x) = (x + 1)^2$ on the interval [0,1] described in Example 1.12 is established. On the other hand, the functions in Examples 1.3 and 1.5 are non-convex, and the function in Example 1.6 is convex, but not strictly convex; see Figure 1.8. As a result, it is clear why the problems of their minimization have not a unique solution. However, the function $f(x) = |x|$ from Example 1.11, being not strictly convex, has a unique minimum point[45]. A unique minimum exists both for the non-convex function shown in Figure 1.9 and for the non-convex function $f(x) = 3x^4–4x^3–12x^2$ considered in Example 1.2. However, in the latter case, there is additionally a local minimum. An example of a non-convex function with a unique minimum point is shown in Figure 1.16[46].

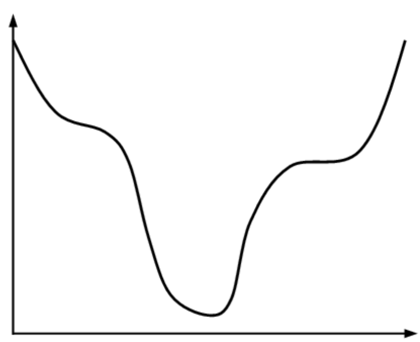

Figure 1.16 Non-convex function with a unique minimum.

1.2.3 Tikhonov well-posedness of problems

In Example 1.7, we considered the Tikhonov ill-posed problem of minimizing a function. A minimizing sequence was specified that did not converge to the minimum point of the given function. Below, we present one result on the well-posedness of the function minimization problem of a general form. In this case, the following property will be used, which is a strengthened version of the convexity of the function[47].

Definition 1.7 *A function f determined on the interval* $[a, b]$ *is* ***strongly convex*** *if the following inequality holds*

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y) - c\alpha(1-\alpha)(x-y)^2 \ \ \forall x, y \in [a, b],\ \alpha \in (0, 1),$$

where c is a positive number.

The following assertion is true[48].

Theorem 1.5 *If the function f strongly convex, then the problem of its minimization is Tikhonov well-posed.*

Proof. Let x be the minimum point of the given function, and let $\{x_k\}$ be an arbitrary minimizing sequence, i.e., $f(x_k) \to \inf f$. To prove the theorem, it suffices to establish the convergence $x_k \to x$.

We have

$$f(\alpha x_k + (1-\alpha)x) \le \alpha f(x_k) + (1-\alpha)f(x) - c\alpha(1-\alpha)(x_k - x)^2.$$

This implies the inequality

$$f(\alpha x_k + (1-\alpha)x) - f(x) \le \alpha[f(x_k) - f(x)] - c\alpha(1-\alpha)(x_k - x)^2.$$

By optimality of the number x, the value at the left-hand side of this inequality is non-negative. After division by α, we obtain

$$c(1-\alpha)(x_k - x)^2 \le f(x_k) - f(x).$$

The parameter α is arbitrary, so we can pass to the limit as $\alpha \to 0$. Thus, we obtain the inequality

$$0 \le c(x_k - x)^2 \le f(x_k) - f(x).$$

The sequence $\{x_k\}$ is minimizing sequence. Therefore, the value on the right-hand side here tends to zero. Now we have the convergence $(x_k - x)^2 \to 0$, i.e., the sequence $\{x_k\}$ tends to the point of minimum x. Thus, this problem is Tikhonov well-posed. □

Example 1.1 considers the function $f(x) = x^2$. Check its convexity. For all values x, y and $\alpha \in (0, 1)$ we find

$$f(\alpha x + (1-\alpha)y) - \alpha f(x) - (1-\alpha)f(y) = [\alpha x + (1-\alpha)y]^2$$
$$= -\alpha(1-\alpha)(x^2 - 2xy + y^2) = -\alpha(1-\alpha)(x-y)^2.$$

Thus, the strong convexity property is true in the form of equality with constant $c = 1$, so the problem is well-posed. However, the function from Example 1.7 is non-convex; see Figure 1.9. Therefore, the ill-posedness of the corresponding minimization problem is quite natural.

1.2.4 Sufficient conditions of function minimum

In Example 1.1, the stationary condition was a necessary and sufficient condition for the minimum of functions, while in the following examples, individual stationary points were not solutions to the minimization problem. In these examples, we found all solutions of the stationary condition, calculated the values of the function at each of these points, and then found the solution to the problem, leaving only those points at which the value of the function turned out to be the smallest. Naturally, such a procedure is possible only in exceptional situations, when the considered function is sufficiently simple, and we can find all stationary points. A more realistic case is when a stationary point is found, and one would like to know whether it is a point, if not of an absolute, then at least of a local minimum of a given function. Below is one result in this direction[49].

Theorem 1.6 *If at a stationary point the second derivative of the function is positive (respectively, negative), then at this point a strict local minimum (respectively, maximum) of this function is reached.*

Proof. Suppose that for the function $f = f(x)$ at some point x_0, the conditions $f'(x_0) = 0$ and $f''(x_0) > 0$ are true. By the last inequality, the first derivative of the function at the point x_0 increases. Considering that the value of the derivative at the point itself is equal to zero, we conclude that for a sufficiently small positive number ε, the derivative $f'(x)$ is negative if $x \in (x_0 – \varepsilon, x_0)$ and positive if $x \in (x_0, x_0 + \varepsilon)$. Using the Lagrange finite-increments formula, for any point $x \in (x_0 – \varepsilon, x_0)$ the following equality holds $f(x) = f(x_0) – f'(\xi)(x_0 – x)$, where $\xi \in (x, x_0)$. Taking into account the negativeness of the derivative, we conclude that $f(x) > f(x_0)$. Similarly, for any point $x \in (x_0, x_0 + \varepsilon)$ we have $f(x) = f(x_0) + f'(\xi)(x – x_0)$. Taking into account the positiveness of the derivative, we conclude that $f(x) > f(x_0)$. Thus, for sufficiently small ε, the last inequality is true for any point from the interval $(x_0 – \varepsilon, x_0 + \varepsilon)$, other than x_0. Therefore, x_0 is indeed a strict local minimum point of this function. The properties of the local maximum are established similarly. □

Let us use this theorem for further analysis of the previously described examples. Example 1.1 considered the function $f(x) = x^2$, which has a unique stationary point $x = 0$. Its second derivative is 2, i.e., positive. Then, by virtue of Theorem 1.5, this point reaches a minimum, which was established earlier. Example 1.2 considered the function $f(x) = 3x^4 – 4x^3 – 12x^2$ with three stationary points $x_1 = –1$, $x_2 = 0$, $x_3 = 2$. Find the second derivative $f''(x) = 12(3x^2 – 2x – 2)$. Determine the corresponding values at the indicated points: $f''(–1) = 36$, $f''(0) = –24$, $f''(2) = 72$. As a result, we conclude that a strict local minimum is realized at the points x_1 and x_3, and a strict local maximum is realized at the point x_2, which is consistent with the results obtained earlier; see Figure 1.2. Similar reasoning can be carried out for Examples 1.3 and 1.5.

Theorem 1.6 does not consider the case when the second derivative of a function vanishes at its critical point, which is the case for Example 1.10. However, the corresponding results can be easily obtained using higher-order derivatives[50].

1.2.5 Minimization of non-smooth functions

In almost all the cases considered earlier, differentiable functions were investigated. However, in Example 1.11, a non-smooth function $f(x) = |x|$ was given. The question arises whether the stationary condition can be somehow extended to problems of this nature. A similar result could be obtained by generalizing the concept of a derivative.

Definition 1.8 *A number p is called a* ***subgradient*** *of the function f at the point x_0 if for any number x the following inequality holds*

$$f(x) \geq f(x_0) + p(x - x_0).$$

The set of all subgradients of a function at a given point is called its ***subdifferential*** *and is denoted by $\partial f(x_0)$. If the subdifferential of a function at a point is not empty, then the function is said to be* ***subdifferentiable*** *at the given point.*

The function subgradient has a natural geometric meaning. Consider the line $l(x) = f(x_0) + p(x - x_0)$. According to the above inequality, the line l does not lie above the curve f, and they coincide at the point x_0. Then p is the tangent of the angle between the line l and the x-axis; see Figure 1.17. We recall that the derivative is the tangent of the slope of the tangent at a given point.

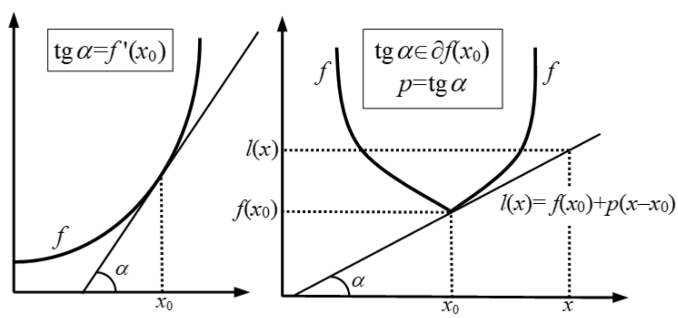

Figure 1.17 Derivative and subgradient.

If the function f is differentiable at the point x_0, then, obviously, its derivative $f'(x_0)$ is a subgradient (it suffices to compare the graphs in Figure 1.17). It is easy to verify that for continuously differentiable functions the subdifferential consists of a unique element that is the derivative, and in the case when there is a unique subgradient at some point, then the derivative of the function at this point exists and is equal to the subgradient[51]. Thus, the subgradient turns out to be a generalization of the concept of derivative[52].

Consider, as an example, the function $f(x) = |x|$. At any point other than zero, this function is differentiable, which means that its subdifferential consists exclusively of a derivative equal to −1 for negative values of the argument and 1 for its positive values. Hence, it follows that $\partial f(x) = \{-1\}$ if $x < 0$ and $\partial f(x) = \{1\}$ if $x > 0$. It remains to define the subdifferential at zero, where the given function is not differentiable. For any $p \in [-1, 1]$ we have $px \leq |p||x| \leq |x|$ for any x, i.e., $f(x_0) + p(x–x_0) \leq f(x)$ for $x_0 = 0$. Thus, the inclusion $[-1, 1] \subset \partial f(0)$. Suppose now that for $|p| > 1$ we have $p \in \partial f(0)$. Then, using the definition of a subgradient, we establish that $|x| \geq px$ for any x. Assuming $x = p$, we get $|p| \geq |p|^2$, which cannot be. Thus, we obtain the equality $\partial f(0) = [-1, 1]$; see Figure 1.18.

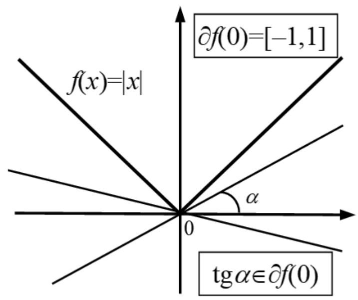

Figure 1.18 Subdifferential of the function $f(x) = |x|$ at zero.

Theorem 1.7 *If a subdifferentiable function f at a point x_0 has a minimum at this point, then the following inclusion holds $0 \in \partial f(x_0)$.*

Indeed, the inequality $f(x) \geq f(x_0) + p(x–x_0)$ for $p = 0$ is $f(x) \geq f(x_0)$. Since, according to the definition of the subgradient, this relation must hold for all x, we conclude that at the point x_0 this function has a minimum[53].

Theorem 1.7 is a generalization of Theorem 1.1, because if a function is differentiable, its only subgradient is the derivative, which, by virtue of Theorem 1.7, is equal to zero[54]. As an illustration, let us return to Example 1.11, for which $f(x) = |x|$. It was previously established that the subdifferential of this function consists of a unique value –1 for negative x, a unique value of 1 for positive x, and is a segment [–1,1] at the zero point. Obviously, the zero value of the subgradient can only be reached at zero. Thus, only the point $x_0 = 0$ can be the minimum point of the considered function. Naturally, the function $f(x) = |x|$ reaches its minimum at this point.

1.2.6 Minimization of functions of many variables

A natural generalization of Problem 1.1 is its vector analog, i.e., the problem of unconditional minimization of a function of many variables.

Problem 1.2 *Find a point of minimum for a function $f = f(x_1, x_2, ..., x_n)$.*

Having defined the vector $x = (x_1, x_2, ..., x_n)$, we can formally reduce Problem 1.2 to Problem 1.1 of minimizing the function $f = f(x)$. Almost all results related to the minimization of the function of one variable remain valid here.

The vector analog of the stationary condition $f'(x) = 0$ is the equality

$$\nabla f(x) = 0. \tag{1.5}$$

Here $\nabla f(x)$ is the ***gradient*** of the function f at the point x, i.e., a vector whose components are all partial derivatives of the function under consideration at a point x of the n-dimensional Euclidean space $\mathbb{R}^n$, and 0 is the zero element of this space, i.e., an n-dimensional vector with all components equal to zero. The vector analog of Fermat theorem is the statement that in order for a differentiable function of many variables to have a minimum at some point, it is necessary that its gradient at this point vanishes[55], i.e., relation (1.5) was satisfied. More precisely, it is a necessary condition for a local extremum of a function of many variables. We know that the stationary condition is an algebraic equation with respect to the extremum point. Analogically, the formula (1.5) is a system of algebraic equations.

A continuous function of many variables in a closed bounded set of Euclidean space reaches its maximum and minimum. Moreover, a continuous coercive function of several variables bounded from below on a closed subset of the Euclidean space reaches its minimum. The function f is called ***coercive*** here if the convergence $f(x) \to +\infty$ takes place as $|x| \to \infty$, where $|x|$ is the modulus of the vector.

The formulation of the uniqueness theorem for a function of many variables uses an extremely important property of the domain of the considered function[56].

Definition 1.9 *A subset U of Euclidean space is called* ***convex*** *if the following condition holds*

$$\alpha x + (1-\alpha)y \in U \quad \forall x, y \in U,\ \alpha \in (0,1).$$

The geometric meaning of this definition is as clear as possible for flat sets. In particular, a set is convex if, with any two of its points, it completely contains the line segment connecting these points; see Figure 1.19. Naturally, the segment $[a, b]$ is a convex subset of the real line.

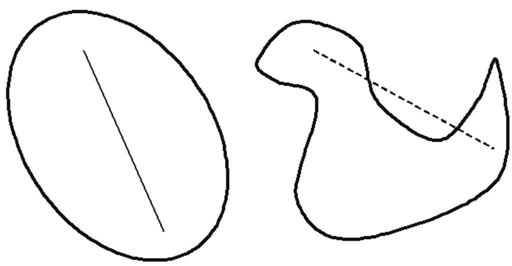

Figure 1.19 Convexity of sets on the plane.

As is known, sufficient conditions for the extremum of a function of one variable at the stationary point were determined by the sign of its second derivative. For a function of n variables, the analog of the second derivative is the ***Hessian***, which is the square matrix of the n order, the elements of which are all possible partial derivatives of the second order of the given function at the point under consideration. If the Hessian of the function at the stationary point, i.e., on the solution of the system of equations (1.5), is positive definite, then this point is the point of the local minimum of the function, and if it is negative definite, then this is the point of its local maximum. The matrix A is called ***positive definite*** here if the following inequality holds

$$\langle Ax, x\rangle > 0 \;\; \forall x \in \mathbb{R}^n, \; x \neq 0,$$

and ***negative definite***, if here the $>$ sign is replaced by $<$, and the ***dot product*** of the corresponding vectors (the sum of the products of their components) is on the left side of this inequality.

If a function of many variables has a minimum at some point, then the zero vector belongs to the ***subdifferential*** of the function at this point, where the subdifferential of the function f at the point x_0 consists of all vectors p (***subgradients***) satisfying the following inequality[57].

$$f(x) \geq f(x_0) + \langle p, x - x_0\rangle \;\; \forall x \in \mathbb{R}^n.$$

Additional conclusions

Based on the results presented in Appendix, some additional conclusions can be drawn about the function minimization problem.

- The existence of a minimum and a maximum of a function is guaranteed if the function itself is continuous and the area on which it is minimized is closed and bounded.
- The existence of a minimum of a continuous function on an unbounded closed set is guaranteed if the function itself is lower bounded and coercive.
- In the absence of the above restrictions, the existence of a minimum of the function is possible, but not guaranteed.
- The uniqueness of the minimum point is guaranteed for a strictly convex function considered on a segment or on the entire real line.
- In the absence of the above properties, the uniqueness of the minimum of the function is possible, but not necessary.
- Tikhonov well-posedness of the function minimization problem is guaranteed in the case of strong convexity of this function.
- If at the stationary point the second derivative of the function is positive (respectively, negative), then at this point a strict local minimum (respectively, maximum) of the function is reached.
- A generalization of the stationary condition to non-smooth functions is the condition of including zero in the subdifferential of a function at the point of its local minimum.

- The previous statement can be used for the problem of minimizing a modulus that is a subdifferentiable but not differentiable function.
- If a function is differentiable at a point, then its derivative is its unique subgradient at that point.
- If a function is subdifferentiable at a point, and the subdifferential consists of a unique element, then the latter is the derivative of the function at that point.
- The results obtained are naturally extended to problems of minimizing a function of many variables.

Notes

1. Here and throughout what follows we focus on those examples in which some unusual effects are observed that are manifestations of certain difficulties. Such examples are usually called ***counterexamples***; see [77], [78].

2. A significant amount of literature is devoted to problems of minimizing functions; see, for example, [26], [41], [42], [49], [65], [70], [69], [79], [112], [132], [139], [141], [149], [180], [193]. One can also consider the problem of finding the minimum of a function of many variables and even functionals, i.e., transformations that associate an object of an arbitrary nature (for example, a function) with a certain number. Calculus of variations (see [37], [61], [208]) is related to the problems of minimizing functionals, as well as the theory of optimal control, which is the direct subject of this book.

3. The stationary condition is classified as a ***first-order extremum condition***, since it uses only the first derivative of the function being minimized

4. In fact, on this idea of comparing the minimized function or functional at the minimum point and on its perturbation, the variational method is based, which underlies the theory of extremum, in particular, the calculus of variations.

5. Critical points of functionals in the calculus of variations are called ***extremals***.

6. The study of critical points of a function is connected with ***Morse theory***; see [15], [94], [131], [153].

7. The above concepts naturally extend to functions of many variables and functionals of general form.

8. ***Neighborhood*** is an essential concept in topology; see, for example, [101]. For the functions, everything is quite obvious, but when passing to functionals, it is necessary to clarify what exactly is meant by neighborhoods. It is intuitively clear that the neighborhood of a point is the set of points sufficiently close to it. However, in the general case, a point can be understood as a function or even a vector function, as a result of which the meaning of the concept of proximity requires clarification. In particular, in the calculus of variations, the conditions of a ***weak*** and ***strong extremum*** are considered, which just differ in the definition of the neighborhoods, see [37], [61].

9. The problem of finding the unconditional minimum of a function of one variable is not necessarily considered here. P can be understood as a problem of minimizing some functional on an arbitrary set.

10. We do not consider here proper sufficient extremum conditions. On sufficient conditions for the calculus of variations; see [37], [61], [208]. Sufficient conditions in optimal control problems are given, for example, in [70], [85], [108], [146].

11. Since the extremum condition, which is both necessary and sufficient, will be equivalent to the original problem of finding an extremum, they have the same degree of difficulty. The fact that the stationary condition is, in principle, a simpler object of study than the original problem requires a certain price. Such is the possibility of obtaining ”extra” solutions.

12. To clarify the question why “extra” points appear in the process of studying the stationary condition, let us return to the proof of Theorem (1.1). If the considered point x delivered the maximum, and not the minimum, of the function f, then the value on the left side of inequality (1.2) would have the opposite sign. As a result, instead of (1.3), we would obtain relation (1.4), and instead of inequality (1.4), condition (1.3). Thus, both inequalities (1.3) and (1.4), and hence the stationary condition (1.1), can be equally obtained in the case when x is the minimum and maximum point of the function f. Therefore, on the basis of equality (1.1) it is impossible to distinguish the minimum from the maximum of the function. If now x turns out to be only a point of only a local minimum of the function, then relation (1.2) will be valid not for all h, but only for small enough values of this parameter. In this case, nothing prevents us from again passing to the limit at $h \to 0$ and obtaining inequality (1.3), and hence the stationary condition. This explains the fact that the extremum condition (1.1) can be satisfied not only by the global extremum point of the function, but also by any point of its local extremum.

13. We will encounter the absence of sufficient optimality conditions for various optimal control problems more than once; see, in particular, Chapters 5, 6, 10, 14, and 15.

14. Pay attention to the fact that the considered function is even, i.e., is invariant under sign change. Indeed, two values of the argument that differ in sign correspond to the same value of the function. This circumstance explains the lack of uniqueness of the solution of the problem. We will encounter similar effects in the analysis of optimal control problems; see Chapters 5, 11, 14, and 15.

15. Examples of optimal control problems in which both the uniqueness of the solution and the sufficiency of the extremum conditions are absent will be considered in Chapters 5, 6, 11, 14, and 15.

16. We will encounter this property more than once when studying optimal control problems.

17. We will meet the considered function in Chapter 5 when studying an optimal control problem that has three solutions; see Example 5.3.

18. Example 5.1 will consider an optimal control problem for which the optimality conditions also have an infinite set of solutions. However, only two of them are optimal. Other examples with similar properties are also discussed in Chapters 6, 11, 15.

19. Here, we have a set with ***continuum*** cardinality. In Section 1.2.2, we consider an optimal control problem whose solution set is also a continuum set; see Example 6.1. Characteristically, in both cases we are dealing with a function and a functional that are convex but not strictly convex. Qualitatively different examples of optimal control problems with a continuum set of solutions are given in Chapters 11 and 15.

20. This example is considered in [193]

21. For the same purpose, we can consider the sequence $y_k = -k$, $k = 1, 2, ...$ or a sequence whose elements take the values k and $-k$ through time.

22. Similarly, the concept of Tikhonov well-posedness is introduced for problems of minimization of functionals; see Section 1.2.4. In Chapter 2, the concept of Hadamard well-posedness will also be defined.

23. Any extremal problems with non-unique solutions are Tikhonov ill-posed. As a non-converging minimizing sequence for them, one can choose a sequence in which different solutions of the problem alternate. Examples of optimal control problems that are ill-posed in the sense of Tikhonov are considered in Chapters 8, 12, and 15; see also [194].

24. Approximate methods for solving function minimization problems are described in the next chapter.

25. In this regard, the question of what should be understood as an approximate solution of the function minimization problem becomes relevant; see Chapter 2.

26. The proof of Theorem 1.1 begins with the assumption that x is the minimum point of the function under consideration. If this assumption is not realized, then the subsequent reasoning loses its meaning. However, if a solution to the problem existed, then it would certainly satisfy the stationary condition. Thus, the absence of solutions to the stationary condition is a sure sign that the absolute minimum of the function does not exist. Examples of optimal control problems that have no solution are given in Chapters 7, 11, and 15.

27. This circumstance is connected with the absence of extrema in the considered function. In this case, we are dealing with the ***inflection point*** of the function. Curiously, at this point, not only the first, but also the second derivative of the function vanishes. However, one should not think that the vanishing of the first two derivatives of a function at some point is evidence that we are dealing precisely with an inflection point. A similar situation is realized, for example, for the function $f(x) = x^4$ at zero. However, this feature has an absolute minimum there. For a function of one variable, the solution of the stationary condition turns out to be either an extremum point (local or global) or an inflection point. For functions of many variables, and even more so for functionals of general form, there is a much richer variety of forms of critical points; see [11]

28. This situation is especially unpleasant in the practical solution of extremal problems, which, as a rule, is carried out on the basis of certain approximate methods. Indeed, if the extremum condition has no solution, then in the process of numerically solving the problem, we will not get anything. However, in the situation of Example 1.9, by solving the extremum condition approximately, we can determine its solution. Since nothing else can be found, and the structure of the real problem is rather complicated, a false impression may arise that a solution to the minimization problem has been found. In Chapter 11, an optimal control problem is presented that has no solution, while the solution of the corresponding necessary optimality condition for the solution has; see Example 11.6.

29. Naturally, in the case when the function being minimized is not differentiable, we cannot obtain inequality 1.2 and the stationary condition that follows from it. Nevertheless, the task of minimizing this function is quite meaningful, and therefore, the development of effective methods for solving such problems is of undoubted interest.

30. A method for solving such problems is described in the following chapter

31. Naturally, in the case when the stationary point satisfies the given constraints, it can be a solution to the problem. In particular, if the function $f(x) = (x + 1)^2$ is minimized on the interval $[-2, 0]$, then the only stationary point $x = -1$ turns out to be a solution to this problem.

32. For methods for finding the global minimum of functions; see [39], [69], [70], [92][147], [202]

33. Other methods for finding function extrema are considered in [26], [35], [41], [42], [49], [65], [69], [70], [79], [132], [139], [141], [149], [112], [180], [193]

34. The Weierstrass theorem talks about the solvability of the conditional extremum problem, which will actually be considered in the next chapter.

35. The ***closeness*** of a set is one of the most important topological concepts; see [101]. On the number line, closed sets are the entire line, semi-infinite intervals $(-\infty, b]$, $[a, \infty)$, closed intervals $[a, b]$, sets consisting of individual points $\{x\}$, and unions of the indicated sets. We will meet with the closeness of sets in function spaces in Chapter 7 when proving the existence of a solution to optimal control problems.

36. The concept of ***boundedness*** of a numerical set is quite natural. In passing to general problems, the boundedness of a set will be understood as the uniform boundedness of the norms of its elements.

37. In the transition to functional minimization problems, we will no longer be able to use the Bolzano–Weierstrass theorem in proving the existence of a solution. However, there is a generalization of it, called the Banach–Alaoglu theorem, which allows one to achieve the desired results for a certain class of problems to be solved; see [94], [158].

38. Closed sets contain the limits of all convergent sequences of elements of this set. Thus, the sequence $\{x_k\}$, characterized by the equality $x_k = 1/k$ consists of elements of the non-closed interval (0,1) and converges. However, its limit does not belong to this interval.

39. Chapter 7 generalizes Theorem 1.3 to the functional minimization problem on an unbounded subset of some normed vector space. In this case, the coercivity property of the functionals will be used.

40. For example, a function that takes the value -1 for negative values of the argument and 1 for their positive values reaches its minimum, being discontinuous. A quadratic function on the interval (–1,1) has a minimum point, although it is minimized on a non-closed set.

41. The concept of convexity can be naturally generalized to general functionals; see Chapter 5.

42. Chapter 5 will generalize Theorem 1.4 to the problem of minimizing a strictly convex functional on a convex subset of a vector space.

43. In order to obtain sufficient conditions for the uniqueness of the maximum of a function, it is required in the conditions of the theorem to replace the convexity of the function with the concavity, which is determined by changing the inequality sign in Definition 1.5.

44. In fact, the convexity of the set U is used here. The definition of the convexity of a subset of a Euclidean space is given in Section 1.2.6, and the general definition of a convex set is given in Chapter 5.

45. Chapter 6 will consider an optimal control problem with a convex but not strictly convex functional that has an infinite number of solutions. On the other hand, Chapter 9 describes an optimization problem with a convex but not strictly convex functional whose solution is unique.

46. The function depicted in Figure 1.16 can be called locally strictly convex in the sense that it is strictly convex in a neighborhood of the minimum point. Chapters 3 and 5 will consider optimal control problems that have a unique solution even in the absence of local convexity of the quantity being minimized.

47. The concept of strong convexity of a function of one variable naturally extends to functions of many variables and even to functionals. In this case, the value $(x–y)^2$ on the right side is replaced by the squared norm of the difference $\|x–y\|^2$.

48. A generalization of Theorem 1.5 to problems of functional minimization will be given in Chapter 8.

49. In Theorem 1.6, a ***second order extremum condition*** is given, since the second derivative of the function being minimized is used here; see also [26], [42], [70], [75]. Chapter 6 contains a statement that is a generalization of Theorem 1.6. We are talking about the Kelley condition, which characterizes the optimality of a singular control that is a specific solution of the necessary optimality condition in the form of the maximum principle.

50. Let all derivatives of the function at the critical point up to order $n–1$ vanish, and the nth order derivative be non-zero. If the number n is even, and this derivative is positive (respectively, negative), then the minimum (respectively, maximum) of this function is realized at a given point. If the number n is odd, then the extremum is not reached at this point. In particular, Example 1.10 considers the function $f(x) = x^3$ with the only stationary point $x = 0$. Here, the second derivative is zero, but the third derivative is non-zero. Since we have an odd-order derivative, there is no extremum at this point, which is consistent with previous results; see Figure 1.11. On the other hand, the function $f(x) = x^4$ also has a unique stationary point $x = 0$. However, now we already have an even order of the derivative. The corresponding fourth derivative is positive, and the function has a minimum at this point. Note, however, that for the analysis of optimal control problems, which are considered in the subsequent parts of the book and are the main objects of research, work even with second derivatives, as a rule, turns out to be inefficient. However, Chapter 6 will give one second-order optimality condition.

51. The proof of this statement and other properties of the subgradient; see [60], [109], [193].

52. The extension of the class of differentiable functions in the transition to subdifferentiable functions is provided by refusing the uniqueness property. The derivative is always unique, but the subgradient is not.

53. It also follows from Theorem 1.7 that at the minimum point an arbitrary function is subdifferentiable.

54. There are also more general constructions for non-smooth optimization, such as the **Clarke derivative**; see [47], [48]. For non-smooth optimization methods; see also [54], [55], [60], [68], [70], [133], [132], [139]. Chapter 4 will consider one optimal control problem with a non-smooth functional.

55. The concept of a derivative naturally extends to functionals, which predetermines the possibility of generalizing Fermat theorem to a much wider class of problems of finding an extremum. In particular, a necessary condition for the minimum of a functional defined on a topological vector space is that its Gateaux derivative vanishes at the minimum point. These statements for one specific example will be established in Chapter 4. In problems of minimizing functionals of general form, the stationary condition will no longer be an algebraic equation. In particular, in the calculus of variations, the ***Lagrange problem*** is considered, which involves the minimization of an integral functional depending on the unknown function and its derivative. The corresponding stationary condition is the ***Euler equation***, which is a second-order

ordinary differential equation. For the ***Dirichlet problem***, which consists of minimizing some special integral depending on a function of many variables and its partial derivatives, the stationary condition is the ***Poisson equation***, which belongs to the class of partial differential equations. On the differentiation of functionals in normed spaces; see, for example, [15], [100], [109], and in topological vector spaces; see [16], [71]. On the connection between the theory of extremum and the general theory of differentiation; see [171].

56. The convexity property naturally extends to subsets of an arbitrary vector space.

57. This statement also extends to the functional minimization problem; see [60].

CHAPTER 2

Additions

In the previous chapter, we considered the problem of minimizing a differentiable function of one variable, for which Fermat theorem was applied. Some additional results in this direction are given below. In particular, the stationary condition obtained using Fermat theorem is not applicable to the problem of minimizing a function on an interval. In this case, one can use the variational inequality. Further, in practice, problems often arise that include some parameters known with some error. We would like to find out whether a small error in determining these parameters will really entail a small error in determining the solution of the problem and how much the properties of the minimization problem change depending on the parameter. Finally, in the examples given earlier, fairly simple functions were considered, as a result of which the solution of the corresponding stationary condition did not cause any special difficulties. However, for sufficiently difficult functions, it is not possible to explicitly find a solution to the stationary condition, as a result of which different iterative methods have to be applied. In Appendix, the problem of sufficiency of variational inequality is studied, the function minimization problem with an equality-type constraint is solved, the problem of well-posedness of extremal problems is posed, general methods for the approximate solution of function minimization problem are described, and the types of approximate solutions of such problems are determined.

2.1 LECTURE

In the process of analyzing function minimization problems in the previous chapter, we encountered some problems that remained unresolved. In particular, in Example 1.12, we considered the problem of minimizing a function on an interval, for which Fermat theorem is inapplicable. To study such problems, one uses the variational inequality (Section 2.1.1). Further, in Examples 1.9 and 1.10, the problem of minimizing the function $f(x) = x^3 - \mu x$ with the values of the parameter μ equal to zero and three, respectively, was considered. Despite the obvious similarity of these problems, they have qualitatively different properties: in the second case, the function has two local extrema, while in the first one there are none at all. Below, we study the problem of minimizing a function depending on a parameter (Section 2.1.2). Finally, in all the examples considered earlier, we easily carried out the analysis of the stationary condition. However, for quite difficult functions, the solution of the problem can only be established approximately (Section 2.1.3). Note that all three

DOI: 10.1201/9781003398585-2

questions discussed are independent of each other, so that the following parts of the lecture can be considered in any order.

2.1.1 Variational inequality

We considered the problem of minimizing the function $f(x) = (x+1)^2$ on the segment $[0, 1]$; see Example 1.12. It is a partial case of the following conditional function minimization problem[1].

Problem 2.1 *Find a point of minimum for a function $f = f(x)$ on an interval $[a, b]$.*

To solve this problem, it is not possible to use the Fermat theorem. However, the following statement is true.

Theorem 2.1 *If a differentiable function $f = f(x)$ has a minimum at the point x on the interval $[a, b]$, then it satisfies the inequality*

$$f'(x)(y-x) \geq 0 \ \ \forall y \in [a, b]. \tag{2.1}$$

Proof. If x is the minimum point of the function f on the given segment, then $f(x) \leq f(z)$ for all $z \in [a, b]$. Obviously, for any point y from $[a, b]$, the point $z = x + \sigma(y-x)$ will belong to the same segment for any number σ from the unit segment, i.e., we are dealing with a convex set. Substituting this value into the previous inequality, we get $f(x + \sigma(y-x)) - f(x) \geq 0$. Dividing this inequality by σ and passing to the limit as $\sigma \to 0$, we get formula (2.1). ☐

Definition 2.1 *The formula* (2.1) *is called the* ***variational inequality***[2]

Obviously, any stationary point satisfies the variational inequality, which in this case is fulfilled in the form of an equality. It is easy to see that in the absence of restrictions, the solution of the variational inequality necessarily satisfies the stationary condition[3]. Thus, the variational inequality is a natural generalization of the stationary condition to the case of problems of minimizing a function on an interval.

Let us consider some fairly simple examples of the application of variational inequalities for the analysis of conditional function minimization problems.

Example 2.1 *Find a point of minimum of the function $f(x) = (x + 1)^2$ on the interval $[a, b]$.*

From the inequality (2.1), it follows

$$(x + 1)(y-x) \geq 0 \ \ \forall y \in [a, b]. \tag{2.2}$$

The first multiplier of the left-hand side here has the concrete value $x + 1$ determined by unknown solution of the problem. However, the second multiplier includes a variable number y.

Suppose the sum $x + 1$ is positive, i.e., $x > -1$. After division inequality (2.2) by $x + 1$, determine $y–x \geq 0$, so $x \leq y$ for all y from the interval $[a, b]$. However, the point x is a solution of the given problem; hence it belongs to this interval. This is possible only if x equals to this least possible value, i.e., a. Thus, if $x > -1$, then $x = a$. Finally, we conclude that we have $x = a$ for $a > -1$.

Suppose now that $x < -1$. Then, after dividing inequality (2.2) by $x + 1$, we set $y–x \leq 0$, and hence $x \leq y$ for all values of y from the segment $[a, b]$. Considering the inclusion $x \in [a, b]$, we conclude that x must take the largest of all possible values, i.e., b. Therefore, if $x < -1$, then $x = b$. Thus, we obtain that for $b < -1$ we have $x = b$.

Finally, suppose the equality $x = -1$. Then the condition (2.2) is true. However, this value can be the solution of the given problem only if it is admissible. Thus, the solution of the problem is -1 if $a \leq -1 \leq b$.

Combining all the above results, we conclude that $x = a$ for $a > -1$, $x = -1$ for $a \leq -1 \leq b$, and $x = b$ for $b < -1$; see Figure 2.1. According to the Weierstrass theorem (see Theorem 1.2), the considered problem has a solution that is unique by Theorem 1.3. Since the necessary minimum condition (2.2) has a unique solution, we conclude that it is the solution of this problem. In particular, in Example 1.11 the considered function was minimized in the segment $[0, 1]$. Considering that in this case the inequality $a > -1$ is valid, we conclude that the solution to the problem is the minimum admissible value, i.e., 0.

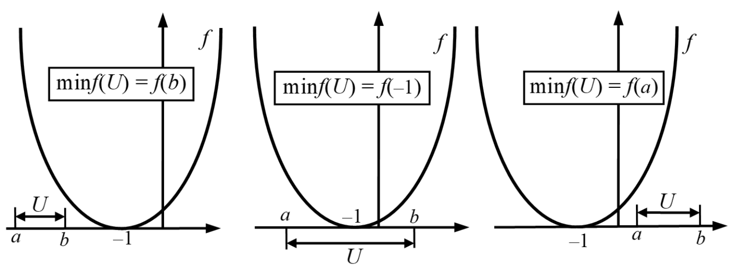

Figure 2.1 Minimum of the function $f(x) = (x + 1)^2$ on the interval $[a, b]$.

In the problem of an unconditional extremum, the transition from the search for a minimum to the search for a maximum can turn a solvable problem into an unsolvable one and vice versa. However, according to the Weierstrass theorem considered in the previous chapter, both a minimum and a maximum of a continuous function on an interval exist. Let us apply Theorem 2.1 to find the maximum of the function from the last example.

Example 2.2 *Find a point of maximum of the function* $f(x) = (x + 1)^2$ *on the interval* $[a, b]$.

Obviously, it is equivalent to the minimization problem on the same interval of the function $f(x) = -(x + 1)^2$. In this case, the variational inequality (2.1) takes the form

$$(x+1)(y-x) \le 0 \ \ \forall y \in [a, b]. \tag{2.3}$$

To solve it, we use the method described above.

Let us assume that $x > -1$. Then, after dividing inequality (2.3) by $x+1$, we get $y-x \le 0$, and hence $x \ge y$ for all values of y from the interval $[a, b]$. This is possible only when x takes the largest of all possible values, i.e., b. So, if $x > -1$, then $x = b$. Thus, we conclude that for $b > -1$, we have $x = b$. If $x < -1$, then after dividing inequality (2.3) by $x+1$, we get $y-x \ge 0$, and hence $x \le y$ for all values of y from the interval $[a, b]$. Thus, x must take the smallest of all possible values, i.e., a. So, if $x < -1$, then $x = a$, i.e., for $a < -1$ we have $x = a$. Finally, if the first multiplier on the left side of inequality (2.3) vanishes, then $x = -1$, but this value can only be a solution to the problem for $a \le -1 \le b$.

Let us summarize. For $b < -1$, among the three situations described above, only the inequality $a < -1$ is realized, and hence $x = a$. For $a > -1$, only the condition $b > -1$ is realized, and therefore, $x = b$. Finally, for $a \le -1 \le b$, all three situations are possible[4], which means that all three values $x = -1$, $x = a$, and $x = b$ satisfy formula (2.3).

In fact, the maximum of the considered function on this segment is reached at point a for $b < -1$ and at point b for $a > -1$. However, for $a \le -1 \le b$, the solution to the problem is one of the two values a or b, which corresponds to the largest value of the given function. Thus, for the considered variational inequality, it turns out to be a necessary but not sufficient condition for an extremum[5].

2.1.2 Dependence of the solution on parameters

In problems of finding an extremum that arises in real applications, the value to be minimized, as a rule, depends on parameters. They are often known with some error. In this regard, it is of interest to evaluate the degree of influence of parameters on the solution of the corresponding problem. We will not present any general results here, confining ourselves to individual examples.

Example 2.3 *Find a point of maximum of the function* $f(x) = (x-\mu)^2$, *where* μ *is a numerical parameter.*

Using Fermat's theorem, we find the stationary point $x = \mu$, which is naturally the unique solution to this problem. In this case, the fact that the solution of the problem continuously depends on the problem parameter is of particular importance.

Definition 2.2 *A problem is called* ***Hadamard well-posed***[6] *if it has a unique solution that continuously depends on the problem parameters*[7].

Thus, the minimization problem considered in Example 2.3 is Hadamard well-posed. However, another situation is possible.

Example 2.4 *Find a point of maximum of the function* $f(x) = 3x^4-16x^3 + 6\mu x^2$ *with parameter* μ *from the open interval*[8] $(0, 4)$.

The corresponding stationary condition is $x^3–4x^2 + \mu x = 0$. The resulting cubic equation has three solutions $x_1 = 0$, $x_2 = 2-\sqrt{4-\mu}$ and $x_3 = 2+\sqrt{4-\mu}$. It is easy to verify that the first and third values correspond to the minimum, and the second to the maximum. Then the solution of the problem will be that of the values x_1 or x_3, on which the function f takes a smaller value. Obviously, $f(0) = 0$. Thus, the solution to the problem is point 0 if $f(x_3)$ is positive, and x_3 if it is negative[9].

Let us determine the value $f(x_3)$ for the parameter μ, which takes the boundary values of the selected interval. For $\mu = 0$ we have $x_3 = 4$, whence it follows that $f(x_3) = –256$. On the other hand, for $\mu = 4$, we have $x_3 = 2$, which implies that $f(x_3) = 16$. Obviously, the dependence of $f(x_3)$ on the parameter μ is continuous on the interval $[0, 4]$. At the same time, at the ends of this segment, it takes on values of different signs. Then there is such a number $\mu^* \in (0, 4)$ that the corresponding value of $f(x_3)$ vanishes[10]. Obviously, for an arbitrarily small number $\varepsilon > 0$ with $\mu = \mu^*–\varepsilon$ the number $f(x_3)$ is negative, which means that the solution to the problem is the point x_3. However, when $\mu = \mu^* + \varepsilon$ the number $f(x_3)$ is positive, and hence the function f has its minimum at zero. Thus, changing the parameter μ by an arbitrarily small value 2ε leads to a large change in the solution of the problem; see Figure 2.2. Therefore, the problem under consideration is not Hadamard well-posed[11].

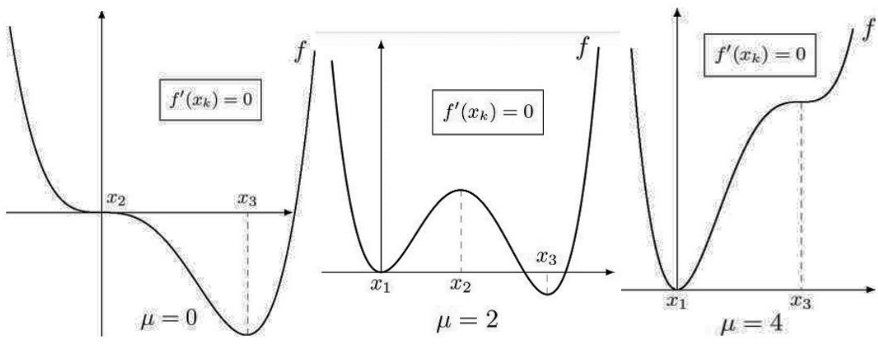

Figure 2.2 Function $f(x) = 3x^4–16x^3 + 6\mu x^2$.

In the last example, a situation was considered when, with a small change in the parameter of the problem, a significant change in its solution occurred. However, the general properties of the considered function with the parameter values $\mu = \mu^*–\varepsilon$ and $\mu = \mu^* + \varepsilon$ remain unchanged. In both cases, the function has a local minimum, a local maximum, and an absolute minimum. However, there are situations when a small change in the problem parameter not only entails a large change in the solution of the problem, but also leads to a qualitative change in the properties of the minimized function.

Example 2.5 *Find a point of maximum of the function* $f(x) = x^3–\mu x$ *with parameter* μ.

The case $\mu = 0$ corresponds to Example 1.9, and when $\mu = 3$ we get Example 1.10. The corresponding stationary condition $3x^2 = \mu$ in the first case has a unique solution $x = 0$, which is the inflection point of the considered function; see Figure 1.11, and in the second case, there are two stationary points $x = -1$ and $x = 1$, the first of which corresponds to a local maximum, and the second to the local minimum. Moreover, with negative values of the parameter μ, for example, with $\mu = -1$, the stationary condition has no solution at all[12]. Thus, the behavior of the function for parameter values from an arbitrarily small neighborhood of zero can be different in quality: for positive values of the parameter μ, there are two local extrema, for zero, there is only an inflection point, and for negative values, there are no stationary points at all; see Figure 2.3.

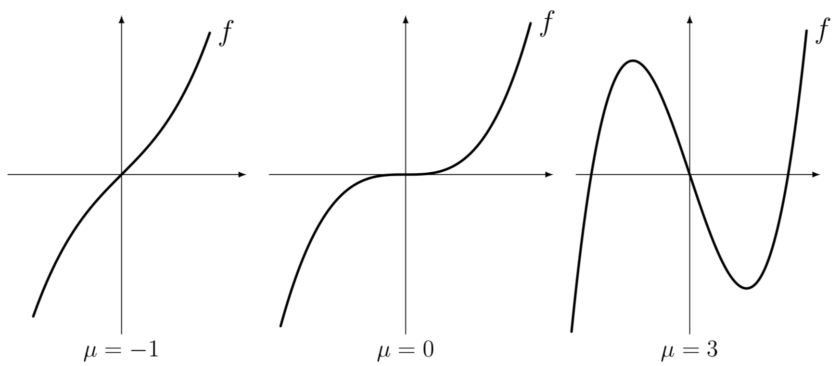

Figure 2.3 Function $f(x) = x^3 - \mu x$.

Definition 2.3 *Changing the number of solutions to a certain problem depending on its parameters is called a* ***bifurcation***[13]*, and the value of the parameter, when passing through which this change occurs, is called a* ***bifurcation point****.*

Thus, in this example, we have a bifurcation of solutions to the stationary condition for the considered function. In the final part of the following section, the behavior of the algorithm for the approximate solution of the problem of minimizing a function depending on a parameter will be considered[14].

2.1.3 Approximate solving of the stationary condition

Earlier, we reduced the function minimization problem using Fermat Theorem to the stationary condition, which is an algebraic equation. Since all the considered functions were quite simple, the practical solution of this equation did not cause difficulties. However, in the general case, finding stationary points is a serious problem, which requires the use of different iterative algorithms. In this case, the question of convergence naturally arises. Below, we describe the corresponding algorithms for sufficiently simple functions, when it is quite easy to establish whether these algorithms converge, and if they converge, then where.

Example 2.6 *Find a point of maximum of the function* $f(x) = x^2/2 + e^{-x}$.

Find the derivative $f'(x) = x - e^{-x}$. The corresponding stationary condition $x = e^{-x}$ is the transcendental equation[15]. It is very difficult to find its solution analytically. However, one can try to solve it approximately using an iterative algorithm. Choose an initial iteration x_0. The simplest iterative process is characterized by the equality $x_{k+1} = e^{x_k}$, where k is the iteration number and x_k is the corresponding k-th approximation.

Definition 2.4 *The described algorithm corresponds to the* ***method of successive approximations***[16] *or the* ***method of simple iteration****.*

It is known [17] that the algorithm $x_{k+1} = F(x_k)$ converges to a unique solution of the algebraic equation $x = F(x)$ for any initial approximation if the following inequality holds[18] $|F'(x_k)| < 1$ for all sufficiently large values of k. In this case, we have $F(x) = e^{-x}$, which means $F'(x) = -e^{-x}$. Then the last inequality reduces to the condition $e^{-x} < 1$, which is realized for all positive values of x. Obviously, for any initial approximation x_0 the value $x_1 = e^{-x_0}$ is positive. All subsequent values of x_k are also positive. Thus, the given convergence condition is certainly satisfied, at least starting from the number $k = 1$. Thus, the iterative process converges for any initial approximation to a unique solution of the stationary condition, which is the only minimum point of this function[19]. The minimized function and the behavior of the algorithm are shown in Figure 2.4.

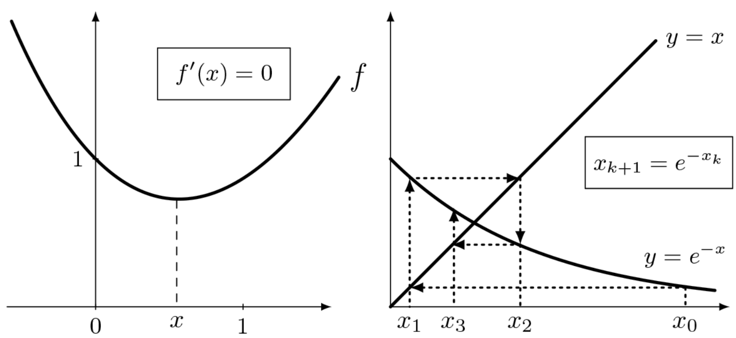

Figure 2.4 Function from Example 2.6 and analysis of the stationary condition.

In the considered example, for any initial approximation, the algorithm converges to the solution of the problem. However, such a situation is not always observed.

Example 2.7 *Find a point of maximum of the function* $f(x) = (x–1)^4/4–x^2/2$.

Determine the derivative $f'(x) = (x–1)^3–x$. Now, the stationary condition is $x = (x–1)^3$. This equation is solved iteratively by the equality $x_{k+1} = (x_k–1)^3$. It is easy to see that, although the stationary condition has a unique solution, the iterative process diverges for any initial approximation[20]; see Figure 2.5.

In addition to the unconditional convergence and unconditional divergence of the algorithm, an intermediate situation is also possible.

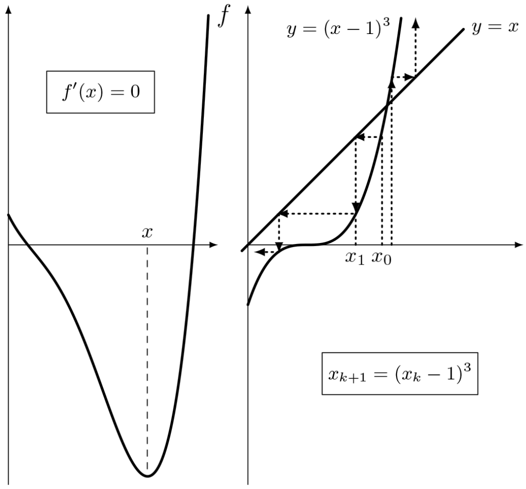

Figure 2.5 Function from Example 2.7 and analysis of the stationary condition.

Example 2.8 *Find a point of maximum of the function f such that $f(x) = x^3/3$ for $x > 0$ and $f(x) = x^2/2 + x/2 + 2/3(1/4–x)^{3/2} + 1/12$ if $x < 0$.*

This function is continuously differentiable, besides $f'(x) = x^2$ if $x > 0$ and $f'(x) = x + 1/2–(1/4–x)^{1/2}$ if $x < 0$. Write the stationary condition as an equality $x = F(x)$, where $F(x) = x^2 + x$ if $x > 0$ and $F(x) = –1/2 + (1/4–x)^{1/2}$ if $x < 0$. The solution of the obtained algebraic equation will again be sought using an iterative process characterized by the equality $x_{k+1} = F(x_k)$. Note that $F'(x) = 2x + 1 > 1$ for $x > 0$ and $0 < F'(x) = 1/2(1/4–x)^{-1/2} < 1$ for $x < 0$. It is easy to see that for negative initial approximations, the algorithm converges to the unique stationary point $x = 0$, which is the solution to this problem, while for positive values of the initial approximation, the process diverges; see Figure 2.6.

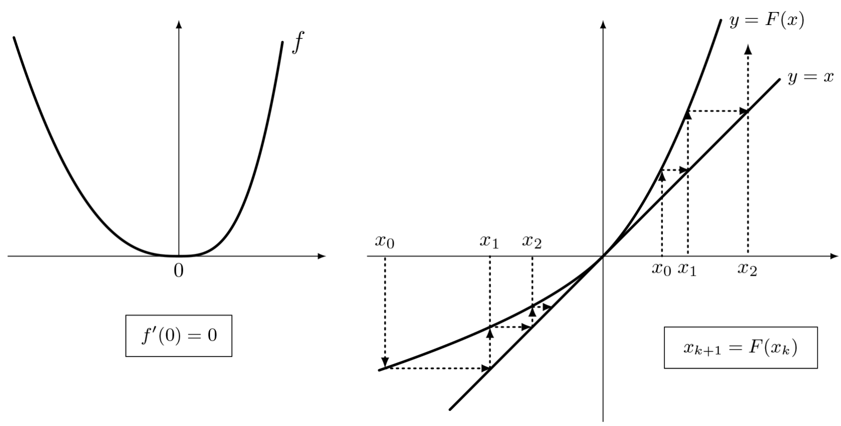

Figure 2.6 Function from Example 2.8 and analysis of the stationary condition.

Thus, a situation is possible when the algorithm converges for some initial approximations and diverges for other initial approximations. Thus, in the case of divergence of the iterative process, the desired result can sometimes be obtained by changing the initial approximation.

Example 2.9 *Find a point of maximum of the function* $f(x) = x^4/4 – x^2/2 – x$.

The function derivative is $f'(x) = x^3 – x – 1$. The corresponding stationary condition $x^3 – x – 1 = 0$ is solved by the algorithm $x_{k+1} = (x_k)^3 – 1$. This diverges for any initial iteration; see Figure 2.7. Therefore, by changing the initial approximation, we do not find a solution to the problem. In this regard, we try to use a different algorithm. Let us define a new variable $y = x^3$. Then $x = y^{1/3}$, and the stationary condition takes the form $y – y^{1/3} – 1 = 0$. To solve the resulting equation, we use the following algorithm $y_{k+1} = (y_k)^{1/3+1}$. Obviously, in this case, the iterative process converges to a single value of y; see Figure 2.7. The corresponding point $x = y^{1/3}$ minimizes the given function. Thus, if the applied algorithm does not converge, and changing the initial approximation does not lead to the desired goal, then to solve the problem, you can try to use another algorithm.

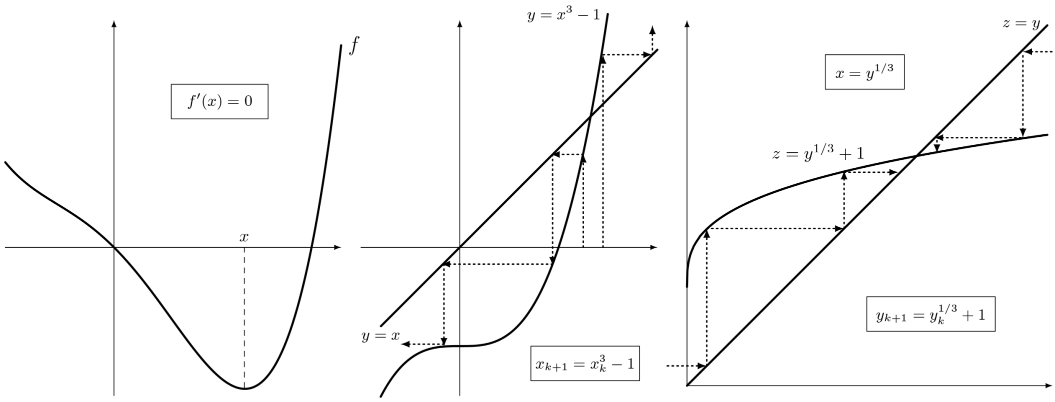

Figure 2.7 Function from Example 2.9 and analysis of the stationary condition.

So far, we have considered iterative methods for solving the stationary condition for the case when the latter has a unique solution. However, as we already know, if the solution of the minimization problem is non-unique or if the extremum condition is insufficient, the corresponding stationary condition has a non-unique solution. Consider the operation of the iterative process in this situation.

Example 2.10 *Find a point of maximum of the function* $f(x) = x^2/2 - 3/4x^{4/3}$.

We have the stationary condition $f'(x) = x – x^{1/3} = 0$. This is solved by the algorithm $x_{k+1} = (x_k)^{1/3}$. Obviously, the process converges to one for positive initial approximations and to minus one for negative approximations[21]; see Figure 2.8. Both limit values are the minimum points of the considered function[22].

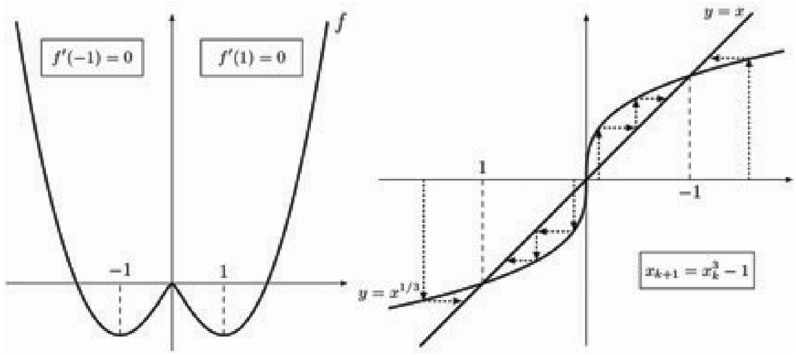

Figure 2.8 Function from Example 2.10 and analysis of the stationary condition.

We conclude that if the solution of the stationary condition is not unique, then different initial iterations can give different solutions. Therefore, if the minimization problem can have many solutions or the stationary condition is not sufficient, then the calculation must be carried out many times with different initial iterations.

Consider now the iterative method of minimizing a function that depends on parameter.

Example 2.11 *Find a point of maximum of the function* $f(x) = 2\mu x^3 – 3(\mu – 1)x^2$ *with positive parameter* μ*.*

The stationary condition here is $\mu x^2 – \mu x + x = 0$. We try to find its solution by the equality $x_{k+1} = \mu x_k(1 – x_k)$. This is called the ***logistic mapping*** or ***discrete Verhulst equation***[23]. Consider the property of the algorithm for different values μ. Obviously, for $\mu < 1$, the algorithm converges to the value $x = 0$ that is the point of local minimum of the function f; see Figure 2.9.

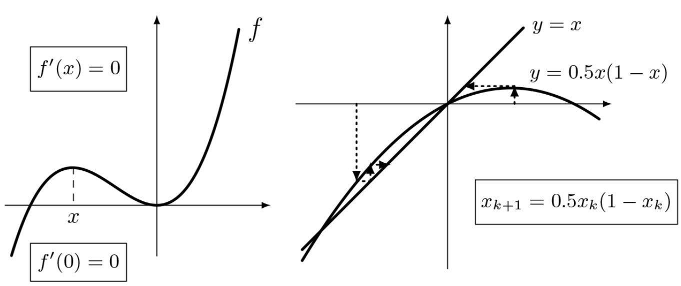

Figure 2.9 The function of Example 2.11 for $\mu = 1/2$ and the algorithm convergence.

For $\mu > 1$, we have an algorithm $x_{k+1} = F(x_k)$ that converges if the inequality $|F'(x)| < 1$ is true for a stationary point x. For this case, $F(x) = \mu x(1 – x)$; so, $F'(x) = \mu(1 – 2x)$. The point $x = 0$ does not satisfy that inequality. However, there exists the second stationary point $x = (\mu – 1)/\mu$. The condition of the algorithm convergence for it is true for $\mu < 3$. Thus, for $1 < \mu < 3$, the algorithm converges to the second stationary point that is the local minimum of the function f, besides the algorithm diverges for the larger values of μ; see Figure 2.10.

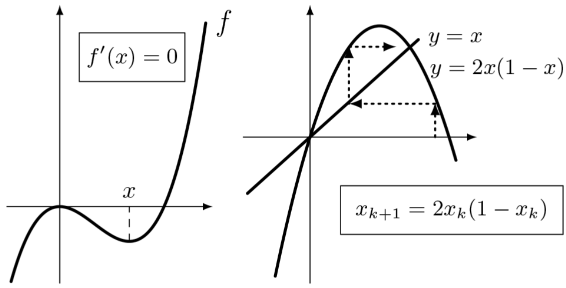

Figure 2.10 The function of Example 2.11 for $\mu = 2$ and the algorithm convergence.

Consider the algorithm for $\mu > 3$. Substituting in the right side of the existing equality $x = F(x)$ instead of x its equal value $F(x)$, we get $x = F(F(x))$, that is, $x = \mu 2x(1–x)[1–\mu x(1–x)]$. This fourth-order algebraic equation has four solutions[24]. It can be shown that in the case of $3 < \mu < 1 + \sqrt{6}$ the algorithm will endlessly oscillate between two of these solutions. With an increase in the parameter μ, the algorithm will fluctuate endlessly between four values, which are solutions to the equation $x = F(F(F(x)))$. Further increase of this parameter causes the algorithm to oscillate around 8, 16, etc. values[25]. Finally, for values of μ equal to approximately 3.57, the behavior of the algorithm becomes chaotic[26].

RESULTS

Here is a list of questions following the results of the lecture, the main conclusions on this chapter, as well as the problems that arise in this case, partially solved in Appendix, partially in the Notes.

Questions

It is required to answer questions related to the previously given lecture subject.

1. Why is the stationary condition a necessary condition for both the minimum and the maximum of a function, while the variational condition turns out to be only a necessary condition for the minimum, but not the maximum?

2. What is the form of the variational inequality for the function maximum problem?

3. In what sense can the variational inequality be considered as a generalization of the stationary condition?

4. Can the stationary point be a solution to the variational inequality?
5. Can the stationary point not be a solution to the variational inequality?
6. Can the solution of the variational inequality be a stationary point?
7. Can the solution of the variational inequality not be a stationary point?
8. Is it possible to use the variational inequality to minimize a function on semi-infinite intervals like $(-\infty, b]$ or $[a, \infty)$?
9. Why is in generally the variational inequality a necessary but, not a sufficient condition for optimality?
10. What can be said about the variational inequality in the case of non-smoothness of the function being minimized?
11. What can be said about the variational inequality in the problem of minimizing a function on a set $[a, b] \cup [c, d]$, where $b < c$?
12. Why can the absence of Hadamard ill-posedness has the serious negative consequence in solving extremal problems that arise in applications even when the problem itself has a unique solution?
13. Why can the absence of Hadamard correctness have serious negative consequences when solving extremal problems, even in the case when all the parameters included in the formulation of the problem are specified exactly?
14. What causes the Hadamard ill-posedness of the problem considered in Example 2.5.
15. What effects can be observed when changing the parameters included in the definition of the minimized function?
16. How, in principle, can an iterative process behave for an approximate solution of the stationary condition?
17. What are the properties of the iterative algorithm for a stationary condition if the function minimization problem has no solution?
18. What are the properties of the iterative algorithm for a stationary condition if it has more than one solution?
19. Why can the iterative process for solving the stationary condition converge to different values with a change in the initial approximation?
20. What should be done if different results are obtained for different initial approximations when iteratively solving optimality conditions?
21. How can one establish the non-uniqueness of the solutions to the stationary condition using the iterative method?

22. What actions are possible with the divergence of the iterative method to solve the stationary condition?

Conclusions

Based on the obtained results, we come to the following conclusions.

- Variational inequality is the necessary condition of minimum for differentiable functions on the interval.
- Variational inequality can also be established for function maximization problems.
- Any stationary point satisfies the corresponding variational inequality.
- Any solution of variational inequality is a stationary point if we minimize the function on the numerical line.
- Sometimes, the variational inequality admits an analytical solution. Function minimization problems may depend on parameters.
- A situation is possible when a small measurement of a problem parameter entails a small change in its solution, which corresponds to Hadamard well-posedness.
- In the absence of Hadamard well-posedness, a small change in a problem parameter can lead to a large change in its solution and even significantly change the general properties of the function.
- In the absence of Hadamard well-posedness, any computational errors (rounding, calculation of special functions, use of series, integrals, derivatives, etc.) can cause significant deviations in determining the solution to the problem.
- For the practical solving stationary conditions, in generally, iterative methods are used.
- The corresponding iterative algorithms may or may not converge.
- If the algorithm diverges, the desired result can be achieved by changing the initial approximation.
- If changing the initial approximation does not lead to the desired goal, then another algorithm should be applied to solve the stationary condition.
- If the stationary condition has no solution, then the iterative process for solving it cannot converge.
- If the solution of the stationary condition is not unique, then for different initial approximations the algorithm may converge to different limits.

Problems

Based on the results obtained above, we arrive at the following problems.

1. **Generalization of variational inequalities**. Theorem 2.1 gives a necessary condition for the minimum of differentiable functions on an interval. This result easily extends to smooth functions of many variables and even to general differentiable functionals defined on a convex set[27].

2. **Sufficiency of the minimum condition in the form of a variational inequality**. In Example 2.1, the variational inequality turns out to be a necessary and sufficient condition for the minimum of this function, while in Example 2.2, sufficiency is not realized. We would like to understand under what conditions sufficiency is realized. This result is given in Appendix.

3. **Practical solving of variational inequalities**. In Examples 2.1 and 2.2, fairly simple functions were considered. As a result, the practical solving of variational inequalities did not cause difficulties. However, in the general case, this is not an easy problem. One result in this direction is given in Appendix.

4. **Minimization of a function with a constraint in the form of equality**. Section 2.1.1 dealt with the problem of minimizing a function on an interval, i.e., conditional extremum problem with an additional constraint in the form of an inequality: the value of the desired value changes within the specified limits. However, in practice, problems of finding an extremum often arise when additional constraints are given in the form of equalities. Appendix describes two methods for studying such problems.

5. **Hadamard well-posedness condition**. In Example 2.3, the function minimization problem was Hadamard will-posed, while in Example 2.4, the problem is ill-posed. It would be desirable to establish under what conditions the function minimization problem turns out to be well-posed. This result is explored in Appendix.

6. **Changing the properties of the problem when changing its parameters**. In Example 2.5, not only was there no continuous dependence of the solution to the function minimization problem, but there was a radical change in the properties of the considered function. This effect is related to the ***singularity theory***[28].

7. **General method for the practical solving of function minimization problems**. In Examples 2.6–2.10, for finding an approximate solution of the stationary condition, the method of successive approximations was used. Writing the corresponding algorithm did not cause any particular difficulties due to the sufficient simplicity of the considered functions. Appendix describes a general method for solving the function minimization problem, which can be easily extended to functions of many variables and even to functionals. General method for the practical solving of function minimization problems.

8. **Convergence of iterative methods for function minimization**. In Example 2.6 the method of successive approximations for solving the function minimization problem converges, in Example 2.7 it diverges, and in Example 2.8 convergence is observed only for a class of initial approximations. We would like to know under what conditions the approximate method for solving these problems converges. About convergence of approximate methods of extremum theory see Notes[29].

9. **Concept of an approximate solution of extremum problem**. In practice, the problem of finding an extremum, as a rule, is solved approximately. However, the very concept of an approximate solution of the problem needs to be clarified. Different types of approximate solutions are defined in Appendix.

2.2 APPENDIX

Some additional results related to the problems raised earlier are considered below. In particular, a necessary condition for the minimum of a function on an interval was obtained in the form of a variational inequality. Section 2.2.1 will establish a statement about the sufficiency of this condition. Sections 2.2.2 and 2.2.3 describe two methods for finding the conditional minimum of a function when there is an equality constraint. The lecture introduced the notion of Hadamard well-posedness of the function minimization problem. Section 2.2.4 provides conditions that guarantee this property. In addition, for a number of examples, approximate methods for solving the corresponding stationary conditions were described. Section 2.2.5 provides a general method for the practical solving of the considered problems. Different types of approximate solutions to minimization problems are described in Section 2.2.6[30].

2.2.1 Sufficiency of the minimum condition in the form of variational inequalities

Theorem 2.1 gives a necessary condition for the minimum of a differentiable function on an interval. As can be seen from Example 2.1, the resulting variational inequality may be a sufficient minimum condition, although Example 2.2 shows that this situation is not always realized. Below is a statement about the sufficiency of the function minimum condition in the form of a variational inequality[31].

Theorem 2.2 *In order for a convex differentiable function f to reach its minimum on the interval $[a, b]$ at a point x, it is necessary and sufficient that it satisfies the variational inequality* (2.1).

Proof. The necessity was proved in Theorem 2.1. Let us show that under the variational inequality

$$f'(x)(y-x) \geq 0 \ \ \forall y \in [a, b]$$

the function f has the minimum at the point x. Using the function convexity, we get

$$f\big((1-\alpha)x + \alpha y\big) \leq (1-\alpha)f(x) + \alpha f(y) \ \ \forall \alpha \in (0, 1)$$

for all $y \in [a, b]$. Then we obtain

$$f(x + \alpha(y - x)) - f(x) \leq \alpha[f(y) - f(x)].$$

Dividing this inequality by α and passing to the limit as $\alpha \to 0$, determine

$$f(y) - f(x) \geq f'(x)(y - x).$$

Using the variational inequality, we establish that $f(y) \geq f(x)$, whence, due to the arbitrariness of y, it follows that x is the minimum point of this function on the considered segment. □

Example 2.1 dealt with the minimization problem of the function $f(x) = x^2$, that is convex. The corresponding extremum condition turned out to be necessary and sufficient by Theorem 2.2. However, in Example 2.2, the problem was to maximize the same function, which is equivalent to minimizing the non-convex function $f(x) = -x^2$. This case is not covered by Theorem 2.2, and the corresponding variational inequality was not a sufficient condition for the minimum of the indicated function[32].

2.2.2 Lagrange multiplier method

The Lecture considered Problem 2.1 on the conditional extremum of a function. At the same time, an additional constraint implied that the desired value must belong to a certain interval. However, often the additional condition is given in the form of equality. We consider the following problem of minimizing a function with an equality-type constraint[33].

Problem 2.2 *Find a point of minimum of the function $f_1 = f_1(x_1, x_2)$ on the set of number pairs $x = (x_1, x_2)$ that satisfy the equality $f_2(x) = 0$.*

For solving this problem, it can find the dependence of x_2 on x_1 from the equality $f_2(x_1, x_2) = 0$, i.e., define a function $x_2 = g(x_1)$ such that $f_2(x_1, g(x_1)) = 0$. After that, the function of one variable $h = h(x_1) = f_1(x_1, g(x_1))$ is determined, which is minimized on the entire number line by known methods. However, for difficult enough functions f_1 and f_2, this approach turns out to be inefficient.

To solve the problem, we use the ***Lagrange multiplier method***[34]. Determine the function

$$L = L(x, \lambda) = L(x_1, x_2, \lambda_1, \lambda_2) = \lambda_1 f_1(x_1, x_2) + \lambda_2 f_2(x_1, x_2).$$

Definition 2.5 *The function L is called the **Lagrange function**, and the numbers λ_1 and λ_2 are the **Lagrange multipliers**.*

To derive the extremum condition, we use the following assertion that is called the ***implicit function theorem***[35].

Theorem 2.3 *Let $F = F(x,y) = (F_1(x_1,x_2,y), F_2(x_1,x_2,y))$ be a continuously differentiable second order vector function of three variables such that $F(x^*,y^*) = 0$, besides the matrix $(\partial F_i(x^*,y^*)/\partial x_j)$, $i,j = 1,2$ is non-degenerate*[36]. *Then for any y close enough to y^* there exists a second order vector $x(y)$ such that $F(x(y),y) = 0$.*

The necessary extremum condition for Problem 2.2 gives the following statement[37].

Theorem 2.4 *If the point $x^* = (x_1^*, x_2^*)$ is a solution of Problem 2.2, then there exists a non-zero vector $\lambda^* = (\lambda_1^*, \lambda_2^*)$ such that*

$$\frac{\partial}{\partial x_j} L(x^*,\lambda^*) = 0, \quad j = 1,2.$$

Proof. The last equalities are as follows

$$\lambda_1 \frac{\partial f_1(x^*)}{\partial x_j} + \lambda_2 \frac{\partial f_2(x^*)}{\partial x_j} = 0, \quad j = 1,2. \tag{2.4}$$

This can be interpreted as a system of linear algebraic equations with respect to λ_1 and λ_2. Assume that the assertions of theorem are not true, i.e., this system has only zero solution. This is only possible if the matrix $\big(\partial f_i(x^*)/\partial x_j\big)$ is not degenerate.

Determine the second order vectors $F(x,y) = (F_1(x,y), F_2(x,y))$ with three arguments by the equalities

$$F_1(x,y) = f_1(x) - f_1(x^*) + y, \ F_2(x,y) = f_2(x).$$

Consider the system of non-linear algebraic equations

$$F(x,y) = 0 \tag{2.5}$$

with respect to the vector x with parameter y. Obviously, $F(x^*,0) = 0$.

Find the derivatives

$$\frac{\partial F(x^*)}{\partial x_j} = \left(\frac{\partial F_1(x^*)}{\partial x_j}, \frac{\partial F_2(x^*)}{\partial x_j}\right) = \left(\frac{\partial f_1(x^*)}{\partial x_j}, \frac{\partial f_2(x^*)}{\partial x_j}\right), \ j = 1,2.$$

Therefore, the matrix $\big(\partial F_i(x^*,0)/\partial x_j\big)$ is equal to $\big(\partial f_i(x^*)/\partial x_j\big)$, so, it is not also degenerate. Using the implicit function theorem, determine that the system (2.5) has a solution $x = x(y)$ for all small enough positive value y, i.e., $F(x(y)) = 0$. By the definition of F, we get

$$f_1(x(y)) = f_1(x^*) - y, \ f_2(x(y)) = 0.$$

Thus, there exists a vector $x(y)$ such that $f_2(x(y)) = 0$, and the following inequality holds

$$f_1(x(y)) = f_1(x^*) - y < f_1(x^*),$$

by positivity of y. We conclude that x^* cannot be a solution of Problem 2.2, because another point was found that satisfies the given constraints, at which the value of the minimized function turned out to be less. Thus, the assumption that relations (2.4) can be satisfied only for the zero vector λ^* is not true. Consequently, these equalities are valid for some non-zero vector of the second order. □

Theorem 2.4 gives an algorithm for solving Problem 2.2. In each specific case, the Lagrange function must first be determined. Then its partial derivatives with respect to the functions x_1 and x_2 are equated to zero. This gives two conditions (2.4) for finding unknown quantities, which also include two unknown Lagrange multipliers λ_1 and λ_2. The third condition is the given equality $f_2(x_1, x_2) = 0$. It seems that we have three equations in four unknowns. However, as can be seen from relations (2.4), there is no need to define the Lagrange multipliers separately. It is enough to determine the ratio[38] λ_2/λ_1.

Example 2.12 *Find a rectangle of maximum area for a given perimeter*[39].

A rectangle is characterized by the lengths x_1 and x_2 of its sizes. The area of the rectangle is determined by the formula $S = x_1x_2$, and its perimeter is equal to $p = 2(x_1 + x_2)$. Thus, we minimize the function $f_1(x_1, x_2) = -S = -x_1x_2$ with condition $f_2(x_1, x_2) = p–2(x_1 + x_2) = 0$. Determine the Lagrange function

$$L(x, \lambda) = \lambda_1 f_1(x_1, x_2) + \lambda_2 f_2(x_1, x_2) = -\lambda_1 x_1 x_2 + \lambda_2[p–2(x_1 + x_2)].$$

The equalities (2.4) have the form

$$\lambda_1 x_1 + 2\lambda_2 = 0,\ \lambda_1 x_2 + 2\lambda_2 = 0.$$

Therefore, $x_1 = x_2$. From the equality $f_2(x_1, x_2) = 0$, we obtain $p = 4x_1$, $x_1 = x_2 = p/4$. Thus, the rectangle of maximum area for a given perimeter is a square. Note that problems of this nature are called ***isoperimetric***[40].

2.2.3 Penalty method

One can find an approximate solution of Problem 2.2 by transformation in to the unconditional function minimization problem with using of the ***penalty method***[41]. Consider the function

$$f_\varepsilon(x_1, x_2) = f_1(x_1, x_2) + \frac{1}{2\varepsilon}[f_2(x_1, x_2)],$$

where ε is a small positive parameter. If the value under the square brackets is large enough, then the corresponding value of the function F_ε is large. Therefore, we can have its minimum only for the case with small enough value under that brackets that corresponds to the approximate realization of the equality $f_2(x_1, x_2) = 0$. Using this idea, we can find the approximate solution of the initial problem[42].

Find the point of minimum for the function F_ε using the Fermat theorem by equaling to zero the partial derivatives of this function or by approximate methods of non-conditional optimization; see Section 2.2.5. Particularly, the stationary condition for this problem gives us two equalities.

$$\frac{\partial F_\varepsilon(x_1,x_2)}{\partial x_j} = \frac{\partial f_1(x_1,x_2)}{\partial x_j} + \frac{1}{\varepsilon}\frac{\partial f_2(x_1,x_2)}{\partial x_j} = 0,\ j=1,2.$$

Denote $\lambda = \varepsilon^{-1} f_2(x_1,x_2)$. Now we have three equations

$$\frac{\partial f_1(x_1,x_2)}{\partial x_j} + \lambda\frac{\partial f_2(x_1,x_2)}{\partial x_j} = 0,\ j=1,2;\ \ f_2(x_1,x_2) = \varepsilon\lambda$$

with respect to the three unknown values x_1, x_2, and λ. Two first equalities here correspond to (2.4) for $\lambda = \lambda_2/\lambda_1$, and the third one with small enough value ε can be interpreted as an approximation of the equality $f_2(x_1,x_2) = 0$. Thus, the penalty method really allows one to find an approximate solution to the function minimization problem with a constraint in the form of equality[43].

For Example 2.12, we have $f_1(x_1,x_2) = -x_1x_2$, $f_2(x_1,x_2) = p–2(x_1+x_2)$. Then the above three equalities take the form

$$x_2 + 2\lambda = 0,\ x_1 + 2\lambda = 0,\ p–2(x_1+x_2) = \varepsilon\lambda.$$

From the first two conditions, follows $x_1 = x_2 = –2\lambda$. Then the third equality is written as $p + 8\lambda = \varepsilon\lambda$. Now we obtain $\lambda = p/(\varepsilon–8)$; so $x_1 = x_2 = p/(4–\varepsilon/2)$. Obviously, for small enough values of ε, we actually obtain values that are close enough to the previously established exact solution of the problem.

2.2.4 Gradient methods

In all the examples described earlier, rather simple functions were considered. As a result, finding a solution to the corresponding optimality condition did not cause any particular difficulties. However, in the general case, we are dealing with a non-linear algebraic equations, the analysis of which is far from obvious. Here one can use the general methods for solving algebraic equations.

Suppose we have the equation

$$F(x) = 0.$$

For its approximate solving, one can use the following algorithm[44]:

$$x_{k+1} = x_k + \theta_k F(x_k), \tag{2.6}$$

where θ_k is a numerical parameter. It is chosen in such a way that convergence $\theta_k \to \theta$ takes place. Assume that the sequence $\{x_k\}$ is convergent [45], i.e., has a limit x. Then, as a result of passing to the limit in equality (2.4), in the case of continuity of the function F, we obtain $x = x + \theta F(x)$. Thus, this limit turns out to be a solution of this equation.

As is known, the problem of minimizing a function f is reduced to solving the algebraic equation $f'(x) = 0$. Then relation (2.6) is written as follows

$$x_{k+1} = x_k + \theta_k f'(x_k).$$

However, note a circumstance that is not typical for general algebraic equations. This is related to the fact that in reality the object of study is not an equation, but the function minimization problem; see Figure 2.11. If the derivative of the function is positive at a point x_k, then the function itself is increasing at that point. Therefore, at the next step of the algorithm, we would like to obtain the value x_{k+1} less than x_k. If the derivative $f'(x_k)$ is negative, then the function at the point x_k decreases. Then it is desirable to get the value x_{k+1} greater than x_k. Both options are implemented when the parameter θ_k is negative. Thus, we have the following algorithm

$$x_{k+1} = x_k - \beta_k F(x_k), \tag{2.7}$$

where $\beta_k > 0$, which corresponds to the ***gradient method***[46].

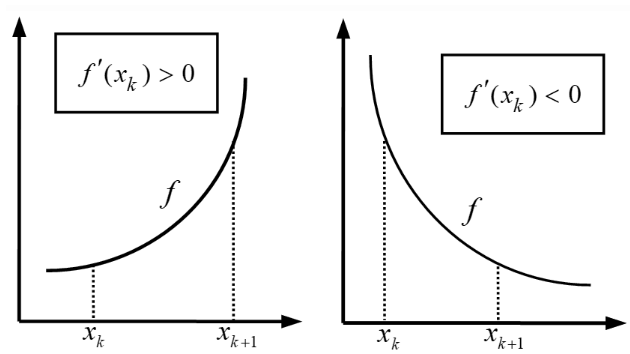

Figure 2.11 Relations between the current and subsequent approximations.

Example 2.6 sets the task of minimizing the function $f(x) = x^2/2 + e^{-x}$. The corresponding derivative is $f'(x) = x - e^{-x}$. Then equality (2.7), which underlies the gradient method, takes the form $x_{k+1} = x_k - \beta_k\big(x_k - e^{-x_k}\big)$.

For the problem of minimization the function f on the interval $[a, b]$, we can use of the gradient method, because the value x_{k+1} determined in accordance with formula (2.7) may go beyond the considered interval. Then one can do the following. First, at the kth iteration, the value $y_k = x_k - \beta_k f'(x_k)$. As a subsequent approximation x_{k+1}, the value of y_k is chosen if it belongs to the given segment, and the boundary of this segment closest to y_k if this value lies outside it. Thus, we obtain the equality

$$x_{k+1} = \begin{cases} a, & \text{if } y_k < a, \\ y_k, & \text{if } a \le y_k \le b, \\ b, & \text{if } y_k > b. \end{cases} \tag{2.8}$$

This is the ***gradient projection method***[47].

Example 2.1 considered the minimization problem for the function $f(x) = (x+1)^2$ on the segment $[a, b]$. Then, in accordance with the gradient projection method, at the kth step, we determine the value $y_k = x_k - 2\beta_k(x_k + 1)$. Further, this value is chosen as a new approximation if it belongs to the given segment, the number a, if the resulting value is less than a, and the number b, if y_k is greater than b. According to Theorem 2.1, the solution of the problem of minimizing a function on an interval satisfies the variational inequality. Thus, the gradient projection method, in the case of its convergence, gives a solution to the corresponding variational inequality.

2.2.5 Hadamard well-posedness of extremum problems

Examples 2.3 and 2.4 dealt with minimization problems for functions that depend on a parameter. Moreover, in the first case, the solution of the problem continuously depended on the parameter, which corresponds to the Hadamard well-posedness of the problem. In the second case, this property did not take place. Below, we present one result on the well-posedness of the function minimization problem.

Consider the problem minimization for the function $f = f(\mu, x)$ for a fixed value of the parameter μ. The following assertion is true[48].

Theorem 2.5 *Suppose that the function $f = f(\mu, x)$ is continuous with respect to the first argument uniformly in x and strongly convex in the second argument for any μ, and the considered problem has a solution for all μ. Then this problem is Hadamard well-posed.*

Proof. Consider an arbitrary convergent sequence of parameters $\{\mu_k\}$. Thus, there is such a value μ, that $\mu_k \to \mu$. Denote by x_k and x the solutions to the minimization problems for the functions $g_k = g_k(y) = f(\mu_k, y)$ and $g = g(y) = f(\mu, y)$. To prove theorem, it suffices to establish the convergence $x_k \to x$.

Consider the equality

$$g(x_k) - g(x) = [g(x_k) - g_k(x_k)] + [g_k(x_k) - g_k(x)] + [g_k(x) - g(x)].$$

Considering that x_k and x are solutions of the corresponding extremal problems, we have the inequalities

$$0 \leq g(x_k) - g(x), \quad g_k(x_k) - g_k(x) \leq 0.$$

Now we obtain

$$0 \leq g(x_k) - g(x) \leq \big|g(x_k) - g_k(x_k)\big| + \big|g_k(x) - g_k(x)\big| \leq 2 \sup_y \big|g_k(y) - g_k(y)\big|.$$

Since the mapping $\mu \to f(\mu, y)$ is continuous uniformly in y, the value on the right-hand side of the last inequality tends to zero. Thus, we get $f(\mu, x_k) \to f(\mu, x)$. This means that the sequence $\{x_k\}$ is minimizing for the function g. However, if the function f is strongly convex in the second argument, the minimization problem for the function g is Tikhonov well-posed. As a result, we obtain the convergence $x_k \to x$, and hence the problem Hadamard well-posedness. □

Example 2.3 considered the function $f(\mu,x) = (x-\mu)^2$. We check the convexity property by the second argument. For any values of μ, x, y, and $\alpha \in (0,1)$ we find the value

$$f(\mu,\alpha x+(1-\alpha)y)-\alpha f(\mu,x)-(1-\alpha)f(\mu,y) = [\alpha x+(1-\alpha)y-\mu]^2-\alpha(x-\mu)^2-(1-\alpha)(y-\mu)^2$$
$$= -\alpha(1-\alpha)\big[(x-\mu)^2-2(x-\mu)(y-\mu)+(y-\mu)^2\big] = -\alpha(1-\alpha)(x-y)^2.$$

Thus, the condition of strong convexity in the form of equality with constant $c = 1$ is true. Now check the property of continuity with respect to the first argument. For any values of μ, ν, and x we have

$$|f(\mu,x)-f(\nu,x)| = |(x-\mu)^2-(x-\nu)^2| \le (2|x|+|\mu|+|\mu|)|\mu-\nu|.$$

Therefore, this function is continuous in the second argument for any x. This is not enough to apply Theorem 2.5, which requires uniform continuity[49]. However, in the case when the considered function is minimized on a limited set, for example, on a segment, the last inequality already implies uniform continuity , and hence the well-posedness of the corresponding problem. At the same time, the function $f(\mu,x) = 3x^4-16x^3+6\mu x^2$ considered in Example 2.4 is not convex at all; see Figure 2.2. Thus, the previously established lack of well-posedness of the problem of its minimization seems to be quite natural.

2.2.6 Approximate solutions to the function minimization problem

As already noted, in practice, the problem of finding the function extremum, as a rule, is solved approximately, i.e., some error in determining such a solution is allowed. In this case, different types of approximate solutions are possible, depending on what exactly is meant with error.

Consider first the problem for the unconditional extremum of a function.

Definition 2.6 *A number x is called a **strong approximate solution** to the function minimization problem if, for a sufficiently small number $\delta > 0$, the inequality $|x-x_{opt}| \le \delta$ is true, where x_{opt} is the exact solution of this problem. The number x is called a **weak approximate solution** of the minimization problem for the function f if for a sufficiently small number $\varepsilon > 0$ the inequality $f(x) \le \min f + \varepsilon$ is true*[50].

Obviously, for continuous functions, any strong approximate solution is a weak approximate solution, which justifies the terminology used[51].

Example 2.13 *Find a point of minimum of the function $f(x) = 3x^4-16x^3+12x^2$.*

This corresponds to Example 2.4 with $\mu = 2$. Obviously, the value of this function at zero only slightly exceeds its minimum possible value, while the point $x = 0$ itself differs significantly from the point x_{opt} of the minimum of the considered function. Thus, we are dealing with a weak, but not a strong, approximate solution of this minimization problem; see Figure 2.12. Thus, we have a weak, but not a strong, approximate solution of the considered problem[52].

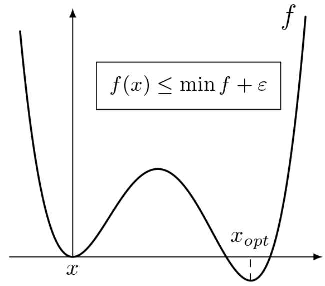

Figure 2.12 Weak approximate solution for Example 2.13 is not strong.

Example 1.7 considered the Tikhonov ill-posed problem of minimizing the function

$$f(x) = \frac{x^2}{1+x^4}.$$

It has a single minimum point $x = 0$, and $f(0) = 0$. Obviously, for sufficiently large values of x, the number $f(x)$ will be arbitrarily close to zero, i.e., to the minimum of this feature. Thus, the point x turns out to be a weak approximate solution of the problem, but by no means its strong approximate solution. Tikhonov-ill-posed problems are natural examples of problems in which the weak approximate solution may differ from their strong approximate solutions.

Example 2.14 *Find a point of minimum of the **Dirichlet function** that takes the value* 1 *for all rational values of the argument and the value* 0 *for irrational values.*

Obviously, the minimum of this function is equal to zero and is reached at all irrational points. As one knows, in an arbitrarily small neighborhood of any irrational point, there is a rational point[53]. Thus, any rational point can be understood as a strong approximate solution of the Dirichlet function minimization problem. At the same time, the value of this function at this rational point differs significantly from its minimum, which means that the latter is not a weak approximate solution. This example shows that for discontinuous functions, a strong approximate solution to the function minimization problem may not be a weak approximate solution.

Let us now turn to the problem of the conditional function extremum. Let it be required to minimize the function f on a set U. A strong approximate solution of this problem is introduced by analogy with Definition 2.6. However, when characterizing a weak approximate solution, one can introduce some refinement, the meaning of which is illustrated by the following example.

Example 2.15 *Find a point of minimum of the function $f(x) = x$ on the set U of positive numbers.*

Obviously, the solution of this problem does not exist, because of the absence of a minimum positive number. However, an arbitrarily small positive number makes sense, i.e., such an element of the set U, on which the value of the minimized function is arbitrarily close to the lower bound of this function on the given set. It is quite reasonable and can be chosen as a weak approximate solution, although the exact solution of the problem does not exist[54].

Definition 2.7 *The number $x \in U$ is called a **strong approximate solution** to the problem of minimizing a function on a given set if the inequality $|x – x_{opt}| \leq \delta$ is true for a small enough number $\delta > 0$, where x_{opt} is the exact solution of this problem. The number $x \in U$ is called a **weak approximate solution** of the problem of minimizing the function f on the set U, if the inequality $f(x) \leq \min f + \varepsilon$ is true for a small enough number $\varepsilon > 0$.*

In particular, a sufficiently small positive number can be chosen as a weak approximate solution to the problem from Example 2.14, which does not have an exact solution[55].

Thus, when determining the approximate solution of the problem, a small error is allowed in determining the minimum point of the function or the minimum of this function. However, when solving the conditional extremum problem in Section 2.2.3, a small error in the fulfilment of the given constraints was also allowed, which leads to another type of approximate solution of the problem.

Definition 2.8 *The number x is called a **strong conditional approximate solution** to the problem of minimizing the function f on the set U if the inequality $|x – x_{opt}| \leq \delta$ is true for a small enough number $\delta > 0$, where x_{opt} is the exact solution of this problem. The number x is called a **weak conditional approximate solution** of this problem, if the inequalities $f(x) \leq \min f + \varepsilon$ and $|x – y| \leq \chi$ are true for an element $y \in U$ and small enough numbers $\varepsilon > 0$ and $\chi > 0$.*

A feature of the conditional approximate solution is that the point x does not have to belong to the set U, but it turns out to be close enough to some element of this set. For a strong conditional approximate solution, this is the exact solution of the problem, and for a weak conditional approximate solution, it is some point y of this set, moreover, the exact solution may not exist. Thus, we can assume that for conditional approximate solutions, minor deviations from the implementation of the given restrictions are allowed[56].

In particular, as a result of applying the penalty method to the problem from Example 2.12 (finding a rectangle of maximum area for a given perimeter p), the sides of the corresponding rectangle $x_1 = x_2 = p/(4 – \varepsilon/2)$ were determined. The perimeter of this rectangle is $4p/(4 – \varepsilon/2)$, i.e., not at all p. Thus, the given restriction is violated. However, the found values for small ε turn out to be quite close to the exact solution of the problem equal to $p/4$. Consequently, we have indeed found a strong conditional approximate solution to the problem.

2.2.7 Minimization of functions of many variables

The obtained results are naturally extended to minimization problems for functions of many variables. First, consider the conditional minimization problem for a function of many variables[57].

Problem 2.3 *Find a point of minimum of the function $f = f(x_1, x_2, ..., x_n)$ on a set U of n-dimensional Euclidean space.*

In order for a differentiable function of many variables $f = f(x)$ to have its minimum at a point $x = (x_1, x_2, ..., x_n)$ on a set U, it is necessary, and if the function is convex, it is sufficient that it satisfies the variational inequality

$$\langle \nabla f(x), y - x \rangle \geq 0 \ \ \forall y \in U.$$

Problem 2.4 [58] *Find a point of minimum for the function of many variables $f_0 = f_0(x)$ on the set of such vectors x of the nth order that satisfy the equalities $f_i(x) = 0,\ i = 1, ..., s,$ where $s < n$.*

For $n = 2$, $s = 1$, we have Problem 2.2. Using the ***Lagrange multiplier method***, determine the ***Lagrange function***

$$L = L(x, \lambda) = L(x_1, ..., x_n; \lambda_1, ..., \lambda_s) = \sum_{i=0}^{s} \lambda_i f_i(x_1, ..., x_n).$$

Assume that all functions f_i are continuously differentiable. The n-th order vector x^* is the point of the local extremum of the considered function under the given constraints if there is such a non-zero vector $\lambda^* = (\lambda_1^*, ..., \lambda_n^*)$ that the following equalities hold

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_j} = 0, \ \ j = 1, ..., n.$$

An approximate solution to Problem 2.4 can also be found using the ***penalty method***. In this case, the minimization problem for the function of many variables

$$F_\varepsilon(x) = f_0(x) + \frac{1}{2\varepsilon} \sum_{i=0}^{s} [f_i(x)]^2$$

is solved, where ε is a small positive parameter. We have a concrete implementation of Problem 1.2, described in Chapter 1.

Finally, consider the multivariate analogues of the previously described gradient methods. An approximate solution to Problem 1.2, consisting in the unconditional minimization of a function of many variables $f = f(x)$, can be found using the ***gradient method***

$$x_{k+1} = x_k - \beta_k \nabla f(x_k), \ \ k = 0, 1, ...,$$

where $\beta_k > 0$. The presence of the function gradient on the right side of this equality explains the name of this algorithm. To solve Problem 2.3 of minimizing a function of many variables $f = f(x)$ on a subset U of the Euclidean space, the ***gradient projection method*** is used, characterized by the equality

$$x_{k+1} = P\Big[x_k - \beta_k \nabla f(x_k)\Big], \quad k = 0, 1, ...,$$

where the operator P is called a ***projector*** and maps the given point to the point of the set U nearest to it, i.e., implements the projection of a vector obtained in accordance with the usual gradient method onto a given set.

Additional conclusions

Based on the results presented in Appendix, some additional conclusions can be drawn about the function minimization problem.

- The variational inequality is a necessary and sufficient condition for the minimum of a convex differentiable function on a segment.
- For non-convex functions, the solution of the variational inequality may not be its minimum point on the segment.
- To solve the problem of minimizing a function of two variables in the presence of an equality constraint, one can use the Lagrange multiplier method, and the unknown arguments of the function and the Lagrange multipliers are determined from the equality to zero of the partial derivatives of the corresponding Lagrange function and the given constraint.
- The approximate solution of the conditional extremum problem obtained using the penalty method satisfies the given constraint with some error.
- The problem of minimizing a strongly convex function, uniformly continuous in a parameter, solvable for all values of this parameter, is Hadamard well-posed.
- In the absence of a strong convexity of a function, the problem of its minimization may turn out to be Hadamard ill-posed.
- For an approximate solution of the problem of minimizing a differentiable function, one can use the gradient method.
- For an approximate solution of the problem of minimizing a differentiable function on a segment, one can use the gradient projection method.
- The gradient projection method can be an iterative method for solving a variational inequality.
- A strong and weak approximate solution of a function minimization problem is, respectively, as an admissible value that is close enough to the exact solution of the problem, and one on which the value of the minimized function is close enough to its lower bound on this set.

- For continuous functions, any strong approximate solution is a weak approximate solution, although this may not be the case for discontinuous functions.
- A weak approximate solution makes sense even if there is no exact solution to the problem, if the infimum of the function exists on the given set.
- In problems with constraints, the concept of a conditional approximate solution makes sense, which can satisfy these constraints not exactly, but approximately.
- The obtained results can be extended to problems of minimizing functions of many variables.

Notes

1. Problem 2.1 will be used to formulate the maximum condition for the optimal control problem in Chapter 3.

2. Variational inequalities will be used in what follows to solve optimal control problems; see Chapters 4, 5, 6, 10, and 13.

3. Indeed, in the absence of restrictions, the relation $f'(x)(y–x) \geq 0$ is realized for all points y of the real line. For any non-zero number h, choosing $y = x + h$, we get $f'(x)h \geq 0$. Having now determined $y = x–h$, we establish that $-f'(x)h \geq 0$. From the obtained two inequalities, due to the arbitrariness of h, it follows that $f'(x) = 0$. Thus, in the problem of minimizing a differentiable function on the entire set of real numbers, the variational inequality and the stationary condition are equivalent. Thus, the stationary condition can be understood as a special case of variational inequality corresponding to the absence of restrictions. In fact, the established relationship between the two considered extremum conditions remains valid for the much more general problem of minimizing a differentiable functional on a convex subset of a topological vector space; see, for example, [60], [70], [116], [171].

4. For $a = –1$ and for $b = –1$, there are actually only two solutions.

5. Examples of optimal control problems for which variational inequality is not a sufficient optimality condition are considered in Chapters 4, 5, and 10.

6. In the previous chapter, the notion of the Tikhonov well-posedness of an extremal problem was considered, which has a different meaning. Both types of well-posedness for optimal control problems are studied in Chapter 8; see also [170], [194], [211].

7. The concept of Hadamard well-posedness makes sense not only for extremum problems; see for example, Tikhonov. In particular, the famous ***Hadamard example*** considers the Cauchy problem for the Laplace equation, which is ill-posed. In particular, there is no continuous dependence of the solution on the data at the boundary for it. The boundary value problem for the heat equation with data at the final moment of time is ill-posed too. Note that well-posedness for the Cauchy problem for an ordinary differential equation is used in Chapter 3 when deriving an optimality condition.

8. For $\mu = 0$ and $\mu = 4$, we get two stationary points instead of three. The resulting properties are not of interest to us in this case, although Example 2.10 considers similar effects.

9. For $f(x_3) = 0$, the value of the function at the points x_1 and x_3 are the same, and the problem has two solutions.

10. This conclusion is a consequence of the well-known statement of mathematical analysis, according to which a continuous function that takes values of different signs at the ends of a certain segment vanishes at some point inside this segment.

11. In this case, Hadamard ill-posedness is due to the fact that the function has two local minima. In this case, with a small change in the problem parameter, the value of the function at the point that initially corresponded to the absolute minimum may turn out to be greater than the value of the function at the point of the second local minimum, despite the fact that these points themselves can be quite far from each other. Examples of optimal control problems that are not Hadamard well-posed are given in Chapters 8, 12, 15, and 17.

12. This, of course, is about real solutions, since we consider only functions of a real variable.

13. On the theory of bifurcations, see [96], [107], [184], [201].

14. The bifurcation phenomenon for optimal control problems will be established in Chapters 12 and 15.

15. A ***transcendental equation*** is an equation that is not algebraic, i.e., does not correspond to the equality to zero of some polynomial. On solving transcendental equations, see [41].

16. Various versions of the method of successive approximations will be used in subsequent parts of this book to solve different problems of optimal control.

17. On the convergence of numerical methods for solving non-linear algebraic equations, see, for example; [17], [189].

18. The above condition is related to the concept of a ***contraction mapping*** (a transformation that transforms a segment into a segment of smaller length) and goes back to ***Banach fixed point theorem***; see [94], [106], [158]. In this case, the solution of the equation $x = F(x)$ corresponds to the point, which is translated by the transformation F into itself, i.e., remains unchanged under the action of this transformation. The analysis of the equation thus reduces to the search for ***fixed points*** of this transformation. We also note that the equality $x_{k+1} = F(x_k)$ can be understood as a discrete dynamical system, and the solution of the equation $x = F(x)$ can be interpreted as the ***equilibrium position*** of this system. Moreover, if the sequence $\{x_k\}$ converges to a solution x for any value of x_0 close enough to x, then this equilibrium is called ***stable***. Otherwise, it is unstable. We will encounter equilibrium positions in Chapters 3, 5, and 7. On the theory of dynamical systems; see, for example, [27], [86]

19. It can be established that the considered function is strictly convex, so that the uniqueness of its minimum point follows from Theorem 1.4. On the other hand, the existence of a minimum point can be established using Theorem 1.3, since this function is continuous, lower bounded, and coercive.

20. Chapter 7 will describe an iterative process for solving a system of optimality conditions for an optimal control problem that diverges for any initial approximation.

21. For any initial iteration the algorithm does not converge to the stationary point $x = 0$ that is the point of local maximum of this function.

22. Similar effects will be encountered in Chapter 5 in the analysis of an optimal control problem.

23. About the logistic equation and its properties, see [184]. Continuous analogue of the ***Verhulst equation*** $x' = \mu x(1–x)$ describes the change in the abundance of a biological species under conditions of limited food intake; see, for example, [173].

24. This follows from the ***fundamental theorem of algebra***, according to which an nth order algebraic equation has n solutions on the set of complex numbers; see [192].

25. We will encounter a related phenomenon of multiple bifurcation of solutions in Chapter 12 when analyzing the Chafee–Infante problem; see also Chapter 15.

26. About ***chaotic dynamics***, see [184].

27. This statement remains valid for the problem of minimizing a function on a semi-infinite interval and on the entire real axis. A necessary optimality condition in the form of variational inequalities for general extremal problems is established, for example, in Lions. The mathematical theory of variational inequalities is given in [102], [117]. Note also that variational inequalities can be mathematical models of physical and not only physical processes; see [57].

28. For the ***singularity theory***; see [11].

29. For the convergence of approximate methods for solving extremal problems; see [49], [149].

30. In Chapter 8, the concepts introduced will be extended to optimal control problems.

31. A more general statement that a variational inequality is a necessary and sufficient condition for a minimum of a convex Gateaux differentiable functional on a convex set is given; for example, in [116], [171].

32. We will meet with examples of the insufficiency of the optimality condition in the form of variational inequalities in Chapter 5.

33. Lagrange multiplier method naturally extends to the problem of minimizing a function of many variables with several constraints in the form of equality.

34. Lagrange multiplier method will be used substantially in Chapter 3 in deriving necessary optimality conditions in the form of the maximum principle. It will also be applied in Chapter 9 when analyzing systems with a fixed final state. In this case, L will no longer be a function, but a functional. In Chapter 13, the Lagrange multiplier method is used to solve optimal control problems in the presence of an additional isoperimetric condition.

35. The implicit function theorem is also true in the multidimensional case, and even for operators; see, for example, in [100].

36. The non-degeneracy of a matrix means that its determinant is non-zero.

37. In fact, a necessary condition for a local extremum of a given function with the considered constraint is given here. A generalization of Theorem 2.4 to the problem of minimization a function of many variables with equality constraints is given in [5], [33], [49], [95], [70], [149], [193], [195]. Its generalization to the functional minimization problem with operator constraints in the form of equalities goes back to ***Lyusternik's theorem*** on the approximation of a smooth manifold [125]; see also [56], [95], [140].

38. In this regard, when determining the Lagrange function, the coefficient of the function being minimized is often, i.e., the first of the Lagrange multipliers is initially assumed to be

equal to one. This is exactly what we will do in Part II when justifying the necessary optimality condition in the form of the maximum principle.

39. In Chapter 13, a more general problem will be considered, in which it is required to determine a curve of a given length in such a way that the corresponding curvilinear trapezoid has a maximum area.

40. In the calculus of variations and the theory of optimal control, ***isoperimetric problems*** are called problems with integral constraints in the form of equalities. Problems of this nature are the subject of Part IV of this book.

41. The penalty method can also be used to minimize functions with constraints in the form of inequalities. For the application of the penalty method to various extremum theory problems; see [42], [49], [70], [118], [193]. We will use the penalty method in Part II when analyzing optimal control problems with interpretation the state equation of the system is as an equality constraint; see Chapter 4. The penalty method can also be used to find singular controls; see Chapter 6. In Chapter 9, the penalty method is used to solve problems of optimal control of systems with a fixed final state, and the restriction removed using this approach is not a state equation, but an equality that characterizes the final state of the system. In Chapter 13, the penalty method is used for solving optimization problems with isoperimetric constraints , and in Chapter 16, we will apply it for analysis systems with a free initial state.

42. In Section 2.9, we clarify what exactly is meant by an approximate solution in this case.

43. To substantiate this assertion, it is required to prove the convergence of the penalty method, i.e., show that the solution of the minimization problem for the function F_ε converges to the solution of the considered problem as $\varepsilon \to 0$; see, for example, [49]. In Chapter 8, we will prove the convergence of the penalty method for solving one optimal control problem for a system with a fixed finite state.

44. The solution of the equation $F(x) = 0$ can be interpreted as the ***equilibrium position*** of the dynamical system described by the ordinary differential equation $x' = F(x)$. To solve it, the derivative can be approximated by the corresponding difference, obtaining the recursive formula $x_{k+1} = x_k + \theta_k F(x_k)$, where θ_k has the meaning of the argument change step, and the initial approximation x_0 is the known initial state that complements the differential equation. We will partially encounter the equilibrium position in Chapters 3, 5, and 7. In addition, we will use this algorithm in solving optimal control problems for systems with a fixed finite state; see Chapter 9.

45. Naturally, the convergence of the iterative method is not obvious

46. The meaning of the term "***gradient method***" is clarified in the Section 2.2.7 when considering minimization problems for the functions of many variables. On the gradient method for functions of many variables, as well as for functionals, see, for example, [26] [35], [42], [49], [65], [68], [70], [79], [82], [149], [193], [195]. We will use the gradient method to study optimal control problems; see Chapter 4. Note also the ***conjugate-gradient method***; see [49], [65], [70], [149], [193]. A generalization of the gradient method to non-smooth functions is the ***subgradient method***, in which one of its subgradients is chosen instead of the derivative; see, for example, [193]. Among other approximate non-smooth optimization methods, i.e., that do not require the calculation of derivatives, we note ***genetic algorithms***; see, for example, [70], [132] and the ***Nelder–Mead method***; see [139].

47. For the gradient projection method in problems of finding the conditional extremum of functions of many variables and general functionals; see [26], [49], [65], [68], [70], [149], [193], [195]

48. A generalization of Theorem 2.5 to optimal control problems is given in Chapter 7.

49. Here it suffices to estimate the value $|x|$ by the maximum of the numbers $|a|$ and $|b|$, where $[a, b]$ is the segment on which the minimum of the function is found.

50. There is no point in using the inequality $|f(x) - \min f| \leq \varepsilon$ here, since the value of the function at the point x cannot be less than the minimum of this function.

51. The difference between weak and strong approximate solutions of the problem will be very significant in optimal control problems that are Tikhonov ill-posed; see Chapter 8, and also Chapters 12 and 15.

52. This circumstance is a consequence of the ill-posedness of the considered problem.

53. This means that the set of rational numbers is dense in the set of all real numbers.

54. Extension methods for optimal control problems are based on this idea; see Chapter 7.

55. Naturally, it makes sense to use the concept of a weak approximate solution only under the conditions of the existence of the lower bound of a function on a given set. For example, for the minimization problem for the function $f(x) = -1/x$ on the set of positive numbers, the concept of an approximate solution does not make sense.

56. It would seem that if in the conditions of the problem it is required to ensure the validity of a certain constraint, then we must certainly guarantee the fulfillment of this property. However, in applied optimization problems, the problem statement itself is usually approximate. In this regard, the presence of a small error in the fulfillment of the specified restrictions, as a rule, can be considered acceptable.

57. We will meet with a specific case of Problem 2.3 in Chapter 3 when considering the maximum principle for the vector optimal control problem.

58. The number of constraints must be less than the number of function arguments. Otherwise, we will not be able to vary these arguments, and the task of finding will lose its meaning.

II

OPTIMAL CONTROL PROBLEMS FOR SYSTEMS WITH A FREE FINITE STATE

The second part of the book deals with the simplest optimal control problem. The state of the system is described by a first-order ordinary differential equation with an initial condition. The control function is included on the right side of the equation and satisfies some restrictions corresponding to the set of admissible controls. The optimality criterion contains an integral term that depends on the control and state function of the system at a time interval, as well as a terminal summand that depends on the state at the final time.

This part consists of six chapters. Chapter 3 gives the necessary optimality conditions in the form of the maximum principle for the stated optimal control problem. Examples of problems for which the optimal control can be found analytically or with the help of an iterative process are considered. Chapter 4 describes alternative methods for studying optimal control problems: the gradient methods for minimizing functionals, the variational inequality, the penalty method, and the Bellman equation are considered. In Chapter 5, the uniqueness of the optimal control and the sufficiency of the optimality condition are investigated; examples of problems for which these properties are satisfied or violated are given. In Chapter 6, we study optimal control problems with a degeneracy of the maximum principle, which corresponds to the singular control. The problem of the existence of an optimal control is considered in Chapter 7, where examples of both solvable and unsolvable optimization problems are given. Finally, Chapter 8 discusses the problem of the well-posedness of optimization problems and gives examples of both well-posed and ill-posed problems.

Maximum principle

The simplest problem of the theory of optimal control is considered[1]. To solve it, the necessary optimality conditions in the form of the maximum principle are derived. They consist of maximizing a function that also depends on the state of the system and the solution of the adjoint system. For fairly simple problems, these optimality conditions can be solved analytically. However, as a rule, they are solved approximately. Appendix provides a justification for the maximum principle, investigates a non-linear boundary value problem equivalent to a system of optimality conditions, describes a class of problems that can be directly solved, and proves the convergence of the iterative process for a specific example.

3.1 LECTURE

A system described by an ordinary differential equation is given. The right side of the equation includes a control function, the values of which belong to a certain interval. It is required to minimize the functional that includes the integral and terminal terms. Section 3.1.1 gives a complete statement of the optimal control problem. To solve this problem, Section 3.1.2 derives an optimality condition in the form of the maximum principle. Section 3.1.3 considers an example of an optimal control problem for which the optimality condition has an analytical solution. For the more difficult problem in Section 3.4, an iterative process is used to find the optimal control.

3.1.1 Statement of the optimal control problem

The optimal control problem, first of all, implies the existence of a ***mathematical model***[2] of the process under study. Any mathematical model is a problem of determining some quantities that characterize the state of the considered system. We are dealing with ***dynamic systems***[3] whose state changes with time[4]. In this case, we restrict ourselves to the study of ***systems with lumped parameters***, the state of which at each moment of time can be described by a finite set of numbers[5].

For simplicity, we assume that there is a unique ***state function*** of the system[6] $x = x(t)$. Since the state of the system changes over time, we can try to estimate how quickly x changes. The rate of change of the function is its derivative. In this regard,

DOI: 10.1201/9781003398585-3

we assume that the basis of the mathematical model is an equality that includes the derivative x' of the system state function. Thus, the system is described by an ***ordinary differential equation***.

In this case, a ***controlled process*** is considered, which implies the presence in the model of some quantities called ***controls***, which can be changed at the will of the researcher. For simplicity, we assume that there is a unique control[7], which is characterized by a function $u = u(t)$. Thus, this system is described by a differential equation that includes control. This equation is considered on a certain time interval. For definiteness, we choose the value $t = 0$ as the initial moment of time, and denote the final moment of time by T.

Based on the assumptions made, we conclude that the considered system is described by a differential equation

$$x'(t) = f(t, u(t), x(t)),\ t \in (0, T), \tag{3.1}$$

which is called the ***state equation***, where f is a given function of its arguments. It is known that the solution of a first-order differential equation is determined up to one arbitrary constant. For the uniqueness of the solution, it is assumed that the state of the system at the initial moment of time is known and is equal to some value x_0. Thus, the function x also satisfies the initial condition

$$x(0) = x_0. \tag{3.2}$$

Thus, equation (3.1) with initial condition (3.2), which constitutes the ***Cauchy problem***, is chosen as the mathematical model of the controlled system.

Suppose that the considered Cauchy problem has a unique solution[8] for an arbitrary control. Note, however, that in practice the possibilities for changing control are usually limited. This means that only ***admissible controls*** are subject to consideration. We will assume that at each moment of time the control can change only within the given limits[9], i.e., the set of admissible controls is characterized by the equality[10]

$$U = \{u \,|\, a(t) \leq u(t) \leq b(t),\ t \in (0, T)\},$$

where a and b are given functions; besides, the the first of them can take the value $-\infty$, and the second $+\infty$.

By specifying different admissible controls, we obtain different variants of the system evolution. Among all such possible variants, one should choose the one that is optimal in some sense. Therefore, an ***optimality criterion*** I should be set, according to the quantitative value of which $I(u)$ in each specific case, one can judge the effectiveness of the selected control u. Since each admissible control (function) u is associated here with a number $I(u)$, the transformation I turns out to be a ***functional***.

Suppose that the optimality criterion is determined by the equality

$$I(u) = \int_0^T g\big(t, u(t), x(t)\big)dt + h(x(T)),$$

where g and h are known functions, and x is a solution of problem (3.1) and (3.2) that depends on the control u. The presence of the integral makes it possible to take into account the influence on the values of the control and the state function at each moment of time at which the system is considered. The presence here of the terminal summand, i.e., the second term on the right side of the last equality emphasizes the special role of the final state of the system, which can be extremely important for assessing the degree of efficiency of the chosen control.

As a result, have the following formulation of the ***optimal control problem with a free final state***[11].

Problem 3.1 *Find a function u (the **optimal control**) from the set U that minimizes on this set the functional I whose definition includes the function x that is the solution of the Cauchy problem* (3.1), (3.2) *for the given function u.*

Let us describe one of the most effective methods for solving this problem.

3.1.2 Maximum principle

In Problem 3.1, the functional is minimized when the equation of state is satisfied, which is an equality. In the previous chapter, the Lagrange multiplier method was used to solve the problem of minimizing a function with a constraint in the form of equality. At the same time, the Lagrange function was introduced as the sum of the minimized function and the function characterizing the given constraint, multiplied by a numerical Lagrange multiplier[12]. If there are several constraints in the form of equalities, then the Lagrange multiplier turns out to be a vector, and the definition of the Lagrange function uses the sum of the products of the components of the vector Lagrange multiplier by the functions characterizing the constraints, which corresponds to the ***dot product*** of the corresponding vectors[13]. Now, the given equality is not numerical, but functional, since the state equation is considered for all values of $t \in (0, T)$. As a result, the Lagrange multiplier is also considered a function, and to take into account the state equation at all points, when determining the function (now the functional) of Lagrange, it is no longer the sum, but the integral[14].

Thus, to solve the problem in accordance with the ***Lagrange multiplier method***[15], the ***Lagrange functional*** is introduced

$$L(u, x, p) = I(u) + \int_0^T p(t)\big[x'(t) - f(t, u(t), x(t))\big]dt,$$

where p is an arbitrary function (***Lagrange multiplier***). Obviously, in the case when the function x satisfies equation (3.1), for any function p, the functionals L and I coincide.

We combine all the terms under the integral in the definition of the functional L, which depends explicitly on the control. To this end, we consider the function[16]

$$H(t, u, x, p) = pf(t, u, x) - g(t, u, x). \tag{3.3}$$

Now, the functional L is

$$L(u,x,p) \;=\; \int_0^T \big[p(t)x'(t) - H(t,u(t),x(t),p(t))\big]dt \;+\; h(x(T)).$$

Let us assume that the function u is an optimal control, i.e., the following inequality holds

$$\Delta I \;=\; I(v)–I(u) \geq 0 \;\; \forall v \in U. \tag{3.4}$$

Therefore, we obtain

$$\Delta L \;=\; L(v,y,p)–L(u,x,p) \geq 0 \;\; \forall v \in U,\, \forall p, \tag{3.5}$$

where x and y are the solutions of the Cauchy problem (3.1) and (3.1) for the controls u and v. Find the functional increment

$$\Delta L \;=\; \int_0^T p(t)[y'(t) - x'(t)]dt \;-\; \int_0^T \Delta H dt \;+\; \Delta h,$$

where

$$\Delta H = H(t,v,y,p)–H(t,u,x,p), \;\; \Delta h = h(y(T))–h(x(T)).$$

Suppose all known functions included in the problem statement are sufficiently smooth. Then, denoting $\Delta x = y–x$ and using the Taylor series, we obtain

$$H(t,v,y,p) \;=\; H(t,v,x+\Delta x,p) \;=\; H(t,v,x,p) + H_x(t,v,x,p)\Delta x + \eta_1,$$

where $H_x = \partial H/\partial x$, η_1 is a value of a higher (greater than the first) order in increment Δx. Now we get

$$H(t,v,y,p) \;=\; H(t,v,x,p) + H_x(t,u,x,p)\Delta x + \eta_1 + \eta_2,$$

where

$$\eta_2 \;=\; \big[H_x(t,v,x,p) - H_x(t,u,x,p)\big]\Delta x.$$

Note that the value η_2 also has a second order with respect to the increments[17]. The equality

$$h(y(T)) \;=\; h\big(x(T) + \Delta x(T)\big) \;=\; h(x(T)) + h_x(x(T))\Delta x(T) + \eta_3,$$

is established in a similar way, where $h_x = dh/dx$; and η_3 is a value of a higher order in increment $\Delta x(T)$. Now, we reduce inequality (3.5) to the following form

$$\int_0^T p(t)\Delta x'(t)dt - \int_0^T \big[\Delta_u H + H_x(t,u,x,p)\big]dt + h_x(x(T)) + \eta \;\geq\; 0 \; \forall v \in U, \;\; \forall p, \tag{3.6}$$

where $\Delta_u H$ is the increment of the function H with respect to the control

$$\Delta_u H \;=\; H(t,v,x,p)–H(t,u,x,p),$$

and η is the ***remainder term***[18], determined by the formula

$$\eta = \eta_3 - \int_0^T (\eta_1 + \eta_2)dt,$$

and having the second order with respect to increments.

Integrating by parts, find the value

$$\int_0^T p(t)\Delta x'(t)dt = p(T)\Delta x(T) - \int_0^T p'(t)\Delta x(t)dt,$$

because $x(0) = y(0) = x_0$. As a result, formula (3.6) is reduced to the inequality

$$-\int_0^T \Delta_u H dt - \int_0^T \big[H_x(t,u,x,p) + p'\big]\Delta x dt + \big[h_x(x(T)) + p(T)\big]\Delta x(T) + \eta \geq 0 \quad (3.7)$$

for all $v \in U$ and arbitrary p.

Taking into account the arbitrariness of the function p, choose it in such a way that formula (3.7) has the simplest possible form. In particular, it is possible to vanish the second and third terms on the left side of this inequality. To do this, it suffices to assume that the function p satisfies the equation

$$p'(t) = -H_x(t,u,x,p),\ t \in (0,T) \quad (3.8)$$

with condition

$$p(T) = -h_x(x(T)). \quad (3.9)$$

Definition 3.1 *Problem* (3.8) *and* (3.9) *is called the* ***adjoint system***[19].

As a result of specifying the function p, inequality (3.7) can be written as follows:

$$-\int_0^T \Delta_u H dt + \eta \geq 0 \ \ \forall v \in U. \quad (3.10)$$

Let τ be a point from the interval $(0,T)$, and w is an arbitrary admissible control. Determine the control (see Figure 3.1)

$$v_{\varepsilon\tau}^w(t) = \begin{cases} u(t), & \text{if } t \notin (\tau-\varepsilon, \tau+\varepsilon), \\ w(t), & \text{if } t \in (\tau-\varepsilon, \tau+\varepsilon), \end{cases}$$

where ε is a small enough positive number. The function $v_{\varepsilon\tau}^w$ is called the ***needle variation***[20] of the control u. Obviously, it is admissible[21], i.e., belongs to the set U.

Substituting the control $v = v_{\varepsilon\tau}^w$ into inequality (3.10) and taking into account that the controls v and u differ only on the interval $(\tau-\varepsilon, \tau+\varepsilon)$, we get

$$\int_{\tau-\varepsilon}^{\tau+\varepsilon} \Big\{H\big[t, w(t), x(t), p(t)\big] - H\big[t, u(t), x(t), p(t)\big]\Big\}dt - \eta_{\varepsilon\tau}^w \leq 0, \quad (3.11)$$

where $\eta_{\varepsilon\tau}^w$ is the remainder term η for $v = v_{\varepsilon\tau}^w$.

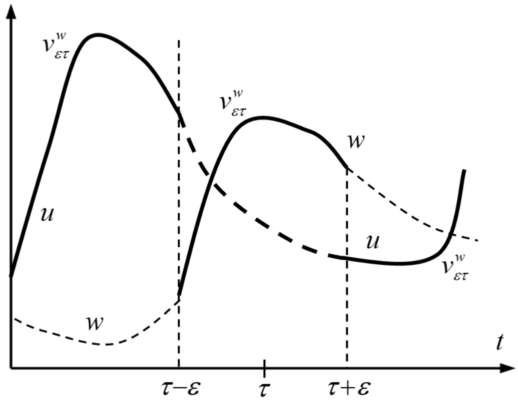

Figure 3.1 Needle variation of control.

For small values of ε, the control $v_{\varepsilon\tau}^{w}$ is close enough[22] to u so that their difference is of the order of ε. It is assumed that the solution of problem (3.1) and (3.2) continuously depends on the control so that the increment Δx also has the order[23] ε. Then the value $\eta_{\varepsilon\tau}^{w}$ has the order of ε^2, i.e., $\eta_{\varepsilon\tau}^{w}/\varepsilon$ tends to zero as $\varepsilon \to 0$. We divide inequality (3.11) by 2ε and pass to the limit as $\varepsilon \to 0$, taking into account the mean value theorem[24]. We obtain

$$H\big[\tau, w(\tau), x(\tau), p(\tau)\big] - H\big[\tau, u(\tau), x(\tau), p(\tau)\big] \leq 0.$$

Taking into account the arbitrariness of the point $\tau \in [0, T]$ and the control value $w(\tau) \in [a(\tau), b(\tau)]$, we have the equality

$$H\big[t, u(t), x(t), p(t)\big] = \max_{v \in [a(t), b(t)]} H\big[t, v, x(t), p(t)\big], \quad t \in [0, T]. \tag{3.12}$$

As a result, we have the following assertion.

Theorem 3.1 *In order for the control u to be a solution to Problem* 3.1, *it is necessary that it satisfies equality* (3.12), *where x is the corresponding solution to problem* (3.1) *and* (3.2), *and p is the solution of the adjoint system* (3.8) *and* (3.9).

Definition 3.2 *The statements of Theorem* 3.1 *are called the* ***maximum principle***[25]*, or more fully, the* ***Pontryagin maximum principle****, and equality* (3.12) *is called the* ***maximum condition***[26].

In accordance with the maximum principle, to solve the problem, it is required to find the functions u, x, and p from relations (3.1), (3.2), (3.8), (3.9), and (3.12). The effectiveness of the maximum principle is due to the fact that we have moved from the problem of minimizing the original functional to the problem of the conditional extremum of the function H, which explicitly depends on the control[27]. However, for the possibility of this transition, one has to pay with the appearance of another unknown function p[28].

The practical application of the maximum principle for solving a specific optimization problem is reduced to the following steps:

1. The function H is introduced in accordance with formula (3.3).
2. The adjoint system (3.8) and (3.9) is determined.
3. From the maximum principle (3.12), which is a problem for the conditional extremum of a function, a control is found that will depend on the functions x and p.
4. Substituting the established dependence into relations (3.1) and (3.8), we obtain a system of two first-order differential equations for the functions x and p, for which there are also two boundary conditions (3.2) and (3.9).
5. Solving the resulting system, we find the functions $x = x(t)$ and $p = p(t)$.
6. Substituting the found functions into the control formula set in step 3, we determine the dependence $u = u(t)$.
7. Check if the found control is a solution to Problem 3.1.

Let us apply the described method to the analysis of specific examples.

3.1.3 Analytical solving of an optimal control problem

Consider an easy optimal control problem.

Example 3.1 *Find a function $u = u(t)$ from the set*

$$U = \{u \big| \; 1 \le u(t) \le 2, \; 0 < t < 1\},$$

which minimizes there the functional

$$I(u) = \int_0^1 \Big(\frac{u^2}{2} - 3x\Big) dt,$$

where x is the solution of the Cauchy problem

$$x'(t) = u(t); \; t \in (0,1); \; x(0) = 0.$$

We have the Problem 3.1 with parameters

$$f(t,u,x) = u, \; T = 1, \; x_0 = 0, \; a(t) = 1, \; b(t) = 2, \; g(t,u,x) = u^2/2–3x, \; h(x) = 0.$$

In accordance with the described method, determine the function

$$H(t,u,x,p) = pf–g = pu–u^2/2 + 3x.$$

Find the solution of the maximum condition (3.12).

First of all, using the stationary condition, equal to zero the derivative H_u of this control function. We get the equality $p – u = 0$, whence it follows that $u = p$. Since the second derivative of the function H with respect to control is equal to −1, i.e., is negative, we conclude that the found stationary point really corresponds to the maximum H. If at a time t the value $p(t)$ belongs to the set of admissible control values, i.e., on the interval [1,2], then it is chosen as a solution to the maximum condition. For $p(t) < 1$, the difference $p(t) – u(t)$ is negative, since, according to the condition of the problem, the number $u(t)$ must belong to the interval [1,2]. Then, the derivative H_u is negative here, which means that H is a decreasing function of control. Therefore, the maximum is reached at the minimum possible point, i.e., $u(t) = 1$. Finally, for $p(t) > 2$, the difference $p(t) – u(t)$ is positive, since $u(t)$ cannot be greater than 2. Then the derivative H_u is positive and H is an increasing function of control. Its maximum is reached at the maximum possible point, i.e., $u(t) = 2$. Thus, according to the maximum principle (3.12), the control is determined by the formula

$$u(t) = \begin{cases} 1, & \text{if } p(t) < 1, \\ p(t), & \text{if } 1 \le p(t) \le 2, \\ 2, & \text{if } p(t) > 2. \end{cases}$$

Now, it is necessary to find the function p. The adjoint system (3.8) and (3.9) now is

$$p'(t) = -3,\ t \in (0,1);\ p(1) = 0.$$

Find its solution $p(t) = 3 – 3t$. It is less than one at $t > 2/3$, greater than two at $t < 1/3$, and takes values from the segment [1,2] at $1/3 \le t \le 2/3$. Thus, the solution of the maximum principle is determined by the formula; see Figure 3.2.

$$u(t) = \begin{cases} 1, & \text{if } 0 < t < 1/3, \\ 3 - 3t, & \text{if } 1/3 \le t \le 2/3, \\ 2, & \text{if } 2/3 < t < 1. \end{cases}$$

It is easy to verify that the obtained value is indeed the solution of the given optimal control problem[29].

Consider now the following example:

Example 3.2 *Find the point of maximum for the functional of Example* 3.1.

Obviously, a control that delivers a minimum to the functional $-I$ will certainly deliver a maximum to the functional I. Therefore, if to minimize the functional I it is required to find the maximum of the function H, then, obviously, the search for the maximum of the same functional will be reduced to minimizing the same function H.

It is clear that the unique solution of the stationary condition for the function H, defined above, provides the maximum of this function. Under these conditions, its minimum can be realized exclusively at the boundaries of the set of admissible control values, i.e., at points 1 or 2. We determine the corresponding values, taking into account the previously found solution of the adjoint system $p(t) = 3 – 3t$. We have

$$H|_{u=1} = p – 1/2 + 3x = 2.5 – 3t + 3x,\ H|_{u=2} = 2p – 2 + 3x = 4 – 6t + 3x.$$

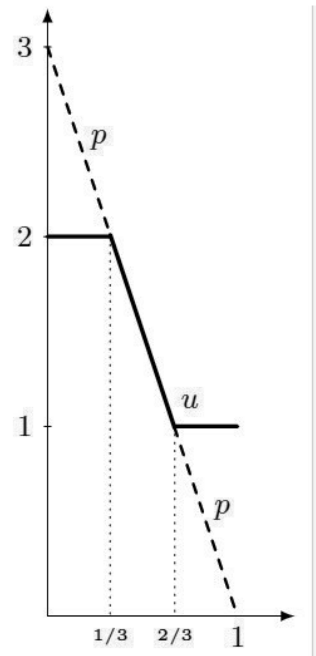

Figure 3.2 Optimal control for Example 3.1.

By the maximum principle, one should choose from the limit values of the control that corresponds to the smaller of these values. As a result, we find the unique solution of the optimality conditions

$$u(t) = \begin{cases} 1, & \text{if } 0 < t < 1/2, \\ 2, & \text{if } 1/2 < t < 1. \end{cases}$$

It is easy to verify[30] that it is optimal for Example 3.2.

The considered examples are in a sense similar to Example 1.1, in which the stationary condition allowed us to find the only minimum point of the given function[31]. However, even for the problem of minimizing a function of one variable in the absence of any restrictions, it is far from always possible to find a solution explicitly. This is all the more true for the much more difficult Problem 3.1. However, as for the problems studied in Part I, the possibility of an approximate solution of the problem with the help of some iterative methods remains here.

3.1.4 Approximate solving of an optimal control problem

In Examples 3.1 and 3.2, both the equation of state and the functional to be minimized were linear with respect to the state function x. As a result, the function H also turned out to be linear with respect to x. As can be seen from equality (3.7), the derivative of H with respect to x includes the right side of the adjoint equation. Thus, in the considered cases, this derivative does not depend on x. In addition, the functional lacked a terminal term defined by the function g. In this connection, the boundary condition (3.8) turned out to be homogeneous. As a result, the adjoint system did not include the functions u and x, and it could be solved independently of the equations of state and the maximum condition. This predetermined the possibility of finding an analytical solution to the formulated problem of optimal control. The situation changes even with a slight complication of the problem statement.

Example 3.3 [32] *Find a function $u = u(t)$ from the set*

$$U = \{u|\ |u(t)| \le 1,\ 0 < t < 1\},$$

which minimizes there the functional

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is the solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{3.13}$$

Compare Examples 3.1 and 3.3. The state of the system in both cases is described by the same Cauchy problem. The fact that in the first case the control values are chosen from the interval [1,2], and in the second case, we have the interval [–1,1], is of no fundamental importance[33]. However, in the first case the functional is linear with respect to the function x, and in the second case it is quadratic, which, as we will soon see, is very important. Nevertheless, we try to solve this problem by a known method[34].

To reduce this problem to the standard form, we define

$$f(t,u,x) = u,\ T = 1,\ x_0 = 0,\ a(t) = 1,\ b(t) = 2,\ g(t,u,x) = (u^2 + x^2)/2,\ h(x) = 0.$$

Determine the function

$$H = pu–(u^2 + x^2)/2.$$

The adjoint system (3.7) and (3.8) takes the form

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0. \tag{3.14}$$

Consider now the maximum condition (3.12).

Equating to zero the derivative of the function H with respect to control, we obtain the same equality $u = p$ as in the previous example. Taking into account the negativity of the corresponding second derivative, we conclude that we are dealing with a maximum point. However, at each time t, the value of $p(t)$ can be located differently relative to the given set of admissible control values [–1,1]; see Figure 3.3. When $p(t) < 1$, the function H decreases with respect to the control on this segment, which means that its maximum is reached at the minimum allowable control value, i.e., $u(t) = –1$. If $p(t) > 1$, then the function H increases on a given interval, and, consequently, its maximum is reached on the maximum admissible control, i.e., $u(t) = 1$. Finally, for $|p(t)| \leq 1$ the value $u(t) = p(t)$ is admissible, which means it turns out to be the solution of the maximum condition. As a result, we get

$$u(t) = \begin{cases} -1, & \text{if } p(t) < -1, \\ p(t), & \text{if } -1 \leq p(t) \leq 1, \\ 1, & \text{if } p(t) > 1. \end{cases} \tag{3.15}$$

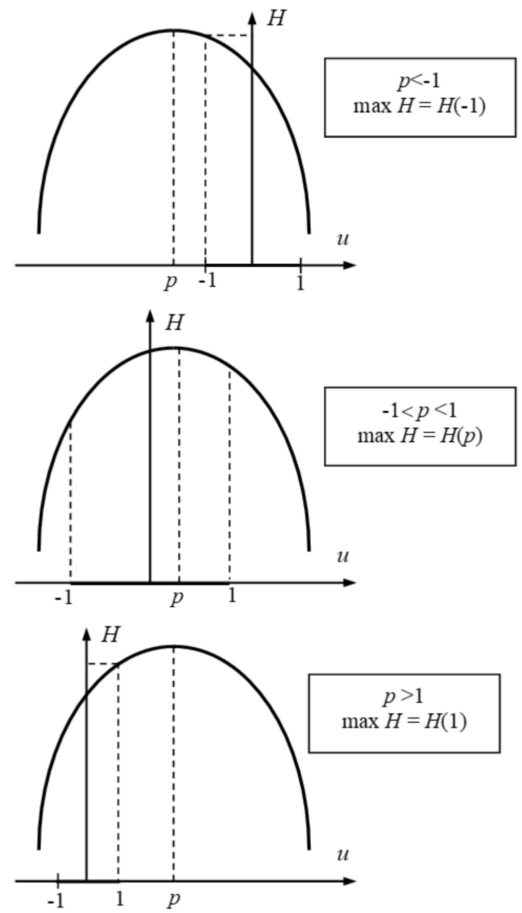

Figure 3.3 Conditional maximum of the function H.

Thus, for three unknown functions u, x, and p, we have obtained a problem that includes the equation of state with initial condition, the adjoint system, and the formula for finding a control. At the same time, unlike Example 3.1, each of these tasks connects two unknown functions, which does not allow finding any of them regardless of the others[35].

Formula (3.15) gives a solution to the maximum condition, so we can find the control if the function p is known; see Figure 3.4. Substituting this value into relation (3.13), we obtain a system (3.13) and (3.14) of two differential equations with two boundary conditions for the unknown functions x and p. Unfortunately, the connection between the functions u and p according to condition 3.15) turns out to be essentially non-linear. As a result, we are unable to find an analytical solution of the resulting system. However, it remains possible to find its approximate solution.

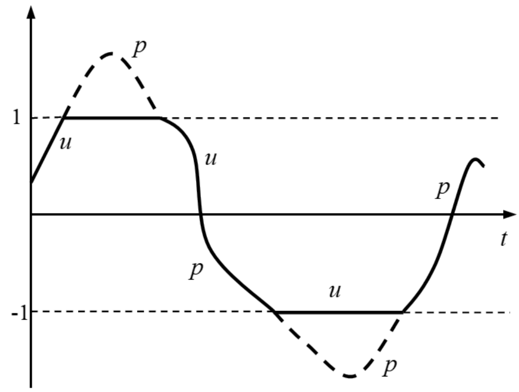

Figure 3.4 Solution of the maximum principle with the given function p.

If the boundary conditions for the functions x and p were given at the same end of the time interval, then the approximate solution of the Cauchy problem for a system of two (even if non-linear) differential equations would not cause any difficulties[36]. However, in reality, the additional condition for the function x is given at the initial time, and the function p has the condition at the end time. In this case, the existing system can be solved exclusively iteratively[37].

Choose an initial approximation control[38] u_0. Suppose that at the kth iteration we know the corresponding approximation u_k. Substituting this value into the equations of state (3.13), we find the function x_k, solving the problem

$$x_k'(t) = u_k(t),\ t \in (0,1);\ x_k(0) = 0.$$

Then, we find the function p_k from the adjoint system

$$p_k'(t) = x_k(t),\ t \in (0,1);\ p_k(1) = 0.$$

Finally, we find the new iteration u_{k+1} be the formula

$$u_{k+1}(t) = \begin{cases} -1, & \text{if } p_k(t) < -1, \\ p_k(t), & \text{if } -1 \le p_k(t) \le 1, \\ 1, & \text{if } p_k(t) > 1. \end{cases}$$

Thus, the solution of the system of optimality conditions can be defined as the limit of the sequence $\{u_k\}$, if, of course, it exists[39].

RESULTS

Here is a list of questions devoted to the simplest problems of optimal control and the optimality condition in the form of the maximum principle, the main conclusions, as well as additional problems that arise in this case, partially solved in Appendix, partially taken out in Notes.

Questions

It is required to answer questions concerning the properties of the optimal control problem and the optimality condition in the form of the maximum principle.

1. What does the optimal control problem include?
2. What is the fundamental difference between the function minimization problem and the optimal control problem?
3. Why, when posing the optimal control problem, did we separately include the term characterizing the final state of the system in the formula for the optimality criterion, and did not include the term characterizing its initial state?
4. Can the optimality criterion not explicitly depend on the control?
5. Can the optimality criterion not depend explicitly on the state function of the system?
6. Can there be no integral term in the formula for the optimality criterion?
7. Can only the integral term be present in the formula for the optimality criterion?
8. Is it possible to select the initial state of the system as control?
9. Why was it necessary to introduce an additional unknown function when deriving the maximum condition?
10. Why does the Lagrange functional combine all terms that explicitly depend on control?
11. In the formula for the increment of the functional (3.7), the function p is arbitrary. Why is the solution of problem (3.8) and (3.9) chosen as p in what follows?
12. Can an adjoint system be characterized by a non-linear differential equation?
13. Why is an additional condition specified for the equation of state at the initial moment of time, and for the adjoint equation we have the final condition?
14. Why and in what sense does the needle variation turn out to be sufficiently close to the optimal control?

15. Needle control variation is a discontinuous function. How justified is the choice of discontinuous control, both from a theoretical point of view (state equation must make sense) and from a practical point of view (control must be technically feasible)?

16. What stage of the proof of Theorem 3.1 actually remained unfounded?

17. What class of problems does the maximum condition belong to?

18. The maximum principle does not solve the stated problem of optimal control, but reduces it to another extremal problem. What then is the meaning of these transformations?

19. Is it possible to use the maximum principle for the optimal control problem in the absence of restrictions on the control?

20. What will the optimality conditions look like if in Problem 3.1 it is required to find not the minimum, but the maximum of the given optimality criterion?

21. How are the methods of minimization and maximization of the same functional related?

22. Why is the optimal control for Example 3.1 turned out to be continuous, and the solution for Example 3.2 can be discontinuous?

23. Why is it not possible to use non-iterative methods for the approximate solution of ordinary differential equations for a system of differential equations that includes an equation of state and an adjoint equation?

24. Why was it possible to solve the optimal control problems for Examples 3.1 and 3.2 without using iterative methods, while in the case of Example 3.3 the solution is found iteratively?

25. To what class of function does the solution of the maximum condition (3.15) for Example 3.3 belong in the general case?

26. Why is it precisely the control that is specified as the initial approximation, and not the state function and the solution of the adjoint system, although the system is being solved with respect to three unknown functions?

27. What type of problem has to be solved at each step of the iterative process in the practical solution of the system of optimality conditions?

28. Is it possible, in the case of convergence of the iterative considered process, to conclude that the corresponding limit is a solution to the optimal control problem?

Conclusions

Based on the study of the general problem of optimal control using the maximum principle and the results of the analysis of specific examples, we can draw the following conclusions.

- The optimal control problem consists in minimizing some functional on a set of admissible controls, and the functional depends on the control both explicitly and implicitly, by means of a state function that satisfies an equation that includes the control.

- The simplest optimal control problem is considered for the case when the system is described by the Cauchy problem for an ordinary differential equation, the optimality criterion is an integral functional, and the set of admissible controls is characterized by a given range of change in control values at each moment of time.

- To solve the optimal control problem, one can use the necessary optimality condition in the form of the maximum principle.

- The maximum principle is based on the problem of maximizing a some function on the set of admissible control, and this function also depends on the state function of the system and the solution of the adjoint system.

- The adjoint system is a linear differential equation with a condition at the final time.

- In the simplest case, the system of optimality conditions can be solved analytically.

- To solve the problem of maximizing a functional, instead of finding the maximum of the function H from the corresponding minimization problem, it is required to minimize the same function.

- As a rule, the system of optimality conditions can be solved approximately, and additional difficulty is due to the fact that the boundary conditions for the state equation and the adjoint equation are set at different ends of a given time interval.

- The approximate solution of the optimality conditions is carried out iteratively, and at each iteration, the solution of the state equation (Cauchy problem), the solution of the adjoint system (Cauchy problem in the reverse direction of time), and the control at the next iteration (problem for the conditional maximum of the function) are successively found.

Problems

Based on the results obtained above, we have extra the following problems.

1. **Substatiation**. In fact, we have not given a complete proof of Theorem 3.1, which is connected with the transformation of inequality (3.11). The justification of the maximum principle is given in Appendix; see also Notes[40].

2. **Generalization of the state equation**. We limited ourselves to consideration of systems described by unique ordinary differential equation. In Appendix, these results are extended to systems of differential equations. Regarding optimal control problems for a wider class of state equations; see Notes[41].

3. **Generalization of the state constraints**. The extension of the results to the case when the control is a vector quantity is trivial. All formulas remain valid, but the maximum principle is a problem of maximizing a function of many variables. Next, we considered the case where restrictions are imposed exclusively on control values. Part III will consider systems with a fixed final state, i.e., only those controls are admissible that not only belong to a given set, but also transfer systems to a given final state. In Part IV, we study the case of isoperimetric constraints given by some integral equalities, which include both the control and the state function of the system. Regarding optimal control problems with general constraints, see Notes[42].

4. **Optimal control for singular systems**. Problem 3.1 was solved under the assumption that the considered system is regular in the sense that, for any admissible control, the equation of state has a unique solution that continuously depends on the control. However, in practice it is not always possible to justify the regularity of the system. Moreover, the system can, in principle, turn out to be singular. Under these conditions, the applied research method is not effective. The problem of optimal control of one singular system is given in Chapter 4, see also Notes[43].

5. **Transformation of the optimality conditions**. The optimality condition is a system of equations in three unknowns. In this regard, one can try to analyze this system using the method of eliminating unknowns. This idea is implemented in Appendix for Example 3.3.

6. **Maximisation problem for the functional from Example 3.3**. Examples 3.1 and 3.2 are related to the study of the same functional, which is minimized in the first case and maximized in the second. Although these problems were studied by the same method, the solution of the first of them turned out to be a continuous, but not differentiable function, and the solution of the second one turned out to be a discontinuous function. Of interest is the problem of finding the maximum of the functional from Example 3.3. It is addressed in Chapter 5.

7. **Solving of maximum condition**. In the considered examples, from the maximum condition, we quite easily found the dependence of the control on the

state function and the solution of the adjoint system. However, in some cases fundamental difficulties arise at this step. Examples of such tasks are discussed in Chapter 6.

8. **Solvability**. In the considered examples, the optimal control exists. However, as noted in Part I, even the much simpler problem of minimizing a function of one variable is not always solvable. Examples of unsolvable optimal control problems and sufficient conditions for the existence of an optimal control are given in Chapter 7.

9. **Uniqueness**. In the considered examples, the optimal control was unique. However, as we know, already in the problem of minimizing a function of one variable, the solution may not be unique. Examples of optimization problems with non-unique solutions and sufficient conditions for the uniqueness of optimal control are given in Chapters 5 and 6.

10. **Sufficiency**. In the considered examples, the optimality condition turned out to be necessary and sufficient. However, as was shown earlier, in the process of solving the necessary condition for the extremum of a function, values that are not minimum points can be found. Chapters 5 and 6 will give examples of problems for which the maximum principle is not a sufficient optimality condition.

11. **Direct solving of the optimality conditions**. In the process of studying Examples 3.1 and 3.2, we managed to find solutions to the problem without using an iterative process. This was explained by the fact that in the case of linearity of both the state equation and the optimality criterion with respect to the state function, the adjoint system did not include the state function and could be solved independently of the state equation. At the same time, for Example 3.3, the functional was quadratic with respect to the state function, and an iterative method had to be used to solve the optimality conditions. It will be shown in Appendix that, in the absence of restrictions on controls for a linear system with a quadratic functional, it is also possible to find a solution to the problem without resorting to the iterative process.

12. **Convergence**. In the analysis of Example 3.3, an iterative algorithm was used. In Appendix, its convergence will be proved. Chapter 7 describes the optimal control problem, for which a similar algorithm diverges for any initial iteration.

13. **General iterative methods**. In the previous chapter, different iterative processes were used to solve problems of minimizing a function on the entire real line and on a segment. Their analog for the optimal control problem will be considered in Chapter 4.

14. **Problems with parameters**. In Chapter 2, the problem of minimizing a function depending on a parameter was considered. Similar questions for the simplest optimal control problem will be considered in Chapter 8.

15. **Variational inequalities**. In Chapter 2, variational inequalities were used to solve the problem on the conditional extremum of a function. In Chapter 4, we will show that this form of the optimality condition is also applicable to optimal control problems.

16. **Penalty method**. In Chapter 2, for the problem of minimizing a function in the presence of equality-type constraints, along with the method of Lagrange multipliers, the penalty method was used. In Chapter 4, it will be used to analyze optimal control problems.

17. **Sufficient optimality conditions**. The maximum principle is a necessary condition for optimality. Of interest is also sufficient optimality conditions, i.e., such relations, the solutions of which (if, of course, they exist) will surely be optimal controls. One such statement is given in Chapter 4.

3.2 APPENDIX

Below, we present some additional results related to the analysis of the simplest optimal control problems. Theorem 3.1 given in the Lecture was not fully substantiated. Section 3.2.1 provides its proof. The optimality conditions obtained earlier are a system of equations for the control, the state function, and the solution of the adjoint system. In Section 3.2.2, to analyze this system in relation to Example 3.3, the elimination method is applied. Earlier, for Examples 3.1 and 3.2, an analytical solution of the problem was found. Section 3.2.3 describes a class of problems that can also be solved without using an iterative process. In the study of Example 3.3, on the contrary, an iterative solution method was used. Section 3.2.4 proves its convergence. The final subsection is devoted to the vector analog of the general optimal control problem.

3.2.1 Existence of function minimum

Theorem 3.1 presented in the lecture cannot be considered strictly substantiated. In particular, we have not substantiated the necessary properties of the remainder term η in the formula for the increment of the functional, leading to inequality (3.10). We had

$$-\int_0^T \Delta_u H dt + \eta \geq 0 \; \forall v \in U. \tag{3.16}$$

This characterizes the increment of the minimized functional, and the remainder term was defined in the Lecture.

From equalities (3.1) and (3.2), it follows that the function Δx satisfies the differential equation

$$\Delta x'(t) = f(t, v(t), x(t) + \Delta x(t)) - f(t, u(t), x(t)), \; t \in (0, T)$$

with homogeneous initial condition. Integrating this equality, we get

$$\Delta x(t) = \int_0^t \Big[f(\xi, v(\xi), x(\xi) + \Delta x(\xi)) - f(\xi, u(\xi), x(\xi)) \Big] d\xi.$$

The following inequality holds

$$|\Delta x(t)| \le \int_0^t \Big| f(\xi, v(\xi), x(\xi) + \Delta x(\xi)) - f(\xi, u(\xi), x(\xi)) \Big| d\xi, \ t \in (0, T).$$

Assume that the function f satisfies the ***Lipschitz condition***, i.e., there exists a constant $L > 0$ such that[44]

$$\big| f(t, v_1, y_1) - f(t, v_2, y_2) \le L\big(|v_1 - v_2| + |y_1 - y_2|\big)$$

for all v_1, v_2, y_1, y_2, and $t \in (0, T)$. Then from the previous inequality we obtain

$$|\Delta x(t)| \le L \int_0^t \Delta x(\xi) d\xi + L \int_0^t |v(\xi) - u(\xi)| d\xi, \ t \in (0, T). \tag{3.17}$$

Further transformations require the following assertion named ***Gronwall lemma***[45]

Lemma 3.1 *If a continuous function $\varphi = \varphi(t)$ satisfies the condition*

$$|\varphi(t)| \le \alpha \int_0^t \varphi(\xi) d\xi + \beta, \ t \in (0, T),$$

where α, β are positive constants, then the following inequality holds

$$|\varphi(t)| \le \beta e^{\alpha t}, \ t \in (0, T).$$

Using the Gronwall lemma from inequality (3.17), we obtain the following estimate

$$|\Delta x(t)| \le c \int_0^t |v(\xi) - u(\xi)| d\xi, \ t \in (0, T).$$

where c is a positive constant. When deriving the maximum condition in Section 3.1.2, the needle variation of the control was chosen as the function v

$$v_{\varepsilon\tau}^w(t) = \begin{cases} u(t), & \text{if } t \notin (\tau - \varepsilon, \tau + \varepsilon), \\ w(t), & \text{if } t \in (\tau - \varepsilon, \tau + \varepsilon), \end{cases}$$

where τ is an arbitrary point of the interval $(0,T)$, w is an arbitrary admissible control, and ε is a small enough positive number. Then, the previous inequality takes the form

$$|\Delta x(t)| \leq c \int\limits_{\tau-\varepsilon}^{\tau+\varepsilon} |w(\xi) - u(\xi)| d\xi, \ t \in (0,T).$$

According to the mean value theorem used in the Lecture, for any bounded integrable function φ on the interval $[a,b]$ there exists a positive constant δ such that

$$\int\limits_a^b \varphi(\xi) d\xi = \delta(b-a).$$

Then the previous inequality implies the estimate

$$|\Delta x(t)| \leq c_1 \varepsilon, \ t \in (0,T), \tag{3.18}$$

where c_1 is a positive constant.

Now it seems possible to analyze the remainder term η, which is determined by the formula

$$\eta = \eta_3 - \int\limits_0^T (\eta_1 + \eta_2) dt,$$

where η_3 is higher order value than $\Delta x(T)$, η_1 higher order value than Δx, and $\eta_2 = [H_x(t,v,x,p) - H_x(t,u,x,p)]\Delta x$. From the inequality (3.18) it follows, that η_1 and η_3 have the higher order of smallness than ε. Using the inequality (3.18), we get

$$\Big| \int\limits_0^T \eta_2 dt \Big| \leq \int\limits_0^T |H_x(t,v,x,p) - H_x(t,u,x,p)| |\Delta x(t)| dt \leq$$

$$c_1 \varepsilon \int\limits_{\tau-\varepsilon}^{\tau+\varepsilon} |H_x(t,w,x,p) - H_x(t,u,x,p)|.$$

The integral on the right side of the obtained formula itself has the order ε due to the existing integration interval. Thus, the value of η_2, and hence the remainder term η has an order of smallness higher than ε.

Let us now divide the inequality (3.18) with $v = v_{\varepsilon\tau}^w$ by 2ε and pass in it to the limit as $\varepsilon \to 0$, taking into account the mean value theorem. We get

$$H\big[\tau, w(\tau), x(\tau), p(\tau)\big] - H\big[\tau, u(\tau), x(\tau), p(\tau)\big] \leq 0,$$

whence, due to the arbitrariness of the point $\tau \in [0,T]$ and $w(\tau) \in [a(\tau), b(\tau)]$, the maximum condition follows

$$H\big[t, u(t), x(t), p(t)\big] = \max_{v \in [a(t), b(t)]} H\big[t, v, x(t), p(t)\big].$$

Thus, the maximum principle can be considered justified.

3.2.2 Elimination method

When studying Example 3.2, the system of optimality conditions (3.13)–(3.15) with three unknown functions u, x, and p was obtained. The most natural way to solve a system of equations is based on the sequential elimination of unknowns from the system, which corresponds to the ***elimination method***[46]. Let us transform the resulting system by excluding two unknown functions from it.

Formula (3.15), which follows from the maximum principle, can be written as $u(t) = F(p(t))$, where $F(p)$ denotes the value on the right side of equality (3.15). Substituting it into the right side of the equation of state (3.13), we have

$$x'(t) = u(t) = F(p(t)).$$

Then, differentiating the adjoint equation (3.14), we obtain

$$p''(t) = x'(t) = F(p(t)).$$

Now the function p turns out to be a solution to the following ***boundary value problem*** for a second-order non-linear differential equation

$$p''(t) = F(p(t)),\ t \in (0,1),\ p(1) = 0,\ p'(0) = 0. \tag{3.19}$$

Thus, the system of optimality conditions can be reduced to the boundary value problem (3.19). In view of the equivalence of this problem to the system of optimality conditions (3.13)–(3.15), we conclude that this problem has a unique solution, which is a function equal to zero.

Consider now the non-linear ***heat equation***

$$\frac{\partial y(\tau,\xi)}{\partial \tau} = \frac{\partial^2 y(\tau,\xi)}{\partial \xi^2} - F(y(\tau,\xi)),\ \tau > 0,\ 0 < \xi < 1 \tag{3.20}$$

with boundary conditions

$$\frac{\partial y(\tau,0)}{\partial \xi} = 0,\ y(\tau,1) = 0,\ \tau > 0 \tag{3.21}$$

and some initial condition, where the function F has the same form as in problem (3.19). We know, that with an unlimited increase in time τ, the solution of an equation of this type can go into an ***equilibrium position***. This is such a state that if the system got there, then it will not leave it[47]. To find the equilibrium position, it is enough to equate the derivative of y with respect to τ to zero. Obviously, the equilibrium position $z = z(\xi)$ for system (3.20), (3.21) is a solution to the boundary value problem

$$z''(\xi) = F(z(\xi)),\ \xi \in (0,1),\ z(1) = 0,\ z'(0) = 0,$$

which, up to notation, coincides with problem (3.19). Then, based on the results obtained earlier, we conclude that the system characterized by equation (3.20) with boundary conditions (3.21) has a unique equilibrium position that is a function identically equal to zero[48].

3.2.3 Decoupling method

In Examples 3.1 and 3.2, we found an explicit solution to the optimal control problem without resorting to an iterative process. This is explained by the fact that the equation of state and the optimality criterion were linear with respect to the state function. As a result, the adjoint system did not include other unknown functions, and its solution could be found independently of the control and the state function. At the same time, in Example 3.3, the optimality criterion was quadratic, and it was not possible to separate the adjoint equation from other components of the optimality conditions. Therefore, some iterative method was used to solve the problem. However, for linear systems with a quadratic functional, in the absence of a control constraint, one can do it without the use of an iterative algorithm by using the ***decoupling method***[49].

Consider the system described by the Cauchy problem

$$x'(t) \;=\; a(t)x(t) + b(t)u(t) + f(t),\; t \in (0,T);\; x(0) = x_0, \tag{3.22}$$

where the functions a, b, f, and the number x_0 are known. Determine the functional

$$I(u) \;=\; \frac{1}{2}\int_0^T \Big\{\alpha\big[x(t) - z(t)\big]^2 + \beta[u(t)]^2\Big\}dt,$$

where a function z and positive constants α, β are known, and x is the solution of the problem (3.22).

Problem 3.2 *The **linear-quadratic optimal control problem** with a free finite state consists in finding a function u that minimizes the functional I whose definition includes the function x, which is the solution of the Cauchy problem* (3.22) *for the given control u.*

In accordance with the method described earlier, we determine the function

$$H = p(ax + bu + f) - [\alpha(x - z)^2 + \beta u^2]/2.$$

Then the adjoint system takes the form

$$p'(t) \;=\; \alpha[x(t) - z(t)] - a(t)p(t),\; t \in (0,T);\; p(T) = 0. \tag{3.23}$$

Unlike the examples considered earlier, there are no restrictions on control here. Thus, the maximum condition implies an unconditional extremum of the function H. Equaling its control derivative to zero, we find

$$u(t) = \beta^{-1}b(t)p(t),\; t \in (0,T). \tag{3.24}$$

Since the corresponding second derivative is negative, we conclude that this equality does give the maximum point of the function H.

Thus, the system (3.22)–(3.24) with respect to three unknown functions u, x, and p, is obtained. From equality (3.24), we substitute the control value into the equations of state

$$x'(t) \;=\; \beta^{-1}b(t)^2p(t) + a(t)x(t) + f(t),\; t \in (0,T);\; x(0) = x_0. \tag{3.25}$$

We have a system of two linear differential equations with two boundary conditions (3.23) and (3.25). As a result, we can assume that the relationship between the functions x and p is also linear too. Then there are some functions r and q, which satisfy the equality[50]

$$p(t) = r(t)x(t) + q(t),\; t \in (0,T). \tag{3.26}$$

To find these functions, we substitute the function p from equality (3.26) into the adjoint equation. We get

$$r'x + rx' + q' \;=\; \alpha(x - z) - a(rx + q).$$

Taking into account the first equality (3.24), we have

$$r'x + r\big(\beta^{-1}b^2p + ax + f\big) + q' \;=\; \alpha(x - z) - a(rx + q).$$

Substituting here the value of the function p from equality (3.22), we obtain

$$\big(r' + \beta^{-1}b^2r^2 + 2ar - \alpha\big)x \;+\; \big(q' + \beta^{-1}b^2rq + aq + fr + \alpha z\big) \;=\; 0.$$

Equalling in equality (3.25) $t = T$ and taking into account the boundary condition for the adjoint system, we have

$$r(T)x(T) + q(T) = 0.$$

The last two formulas represent equalities to zero of some linear functions with respect to the state function, respectively, at an arbitrary and final time. These equalities can be satisfied if both the coefficients in front of x and the free terms in them vanish. As a result, we determine the following problems with respect to the functions r and q:

$$r'(t) + \beta^{-1}b(t)^2r(t)^2 + 2a(t)r(t) \;=\; \alpha,\; t \in (0,T)\; r(T) = 0. \tag{3.27}$$

$$q'(t) + \beta^{-1}b(t)^2r(t)q(t) + a(t)q(t) + f(t)r(t) + \alpha z(t) \;=\; 0\; t \in (0,T)\; q(T) = 0. \tag{3.28}$$

Based on the results obtained, we have the following algorithm for solving the considered problem:

1. The Cauchy problem (3.27) is solved in the backward direction of time for the ordinary differential equation with a quadratic non-linearity, called the ***Riccati equation***[51] with respect to the function r.

2. The Cauchy problem (3.28) is solved in the backward direction of time for the linear ordinary differential equation with respect to the function q.

3. Based on formulas (3.24) and (3.26), an explicit dependence of the control on the system state function is determined[52]

$$u(t) = \beta^{-1}b(t)[r(t)x(t) + q(t)],\ t \in (0,T) \tag{3.29}$$

with known functions r and q.

4. After substituting the control from formula (3.29) into problem (3.22), the state function x is found.

5. By formula (3.29), the solution of the problem is calculated.

Thus, a complete analysis of Problem 3.2 can be carried out explicitly without using an iterative process. Consider one particular case.

Example 3.4 *It is necessary minimize the functional*

$$I(u) = \int_0^1 (u^2 + x^2)dt,$$

where x is the solution of the problem

$$x'(t) = u,\ t \in (0,T); x(0) = 0.$$

The difference from Example 3.3 here is solely in the absence of control restrictions. We have Problem 3.2 with the following parameter values:

$$a = 0,\ b = 1,\ f = 0,\ x_0 = 0,\ z = 0,\ \alpha = 1,\ \beta = 1,\ T = 1.$$

In accordance with the described method, we define the problem (3.28)

$$r'(t) + r(t)^2 = 1,\ t \in (0,1);\ r(1) = 0.$$

It has the solution

$$r(t) = \frac{e^{2t-2} - 1}{e^{2t-2} + 1}.$$

The problem (3.28) take the form

$$q'(t) + r(t)q(t) = 0,\ t \in (0,1);\ q(1) = 0.$$

It has zero solution. Then the dependence of the control on the state function is determined by the formula $u(t) = r(t)x(t)$ with the function r found above. After substituting this control into the equation of state, we again obtain a linear homogeneous equation with a homogeneous initial condition

$$x'(t) + r(t)x(t) = 0,\ t \in (0,1);\ x(0) = 0.$$

It has a zero solution. Taking into account the previously obtained formula for control, we conclude that $u(t) = 0$.

In reality, the optimality criterion takes exclusively non-negative values. Equality to zero here is possible only for $u(t) = 0$. Thus, we actually found the only solution to the problem[53].

3.2.4 Algorithm Convergence for Example 3.3

To solve the optimal control problem described in Example 3.3, an iterative process was used, characterized by the equalities

$$x_k'(t) = u(t),\ t \in (0,1),\ x_k(0) = 0;$$

$$p_k'(t) = x_k(t),\ t \in (0,1),\ p_k(1) = 0,$$

$$u_{k+1}(t) = \begin{cases} -1, & \text{if } p_k(t) < -1, \\ p_k(t), & \text{if } -1 \le p_k(t) \le 1, \\ 1, & \text{if } p_k(t) > 1. \end{cases}$$

Prove its convergence.

Choose some initial approximation u_0. It must certainly be an element of the set of admissible controls, and therefore satisfy the inequality

$$-1 \le u_0(t) \le 1,\ t \in [0,1].$$

Integrating the last relation from zero to an arbitrary value of t and using the equation of state, we have

$$-t \le x_0(t) = \int_0^t u_0(t)dt \le t,\ t \in [0,1].$$

As a result of integrating the resulting inequality from some value of t to unity, we obtain

$$-\frac{1}{2} \le \frac{t^2-1}{2} \le p_0(t) = -\int_t^1 x_0(t)dt \le \frac{1-t^2}{2} \le \frac{1}{2},\ t \in [0,1].$$

Since the values of p_0 do not go beyond the specified interval [–1,1], in accordance with the above formula, we find a new control approximation $u_1(t) = p_0(t)$. In this case, the following inequality holds

$$-1/2 \le u_1(t) \le 1/2,\ t \in [0,1].$$

Integrating the obtained relation, we have

$$-\frac{t}{2} \le x_1(t) = \int_0^t u_1(t)dt \le \frac{t}{2},\ t \in [0,1].$$

As a result of integrating this inequality from some value of t to 1, we obtain

$$-\frac{1}{4} \le \frac{t^2-1}{4} \le p_1(t) = -\int_t^1 x_1(t)dt \le \frac{1-t^2}{4} \le \frac{1}{4},\ t \in [0,1].$$

Then the new control approximation will satisfy the inequality

$$-1/4 \le u_2(t) \le 1/4,\ t \in [0,1].$$

Repeating the calculations above, at the next iteration we get

$$-1/8 \leq u_3(t) \leq 1/8,\ t \in [0,1].$$

In the general case, at the kth iteration, the estimate

$$|u_k(t)| \leq 2^{-k},\ t \in [0,1].$$

Thus, for $k \to \infty$ there is a convergence $u_k(t) \to 0$.

The obtained results show that for any initial approximation of the control chosen from the set U, the sequence $\{u_k\}$, determined in accordance with the method of successive approximations, converges to the function u^*, which is equal to zero.

A natural question arises: is the found value u^* be a solution to the optimal control problem? To answer this question, let us return to the formulation of the considered problem. Since the integrand in the functional to be minimized is non-negative, the inequality $I(u) \geq 0$ holds true for any admissible control u. The zero value of the functional can be achieved only when the equalities $u(t) = 0$ and $x(t) = 0$ are fulfilled for all $t \in [0,1]$. The zero value of the control is admissible, and it corresponds exactly to the value of the state function, which is identically equal to zero. Thus, the zero value of the functional is achieved exclusively on the admissible control u^*, and the negative values of the minimized functional are not realized. Consequently, the optimal control problem under consideration has a unique solution, which was found as a result of an approximate solution of the obtained optimality conditions. In this case, the rate of convergence of the algorithm is exponential[54].

3.2.5 Vector optimal control problem

The above results naturally extend to the vector case, when the control and the state function are vector functions

$$u = (u_1, u_2, ..., u_r),\ \ x = (x_1, x_2, ..., x_n).$$

In this case, the state equation

$$x'(t) = f(t, u(t), x(t)),\ t \in (0,T);\ x(0) = x_0$$

is the Cauchy problem for the system of differential equations, where f is nth order vector function of $r + n + 1$ variables, and x_0 is nth order vector. A vector control u belongs to the set

$$U = \{u | u(t) \in G(t); t \in (0,T)\},$$

where $G(t)$ is a subset of r-dimensional Euclidean space. The optimality criterion is determined by the formula

$$I(u) = \int_0^T g(t, u(t), x(t))dt + h((x(T),$$

where g is a function of $r + n + 1$ variables, h is a function of n variables. We have the following ***vector optimal control problem with free final state***[55].

Problem 3.3 *Find a vector function u from the set U that minimizes on this set the functional I whose definition includes the function x, which is the equation of state for the given function u.*

By analogy with the formula (3.3), the function $r+2n+1$ variables is determined

$$H(t,u,x,p) \;=\; \langle p, f(t,u,x)\rangle - g(t,u,x).$$

On the right side of this equality is the corresponding dot product. By the ***maximum principle***, the optimal control satisfies the condition

$$H\big[t,u(t),x(t),p(t)\big] \;=\; \max_{v\in G(t)} H\big[t,v,x(t),p(t)\big],\;\; t\in[0,T],$$

where p is a solution of the adjoint system

$$p'(t) = -H_x(t,u,x,p),\; t\in(0,T)\; p(T) = -h_x(x(T)).$$

The maximum condition here is a problem for the conditional extremum of the function H with respect to r variables, i.e., controls. The adjoint system includes n equations with corresponding boundary conditions. Here, the vectors H_x and h_x each include n components that are partial derivatives with respect to the variables x_i, $i = 1,\dots,n$.

We can also consider the vector analog of Problem 3.2, i.e., ***vector linear-quadratic optimal control problem with a free final state***.

Problem 3.4 *Find a point of minimum of the functional*

$$I(u) \;=\; \frac{1}{2}\int_0^T \Big[\alpha\big\|x(t)-z(t)\big\|^2 + \beta\big\|u(t)\big\|^2\Big]dt,$$

where the vector function x is a solution of the problem

$$x'(t) \;=\; A(t)x(t) + B(t)u(t) + f(t),\; t\in(0,T);\; x(0) = x_0,$$

besides, under the integral are the norms of the corresponding vectors, α and β are positive constants, z and f are vector functions, and x_0 is a n-order vector, A and B are matrix functions of orders $n\times n$ and $n\times r$.

For this problem, the maximum principle is used first (see above), and then, in accordance with the ***decoupling method***, we can try to find the function p by the formula

$$p(t) = R(t)x(t) + q(t),\; t\in(0,T),$$

where R is a matrix function of order $n\times n$, and q is an n-order vector. They satisfy, respectively, the system of differential equations with quadratic non-linearity of order $n\times n$ (***matrix Riccati equation***[56]) and the system of linear differential equations of order n.

Additional conclusions

Based on the results given in Appendix, we can draw some additional conclusions about the optimal control problem and the considered examples.

- Justification of the maximum principle is reduced to estimating the remainder term in the functional increment formula.
- Using the method of elimination of unknowns, the system of optimality conditions can be reduced to a boundary value problem for a non-linear second-order differential equation.
- For a system that is linear both in control and in the state of the system, in the absence of restrictions on control, the solution to the problem can be found without using an iterative process due to the linearity of the system of optimality conditions.
- The iterative process for solving the system of optimality conditions in Example 3.3 converges to the optimal control for any initial approximation, and the rate of convergence is exponential.
- Almost all the results obtained earlier can be extended to the vector optimal control problem.

Notes

1. General questions of optimal control theory are considered, for example, in [5], [13], [28], [42], [49], [44], [67], [70], [74], [93], [95], [123], [152], [162], [180], [193], [194], [195], [200], [208].

2. For mathematical modeling, see, for example, [30], [173], [203].

3. The general theory of dynamical systems is considered, for example, in [10], [83], [144].

4. The considered systems are also called ***evolutionary***. There are also ***stationary systems***, the state of which does not depend on time, but may depend on spatial coordinates. They can be described by elliptic partial differential equations.

5. For ***systems with lumped parameters***, the corresponding ***phase space*** (the set of possible values of the state function of the system) is ***finite-dimensional***, i.e., it is described by a finite set of numbers at any point in time. ***Systems with distributed parameters*** are also considered, for which the phase space is ***infinite-dimensional***, i.e., the corresponding state of the system at each moment of time cannot be characterized by a finite set of numbers. If systems with lumped parameters are most often associated with ***ordinary differential equations***, then systems with distributed parameters, as a rule, are associated with ***partial differential equations***.

6. The requirement for the uniqueness of the state function is introduced solely for reasons of simplicity. The results obtained naturally extend to the case when the state of the system is characterized by a vector function; see Appendix.

7. Naturally, the uniqueness assumption of the control function is not necessary; see Appendix. The control can be a number, a vector, or a vector function. For systems with distributed parameters, the control can be a function of many variables.

8. The existence and uniqueness of a solution of the considered Cauchy problem is established by means of the ***theory of differential equations***; see, for example, [10], [86]. In principle, optimal control problems are also studied in the absence of the requirement that the state equations be uniquely solvable; see [73], [118]. Chapter 4 considers optimal control problems for systems described by differential equations in the absence of existence or uniqueness of a solution.

9. Optimal control theory also considers problems with restrictions on the state function of the system, i.e., with ***phase constraints***. The simplest problems of this nature are considered in later parts of the book. Significantly more general optimal control problems with phase constraints are considered; for example, in [5], [56], [95], [140].

10. In fact, when specifying the set of admissible controls, it is required to indicate not only the restrictions that are imposed on the controls, but also the functional properties of these controls, i.e., function space to which they must all belong, including the boundary functions a and b. At this stage of the study, we do not need to specify this space, although the functional class of control largely determines the properties of the system state function and is used to justify the optimality condition. In Chapter 7, we will return to this question when we study the problem of the existence of an optimal control.

11. Optimal control problems with a fixed final state are considered in Part III. Functional minimization problems are also studied in the framework of the ***calculus of variations***; see [37], [61], [208]. However, this implies an explicit dependence of the functional on the unknown function. In optimal control problems, this dependence is implicitly specified by means of an equation of state, the solution of which depends on the control and affects the optimality criterion.

12. More precisely, in Chapter 2, when defining the Lagrange function, the minimized function was multiplied by another Lagrange multiplier. However, it was noted that the latter can be assumed to be equal to unity.

13. See in particular the final subsection of Chapter 2.

14. The resulting value can also be interpreted as a ***dot product*** in some function space. Actually, the sum is a discrete analog of the integral, and the integral is the continuous analog of the sum. Behind this is the same construction, namely integration over a ***measure***. The difference here is solely in the type of measure; see [158]. If the state of the system turns out to be a vector quantity, then so will the Lagrange multiplier. In this case, the definition of the Lagrange functional will use the integral of the sum of the product of the vector Lagrange multiplier and the vector of the right-hand sides of the equations of state, which again corresponds to the dot product of the corresponding quantities in the vector functional space.

15. The Lagrange multiplier method is used in the calculus of variations to minimize functionals when there are additional constraints in the form of equalities; see [37], [61], [208].

16. In the calculus of variations, the function H corresponds to the ***Hamiltonian***; see [37], [61], [208]. As applied to physical problems, the Hamiltonian has the sense of the total energy of the system; see [110].

17. We have the difference between the values of two functions on the controls v and u, multiplied by the increment Δx.

18. The form of the remainder term will be used in Chapter 5 when studying the sufficiency of the optimality condition in the form of the maximum principle.

19. In ***Hamiltonian mechanics***, closely related to the calculus of variations, the function (usually a vector function) p corresponds to the generalized momentum of the system; see [110].

20. Needle variation is used in the calculus of variations to derive the ***Weierstrass condition*** of a functional; see [37], [61], [208]. Other types of variations are considered, for example, in [75], where they are used they are used in deriving the necessary optimality conditions for singular controls; see Chapter 6.

21. We pay attention to the fact that the needle variation of the control is a discontinuous function. The choice of a discontinuous function as a control is justified, first of all, by the fact that differential equations with discontinuous parameters make sense; see [10], [86]. This choice also makes sense from a practical point of view, since discontinuous functions are technically realizable. The control system may include a certain switch, which, if necessary, at a certain point in time, transfers the system from one state to another.

22. The concept of proximity of functions is far from obvious. The general notion of proximity is related to ***topology***; see [101]. Various forms of closeness of functions are defined in functional analysis; see [94], [100], [106], [158]. If two continuous functions u and v defined on the interval $[0, T]$ are considered, then they can be considered sufficiently close if the maximum value of the modulus of the difference $|u(t)–v(t)|$ over all values of t from the specified interval is sufficiently small. This quantity is called the ***norm*** of the difference of the considered functions in the space of continuous functions $C[0, T]$. Naturally, in this sense, the optimal control and its needle variation will not be close, unless the functions u and w take different values at the point τ. However, in optimal control problems, continuity of the control is usually not required, and the closeness of the considered functions is understood in the sense of the smallness of the Lebesgue integral of the difference $|u(t)–v(t)|$ (or from the power p of this value, greater than one) over all values of t from the interval [0,T]. This corresponds to the norm in the space $L_1(0, T)$ (respectively, in the space $L_p(0, T)$). It is easy to see (this follows from the mean value theorem, see Note 23) that in this case the difference between the optimal control and its needle variation is of the order of smallness ε. In this sense, in the calculus of variations, strong and weak local extrema of functionals are distinguished. A function $x = x(t)$ is called a ***weak minimum*** of a given functional if its value in it does not exceed the value of the functional at all points $y = y(t)$ sufficiently close to x in the sense of being small as the values $|x(t)–y(t)|$ and $|x'(t)–y'(t)|$. At the same time, for a ***strong minimum***, the corresponding relation holds only if $|x(t)–y(t)|$ is small. Obviously, a strong extremum is always weak, but the converse statement is not always realized. The needle variation is sufficiently close to the optimal control only in the sense of the closeness of the functions, but not their derivatives.

23. Here, we use the assumption that for any function u problem (3.1) and (3.2) has a unique solution that depends continuously on the control. This corresponds to the Hadamard well-posedness of the problem, defined in Chapter 2. This property, under certain restrictions, is established by means of the theory of ordinary differential equations; see [10], [86].

24. By the ***mean value theorem***, the integral of some function is equal to the value of the function at some point in the integration interval, multiplied by the length of this interval; see, for example, [158]. In this case, the integral of the corresponding function is calculated from the point $\tau - \varepsilon$ to the point $\tau + \varepsilon$. The length of this interval is 2ε. Therefore, after dividing by 2ε, we obtain the value of the integrand at some point from the segment $[\tau - \varepsilon, \tau + \varepsilon]$. After passing to the limit at $\varepsilon \to 0$, we obtain the value of this function at the point τ.

25. The concept of the maximum principle is also used in the theory of functions of a complex variable and in problems of mathematical physics. However, there it has a completely different meaning.

26. In the calculus of variations, an analog of the maximum principle is the ***Weierstrass condition***, which gives a necessary condition for a strong minimum of the corresponding functional; see [37], [61].

27. In fact, for each value of t, we have a conditional extremum problem for a function of one variable, i.e., Problem 2.1.

28. Another price to pay for the transition from functional to function is the absence of sufficiency of the optimality condition; see Chapter 5. However, we encountered a similar phenomenon when minimizing the function in Chapter 1.

29. In Chapter 5, we will see that the maximum principle for Example 3.1 is a necessary and sufficient condition for optimality. This implies that the only solution of the maximum principle is the optimal control. Moreover, in Chapter 7 we will show that the considered optimal control problem is solvable, which means that the optimality condition certainly has a solution. Finally, in Chapter 9 we will show that the resulting optimal control also minimizes this functional on the subset U of functions that ensure the transfer of the system to the state $x(1) = 3/2$.

30. This statement is justified in the same way as the similar result for Example 3.1; see also Chapters 5 and 7.

31. In Chapter 9, optimal control problems with a fixed final state will be considered, for which the exact solution will also be found using the maximum principle without involving any iterative methods; see Examples 9.1 and 9.2. There, as for Examples 3.1 and 3.2, the equation of state and the optimality criterion are linear in the state of the system. In Chapter 13, we will find an analytical solution to the problem of optimal control of the system with an additional integral constraint in the form of equality; see Example 13.1. Similar examples for optimal control problems in the absence of initial conditions are considered in Chapter 16.

32. Chapter 5 will show that the functional maximization problem from Example 3.3 has qualitatively different properties; see Example 5.1. The problem given in Example 7.3 from Chapter 7 also has completely different properties, differing from the given one only in the set of admissible controls. Chapter 12 will consider the minimization problem of the functional from Example 3.3 with an additional constraint, when the final state of the system is fixed.

33. Between any two closed finite segments, one can establish a one-to-one correspondence with the preservation of almost all properties that we take into account. Two such objects are called and, in principle, are indistinguishable from the standpoint of the subject area under consideration.

34. In Chapter 4, we will use the variational inequality and the penalty method to analyze this example.

35. This is due to the presence of a quadratic term with respect to the state function in the optimality criterion.

36. For numerical methods for solving ordinary differential equations; see, for example, [84].

37. This algorithm can also be used to solve the optimality conditions corresponding to Problem 3.1 of a general form. In this case, at each iteration, first, from the known control, the function x is found from the equations of state. Then the adjoint system is solved. Finally, a new control approximation is determined from the maximum condition. In Chapter 9 this algorithm will be extended to optimal control problems with a fixed final state, and in Section 13 to problems with isoperimetric constraints. For iterative methods for solving optimality conditions; see, for example, [46], [65], [70], [149].

38. In principle, any of them can be set as the initial approximation. However, only about the control we have initially defined information, namely, its belonging to a given set. In this regard, it is natural to begin calculations with the control. In the next section, we will show that in a specific case, the formula for solving the optimality condition allows us to refine the choice of the initial approximation, which increases the efficiency of the iterative process.

39. Naturally, we must also prove the existence of the limit. In addition, we have to make sure that this limit is not just a solution to the system of optimality conditions, but also turns out to be an optimal control for the considered problem. The corresponding results are given in Appendix. In the next chapter it will be shown that the solution of this problem is unique, and the maximum principle is a necessary and sufficient condition for optimality. From here it will also follow that the limit of the considered sequence is an optimal control.

40. Justification of the maximum principle for systems described by differential equations in the vector case is given in Appendix; see also [5], [42], [62], [70], [74], [95], [103], [152], [162], [182], [193], [195].

41. The obtained results are naturally extended to the case when the state function is a vector quantity, and we have a system of differential equations. In this case, the formula for the function H has the same form as before $H(t,u,x,p) = pf - g$, but here x is the vector of system states (vector function), f is the vector of the right parts of the system of differential equations, p is the adjoint state vector, and pf has the meaning of the dot product, i.e., the sum of products of vector components. In this case, p is determined from the adjoint system, again determined by the equation $p' = -H_x$, where the derivative of the vector function is on the left side, and the vector H_x of partial derivatives of the function H. Optimal control problems for systems of differential equations are studied, for example, in [5], [42], [62], [70], [74], [152], [162], [193]; for partial differential equations, see [9], [24], [36], [59], [73], [104], [111], [116], [118], [119], [121], [124], [136], [137], [138], [155], [165], [171], [177], [178], [186], [185], for integral equations see [20], [50], [53], [164], [176]; for integro-differential equations [2], [135], [209]; for differential inclusions, see [126]; for the system described by variational inequalities see [22], [186], [198], [204]; for discrete dynamical system, see [28], [29], [38], [42], [70], [193], [194], [195]. Optimal control problems under conflict or uncertainty are studied using differential, see [42], [97]. On optimal control of stochastic systems, see [8], [31], [42], [62], [67], [70], [177].

42. Optimal control problems for systems with phase constraints are considered; for example, in [5], [42], [95], [140].

43. Optimal control problems for singular systems are considered in [73], [118].

44. It is easy to verify that the ***Lipschitz condition*** is satisfied by any continuously differentiable function. Thus, the Lipschitz condition is stronger than the continuity and weaker than the differentiability of the function.

45. The name Gronwall–Bellman lemma is also used. The proof of this assertion is given, for example, in [86], [143], [193].

46. The ***Gauss method*** for solving systems of linear algebraic equations is based on this idea; see [14]. The fact that in this case we are dealing with neither algebraic nor linear equations (the dependence of u on p is non-linear) does not play a fundamental role. We will use the method of eliminating unknowns in order to reduce the system of optimality conditions to a problem that includes only one unknown function in the subsequent sections of the book.

47. On the ***equilibrium position*** of dynamical systems; see, for example, [10], [86]. Similar questions for non-linear equations of parabolic type see [88].

48. On boundary value problems for non-linear differential equations; see [10], [61], [81], [86]. The transition to an equation of the heat conduction type is associated with the ***establishment method*** for solving the corresponding boundary value problem. In Chapters 5, 7, and 12, the solutions of the system of optimality conditions for various optimal control problems will also be interpreted as the equilibrium positions of some non-stationary systems. In Chapters 5 and 7, we will consider boundary value problems that differ from (3.19) only in the form of the function F. However, the properties of these problems differ radically from the considered one.

49. On the ***decoupling method*** for systems described by partial differential equations; see [116]. In Chapter 8, the linear-quadratic optimal control problem with a fixed finite state is studied in a similar way and in Chapter 16 this method is used for a problem with a free initial state

50. In the vector case, r is a matrix function, see Section 3.2.5, and in the case of partial differential equations, r is some operator; see [116].

51. On the ***Riccati equation***, see [59], [159], [210]. Extending these results to partial differential equations, at this stage, an operator equation is established, which reduces to the analysis of an integro-differential equation with a quadratic non-linearity; see [116]. On the numerical solution of the Riccati equation; see [149]. We will meet the Riccati equation again in Chapter 4 when using the Bellman equation to analyze a linear-quadratic optimal control problem and also in Chapters 9 and 16 when studying linear-quadratic optimal control problems with a fixed final state and a free initial state, respectively.

52. Defining the dependence of the optimal control on the state function is the subject of the ***synthesis problem*** of great importance in the ***automatic control theory***; see [4], [42], [59], [67], [113], [212]. In a program control problem, optimal control is defined as a function of independent variables (in this case, time). In practice, the control is given directly. The practical application of the synthesis problem is reduced to the consideration of a system with feedback, in which, based on the results of measuring the state of the system, control is determined, which, in turn, regulates the course of development of the controlled process. This approach seems to be especially relevant in the analysis of stochastic systems subject to random influences; see [4], [42], [59], [67], [212].

53. An iterative method for solving the optimal problem from Example 3.4 is described in Chapter 4, and the existence of its solution is proved in Chapter 7.

54. Although the optimal control for Example 3.3 has already been found, we will repeatedly return to its study in subsequent chapters in order to illustrate the features of certain results obtained in the future.

55. A vector optimal control problem with a fixed final state is considered in Chapter 9.

56. The matrix Riccati equation; see [59], [67], [159], [212].

CHAPTER 4

Alternative methods

In the previous chapter, we considered the simplest problem of optimal control theory. It was analyzed using optimality conditions in the form of the maximum principle. However, there are other methods for solving such problems. In particular, we note approximate methods for minimizing functionals, which are natural generalizations of the corresponding methods for minimizing functions, described in Chapter 2. To solve optimal control problems, we will also use variational inequalities and the penalty method, which were previously used to minimize functions under certain restrictions. In addition, sufficient optimality conditions are considered.

4.1 LECTURE

The subject of this lecture is again the problem of minimizing an integral functional for a system described by an ordinary differential equation with an initial condition in the presence of restrictions on the control values. In the previous chapter, the necessary optimality conditions in the form of the maximum principle were used to study it. However, optimization methods are not limited to this. First of all, we recall that for the function minimization problems, along with minimum conditions, some iterative methods were used. In particular, in Chapter 2, gradient methods for minimizing functions were described. We give a generalization of these methods to optimal control problems, confining ourselves to some specific examples; see Section 4.1.1. The variational inequality, which was also applied in Chapter 2 to solve the problem of minimizing a function on an interval, is fairly close to the maximum condition. We use it to analyze considered before Problem 3.1; see Section 4.1.2. Chapter 2 also used the penalty method to solve the function minimization problem in the presence of an additional constraint in the form of equality. In Section 4.1.3, it is applied to the approximate solving of optimal control problems. Finally, in Section 4.1.4, the Bellman equation is derived, which gives a sufficient optimality condition for Problem 3.1.

DOI: 10.1201/9781003398585-4

4.1.1 Iterative methods for solving an optimization problem

In Chapter 2, iterative methods were used to solve the problem of minimizing a differentiable function. Let us show that they can also be used to solve optimal control problems. We confine ourselves easy examples. Let us start with an Example 3.4. It is required to minimize the functional

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

In Chapter 3, its solution was found using the maximum principle and the decoupling method. In this case, we try to use iterative methods.

Let u be a solution of this problem. Consider the control $v = u + \sigma h$, where σ is a positive number, h is an arbitrary function[1]. The following inequality holds

$$I(v) - I(u) = \frac{1}{2}\int_0^1 \Big[(v^2 - u^2) + (y^2 - x^2)\Big]dt,$$

where y is the solution of the considered Cauchy problem corresponding to the control v. The last inequality can be written as

$$\sigma\int_0^1 uhdt + \int_0^1 x\Delta xdt + \frac{1}{2}\int_0^1 (\sigma^2 h^2 + \Delta x^2)dt \geq 0, \tag{4.1}$$

where $\Delta x = y - x$.

The function Δx is the solution of the problem

$$\Delta x'(t) = \sigma h(t),\ t \in (0,1);\ \Delta x(0) = 0.$$

Multiplying the first equality by an arbitrary differentiable function λ and integrating the result using the second equality, we get

$$\sigma\int_0^1 \lambda hdt = \int_0^1 \lambda\Delta x' dt = \lambda(1)\Delta x(1) - \int_0^1 \lambda'\Delta xdt.$$

Choose here as λ the solution of the problem[2]

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0.$$

We obtain

$$\sigma\int_0^1 phdt = -\int_0^1 x\Delta xdt.$$

Then inequality (4.1) takes the form

$$\sigma \int_0^1 (u-p)hdt + \frac{1}{2}\int_0^1 (\sigma^2 h^2 + \Delta x^2) \geq 0.$$

Obviously, the function Δx is

$$\Delta x(t) = \sigma \int_0^t h(\tau)d\tau,$$

i.e., proportional to σ. Then, dividing the previous inequality by σ and passing to the limit as $\sigma \to 0$, we obtain

$$\int_0^1 (u-p)hdt \geq 0.$$

The resulting relation is valid for any function h. Then we can replace h with $-h$ and get the inequality

$$\int_0^1 (u-p)hdt \leq 0.$$

Since both last conditions are true, we have the equality

$$\int_0^1 (u-p)hdt = 0$$

for all functions h.

In fact, we have established the formula

$$\lim_{\sigma\to 0} \frac{I(u+\sigma h) - I(u)}{\sigma} = \int_0^1 (u-p)hdt = 0 \ \forall h. \tag{4.2}$$

Thus, the corresponding limit[3] exists and is linear with respect to the function h. This result is related to some generalization of the concept of the function derivative.

Definition 4.1 *The value under the integral in equality* (4.2) *and multiplied by an arbitrary function h is called the* ***Gateaux derivative***[4] *of the functional I at the point u and is denoted by $I'(u)$.*

It is easy to verify that the Gateaux derivative of a differentiable function of one variable coincides with its usual derivative, and of a function of many variables with its gradient[5]. In our case, we have found the Gateaux derivative of the considered functional, i.e., we have $I'(u) = u - p$.

It follows[6] from the formula (4.2) that $u–p = 0$, and hence $u = p$. In fact, we have proved here that the Gateaux derivative $I'(u)$ of the considered functional at a given point, equal to $u–p$ vanishes if the control u is optimal. This corresponds to the ***stationary condition*** for our problem[7].

Now, we can use the ***method of successive approximations***, similar to that used in the analysis of Examples 2.6–2.11. Assume that at the kth iteration, the control approximation u_k is known. Substituting this value into the equation of state, we find the function x_k as a solution to the Cauchy problem

$$x_k'(t) = u_k(t),\ t \in (0,1);\ x_k(0) = 0.$$

Next, the function p_k is defined as a solution to the problem

$$p_k'(t) = x_k(t),\ t \in (0,1);\ p_k(1) = 0.$$

The new approximate control u_{k+1} is determined by the formula

$$u_{k+1}(t) = p_k(t),\ t \in (0,1). \tag{4.3}$$

Another version of the algorithm corresponds to the ***gradient method***[8]

$$u_{k+1} \ = \ u_k - \beta_k I'(u_k),$$

considered in Chapter 2, where the positive constant β_k is an algorithm parameter. For Example 3.4, the gradient method is characterized by the equality

$$u_{k+1}(t) \ = \ u_k - \beta_k p_k(t),\ t \in (0,1). \tag{4.4}$$

In Example 3.3, we considered the problem of minimizing the same functional on the set of such controls whose values at each point lie on the interval [–1,1]. In Chapter 2, to solve the problem of minimizing a function on a segment, the gradient projection method was used, according to which, at this step of the algorithm, the usual gradient method was first used. If the found value satisfies the given constraints, then it is chosen as a new approximation of the desired value. Otherwise, the closest point to the found value that satisfies these restrictions is chosen as such. In the general case, the ***gradient projection method*** is characterized by the equality

$$u_{k+1} \ = \ P\big[u_k - \beta_k I'(u_k)\big],$$

where P is an operator of projection onto the set of admissible controls, also called a ***projector***. The projector transforms an arbitrary point of the space under consideration into the point of the given set closest to it. Thus, for Example 3.3, formula (4.4) is replaced by the equality

$$u_{k+1}(t) \ = \begin{cases} -1, & \text{if } \ v_k(t) < -1, \\ v_k(t), & \text{if } \ -1 \le v_k(t) \le 1, \\ 1, & \text{if } \ v_k(t) > 1, \end{cases}$$

where $v_k = u_k – \beta_k p_k$. However, one can also use an analog of formula (4.3) that is

$$u_{k+1}(t) = \begin{cases} -1, & \text{if } v_k(t) < -1, \\ p_k(t), & \text{if } -1 \le v_k(t) \le 1, \\ 1, & \text{if } v_k(t) > 1, \end{cases}$$

This result exactly corresponds to what was obtained in Chapter 3 using the iterative method for solving the system of optimality conditions obtained by the maximum principle.

4.1.2 Variational inequality

In Chapter 2, a variational inequality was used to solve the problem of minimizing a function on an interval. Let us show that this extremum condition is also applicable[9] to the solution of Problem 3.1. It consists in minimizing the functional

$$I(u) = \int_0^T g(t, u(t), x(t))dt + h(x(T))$$

on the set of admissible controls

$$U = \Big\{u \big| \, a(t) \le u(t) \le b(t), \ t \in (0, T)\Big\},$$

where x is a solution of the Cauchy problem

$$x'(t) = f(t, u(t), x(t)), \ t \in (0, T); \ x(0) = x_0.$$

In the previous section, the following function was defined

$$H(t, u, x, p) = pf(t, u, x) – g(t, u, x),$$

where p is a solution of the adjoint system

$$p'(t) = -H_x(t, u(t), x(t), p(t)), \ t \in (0, T); \ p(T) = -h_x(x(T)).$$

In doing so, inequality (3.10) was determined that is

$$-\int_0^T \Delta_u H dt + \eta \ge 0,$$

with a remainder term η.

For any function $s \in U$, the control[10] $v_{\sigma s} = u + \sigma(s – u)$ is admissible for any number $\sigma \in (0, 1)$. We assume in the above inequality $v = v_{\sigma s}$. Note that for small enough values of σ, the control $v_{\sigma s}$ is be arbitrarily close to u. Assuming again the continuous dependence of the solution of the considered Cauchy problem on the control, we establish that for small σ the corresponding increment of the state function Δx is of the order of σ. Then the remainder term η has the order σ^2. Assuming the functions f and g to be differentiable with respect to control, after dividing the last inequality by σ and passing to the limit as $\sigma \to 0$, we have

$$\int_0^T H_u(t, u(t), x(t), p(t))(s - u)dt \le 0 \;\; \forall s \in U.$$

For further transformations of the obtained relation, we choose as the control s the needle variation of the control defined in Chapter 3

$$v^w_{\varepsilon\tau}(t) = \begin{cases} u(t), & \text{if } t \notin (\tau - \varepsilon, \tau + \varepsilon), \\ w(t), & \text{if } t \in (\tau - \varepsilon, \tau + \varepsilon) \end{cases}$$

with an arbitrary point τ from the interval $(0, T)$, small enough number ε, and an arbitrary control w. We get

$$\int_{\tau-\varepsilon}^{\tau+\varepsilon} H_u(t, u(t), x(t), p(t))[w(t) - u(t)]dt \le 0.$$

Dividing the resulting inequality by 2ε and passing in it to the limit as $\varepsilon \to 0$, similarly to how it was done in the justification of Theorem 3.1, we obtain

$$H_u(\tau, u(\tau), x(\tau), p(\tau))[w(\tau) - u(\tau)] \le 0.$$

Hence, due to the arbitrariness of the point τ and the control w, we get the ***variational inequality***[11]

$$H_u(t, u(t), x(t), p(t))[v - u(t)] \le 0 \;\; \forall v \in [a(t), b(t)], \; t \in (0, T). \tag{4.5}$$

We formulate the results obtained in the form of a theorem[12].

Theorem 4.1 *Under the assumptions made, in order for the control u to be optimal, it is necessary that it satisfies the variational inequality* (4.5).

Now, we check this theorem in a concrete situation. Let us turn to Example 3.1, which consists in minimizing the functional

$$I(u) = \int_0 \Big(\frac{u^2}{2} - 3x\Big)dx$$

on the control set

$$U = \Big\{u \,\big|\, 1 \le u(t) \le 2, \; 0 < t < 1\Big\},$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t), \; t \in (0, 1); \; x(0) = 0.$$

The function H is determined here by the formula

$$H(t, u, x, p) = pu - u^2/2 + 3x.$$

Then the variational inequality (4.5) takes the form

$$[p(t)–u(t)][v–u(t)] \leq 0 \ \ \forall v \in [1,2], \ t \in (0,1).$$

To solve it, we use the technique described in Chapter 2. On the left side of this inequality is the product of two quantities, the first of which is fixed, and the second varies. The first multiplier can have any sign. Consider all three possible cases.

Suppose $p(t)–u(t) > 0$. Then, dividing the inequality by the first multiplier, we get $v–u(t) \leq 0$ for all $v \in [1,2]$, i.e., $u(t)$ must be no less than all numbers from the interval [1,2]. However, the control value itself belongs to this segment. Then the preceding assertion can only be true for $u(t) = 2$. Therefore, if the value of $p(t)$ is greater than $u(t)$, equal to 2, then $u(t) = 2$. Thus, this equality is satisfied for $p(t) > 2$.

If $p(t)–u(t) < 0$, then after dividing the variational inequality by the first multiplier, we obtain that $v–u(t) \geq 0$ for all $v \in [1,2]$, i.e., $u(t)$ must not be greater than all numbers from the interval [1,2]. This is possible only for $u(t) = 1$. Thus, if $p(t)$ is less than $u(t)$, which is equal to 1, then $u(t) = 1$. Thus, this equality holds for $p(t) < 1$.

Finally, the equality $p(t)–u(t) = 0$ is possible, and hence $u(t) = p(t)$. However, this control is admissible only for $p(t) \in [1,2]$.

Previously, it was found that $p(t) = 3–3t$. Then the solution of the variational inequality (4.5) for the considered example leads to the formula

$$u(t) = \begin{cases} 2, & \text{if } 0 < t < 1/3, \\ 3 - 3t, & \text{if } 1/3 \leq t \leq 2/3, \\ 1, & \text{if } 2/3 < t < 1. \end{cases}$$

This result is exactly the same as what was obtained earlier using the maximum principle, and gives a solution to the maximum principle. Thus, the solution of the optimal control problem for Example 3.1 based on the variational inequality and the maximum principle leads to the same results.

For Example 3.3, the function H is equal to $pu–(u^2 + x^2)/2$. Then we obtain the variational inequality

$$[p(t)–u(t)]\,[v–u(t)] \leq 0 \ \ \forall v \in [-1,1]. \tag{4.6}$$

If $p(t)–u(t) > 0$, then, dividing inequality (4.6) by the first multiplier of its left side, we get $v–u(t) \leq 0$ for all $v \in [–1,1]$. Taking into account that the value $u(t)$ itself belongs to the segment [–1,1], from the last inequality, we deduce $u(t) = 1$. Thus, if $p(t)$ is greater than $u(t)$, equal to 1, then $u(t) = 1$. If $p(t)–u(t) < 0$, then, dividing inequality (4.6) by the first multiplier of its left side, we obtain $v–u(t) \geq 0$ for all $v \in [–1,1]$. This is possible only $u(t) = –1$. Thus, if $p(t)$ is less than $u(t)$, equal to –1, then $u(t) = –1$. Finally, inequality (4.6) is satisfied for $u(t) = p(t)$. However, this value is only true if $p(t) \in [–1,1]$. As a result, we have the formula

$$u(t) = \begin{cases} -1, & \text{if } p(t) < -1, \\ p(t), & \text{if } -1 \leq p(t) \leq 1, \\ 1, & \text{if } p(t) > 1 \end{cases}$$

determined in the previous chapter by the maximum principle. Thus, for Example 3.3, the variational inequality and the maximum condition also lead to the same result[13].

4.1.3 Penalty method

In Chapter 2, to find an approximate solution to the problem of minimizing the function in the presence of a constraint in the form of equality, the penalty method, a former specific alternative to the Lagrange multiplier method. It suggested a transition to an unconditional extremum problem to with a small parameter. In the optimal control problem by restriction in the form of equality, one can understand the state equation, which opens up certain prospects for the application of the penalty method. Apply it for analysis of Example 3.4. In this case, it is required to minimize the functional

$$I(u) \;=\; \frac{1}{2}\int\limits_0^1 (u^2 + x^2)dt,$$

where the function x is a solution of the problem

$$x'(t) = u(t),\; t \in (0,1);\; x(0) = 0.$$

Using ***penalty method***, we determine the functional of two variables

$$J_\varepsilon(u,x) \;=\; \frac{1}{2}\int\limits_0^1 (u^2 + x^2)dt \;+\; \frac{1}{2\varepsilon}\int\limits_0^1 (x' - u)^2,$$

where ε is a small positive parameter. Consider its minimization problem with additional condition $x(0) = 0$.

Suppose the pair $(u_\varepsilon, x_\varepsilon)$ is a solution of this problem, i.e., its ***optimal pair***[14]. To find it, we use the method described in Section 4.1.1, taking into account the fact that the functional J_ε depends on the two variables. We have the inequality

$$J_\varepsilon(u_\varepsilon + \sigma h, x_\varepsilon) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \;\geq\; 0$$

for any function h and all positive numbers σ. The following inequality holds

$$\frac{\sigma}{2}\int\limits_0^1 (2u_\varepsilon h + \sigma h^2)dt \;+\; \frac{\sigma}{2\varepsilon}\int\limits_0^1 \big[-2(x'_\varepsilon - u_\varepsilon)h + \sigma h^2\big]dt \;\geq\; 0.$$

Dividing this formula by σ and passing to the limit as $\sigma \to 0$, we get

$$\int\limits_0^1 \Big(u_\varepsilon - \frac{x'_\varepsilon - u_\varepsilon}{\varepsilon}\Big)hdt \;\geq\; 0.$$

The function h is arbitrary, so we obtain

$$u_\varepsilon - \varepsilon^{-1}(x'_\varepsilon - u_\varepsilon) \;=\; 0.$$

Defining the function

$$p_\varepsilon = \varepsilon^{-1}(x'_\varepsilon - u_\varepsilon), \tag{4.7}$$

we have the equality

$$u_\varepsilon = p_\varepsilon. \tag{4.8}$$

Similarly for any continuously differentiated function h, equal to zero at the initial time moment[15], and positive numbers σ the following inequality holds

$$J_\varepsilon(u_\varepsilon, x_\varepsilon + \sigma h) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0.$$

Now we have

$$\frac{\sigma}{2}\int_0^1 (2x_\varepsilon h + \sigma h^2)dt + \frac{\sigma}{2\varepsilon}\int_0^1 \big[-2(x'_\varepsilon - u_\varepsilon)h' + \sigma h'^2\big]dt \geq 0.$$

Dividing this inequality by σ and passing to the limit as $\sigma \to 0$ using formula (4.8), we have

$$\int_0^1 (x_\varepsilon h - p_\varepsilon h')dt \geq 0.$$

Integrating by parts, we get

$$\int_0^1 p_\varepsilon h' dt = -\int_0^1 p'_\varepsilon h dt + p_\varepsilon(1)h(1),$$

because $h(0) = 0$. Now the previous inequality takes the form

$$\int_0^1 (x_\varepsilon - p'_\varepsilon)h dt + p_\varepsilon(1)h(1) \geq 0. \tag{4.9}$$

This formula is true for all functions h, including those that satisfy equality $h(1) = 0$. For them, this inequality is

$$\int_0^1 (x_\varepsilon - p'_\varepsilon)h dt \geq 0.$$

Using the previously described technique, taking into account arbitrariness of h, we establish that $p'_\varepsilon = x_\varepsilon$. Then from inequality (4.9) it follows that $p_\varepsilon(1)h(1) \geq 0$. From here, due to arbitrariness of h, it follows that $p_\varepsilon(1) = 0$. Therefore, the function p_ε is a solution to the problem

$$p'_\varepsilon(t) = x_\varepsilon(t),\ t \in (0,1);\ p_\varepsilon(1) = 0. \tag{4.10}$$

Using the formula (4.8) and given initial condition, determine the Cauchy problem

$$x'_\varepsilon(t) = u_\varepsilon(t) + \varepsilon p\varepsilon(t),\ t \in (0,1);\ x_\varepsilon(0) = 0. \tag{4.11}$$

Now, we have the system[16] (4.8), (4.10), and (4.11) for finding three unknown functions u_ε, p_ε, and x_ε. At the same time, the problem (4.10) is a standard adjoint system, and equality (4.8) is a solution to the conditions of the maximum principle and variational inequality for the example considered. The same result was obtained in Section 4.1.1. Finally, the Cauchy problem (4.11) with small values ε can be interpreted as an approximation of a given state equation. We have established formulas, which in a certain sense an approximation of the previously obtained system of optimality conditions for this example. Thus, using the penalty method, we can really get the approximate value of optimal control[17] for Example 3.4.

4.1.4 Bellman equation

We return to the general optimal control problem discussed earlier in Chapter 3 and Section 4.1.2. In this case, this problem is analyzed using ***dynamic programming***[18]. It is based on the ***Bellman optimality principle***, according to which the optimal control does not depend on the history of the system and is determined by the state of the system at a given moment. Thus, if $u = u(t)$ is the optimal control of the system on the time interval $(0, T)$, then for any point ξ of this interval, the same function will be the optimal control of the system on the interval (ξ, T), i.e., any final part of the optimal control of the system is itself optimal[19].

When solving a problem using dynamic programming, the following function of two variables is definite

$$B(t,x) \;=\; \min_{v\in U(t)} \Big[\int\limits_t^T g(\tau, v(\tau), y(\tau))d\tau \;+\; h(y(T))\Big] \tag{4.12}$$

for all numbers $t \in (0,T)$ and x, where

$$U(t) \;=\; \Big\{u \big|\; a(\tau) \le u(\tau) \le b(\tau),\; \tau \in (t,T)\Big\},$$

and $y = y(\tau)$ is a solution of the Cauchy problem

$$y'(\tau) \;=\; f(\tau, v(\tau), y(\tau)),\; \tau \in (t,T);\; y(t) = x. \tag{4.13}$$

Thus, there is a family of optimal control problems of the type of Problem 3.1, differing only in the starting point t and the state of the system x at this moment in time, and the function B characterizes the minimum value of the corresponding functional.

Definition 4.2 *The function B is called the* ***Bellman function****.*

Consider a small enough time interval Δt. We have

$$B(t,x) = \min_{v\in U(t)} \Big[\int\limits_t^{t+\Delta t} g(\tau, v(\tau), y(\tau))d\tau + \int\limits_{t+\Delta t}^{T} g(\tau, v(\tau), y(\tau))d\tau + h(y(T))\Big].$$

Because of the Bellman optimality principle, the solution of the optimal control problem on the segment (t, T) turns out to be optimal on the interval $(t + \Delta t, T)$. Then the last equality can be written as

$$B(t,x) = \min_{v \in U(t)} \Big\{ \int_t^{t+\Delta t} g(\tau, v(\tau), y(\tau)) d\tau + \min_{v \in U(t+\Delta t)} \Big[\int_{t+\Delta t}^{T} g(\tau, v(\tau), y(\tau)) d\tau + h(y(T)) \Big] \Big\}.$$

Using the definition of the Bellman function, we get

$$B(t,x) \; = \; \min_{v \in U(t)} \Big[\int_t^{t+\Delta t} g(\tau, v(\tau), y(\tau)) d\tau \; + \; B(t + \Delta t, x + \Delta x) \Big],$$

where $\Delta x = y(t + \Delta t) - y(t)$ is determined by the control on the interval $(t, t + \Delta t)$. Using Taylor formula[20], we have

$$B(t + \Delta t, x + \Delta x) \; = \; B(t,x) + B_t(t,x)\Delta t + B_x(t,x)\Delta x + \eta \; =$$

$$B(t,x) \; + \; \Big[B_t(t,x) + B_x(t,x) \frac{\Delta x}{\Delta t} \Big] \; + \; \eta,$$

where η is a value that has a higher order of smallness with respect to increments. Then the previous equality takes the form

$$B(t,x) \; = \; \min_{v \in U(t)} \Big\{ \int_t^{t+\Delta t} g(\tau, v(\tau), y(\tau)) d\tau + B(t,x) + \Big[B_t(t,x) + B_x(t,x) \frac{\Delta x}{\Delta t} \Big] \; + \; \eta \Big\}.$$

The value $B(t, x)$ here does depend from the control. Therefore, we can replace it before the minimum at the right-hand side of this formula.

Divide this equality by Δt and pass to the limit as $\Delta t \to 0$. Using the mean value theorem, we get

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} g(\tau, v(\tau), y(\tau)) d\tau \; = \; g(t, v(t), y(t)) = g(t, w, x),$$

where $w = v(t)$. Using the equation (4.13), determine

$$\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} \; = \; \lim_{\Delta t \to 0} \frac{y(t + \Delta t) - y(t)}{\Delta t} \; = \; y'(t) = f(t, v(t), y(t)) = f(t, w, x).$$

Finally, $\eta / \Delta t \to 0$ as $\Delta t \to 0$. Now we obtain

$$0 \; = \; \min \Big[g(t, w, x) + B_t(t,x) + B_x(t,x) f(t, w, x) \Big].$$

On the right side of this equality, the minimum is taken over the control. However, the value under square brackets depends only on w, which is the control at time t. It follows from the definition of the set of admissible controls that it belongs to the interval $[a(t), b(t)]$. Then, taking into account that the derivative $B_t(t, x)$ does not depend on the control, we reduce the last equality to the form

$$B_t(t,x) \; + \; \min_{w \in [a(t), b(t)]} \Big[g(t, w, x) + B_x(t,x) f(t, w, x) \Big] \; = \; 0. \qquad (4.14)$$

Definition 4.3 *The equality* (4.14) *is called the* ***Bellman equation***[21]

Obviously, this is a first-order partial differential equation. Passing at (4.12) $t = T$, we obtain the final condition for the Bellman equation[22]

$$B(T, x) = h(x). \tag{4.15}$$

The following assertion holds[23].

Theorem 4.2 *Suppose the exists a Bellman function $B = B(t, x)$, which is a solution of problem* (4.14), (4.15), *and the control $u = u(t)$, which minimizes the correspondence function of the Bellman equation. Then this control is the solution of Problem* 3.1.

Theorem 4.2 gives the ***sufficient optimality conditions***[24] for Problem 3.1, because by this result the control determining after analysis of the Bellman equation is optimal. However, this is not guaranty that any optimal control can be determined by this method[25].

Use Theorem 4.2 for the analysis of the considered before ***linear-quadratic optimal control problem***; see Problem 3.2. We have the problem of minimization of the quadratic functional

$$I(u) = \frac{1}{2}\int_0^T \Big\{\alpha\big[x(t) - z(t)\big]^2 + \beta\big[u(t)\big]^2\Big\}dt,$$

where x is a solution of the linear problem

$$x'(t) = a(t)x(t) + b(t)u(t) + f(t),\ t \in (0, T);\ x(0) = x_0.$$

The constraints with respect to the control are absent here. Consider an easy case with $z(t) = 0,\ f(t) = 0$.

The value under the square brackets of the Bellman equation is

$$\Psi(w) = B_x(t, x)\big[a(t)x) + b(t)w\big] + \alpha x^2 + \beta w^2.$$

Differentiate the function Ψ and equal the result to zero. We find the corresponding stationary point that is

$$u(t) = -(2\beta)^{-1}Bx(t, x)b(t).$$

The second derivative of Ψ is positive, so this is the point of minimum. Putting it to the Bellman equation, we get

$$B_t(t, x) + B_x(t, x)a(t)x + \alpha x^2 - (4\beta)^{-1}\Big[B_x(t, x)b(t)\Big]^2 = 0. \tag{4.16}$$

Try to find the solution of this equation by the formula

$$B(t, x) = r(t)x^2,$$

where the function r will be chosen later. Find the partial derivatives

$$B_t(t,x) = r'(t)x^2, \quad B_x(t,x) = 2r(t)x.$$

Now the control is determined by the formula

$$u(t) \;=\; \beta^{-1}r(t)b(t)x; \tag{4.17}$$

and the equality (4.16) takes the form

$$\Big[r'(t) - \beta^{-1}r(t)^2b(t)^2 + 2a(t)r(t) + \alpha\Big]x^2 \;=\; 0.$$

Besides, from the formula (4.15) it follows

$$r(T)x^2 = 0.$$

Two previous equalities are true if the coefficients before x^2 for both cases are zero. This is possible if the function r is a solution of the problem

$$r'(t) - \beta^{-1}r(t)^2b(t)^2 + 2a(t)r(t) + \alpha \;=\; 0, \; t \in (0,T); \; r(T) = 0. \tag{4.18}$$

We have the ***Riccati equation***, which is exactly the same as equation (3.28), which was obtained in Chapter 3.

Now the algorithm for solving the problem is as follows.

1. The function r is found from system (4.18).
2. The explicit dependence of the control on the state function is determined by formula (4.17), which corresponds to the solution of the ***synthesis problem***.
3. After substituting the control from formula (4.17) into the state equation and its solving, the function x is found.
4. Optimal control as a function of time is determined by formula (4.17) with previously defined functions r and x.

Consider, in particular, Example 3.4, which consists in minimizing the functional

$$I(u) \;=\; \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is a solution of the system

$$x'(t) \;=\; u(t), \; t \in (0,1); \; x(0) = 0.$$

The problem (4.18) takes the form

$$r'(t) + r(t)^2 \;=\; 1; \; r(1) = 0.$$

Its solution was found in the previous chapter. After its substitution, first into formula (4.18), and then into the state equation, the Cauchy problem is obtained

$$x'(t) \;=\; r(t)x(t), \; t \in (0,1); \; x(0) = 0.$$

which has zero solution. Then it follows from equality (4.18) that $u(t) = 0$. This control, as we already know, is the solution of the considered optimal control problem.

RESULTS

Here is a list of questions devoted to the described methods for solving optimal control problems and the examples considered, the main conclusions on this topic, as well as the additional problems that arise in this case, partially solved in Appendix, partially explained by the Notes.

Questions

It is required to answer questions concerning the properties of the optimal control problem and the optimality condition in the form of the maximum principle.

1. In the process of analyzing Example 3.4, inequality (4.1) was obtained. Subsequently, this formula was divided by σ, and the passage to the limit was carried out as $\sigma \to 0$. Why does the summand, which includes the square of the state increment Δx, tend to zero?
2. Why, after obtaining inequality (4.1), it was necessary to consider the problem for the increment Δx?
3. How did the function λ appear when calculating the limit (4.2)?
4. For what purpose, in the study of Example 3.4 in Section 4.1.1, the function λ specially chosen?
5. On the basis of what, after passing to the limit in the corresponding inequality in Section 4.1.1, was it concluded that the equality $u = p$ is true?
6. What is the Gateau derivative of the differentiable function?
7. How can the stationary conditions for functions of one and many variables be derived from the equality to zero of the Gateaux derivative of an arbitrary functional at an extremum point?
8. Why was the formula $v = u + \sigma h$ used in the analysis of Example 3.4 replaced by another formula in the analysis of Example 3.3?
9. Why are the algorithms characterized by equalities (4.3) and (4.4) not suitable for the analysis of Example 3.3?
10. Is there any difference between the iterative method of solving the optimality condition in the form of the maximum principle for Example 3.3 used in Chapter 3 and the gradient projection method described in the Lecture?
11. Whence does it follow that the control $v_{\sigma s}$ used in the derivation of the variational inequality is admissible?
12. What property of the state equation was used in the derivation of the variational inequality?

13. What requirements and why were imposed on the functions f and g when deriving the variational inequality?
14. Which of the methods (maximum principle or variational inequality) is more effective in the analysis of Example 3.1?
15. Can variational inequalities be used to analyze Example 3.4?
16. Why is the equality $u(t) = p(t)$ obtained during the study of Example 3.1 only meaningful for $p(t) \in [1, 2]$?
17. What is the point of using the penalty method for solving optimal control problems?
18. When analyzing Example 3.4 using the penalty method, inequality (4.9) was obtained. How was problem (4.10) derived from it?
19. What is the difference between the system (4.8), (4.10), and (4.11) obtained in the analysis of Example 3.4 and similar relations for the same example obtained in Section 4.1.1?
20. How can the approximate value of the optimal control be found using the penalty method?
21. Is it possible to find an exact, rather than an approximate, solution to an optimal control problem using the penalty method?
22. What is the meaning of Bellman optimality principle?
23. How is the Bellman optimality principle used in deriving the Bellman equation?
24. What class of equations does the Bellman equation belong to?
25. The Bellman equation includes the first derivatives of the function B with respect to both variables. Why does it get a boundary condition for t and not a condition for the variable x?
26. From what considerations was problem (4.18) obtained with respect to the function r in the considered example?
27. What happens if one applies the Bellman equation to analyze Example 3.4?

Conclusions

Based on the study of the considered optimization methods, the following conclusions can be drawn.

- The stationary condition can be extended to problems of minimizing functionals using Gateaux derivatives.

- Gradient methods can be extended to minimization problems for functionals using Gateaux derivatives.
- For finding the optimal control, iterative methods for solving the corresponding system of optimality conditions can be applied.
- The system of optimality conditions for Example 3.4 can be solved iteratively using the method of successive approximations or the gradient method.
- The system of optimality conditions for Example 3.3 can be solved iteratively using the gradient projection method.
- To solve optimal control problems, along with the maximum principle, variational inequalities can be used.
- Variational inequality is a necessary condition for optimality for the considered class of problems.
- For Example 3.1, the maximum principle and variational inequality are equivalent.
- For Example 3.3, the maximum principle and variational inequality are equivalent.
- For an approximate solving of optimal control problems, one can use the penalty method, which involves solving the problem of unconstrained minimization of a functional with a small parameter.
- Application of the penalty method for Example 3.4 results in a system that differs from the system of optimality conditions obtained for it in Chapter 3 only by a relation that can be interpreted as an approximate equation of state.
- Dynamic programming can be used to solve optimal control problems.
- Dynamic programming provides a sufficient optimality condition based on the Bellman equation.
- The Bellman equation is a first-order partial differential equation that includes the solution of a function minimization problem.
- Using the Bellman equation, one can obtain a solution to a program control problem and a synthesis problem for a linear-quadratic problem.
- In the process of applying the Bellman equation for a linear-quadratic optimal control problem, the Riccati equation is obtained.

Problems

Based on the described optimization methods and the considered examples, we have the following problems.

1. **Application of iterative methods for the practical solving of optimality conditions**. In the Lecture, iterative methods were used to study the optimality conditions for Examples 3.3 and 3.4. It is of interest to use them for a wider class of problems to be solved. Similar methods will be used in Chapters 5, 6, and 7.

2. **Convergence of iterative methods**. In connection with the practical application of iterative methods for solving optimal problems, the problem of justifying their convergence arises, see Notes[26]

3. **Applicability of iterative methods**. We would like to know the limits of applicability of the iterative methods described in Section 4.1.1. One result in this direction is given in Appendix.

4. **Justification of the variational inequality**. The Lecture actually lacked the substantiation of Theorem 4.1. For justification of optimality conditions in the form of variational inequalities, see Notes[27].

5. **Relationship between the maximum condition and the variational inequality**. For the considered examples, the variational inequality and the maximum condition turned out to be equivalent. However, it remains unclear whether this equivalence holds in the general case. The answer to this question is given in Appendix.

6. **Application of variational inequality for the analysis of different optimal control problems**. We limited ourselves to applying the variational inequality to solve the problems from Examples 3.1 and 3.3. In Appendix, this optimality condition will also be used to analyze Example 3.2. In Chapters 5, 6, 9, and 13, variational inequalities are applied to study other problems.

7. **Application of the penalty method to solve general optimal control problems**. In the Lecture, we limited ourselves to applying the penalty method to the analysis of Example 3.4. In Appendix, it is used to study two more difficult problems. Regarding the application of this method to solving a wide class of optimal control problems, see Notes[28].

8. **Penalty method for problems with constraints on control values**. We used the penalty method for the analysis, in which there were no restrictions on the control values. However, problems with control constraints are more interesting. In Appendix, this method is used for a problem with a control constraint.

9. **Optimal control of singular systems**. The solution of optimal control problems was carried out under the assumption that the system under consideration

is regular in the sense that for any admissible control the equation of state has a unique solution that continuously depends on the control. However, in practice it is not always possible to justify the regularity of the system. Moreover, the system can, in principle, turn out to be singular. Under these conditions, the applied research methods are not effective. Optimal control problems for two qualitatively different singular systems are studied in Appendix using the penalty method, see also Notes[29].

10. **Justification of the Bellman equation**. In Lecture, Theorem 4.2 remained virtually unfounded. Justification of the Bellman equation is given in Appendix, see also Notes[30].

11. **Applicability of Bellman optimality principle**. The derivation of the Bellman equation was carried out on the basis of the Bellman optimality principle. It is of interest to identify the limits of its application. Appendix shows that it is being readied for Example 3.1. In Chapter 10, we will show the applicability of this statement to one optimal control problem for a system with a pinned finite state. An example of an optimal control problem for which the Bellman optimality principle does not apply is given in Chapter 15.

12. **Application of dynamic programming for solving optimal control problems**. For practical applications of dynamic programming, see Notes[31].

13. **Relationship between the maximum principle and dynamic programming**. To analyze the general problem of optimal control, use the maximum principle and dynamic programming. The connection between these approaches is discussed in Appendix, see also Notes[32].

4.2 APPENDIX

Some additional information about the methods for solving optimal control problems described in the Lecture is given below. In particular, Section 4.2.1 considers an optimal control problem for which the methods of Section 4.1.1 turn out to be inapplicable due to the non-differentiability of the optimality criterion. Section 4.2.2 analyzes Example 3.2 using a variational inequality, which leads to a different result from that obtained earlier using the maximum principle. In Section 4.2.3, the penalty method is used to solve a constrained optimal control problem, in which case it is used in conjunction with a variational inequality. In Section 4.2.4, we consider the optimal control problem of the system, for which it is impossible to guarantee the existence of a solution to the state equation on a given time interval for an arbitrary admissible control, and in Section 4.2.5, there is no uniqueness of the solution to the Cauchy problem. Under these conditions, the penalty method turns out to be quite effective. Section 4.2.6 provides a justification for Theorem 4.2 on sufficient optimality conditions, and Section 4.2.7 establishes a connection between dynamic programming and the maximum principle. Section 4.2.8 shows that for Example 3.1 the Bellman optimality principle is implemented.

4.2.1 Problem with a non-smooth functional

In Section 4.1.1, we considered optimal control problems that were solved using different iterative methods. However, it is not always possible to use them. Consider the following problem[33].

Example 4.1 *Find the minimum of the functional*

$$I(u) = \frac{1}{2}\int_0^1 \big[x(t) - z(t)\big]^2 dt,$$

where z is a given function, and x is a solution of the Cauchy problem

$$x'(t) = f(t,u),\ t \in (0,1);\ x(0) = 0,$$

besides

$$f(t,u) = \begin{cases} f_1, & \text{if } 0 < t < u, \\ f_2, & \text{if } u < t < 1, \end{cases}$$

with known values f_1 and f_2.

To solve the problem, we try to use the iterative methods described in Section 4.1.1. Suppose u is the solution to this problem. Let us find the Gateaux derivative of the minimized functional at this point. We define $v = u + \sigma h$, with some numbers h and σ. Determine the difference

$$I(v) - I(u) = \frac{1}{2}\int_0^1 \big[(y-z)^2 - (x-z)^2\big]dt,$$

where y is a solution of the same Cauchy problem for the control v. The previous equality takes the form

$$I(v) - I(u) = \int_0^1 (x-z)\Delta x dt + \frac{1}{2}\int_0^1 \Delta x^2 dt, \tag{4.19}$$

where $\Delta x = y - x$. This can be transformed to the equality[34]

$$I(v) - I(u) = \int_0^1 p\big[f(t,v) - f(t,u)\big]dt + \frac{1}{2}\int_0^1 \Delta x^2 dt, \tag{4.20}$$

where p is a solution to the problem

$$p'(t) = x(t) - z(t),\ t \in (0,1);\ p(1) = 0.$$

Using the state equation, we find

$$\Delta x(t) = \int_0^t \big[f(\tau,v) - f(\tau,u)\big]d\tau. \tag{4.21}$$

Determine the integral

$$\int_0^1 p[f(t,v)-f(t,u)]dt = \Big[\int_0^{u+\sigma h} pf_1 dt - \int_0^u pf_1\Big]dt + \Big[\int_{u+\sigma h}^1 pf_2 dt - \int_u^1 pf_2\Big]dt.$$

Further transformation depends on the sign of the number h. Suppose the value of h is positive. Then we get

$$\int_0^1 p[f(t,v)-f(t,u)]dt = (f_1-f_2)\int_u^{u+\sigma h} p(t)dt.$$

From equality (4.21), it follows that the increment $\Delta x(t)$ is zero for $t<u$, the value $(f_1-f_2)(t-u)$ for $u<t<u+\sigma h$ and $(f_1-f_2)\sigma h$ for $t>u+h$. Then the second term on the left side of equality (4.20) does not exceed $(|f_1-f_2|\sigma h)^2/2$. As a result, condition (4.20) implies

$$\frac{I(u+\sigma h)-I(u)}{\sigma} = (f_1-f_2)\frac{1}{\sigma}\int_u^{u+\sigma h} p(t)dt + \frac{1}{2\sigma}\int_0^1 \Delta x^2 dt.$$

Passing here to the limit as $\Delta x \to 0$, using the mean value theorem, we obtain

$$\lim_{\sigma\to 0}\frac{I(u+\sigma h)-I(u)}{\sigma} = (f_1-f_2)p(u)h. \tag{4.22}$$

For negative values h, we have

$$\int_0^1 p[f(t,v)-f(t,u)]dt = (f_2-f_1)\int_{u+\sigma h}^{u} p(t)dt.$$

From inequality (4.20), after dividing by h, we have

$$\frac{I(u+\sigma h)-I(u)}{\sigma} = (f_2-f_1)\frac{1}{\sigma}\int_{u+\sigma h}^{u} p(t)dt + \frac{1}{2\sigma}\int_0^1 \Delta x^2 dt.$$

For the increment $\Delta x(t)$ in this case, the same estimate is set as in the previous case. Then after passing to the limit, we have

$$\lim_{\sigma\to 0}\frac{I(u+\sigma h)-I(u)}{\sigma} = (f_2-f_1)p(u)h. \tag{4.23}$$

Combining equalities (4.22) and (4.23), we obtain

$$\lim_{\sigma\to 0}\frac{I(u+\sigma h)-I(u)}{\sigma} = (f_1-f_2)p(u)|h|.$$

Thus, the limit of the value, which is here on the left side of the equality, exists, but is not linear with respect to h. This means that the functional to be minimized is not Gateaux differentiable. As a result, it is not possible to use the iterative methods described in Section 4.1.1 for the analysis of Example 4.1. At the same time, it is possible to determine the subgradient of the functional (see Chapter 1) and use non-smooth optimization methods[35].

4.2.2 Non-equivalence of the variational inequality and maximum condition

As we already know, along with the maximum condition, variational inequality can be used to solve optimal control problems. The corresponding optimality conditions are quite close, including the state equation, the adjoint system, and the optimality condition itself. In the Lecture, it was shown that, in relation to Examples 3.1 and 3.3, these methods turned out to be equivalent, since they led to the same result. We now apply the variational inequality to study Example 3.2, in which it is required to find the maximum of the functional from Example 3.1. Thus, it is required to find the maximum of the functional

$$I(u) = \int_0^1 \Big(\frac{u^2}{2} - 3x\Big)dt$$

on the control set

$$U = \Big\{u \big| \; 1 \leq u(t) \leq 2, \; 0 < t < 1\Big\},$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t), \; t \in (0,1); \; x(0) = 0.$$

As already noted, the problem of maximizing the functional I is equivalent to the problem of minimizing the functional $-I$. With regard to the use of the variational inequality, this only leads to the replacement of the relation $\leq$ included in it by $\geq$. We define the function H by the formula

$$H(t,u,x,p) = pu - u^2/2 + 3x,$$

same as for Example 3.1 where, as established in Chapter 3, $p(t) = 3-3t$. Then the optimal control will satisfy the variational inequality

$$[p(t) - u(t)][v - u(t)] \geq 0 \;\; \forall v \in [1,2], \; t \in (0,1).$$

The unique difference from the analogous inequality for Example 3.1 is the presence here of the relation $\geq$ instead of $\leq$. We again have a product of two quantities, the first of which is concrete, and the second varies.

Assume that $p(t) - u(t) > 0$. Then, dividing the inequality by the first multiplier, we establish that $v - u(t) \geq 0$ for all $v \in [1,2]$, i.e., $u(t)$ must not be greater than all numbers from the interval [1,2]. Considering that the value of the control itself belongs to this segment, we conclude that the last inequality is possible only for $u(t) = 1$. Therefore, if the value of $p(t)$ is greater than $u(t)$, equal to one, then $u(t) = 1$. Thus, this equality is satisfied for $p(t) > 1$. If $p(t) - u(t) < 0$, then after dividing the variational inequality by the first multiplier, we obtain that $v - u(t) \leq 0$ for all $v \in [1,2]$, i.e., $u(t)$ must be no less than all numbers from the interval [1,2], which is possible for $u(t) = 2$. Thus, if $p(t)$ is less than $u(t)$, which is equal to 2, then $u(t) = 2$. Thus, this equality holds for $p(t) < 2$. Finally, the equality $p(t) - u(t) = 0$ is possible, and hence $u(t) = p(t)$. However, this equality is admissible only for $p(t) \in [1,2]$.

Previously, it was found that $p(t) = 3–3t$. Obviously, $p(t)$ is greater than 2 for $t < 1/3$, belongs to the interval [1,2] for $1/3 \leq t \leq 2/3$, and is less than 1 for $t > 2/3$. Based on the results obtained earlier, we conclude that $u(t) = 1$ for $t < 1/3$, $u(t) = 2$ for $t > 2/3$, and can take one of the three values 1, 2, or $3–3t$ for $1/3 \leq t \leq 2/3$. At the same time, using the maximum principle, the optimal control is

$$u(t) \;=\; \begin{cases} 1, & \text{if } \; 0 < t < 1/2, \\ 2, & \text{if } \; 1/2 < t < 1. \end{cases}$$

Obviously, the previous result coincides with this one for $t < 1/3$ and $t > 2/3$. However, for $1/3 \leq t \leq 2/3$, the maximum principle allows one to determine the solution of the problem uniquely, while, by the variational inequality, one of the three specified values can be optimal, but it is not clear in advance which one. Thus, for this example, the considered optimality conditions are not equivalent, and the variational inequality turns out to be only a necessary, but not sufficient, optimality condition[36].

4.2.3 Penalty method in the optimal control problem with constraints

In the Lecture, the penalty method was applied to the study of Example 3.4, in which there were no restrictions on the control values. Let us show that this method is also applicable to problems with constraints. Consider, in particular, Example 3.3, which is a generalization of Example 4.1 to problems of optimal control with constraints. In this case, it is required to minimize the functional

$$I(u) \;=\; \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

on the set

$$U \;=\; \Big\{u \big|\; |u(t)| \leq 1,\; t \in (0,T)\Big\},$$

where the function x is a solution of the problem

$$x'(t) \;=\; u(t),\; t \in (0,1);\; x(0) = 0.$$

As in the case of Example 3.4, we introduce a functional of two variables

$$J_\varepsilon(u,x) \;=\; \frac{1}{2}\int_0^1 (u^2 + x^2)dt \;+\; \frac{1}{2\varepsilon}\int_0^1 (x' - u)^2 dt,$$

where ε is a small positive number. We consider the problem of its minimization on the set of ***admissible pairs***, characterized by the control belonging to the set U and the fulfillment of the initial condition $x(0) = 0$. Due to the presence of restrictions on the control, we use the variational inequality to solve this problem.

Assume that a pair of functions $(u_\varepsilon, x_\varepsilon)$ is a solution to this problem. To analyze it, we will use the previously described method, corrected for the fact that the functional J_ε depends on two variables. The following inequality holds

$$J_\varepsilon(u_\varepsilon + \sigma(s - u_\varepsilon), x_\varepsilon) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0$$

for any function $s \in U$ and all numbers $\sigma \in (0, 1)$. We get

$$\frac{\sigma}{2}\int_0^1 \Big[2u_\varepsilon(s - u_\varepsilon) + \sigma(s - u_\varepsilon)^2\Big]dt + \frac{\sigma}{2\varepsilon}\int_0^1 \Big[-2(x'_\varepsilon - u_\varepsilon)(s - u_\varepsilon) + \sigma(s - u_\varepsilon)^2\Big]dt \geq 0.$$

Dividing this inequality by σ and passing to the limit as $\sigma \to 0$, we have

$$\int_0^1 \Big(u_\varepsilon - \frac{x'_\varepsilon - u_\varepsilon}{\varepsilon}\Big)(s - u_\varepsilon)dt \ \geq 0 \ \ \forall s \in U.$$

Denoting

$$p_\varepsilon \ = \ \varepsilon^{-1}(x'_\varepsilon - u_\varepsilon), \tag{4.24}$$

we obtain

$$\int_0^1 (u_\varepsilon - p_\varepsilon)(s - u_\varepsilon)dt \ \geq 0 \ \ \forall s \in U.$$

We have here a variational inequality similar to that established in Section 4.1.2. Using here the needle variation of the control (see Section 4.1.2), we obtain the pointwise variational inequality

$$[u_\varepsilon(t) - p_\varepsilon(t)][v - u_\varepsilon(t)] \geq 0 \ \ \forall v \in [-1, 1], \ t \in (0, T), \tag{4.25}$$

which is the partial case of the formula (4.5). Moreover, up to notation and the form of the set of admissible controls, it coincides with the variational inequality obtained in the Lecture when analyzing Example 3.1. Using the well-known method, we find[37]

$$u_\varepsilon(t) \ = \ \begin{cases} -1, & \text{if } \ p_\varepsilon(t) < -1, \\ p_\varepsilon(t), & \text{if } \ -1 \leq p_\varepsilon(t) \leq 1, \\ 1, & \text{if } \ p_\varepsilon(t) > 1. \end{cases} \tag{4.26}$$

Further, for any continuously differentiable function h equal to zero at the initial time moment, and positive numbers σ, we have the inequality

$$J_\varepsilon(u_\varepsilon, x_\varepsilon + \sigma h) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0.$$

Repeating here the same transformations as for the analogous inequality from Section 4.1.3, we arrive at the adjoint system

$$p'_\varepsilon(t) \ = \ x_\varepsilon(t), \ t \in (0, 1); \ p_\varepsilon(1) = 0, \tag{4.27}$$

which coincides with the problem (4.10).

Equality (4.24), taking into account the given initial condition, implies the Cauchy problem

$$x_\varepsilon'(t) \;=\; u_\varepsilon(t) + \varepsilon p_\varepsilon(t),\; t \in (0,1);\; x_\varepsilon(0) = 0, \tag{4.28}$$

which coincides with problem (4.11). Now, with respect to three unknown functions u_ε, p_ε and x_ε, the system (4.26)–(4.28) is obtained. Comparing it with the system of optimality conditions (3.13)–(3.15), we note that the only difference here is equation (4.28), which, for small enough ε, can be interpreted as an approximation of the initially given equation of state. Practical finding of the solution of the obtained system can be carried out using the iterative process described in Chapter 3.

4.2.4 Optimal control of a singular system

All previously obtained results were established under the assumption that the state equation with the corresponding initial condition has a unique solution that depends continuously on the control[38]. However, singular systems for which these requirements are not met can also be the object of study.

Example 4.2 *Find a minimum of the functional*

$$I(u) \;=\; \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = x(t)^2 + u(t),\; t \in (0,1);\; x(0) = 2. \tag{4.29}$$

The main difference from Example 3.4 here is the presence of a non-linear term x^2 in the state equation.

At first, we analyze the properties of the problem (4.29). Let us define, for example, $u(t) = 0$. Then this equation is reduced to the form

$$\frac{dx}{x^2} \;=\; dt.$$

After integration, we get

$$-\frac{1}{x} \;=\; t + c,$$

where c is an arbitrary constant. Thus, the general solution of the equation is

$$x(t) \;=\; -\frac{1}{t+c}.$$

Obviously, $x(t) \to \infty$ as $t \to 1/2$. Thus, there exists a control for which problem (4.29). has no solution[39] on the time interval (0,1).

Let us now set the control $u(t) = 4$. Equating to zero the value on the right side of the first equality (4.29), we find the value $x(t) = 2$ corresponding to the

equilibrium position of the system. However, at the initial time, the system is already in equilibrium, which means that it will be in it subsequently. Consequently, there exists a control for which problem (4.29) has solutions on the time interval (0,1).

Thus, the state equation, depending on the choice of control, may or may not have a solution on a given interval. Now we have the following definition.

Definition 4.4 *The function pair* (u, x) *is called* ***admissible*** *for system* (4.29) *if it satisfies these equalities.*

We would like to minimize the given functional for system (4.29). Naturally, if under some control the Cauchy problem has no solution, then it makes no sense to consider such a control as a possible solution to the optimization problem. In this connection, we can assume that our problem is to minimize the functional I on the set of admissible pairs of the system (4.29). At the same time, we do not know the set of controls under which the equation of state makes sense. As a result, it is not possible to use the maximum principle or variational inequality to solve the problem, since in the process of variation of the control, when deriving optimality conditions, we can reach such a control for which the equation of state does not make sense. However, we can understand control and states as equal arguments of the minimized functional and interpret the equation of state as a constraint on the system, given as an equality. In this case, the penalty method described above makes it possible to find an approximate solution to the problem.

Consider the functional

$$J_\varepsilon(u, x) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt + \frac{1}{2\varepsilon}\int_0^1 (x' - x^2 - u)^2 dt,$$

where ε is a small positive parameter. The problem of its minimization under the additional condition $x(0) = 2$ is considered. To solve it, we use the technique described in the previous subsection.

Suppose the pair $(u_\varepsilon, x_\varepsilon)$ is a solution of this problem. We have the inequality

$$J_\varepsilon(u_\varepsilon + \sigma h, x_\varepsilon) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0$$

for any function h and all positive number σ. We get

$$\frac{\sigma}{2}\int_0^1 (u_\varepsilon^2 + x_\varepsilon^2)dt + \frac{\sigma}{2\varepsilon}\int_0^1 \Big[-2(x_\varepsilon' - x_\varepsilon^2 - u_\varepsilon)h + \sigma h^2\Big]dt \geq 0.$$

After division by σ and passing to the limit, as was done in Section 4.1.3, we have

$$\int_0^1 \Big(u_\varepsilon - \frac{x_\varepsilon' - x_\varepsilon^2 - u_\varepsilon}{\varepsilon}\Big)h dt \geq 0.$$

Hence it follows

$$u_\varepsilon - \varepsilon^{-1}(x'_\varepsilon - x_\varepsilon^2 - u_\varepsilon) \;=\; 0.$$

By analogy with equality (4.7), having defined the function

$$p_\varepsilon \;=\; \varepsilon^{-1}(x'_\varepsilon - x_\varepsilon^2 - u_\varepsilon), \tag{4.30}$$

we obtain the equality

$$u_\varepsilon \;=\; p_\varepsilon. \tag{4.31}$$

Similarly, for any continuously differentiable function h, equal to zero at the initial time, and positive numbers σ, the following inequality holds

$$J_\varepsilon(u_\varepsilon, x_\varepsilon + \sigma h) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0.$$

Now we obtain

$$\frac{\sigma}{2}\int_0^1 (2x_\varepsilon h + \sigma h^2)dt \;+\; \frac{\sigma}{2\varepsilon}\int_0^1 \Big[2(x'_\varepsilon - x_\varepsilon^2 - u_\varepsilon)(h' - 2x_\varepsilon h) + \sigma h'^2 + \sigma h^2\Big]dt \;\geq\; 0.$$

Divide this inequality by σ and pass to the limit, taking into account equality (4.31). We have

$$\int_0^1 (x_\varepsilon h - 2p_\varepsilon x_\varepsilon h + p_\varepsilon h')dt \;\geq\; 0.$$

Integrating by parts and performing the same transformations as in the previous chapter, we have the problem

$$p'_\varepsilon(t) \;=\; x_\varepsilon(t) - 2x_\varepsilon(t)p_\varepsilon(t),\; t \in (0,1);\; p_\varepsilon(1) = 0, \tag{4.32}$$

which is analog of (4.10). Finally, from equality (4.30), taking into account the given initial condition, it follows

$$x'_\varepsilon(t) \;=\; x_\varepsilon(t)^2 + u_\varepsilon(t) + \varepsilon p_\varepsilon(t),\; t \in (0,1);\; x_\varepsilon(0) = 2. \tag{4.33}$$

Thus, with respect to three unknown functions u_ε, p_ε and x_ε, the system (4.31)–(4.33) is obtained, and for a sufficiently small ε, problem (4.33) is an approximation of the equations of state (4.29). Thus, the penalty method really provides an opportunity to find an approximate solution to the problem of optimal control of a singular system[40].

4.2.5 Optimal control of a singular system with constraints

Consider an optimal problem of control for a qualitatively different singular system[41].

Example 4.3 *Find a minimum of the functional*

$$I(u) \;=\; \frac{1}{2}\int_0^1 (u^2 + x^2)dt$$

on the set

$$U = \Big\{u \big|\ 0 \le u(t) \le 1,\ t \in (0,1)\Big\},$$

where x is a solution of the Cauchy problem[42]

$$x'(t) = \sqrt{x(t)} + u(t),\ t \in (0,1);\ x(0) = 0. \tag{4.34}$$

Let us establish the properties of the given Cauchy problem[43]. Choose a concrete admissible control, for example, identically equal to zero. By a direct check, one can verify that the function

$$x_\xi(t) = \begin{cases} 0, & \text{if } 0 \le t \le \xi, \\ (t-\xi)^2/4, & \text{if } t > \xi \end{cases}$$

for any non-negative ξ is the solution to problem (4.34). Thus, a situation is possible when the state equation has an infinite set of solutions[44]; see Figure 4.1. This circumstance cannot serve as an obstacle to solving the problem of minimizing this functional on the set of admissible pairs of the system using the penalty method.

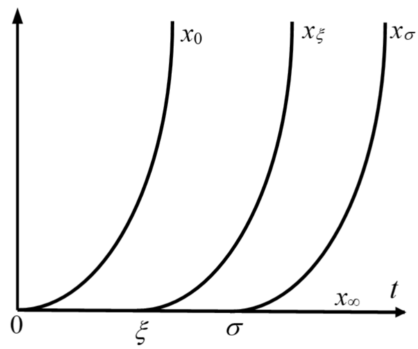

Figure 4.1 Solutions of problem (4.34) for $u = 0$.

Determine the functional

$$J_\varepsilon(u,x) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt + \frac{1}{2\varepsilon}\int_0^1 \left(x' - \sqrt{x} - u\right)^2 dt,$$

where ε is a small positive parameter. We consider the problem of its minimization under the initial condition $x(0) = 0$.

Let $(u_\varepsilon, x_\varepsilon)$ be a solution of this problem. The following inequality holds

$$J_\varepsilon(u_\varepsilon + \sigma h, x_\varepsilon) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \ge 0$$

for any function h and all positive number $\sigma \in (0,1)$. We get

$$\sigma\int_0^1 (u_\varepsilon(s - u_\varepsilon)dt + \frac{\sigma}{\varepsilon}\int_0^1 (x'_\varepsilon - \sqrt{x_\varepsilon} - u_\varepsilon)(s - u_\varepsilon)dt + o(\sigma) \ge 0,$$

where $o(\sigma)/\sigma \to 0$ as $\sigma \to 0$. Denote

$$p_\varepsilon(t) = \varepsilon^{-1}\big(x_\varepsilon' - \sqrt{x_\varepsilon} - u_\varepsilon\big). \tag{4.35}$$

Then, after dividing the previous inequality by σ and passing to the limit at $\sigma \to 0$, we obtain the variational inequality

$$\int_0^1 (u_\varepsilon - p_\varepsilon)(s - u_\varepsilon)dt \geq 0.$$

Up to the meaning of the quantities included in it, it coincides with the analogous relation from Section 4.2.4. After using the needle variation, we obtain the variational inequality

$$\big[u_\varepsilon(t) - p_\varepsilon(t)\big]\big[v - u_\varepsilon(t)\big] \geq 0 \;\; \forall v \in [0,1],\; t \in (0,1),$$

which is an analog of (4.25). Repeating the same reasoning as in solving the latter, we have the formula

$$u_\varepsilon(t) = \begin{cases} 0, & \text{if } p_\varepsilon(t) < 0, \\ p_\varepsilon(t), & \text{if } 0 \leq p_\varepsilon(t) \leq 1, \\ 1, & \text{if } p_\varepsilon(t) > 1, \end{cases} \tag{4.36}$$

which is an analog of (4.25). Now, for any continuously differentiable function h equal to zero at the initial time, and positive numbers σ, we have the inequality

$$J_\varepsilon(u_\varepsilon, x_\varepsilon + \sigma h) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0.$$

Using the equality

$$\sqrt{x_\varepsilon + \sigma h} = \sqrt{x_\varepsilon} + \frac{1}{2\sqrt{x_\varepsilon}}\sigma h + o(\sigma),$$

transform the previous inequality to the form

$$\sigma \int_0^1 x_\varepsilon h dt + \frac{\sigma}{\varepsilon}\int_0^1 \big(x_\varepsilon' - \sqrt{x_\varepsilon} - u_\varepsilon\big)\Big(h' - \frac{h}{2\sqrt{x_\varepsilon}}\Big)dt + o(\sigma) \geq 0.$$

After division by σ and passage to the limit as $\sigma \to 0$ using equality (4.35), we get

$$\sigma \int_0^1 x_\varepsilon h dt + \int_0^1 p_\varepsilon\Big(h' - \frac{h}{2\sqrt{x_\varepsilon}}\Big)dt \geq 0.$$

Carrying out the standard transformations here (see Section 4.1.3), taking into account the arbitrariness of the function h, we obtain the adjoint system

$$p_\varepsilon'(t) + \frac{1}{2\sqrt{x_\varepsilon(t)}}p_\varepsilon(t) = x_\varepsilon(t),\; t \in (0,1);\; p_\varepsilon(1) = 0. \tag{4.37}$$

Now from equality (4.35) follows the approximate state equation

$$x_\varepsilon'(t) = \sqrt{x_\varepsilon(t)} + u_\varepsilon(t) + \varepsilon p_\varepsilon(t),\; t \in (0,1);\; x_\varepsilon(0) = 0. \tag{4.38}$$

Thus, we get the system (4.36)–(4.38) for finding three unknown functions u_ε, p_ε, and x_ε. Solving this system for sufficiently small values of ε, it is possible, in principle, to find an approximate solution to the formulated problem of optimal control[45].

4.2.6 Justification of the sufficient optimality condition

The Lecture considered the Bellman equation as a sufficient optimality condition for Problem 3.1. However, Theorem 4.2 given there was not substantiated. We give a proof of this assertion.

Let the function $B = B(t,x)$ be a solution to the Bellman equation

$$B_t(t,x) + \min_{w\in[a(t),b(t)]} \Big[B_x(t,x)f(t,w,x) + g(t,w,x)\Big] = 0$$

with the final condition

$$B(T,x) = h(x).$$

Then for any control v from the set U the following inequality holds

$$B_t(t,x) + B_x(t,x)f(t,v(t),x) + g(t,v(t),x) \geq 0.$$

Let us define here as x the solution $y(t)$ of the problem

$$y'(t) = f(t,v(t),y(t)),\ t \in (0,T);\ y(0) = x_0.$$

As a result, the preceding inequality takes the form

$$B_t(t,x) + B_x(t,y(t))y'(t) + g(t,v(t),y(t)) \geq 0.$$

We have the equality

$$\frac{d}{dt}B(t,y(t)) = B_t(t,y(t)) + B_x(t,y(t))y'(t).$$

Thus, the following inequality holds

$$\frac{d}{dt}B(t,y(t)) + g(t,v(t),y(t)) \geq 0.$$

Integrating it from $t = 0$ to $t = T$, we get

$$B(T,y(T)) - B(0,y(0)) + \int_0^T g(t,v(t),y(t)) \geq 0,$$

where $B(T,y(T)) = h(y(T))$. The sum of this value with the integral from the last inequality is the value $I(v)$ of the functional to be minimized on an arbitrarily chosen admissible control v. Taking into account the initial condition $y(0) = x_0$, we obtain the inequality

$$I(v) \geq B(0,x_0). \tag{4.39}$$

Now let u be the control on which the minimum in the Bellman equation is reached. Then we obtain the equality

$$B_t(t,x) + g(t,u(t),x) + B_x(t,x)f(t,u(t),x) = 0.$$

Choose as x the solution $x(t)$ of the problem

$$x'(t) = f(t, u(t), x(t)),\ t \in (0,T);\ x(0) = x_0.$$

We have the problem

$$B_t(t, x(t)) + B_x(t, x(t))x'(t)) + g(t, u(t), x(t)) = 0.$$

Then the following equality holds

$$\frac{d}{dt}B(t, (t)) + g(t, v(t), y(t)) = 0.$$

After integration, we get

$$B(T, y(T)) - B(0, y(0)) + \int_0^T g(t, u(t), x(t)) = 0.$$

This equality is reduced to the form

$$I(u) = B(0, x_0).$$

Taking into account inequality (4.39), we obtain $I(v) \geq I(u)$. Hence, since the admissible control v is arbitrary, it follows that the control u is optimal. This completes the proof of Theorem 4.2.

4.2.7 Relationship between dynamic programming and the maximum principle

The maximum principle and dynamic programming were used to solve optimal control problems. Let us establish a connection between the corresponding optimality conditions for the considered Problem 3.1.

Assume that there exists a function $B = B(t, x)$ satisfying the Bellman equation (4.14) with boundary condition (4.15). Define the function

$$R(t, v, y) = B_t(t, y) + B_y(t, y)f(t, v, y) + g(t, v, y).$$

Let $u = u(t)$ be an optimal control, and $x = x(t)$ be a corresponding state system that is a solution of the problem

$$x'(t) = f(t, u(t), x(t)),\ t \in (0,T);\ x(0) = x_0. \tag{4.40}$$

From Bellman equation, it follows the relations

$$R(t, u(t), x(t)) = 0, \tag{4.41}$$

$$R(t, w, y) \geq 0\ \ \forall w \in [a(t), b(t)],\ \forall y. \tag{4.42}$$

Thus, the function R has its minimum in the second and third arguments at the point $(u(t), x(t))$. In accordance with the stationary condition, the derivative of R with respect to the third argument at the point $x(t)$ is equal to zero. Find the value[46]

$$R_x(t, u(t), x(t)) = B_{tx}(t, x(t)) + \frac{\partial}{\partial x}\Big[B_x(t, x(t)) f(t, u(t), x(t))\Big] + g(t, u(t), x(t)) = 0.$$

Calculating the derivative on the right side of this equality and using the state equation, we obtain

$$B_{tx}(t, x(t)) + B_{xx}(t, x(t))x'(t) + B_x(t, x(t)) f_x(t, u(t), x(t)) + g_x(t, u(t), x(t)) = 0.$$

Now we have

$$\frac{d}{dt}\Big[B_x(t, x(t))\Big] + f_x(t, u(t), x(t))B_x(t, x(t)) + g_x(t, u(t), x(t)) = 0.$$

Introduce the notation

$$p(t) = -B_x(t, x(t)).$$

Then the previous equality takes the form

$$p'(t) = g_x(t, u(t), x(t)) - f_x(t, u(t), x(t))p(t).$$

Determine the function

$$H(t, v, y, p) = pf(t, v, y) - g(t, v, y).$$

Thus the function p satisfies the equation

$$p'(t) = -H_x(t, u(t), x(t), p(t)). \tag{4.43}$$

Besides, from the equality $B(T, x) = h(x)$ and the definition of the function p it follows the equality

$$p(T) = -h_x(x). \tag{4.44}$$

Problem (4.43) and (4.44) exactly coincides with the adjoint system obtained in accordance with the maximum principle. Conditions (4.41) and (4.42) also imply the inequality

$$R(t, u(t), x(t)) \geq R(t, w, x(t)) \;\; \forall w \in [a(t), b(t)]$$

that is,

$$B_t(t, x(t)) + B_x(t, x(t)) f(t, u(t), x(t)) + g(t, u(t), x(t))$$
$$\geq B_t(t, x(t)) + B_x(t, x(t)) f(t, w, x(t)) + g(t, w, x(t)) \;\; \forall w \in [a(t), b(t)].$$

Using the definition of the functions H and p, we obtain the equality

$$H(t, u(t), x(t), p(t)) = \max_{w \in [a(t), b(t)]} H(t, u(t), x(t), p(t)), \tag{4.45}$$

i.e., the maximum principle.

Thus, the optimal control is determined by relations (4.40), (4.42)–(4.45), which exactly coincide with the system of optimality conditions obtained in Chapter 3 based on the maximum principle.

4.2.8 Applicability of Bellman optimality principle

Let us check the validity of the Bellman optimality principle for Example 3.1. Here, we minimize the functional

$$I(u) = \int_0^1 \Big(\frac{u^2}{2} - 3x\Big)dt$$

on the set

$$U = \Big\{u \big| \ 1 \le u(t) \le 2, \ 0 < t < 1\Big\},$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t), \ t \in (0,1); \ x(0) = 0.$$

Chapter 3 found its unique solution

$$u_\varepsilon(t) = \begin{cases} 2, & \text{if } 0 < t < 1/3, \\ 3 - 3t, & \text{if } 1/3 \le t \le 2/3, \\ 1, & \text{if } 2/3 < t < 1. \end{cases} \tag{4.46}$$

Let us now consider a family of problems that differ from the one above only in the initial time ξ and the initial state x. In this case, it is required to minimize the functional

$$I_\xi^x(v) = \int_\xi^1 \Big(\frac{v^2}{2} - 3y\Big)dt$$

on the set

$$U(\xi) = \Big\{v \big| \ 1 \le v(t) \le 2, \ \xi < t < 1\Big\},$$

where y is a solution of the Cauchy problem

$$y'(t) = v(t), \ t \in (\xi,1); \ y(\xi) = 0.$$

Using the maximum principle, we determine the function

$$H = pv - v^2/2 + 3y,$$

where p is a solution of the adjoint system

$$p'(t) = -3, \ t \in (\xi,1); \ p(1) = 0.$$

The corresponding optimal control is found from the maximum condition for the function H with respect to the control. The result is the formula

$$v(t) = \begin{cases} 1, & \text{if } p(t) < 1, \\ p(t), & \text{if } 1 \le p(t) \le 2, \quad t \in (\xi,1), \\ 2, & \text{if } p(t) > 2, \end{cases}$$

similar to the one that was obtained in Chapter 3 when analyzing Example 3.1. Determine $p(t) = 3–3t$. Obviously, this function decreases, and takes the value 2 at $t = 1/3$ and the value 1 at $t = 2/3$. Thus, we find

$$v(t) = \begin{cases} 2, & \text{if } 0 < t < 1/3, \\ 3 - 3t, & \text{if } 1/3 \le t \le 2/3, \quad t \in (\xi, 1), \\ 1, & \text{if } 2/3 < t < 1. \end{cases}$$

Naturally, for $\xi > 2/3$ the control can only take the value 1, for $1/3 < \xi < 2/3$ the control is equal to $3–3t$ for $\xi < t \le 2/3$ and equal to 1 for $2/3 < t < 1$. Finally, for $\xi < 1/3$ the control takes the value 2 for $\xi < t < 1/3$, $3–3t$ for $1/3 \le t \le 2/3$ and 1 for $2/3 < t < 1$. Thus, the control u, which is optimal in Example 3.1 with a time interval (0,1), remains optimal when considering the problem on an arbitrary finite interval $(\xi, 1)$. Thus, the Bellman optimality principle is indeed realized[47].

Additional conclusions

Based on the results presented in Appendix, some additional conclusions can be drawn about methods for solving optimal control problems.

- A situation is possible when the optimality criterion turns out to be Gateaux non-differentiable, as a result of which there is a need to use methods of non-smooth optimization.
- The variational inequality is, in generally, a necessary but not sufficient optimality condition, i.e., a control that satisfies a variational inequality is not always optimal.
- For Example 3.2, the maximum principle turns out to be more efficient than the variational inequality, since the set of solutions of the corresponding variational inequality turns out to be much wider and includes obviously non-optimal controls.
- The penalty method can also be used to solve optimal control problems with constraints on control values, and in this case it is used together with the variational inequality.
- In the absence of a unique solvability of the equation of state on the entire set of admissible controls, the optimal control problem makes sense, and the optimality criterion is minimized on the set of admissible pairs of the system, i.e., "control-state" pairs, on which the equation of state makes sense and the specified constraints on the system are satisfied.
- To solve problems of optimal control of singular systems in the absence of a unique solvability of the equation of state on the entire set of admissible controls, it is not possible to use the maximum principle or variational inequalities. This is explained by the fact that when deriving the optimality conditions in

the process of variation of the control, we can reach a control under which the equation of state either does not have, or has a non-unique solution.

- For an approximate solution of problems of optimal control of singular systems, one can use the penalty method, in which the equation of state is interpreted as an equality constraint imposed on the system.
- The Bellman equation can be justified.
- There exists a relation between the dynamical programming and maximum principle.
- The Bellman optimality principle holds for Example 3.1.

Notes

1. In the calculus of variations, the function v is called the ***extremal variation***; see [37], [61], [208]. An ***extremal*** is a smooth solution of the most important necessary condition for the extremum of the calculus of variations that is the ***Euler equation***.

2. In fact, we have here the adjoint system (3.14).

3. In the calculus of variations, this limit is called the ***functional variation***; see [37], [61], [208].

4. More precisely, the **Gateaux derivative** of an arbitrary functional I defined on a normed vector space V at a point u is a linear continuous functional $I'(u)$ (this is an element of the ***adjoint space*** of V; see [94], [106], [158]), which satisfies equality

$$\lim_{\sigma \to 0} \frac{I(u + \sigma h) - I(u)}{\sigma} = I'(u)h \;\; \forall h \in V.$$

5. For differentiation of functionals and general operators; see [94], [158], [171]. Differentiation in non-normed spaces is considered in [71].

6. This follows from the ***Lagrange–Euler lemma***; see [37], [61], [208], or the statement that a linear continuous functional that maps an arbitrary function to zero is the zero element of the corresponding dual space; see [94], [158].

7. A necessary condition for a local extremum of an arbitrary differentiable functional at some point is that its Gateaux derivative vanishes at that point, which corresponds to the general form of the ***stationary condition***; see, for example, [60], [116], [171]. From this, one can derive the stationary condition for a function of one and many variables, obtained in Chapter 1, as well as a significant number of extremum conditions obtained in the calculus of variations. The stationary condition for the functionals will be used in Chapter 10 to solve one optimal control problem for a system with a fixed final state.

8. For the gradient method for solving functional minimization problems; see [49], [70], [171], [194].

9. This is true, of course, under the differentiability of the functions included in the problem statement.

10. In fact, we use here the convex property of the set U. It is natural to call the corresponding form of control variation the ***convex variation***.

11. We again actually obtain a relation that is valid almost everywhere on the considered time interval.

12. The variational inequality naturally extends to the problem of minimizing a Gateaux differentiable functional I in the on a convex subset U of a topological vector space; see [60], [70], [116], [171]. In particular, for the point u to be a solution to this problem, it is necessary that the derivative $I'(u)$ satisfy the ***variational inequality*** $I'(u)(v–u) \geq 0$ for any point v from the set U. Of course, we did not give the complete proof of the considered theorem.

13. Other examples of the coincidence of optimality conditions in the form of the maximum principle and variational inequalities will be described in Chapters 6, 9, and 13. In Appendix, as well as in the subsequent section, we will make sure that this situation is not always observed.

14. We will meet the concepts of an optimal and admissible pair (in this case, any pair of functions $(u_\varepsilon, x_\varepsilon)$ satisfying the equality $x(0) = 0$) in Part V when considering optimal control problems for systems with a free initial state.

15. The conditions imposed on the function h are such that the function $x_\varepsilon + \sigma h$ has the properties of the state function of the system, in particular, satisfying the given initial conditions. For this, it is required that $h(0) = 0$.

16. In fact, this system was obtained from the condition that the Gateaux derivative of the functional J_ε vanishes, which corresponds to the stationary condition for the problem of its minimization.

17. To substantiate this assertion, it is necessary, of course, to prove that for $\varepsilon \to 0$ the solution to the problem of minimizing the functional J_ε converges to the optimal control for Example 3.4. We also note that an approximate solution of the problem here means a pair (u, x), that satisfies the equation of state with a sufficiently high degree of accuracy. For various forms of approximate solution of optimal control problems; see Chapter 8. The convergence of the penalty method for one optimal control problem is proved in Chapter 10. In addition, to find the optimal control in practice, one should also find a solution to system (4.8), (4.10), and (4.11). To do this, you can use the method of successive approximations described in Chapter 3.

18. For dynamic programming; see, for example, [28], [29], [34], [42], [62], [70], [103], [179], [193], [195].

19. In other words, the restriction of an optimal control to any finite part of its domain of definition is itself optimal. Chapter 13 gives an example of an optimal control problem for which the Bellman optimality principle does not hold.

20. We suppose here the differentiability of the Bellman function that is obviously.

21. The Bellman equation is an analog of the well-known ***Hamilton–Jacobi equation*** of the calculus of variations; see [37], [61], [208].

22. When we solved the same Problem 3.1 using maximum principle, we had the adjoint equation with final condition too.

23. The proof of the Theorem 4.2 is given in Appendix; see also [28], [29], [34], [70], [193], [195].

24. In the calculus of variations, ***Weierstrass*** and ***Legendre sufficient extremum conditions*** are used; see [37], [61]. For ***sufficient optimality conditions*** for optimal control theory; see [108], [193]. Note also that Theorem 4.2 assumes the existence of the Bellman function and its smoothness.

25. This corresponds to the absence of the need for an optimality condition in the general case.

26. For the convergence of iterative methods for solving optimal control problems; see [49], [46], [65], [149].

27. Justification of the variational inequality for abstract functional minimization problems is given in [60], [116], [171]. In [73], [116], [118], [171], variational inequalities are used to solve optimal control problems for systems with distributed parameters.

28. The ***penalty method*** for optimal control problems is for systems described by ordinary differential equations (see [32], [54]), for partial differential equations (see [118], [138], [168]), for integral equations (see [206]), for systems described by variational inequalities (see [22]), for multiextremal problems (see [207]).

29. For application of the penalty method to the analysis of distributed singular systems, see [118]. Optimal control problems for singular systems are also considered in [73], [72], [168].

30. Justification of the Bellman equation is given; for example, in [28], [29], [67], [193], [195].

31. References [28], [29], [67], [193], [195] give examples of applying dynamic programming to solve various optimal control problems.

32. For the connection between dynamic programming and the maximum principle; see, for example, [193], [195].

33. Example 4.1 goes back to one inverse problem in geophysics; see [175]. In this case, the system is described by the ***Poisson equation*** with respect to the potential of the gravitational field. The right side of the equation includes the distribution of the medium density. In this case, the medium is inhomogeneous and consists of two different materials, as a result of which the density turns out to be a piecewise constant function (an analog of the function f considered in the example with the values f_1 and f_2). It is required to determine the interface between these materials based on the results of measuring the derivative of the potential field of the region. The desired boundary ("control" u) is determined from the problem of minimizing the corresponding root-mean-square deviation (functional I).

34. Indeed the function Δx is a solution of the Cauchy problem

$$\Delta x'(t) \;=\; f(t,v) - f(t,u),\ t \in (0,1);\ \Delta x(0) = 0.$$

Multiplying the first equality by an arbitrary differentiable function λ and integrating the result, taking into account the second equality, we obtain

$$\int\limits_0^1 \lambda\big[f(t,v) - f(t,u)\big]dt \;=\; \int\limits_0^1 \lambda\Delta x' dt \;=\; \lambda(1)\Delta x(1) - \int\limits_0^1 \lambda'\Delta x dt.$$

Choose here as λ the solution of the problem

$$p'(t) \;=\; x(t) - z(t),\ t \in (0,1);\ p(1) = 0.$$

We get

$$\int_0^1 p\big[f(t,v)-f(t,u)\big]dt = -\int_0^1 (x-z)\Delta x dt.$$

Then equality (4.19) takes the form (4.20).

35. For methods of non-smooth optimization; see, for example, [47], [48], [60], [133], [132], [139]. Note that the lack of differentiability of the functional in the sense of Gateaux can be due not only to the presence of non-smooth terms of the module type, but also to the insufficiently strong convergence of the corresponding value for $\sigma \to 0$; see [171].

36. The variational inequality is a necessary and sufficient optimality condition for convex functionals; see [116]. In particular, the minimized functional for Example 3.1 turns out to be convex (this will be established in Chapter 5), while the corresponding functional for Example 3.2, which differs from the previous one only in sign, is not convex. In Sections 9 and 10, for analog of Examples 3.1 and 3.2 in the presence of a fixed final state, the maximum principle and variational inequality will be applied. At the same time, both approaches turn out to be equivalent in the minimization problem, and in the maximization problem, the maximum principle turns out to be more effective.

37. When $u_\varepsilon(t)–p_\varepsilon(t) > 0$, it follows from (4.25) that $v \geq u_\varepsilon(t)$ for all $v \in [-1,1]$. Then $u_\varepsilon(t) = -1$, and this equality is true for $p_\varepsilon(t) < -1$. When $u_\varepsilon(t)–p_\varepsilon(t) < 0$, it follows from (4.25) that $v \leq u_\varepsilon(t)$ for all $v \in [-1,1]$. Then $u_\varepsilon(t) = 1$, and this equality is valid for $p_\varepsilon(t) > -1$. Finally, the case $u_\varepsilon(t)–p_\varepsilon(t) = 0$ is possible, which means that $u_\varepsilon(t) = p_\varepsilon(t)$. However, this case is admissible for $p_\varepsilon(t) \in [-1,1]$. As a result, we get the formula (4.26).

38. This corresponds to the well-posedness of the state equations. We will consider the well-posedness of optimal control problems in Chapter 8.

39. In this situation, we have only about the existence of a ***local solution*** for the Cauchy problem; see [86].

40. Naturally, this also requires justification of the penalty method, i.e., proof that the solution to the problem of minimizing the functional J_ε converges to the optimal control for Example 3.4. It is also desirable to show that both of these problems have a solution. For the considered equation of state, one could consider an optimal control problem with restrictions on the control values. In this case, to minimize the functional J_ε, a variational inequality is established similarly to how it was done in Section 4.2.2. The penalty method can also be used to find the singular control in practice; see Chapter 6. In Chapter 10, the penalty method is used to solve the optimal control problem for a system with a fixed final state, and in this way not the equation of state is removed, but an additional condition characterizing the final state of the system. At the same time, the convergence of the method for one specific example is proved. A justification for the penalty method for a wide class of systems described by singular partial differential equations is given in [118].

41. A similar problem in the absence of an initial condition will be considered in Chapter 14.

42. This differential equation is considered in [88] without regard to optimal control theory.

43. Note, first of all, that from the definition of the set of admissible controls it follows that the derivative of the function x is not negative. From here and from the initial condition it follows that this function cannot take negative values, as a result of which the value under the root in the state equation is not negative.

44. Obviously, the state equation on a given control has even a non-countable set of solutions. In Part V, optimal control problems will be considered for systems described by ordinary differential equations in the absence of initial conditions. In this case, each control corresponds to a general solution to the equation, the definition of which includes arbitrary constants. Thus, the optimality criterion is also minimized on a set of pairs of "control-state" functions related by the state equation.

45. Naturally, all this needs justification. However, based on the formulation of the problem, it is easy to see that it has a unique solution equal to zero.

46. Here, it is assumed that the corresponding derivatives exist.

47. In Chapter 10, an example will be given where the validity of the Bellman principle of optimality is realized for a problem with a fixed final state, and in Chapter 15, examples of violation of the Bellman principle for problems with an isoperimetric condition will be given.

CHAPTER 5

Uniqueness of the optimal control and sufficiency of optimality conditions

Part II discusses methods for studying optimal control problems. In particular, in Chapter 3, the maximum principle was used, which consists in maximizing, on the set of admissible control values, a certain function depending on the control, the state function, and the solution of the ad-joint system. The optimal control there was found directly or by some iterative process. In this chapter, we study a rather simple example, for which the described technique encounters certain difficulties. In particular, we do not have the uniqueness of the optimal control and the sufficiency of the optimality conditions. The Lecture analyzes this example, and Appendix provides some additional information about the studied problems.

5.1 LECTURE

In Chapter 3, the simplest optimal control problem was considered, for which the maximum principle was applied. The corresponding optimal controls were found analytically in the case of Examples 3.1, 3.2, and 3.4 or by iterative process for Example 3.3. Below we consider a problem that differs from Example 3.3 only in the type of extremum; see Section 5.1.1. In Section 5.1.2, a system of optimality conditions is obtained. For its analysis, Section 5.1.3 uses the iterative process described in Chapter 3. In this case, two solutions are found corresponding to different initial approximations. Section 5.1.4 gives sufficient conditions for the uniqueness of optimal control and shows that the examples of optimal control problems considered in Chapter 3 have unique solutions. The study of the considered example is finished in Section 5.1.5, where the complete set of solutions of the corresponding system of optimality conditions is determined and optimal controls are found.

DOI: 10.1201/9781003398585-5

5.1.1 Problem statement

Continuing the study of optimal control problems based on the maximum principle, consider an example[1] that is fairly close to Example 3.3.

Example 5.1 *The optimal control problem is to find such a function $u = u(t)$ from the set*

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\}$$

that maximizes there the functional

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{5.1}$$

This problem differs from Example 3.3 only in the type of extremum. To study it, we will use the same technique as in the previous chapter. As already noted, a control that minimises the functional $-I$ will certainly maximises the functional I. Therefore, if to minimize the functional I it is required to find the maximum of the function H, then the search for the maximum of the same functional is reduced to minimizing the same function H.

5.1.2 Maximum principle

To bring the problem to the standard form described in Chapter 3, we define

$$f(t,u,x) = u,\ T = 1,\ x_0 = 0,\ a(t) = 1,\ b(t) = 2,\ g(t,u,x) = (u^2 + x^2)/2,\ h(x) = 0,$$

which is exactly the same as in Example 3.3. In accordance with formula (3.3), we determine the function

$$H = pu – (u^2 + x^2)/2.$$

Then the adjoint system (3.8), (3.9) takes the form

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0. \tag{5.2}$$

This is the problem (3.14). Since in this case the maximization of the given functional is carried out, and not its minimization, as in the previous case, the maximum condition (3.12) represents equality

$$H(u) = \min_{|v|\leq 1} H(v). \tag{5.3}$$

Thus, to find three unknown functions u, x, p, we have three conditions[2] (5.1) – (5.3). They differ from the system (3.13) – (3.15) considered in the previous section only in the type of extremum of the function H.

We find a solution to problem (5.3). Equating to zero the derivative of the function H with respect to control, we have

$$\frac{\partial H}{\partial u} = p–u = 0.$$

It follows that the function H has a local extremum point[3] $u = p$. Since the second derivative of H is negative, we have its maximum, not minimum. This means that the given function has no local minima at all. Thus, the minimum of the function H can be achieved only on the boundary of the set of admissible controls[4].

Check the values of this function on the boundary of this set, i.e., at points 1 and −1. We find

$$H(1) = p–(1+x^2)/2, \; H(-1) = -p–(1+x^2)/2.$$

The solution of minimum condition (5.3) corresponds to the smallest of these two values. They differ only in the sign of the first term, which is a variable, in particular, the solution of adjoint system (5.2). Thus, we obtain the following formula

$$u(t) = \begin{cases} 1, & \text{if } p(t) < 0, \\ -1, & \text{if } p(t) > 0. \end{cases} \tag{5.4}$$

Formula (5.4) allows us to find the control by the known solution of the adjoint system; see Figure 5.1.

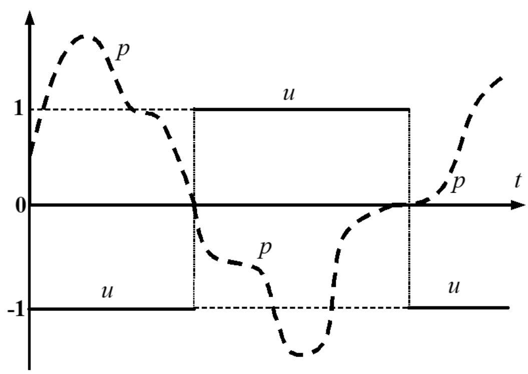

Figure 5.1 Solution of minimum condition (5.3).

Thus, to find three unknown functions u, x, and p we have system (5.1), (5.2), and (5.4). Its analytical solution is difficult in view of the non-linear dependence of the function u on p by equality (5.4). However, this problem can be solved iteratively in the same way as it was done in the analysis of Example 3.3.

5.1.3 Analysis of optimality conditions

To solve the resulting system of optimality conditions, we use the method of successive approximations[5] described in the previous section. The initial approximation of the control is set. At the current iteration, given the known control from the Cauchy problem (5.1), the corresponding system state function is found. Then the adjoint system (5.2) is solved. After that, according to formula (5.4), a new control approximation is determined.

In Chapter 3, when analyzing Example 3.3, it was noted that the choice of the initial approximation is due to the fact that the control must belong to a given set, i.e., we have some a priori information about it. In this case, we have additional information, since according to formula (5.4) the control can take only two values 1 and -1, the choice of which is determined by the sign of the function p.

We choose one of the possible control values, for example, the first of them as an initial approximation, i.e., determine

$$u_0(t) = 1,\ t \in (0,1).$$

Substituting this control into a problem (5.1), we find the state function

$$x_0(t) = \int_0^t u_0(\tau)d\tau = t.$$

In accordance with problem (5.2), we integrate this equality from an arbitrary value t to 1. We get

$$p_0(t) = -\int_t^1 x_0(\tau)d\tau = -\int_t^1 \tau d\tau = \frac{t^2-1}{2}.$$

It follows that the function $p_0 = p_0(t)$ at $t \in (0.1)$ takes exclusively negative values. Then, according to formula (5.4), at the next iteration, we obtain the control u_1, which is identically equal to 1. Thus, the considered iterative process converged in one iteration. The results obtained indicate that the triple of functions

$$u(t) = 1,\ x(t) = t,\ p(t) = (t^2{-}1)/2,\ t \in (0,1) \tag{5.5}$$

is the solution of problem[6] (5.1), (5.2), (5.4).

Having determined the solution of the system of optimality conditions, it would seem that we can complete the research. However, in Chapter 2, we encountered a situation where a change in the initial approximation led to finding a new solution to the problem[7]. Note that, according to formula (5.4), the control could also be equal to the value -1.

Thus, now let us try to choose as the initial approximation the function

$$u_0(t) = -1,\ t \in (0,1).$$

The corresponding solution of Cauchy problem (5.1) is

$$x_0(t) = \int_0^t u_0(\tau)d\tau = -t.$$

Solving the problem (5.2), we find

$$p_0(t) = -\int_t^1 x_0(\tau)d\tau = \int_t^1 \tau d\tau = \frac{1-t^2}{2}.$$

It follows that the function $p_0 = p_0(t)$ at $t \in (0.1)$ takes exclusively positive values. Then, in accordance with formula (5.4), a new control approximation $u_1(t) = -1$ is found. Considering that it coincides with the value of the control at the previous iteration, we conclude that the iterative process has converged again. Therefore, the functions

$$u(t) = -1,\ x(t) = -t,\ p(t) = (1-t^2)/2,\ t \in (0,1) \tag{5.6}$$

satisfy the system of optimality conditions (5.1), (5.2), (5.4).

Thus, we have found two solutions to the optimality conditions. We have already encountered a similar situation in Chapter 1 when studying function minimization problems. In particular, for Example 1.2, only one of the solutions to the stationary condition minimized the considered function, which means that the extremum condition turned out to be insufficient. At the same time, in Example 1.3, there were two minimum points of the function, i.e., the solution to the problem was not unique. Thus, we have to find out whether the two found solutions of the optimality conditions are equal, and then we may be dealing with the absence of uniqueness of the solution to the problem, or these solutions are not equal, which means that the optimality condition is certainly not sufficient.

Calculate the corresponding values of the functional

$$I(1) = \frac{1}{2}\int_t^1 (1+t^2)dt = \frac{2}{3},\ \ I(-1) = \frac{1}{2}\int_t^1 [(-1)^2 + (-t)^2]dt = \frac{2}{3}.$$

Thus, the value of the functional on both found controls is the same, i.e., both of these values are completely equal. The results obtained suggest that the solution of the optimal control problem for the considered example is not the only unique[8].

5.1.4 Uniqueness of the optimal control

The very fact that some extremal problems have a unique solution, while others do not, is not particularly surprising[9]. A kind of analog of the optimal control problems from Examples 3.3 and 5.1 is the study of the extremum of the quadratic function $f(x) = x^2$. It has one minimum point, but two maximum points on the segment $[-1,1]$; see Figure 5.2.

In Chapter 1, a condition was established that guarantees the uniqueness of the minimum point of the function. According to Theorem 1.4, a strictly convex function on a segment cannot have two minimum points, i.e., the solution to the problem, if it exists, is necessarily unique. By Definition 1.5, a function f is strictly convex on the segment $[a,b]$ if the following relation holds

$$f[\alpha x + (1-\alpha)y] \le \alpha f(x) + (1-\alpha)f(y)\ \ \forall x,y \in [a,b],\ \alpha \in (0,1), \tag{5.7}$$

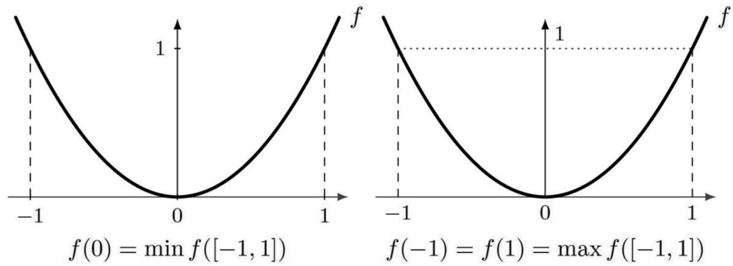

Figure 5.2 The parabola on the segment [-1,1] has one minimum, but two maximums.

and equality here is realized only when $x = y$. In particular, a strictly convex function $f(x) = x^2$ has a unique minimum point, while the minimization problem for a non-convex function $g(x) = -x^2$ on the interval $[-1, 1]$, which is equivalent to the maximization problem there function f, has two solutions; see Figure 5.2.

It was noted earlier that Theorem 1.4 can be extended to functions of many variables. We would like to establish conditions that guarantee the uniqueness of the minimum point of an arbitrary functional I on a set U. It is necessary first define strict convexity for functionals. However, in order to obtain some analog of inequality (5.7), it is necessary to reveal certain properties of the given set under which this inequality makes sense.

Thus, we pass from the segment $[a, b]$ to the set U of a general form, i.e., points x and y are chosen from this set. On the left side of inequality (5.7), they are first multiplied by the numbers α and $(1 - \alpha)$, and then the results are added. Therefore, on the set U, the operations of addition and multiplication by a number must at least make sense. This is certainly true for the Euclidean space, where the operations of vector addition and multiplication of vectors by numbers are defined. Sets with similar properties are called ***vector*** or ***linear spaces***[10].

Note that the segment $[a, b]$ itself is not a vector space[11]. It represents only a subset of the number line, which is the vector space. Thus, in order to extend the uniqueness theorem to the consideration problems, the corresponding set U must be a subset of the vector space. Moreover, in order for us to replace the segment in inequality (5.7) with a set of a general nature, the latter must have an additional property: the argument of the function (now a functional) on the left side of the relation under consideration must surely belong to this set. As a result, we have the following definition, which is a generalization of Definition 1.9.

Definition 5.1 *A subset U of a vector space is called* ***convex***[12] *if for all its points x and y it contains the points $\alpha x + (1–\alpha)y$ for any $\alpha \in [0, 1]$.*

Now we can give the definition of the functional convexity[13].

Definition 5.2 *A functional I defined on a convex subset U of a vector space is called* ***convex*** *if the following inequality holds*

$$I[\alpha z + (1 - \alpha)y] \leq \alpha I(x) + (1 - \alpha)I(y) \;\; \forall x, y \in U, \; \alpha \in (0, 1).$$

If this condition is satisfied in the form of equality only for $x = y$, then the functional is called ***strictly convex****.*

Now we can formulate the uniqueness theorem for the minimum of the functional[14].

Theorem 5.1 *A strictly convex functional on a convex subset of a vector space cannot have two minimum points.*

Proof. Suppose that there are two distinct points x and y of the minimum of a strictly convex functional I on a convex set U. Then the element $\alpha x + (1-\alpha)y$ for any number $\alpha \in (0,1)$ due to the convexity of the set U belongs to this set. Taking into account the strict convexity of the functional, we establish the inequality

$$I[\alpha x+(1-\alpha)y] \; < \; \alpha I(x)+(1-\alpha)I(y) \; = \; \alpha \min I(U)+(1-\alpha)\min I(U) \; = \; \min I(U).$$

Therefore, there is such an element of the set U, on which the value of the functional is less than the minimum possible value. Thus, the assumption of the existence of two different minimum points led to a contradiction. □

To use this result in the study of optimal control problems, it must be borne in mind that the dependence of the optimality criterion on the control is characterized both explicitly and by the state equation[15]. This circumstance introduces certain difficulties, but does not serve as an insurmountable obstacle. However, to obtain the desired result, it is required here to set some properties of the dependence of the system state function on the control.

Let u_1 and u_2 be diverse admissible controls, and x_1 and x_2 are the corresponding solutions of the system (5.1). Then we obtain

$$x_i'(t) \; = \; u_i(t), \; t \in (0,1), \; x_i(0) = 0, \; i = 1,2.$$

Denoting by y the solution of problem (5.1) corresponding to the control $v = \alpha u_1 + (1-\alpha)u_2$, we have

$$y'(t) \; = \; v(t), \; t \in (0,1), \; y(0) = 0.$$

Based on the results obtained, determine the following equality[16] $y = \alpha x_1 + (1-\alpha)x_2$.

Find the integral

$$I(v) \; = \; \frac{1}{2}\int_0^1 (v^2 + y^2)dt.$$

Due to the strict convexity of the quadratic function, for any value of $t \in (0,1)$, the following inequalities hold

$$v(t)^2 \; = \; [\alpha u_1(t) + (1-\alpha)u_2(t)]^2 \; < \; u_1(t)^2 + (1-\alpha)u_2(t)^2,$$

$$y(t)^2 \; = \; [\alpha x_1(t) + (1-\alpha)x_2(t)]^2 \; < \; x_1(t)^2 + (1-\alpha)x_2(t)^2.$$

After integration, we get

$$I[\alpha u_1 + (1-\alpha)u_2] = \frac{1}{2}\int_0^1 (v^2 + y^2)dt <$$

$$< \frac{\alpha}{2}\int_0^1 \Big[(u_1)^2 + (x_1)^2\Big]dt + \frac{1-\alpha}{2}\int_0^1 \Big[(u_2)^2 + (x_2)^2\Big]dt = \alpha I(u_1) + (1-\alpha)I(u_2).$$

Thus, the functional $I = I(u)$ is indeed strictly convex. Then it follows from Theorem 5.1 that the problem of its minimization on the set U, i.e., the optimal control problem considered in Example 3.3 does indeed have a unique solution[17]. The problem from Example 3.4 also has a similar property, differing from the previous one only in the absence of restrictions on control. At the same time, the problem of its maximization, which is the subject of Example 5.1, no longer has this property. As a result, the result obtained earlier on the coincidence of the values of the functional on two solutions of the optimality conditions seems to be quite natural[18].

In Example 3.1, the state equation is the same, but the functional has the form

$$I = \int_0^1 \Big(\frac{u^2}{2} - 3x\Big)dt.$$

It is the sum of a quadratic functional with respect to the control and a linear functional with respect to the state. We determined before the relations

$$v(t)^2 < \alpha u_1(t)^2 + (1-\alpha)u_2(t)^2, \; y(t) = \alpha x_1(t) + (1-\alpha)x_2(t).$$

Dividing the first inequality by 2 and subtracting the second equality multiplied by 3, after integration we come to the conclusion that the functional to be minimized is strictly convex, and hence that the solution of the problem is unique.

The optimality criterion in Example 3.2 differs only in sign from the one considered above. Then it follows from the last inequality that this functional is not convex. Nevertheless, the optimal control here is unique, which does not contradict the Theorem 5.1, which gives only sufficient conditions for the uniqueness of the solution of the problem[19].

5.1.5 Completion of the analysis of optimality conditions

Having established the absence of uniqueness of the optimal control for Example 5.1, we again return to system (5.1), (5.2), and (5.4). It is natural to ask the question: does this system have other solutions?

As can be seen from formula (5.4), the control can take only two values, depending on the sign of the function p. At first glance, we have already considered all possible cases[20]. However, we recall that p is a function, which means that, in principle, it can change sign somewhere. Suppose that there is such a point ξ from the interval (0,1) that for $t < \xi$ the function p is positive, and for $t > \xi$ it is negative. According to equality (5.4), we choose the corresponding function u as the initial control. Thus, we set the function

$$u_0(t) = \begin{cases} 1, & \text{if } t < \xi, \\ -1, & \text{if } t > \xi, \end{cases}$$

The solution of the problem (5.1) for $u = u_0$ is $x_0(t) = t$ for $t < \xi$. Then $x_0(\xi) = \xi$. For $t > \xi$, we get

$$x_0(t) = x_0(\xi) - \int_\xi^t d\tau = 2\xi - t.$$

Therefore, the state of the system at the zero iteration is determined by the formula

$$x_0(t) = \begin{cases} t, & \text{if } t < \xi, \\ 2\xi - t, & \text{if } t > \xi. \end{cases}$$

Since the adjoint state is known at the final time, the solution to the problem (5.2) is first determined for $t > \xi$.

$$p_0(t) = -\int_t^1 (2\xi - \tau)d\tau = 2\xi(t-1) + (1-t^2)/2 = (1-t)\big[(1+t)/2 - 2\xi\big].$$

We are interested in such a function p, which at the point $t = \xi$ changes sign, i.e., goes to zero. As a result, we obtain the formula

$$p_0(\xi) = (1\text{–}\xi)(1\text{–}3\xi)/2 = 0.$$

This is the square equation with respect to the parameter ξ. It has two solutions: $\xi = 1$ and $\xi = 1/3$. The first of these is trivial, since it is known from the outset that the function p vanishes at a finite time. Of particular interest is the value $\xi = 1/3$, which belongs to the given time interval.

Thus, there is a unique point $\xi = 1/3$ at which the solution to problem (5.2) can vanish. Solving this problem from an arbitrary time $t \in (0, 1/3)$ to 1/3, we find the solution of the adjoint system

$$p_0(t) = \begin{cases} -(1/3 - t)(t + 1/3)/2, & \text{if } t < 1/3, \\ (1-t)(t+1/3)/2, & \text{if } t > 1/3. \end{cases}$$

Obviously, the function p_0 takes negative values for $t < 1/3$ and positive values for $t > 1/3$. Then, in accordance with formula (5.4), the control at the next iteration is

$$u_1(t) = \begin{cases} 1, & \text{if } t < 1/3, \\ -1, & \text{if } t > 1/3. \end{cases}$$

This value coincides with the function u_0 at $\xi = 1/3$. Thus, the iterative process again converged in one iteration. This means that the system of optimality conditions (5.1), (5.2), (5.4) has one more solution, already the third in a row. Characteristically, unlike the first two solutions, the corresponding control is a discontinuous function.

The appearance of the third solution of the optimality conditions, which differs significantly from the two previous ones, leads to certain thoughts. Moreover, it is easy to verify that the control, which differs from the latter only in sign, is also the solution to the system of optimality conditions. Thus, for problem (5.1), (5.2), and (5.4) there are four solutions. We already know that the value of the functional I on the first two found controls (identically equal to 1 and −1) coincide with each other. There is no doubt that the functionals on the last two controls coincide, as long as they differ only in sign. However, it is not at all clear whether the values of the optimality criterion coincide on the first and third of the found controls?

Let us find the value of the functional on the resulting discontinuous control. We have

$$I(u) = \frac{1}{2}\Big(\int_t^1 u^2 dt + \int_t^1 x^2 dt\Big) = \frac{1}{2}\Big[1 + \int_1^{1/3} t^2 dt + \int_0^{1/3} \Big(\frac{2}{3} - t\Big)^2 dt\Big] = \frac{1}{2}\Big(1 + \frac{1}{27}\Big) = \frac{14}{27}.$$

This value does not coincide with what was established earlier, which means that on different solutions of the system of optimality conditions, this functional takes on different values. Thus, some functions that satisfy the maximum principle (in particular, the third and fourth ones) turn out to be worse than others (the first and second ones)[21]. As a result, we can conclude that the maximum principle is not sufficient optimality condition[22].

We could conclude that the considered optimal control problem has two solutions (functions identically equal to 1 and −1) if we were sure that all solutions of the system of optimality conditions were found. However, is this true? We have found the first pair of solutions to this system, assuming that the function p is of constant sign. The second pair of solutions corresponds to the case where p changes sign somewhere. However, the situation is not excluded that this function changes sign twice. The corresponding control has two break points. Consider now a piecewise constant control having two discontinuity points ξ and η such that $0 < \xi < \eta < 1$. Thus, the function

$$u(t) = \begin{cases} 1, & \text{if } 0 < t < \xi, \\ -1, & \text{if } \xi < t < \eta, \\ 1, & \text{if } \eta < t < 1 \end{cases}$$

can be chosen as an initial iteration. The corresponding solution of problem (5.1) is

$$x(t) = \begin{cases} t, & \text{if } 0 < t < \xi, \\ 2\xi - t, & \text{if } \xi < t < \eta, \\ 2\xi - 2\eta + t, & \text{if } \eta < t < 1. \end{cases}$$

As a result of integrating the adjoint equation by the intervals (ξ, η) and $(\eta, 1)$, taking into account the fact that the function p must vanish at each of their boundaries, we obtain equalities

$$\int_\xi^\eta (2\xi - t)dt = 0, \quad \int_\eta^1 (2\xi - 2\eta + t)dt = 0.$$

As a result, we obtain

$$2\xi(\eta-\xi)-(\eta^2-\xi^2)/2 = 0, \quad (2\xi-2\eta)(1-\eta)+(1-\eta^2)/2 = 0.$$

Considering that the trivial solutions $\xi=\eta$ and $\eta=1$ do not suit us (under these conditions, we will not get two control break points), we establish the following system of equations for the control break points:

$$2\xi(\eta+\xi)/2 = 0, \quad (2\xi-2\eta)+(1+\eta)/2 = 0.$$

Therefore, there is a unique pair of points $\xi = 1/5$ and $\eta = 3/5$ that have the necessary properties.

Determine the solution of the adjoint system on each of the three obtained intervals. For $t \in (0, 1/5)$ we have[23]

$$p(t) = \int_t^{1/5} \tau d\tau = \frac{1}{2}\Big(t^2 - \frac{1}{25}\Big).$$

Obviously, all these values are negative. Further, at $t \in (1/5, 3/5)$ we get

$$p(t) = -\int_t^{3/5} \Big(\frac{2}{5}-\tau\Big)d\tau = \frac{1}{2}\Big(\frac{3}{5}-t\Big)\Big(t-\frac{1}{5}\Big).$$

Therefore, the function p is positive on the second interval. Finally, at $t \in (3/5, 1)$ we have

$$p(t) = -\int_t^{1} \Big(\tau-\frac{4}{5}\Big)d\tau = \frac{1}{2}\Big(t-\frac{3}{5}\Big)\Big(t-\frac{1}{5}\Big).$$

All obtained values of this function are negative.

Thus, with the above choice of control discontinuity points, the function sign of the corresponding function p is consistent with the behavior of the control, i.e., equality (5.4) actually holds. This means that when this control is chosen as the initial approximation, the iterative process converges in one iteration, and we again obtain the solution of the system of optimality conditions, already the fifth in a row. Naturally, the sixth solution is obtained from the fifth by changing the sign.

Similar reasoning can be carried out in a general form. For an arbitrary number k, we divide the segment [0,1] into $2k+1$ equal parts, $k = 0,1,...$ On the first of them, the control is chosen equal to 1, on the next two sections we set it equal to -1, on the next two it is equal to 1, then -1, etc. The resulting function is denoted by u_k^+. The corresponding solution x_k^+ to problem (5.1) is a piecewise linear function, and the solution p_k^+ to problem (5.2) is a differentiable function that changes sign at control discontinuity points; see Figure 5.3. The resulting triple of functions satisfies the considered system of optimality conditions. Along with this, there is also a solution $u_k^- = u_k^+$. Thus, each value of k (the number of discontinuity points) corresponds to exactly two controls that satisfy the optimality conditions. Thus, problem (5.1), (5.2), and (5.4) has an infinite set of solutions[24].

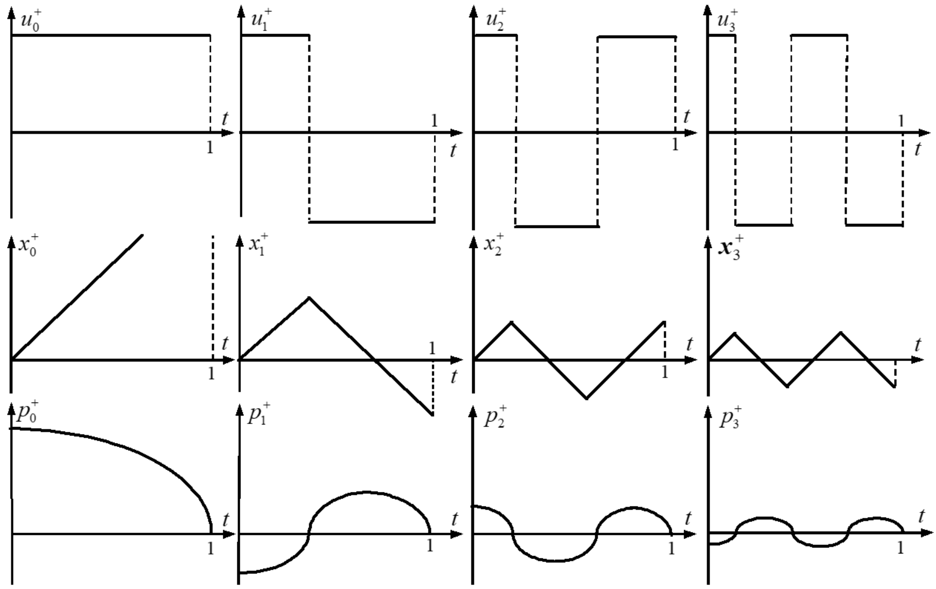

Figure 5.3 Solutions of optimality conditions for Example 5.1.

In order to find out which of these solutions is optimal, we find the values of the functional

$$I(u_k^+) = I(u_k^-) = \frac{1}{2}\int\limits_0^1 \Big[\Big(u_k^+\Big)^2 + \Big(x_k^+\Big)^2\Big]dt = \frac{1}{2} + \frac{2k+1}{2}\int\limits_0^{1/(2k+1)} \Big(x_k^+\Big)^2 dt =$$

$$= \frac{1}{2} + \frac{2k+1}{2}\frac{1}{3(2k+1)^3} = \frac{1}{2} + \frac{1}{6(2k+1)^2},\ k = 0, 1, \dots .$$

The results obtained indicate that with the growth of the parameter k, the value of the functional decreases. Therefore, the considered problem has exactly two solutions[25] that are the controls, identically equal to 1 and -1. Thus, despite the fact that the maximum principle (in this case, the minimum condition for the function H) is not a sufficient condition for optimality, and the set of its solutions is infinite, we were able to find the optimal control by choosing among all such solutions the one on which this criterion optimality takes the largest value[26].

RESULTS

Here is a list of questions on the uniqueness of optimal control and the sufficiency of the optimality condition in the form of the maximum principle, the main conclusions on this topic, as well as additional problems that arise in this case, partially solved in Appendix, partially taken out in the Notes.

Questions

It is required to answer questions concerning the properties of the optimal control problem and the maximum principle based on the results of the analysis of Example 5.1.

1. Why does the minimum condition for the function H turn out to be a necessary condition for the maximum of the functional?
2. Is it possible, without solving the problem (5.3), to establish the existence of its solution?
3. Is it possible, without solving problem (5.3), to establish the uniqueness of its solution?
4. Why, for Example 5.1, the unique stationary point of the function H is excluded from consideration, and its minimum can be reached only on the boundary of the set of admissible controls?
5. What class of functions does the solution of the optimality condition for Example 5.1 refer to in the general case?
6. What determines the choice of the initial approximation when solving the system of optimality conditions for Example 5.1?
7. On what basis was it concluded that the iterative process for the considered example converged in one iteration?
8. Why, having established the convergence of the iterative process to the solution of the system of optimality conditions, should we continue further analysis of the problem?
9. Why do the found solutions (5.5) and (5.6) of the system of optimality conditions differ only in signs?
10. Why, by changing the sign in solving the system of optimality conditions for Example 5.1, we get a new solution, but we do not get a new solution for Example 3.3, although the formulations of both optimization problems are invariant under the sign change for the "control-state" pair?
11. On the basis of what is the conclusion made that the solution of the considered problem is not unique, and how justified is this conclusion?
12. Why is it important in Theorem 5.1 that the functional is minimized on a subset of the vector space?
13. Why is there a requirement that the set of admissible controls be convex in the conditions of Theorem 5.1?
14. Why is the set of admissible controls from Example 5.1 convex?

15. Can an optimal control problem with a non-convex functional have a unique solution?

16. Why is the assumption of discontinuity of optimal control acceptable from both theoretical and practical points of view?

17. Why do we have the opportunity to find a solution to the considered optimal control problem, despite the fact that the system of optimality conditions has an infinite set of solutions?

18. Is it possible to conclude that the maximum principle for the considered example turned out to be effective?

Conclusions

Based on the study of the considered optimization methods, the following conclusions can be drawn.

- The optimal control problem for Example 5.1 differs from that considered in Example 3.3 only by the type of extremum.
- The solution of the functional maximization problem is reduced to the minimization of the corresponding function H.
- According to the maximum principle, the solution of this optimal control problem at any time can take only the values 1 and -1, corresponding to the boundaries of the set of admissible controls.
- The choice of the initial approximation for the iterative process for the optimality conditions should be carried out taking into account the results of the analysis of the maximum principle, i.e., from the class of functions that can take only the values 1 and -1.
- The iterative process for solving the system of optimality conditions with any initial approximation converges in one iteration.
- If a triple of functions u, x, p is a solution of the obtained optimality conditions, then the triple of functions that differ from the original one only in signs also turns out to be a solution of the optimality conditions.
- Any two admissible controls for Example 5.1, which differ only in signs, turn out to be equal in the sense that they correspond to the same value of the optimality criterion.
- The uniqueness of the optimal control is guaranteed for the problem of minimization of a strictly convex functional on a convex subset of a vector space.
- The system of optimality conditions for the considered example has an infinite set of solutions, two of which are continuous, two have one discontinuity point, two have two dis-continuity points, and so on.

- The choice of optimal controls from the set of solutions of the system of optimality conditions is carried out by comparing the values of the optimality criterion on these solutions.
- Optimal controls for the considered example are functions that are identically equal to 1 and −1.

Problems

Based on the results obtained above, we can have the following problems.

1. **Transformation of solutions of optimality conditions**. In the process of analyzing the optimality conditions for Example 5.1, it turned out that as a result of a sign change in their solution, a new solution to the system of optimality conditions is obtained[27]. Appendix explains why this result is true. In Chapters 11 and 14 we will have the problem of finding non-trivial transformations that translate one solution of a system of optimality conditions into another.

2. **Sufficiency of optimality conditions**. For the considered example, the maximum principle turns out to be a necessary but not sufficient condition for optimality. In Appendix, properties are established under which the optimality condition is both necessary and sufficient, and an example is given when sufficiency is realized even if these properties are violated. Some additional results in this direction will be given in the next chapter, see also Notes[28].

3. **Non-optimal solutions of optimality conditions**. As we already know, the stationary condition is also in the general case a necessary but not sufficient condition for the minimum of a function. At the same time, at stationary points that are not solutions of the minimization problem, the function invariably has some special properties. Appendix gives a result that characterizes the properties of non-optimal solutions of the optimality condition for the considered example.

4. **Sufficiency and uniqueness**. The optimal control problems considered in Chapter 3 had a unique solution, and the corresponding optimality conditions were necessary and sufficient. In Example 5.1, on the contrary, both the uniqueness of the solution and the sufficiency of the optimality conditions were absent. It would be interesting to give examples of problems where only one of the indicated properties was violated. An optimal control problem that admits non-uniqueness of the solution with sufficient optimality conditions is considered in Chapter 6. Chapters 6 and 14 describe examples of problems for which the optimal control is unique, but the optimality conditions are not necessary and sufficient.

5. **Variational inequalities**. In Chapter 2, for the problem of minimizing a functional on an interval, a variational inequality was used. In Chapter 3, this form of the optimality condition was applied to the analysis of Examples 3.1 and

3.3, for which it turned out to be equivalent to the maximum principle. Appendix presents the results of using the variational inequality for Example 5.1 under conditions of insufficiency of the maximum principle. Regarding the use of variational inequalities to study more difficult optimal control problems, see Notes[29].

6. **Boundary value problems for non-linear differential equations**. The system of optimality conditions for Example 5.1 includes a state equation, an adjoint equation, and control formulas derived from the maximum principle with respect to three unknown functions. In accordance with the method of eliminating unknowns, one can eliminate two unknown functions and establish a boundary value problem for a second-order differential equation with respect to the third function, as was done in the previous chapter. In Appendix, it is established that the resulting boundary value problem has very unexpected properties. An even more unexpected result for boundary value problems, to which optimality conditions are reduced, is obtained in Chapters 7, 11, and 12.

7. **Uniqueness and sufficiency when the conditions of the uniqueness and sufficiency theorems are violated**. In Chapter 1, it was noted that the properties established there that guarantee the uniqueness of the minimum point of the function and the sufficiency of the stationary condition are not necessary for the fact that the minimum point was unique, and the stationary condition was a necessary and sufficient condition for the minimum of the function. Appendix gives an example showing that similar results can be observed for optimal control problems.

5.2 APPENDIX

Some of the problems that arose during the analysis of Example 5.1 and need additional research are considered below. In particular, Section 5.2.1 establishes the reason why changing the sign in solving the system of optimality conditions leads to its new solution. Section 5.2.2 provides conditions that guarantee the sufficiency of the maximum principle and explains why in Examples 3.1, 3.2, and 3.3 the optimality condition was sufficient, but in Example 5.1 it was not. One result concerning the properties of non-optimal solutions of the maximum principle for Example 5.1 is given in Section 5.2.3. In Section 5.2.4, the analysis of the considered example is carried out on the basis of a variational inequality. In Section 5.2.5, the system of optimality conditions is reduced to a boundary value problem for a non-linear second-order differential equation with non-trivial properties. These solutions also turn out to be equilibrium positions for some non-linear heat conduction equation with the corresponding boundary conditions. In Section 5.2.6, we consider the optimal control problem, which differs from Example 5.1 only in the set of admissible control values. Nevertheless, for it the optimal control turns out to be unique, and the optimality conditions are necessary and sufficient, although the conditions of both theorems considered in this chapter are violated. Finally, the final subsection gives an example of a problem that has three solutions.

5.2.1 Invariance of the solution under sign change

Attention is drawn to the fact that the functions defined by equalities (5.5) and (5.6) and characterizing the first two solutions of the system of optimality conditions differ only in signs. Let us try to assess how natural this fact is. Assume that a triple of functions u, x, p is a solution to system (5.1), (5.2), and (5.4). Then the equalities hold

$$(-x)'(t) = (-u)(t),\ t \in (0,1);\ (-x)(0) = 0;$$
$$(-p)'(t) = (-x)(t),\ t \in (0,1);\ (-p)(1) = 0;$$
$$(-u)(t) = \begin{cases} 1, & \text{if } (-p)(t) < 0, \\ -1, & \text{if } (-p)(t) > 1. \end{cases}$$

It follows that the functions $-u$, $-x$, and $-p$ also satisfy this relation. Thus, if there is some solution to the system of optimality conditions, then, by changing the sign, we get a new solution of the same system[30]. This circumstance explains the fact that each previously obtained solution u_k^+ corresponds to a solution u_k^- that differs from it only in sign.

However, the question arises, why does the system of optimality conditions have this property? To do this, we should turn to the formulation of the optimal control problem. Consider an arbitrary admissible control u, which corresponds to a solution x of Cauchy problem (5.1). Then the control $-u$ is the element of the set of admissible controls. In this case, as already noted, the function $-x$ turns out to be a solution to problem (5.1) corresponding to the control $-u$. Then the following equalities hold

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt = \frac{1}{2}\int_0^1 \left[(-u)^2 + (-x)^2\right]dt = I(-u).$$

Thus, on the controls u and $-u$, the considered functional takes the same value. Thus, if the maximum of the functional I on the set U is reached on the control u, then on the control $-u$ this functional takes the same value, and hence $-u$ is also a solution to the problem. Therefore, the invariance of solutions of the optimality conditions with respect to sign change is a consequence of the very formulation of the problem[31].

The determining factors that ensure the above property of setting the optimal control problem under study are the invariance of the set of admissible controls with respect to sign change (if the control is admissible, then the control taken with the opposite sign also turns out to be admissible), the invariance of the equations of state with respect to sign change (if the ”control-state” pair satisfies the state equation, then the pair that differs from it only in sign also satisfies this equation) and the independence of the value of the functional from the sign of the control[32].

5.2.2 Sufficiency of the maximum principle

As we already know, in the examples considered in Chapter 3, the necessary optimality condition in the form of the maximum principle also turned out to be sufficient, i.e., any solution of the maximum principle is necessarily optimal. However, for Example 5.1 this is no longer the case, since only two of an infinite number of solutions to the system of optimality conditions turn out to be optimal. We have already encountered a similar problem when analyzing the stationary condition in the function minimization problem. Let us try to establish why the fulfillment of the maximum condition for an arbitrary control does not guarantee its optimality, and why in some cases sufficiency is nevertheless realized.

The derivation of optimality conditions in the form of the maximum principle in the previous chapter was carried out by us according to the following scheme. If an admissible control is optimal, then the corresponding increment of the functional must be non-negative. As a result of the transformation of the formula for the functional increment the maximum condition was obtained. Thus, from the optimality of a control, the validity of the maximum principle followed for it, which corresponds to the necessity of optimality conditions. The sufficiency of the optimality condition, on the contrary, implies that any of its solutions is necessarily an optimal control[33]. Thus, the absence of the sufficiency of the maximum condition means the irreversibility of the chain of reasoning connecting the assumption about the optimality of the control (reason) with the statement about the validity of the maximum principle (consequence). The question arises, at what stage of the derivation of optimality conditions was the reversibility of cause and effect violated? We minimize the functional

$$I(u) = \int_0^T g\big(t, u(t), x(t)\big)dt + h\big(x(T)\big)$$

on the set

$$U = \big\{u\big|\ a(t) \le u(t) \le b(t),\ t \in (0,T)\big\},$$

where x is a solution of the problem

$$x'(t) = f\big(t, u(t), x(t)\big),\ t \in (0,T);\ x(0) = x_0.$$

Let us return to the maximum principle derivation scheme described in the previous chapter.

The assumption of the optimality of the control u means the fulfillment of the inequality

$$I(v) - I(u) \ge 0\ \ \forall v \in U. \tag{5.8}$$

This is equivalent to formula (3.5), i.e.,

$$L(v, y, \lambda) - L(u, x, \lambda) \ge 0\ \ \forall v \in U,\ \forall \lambda,$$

where L is a Lagrange functional, x and y are the state functions for the controls u and v respectively. Using equivalent transformations, then it was transformed to inequality (3.7), i.e.,

$$-\int_0^T \Delta_u H dt - \int_0^T (H_x + \lambda')\Delta x dt + \big[h_x + \lambda(T)\big]\Delta x(T) + \eta \ge 0\ \ \forall v \in U,\ \forall \lambda,$$

where $\Delta_u H$ is the increment of the function

$$H(t,u,x,\lambda) \;=\; \lambda f(t,u,x) - g(t,u,x)$$

with respect to the control, and η is the remainder term determined in the previous chapter. After choosing the solution p of the adjoint system (3.8), (3.9) as λ, we get the inequality

$$-\int\limits_0^T \Delta_u H dt \;+\; \eta \;\geq\; 0 \;\; \forall v \in U. \tag{5.9}$$

So far, all reasoning has been reversible. Indeed, if a certain function u satisfies inequality (5.9), then the increment of the Lagrange functional on this control for the chosen value of the function λ will not be negative. However, by definition, the Lagrange functional at any value of the function λ, including at $\lambda = p$, coincides with the optimality criterion.This implies the fulfillment of inequality (5.8), and hence the optimality of the control u.

Further transformations consisted in choosing as an arbitrary control v the needle variation

$$v^w_{\xi\tau}(t) \;=\; \begin{cases} u(t), & \text{if } \; t \notin (\tau-\xi, \tau+\varepsilon), \\ w(t), & \text{if } \; t \in (\tau-\xi, \tau+\varepsilon), \end{cases}$$

where w is an arbitrary admissible control, τ is an arbitrary point of the interval $(0,T)$, and ε is a small enough positive number. Then dividing by ε and passing to the limit as $\varepsilon \to 0$ we get

$$H\big[\tau, w(\tau), x(\tau), p(\tau)\big] \;-\; H\big[\tau, u(\tau), x(\tau), p(\tau)\big] \;\leq\; 0. \tag{5.10}$$

Hence, due to the arbitrariness of the point τ and the admissible control w, the maximum condition follows

$$H\big[t, u(t), x(t), p(t)\big] \;=\; \max_{v\in[a(t),b(t)]} H\big[t, v, x(t), p(t)\big], \; t \in (0,T), \tag{5.11}$$

If now the control u satisfies the maximum condition (5.11), then inequality (5.10) also holds. However, the transition from here to formula (5.9) is far from obvious, and in the general case, it is not realizable. Indeed, as a result of integrating inequality (5.10), we obtain

$$-\int\limits_0^T \Delta_u H dt \;\geq\; 0 \;\; \forall v \in U.$$

This does not guarantee the fulfillment of condition (5.9), and hence the optimality of the control u. However, if now the remainder term η is not negative, then by adding the value η to the left side of the last inequality, we actually have the relation (5.9), which implies that the control u is optimal. Thus, the maximum principle turns out to be a necessary and sufficient condition for optimality in the case of non-negativity of the remainder term. Let us try to find out what properties of the problem statement guarantee the validity of the inequality $\eta \geq 0$.

In the previous chapter, the following formula was established

$$\eta = \eta_3 - \int_0^T (\eta_1 + \eta_2)dt. \tag{5.12}$$

Here, the value η_3 corresponds to the second-order term obtained as a result of the transformation of the part of the optimality criterion, which characterizes the state of the system at the final time. The value of η_1 is associated with terms of the second order in the expansion of the value $H(t, u, x + \Delta x, p)$ in a series of Δx. Finally, we have $\eta_2 = [H_x(t, v, x, p) - H_x(t, u, x, p)]\Delta x$.

Suppose the following equalities hold

$$f(t, u, x) = f_1(t, u) + cx, \quad g(t, u, x) = g_1(t, u) + g_2(t, x), \tag{5.13}$$

where a constant c and a function f_1 are arbitrary, and functions g_1, g_2 satisfy the inequalities

$$\frac{d^2 h}{dx^2} \geq 0, \quad \frac{\partial^2 g}{\partial x^2} \geq 0. \tag{5.14}$$

The first of the inequalities (5.14) guarantees the condition $\eta_3 \geq 0$. When equalities (5.13) are fulfilled, the derivative H_x does not depend on the control, and hence $\eta_2 = 0$. Finally, equalities (5.13) together with the second inequality (5.14) lead to the condition $\eta_1 \leq 0$. As a result, from formula (5.12), it follows that $\eta \geq 0$, which means that the maximum principle gives a necessary and sufficient optimality condition. We obtain the following result

Theorem 5.2 *Suppose $\eta \geq 0$, particularly, the conditions* (5.13), (5.14) *hold. Then the maximum principle gives the necessary and sufficient optimality condition.*

Example 3.1 satisfies the equalities

$$f(t, u, x) = u, \ T = 1, \ x_0 = 0, \ a(t) = 1, \ b(t) = 2, \ g(t, u, x) = u^2/2–3x, \ h(x) = 0.$$

This implies the validity of equalities (5.13) with the values $f_1(t, u) = u$, $g_1(t, u) = u^2/2$, $g_2(t, x) = –3x$. In this case, formulas (5.14)) are satisfied in the form of equalities. Thus, the conditions of Theorem 5.2 are satisfied, which means that the maximum principle for this example gives a necessary and sufficient optimality condition. Previously, it was shown that the optimality condition here has a unique solution, which is optimal. Example 3.2 differs from Example 3.1 only in the sign of the function g, which does not affect the validity of relations (5.14) in the form of equality. As we know, here too the maximum principle gives a necessary and sufficient condition for optimality.

For Example 3.3 we have

$$f(t, u, x) = u, \ T = 1, \ x_0 = 0, \ a(t) = 1, \ b(t) = 2, \ g(t, u, x) = (u^2 + x^2)/2, \ h(x) = 0.$$

Therefore, the equalities (5.13) hold such that $f_1(t,u) = u$, $g_1(t,u) = u^2/2$, $g_2(t,x) = x^2/2$. In this case, the second derivative of the function h is equal to zero, and the second derivative of g_2 is equal to 1. Thus, inequalities (5.14) are also satisfied, which means that the optimality condition for Example 3.3 is necessary and sufficient. As we know from the previous chapter, it has a unique solution, which is the optimal one. Example 3.4 also has similar properties, differing from the previous one only in the absence of constraints.

Example 5.1 differs from Example 3.3 only in the type of extremum. To use Theorem 5.2, we pass here to the problem of minimizing the functional from the previous example, taken with the opposite sign. In this case, the only difference from the example considered above is the function $g(t,u,x) = -(u^2+x^2)/2$. Then equalities (5.13) are again valid, but with the function $g_2(t,x) = -x^2/2$. Its second derivative is equal to -1, which means that the second inequality (5.14) does not hold. Thus, the conditions of Theorem 5.2 are not satisfied, and it is not possible to establish the sufficiency of the optimality condition[34]. Earlier it was shown that sufficiency does not really hold. Thus, Theorem 5.2 provides a fairly effective tool for analyzing optimality conditions in the form of the maximum principle[35].

Note that the first equality in (5.13) corresponds to linear equations. Both equalities (5.13) assume that the control and the state function enter both the equation and the optimality criterion separately, i.e., the case of the presence of control in the coefficients at the state function is excluded[36]. Finally, inequality (5.14) is realized in the case of convexity of the functional with respect to the state function. It is clear that all these properties are realized only for a rather narrow class of problems[37].

5.2.3 Properties of non-optimal solutions of the maximum principle

As we have already seen, the maximum principle is, in generally, a necessary condition for optimality, i.e., not every its solution is optimal. We also encountered a similar situation when analyzing the stationary condition. At the same time, those solutions of the stationary condition that do not minimize the considered function, nevertheless, have some important properties. Perhaps they correspond not to the minimum, but to the maximum of the function, they deliver a local, and not an absolute minimum; they are inflection points, etc. It can be expected that non-optimal solutions of the maximum principle also have some interesting properties with respect to the considered functional. We present one result in this direction.

It was previously established that all non-optimal solutions of the maximum condition are piecewise constant functions that take the value of unity modulo. Consider, for example, an arbitrary function of this form with one break point

$$u(t) = \begin{cases} 1, & \text{if } t < \xi, \\ -1, & \text{if } t > \xi, \end{cases}$$

where ξ is a constant from the interval (0,1). Consider a newly controlled system described by the Cauchy problem (5.1), where the above function u is chosen as the control. The problem is to study the dependence on ξ of the given optimality criterion

$$I = \frac{1}{2}\int_0^1 (u^2 + x^2)dt.$$

As a result of integrating the state equation, we find the function $x(t) = t$ for $t \in (0, \xi)$. In this case, $x(\xi) = \xi$. Now integrating the equation of state from ξ to an arbitrary value t, we obtain $x(t) = 2\xi – t$ for $t \in (\xi, 1)$. Now we calculate the value

$$2I = \int_0^1 dt + \int_0^\xi t^2 dt + \int_\xi^1 (4\xi^2 - 4\xi t + t^2)dt = \frac{4}{3} - 2\xi + 4\xi^2 - 2\xi^3.$$

Find the derivative

$$\frac{dI}{d\xi} = -(3\xi^2 - 4\xi + 1).$$

It can be equal to 0 for the values[38] $\xi = 1$ and $\xi = 1/3$. The first of these values does not belong to the considered interval, for which we do not obtain a discontinuous function. Thus, of interest is the second value, which exactly coincides with the discontinuity point of the corresponding solution of the maximum principle. However, the second derivative of I at this point is equal to 2, i.e., positive. Therefore, at this point, the minimum of the optimality criterion is reached.

Thus, on the third solution of the system of optimality conditions (see Section 5.1.5), the minimum of this functional is achieved in the class of all functions that change the value from 1 to -1 once. Obviously, the corresponding fourth solution minimizes the functional in the class of all functions that change the value from -1 to 1 once. Similar results can be obtained for other solutions of the optimality conditions obtained earlier. Thus, non-optimal solutions of the maximum principle have some special properties, like non-optimal solutions of the stationary condition[39].

5.2.4 Variational inequality in case of insufficiency of the maximum principle

In the previous chapter, to study Examples 3.1 and 3.3, a necessary optimality condition was applied in the form of a variational inequality. These results coincided with the result of applying the maximum principle. For Example 3.2, the maximum principle gave a stronger result. All these examples are united by the circumstance that the maximum principle for them was a necessary and sufficient condition for optimality. In this regard, the question arises, what result will we come to by using the variational inequality to solve Example 5.1, when the maximum principle is not a sufficient condition for optimality, and the problem being solved has non-unique solution.

According to Theorem 4.1, the optimal control in the functional maximization problem satisfies the variational inequality

$$H_u(t, u(t), x(t), p(t))[v – u(t)] \geq 0 \;\; \forall v \in [a(t), b(t)].$$

For Example 5.1, the function H is defined by the formula

$$H \ = \ pu – (u^2 + x^2)/2.$$

Then the corresponding variational inequality takes the form

$$[p(t) – u(t)][v – u(t)] \geq 0 \ \ \forall v \in [–1, 1]. \tag{5.15}$$

Find its solutions.

Assume again that $p(t) – u(t) > 0$. Then dividing inequality (5.15) by the first multiplier of its left side, we establish that $v – u(t) \geq 0$ for all $v \in [–1, 1]$. This is possible only for $u(t) = -1$. Thus, if $p(t)$ is greater than $u(t)$, equal to -1, then $u(t) = -1$. If $p(t) - u(t) \leq 0$, then, dividing inequality (5.15) by the first multiplier of its left side, we establish that $v - u(t) \leq 0$ for all $v \in [–1, 1]$. This implies that $u(t) = 1$. Therefore, in the case when $p(t)$ is less than $u(t)$, equal to 1, then $u(t) = 1$. Finally, inequality (5.15) is satisfied for $u(t) = p(t)$, which is admissible only for $p(t) \in [–1, 1]$. Thus, for $p(t) < -1$ we have $u(t) = 1$, and for $p(t) > 1$ we have $u(t) = -1$. These results correspond exactly to those obtained on the basis of the maximum principle; see formula (5.4). However, at $p(t) \in [-1, 1]$ three values are allowed at once: 1, -1 and $p(t)$. It follows from this that the variational inequality turns out to be less effective for this example than the maximum principle, since in the general case it does not give a unique dependence of the control on the solution of the adjoint system[40].

5.2.5 Elimination method

Let us return to the consideration of the system of optimality conditions (5.1), (5.3), (5.4). We apply for it the ***elimination method***, considered in the previous section. Let us try to reduce this system to a problem with respect to a single unknown function. Eliminating the functions u and x from it, we obtain the ***boundary value problem***

$$p''(t) \ = \ F(p(t)), \ t \in (0, 1); \ p(1) = 0, \ p'(0) = 0, \tag{5.16}$$

where $F(p)$ denotes the value at the right-hand side of the equality (5.4). Due to the equivalence of problem (5.16) to the system of optimality conditions, we come to the following interesting conclusion: the boundary value problem for the second-order differential equation (5.16) has an infinite set of solutions[41]. This statement has one curious consequence.

Consider now the non-linear ***heat equation***

$$\frac{\partial y(\tau, \xi)}{\partial \tau} \ = \ \frac{\partial^2 y(\tau, \xi)}{\partial \xi^2} \ - \ F(y(\tau, \xi)), \ \tau > 0, \ 0 < \xi < 1 \tag{5.17}$$

with the boundary conditions

$$\frac{\partial y(\tau, 0)}{\partial \xi} \ = \ 0, \ y(\tau, 1) = 0, \ \tau > 0 \tag{5.18}$$

and an initial condition, where the function F has the same form as in problem (5.16). To find the equilibrium position of the considered system, it is sufficient to equate the derivative of y with respect to τ to zero. Obviously, the equilibrium position $z = z(\xi)$ for system (5.17), (5.18) is a solution to the boundary value problem

$$z''(\xi) \;=\; F(z(\xi)),\; \xi \in (0,1);\; z(1) = 0,\; z'(0) = 0,$$

which, up to notation, coincides with problem (5.16). Then, based on the results obtained earlier, we conclude that the system characterized by equation (5.17) with boundary conditions (5.18) has an infinite set of equilibrium positions[42].

5.2.6 Modification of Example 5.1

Consider now a modification of Example 5.1

Example 5.2 *Find the maximum of the functional*

$$I(u) \;=\; \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

on the set

$$U \;=\; \{u\big|\; 0 \le u(t) \le 1,\; t \in (0,1)\},$$

where x is a solution of the problem

$$x'(t) \;=\; u(t),\; t \in (0,1);\; x(0) = 0.$$

This example differs from the previous one only in the set of admissible controls. As a result, the function H has the same form as before, i.e., $H = pu–(u^2 + x^2)/2$. Remains unchanged and the adjoint system

$$p'(t) \;=\; x(t),\; t \in (0,1);\; p(1) = 0.$$

As in Example 5.1, the unique stationary point $u = p$ of the function H corresponds to its minimum, which means that the solution of the optimality condition is achieved on the boundary of the set of admissible controls. Find the values

$$H(0) \;=\; -x^2/2,\; H(1) \;=\; p - (1 + x^2)/2.$$

The solution of the optimality condition corresponds to the smallest of them. As a result, we determine the control

$$u(t) \;=\; \begin{cases} \;\;\,1, & \text{if } p(t) < 1/2, \\ -1, & \text{if } p(t) > 1/2 \end{cases} \tag{5.19}$$

that is the analogue of formula (5.4).

To find a solution to the resulting system of optimality conditions, one could use the previously described algorithms. However, we try to estimate the set of possible solutions of the adjoint system, which have the form

$$p(t) = -\int_t^1 x(\tau)d\tau.$$

The system state function is found by the formula

$$x(t) = \int_0^t u(\tau)d\tau.$$

Integrating the inequality $0 \leq u(t) \leq 1$, which follows from the definition of the set of admissible controls, from zero to an arbitrary value t, we conclude that $0 \leq x(t) \leq t$ for any admissible control. Integrating the resulting relation from an arbitrary value of t to 1, we have

$$0 \leq \int_t^1 x(\tau)d\tau \leq \frac{1-t^2}{2}.$$

This implies the estimate $(t^2–1)/2 \leq p(t) \leq 0$. Thus, for any admissible control and for all values $t \in (0,1)$ the function p can take exclusively negative values. Then formula (5.19) implies that the optimality conditions have a unique solution $u(t) = 1$ for all t.

Let us verify that this is indeed a solution to the optimal control problem. Indeed, it follows from the inequalities $u(t) \leq 1$ and $x(t) \leq t$ that for any admissible control the estimate

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt \leq \frac{1}{2}\int_0^1 (1+t^2)dt \leq \frac{2}{3}.$$

In this case, the value $I(u) = 2/3$ is possible only when $u = 1$. Thus, the found solution of the optimality conditions is indeed optimal[43].

Thus, in the considered example, the uniqueness of the optimal control and the sufficiency of the optimality conditions are realized. It is characteristic that the general properties of the optimality criterion and the set of admissible controls for Examples 5.1 and 5.2 coincide. In particular, when passing to the corresponding minimization problems, one obtains a non-convex functional with the same remainder term. However, in the second case, the optimal control is unique, and the optimality conditions are necessary and sufficient, while in the first case, both of these properties are violated. Thus, the uniqueness of the optimal control for Example 5.2 is obtained when the conditions of the previously given uniqueness theorem are violated, and the sufficiency of the maximum principle is realized when the conditions of the theorem proved earlier on the sufficiency of optimality conditions are violated[44].

Additional conclusions

Based on the results given in Appendix, we have additional conclusions about the uniqueness of the solution of optimal control problems and the sufficiency of the optimality conditions.

- The system of optimality conditions for Example 5.1 is sign-invariant in the sense that if a triple of functions u, x, p is a solution to a system of optimality conditions, then a triple of functions that differs from it only in sign is also a solution to this system.

- The optimal control problem from Example 5.1 is sign-invariant, i.e., the values of the optimality criterion for two controls that differ only in signs are the same.

- The maximum principle is a necessary and sufficient condition for optimality in the case of non-negativity of the remainder term in the functional increment formula.

- The sufficiency of the maximum principle is guaranteed in the case when the equation of state is linear and does not include control in the coefficients before the state function, the control and the state function are included in separate terms of the integrand of the optimality criterion, and the term depending on the state and the terminal term of the optimality criterion have non-negative second derivatives.

- Non-optimal solutions of the maximum principle may have some special properties. The variational inequality for Example 5.1 turns out to be less efficient than the maximum principle, since it has a wider set of solutions.

- The system of optimality conditions for Example 5.1 is reduced to a boundary value problem for a non-linear second-order ordinary differential equation with an infinite set of solutions.

- There is a boundary value problem for a non-linear heat equation with an infinite set of equi-librium positions corresponding to the solutions of the system of optimality conditions for Example 5.1.

- Despite the fact that the optimal control problems considered in Examples 5.1 and 5.2 differ only in the sets of admissible control values (the segment $[-1, 1]$ in the first case and [0,1] in the second), these problems themselves differ significantly according to its properties (in the first case there are two optimal controls with an infinite set of solutions to the optimality conditions, while in the second case both the problem itself and the optimality condition have a unique solution).

- Optimal control problems can have any number of solutions.

- The optimal control problem from Example 5.3 has three solutions.

- A situation is possible when the optimality criterion turns out to be Gateaux non-differentiable, as a result of which there is a need to use methods of non-smooth optimization.

Notes

1. Example 5.1 is provided in [170]. Chapters 11 and 12 will consider the functional maximization problems from Example 5.1 in the presence of an additional boundary condition for the state of the system; see Examples 11.1 and 12.7.

2. It is interesting that we will meet the system of optimality conditions (5.1), (5.2), and (5.4) in the subsequent section, where it will characterize the set of non-singular solutions of optimality conditions for another optimal control problem. In Chapters 6 and 7, two different optimal control problems will be considered, in which a system is obtained that differs from the given one solely by the sign in the control formula. It turns out that this circumstance qualitatively changes the properties of the optimality conditions.

3. Naturally, this is the same control as in Example 3.3

4. Indeed, formula (5.3) is the problem of finding the minimum of a control-continuous function H on a closed set (segment) $[-1, 1]$. By the Weierstrass theorem (see Chapter 1), this problem has a solution. In the absence of local minimum points, only a boundary point of a given set can be such.

5. Interestingly, for the analog of Example 5.1 considered in Chapter 12, in the presence of an additional boundary condition, the desired result is obtained without using an iterative process, although the optimal control problem of a system with a fixed final state is considered there is more difficult than the corresponding system with a free final state.

6. This can be easily verified by substituting these values into the given system.

7. See, in particular, Example 2.10.

8. Other examples of the non-uniqueness of optimal controls are considered in Chapters 6, 11, 14, 15, and 17.

9. We have already encountered the absence of the uniqueness of minimum points for functions in Examples 1.3, 1.5, and 1.6.

10. More precisely, a set X is called a ***vector*** or ***linear space*** if the sum of any two elements from X, as well as the product of any element from X by an arbitrary number (scalar), belong to the same set. Moreover, the set X forms an ***Abelian group*** with respect to addition (the associativity and commutativity conditions are satisfied, there is a neutral element, and each element is invertible), multiplying any element by 1 gives this element itself, and some distributivity conditions are satisfied; see [94], [100], [106], [158].

11. Naturally, the sum of elements of a numerical segment cannot to be an element of the same segment.

12. Figure 1.18 shows convex and non-convex sets in the plane.

13. The properties of convex sets and convex functionals are the subject of ***convex analysis***; see [60], [91], [161].

14. The following chapter will give an example of an optimal control problem with a convex but not strictly convex functional that has an infinite number of solutions; see Example 6.1. We already encountered a similar situation in Example 1.6, where a convex but not strictly convex function has an infinite set of minimum points. On the other hand, a convex but not strictly convex function $f(x) = |x|$ has a unique minimum point. Theorem 5.1 gives conditions that guarantee the uniqueness of the solution to the problem, but if these conditions are violated, the solution may also turn out to be unique. An example of a uniquely solvable optimal control problem with a convex but not strictly convex functional is given in Chapter 9; see Example 9.1.

15. Actually, this distinguishes the problems of optimal control from the problems of the calculus of variations, which are characterized by a direct dependence of the functional on the unknown function.

16. This means that the dependence of the state function on the control is an ***affine operator***. Naturally, any ***linear operator***, i.e., such an operator L for which the equality $L(\alpha x + \beta y) = \alpha Lx + \beta Ly$ is valid for any elements x and y from its domain and any numbers α and β is affine. Any affine operator A, defined on a vector space, is characterized by the equality $Av = Bv + c$, where B is a linear operator acting in the same spaces as A, and c is some point from the set of values of the considered operator.

17. More precisely, it follows from the theorem that this problem cannot have two solutions. However, we already know the existence of a solution to the problem. Theorem 5.1 will be used to prove the uniqueness of the optimal control for a system with a fixed final state in Chapter 9, for a system with an isoperimetric condition in Chapter 13 and for a system with a free initial state in Chapter 16.

18. One should not think that checking the convexity of a functional for general optimal control problems is such a simple problem. In this case, the achievement of the desired goal is associated with a linear (more precisely, affine) property of the dependence of the state function on the control. This, in turn, is explained by the fact that both the control and the state of the system enter problem (5.1) linearly. At the same time, the equation of state for Example 4.1 is non-linear with respect to control, and for Examples 4.2 and 3.3, it is non-linear with respect to the state of the system. Checking the convexity of the functional criterion, even for sufficiently simple non-linear equations, is extremely difficult. In this regard, the justification of the uniqueness of the optimal control for non-linear systems can only be carried out in exceptional cases; see, in particular, [73], [148], [156], [165]. It should also be borne in mind that any positive property, in particular, the uniqueness of the solution of the problem, is usually an exception.

19. Another example of a uniquely solvable optimal control problem when the conditions of Theorem 5.1 are violated is given in Appendix, but the solution of the problem is continuous there; see also Example 11.1.

20. The zero value of the function p does not give new results. We can determine the control equal to 1 for both negative and non-positive values of this function.

21. We have already met with a similar situation in Examples 1.2, 1.3, and 1.5, when minimizing various functions of one variable.

22. Other examples of insufficiency of optimality conditions are given in Chapter 6, where insufficient solutions of the maximum principle are singular controls, in Chapter 11, where systems with fixed finite states are studied, in Chapters 14 and 15, when studying optimal control problems with isoperimetric conditions, and in Chapter 17, for systems with a free initial state.

23. Usually, the adjoint system is solved in the reverse direction of time. However, in our case, this is not essential, since the values of the function p at the three points $1/5$, $3/5$, and 1 are known.

24. Curiously, for the analog of this problem considered in Section 12 for a system with a fixed final state, the set of solutions to the optimality condition is also infinite. However, it does not include continuous controls, since they do not guarantee the output of the system to the desired final state. In the following section, we will encounter a situation where the optimality conditions have not just an infinite, but not even a countable set of solutions.

25. For the analogue of Example 5.11 considered in Chapter 12 for a system with a fixed final state, in which the set of solutions of the optimality condition is also infinite, the controls with one discontinuity point turn out to be optimal. The following section gives an Example of an optimal control problem with an infinite number of solutions. Examples of optimization problems that have a non-unique solution are given in [75], [116], [165], [170]. Examples of optimal control problems for which the necessary optimality conditions are not sufficient are also given in Chapters 6, 11, 14, 15 and 17.

26. A similar example will be considered in Chapter 11 for the problem of optimal control of a system with a fixed final state. Chapter 15 gives examples of optimal control problems with an isoperimetric condition that have an infinite number of solutions.

27. We already met with a similar effect in Chapter 1 when analyzing the function $f(x) = x^4 – 2x^2$.

28. For the sufficiency of optimality conditions in the form of the maximum principle; see [74].

29. Using variational inequalities to analyze a wide class of optimal control problems for systems described by partial differential equations; see [116], [171].

30. From an algebraic point of view, this means that on the set of solutions of the optimality conditions, a first-order operation is defined, which consists of changing the sign of the function. On the other hand, we can consider the set of transformations $X = \{e, s\}$, which translates the set of solutions of the optimality conditions into itself, where e is the identity transformation, and s is the sign change. On the set X, one can define a second-order operation $\bullet$ consisting in successive realization of two transformations, i.e., superposition of transformations. Obviously, the equalities $e \bullet e = e$, $s \bullet s = e$, $e \bullet s = s$, $s \bullet e = s$ are valid. It follows from these equalities that the given operation is associative, and e is the unit with respect to the given operation. Moreover, both elements of the set X are invertible, and this element itself turns out to be the inverse element. Thus, we are dealing with a ***group*** of transformations of the set of solutions of optimality conditions. The search for transformations that ensure the transition from one solution of the system to another makes it possible to find new solutions to the problem based on the existing ones. We will use this technique in Chapters 12 and 14.

31. The result obtained is a manifestation of the ***Curie principle***, according to which, when certain causes cause certain consequences, then the symmetry elements of the causes should appear in the consequences caused by them. This statement was formulated by Pierre Curie in connection with the problems of crystallography. Note that the phenomenon of symmetry is invariably investigated using ***group theory***; see [163], [192].

32. The optimal control problem for a system with a fixed finite state, considered in Chapter 9; see Example 9.1, and an optimal control problems with an isoperimetric condition; see Chapters 14 and 15, have similar properties. Example 3.2, the problem of minimizing this functional also has all the described properties. At the same time, it has a unique solution. The point is that this solution is zero, so changing the sign does not give a new solution. It is characteristic that

the problem from Example 9.1 has two solutions that differ in signs, however, the corresponding optimality conditions also have a non-optimal zero solution.

33. We recall that it was in this way that the sufficiency of the optimality condition obtained in accordance with dynamic programming in the previous chapter was established.

34. Theorem 5.2 only gives conditions under which the maximum principle is certainly a sufficient condition for optimality. It is clear that if these conditions are violated, the maximum principle can, nevertheless, also turn out to be a sufficient optimality condition, which is realized for Example 5.2.

35. It should be borne in mind that, as a rule, it is not possible to establish the sign of the remainder term in the formula for the increment of the functional for sufficiently difficult optimal control problems. Establishing the sufficiency of the maximum principle for systems described by non-linear equations is a problem of an exceptionally high degree of difficulty.

36. Coefficient control problems often arise in coefficient inverse problems; see [43], [99], [120] and in area control problems; see [119], [136], [137], [178], [185]. Applied control problems in coefficients for distributed parameter systems are considered; for example, in [9], [119], [124]. Justification of optimality conditions for control problems in coefficients for such systems is carried out; for example, in [116], [119], [124], [136], [137], [155], [171], [178].

37. The following chapter gives an Example of an optimal control problem with qualitatively different properties, for which the maximum principle is a necessary and sufficient optimality condition in accordance with Theorem 5.2. In Chapter 9, this theorem is used to prove the sufficiency of optimality conditions for an optimal control problem with a fixed final state, in Chapter 13 the sufficiency of optimality conditions for a problem with an isoperimetric condition is similarly proved, and in Chapter 16, for a problem with a free initial state.

38. It is at these points that the derivative of the solution of the adjoint system vanishes; see Section 5.1.5.

39. Chapter 15 will consider an optimal control problem in which non-optimal solutions of the optimality conditions form a minimizing sequence for a given optimality criterion.

40. The sufficiency of the necessary optimality condition in the form of a variational inequality for Examples 3.1 and 3.3 is due to the fact that the minimized functional for these problems is convex. At the same time, for Example 5.1, the corresponding minimized functional is non-convex. As a result, the set of solutions to the optimality condition turns out to be much wider than the set of optimal controls. Despite the obtained result, it cannot be said that the variational inequality is generally inapplicable for this example. Indeed, if the value $p(t)$ does not belong to the interval $[-1, 1]$, then the dependence of the function u on p is uniquely established. If $p(t)$ belongs to this interval, then we still got serious information about the optimal control. If initially we knew that at any moment of time it could take any values from the interval $[-1, 1]$, now we conclude that only one of the three specified values is possible, i.e., as a result of applying the variational inequality, the measure of uncertainty is significantly reduced.

41. It is curious that, by changing the sign in front of the function F, we get a boundary value problem for a non-linear differential equation of the second order, which has no solution at all; see Chapter 7.

42. The tendency of the system to a specific equilibrium position is due to the choice of the initial state of the system, just as the output of the iterative process for an ambiguously solvable equation to a specific solution is due to the choice of the initial approximation; see Example 5.1.

43. The optimality of the control $u = 1$ for Example 5.2 can be proved in another way. In Example 5.1, it was shown that this control maximizes this functional on the set of controls with values from the interval $[-1, 1]$. Then, all the more, it minimizes the functional on the set of functions with values from the narrower set [0,1], which means that it is also an optimal control for Example 5.2. The second optimal control $u = -1$ in Example 5.1 is not admissible in this case, and therefore cannot be a solution to the considered problem. Thus, the control $u = 1$ is the only solution to the problem for Example 5.2.

44. In Example 3.2, the uniqueness of the optimal control was also observed when the conditions of Theorem 5.1 are violated. However, in this case, the conditions of Theorem 5.2 were satisfied, as a result of which the corresponding optimality condition was necessary and sufficient. In this case, however, the conditions of Theorem 5.2 are violated. However, the optimality conditions are also sufficient.

CHAPTER 6

Singular controls

To solve optimal control problems in the previous chapters, we used the maximum principle, according to which the optimal control delivers the maximum of some function on the set of admissible control values. However, sometimes the maximum principle degenerates, as a result of which it is difficult to determine control explicitly from the optimality condition. The corresponding solutions to the optimality conditions are called singular controls and may (but need not) also be optimal[1]. This chapter provides examples of such problems, studies the problems of the uniqueness of optimal control and the sufficiency of the optimality condition when the maximum principle degenerates, establishes conditions for the existence of singular controls, estimates their number and provides necessary conditions for the optimality of singular controls.

6.1 LECTURE

In Chapter 3, for the optimal control problem a necessary optimality condition in the form of the maximum principle was established. The resulting system can be solved analytically or by some iterative process. Below, we study an example of an optimal control problem for which the maximum principle degenerates. In this case, the optimality conditions have specific solutions, called singular controls. The problem of the existence of singular controls, as well as the size of the set of singular controls, is condemned.

6.1.1 Problem statement

In the previous chapters, we considered optimal control problems, for which the maximum principle was used. The system of optimality conditions included a state equation with an initial condition, an adjoint system, and a maximum condition, which is the problem of maximizing a certain function on a given set. Next, we found the control from the maximum condition and substituted the result into the state equation and the adjoint system, after which we found the solution of the problem either directly from the optimality conditions (Examples 3.1 and 3.2), or using the decoupling method (Example 3.4), or using an iterative algorithm (Examples 3.3

DOI: 10.1201/9781003398585-6

and 5.1). However, there are situations when all these techniques do not lead to the desired goal.

Example 6.1 *It is required to minimize the functional*[2]

$$I(u) = \int_0^1 u(t)x(t)dt$$

on the set

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\}$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{6.1}$$

Note that the such state equation and the set of admissible controls were previously considered in Examples 3.3 and 5.1, and the optimality criterion is extremely simple. These circumstances do not seem to portend any surprises. To study this problem, we use the previously described method.

The characteristics of the general optimal control problem from Chapter 3 are given as follows:

$$f(t,u,x) = u,\ T = 1,\ x_0 = 0,\ a(t) = -1,\ b(t) = 1,\ g(t,u,x) = ux,\ h(x) = 0.$$

In accordance with formula (3.3), we determine the function

$$H = pu - ux.$$

Then the adjoint system (3.8), (3.9) takes the form

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0. \tag{6.2}$$

According to the maximum principle, the optimal control must maximize the function H on the set U, which corresponds to the equality

$$H(u) = \min_{|v|\leq 1} H(v). \tag{6.3}$$

Thus, to find the optimal control, we have the system (6.1)–(6.3). To analyze this system, we will use the methods described before[3].

6.1.2 Analysis of optimality conditions

The first step in the analysis of the system of optimality conditions is to find the control from the system of optimality conditions. Using the standard method, we find the derivative

$$\partial H/\partial u = p - x.$$

Equating this derivative to zero, we could find the corresponding stationarity points. However, unlike the previous cases, the resulting value does not explicitly depend on the control[4]. Therefore, we can conclude that the function H has no local extrema. In view of the absence of stationary points, the conditional extremum of the function H can be achieved only on the boundaries of the set of admissible controls. We encountered a similar situation in Example 5.1, where the stationary point existed but was not a solution to the optimality condition.

Find the boundary values

$$H(1) = p - x, \quad H(-1) = x - p.$$

Then the solution of the maximum principle is determined by the formula

$$u(t) = \begin{cases} 1, & \text{if } p(t) - x(t) > 0, \\ -1, & \text{if } p(t) - x(t) < 0. \end{cases} \tag{6.4}$$

The resulting relation is similar to equality (5.4) from the previous chapter and means that the optimal control must be constant or piecewise constant. Consider the system of equations (6.1), (6.2), and (6.4).

Since the dependence of the control on the functions x and p is essentially nonlinear, we cannot find an analytical solution, as it was for Examples 3.1 and 3.2. We also do not have the possibility of using the decoupling method, as for Example 3.4, due to the non-linear dependence (6.4) between unknown functions. However, there is no obstacle to applying the iterative method successfully used earlier in the analysis of Examples 3.3 and 5.1. In particular, at the kth iteration, with a known control u_k, the system state function is found from the relations

$$x_k'(t) = u_k(t), \ t \in (0,1); \ x_k(0) = 0.$$

Then the adjoint system is solved

$$p_k'(t) = x_k(t), \ t \in (0,1); \ p_k(1) = 0.$$

Finally, the new control approximation is determined by the formula

$$u_{k+1}(t) = \begin{cases} 1, & \text{if } p_k(t) - x_k(t) > 0, \\ -1, & \text{if } p_k(t) - x_k(t) < 0. \end{cases}$$

Let us try to prove the convergence of the described algorithm. From formula (6.4) it follows that at any t the control can take only one of two values: -1 or 1. In this case, the variant is not excluded, in which at some moments the control switches from one of these values to another, as it was for Example 5.1.

To begin with, we choose the constant control $u_0(t) = 1$ as the initial approximation. The corresponding state function is

$$x_0(t) = \int_0^t u_0(\tau) d\tau = t.$$

Find the solution of the adjoint system

$$p_0(t) = -\int_t^1 u_0(\tau)d\tau = t - 1.$$

We calculate the difference $p_0(t) - x_0(t) = -1$. Since this value is negative, in accordance with formula (6.4), we determine a new approximation control $u_1(t) = -1$. It corresponds to the state function $x_1(t) = -t$ and the solution of the adjoint system $p_0(t) = 1-t$. Calculating the difference $p_0(t)-x_0(t) = 1$, we conclude that $u_2(t) = 1$. This value coincides with the initial approximation of the control. It is clear that at the next iteration the control is equal to -1, and so on. Thus, for the chosen initial approximation, the iterative process does not converge.

The divergence of the algorithm for a particular initial approximation is, of course, an unpleasant situation. However, this can hardly be considered fatal, since we can use the algorithm with other initial approximations[5]. It is clear that by choosing $u_0 = -1$ as the initial approximation, we obtain the value $u_1 = 1$ at the next iteration. Then we get -1, then again 1, etc., which again means the divergence of the iterative process[6]. Of course, we still have the possibility to choose discontinuous controls as the initial approximation. For Example 5.1, such a choice made it possible to find new solutions to the system of optimality conditions. However, we will do otherwise.

Let us try to exclude two of the three unknown functions from the system of optimality conditions[7] (6.1), (6.2), and (6.4). Obviously, the solution of the Cauchy problem (6.1) is

$$x(t) = \int_0^t u(\tau)d\tau.$$

The solution to problem (6.2) is found similarly

$$p(t) = -\int_t^1 u(\tau)d\tau.$$

Now we determine the difference

$$p(t) - x(t) = -\int_t^1 u(\tau)d\tau - \int_0^t u(\tau)d\tau = -\int_0^1 u(\tau)d\tau.$$

Substituting this value into equality (6.4), we obtain the peculiar equation for the control[8]

$$u(t) = \begin{cases} 1, & \text{if } \int_0^1 u(\tau)d\tau < 0, \\ -1, & \text{if } \int_0^1 u(\tau)d\tau > 0. \end{cases} \tag{6.5}$$

Obviously, the value on the right side of the resulting formula does not depend on t. Thus, the control turns out to be a constant, which means that it is always equal to either 1 or –1. However, for $u(t) = 1$ the integral of the control is positive, and for $u(t) = -1$ it is negative. In both cases, equality (6.5) is not valid. Therefore, equation (6.5) has no solution[9]. However, it is equivalent to system (6.1), (6.2), and (6.4). As a result, we come to the conclusion that this system also has no solution[10].

Based on the results obtained, it would seem that one could conclude that the optimality conditions have no solution at all. We already encountered a similar situation in Chapter 1 when studying the function $f(x) = x$. The stationary condition for it reduces to the equality 1=0, which is meaningless. However, the solution of the minimization problem must surely satisfy the necessary minimum condition. The absence of an object that satisfies this condition is a sure sign of the absence of a solution to the minimization problem itself. In particular, the above function f has no minimum points. Having established the absence of a solution to problem (6.1), (6.2), and (6.4), it would seem that we should conclude that the considered optimal control problem has no solution. The more striking is the fact that the optimal control for this problem still exists[11].

6.1.3 Singular controls

We have a paradoxical situation. On the one hand, the solution of the optimal control problem exists[12], and, on the other hand, the resulting system (6.1), (6.2), and (6.4) has no solution. However, an optimal control must always satisfy the necessary optimality condition. The way out of this situation is possible only in the case when this system is not the necessary condition of optimality. The only way out of this paradox is to admit that formula (6.4) is not equivalent to the maximum condition (6.3), which, without a doubt, gives the necessary optimality condition.

The maximum condition (6.3) has the following form

$$[p(t) - x(t)]u(t) = \max_{|v|\leq 1} [p(t) - x(t)]v, \ t \in [0, 1].$$

Obviously, this equality will certainly hold if the value in square brackets vanishes. Indeed, in this case, the maximum principle is fulfilled in the trivial form 0=0. Characteristically, this value corresponds to the derivative of the function H with respect to control. Consequently, its equality to zero is a stationary condition for this function.

Definition 6.1 *The situation in which the maximum principle is satisfied in a trivial form is called its* ***degeneration****, and any admissible control on which the maximum principle degenerates is called a* ***singular control****. Non-singular solutions of the maximum principle are called* ***regular****.*

Previously, it was established that the maximum principle for the considered example does not have a regular solution, i.e., if the optimality conditions have a solution, then it must necessarily be a singular control. Obviously, singular control, being a specific form of solving the maximum principle, may turn out to be optimal, but it does not have to be that way. Let us first try to determine which admissible control turns out to be singular. The p–x difference was determined earlier. Then the degeneration of the maximum principle is realized when the following equality holds[13]

$$\int_0^1 u(\tau)d\tau = 0. \tag{6.6}$$

Thus, the set of all singular controls for the considered example is characterized by the equality

$$U_0 = \Big\{u \in U \Big| \int_0^1 u(\tau)d\tau = 0\Big\}.$$

The question arises, how wide is the set of singular controls U_0? Obviously, it includes a function identically equal to zero; any function equal to some number a from the segment $[-1, 1]$ at $t \in (0, 1/2)$ and equal to $-a$ at $t \in (1/2, 1)$; a function of the form $a \sin 2k\pi t$, where k is a natural number and much more; see Figure 6.1.

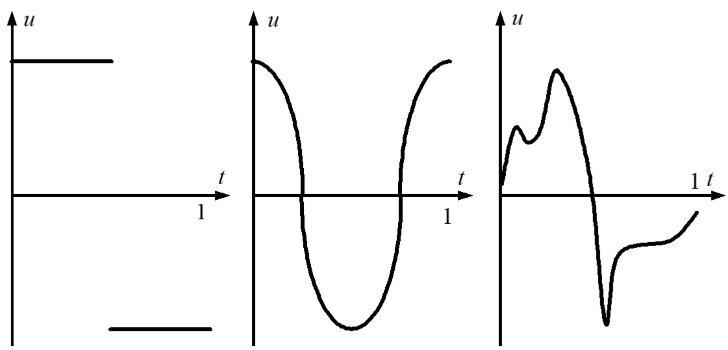

Figure 6.1 Singular controls for Example 6.1).

Thus, the maximum principle has an infinite and not even countable set of solutions[14], all of which are singular controls. Now a natural question arises, which of them is optimal? In Chapter 5, we have already encountered a situation where the optimality condition has an infinite number of solutions. To choose the best of them, i.e., optimal control, we then calculated the value of the optimality criterion on an arbitrary solution of the optimality conditions, which made it possible, as a result, to find a solution to the original problem. In this case, such a technique, it would seem, should not be effective, since, unlike Example 5.1, we do not have a formula for an arbitrary singular control. However, we will try to use it.

Let us find the value of the functional to be minimized on an arbitrary singular control, i.e., on any admissible control satisfying equality (6.6). Substituting the value of control into the optimality criterion, we have

$$I = \int_0^1 xudt = \int_0^1 xx'dt = \frac{1}{2}\int_0^1 \frac{d}{dt}x^2dt = \frac{x(1)^2}{2}. \tag{6.7}$$

Taking into account the well-known formula for solving problem (6.1), we find the value

$$x(1) = \int_0^1 u(\tau)d\tau.$$

As a result, we determine the value of the functional on an arbitrary control

$$I(u) = \frac{1}{2}\Big[\int_0^1 u(\tau)d\tau\Big]^2. \tag{6.8}$$

Obviously, this value is not negative, and here the equality to zero is possible if and only if the integral is equal to zero. However, this exactly corresponds to equality (6.6), i.e., description of the set U_0 and definition of singular controls. As a result, we conclude that the considered problem has an infinite and even non-countable set of optimal controls, all of which are singular[15].

6.1.4 Existence of singular control

In Example 6.1, we encountered singular controls, while for the optimization problems discussed in the previous chapters, there seemed to be no singular controls. The question arises, why does singular control exist in some problems, but not in others? Let us return to the general optimal control problem posed in Chapter 3.

We consider the general problem of minimizing the functional

$$I(u) = \int_0^T g(t, u(t), x(t))dt + h(x(T))$$

on the set

$$U = \{u|\ a(t) \le u(t) \le b(t),\ t \in (0, T)\},$$

where x is a solution of the Cauchy problem

$$x'(t) = f(t, u(t), x(t)),\ t \in (0, T);\ x(0) = x_0$$

with known functions a, b, f, g, h and numbers T, x_0. In accordance with Theorem 3.1, the optimal control satisfies the maximum condition

$$H\big(t, u(t), x(t), p(t)\big) = \max_{v \in [a(t), b(t)]} H\big(t, v, x(t), p(t)\big),$$

where p is a solution of the corresponding adjoint system, and the function H is determined by the equality

$$H(t, u, x, p) = f(t, u, x)p - g(t, u, x).$$

Suppose the functions f and g satisfy the conditions

$$f(t, u, x) = f_1(t, x)\varphi(u) + f_2(t, x),\ \ g(t, u, x) = g_1(t, x)\varphi(u) + g_2(t, x), \tag{6.9}$$

where f_1, f_2, g_1, g_2, and φ are some functions of their arguments. Then the function H takes the following form

$$H(t,u,x,p) = \big[f_1(t,x)p–g_1(t,x)\big]\varphi(u) + \big[f_2(t,x)p–g_2(t,x)\big].$$

As a result, we obtain the maximum condition

$$[f_1(t,x(t))–g_1(t,x(t))]\varphi(u(t)) = \max_{v\in[a(t),b(t)]} [f_1(t,x(t))–g_1(t,x(t))]\varphi(v),\ t\in(0,T). \tag{6.10}$$

If now there is a control u from the set U such that

$$f_1(t,x(t))p(t)-g_1(t,x(t)) = 0,\ t\in(0,T), \tag{6.11}$$

then equality (6.10) is certainly satisfied in a trivial way. This corresponds to the degeneration of the maximum principle. Thus, the following assertion is true.

Theorem 6.1 *A singular control exists if equalities* (6.9) *are satisfied, and the set of admissible controls for which conditions* (6.11) *are valid is not empty*[16].

Note that, according to equalities (6.9), the control enters equally into both the state equation and the optimality criterion. In particular, for Example 6.1, we have $f(t,u,x) = u$, $g(t,u,x) = ux$. Thus, the functions entering into conditions (6.9) are $f_1(t,x) = 1$, $f_2(t,x) = 0$, $g_1(t,x) = x$, $g_2(t,x) = 0$, $\varphi(u) = u$. Equality (6.11) here takes the form $p(t) - x(t) = 0$ and determines the set U_0 of singular controls for this example; see Section 6.2.1. However, for Example 3.1 we have $f(t,u,x) = u$, $g(t,u,x) = u^2/2–3x$, and for Examples 3.3 and 3.4 we have $f(t,u,x) = u$, $g(t,u,x) = (u^2+x^2)/2$. Examples 3.2 and 5.1 differ from Examples 3.1 and 3.3, respectively, only in the type of extremum. Thus, the function f is the same in them, while g differs only in sign. In all five cases, equalities (6.9) do not hold, and the function H includes the square of the control without any coefficients that could vanish. Consequently, the degeneration of the maximum principle is impossible here, and there are no singular controls.

6.1.5 Finiteness of the set of singular controls

In Example 6.1 the set of singular controls is infinite, while in Examples 3.1–3.4 and 5.1 this set is empty. We would like to know whether this set can be finite and non-empty. Consider another example[17].

Example 6.2 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 x(t)^2 dt$$

on the set

$$U = \{u|\ |u(t)| \le 1,\ t\in(0,1)\}$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t\in(0,1);\ x(0) = 0. \tag{6.12}$$

Determine the function $H = pu – x^2/2$. Then the adjoint system takes the form

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0. \tag{6.13}$$

The corresponding maximum condition is written as follows

$$p(t)u(t) = \max_{|v|\le 1} p(t)v,\ t \in (0,1). \tag{6.14}$$

Due to the linearity of the function H, its maximum, as if it exists, can only be located on the boundary of the set of admissible controls, which leads to the relation

$$u(t) = \begin{cases} 1, & \text{if } p(t) > 0, \\ -1, & \text{if } p(t) < 0. \end{cases} \tag{6.15}$$

Thus, it seems that the optimal control is identically modulo one.

However, the considered problem is so simple that its solution can be found without recourse to the maximum principle[18]. Indeed, by definition, the functional to be minimized is non-negative. Equality to zero here is possible only in the case when the state of the system is identically equal to zero, which, in turn, is realized on a single control $u = 0$. The latter, being an element of the set of admissible controls, turns out to be optimal. Therefore, the optimal control problem considered in Example 6.2 has a unique solution $u = 0$.

The obtained result comes into direct conflict with equality (6.15). However, this relation characterizes exclusively regular solutions of the maximum principle, provided that they exist[19]. The discrepancy between representation (6.15) and the established type of optimal control suggests that the maximum principle should determine some singular controls.

If we exclude the controls defined by formula (6.15), then the maximum condition (6.14) is satisfied only when the control coefficient vanishes. As a result, we obtain the equality $p = 0$. As can be seen from problem (6.13), the equality to zero of the solution of the adjoint system is possible only for $x = 0$. Substituting this value into equality (6.12), we find a singular control $u = 0$, which, as we already know, is optimal[20].

In Example 6.1, there were an infinite number of singular controls; in Example 6.2, there was only one singular control[21], and in other examples there were no singular controls at all. The question arises whether the situation is possible when the set of singular controls is finite, but their number is greater than one. Consider another example[22].

Example 6.3 *It is required to minimize the functional*

$$I(u) = \int_0^1 \Big(\frac{x^3}{3} - \frac{x^2 t}{2}\Big) dt$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Find the function

$$H = pu - \frac{x^3}{3} + \frac{x^2 t}{2},$$

where p is a solution of the adjoint system

$$p'(t) = x(t)^2 - tx(t),\ t \in (0,1);\ p(1) = 0.$$

Based on the form of the function H, we conclude that the degeneration of the maximum principle is possible only at $p = 0$. Then the derivative of p also vanishes. This is possible in two cases: for $x(t) = 0$ and for $x(t) = t$. The obtained states of the system are implemented, respectively, under the controls $u(t) = 0$ and $u(t) = 1$. Thus, for Example 6.3, there are two singular controls[23]. Note that in this case, we do not ask ourselves whether these controls are optimal or not[24]. It was important for us to establish that for the considered example, there are two singular controls.

RESULTS

Here is a list of questions devoted to the degeneration of the maximum principle and singular controls, the main conclusions on this topic, as well as additional problems that arise in this case, partially solved in Appendix, partially taken out in the Notes.

Questions

It is required to answer questions concerning the degeneration of the maximum principle.

1. Why is the adjoint system for Example 6.1 independent of the state function, unlike the examples described in previous chapters?
2. What is unusual about the properties of the maximum principle for Example 6.1?
3. Why does the maximum principle for this example have properties that essentially distinguish it from the examples considered earlier?
4. Can the iterative process for the system (6.1), (6.2), and (6.4) converge for some initial approximation?
5. Why did the iterative process for the considered example not lead to the desired results?
6. Why are relations (6.3) and (6.4) not equivalent?
7. What is the difference between singular and regular solutions of the maximum principle?
8. What class of equations does relation (6.5) belong to?

9. On what basis was it concluded that problem (6.5) has no solution?
10. What class of problems does relation (6.6) belong to?
11. Is it possible to find solutions to equation (6.6) different from those given earlier?
12. Why is an arbitrary singular control for the considered example turned out to be optimal?
13. How many solutions does the considered optimal control problem have?
14. Can we say that the maximum principle for this example turned out to be ineffective?
15. Is the maximum principle for this example a necessary and sufficient optimality condition for why?
16. What happens if we use a variational inequality for Example 6.1?
17. Is it possible to use the penalty method for an approximate solution of the considered problem?
18. Why does singular control exist in the examples described above, but not in the problems from the previous chapters?
19. Under what conditions can the maximum principle degenerate?
20. How many singular controls can there be?
21. How to obtain an optimization problem in which there would be exactly three singular controls?

Conclusions

Based on the results obtained, the following conclusions can be drawn.

- To solve the considered problem, one can use the maximum principle.
- A feature of this problem is the linearity of the function H with respect to control, as well as the absence of a state function in the adjoint system.
- Choosing the boundary values of the control as a solution to the maximum condition, we obtain a system with respect to the functions u, x, and p, which, in principle, can be solved iteratively.
- The resulting iterative process does not converge for any initial approximation.
- The convergence of the iterative process is impossible due to the absence of a solution to the resulting system.

- For the considered problem, the maximum principle can be fulfilled in a degenerate form, when the pre-control coefficient in the function H definition is zero.
- Admissible controls for which the maximum principle degenerates are singular.
- The set of singular controls for the considered example is infinite and not even countable.
- All singular controls for this example are optimal, which means that the set of optimal controls is infinite and not even countable here.
- Singular controls can exist if the control is equally included in the state equation and in the optimality criterion.
- Singular controls exist if the value in the definition of the function H, which is a multiplier in the term that takes into account all dependence on the control, vanishes. The set of singular controls can be empty, consist of an arbitrary number of elements, or be infinite.
- The set of singular controls can be empty, consist of an arbitrary number of elements, or be infinite.
- The maximum principle for the considered example is a necessary and sufficient optimality condition.

Problems

Based on the results obtained above, we have the following problems related to the subject under study, the solution of which is of undoubted interest.

1. **Coexistence of singular and regular solutions of the maximum principle**. In Example 6.1, there are exclusively singular controls, while in all previous problems the solutions of the maximum principle were not singular controls. We would like to know whether a situation is possible when the maximum principle has both singular and regular solutions at the same time. The answer to this question is given in Appendix.
2. **Finding singular controls**. For Example 6.1, the set U_0 of all singular controls was allocated. It is clear that this is possible only in exceptional cases. The question arises, how can one find special singular controls in a particular situation? For this problem, see Notes[25].
3. **Optimality of singular control**. In Example 6.1, any singular control turns out to be optimal. We would like to know if this is always done? The answer to this question is given in Appendix.
4. **Uncountable set of non-singular optimal controls**. In Example 6.1, there is an uncountable set of optimal controls, all of which are singular. It would be

interesting to establish such a property for the case when the optimal controls are not singular. Such an example is given in Chapter 15.

5. **Existence of singular controls located on the boundary of the set of admissible controls**. In the example considered, the special control turns out to be an internal point of the set of admissible controls. It would be interesting to consider the case when it belongs to the boundary of this set. Such an example is given in Appendix.

6. **Optimality of part of singular controls**. We would like to find out whether some of the singular controls can be optimal, and some can be non-optimal. The answer to this question is also given in Appendix.

7. **Optimality condition for singular control**. For Example 6.1, the proof of the optimality of singular controls did not cause serious difficulties. However, it is clear that such a result is explained by the simplicity of the problem and is established only in exceptional cases. We would like to be able to check in particular case whether a particular singular control is optimal or not. Such a condition is given in Appendix.

6.2 APPENDIX

In the Lecture, an example of the optimal control problem was given, for which there is a set of singular controls that are specific solutions of the maximum principle, and all of them turned out to be optimal. Below is some additional information about singular controls. In particular, the theorems of the uniqueness of optimal control and the sufficiency of optimality conditions for the considered example are used (Section 6.2.1), the control rejected in the process of analyzing regular solutions of the maximum principle turns out to be its singular solution, and it is achieved on the boundary of the set of admissible controls (Section 6.2.2), examples are considered in which the singular controls are not optimal (Section 6.2.3), necessary conditions for the optimality of the singular control are given (Sections 6.2.4 and 6.2.5).

6.2.1 Application of uniqueness and sufficiency theorems

The analysis of Example 6.1, in principle, could be considered complete. However, it is interesting to check the effect in this case of the statements given in the previous chapter. We are talking about Theorems 5.1 and 5.2, which establish the uniqueness of the optimal control and the sufficiency of the optimality condition.

According to Theorem 5.1, the problem of minimizing a strictly convex functional on a convex set cannot have two solutions. The convexity of the set U of admissible controls for Example 6.1 is beyond doubt[26]. Let us check the properties of the optimality criterion whose definition reduces to equality (6.8). Determine the value

$$I\big(\alpha u + (1-\alpha)v\big) \;=\; \frac{1}{2}\Big[\alpha \int_0^1 u(\tau)d\tau + (1-\alpha)\int_0^1 v(\tau)d\tau\Big]^2.$$

We noted earlier that the squaring function is convex, which allows us to establish the inequality

$$I(\alpha u + (1-\alpha)v) \le \frac{\alpha}{2}\Big[\int_0^1 u(\tau)d\tau\Big]^2 + \frac{1-\alpha}{2}\Big[\int_0^1 v(\tau)d\tau\Big]^2 = \alpha I(u) + (1-\alpha)I(v).$$

Thus, the functional I is convex.

Now determine the function u and v such that $u(t) = 1$, $v(t) = -1$ for $t < 1/2$ and $u(t) = -1$, $v(t) = 1$ for $t > 1/2$. Obviously, they are the admissible controls. Then for any α, we get

$$I(\alpha u + (1-\alpha)v) = \alpha I(u) + (1-\alpha)I(v).$$

Thus, the considered functional is not strictly convex, as a result of which we cannot use Theorem 5.1 on the uniqueness of the optimal control. Thus, the absence of uniqueness of the problem solution for Example 6.1 seems quite natural[27].

Let us now turn to Theorem 5.2 on the sufficiency of the optimality condition in the form of the maximum principle. This property is satisfied in the case of non-negativity of the remainder term in the functional increment formula, which has the following form

$$\eta = \eta_3 - \int_0^1 (\eta_1 + \eta_2)dt.$$

Here, the value η_3 is a second-order term obtained as a result of expansion into a Taylor series of a part of the optimality criterion that characterizes the state of the system at the final time, η_1 corresponds to a second-order term in the expansion of the function $H(t,u,x+\Delta x,p)$ in a series in Δx, and $\eta_2 = [H_x(t,v,x,p) - H_x(t,u,x,p)]\Delta x$.

For the considered example, the optimality criterion is determined exclusively by the integral, which means that $\eta_3 = 0$. The function $H = pu - ux$ is linear in x. Then $H(t,u,x+\Delta x,p) = pu - u(x+\Delta x)$. In the right part of this equality, there are no second-order terms with respect to Δx, which means that $\eta_1 = 0$. Finally, we find the value $\eta_2 = (u - v)\Delta x$. Thus, the remainder term is

$$\eta = \int_0^1 (v-u)\Delta x dt.$$

Equation (6.1) implies that the difference $v - u$ is equal to the derivative of Δx. As a result, we find

$$\eta = \int_0^1 \Delta x' \Delta x dt = \frac{1}{2}\int_0^1 \frac{d}{dt}(\Delta x)^2 dt = \frac{1}{2}\big[\Delta x(1)\big]^2.$$

Thus, the remainder term is non-negative, which means that for the considered optimal control problem, the maximum principle gives the necessary and sufficient optimality conditions by virtue of Theorem 5.2. This is consistent with the previously established results.

Return now to the consideration Example 6.2. Of course, we have the convex set of admissible controls, and our functional is strongly convex[28]. Therefore, the corresponding optimal control problem has a unique solution. Determine the sign of the remainder term for this example. We determine again the equality $\eta_3 = 0$ because of the absence of the terminal term at the optimality criterion. For the considered problem, we have the equality $H = pu – x^2/2$. Therefore, $H_x = -x$. This value does not depend from the control, so $\eta_2 = 0$. Besides, $\eta_1 = \Delta x^2$. Thus, the remainder term is non-negative. Using Theorem 5.2, we conclude that the necessary conditions of optimality are sufficient here. However, we have some problems with analysis of Example 6.3 by means the theorems from the previous chapter because of properties of the optimality criterion[29].

6.2.2 Control is optimal as singular and not optimal as regular

Let us consider another example, quite close to the ones above[30].

Example 6.4 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int\limits_0^1 \big[x(t)^2 - 2x(t)t\big]dt$$

on the set

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\},$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Define the function

$$H = pu - (x^2 – 2xt)/2.$$

The adjoint system takes the form

$$p'(t) = x(t) - t,\ t \in (0,1);\ p(1) = 0.$$

Optimal control is determined from the condition of maximum function H on the interval $[-1, 1]$. Here, in principle, two situations are possible: either the solution is achieved on the boundary of a given set, or it is singular.

For the regular case, control is determined by the formula

$$u(t) = \begin{cases} 1, & \text{if } p(t) > 0, \\ -1, & \text{if } p(t) < 0. \end{cases} \tag{6.16}$$

Let us assume that $u(t) = 1$ for any t, which, by virtue of equality (6.16), is possible if the function p is positive everywhere. Then we find $x(t) = t$. In this case, the adjoint system has a solution $p(t) = 0$ for all t. However, this contradicts the previously accepted assumption. Therefore, a control identically equal to one cannot be a solution to the maximum principle[31].

Let now $u(t) = -1$, which is possible if the function p takes only negative values. Then $x(t) = -t$. Consequently, the adjoint equation takes the form $p' = -2t$. Thus, the function p is strictly decreasing. However, at the end point it is zero. Then before this it was positive, which contradicts equality (6.16). However, the case has not yet been ruled out when the control is piecewise constant.

Thus, if there exists a regular solution to the maximum condition, then it cannot be constant. Somewhere there is a discontinuity, which, by virtue of (6.16), means that the function p vanishes there. Let us assume that ξ is the first control discontinuity point, and it does not matter whether there are other such points. Suppose the equality $u(t) = 1$ is initially true that is possible if the function p is positive on the interval $(0, \xi)$. Solving the equation of state, we find $x(t) = t$ at $t \in (0, \xi)$. Consequently, on this interval the adjoint equation takes the form $p'(t) = 0$. Considering that $p(\xi) = 0$, we conclude that over the entire interval where the control is 1, the function p must be equal to zero, which cannot be the case.

The case when the control, when passing through the first break point ξ, switches from the value -1 to 1, also assumes the equality $p(\xi) = 0$. However, now at all previous points the function p must take negative values, and at subsequent points[32] positive. In this case, the equality $x(t) = -t$ at $t \in (0, \xi)$ is valid. Consequently, the adjoint equation here takes the form $p' = -2t$. Then the function p on the specified interval must decrease. However, at the end of this interval it becomes zero, which is possible if it was positive before. Thus, this option is not implemented either.

Thus, we conclude that the maximum condition cannot have regular solutions. Singular control here is possible only if the function p is identically equal to zero. Then the conjugate equation implies the equality $x(t) = t$. This state corresponds to control $u(t) = 1$, which is special. Obviously, this is the unique optimal control[33]. However, what is important for us here is the fact that the control u, identically equal to 1, which we had previously rejected as a possible regular solution to the maximum condition, turned out to be its singular solution, and the optimal one.

6.2.3 Non-optimal singular controls

In Examples 6.1 and 6.2, all singular controls are optimal. A natural question arises: is a singular control always optimal? If this turned out to be the case, then the solution of the optimization problem would be reduced to finding a singular control, if, of course, one exists. Consider the following example[34].

Example 6.5 *It is required to maximize the optimality criterion defined in Example* 6.2 *on the same set of admissible controls.*

As we noted earlier, to solve such a problem, one should define the same function H as for the corresponding minimization problem, but at the same time look for the optimal control from the minimum condition for this function on the set of admissible control values. Thus, the solution of this problem satisfies the equality

$$p(t)u(t) = \min_{|v|\leq 1} p(t)v, \ t \in (0, 1). \tag{6.17}$$

where the function p again is the solution of problem (6.13).

Obviously, in the case when the function p is identically equal to zero, equality (6.17) is satisfied for sure, which corresponds to the degeneration of the optimality condition. However, the state equation and the adjoint system in Examples 6.2 and 6.5 are the same. Therefore, for this problem there is a unique singular control $u = 0$, the same as in the problem of minimizing the considered optimality criterion. It is already known that this control delivers a minimum rather than a maximum to the functional, which means that it cannot be a solution to this optimal control problem[35].

Note that in the general case, the singular controls in the problems of minimization and maximization of the same functional are the same[36]. Consider, in particular, the following example[37].

Example 6.6 *It is required to maximize the optimality criterion defined in Example* 6.1 *on the same set of admissible controls.*

For this problem, there is an infinite set of singular controls U_0 defined in the Lecture. However, these controls again deliver a minimum, and not a maximum, to the optimality criterion, which means that the singular controls are no longer be optimal for the problem being solved in this case. Thus, the maximum principle for this example is not a sufficient condition for optimality[38] We conclude that a singular control can be, but not necessarily optimal[39].

The question arises, what is the solution of the optimization problem considered in Example 6.6? To answer this, let us return to the optimality condition. Taking into account the previously defined form of the function $H = pu–ux$, we have the optimality condition

$$[p(t) - x(t)]u(t) = \min_{|v|\leq 1} [p(t) - x(t)]v, \ t \in (0, 1).$$

Eliminating obviously non-optimal singular controls, we conclude that the solution of the last relation has the form

$$u(t) = \begin{cases} -1, & \text{if } \ p(t) - x(t) > 0, \\ \ \ 1, & \text{if } \ p(t) - x(t) < 0, \end{cases}$$

which differs from formula (6.4) only in sign[40].

Section 6.1.3 established the equality

$$p(t) - x(t) = -\int_0^1 u(\tau)d\tau.$$

Then the previous formula takes the following form

$$u(t) = \begin{cases} -1, & \text{if } \int_0^1 u(\tau)d\tau < 0, \\ \ \ 1, & \text{if } \int_0^1 u(\tau)d\tau > 0, \end{cases} \tag{6.18}$$

which again differs only in sign from relation (6.5). Since the value on the right side of equality (6.18) does not depend on t, we conclude that we are dealing with a constant function. It can only take the value 1 or -1. Obviously, both of them satisfy equality (6.18), and hence are solutions of the optimality condition. Considering that the remaining solutions of the maximum principle, i.e., singular controls are not optimal, we conclude that the optimal control problem for Example 6.7 has exactly two solutions[41]. There are the functions that are identically equal to 1 or -1.

In Examples 3.1–3.4 and 5.1, all solutions of the maximum principle were not singular controls. In contrast, in Examples 6.1 and 6.2 the maximum principle had only singular solutions. Example 6.6 is also interesting because in it the maximum principle has singular and regular solutions at the same time.

Example 6.5 has similar properties. In particular, the conditions of Theorems 5.1 and 5.2 are violated for it, i.e., it is not possible to establish on their basis the sufficiency of optimality conditions and the uniqueness of the maximum principle. Finding optimal controls for this example is not difficult. Since the only singular control is obviously not optimal, from relation (6.17), we can find the function

$$u(t) = \begin{cases} 1, & \text{if } p(t) < 0, \\ -1, & \text{if } p(t) > 0. \end{cases} \tag{6.19}$$

Thus, the optimal control for Example 6.5 satisfies the system (6.12), (6.13), and (6.19). However, this is consistent with the system (5.1), (5.2), and (5.4) for Example 5.1 of previous chapter.

Using the results of Chapter 5, we conclude the set of all regular solutions of the optimality conditions is infinite. For determining the arbitrary solution, we divide the interval [0,1], i.e., the domain of the considered functions, by $2k+1$ equal parts, where $k = 0, 1, \ldots$. Let the control be 1 on the first interval, -1 on two next intervals, 1 on two next intervals, etc. The resulting control u_k^+ and the function $u_k^- = -u_k^+$ satisfy the equalities (6.12), (6.13), and (6.19). The corresponding state functions x_k^+ and $x_k^- = -x_k^+$ are piecewise linear; see Figure 5.3. Calculate the corresponding values of the optimality criterion similar to how it was done in Chapter 5.

$$I(u_k^+) = I(u_k^-) = \frac{1}{2}\int_0^1 (x_k^+)^2 dt = \frac{2k+1}{2}\int_0^{1/(2k+1)} (x_k^+)^2 dt = \frac{1}{6(2k+1)^2}, \; k = 0, 1, \ldots .$$

The maximal values here correspond to the functions which are equal to 1 and -1. There are the solutions of Example 6.5.

6.2.4 Kelley condition

We know that a singular control can be optimal (Examples 6.1, 6.2, and 6.4) or non-optimal (Examples 6.5 and 6.6). The sufficiency of optimality conditions or its absence was determined directly or by means of Theorem 5.2 after finding the sign of the remainder term. This analysis is possible only for easy enough problems. Besides, it does not use properties of singular controls. However, there exists an optimality condition of singular controls[42].

Theorem 6.2 *Suppose the functions of the problem statement are smooth enough. Then the singular control is optimal if it satisfies the* ***Kelley condition***

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} \geq 0. \tag{6.20}$$

By this result, if a singular control satisfies the Kelley condition, then it can be optimal. However, if the inequality (6.20) is false, then the corresponding singular control is non-optimal[43].

Use Theorem 6.2 for the analysis of Example 6.2. The function H is determined by the formula $H = pu – x^2/2$. Find the derivative

$$\frac{\partial H}{\partial u} = p.$$

Using the adjoint system (6.13), we get

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' = x.$$

Using the state equation (6.12), we obtain

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = x' = u.$$

Now we calculate

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 1.$$

Thus, the Kelley condition is true; so the singular control $u = 0$ can be optimal. We know, that this is optimal control, in reality.

Example 6.5 differs from Example 6.2 only in the type of the extremum. The corresponding function H differs from considered before only by the sign. Besides, the state equation and the adjoint system coincide for both examples. Therefore, we obtain the equality

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = -1.$$

The inequality (6.20) gets broken, and the corresponding singular control is not optimal. We determined this result before directly. Thus, by means of the Kelley condition we can exclude non-optimal singular controls.

Try to use Theorem 6.2 for the analysis of Example 6.1 with function $H = pu – xu$. Using equalities (6.1) and (6.2), we obtain

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' - x' = u - u = 0.$$

After differentiation by t and then by u, we prove that the relation (6.20) is true for any singular control in the form of equality. Therefore, all singular controls can be optimal. We know that they are optimal, in reality.

Example 6.6 differs from Example 6.1 only by the type of the extremum. The corresponding functions H differ by the sign. Then we obtain again Kelley condition in the form of equality; and the singular controls can be optimal. However, we know that they are not optimal. This result does not contradict Theorem 6.2, according to which a special control that satisfies the Kelley condition may be optimal, but need not be so.

For Example 6.4, find the derivative

$$\frac{\partial H}{\partial u} = p.$$

Taking into account the adjoint equation, we obtain

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' = x - t.$$

Using the state equation, determine

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = x' - 1 = u - 1.$$

As a result, we get

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 1.$$

Thus, the Kelley condition is satisfied, which means that the special control $u = 1$ can be optimal, and indeed it is.

Verify the validity of the Kelley condition for singular controls from Example 6.3. In this case, the function H is defined by the formula

$$H = pu - \frac{x^3}{3} + \frac{x^2 t}{2}.$$

Its derivative with respect to the control is p. Taking into account the form of the adjoint system, we find

$$\frac{d}{dt}\frac{\partial H}{\partial u} = x^2 - tx.$$

Using the state equation, we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = (2x - t)x' = (2x - t)u - x.$$

Finally, we obtain

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 2x - t.$$

Check the sign of this value for the two singular controls $u = 0$ and $u = 1$. For the first of them, the equality $x(t) = 0$ is valid. This means $2x–t = –t$, which is negative for $t > 0$. Thus, the Kelley condition is not satisfied, which means that the singular control $u = 0$ is certainly not optimal. On the other hand, for $u = 1$ we have $x(t) = t$. This means $2x–t = t$, which is positive. Thus, the singular control $u = 1$ can be optimal[44].

Check the validity of Kelley condition for Example 6.4. Indeed, the derivative of the function H with respect to the control is equal to p. Its derivative with respect to t, by virtue of the adjoint equation, gives $x–t$. As a result of repeated differentiation taking into account the equation of state, we have $u–1$. Finally, the new differentiation with respect to control gives 1, which means Kelley condition is satisfied.

Consider another example.

Example 6.7 *It is required to minimize the functional*

$$I(u) = \int_0^1 (x^6 - 2x^4t^2 + x^2t^4)dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Determine the function

$$H = pu – x^6 + 2t^2x^4 – t^4x^2.$$

Then the adjoint system takes the form

$$p'(t) = 6x(t)^5 - 8t^2x(t)^3 + 2t^4x(t),\ t \in (0,1);\ p(1) = 0.$$

Since the function H is linear with respect to the control and there are no restrictions on the control, there are no regular solutions of the maximum principle.

Singular controls are obtained if the value on the right side of the adjoint equation is equal to zero. Fifth-order algebraic equation

$$6x^5 - 8x^3t^2 + 2xt^4 = 0$$

has five solution

$$x_1 = 0,\ x_2 = t,\ x_3 = -t,\ x_4 = t/\sqrt{3},\ x_5 = -t/\sqrt{3}.$$

They are state functions corresponding to controls

$$u_1 = 0,\ u_2 = 1,\ u_3 = -1,\ u_4 = 1/\sqrt{3},\ u_5 = -1/\sqrt{3}.$$

Thus, for the considered example, the maximum condition has five solutions, which are singular controls.

Let us check the validity of the Kelley condition. Find the derivative

$$\frac{\partial H}{\partial u} = p.$$

Taking into account the adjoint equation, we determine

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' = 6x^5 - 8t^2x^3 + 2t^4x.$$

Differentiate this equality taking into account the state equation. We get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = (30x^4 - 2t^2x^2 + 2t^4)u - 16tx^3 + 8t^3x.$$

Now we obtain

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 30x^4 - 2t^2x^2 + 2t^4.$$

We calculate this value for all five singular controls:

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=0} = 2t^4,\ \frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=\pm t} = -8t^4/3,\ \frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=\pm t/\sqrt{3}} = -8t^4/3.$$

The first three singular controls satisfy the Kelley condition, which means that they can be optimal. The last two do not satisfy this condition, which means that they are obviously not optimal.

Let us estimate the value of the optimality criterion for three singular controls that may turn out to be optimal. Note that the functional to be minimized can be written in the form

$$I(U) = \int_0^1 \big[x(x-t)(x+t)\big]^2 dt.$$

Obviously, all its values are non-negative, and equality to zero is realized only in three cases, which correspond to special controls that satisfy the Kelley condition. Therefore, the optimal control problem under consideration has three solutions, which were previously defined.

Now we consider the following example.

Example 6.8 *It is required to maximize the optimality criterion defined in Example* 6.7 *on the same set of admissible controls.*

Due to the linearity of the function H and the absence of restrictions on the maximum principle solution controls, here there can also be only singular controls, which are the same as for Example 6.7. In this case, in the Kelley condition, it is required to change the sign $\geq$ to $\leq$. Then the previously defined singular controls u_0, u_1, and u_2 do not satisfy this condition and, therefore, are not optimal. The controls u_3 and u_4 satisfy the Kelley condition, which means that they can be optimal, and would turn out to be so if the problem is solvable[45].

We have carried out a complete analysis of all seven examples of this chapter[46].

6.2.5 Kopp–Moyer condition

As seen in Examples 6.1 and 6.6, there are situations where the Kelley condition is true in the form of equality. Then we do not get any information about the optimality of the singular control, as a result of which this case can be interpreted as a degeneration of the Kelley condition. However, sometimes it is possible to establish an additional optimality condition for a singular control.

Example 6.9 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 x_1(t)^2 dt$$

on the set

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\}$$

where x_1 is a solution of the Cauchy problem

$$x_1'(t) = x_2(t),\ x_2'(t) = u(t),\ t \in (0,1);\ x_1(0) = 0,\ x_2(0) = 0. \tag{6.21}$$

Consider the function

$$H = p_1 x_2 + p_2 u - x_1^2/2,$$

where p_1 and p_2 are solutions of the adjoint system

$$p_1'(t) = x_1(t),\ p_2'(t) = -p_1(t),\ t \in (0,1);\ p_1(1) = 0,\ p_2(1) = 0. \tag{6.22}$$

The degeneration of the maximum principle here is possible for $p_2 = 0$. Then it follows from the second equation (6.22) that $p_1 = 0$, and from the first equation that $x_1 = 0$. Substituting this value into the first equation (6.21), we find $x_2 = 0$. It follows from the first equation (6.21) that $u = 0$. This function is the unique special singular for this example.

Let us check the validity of the Kelley condition. We find

$$\frac{\partial H}{\partial u} = p_2.$$

Taking into account the second adjoint equation, we determine

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p_2' = -p_1.$$

Differentiate this equality taking into account the second adjoint equation of state. We get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = -p_1' = -x_1. \tag{6.23}$$

This implies

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 0.$$

Thus, the Kelley condition is satisfied in the form of equality (i.e., degenerates), as a result of which we cannot conclude that the singular control is optimal. However, one can try to use the following assertion[47].

Theorem 6.3 *If for a singular control the Kelley condition degenerates, then for its optimality, it is necessary to fulfill the* ***Kopp–Moyer condition***

$$\frac{\partial}{\partial u}\frac{d^4}{dt^4}\frac{\partial H}{\partial u} \leq 0. \tag{6.24}$$

Let us check the validity of the Kopp–Moyer condition for Example 6.9. We differentiate equality (6.23) taking into account the first equation (6.21). We have

$$\frac{d^3}{dt^3}\frac{\partial H}{\partial u} = -x_1' = -x_2.$$

After repeated integration, taking into account the second equation (6.21), we obtain

$$\frac{d^4}{dt^4}\frac{\partial H}{\partial u} = -x_2' = -u.$$

Finally, we determine the value

$$\frac{\partial}{\partial u}\frac{d^4}{dt^4}\frac{\partial H}{\partial u} = -1.$$

Thus, for Example 6.9, the Kopp–Moyer condition is satisfied, which means that the singular control $u = 0$ may turn out to be optimal. Indeed, the minimized functional is non-negative, and its equality to zero is possible only for $x_1 = 0$. The resulting equality is realized for $x_2 = 0$, which is true only for $u = 0$. Thus, the singular control is the only optimal control.

It is now natural to consider the functional maximization problem from Example 6.9.

Example 6.10 *It is required to maximize the optimality criterion defined in Example* 6.9 *on the same set of admissible controls.*

As is known, in the problems of minimization and maximization of the same functional, the singular controls are the same. However, when using the Kopp–Moyer condition for the maximization problem in relation (6.24), the sign should be changed. Considering that the function H is the same in both problems, we conclude that for Example 6.10 the Kopp–Moyer condition is not satisfied. Thus, the only singular control $u = 0$ is certainly not optimal, which is not surprising, since it delivers a minimum, not a maximum, to the given functional. Thus, when examining Example 6.10, Theorem 6.3 proved to be effective. The optimal controls here are regular solutions of the maximum principle.

We can now try to use Theorem 6.3 to analyze Examples 6.1 and 6.6. However, in the previous subsection, the inequality was established for them

$$\frac{d}{dt}\frac{\partial H}{\partial u} = 0.$$

This implies the degeneration of the Kopp–Moyer condition. Thus, using this result, we cannot distinguish the problem of minimizing a functional from the problem of its maximization.

The following examples are natural generalizations of the problems considered earlier.

Example 6.11 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 x_1(t)^2 dt$$

on the set

$$U = \{u|\ |u(t)| \le 1,\ t \in (0,1)\},$$

where x_1 is determined from the system

$$x_1'(t) = x_2(t),\ x_2'(t) = x_3(t),\ x_3'(t) = u(t),\ t \in (0,1);\ x_i(0) = 0,\ i = 1,2,3.$$

Example 6.12 *It is required to maximize the optimality criterion defined in Example* 6.11*.*

Consider the function

$$H = p_1 x_2 + p_2 x_3 + p_3 u - x_1^2/2,$$

where p_i are solutions of the adjoint system

$$p_1'(t) = x_1(t),\ p_2'(t) = -p_1(t),\ p_3'(t) = -p_2(t),\ t \in (0,1);\ p_i(1) = 0,\ i = 1,2,3.$$

The maximum principle for these examples degenerates at $p_3 = 0$. Substituting this value into the third adjoint equation, the result into the second, and the result into the first, we conclude that $x_1 = 0$. Substituting this value into the first equation of state, the result into the second, and the result into the third, we conclude that $u = 0$. The found function is the unique singular control of the considered examples.

Calculate the value

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = \frac{d^2 p_3}{dt^2} = -\frac{dp_2}{dt} = p_1.$$

Hence, after differentiation with respect to the control, the Kelley condition degenerates. We continue to calculate the derivatives

$$\frac{d^4}{dt^4}\frac{\partial H}{\partial u} = \frac{d^2 p_1}{dt^2} = \frac{dx_1}{dt} = x_2.$$

As a result of differentiation of this equality with respect to control, we now establish the degeneration of the Kopp–Moyer condition too. However, the following statement is true[48]:

Theorem 6.4 *If the following conditions are true for a singular control*

$$\frac{\partial}{\partial u}\frac{d^s}{dt^s}\frac{\partial H}{\partial u} = 0,\ s = 0, ..., 2r-1;\ \frac{\partial}{\partial u}\frac{d^{2r}}{dt^{2r}}\frac{\partial H}{\partial u} \ne 0,$$

then for its optimality it is necessary to fulfill the ***generalized Kopp–Moyer condition*** *of order r*

$$(-1)^r \frac{\partial}{\partial u}\frac{d^{2r}}{dt^{2r}}\frac{\partial H}{\partial u} \le 0.$$

We continue the analysis of the considered examples. Then

$$\frac{d^6}{dt^6}\frac{\partial H}{\partial u} = \frac{d^2 x_2}{dt^2} = \frac{dx_3}{dt} = u.$$

Now we obtain

$$\frac{\partial}{\partial u}\frac{d^6}{dt^6}\frac{\partial H}{\partial u} = 1.$$

Thus, for Example 6.11 the generalized Kopp–Moyer condition of the third order is satisfied, and for Example 6.12 it is violated. Then, according to Theorem 6.4, the control $u = 0$ is certainly not optimal for Example 6.12, but may be optimal for Example 6.11. Obviously, in the latter case it is the unique optimal control. As for Example 6.12, there are two regular optimal controls $u = 1$ and $u = -1$ for it. It is easy to see that for Examples 6.12 and 6.6, Theorem 6.4 also turns out to be ineffective.

Let us consider generalizations of the last problems.

Example 6.13 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 x_1(t)^2 dt$$

on the set

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\},$$

where x_1 is determined from the system

$$x_i'(t) = x_{i+1}(t),\ i = 1,...,r-1,\ x_r'(t) = u(t);\ t \in (0,1);\ x_i(0) = 0,\ i = 1,...,r.$$

Example 6.14 *It is required to maximize the optimality criterion defined in Example 6.13.*

Determine the function

$$H = \sum_{i=1}^{r-1} p_i x_{i+1} + p_r u - x_1^2/2,$$

where the functions p_i satisfy the equalities

$$p_1'(t) = x_1(t),\ p_i'(t) = -p_{i-1}(t),\ i = 2,...,r,\ t \in (0,1);\ p_i(1) = 0,\ i = 1,...,r.$$

The maximum principle now degenerates at $p_r = 0$. Substituting this value into the last adjoint equation, the result obtained into the penultimate one, etc., having reached the first equation, we establish that $x_1 = 0$. Substituting this value into the first state equation, the result into the second one, etc., having reached the last one, we conclude that $u = 0$. The found function is the only singular control for the examples under consideration.

Using adjoint equations, we sequentially calculate the values

$$\frac{\partial H}{\partial u} = p_r,\ \frac{d}{dt}\frac{\partial H}{\partial u} = -p_{r-1}, ...,\ \frac{d^{r-1}}{dt^{r-1}}\frac{\partial H}{\partial u} = (-1)^{r-1}p_1,\ \frac{d^r}{dt^r}\frac{\partial H}{\partial u} = (-1)^r x_1.$$

We note that so far no explicit control has been encountered anywhere. Thus, as a result of differentiation with respect to the control of the value on the left side of any of these equalities, we obtain zero. This means that while the Kelley condition and its generalizations degenerate.

We continue differentiation taking into account the equations of state until the first appearance of the control. We have

$$\frac{d^{r+1}}{dt^{r+1}}\frac{\partial H}{\partial u} = (-1)^r x_2,\ \frac{d^{r+2}}{dt^{r+2}}\frac{\partial H}{\partial u} = (-1)^r x_3, ...,$$

$$\frac{d^{2r-1}}{dt^{2r-1}}\frac{\partial H}{\partial u} = (-1)^r x_r,\ \frac{d^{2r}}{dt^{2r}}\frac{\partial H}{\partial u} = (-1)^r u.$$

As a result, we find

$$(-1)^r \frac{\partial}{\partial u}\frac{d^{2r}}{dt^{2r}}\frac{\partial H}{\partial u} = 1.$$

Thus, for Example 6.13, the generalized Kopp–Moyer condition of order r is satisfied, which means that the singular control $u = 0$ found earlier can be optimal. It really is such, since this functional cannot have negative values, and zero is reached exclusively on the indicated control. At the same time, for the maximization problem, under the condition of optimality of the singular control, the inequality sign should be changed to the opposite ones. As a result, it turns out that the generalized Kopp–Moyer condition of order r for Example 6.14 is not satisfied, which means that the singular control $u = 0$ is obviously not optimal here. The solutions of this problem turn out to be non-singular solutions of the maximum condition $u = 1$ and $u = -1$.

Additional conclusions

Based on the results presented in Appendix, some additional conclusions can be drawn about the properties of singular controls.

- The optimality criterion for Example 6.1 is convex, but not strictly convex, so the optimal control uniqueness theorem is not applicable.
- The remainder term in the functional increment formula for Example 6.1 is nonnegative, which allows us to establish the sufficiency of the maximum principle for it using the corresponding theorem.
- The singular control for Example 6.4 belongs to the boundary of the set of admissible controls.
- An optimal control problem can have no singular controls at all, have only singular solutions of the maximum principle, or have both singular and regular solutions of the maximum principle at the same time.

- The singular controls in the problem of minimization and maximization of an arbitrary functional are the same.
- The singular controls can be both optimal and non-optimal.
- A situation is possible when some singular controls are optimal, while others are not optimal.
- If a singular control does not satisfy the Kelley condition, then it is not optimal.
- If a singular control satisfies the Kelley condition, then it may be optimal, but not necessarily so.
- In the case of degeneracy of the Kelley condition, one can use the Kopp–Moyer condition to test for the optimality of the singular control.
- If a singular control does not satisfy the Kopp–Moyer condition, then it is not optimal.
- If a singular control satisfies the Kopp–Moyer condition, then it may be optimal, but it need not be.
- If the Kopp–Moyer condition degenerates to estimate the optimality of a singular control, then in some cases, it is possible to use the generalized Kopp–Moyer condition.
- If a singular control does not satisfy the generalized Kopp–Moyer condition, then it is not optimal.
- If a singular control satisfies the generalized Kopp–Moyer condition, then it may be optimal, but it is not necessarily so.

Notes

1. The book [75] is devoted directly to the theory of singular controls; see also [42], [160], [190], [205]. For singular controls for distributed parameters systems see [12], for minimax problems see [6].

2. This example is given in [75]; see also [170]. In Section 15, we will consider the problem of minimizing the same functional in the presence of an additional isoperimetric condition; see Example 15.8

3. For the considered problem, one could also use the necessary optimality conditions in the form of variational inequalities too; see Chapter 4. The corresponding variational inequality takes the form $H_u(v-u) \leq 0$ for all $v \in [-1,1]$, where H_u is the derivative of the function H with respect to the control. This function for the considered example is equal to $H = pu – xu$, so we conclude the variational inequality $(p-x)(v-u) \leq 0$ for all $v \in [-1,1]$, which in coincides with relation (6.3). Thus, for Example 6.1, the maximum principle and the variational inequality are coincide. Example of the coincidence of these optimality conditions is given in Chapter 9, which deals with optimal control problems for systems with a fixed final state.

4. This circumstance should not surprise us, since the function H is linear with respect to the control.

5. We encountered a similar situation in Example 2.8, for which the algorithm for solving the stationary condition diverged for some initial approximations, but converged for others.

6. In fact, we could not check the convergence of the iterative process with the initial approximation of the control $u_0 = -1$, since we received the value -1 as a control at the first iteration with the previous choice of the initial approximation.

7. The same technique was used in Chapters 3 and 5, when, having excluded the control and the state function from the system of optimality conditions, we obtained a boundary value problem for the function p for a non-linear second-order differential equation. This corresponds to the elimination method.

8. This equation, apparently, should be classified as integral, since the unknown function here is under the integral sign. However, this is clearly not the object that is usually studied in the standard theory of ***integral equations***; see [107], [151].

9. This is not surprising, since otherwise the previously described iterative process would certainly converge.

10. This circumstance explains the fact that the iterative process for solving this system does not converge. It can be seen that by choosing a discontinuous control as the initial approximation, as was done in the study of Example 5.1, we again obtain the oscillation of the algorithm. However, it cannot converge, because there is nowhere to converge.

11. We will return to proving the existence of an optimal control for Example 6.1 in a later chapter.

12. The existence of a solution is not yet obvious, but we will find it soon.

13. Formula (6.6) corresponds to the ***Fredholm integral equation of the first kind***. In general, such an equation is

$$\int_a^b K(t,\tau)u(\tau)d\tau = f(t), \ t \in (a,b).$$

For our case, $a = 0$, $b = 1$, $K(t,\tau) = 1$, $f(t) = 0$ for all $t,\tau \in (0,1)$. About integral equations; see [107], [151].

14. In Chapter 1, we considered a function that has a continuum of minimum points.

15. In Chapter 11, we will consider the problem of minimizing the considered functional on a subset of such controls from U that ensure the fulfillment of some additional condition. Chapter 15 gives an example of an optimal control problem in which the set of optimal controls is also infinite and uncountable. However, they are all discontinuous and are not singular controls. Chapter 17 will consider the problem of minimizing the functional from Example 6.1 in the absence of an initial condition. Moreover, if in this case all singular controls are optimal, and the optimality conditions turn out to be necessary and sufficient, then in Example 17.9 the same singular controls are not optimal, and a solution to the problem does not exist.

16. Chapter 9 gives an example for which conditions (6.9) are implemented; see Example 9.1 However, there is no admissible control under which equality (6.11) is satisfied. As a result, there is no singular control there.

17. Chapter 8 will show that the considered optimal control problem is Tikhonov ill-posed. Chapter 12 will consider the problem of minimizing the functional from Example 6.2 with an additional boundary condition (see Example 12.3), and Chapter 15 with an isoperimetric condition (see Example 15.4) and Chapter 17 without an initial condition.

18. Chapter 11 will consider a fairly close optimal control problem for a system with a fixed final state. Moreover, it is the presence of an additional condition at a finite time that makes it quite easy to prove the absence of a non-singular solution of the maximum condition.

19. It is easy to see that problem (6.12), (6.13), and (6.15) has no solution, and the corresponding iterative process does not converge. Let, for example, in accordance with formula (6.15) the initial approximation $u_0 = 1$ be chosen. The solution to problem (6.12) here is equal to $x_0(t) = t$. We find the solution of the adjoint system $p_0(t) = (t^2 – 1)/2$. Since this value is negative, in accordance with formula (6.15) we determine a new approximation control $u_1 = 1$. It corresponds to the solution $x_1(t) = –t$ of problem (6.12) and the solution to the adjoint system $p_1(t) = (1 – t^2)/2$. Therefore, $u_2 = 1$, i.e., we are back to the initial approximation. It is easy to see that a similar situation is observed for a different choice of the initial approximation. We will meet with the system (6.12), (6.13), and (6.15) in the next chapter when analyzing another optimal control problem with fundamentally different difficulties; see Example 7.1.

20. For the problem under consideration, the equalities $f(t, u, x) = u$, $g(t, u, x) = x^2/2$ are valid. This implies the fulfillment of relations (6.9) for $f_1(t, x) = 1$, $f_2(t, x) = 0$, $g_1(t, x) = x^2/2$, $g_2(t, x) = 0$, $\varphi(u) = u$. Equality (6.10) takes the form $p(t) = 0$. This is possible only for the control $u(t) = 0$, which is the unique solution to this problem. It is easy to see that for Example 6.2 the conditions of Theorems 5.1 and 5.2 are satisfied, which implies, respectively, the uniqueness of the optimal control and the sufficiency of the optimality conditions. Indeed, the considered Example differs from Example 3.3 only in a simpler form of the minimized functional. In the previous Chapter, the uniqueness of the optimal control and the sufficiency of the optimality condition for Example 3.3 were proved. Rejecting the square of the control under the integral in the optimality criterion does not change its properties of lower boundedness, continuity and strict convexity, as well as the sign of the remainder term. Thus, the uniqueness of the solution of the problem and the sufficiency of the optimality condition for Example 6.2 can be established quite easily. In the next chapter, we will prove the existence of an optimal control for it. Moreover, it is easy to show that for this example, the variational inequality leads to the same result as the maximum principle. Further exploration of this example will be continued in Chapter 8, where it will be established that this problem has, however, one unusual property.

21. In the next chapter, another example will be considered with a single special control, but with qualitatively different properties.

22. Chapter 15 will consider a similar problem with an additional isoperimetric constraint; see Example 15.6.

23. It is easy to establish that with a given equation of state it is possible to choose such an integrand function $f = f(t, x)$ so that its derivative with respect to the second argument vanishes for n different functions x vanishing at $t = 0$. Then the derivatives of all these functions turn out to be singular controls if the set of admissible controls is chosen so that it includes them all. Thus, the optimal control problem can include an arbitrary number of singular controls. In particular, Example 6.7 considers a problem with five singular controls, and Chapter 11 deals with three singular controls, of which only two are optimal. In another example from Chapter 11 there are also three singular controls, but none of them is optimal.

24. This problem is clarified in Section 6.2.4.

25. To find the singular control in practice, ***regularization methods*** can be used. In this case, a certain term with a small parameter is introduced into the equation of state or into the optimality criterion so that the regularized problem can no longer contain singular controls (the conditions of Theorem 6.1 are violated). Due to the smallness of the corresponding parameter, the regularized problem turns out to be quite close to the original one, which in some cases allows one to choose the solution of the regularized problem as an approximate solution to the original problem; see [205]. One such algorithm will be discussed in Chapter 8 to overcome another difficulty. For general regularization methods; see [79], [187], [194]. We also note that an approximate solution to the problem can be found using the previously described penalty method, which is not related to the problem of degeneracy of the maximum principle.

26. It exactly coincides with the set of admissible controls for the previously considered Examples 3.3 and 5.1.

27. Recall that Chapter 1 dealt with the problem of minimizing a convex but not strictly convex function $f(x) = |x|$, which nevertheless has a unique solution. This is quite natural, since Theorem 5.1 gives only sufficient conditions for the uniqueness of the optimal control, i.e., violation of the conditions of the theorem does not necessarily lead to violation of its statements.

28. These results are obtained, in reality, in Chapter 5.

29. This functional is non-convex, besides we cannot determine the sign of the corresponding remainder term. In reality, the properties of Example 6.3 are not very easy; see Appendix.

30. Chapter 17 considers the same example but without the initial condition; see Example 17.8.

31. In fact, we have only proved that it cannot be a regular solution of the maximum principle.

32. More precisely, to the next break point, if it exists.

33. Let us note that the integrand of the functional being minimized can be written in the form $[x(t)–t]^2 + t^2$. Here, the second term does not depend on the control, and the first, being non-negative, can be identically equal to zero only if the equality $x(t) = t$ is satisfied for all t. According to the state equation, this corresponds to a control identically equal to unity, which thereby turns out to be the unique solution to the considered optimal control problem.

34. Chapter 15 considers the problem of maximizing the same functional in the presence of an additional isoperimetric condition; see Example 15.5.

35. Obviously, the optimal controls here are functions that are identically equal to 1 or –1 and are regular solutions of the maximum principle.

36. This is explained by the fact that in both cases the same function H is considered, only in one case it is minimized, and in the other it is maximized. However, the presence or absence of degeneracy of the optimality condition is not related to the type of extremum.

37. Chapter 15 considers the problem of maximizing the same functional in the presence of an additional isoperimetric condition; see Example 15.9.

38. The absence of sufficiency of the maximum principle for Examples 6.3 and 6.6 can be investigated using Theorem 5.2. According to this statement, the sufficiency of the optimality condition is guaranteed if the remainder term in the functional increment formula is non-negative and is not guaranteed otherwise. It is clear that the remainder term in the maximization prob-

lem will be equal to the remainder terms in the corresponding minimization problem, taken with the opposite sign. In particular, for Example 6.1, an explicit form of the remainder term was found $\eta = \frac{1}{2}\big[\Delta x(1)\big]^2$. Then for Example 6.6 we get $\eta = -\frac{1}{2}\big[\Delta x(1)\big]^2$. Since this value is, generally speaking, negative, the condition of the sufficiency theorem is violated, which means that we cannot use it to establish the sufficiency of the optimality conditions. In fact, sufficiency is not realized, as we established above.

39. Actually, there is nothing surprising in this. A singular control is a specific solution of the maximum principle, which is a necessary, but generally speaking, not a sufficient condition for optimality. Being a solution of the maximum principle, a singular control may turn out to be optimal, but does not have to be.

40. We recall again that the problem of maximizing a certain functional is reduced to the problem of minimizing the same functional taken with the opposite sign.

41. In Section 6.2.1, it was stated that the minimized functional in Example 6.1 is convex, but not strictly convex. If we go to Example 6.6 and change the sign, then the corresponding functional will not even be convex. As a result, the conditions of Theorem 5.1 on the uniqueness of the optimal control are violated. In this regard, the presence of two solutions for Example 6.6 is not surprising.

42. The proof of Theorem 6.2 is given in [42], [75].

43. There is some analogy here with Theorem 1.6 given in Chapter 1. If at a stationary point the second derivative of the minimized function is positive, then at this point the local minimum of this function is reached. Thus, the violation of this condition indicates that the stationary point does not minimize the function under consideration. In both cases, we are dealing with second-order extremum conditions. Note also the ***Legendre condition*** of the calculus of variations, where the second derivative also appears; see [37], [61]. In a sense, we can consider that the Kelley condition is a generalization of the Legendre condition.

44. It is also necessary to prove that the optimal control problem has a solution. This problem is considered in the next chapter. Due to the necessity of the maximum principle, the optimal control satisfies it. The singular control $u = 0$ is not optimal due to the Kelley condition. Then the second singular control $u = 1$ is surely optimal. In particular, one can find the value of the optimality criterion for both singular controls: $I(0) = 0$, $I(1) = -1/24$.

45. In reality, the maximized functional can turn out to be arbitrarily large with a proper choice of control, which is possible due to the absence of restrictions on the values of the control.

46. We also note that in all considered problems, except for Example 6.3, they are invariant under control sign change, i.e., changing any control to a function with the opposite sign leads to exactly the same value of the optimality criterion. In Example 6.3, changing the sign of the control leads out of the set of valid controls and leads to a change in the value of the criterion.

47. The proof of this theorem is given in [42], [75].

48. The proof of this theorem is given in [42], [75].

CHAPTER 7

Unsolvability of optimal control problems

To solve optimal control problems, earlier, as a rule, the maximum principle was used, which is a system that includes a state equation, an adjoint system, and a maximum condition for a certain function. The practical solving of this system is most often carried out iteratively. In the process of using this technique to solve even fairly simple problems, certain difficulties may arise. In particular, the system of optimality conditions can have too many solutions, which are not always optimal controls. However, the opposite situation is also possible. The iterative process for solving the corresponding system of optimality conditions sometimes does not converge for any initial approximation. We encountered a similar phenomenon when the maximum principle degenerates, when its solution turns out to be singular. However, this situation is also possible in the absence of singular control. The algorithm simply has nowhere to converge, if the problem has no solution at all. The subject of this chapter is the study of the solvability of optimal control problems.

7.1 LECTURE

Below is an example of a fairly simple optimal control problem for which the iterative process for a system of optimality conditions does not converge for any initial approximation. The reason for this situation is the absence of optimal control[1]. An existence theorem for a solution to an optimization problem is proved, which is used to justify the solvability of the problems considered earlier.

7.1.1 Statement of the problem and its analysis

To solve optimal control problems in the previous chapters, as a rule, we applied the maximum principle. The resulting system of optimality conditions is generally solved iteratively. In the process of studying this system, we encountered difficulties due to the absence of a sufficiency condition for the optimality or uniqueness of the solution to the problem, as well as the degeneration of the maximum principle. However, there

DOI: 10.1201/9781003398585-7

is another serious nuisance that we already encountered in Part I when studying the problem of minimizing a function of one variable[2].

Example 7.1 *It is required to minimize the functional*

$$I(u) = \int_0^1 \left[x(t)^2 - u(t)^2\right] dt$$

on the set

$$U = \{u \mid |u(t)| \leq 1,\ t \in (0,1)\},$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{7.1}$$

We considered this system and the set of admissible controls for Examples 3.3, 5.1, 6.1, 6.2, 6.5, and 6.6. However, the optimality criterion is another. For example, the both terms under the integral have the sign plus in Example 3.3 and the sign minus in Example 5.1. For the first case, the problem has a unique solution, and the optimality conditions are necessary and sufficient, but for the second one, the solution of the problem is non-unique, and the optimality conditions are only necessary. We will prove that under the presence of the terms with different signs, the optimal control problem has another property.

Try to analyze this problem using the standard method. Determine the function $H = pu - (x^2 – u^2)/2$. Then the adjoint system takes the form

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0. \tag{7.2}$$

Now the optimal control u satisfies the maximum condition

$$H(u) = \max_{|v| \leq 1} H(v). \tag{7.3}$$

Equaling to zero the derivative of the function H with respect to the control, determine $u = p$. The corresponding second derivative is 1, so we have the point of minimum. In this situation, we can have the maximum of the function at the boundary points. Find the values

$$H(1) = p – (x^2 - 1)/2,\ \ H(-1) = -p - (x^2 - 1)/2.$$

Choosing the maximum of them, we get

$$u(t) = \begin{cases} 1, & \text{if } p(t) > 0, \\ -1, & \text{if } p(t) < 0. \end{cases} \tag{7.4}$$

Thus, we can find the solution of the problem from the system (7.1), (7.2), (7.4) that is an analog of the optimality conditions for the examples, which were considered before[3]. Try to solve it using the iterative method.

The function u of formula (7.4) can have only two values. Choose the control $u_0 = 1$ as the initial iteration. This corresponds to the function p that is positive everywhere. Putting this control to the state equation, we find $x_0(t) = t$. Then the solution of the adjoint system is

$$p_0(t) = -\int_t^1 x_0(\tau)d\tau = -\int_t^1 \tau d\tau = \frac{t^2-1}{2}.$$

This function is negative. Then the next iteration is $u_1 = -1$. After solving of the problem (7.1), we find $x_1(t) = -t$. The corresponding solution of problem (7.2) is

$$p_1(t) = -\int_t^1 x_1(\tau)d\tau = \int_t^1 \tau d\tau = \frac{1-t^2}{2}.$$

This function has positive values only, so the next iteration of control is 1. Thus, we return to our initial iteration. If we choose −1 as initial iteration of control, then we get $u_1 = 1$, and after iteration we return again to the value −1.

However, it is possible that the control has jumps; see Example 5.1. Then the function p changes the sign at the jump points of the control. Suppose ξ is a minimal point with these properties. Let the control be 1 before this value, and −1 after that. This corresponds to the inequality $p(t) > 0$ for $t < \xi$ and $p(t) < 0$ for $t > \xi$. Then the solution of the problem (7.1) is $x(t) = t$ for $t < \xi$. Solve the adjoint equation with boundary condition $p(\xi) = 0$; we find $p(t) = (t^2 – \xi^2)/2$. This value is negative for $t < \xi$, but we would like to obtain positive values here. Therefore, the control that changes the sign from 1 to −1 does not exist. Suppose now the $u = -1$ before the point ξ and $u = 1$ after this point. Then $t(t) = -t$ for $t < \xi$. The corresponding solution of the adjoint equation with boundary condition $p(\xi) = 0$ is $p(t) = (\xi^2 – t^2)/2$. This is positive on the considered interval, not negative. Therefore, our supposition is not realized[4], so the control determined by formula (7.4) cannot to be discontinuous too. Thus, the problem (7.1), (7.2), and (7.4) has no solution, and the corresponding iterative method does not converge for any initial iteration[5].

At first glance, the conclusion drawn is not so catastrophic. We had analogical situation in the previous chapter, when we analyzed Example 6.1 and 6.2. Particularly, the system (6.1), (6.2), and (6.4) is very similar to the above system (7.1), (7.2), and (7.4), and the system (6.12), (6.13), and (6.15) is exactly the same with it. For both these cases, the iterative method does not converge because of the absence of the solutions for the considered system. However, we found the corresponding optimal controls. This is clear because these systems were not equivalent to the maximum principle, and the solutions of the problems were singular control. The principle difference of Example 7.1 from previous examples is the absence of singular control. We can prove it using Theorem 6.1. Particularly, the function H includes the square of control without any coefficient that can be zero[6]. Therefore, formula (7.4) is equivalent to the maximum condition (7.3).

Thus, no admissible control satisfies the necessary optimality conditions. We had an analogical situation in Chapter 1 for the minimization problem of the function

$f(x) = x$. The corresponding stationary condition gives the false equality 1=0. The meaning of the necessary extremum condition is that any solution of an extremal problem must satisfy it, which means that the set of solutions of the optimization problem is embedded in the set of controls that satisfy the necessary optimality condition. If the last set turns out to be empty, then, generally speaking, a narrower set of solutions to a given optimal control problem is inevitably empty. Thus, this problem turns out to be unsolvable[7]. Under these conditions, the maximum principle and the original optimal control problem are equivalent, i.e., the optimality conditions turn out to be necessary and sufficient[8].

In this case the divergence of the algorithm for solving the optimality conditions for any initial approximation is due to the fact that there is simply nowhere to converge[9].

We proved the unsolvability of the problem, starting from the optimality condition. However, it would be interesting to establish this result directly without using to the maximum principle.

7.1.2 Unsolvability of the optimization problem

Let us try to give a direct proof of the unsolvability of the optimization problem. Taking into account the definition of the set of admissible controls, we establish the inequalities

$$x(t)^2 \geq 0,\ u(t)^2 \leq 1,\ t \in (0,1).$$

Then, for any admissible control, we get

$$I = \frac{1}{2}\int_0^1 x^2 dt - \frac{1}{2}\int_0^1 u^2 dt \geq \frac{1}{2}. \tag{7.5}$$

Inequality (7.5) gives a lower bound for the value of the functional to be minimized on the set of admissible controls.

Consider a sequence of controls characterized by equality; see Figure 7.1

$$u_k(t) = \begin{cases} 1, & \text{if } 2j/2k \leq t < (2j+1)/2k, \\ -1, & \text{if } (2j+1)/2k \leq t < (2j+2)/2k, \end{cases} \tag{7.6}$$

where $j = 0, 1, ..., k–1$. Obviously, all these functions belong to the set of admissible controls.

Find the solution x_k of problem (7.1) for $u = u_k$. For $2j/2k \leq t < (2j+1)/2k$ we get

$$x_k(t) = \int_0^t u_k(\tau)d\tau = \sum_{i=0}^{j-1}\Big[\int_{2i/2k}^{(2i+1)/2k} u_k(\tau)d\tau + \int_{2i+1/2k}^{(2i+2)/2k} u_k(\tau)d\tau\Big] +$$

$$+ \int_{2j/2k}^{t} u_k(\tau)d\tau = \sum_{i=0}^{j-1}\Big(\frac{1}{2k} - \frac{1}{2k}\Big) + \Big(t - \frac{2j}{2k}\Big) = t - \frac{2j}{2k}.$$

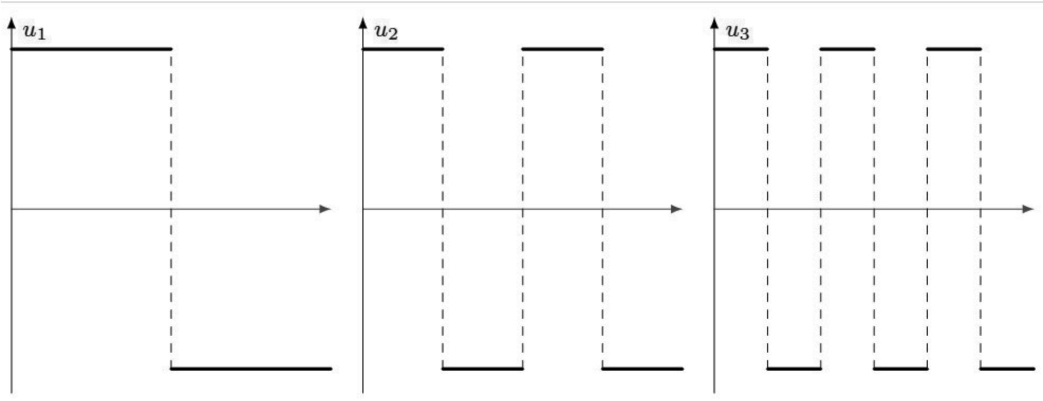

Figure 7.1 Minimizing sequence for Example 7.1.

Analogically, for $(2j+1)/2k \le t < (2j+2)/2k$ we obtain

$$x_k(t) = \frac{1}{2k} - \Big(\frac{2j+1}{2k} - t\Big) = \frac{2j+2}{2k} - t.$$

The elements of the state function sequence[10] are shown in Figure 7.2.

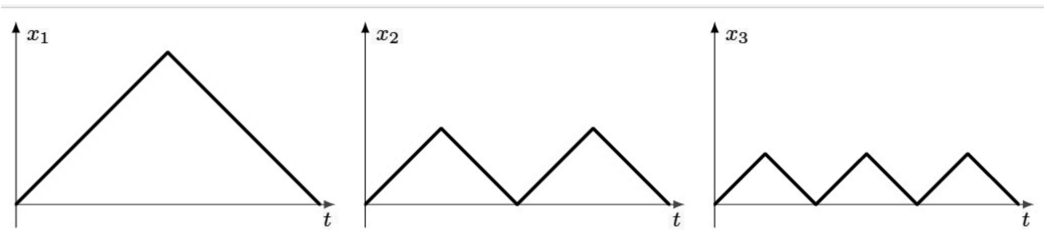

Figure 7.2 State sequence for Example 7.1.

Based on the established relations, we obtain the inequality

$$0 \le x_k(t) \le 1/2k,\ t \in (0,1),\ k = 1, 2, \dots.$$

Using the inequality (7.5), we get

$$-\frac{1}{2} \le I_k = \frac{1}{2}\int\limits_0^1 (x_k^2 - u_k^2)dt \le \frac{1}{8k^2} - \frac{1}{2},\ k = 1, 2, \dots.$$

Passing to the limit as $k \to \infty$, determine $I_k \to -1/2$.

By inequality (7.5), on any admissible control, the value of the minimized functional cannot be less than $-1/2$. At the same time, for a sufficiently large number k on the control u_k defined by formula (7.6), the value of the functional is arbitrarily close to the value of $-1/2$. This means that the number $-1/2$ is the lower bound of the considered functional on the set of admissible controls.

Definition 7.1 *A sequence of admissible controls is called **minimizing** if the values of the minimized functional on it tend to its lower bound on the set of admissible controls.*

Thus, the sequence of controls determined by formula (7.6) is minimizing for the considered problem.

It follows from condition (7.5) that the admissible control, on which the lower bound of the functional on the set of admissible controls is reached, must be such that the following two equalities are simultaneously satisfied

$$\int_0^1 x^2 dt = 0, \quad \int_0^1 u^2 dt = 1. \tag{7.7}$$

The function x here is a solution to the Cauchy problem (7.1) and is defined by the formula

$$x(t) = \int_0^1 u(\tau) d\tau.$$

In the case of control integrability, this function turns out to be continuous. Then it follows from the first formula (7.7) that the function x is identically equal to 0. Turning to equation (7.1), we conclude that the corresponding control must also be identically equal to 0. However, in this case, it can no longer satisfy the second equality (7.7). Thus, equalities (7.7) cannot hold simultaneously. However, the simultaneous fulfillment of these relations is the only way to achieve the lower bound of the functional. As a result, we conclude that the lower bound of the functional is not achievable[11], which means that there is no optimal control for Example 7.1.[12].

The result obtained, it would seem, is surprising. By setting a concrete control, we obtain a value of the optimality criterion. Another control corresponds to another value of the minimized functional. Thus, one of the selected controls will be better than the other in terms of the chosen criterion. It is natural to assume that, due to the limited set of admissible controls, one of the controls will be better than all the others[13].

The absence of an optimal control for the considered example means that no matter which control is chosen, there will always be an admissible control on which the value of the functional will be even less. In particular, for any admissible control v there is a number k such that the value of the functional I on the control u_k defined by formula (7.6) is less than on the control v[14].

The question arises why some optimization problems are solvable, while others are not? In order to answer this question, we have to establish how it is possible to find out in advance (that is, before obtaining the optimality conditions) that the optimal control problem has a solution.

7.1.3 Existence of optimal control

Chapter 1 considered the problem of solvability of the function minimization problem; see Theorem 1.2 and 1.3. Try to extend it to the general extremum problem.

Problem 7.1 *It is required to minimize a functional I on a set U.*

Determine sufficient conditions of solvability of this problem using the method of Chapter 1. In the course of the reasoning, we will impose restrictions on the functional I and the set U necessary to go through the entire proof path, and we will give the formulation of the corresponding theorem after that.

Theorem 1.2 (Weierstrass theorem) analyzes the minimization problem for the function $f = f(x)$ on a numerical set U. The first step there was the proof of the lower boundedness of the value set $f(U)$, which is true if the function is continuous, and the given set is bounded. We determine the boundedness of the numerical set $I(U)$ for Problem 7.1 if the functional I is lower bounded on the given set[15]. Then there exists a minimizing sequence, i.e., a sequence $\{u_k\}$ of the set U such that $I(uk) \to \inf I(U)$. However, we do not know the properties of the sequence $\{u_k\}$.

Weierstrass theorem uses the boundedness of the set U. Under this supposition, the minimizing sequence $\{x_k\}$ for the function f is bounded, because its elements belong to the bounded set. Therefore, this sequence has a convergent subsequence $\{x_s\}$ by the Bolzano−Weierstrass theorem, so there exists a number x such that $x_s \to x$. Some difficulties arise if we try to use this idea for Problem 7.1.

At first, it is necessary to specify what is the boundedness of the arbitrary set U. Suppose this includes to a ***control space*** V that is a ***normed vector space***, i.e., a vector space such that for any its element u one determines a non-negative number $\|u\|$ called the ***norm*** with some properties[16]. A subset U of the normed vector space is called ***bounded***[17] if there exists a number $c > 0$ such that $\|u\| \leq c$ for all $u \in U$.

We used the boundedness of the set of the function minimization problem for determining the boundedness of the minimizing sequence and using of the Bolzano−Weierstrass theorem for determination of its convergent subsequence. Unfortunately, this theorem has very limited application area[18]. However, there exists a very important class of spaces for which we can use an extension of the Bolzano−Weierstrass theorem called the Banach−Alaoglu theorem. Suppose V is ***Hilbert space***, i.e., complete vector space with dot product[19]. The norm is a root of the scalar square here, i.e., the following equality holds $\|u\|^2 = (u, u)$. By the ***Banach−Alaoglu theorem***, for any bounded sequence $\{u_k\}$ of the Hilbert space[20] one can extract a ***weakly convergent*** subsequence, i.e., a subsequence $\{u_s\}$ such that there is convergence in the sense of the dot product $(u_s, v) \to (u, v)$ for all elements v of V.

Thus, assuming that the set of admissible controls U is a bounded subset of a Hilbert space V, we conclude that the minimizing sequence $\{u_k\}$ for Problem 7.1 is bounded, and, by virtue of the Banach–Alaoglu theorem, we can extract from it a subsequence $\{u_s\}$ such that $u_s \to u$ weakly in V. At this stage of the proof of the existence theorem for a solution to Problem 7.1, a certain element u appeared, which is the weak limit of the indicated subsequence. It certainly belongs to the space V, but generally speaking, it may not belong to the set U, which means that it may not be the subject of our consideration[21].

At the next step of the proof, we used the assumption that the set on which the function is minimized is closed. The ***closedness*** of a set implies that any convergent sequence of elements of this set necessarily includes an element of this sequence. The concept of closed set also makes sense for Hilbert spaces[22]. We can require the set U to be closed under the conditions of Problem 7.1, but this will not be enough to justify the inclusion of the above element u in this set, since we are not dealing with strong convergence (i.e., convergence in the sense of the norm), but with weak convergence[23]. There are significantly more weakly convergent sequences than strongly convergent ones, as a result of which it is much more difficult to ensure that all weak limits belong to this set. However, the desired goal can be achieved if, in addition to the closedness of the set U, we also require its convexity. A convex closed set is ***weakly closed***[24], which means it contains all the limits of its weakly convergent sequences. This implies the inclusion $u \in U$.

We have established that the subsequence of the minimizing sequence for Problem 7.1 has a weak limit that belongs to the set of admissible controls. However, we do not yet know whether it is really optimal, i.e., it is on it that the infimum of the minimized functional is attained.

We return again to the Weierstrass theorem, more precisely, to its final stage. We had the convergence of a subsequence of the minimizing sequence $x_s \to x$, and the point x belonged to the set U. Assuming the continuity of the minimized function f, we deduced from this the convergence of the sequence of values of the function $f(x_s) \to f(x)$. However, since the value of the function on the minimizing sequence, and hence on any of its subsequences, converges to the lower bound of this function on the given set, then $f(x_s) \to \inf f(U)$. As a result, we conclude that $f(x) = \inf f(U)$, i.e., at the point x, the infimum of the given function is attained on the considered set.

For Problem 7.1, we can also assume that the functional I is continuous on the set U. However, we have established only a weak convergence $u_s \to u$ in the space V. Since this convergence is much weaker than convergence in the sense of the norm, there are many more weakly convergent sequences than strongly converging. As a result, it is much more difficult to establish the convergence of the values of the functional. In particular, the continuity of the functional alone is not enough for this, and it is required to impose an additional restriction on it. This is the property of convexity. In particular, a convex continuous functional turns out to be ***weakly lower semicontinuous***[25]. This means that from the weak convergence of the sequence of space elements it follows that the value of the functional at the limit of this sequence does not exceed the lower limit[26] of the sequence of values of the functional. In our case, this corresponds to the inequality

$$I(u) \leq \inf \lim I(u_s),$$

on the right side of which is the lower limit of the considered sequence.

According to the obtained inequality, the limit of any subsequence of the sequence $\{I(u_s)\}$ is not less than the value of $I(u)$. However, as we already know, the entire sequence[27] $\{I(u_s)\}$ and hence any of its subsequences, has the value $\inf f(U)$ as its limit. Therefore, the last inequality takes the form

$$I(u) \leq \inf f(U).$$

However, the element u belongs to the set U. The value of the functional on it cannot be less than the lower bound of the functional on this set. Then the last condition can only be satisfied in the form of equality. We have proved the existence of such an element u of the set U, on which the infimum of the functional on the given set is attained. Thus, this element turns out to be a solution to Problem 7.1, which means that this problem has a solution.

The results obtained can be formulated as a theorem.

Theorem 7.1 *The problem of minimizing a convex continuous lower bounded functional on a convex closed bounded subset of a Hilbert space has a solution.*

Naturally, this theorem gives only sufficient conditions for the solvability of the optimal control problem, i.e., the existence of an optimal control is also possible if its conditions are violated[28]. The uniqueness theorem for the solution of the problem was given earlier; see Theorem 5.1. If the conditions of both these theorems are realized, then the existence of a unique solution to the problem will be guaranteed. To do this, it suffices, under the conditions of Theorem 7.1, to require the strict convexity of the functional. Let us use this theorem to analyze the optimization problems considered earlier.

7.1.4 Application of the existence theorem

Let us return to the consideration of Example 3.3, the analysis of which has long been completed, and the corresponding optimal control has been found. Consider the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

It is required to find such a function $u = u(t)$ from the set

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\},$$

which minimizes on this set the functional

$$I = \frac{1}{2}\int_0^1 (u^2 + x^2)dt.$$

In order to use Theorem 7.1 to prove the solvability of this problem, one should first of all take into account that the dependence of the functional on the control here (as in any other optimal control problem) is determined not only directly, but also through the state of the system. We have already taken this circumstance into account when we applied the uniqueness theorem for optimal control to the same problem; see Chapter 5.

So far, we have not been interested in what functional properties the control should satisfy. The existence theorem for an optimal control confronts us with the need to concretize the set U by specifying the functional class to which any admissible control must belong. Let us pay attention to the fact that the optimality criterion includes the integral of the squared control. In this connection, as the control space V, we will consider the ***space*** $L_2(0,1)$ that is square-integrable in the sense of Lebesgue on a given time interval, i.e., those for which the following inequality holds

$$\int_0^1 |u(t)|^2 dt < \infty.$$

This is the Hilbert space with the dot product

$$(u,v) = \int_0^1 u(t)v(t)dt$$

and the norm determined by the equality[29]

$$\|u\|^2 = \int_0^1 |u(t)|^2 dt.$$

To prove the solvability of the problem under consideration using Theorem 7.1, it is required to establish that the set of admissible controls and the functional to be minimized satisfy all the necessary properties. In Chapter 5, in the process of proving the uniqueness of the optimal control, the convexity of the set of admissible controls was already established. Let us show that this set is closed. It suffices to show that any convergent sequence of elements of the set of admissible controls has as its limit an element of the same set.

Consider an arbitrary sequence $\{u_k\}$ of elements of the set U such that $u_k \to u$ in $L_2(0,1)$, i.e., the norm of the difference between the elements of the sequence and its limit tends to zero. As is known, if a sequence of elements of the space $L_2(0,1)$ converges in the sense of the norm of this space, then one can extract from it a subsequence that converges almost everywhere[30]. Therefore, from $\{u_k\}$ one can extract a subsequence $\{u_s\}$ such that we have the numerical sequence $u_s(t) \to u(t)$ for almost values of t from the interval (0,1).

Since the function u_s belongs to the set U, we have the inequality

$$|u_s(t)| \leq 1, \; t \in (0,1).$$

Passing in it to the limit, taking into account the established convergence, we obtain

$$|u(t)| \leq 1, \; t \in (0,1).$$

Then the inclusion $u \in U$ is true, which means that the considered set is closed[31].

Thus, the set of admissible controls has all the necessary properties. We now turn to the functional to be minimized. Obviously, it takes exclusively non-negative values, which means that it is bounded from below. The convexity of the functional was established earlier when proving the uniqueness of the solution to the problem[32]; see Chapter 5. Thus, it remains only to establish its continuity.

Assume that $u_k \to u$ converges in $L_2(0,1)$, i.e., the norm of the difference between the elements of the sequence and its limit tends to zero. Using the equation of state, we establish the formula

$$\big|x_k(t) - x(t)\big| = \Big|\int_0^t \big[u_k(\tau) - u(\tau)\big]d\tau\Big| \le \int_0^1 \big|u_k(\tau) - u(\tau)\big|d\tau,$$

where x_k and x are the system states for the controls u_k and u. For any elements a and b in Hilbert spaces, the ***Schwarz inequality*** holds

$$\big|(a,b)\big| \le \|a\|\,\|b\|.$$

We choose as a a function identically equal to 1, and as b the function $|u_k - u|$, from the preceding inequality, we obtain

$$\big|x_k(t) - x(t)\big| \le \Big(\int_0^1 1^2 dt\Big)^{1/2}\Big[\int_0^1 \big|u_k(\tau) - u(\tau)\big|^2 dt\Big]^{1/2} = \|u_k - u\|.$$

This implies that for any t convergence $x_k(t) \to x(t)$, and hence, $x_k \to x$ in $L_2(0,1)$ and even in the class of continuous functions.

Using the Schwartz inequality, we get

$$\big|I(u_k) - I(u)\big| \le \frac{1}{2}\Big[\int_0^1 (u_k - u)(u_k + u)d\tau + \int_0^1 (x_k - x)(x_k + x)d\tau\Big] \le$$

$$\frac{1}{2}\Big[\|u_k - u\|\,\|u_k - u\| + \|x_k - x\|\,\|x_k - x\|\Big].$$

Any admissible control v at an arbitrary point does not exceed 1, whence follows the estimate

$$\|v\|^2 = \int_0^1 \big|v(\tau)\big|^2 dt \le 1.$$

The corresponding system state y satisfies the inequality

$$\big|y(t)\big| = \Big|\int_0^t v(\tau)d\tau\Big| \le \int_0^t |v(\tau)|d\tau \le t \le 1,$$

whence we obtain the condition $\|y\| \le 1$. Thus, arbitrary admissible controls and their corresponding state functions do not exceed unity in norm.

It is known that the norm of the sum of two elements does not exceed the sum of norms[33]. Then we obtain

$$\|u_k + u\| \le \|u_k\| + \|u\| \le 2, \quad \|x_k + x\| \le \|x_k\| + \|x\| \le 2.$$

As a result, we get

$$\big|I(u_k) - I(u)\big| \le \frac{1}{2}\Big[\|u_k - u\|\,\|u_k - u\| + \|x_k - x\|\,\|x_k - x\|\Big] \le \|u_k - u\| + \|x_k - x\|.$$

Hence, the convergence $I(u_k) \to I(u)$ takes place. Thus, all the conditions of Theorem 7.1 for Example 3.3 are satisfied, so the corresponding optimal control problem has a solution.

Similarly, the solvability of the optimization problem for Example 6.2 can be proved, which differs from Example 3.3 only in the absence of a control-dependent term in the optimality criterion. It is clear that the exclusion of this term under the integral does not affect the proof of the boundedness from below, convexity and continuity of the functional.

In Example 3.1, the equation of state is the same, the set of admissible controls is characterized by a segment, and the functional is quadratic with respect to the control and linear with respect to the state of the system. The convexity of the functional was proved in Chapter 5, the boundedness is derived from the boundedness of the set of admissible controls, and the continuity of the functional is proved in the same way as for Example 3.3.

Let us now turn to Example 6.1, which differs from the one considered above only in the form of the optimality criterion, which is defined as follows

$$I = \int_0^1 u(t)x(t)dt.$$

To prove the existence of an optimal control using Theorem 7.1, it is required to establish its lower boundedness, convexity, and continuity. We recall that the convexity of the functional was established in Chapter 6. It was also shown there that it can be converted to the form

$$I(u) = \frac{1}{2}\Big[\int_0^1 u(t)dt\Big]^2.$$

Thus, the functional takes only non-negative values, i.e., bounded below by zero. Thus, it remains only to establish its continuity.

Let the convergence $u_k \to u$ again take place in $L_2(0,1)$. Since the state equation in this case is the same as in the previous one, we obtain the convergence $x_k \to x$ in $L_2(0,1)$. Let us estimate the value

$$\big|I(u_k) - I(u)\big| \leq \Big|\int_0^1 (u_k x_k - ux)dt\Big| \leq \Big|\int_0^1 (u_k - u)xdt\Big| + \Big|\int_0^1 u(x_k - x)dt\Big|.$$

Using the Schwartz lemma, we establish the inequality

$$\big|I(u_k) - I(u)\big| \leq \|u_k - u\|\,\|x\| + \|u\|\,\|x_k - x\| \leq \|u_k - u\| + \|x_k - x\|$$

because of the previously established estimates of the control norms and the state function. As a result, we conclude that $I(u_k) \to I(u)$. Thus, according to Theorem 7.1, the optimal control problem for Example 6.1 does indeed have a solution[34]. Let us now try to apply the existence theorem for Example 7.1. For it, the state equation and the set of admissible controls are defined in the same way as for the examples considered above, but the optimality criterion has the form

$$I = \int_0^1 (x^2 - u^2)dt.$$

Its lower boundedness was established earlier, and the continuity is proved in the same way[35] as for Example 3.3. Since we know for sure that in this example the optimal control does not exist, we come to the conclusion that the given optimality criterion is not convex[36].

The question arises whether the absence of convexity of the minimized functional always corresponds to an unsolvable optimal control problem? Recall that Example 5.1 differs from Example 3.3 considered above only in the type of extremum. For the latter, the functional is strictly convex. Therefore, the problem of maximizing a strictly convex functional considered in Example 5.1 turns out to be equivalent to minimizing a strictly concave, and therefore certainly not convex, functional. Thus, the conditions of Theorem 7.1 are violated for it. However, the corresponding optimization problem has a solution. The same can be said about Examples 3.2 and 6.2. Thus, the optimal control problem can also have a solution in the absence of convexity of the minimized functional[37].

RESULTS

Here is a list of questions in the field of the existence of optimal control, the main conclusions on this topic, as well as the problems that arise in this case, partially solved in Appendix, partially taken out in the Notes.

Questions

It is required to answer questions concerning the example considered in the lecture and the existence of a solution to optimal control problems.

1. Why is formula (7.4) obtained from the maximum condition (7.3), although the stationary condition for the function H has a solution?

2. Why does the iterative process for solving the system of optimality conditions in Example 7.1 not converge for any initial iteration?

3. Why for different Examples 5.2 and 7.1 the same system (7.1), (7.2), (7.4) is obtained, derived from the optimality conditions?

4. Why are there qualitatively different conclusions based on the divergence of iterative processes for Examples 5.1, 5.2, and 7.1?

5. Why is it impossible for Example 7.1 to have a singular control?

6. On what basis can we conclude that the system of optimality conditions for Example 7.1 has no solution?

7. Based on what can we conclude that the optimal control problem considered in Example 7.1 has no solution?

8. Whence follows the existence of the lower bound of the functional for Example 7.1?

9. Why is the sequence of controls defined by equalities (7.6) minimizing?

10. Why is the sequence of controls defined by equalities (7.6) not convergent in the natural sense?

11. What properties does the sequence of states corresponding to the considered minimizing sequence of controls have?

12. Why cannot equalities (7.7) and (7.8) hold simultaneously?

13. The derivation of the optimality condition in the form of the maximum principle in Chapter 3 began with the assumption of the existence of optimal control. What is the point of the above reasoning if in reality optimal control for the example considered does not exist?

14. What happens if for Example 7.1 we try to apply the unique optimal control theorem from Chapter 5?

15. What happens if, for Example 7.1, we try to apply the theorem on the sufficiency of optimality conditions from Chapter 5?

16. Can Theorem 7.1 be considered as a generalization of the Weierstrass theorem from Chapter 1?

17. Why does Theorem 7.1 require boundedness of the set of admissible controls?

18. Why does Theorem 7.1 require the convexity of the set of admissible controls?

19. Why does Theorem 7.1 require the closedness of the set of admissible controls?

20. Why does Theorem 7.1 require the lower boundedness of the minimized functional?

21. Why does Theorem 7.1 require convexity of the functional to be minimized?

22. Why does Theorem 7.1 require the continuity of the functional to be minimized?

23. Is it possible to have an optimal control if the conditions of Theorem 7.1 are violated?

24. Why is $L_2(0,1)$ chosen as the control space to prove the existence of a solution to the problem from Example 3.3?

25. Why is $L_2(0,1))$ chosen as the control space to prove the existence of a solution to the problem from Example 6.1?

26. Why is Theorem 7.1 not applicable to the analysis of Example 7.1?

Conclusions

Based on the analysis carried out, the following conclusions can be drawn about the existence of a solution to the optimal control problems.

- A situation is possible when the iterative process for solving the optimality conditions diverges for any initial approximation, which can be explained by the absence of a solution to the optimality conditions.
- In the process of solving of optimal control problems, a situation is possible when the iterative process diverges for any initial approximation, which can be explained by the absence of solution of the optimality conditions.
- The absence of a solution to the optimality conditions can be caused by the unsolvability of the optimization problem.
- The absence of a solution to the optimal control problem lies in the fact that there is no admissible control on which the lower bound of the functional is reached.
- The absence of a solution to the optimal control problem lies in the fact that for any admissible control it is possible to find another admissible control with a smaller value of the minimized functional.
- The optimal control problem for Example 7.1 has no solution.
- The optimality conditions for Example 7.1 are necessary and sufficient.
- The presence or absence of optimal control, in principle, can be established in advance, before the start of the direct solution of the optimization problem.
- When justifying the solvability of an optimal control problem, it is required to indicate the space to which the control must belong.
- When proving the existence of an optimal control, the properties of lower boundedness, convexity and continuity of the functional, as well as convexity, closure and boundedness of the set of admissible controls were used.
- All the conditions of Theorem 7.1 are satisfied for Example 3.3, which allows us to establish the solvability of the corresponding optimal control problem.
- All the conditions of Theorem 7.1 are satisfied for Example 6.1, which allows us to establish the solvability of the corresponding optimal control problem.
- The reason for the unsolvability of the problem from Example 7.1 is the non-convexity of the minimized functional.

- The optimal control problem can also be solvable in the absence of convexity of the functional, which is observed for Example 5.1.

Problems

Based on the results obtained above, we get the following problems related to the solvability of optimal control problems.

1. **Generalization**. In the lecture, one statement was given that guarantees the existence of an optimal control, which was used to prove the solvability of the problems considered earlier. In Appendix, a result will be given that allows one to establish the existence of an optimal control in the case of an unbounded set of admissible controls, in particular, for Example 3.4. Additional results in this direction are given in the Notes[38].

2. **Unsolvability and insufficiency**. When studying Example 7.1, it was noted that there is no solution to both the original optimal control problem and the system of optimality conditions, i.e., the optimality conditions turned out to be necessary and sufficient. However, when studying the problem of minimizing the function $f(x) = x^3$ in Chapter 1, we encountered a situation where the stationary condition has a solution $x = 0$, although the problem itself has no solution. Thus, we simultaneously encounter the unsolvability of the problem and the insufficiency of the extremum condition. It would be interesting to give an example of an optimal control problem with similar properties. Such an example is given in Appendix.

3. **Unsolvability of the problem with a convex functional**. In Example 7.1, the absence of optimal control is due to the non-convexity of the minimized functional. We would like to consider the unsolvable problem of minimizing a convex functional. Such an example is given in Appendix.

4. **Analysis of unsolvable optimal control problems**. At first glance, unsolvable optimal control problems do not make sense at all. However, the lower bound of the functional exists even in the absence of an admissible control on which it is attained. This suggests that the absence of optimal control cannot serve as a basis for refusing to analyze the problem at hand. Some ideas in this direction are presented in Appendix.

5. **Features of non-linear boundary value problems**. In Chapter 5, we considered an optimal control problem with an infinite set of solutions to a system of optimality conditions. It was shown that this system is equivalent to a boundary value problem for some non-linear second-order differential equation. As a result, the corresponding boundary value problem has completely non-trivial properties. It would be interesting to try in a similar way to reduce the unsolvable system of optimality conditions for Example 7.1 to some boundary value problem and establish its properties. These properties are addressed in Appendix.

7.2 APPENDIX

Below, we present some additional results related to Example 7.1 and the problem of solvability of optimal control problems. In particular, Section 7.2.1 gives an optimal control existence theorem for problems in an unbounded domain, which is a natural generalization of a similar result from Chapter 1. Section 7.2.2 gives an example of an unsolvable optimal control problem for which the maximum principle nevertheless has a solution, and in Section 7.2.3 there is no optimal control, although the problem of minimizing a convex functional is considered. Some considerations regarding the analysis of unsolvable problems are given in Subsection 7.2.4. Finally, in Subsection 7.2.5, the system of optimality conditions for Example 7.1 is reduced to a boundary value problem for a second-order differential equation with very non-trivial properties.

7.2.1 Existence of an optimal control when the set of admissible controls is unbounded

Theorem 7.1 gives sufficient conditions for the solvability of the optimal control problem for a bounded set of admissible controls. However, the optimal control problem considered in Example 3.4 has a solution, although the set of admissible controls is unbounded there. Note that in Chapter 1, a statement was made about the existence of a minimum of a function in an unbounded area. This result was obtained in the case when an additional requirement of coercivity is imposed on the function; see Theorem 1.3. To generalize this theorem to optimal control problems, it is necessary to define the concept of functional coercivity.

Definition 7.2 *A functional I defined on a normed space is called **coercive** if, for $\|v_k\| \to \infty$ we have $I(v_k) \to +\infty$.*

The following assertion is true[39].

Theorem 7.2 *The problem of minimizing a convex continuous coercive lower bounded functional on a convex closed subset of a Hilbert space has a solution.*

Proof. In this case, all the conditions of Theorem 7.1 are satisfied, except for the boundedness of the set of admissible controls, which was used earlier only to justify the boundedness of the minimizing sequence. Therefore, to prove the theorem, it suffices to show that the condition for the coercivity of the functional guarantees the boundedness of the minimizing sequence. Let there be a sequence $\{u_k\}$ of elements of the set U such that $I(u_k) \to \inf I(U)$. Assume that this sequence is not bounded. Then the condition $\|uk\| \to \infty$ is true, whence, due to the coercivity of the functional, it follows that $I(u_k) \to +\infty$, which contradicts the convergence of the sequence $\{I(u_k)\}$. As a result, we conclude that the sequence $\{u_k\}$ is bounded, which allows us to complete the proof in the same way as in Theorem 7.1. □

As an application of Theorem 7.2, consider Example 3.4, which consists in minimizing the functional

$$I = \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is the solution of the Cauchy problem

$$x'(t) = u(t), \; t \in (0,1); \; x(0) = 0.$$

The only difference from Example 3.3 considered earlier is the absence of restrictions imposed on the control. This does not allow using Theorem 7.1 to prove the existence of an optimal control. However, in view of the obvious coercivity of the functional defined by the squares of the control and state norms, the solvability of this problem follows from Theorem 7.2[40]. Its unique solution is a function identically equal to zero, on which the functional takes on a zero value[41].

7.2.2 Unsolvability of a problem with insufficient optimality conditions

In Example 7.1, we considered an optimal control problem that has no solution. In this case, the optimality conditions also turned out to be unsolvable. This situation, of course, should be recognized as highly undesirable. However, a much worse case is possible. In practice, it is often necessary to solve problems in the absence of their complete mathematical analysis. Failure in the practical implementation of the system of optimality conditions may lead to the idea that the problem lies in the problem statement itself, and, therefore, it is necessary either to correct the problem statement itself or to fundamentally change the research method (see Subsection 7.2.4). However, in Chapter 1, when solving the problem of minimizing the function $f(x) = x^3$, we encountered a much more unpleasant situation, when the extremum condition has a solution, although the problem itself turns out to be unsolvable. In such a case, in the absence of information about the possible unsolvability of the problem, one can come to the erroneous conclusion that the found solution of the optimality conditions is optimal. Let us consider one such example.

Example 7.2 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 \Big[x(t)^2 + (1-t^2)u(t)\Big]dt,$$

in the absence of control constraints, where x is the solution of the Cauchy problem

$$x'(t) = u(t), \; t \in (0,1); \; x(0) = 0.$$

Determine the function

$$H(u,x,p) = pu - \frac{1}{2}x^2 - \frac{1}{2}(1-t^2)u.$$

The function p here is a solution of the adjoint system

$$p'(t) = x(t), \; t \in (0,1); \; p(1) = 0.$$

Optimal control u is determined from the minimum condition for the function H

$$\Big[p - \frac{1}{2}(1-t^2)\Big]u - \frac{1}{2}x^2 = \min_v \Big\{\Big[p - \frac{1}{2}(1-t^2)\Big]v - \frac{1}{2}x^2\Big\}.$$

Obviously, the function H is linear, and therefore not bounded. Taking into account the absence of restrictions on control, we conclude that the last relation can have a solution exclusively in the form of singular control[42]. It is realized in the case when the coefficient at the term in the function H, which includes the control, vanishes. As a result, we obtain the equality

$$p = \frac{1}{2}(1 - t^2).$$

Substituting this value into the adjoint equation, we conclude that the existence of a singular control is possible only when the equality $x(t) = -t$ is fulfilled. This state is implemented for the control $u(t) = -1$, which is singular.

Thus, the maximum principle in Example 7.2 is satisfied by the unique control that is singular. Chapter 6 noted that the optimal singular control in the optimality criterion maximization problem satisfies the Kelley condition

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} \le 0.$$

Find the value

$$\frac{\partial H}{\partial u} = p - \frac{1}{2}(1 - t^2).$$

Taking into account the adjoint equation, we determine the derivative

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' + t = x + t.$$

We differentiate this equality taking into account the state equation.

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = x' + 1 = u + 1.$$

As a result, we find the value

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 1.$$

Thus, the Kelley condition is not satisfied, which means that the existing singular control is not optimal. However, there are no other controls that satisfy the maximum principle. The non-optimality of the only solution of the necessary optimality condition is possible only in the case when the solution of the problem does not exist, and the optimality condition is not sufficient[43].

7.2.3 Minimization of a functional on a non-convex set

In the considered examples, the absence of a solution to the problem of optimal control was due to the non-convexity of the minimized functional. Let us show that the problem may turn out to be unsolvable even in the case of convexity of the optimality criterion[44].

Example 7.3 *It is required to find a function $u = u(t)$ that minimizes on the set*

$$U = \{u|\ 1 \leq |u(t)| \leq 2;\ t \in (0,1)\}$$

the functional

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where the state function is a solution of the problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{7.8}$$

This problem differs from the one considered in Example 3.3 only by the set of admissible controls. Define the function

$$H = pu - (u^2 + x^2)/2.$$

In this case, the adjoint system has the form

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0. \tag{7.9}$$

Obviously, the only stationary point of the function H is zero. Although it maximizes this function, it does not belong to the set of admissible controls. Therefore, the maximum of this function can be achieved only on the boundary of this set. We find the values of H at all four boundary points.

$$H(1) = p - (1 + x^2)/2,\ H(-1) = -p - (1 + x^2)/2,$$

$$H(2) = p - (4 + x^2)/2,\ H(-2) = -p - (4 + x^2)/2.$$

Choosing the largest of these values, we determine the solution of the maximum condition

$$u(t) = \begin{cases} -2, & \text{if } p(t) < -3/2, \\ -1, & \text{if } -3/2 < p(t) < 0, \\ 1, & \text{if } 0 < p(t) < 3/2, \\ 2, & \text{if } p(t) > 3/2. \end{cases} \tag{7.10}$$

To study the system (7.8) – (7.10), let us estimate the set of possible values of the function p. The definition of the set U implies the inequality $-2 \leq u(t) \leq 2$. As a result of its integration, we obtain an estimate for the solution of problem (7.8) $-2t \leq x(t) \leq 2t$. Integrating the resulting relation from an arbitrary value of t to 1, we establish an estimate for the solution of problem (7.9) $-t^2 \leq p(t) \leq t^2$. Thus, the inequality $|p(t)| \leq 1$ is true, which means that the control values –2 and 2 are not allowed. As a result, formula (7.10) takes the form

$$u(t) = \begin{cases} -1, & \text{if } p(t) < 0, \\ 1, & \text{if } p(t) > 0. \end{cases} \tag{7.11}$$

Let us try to find a solution to the system (7.8), (7.9), (7.11). According to formula (7.11), the control can take only two values. For $u = 1$, the solution to problem (7.8) is determined by the formula $x(t) = t$. The corresponding solution to the adjoint system (7.9) has the form $p(t) = (t^2–1)/2$. This function takes exclusively negative values, which contradicts formula (7.11). If $u = –1$, then we get $x(t) = –t$, which means $p(t) = (1–t^2)/2$. Thus, the solution of the adjoint system in this case turns out to be positive, which again does not agree[45] with equality (7.11).

Suppose that the control takes the value 1 on the interval $(0, \xi)$, after which it changes sign (possibly repeatedly). Then on this interval the solution of problem (7.8) has the form $x(t) = t$. The result is the adjoint equation $p'(t) = t$. Since the derivative is positive, the function p increases on this interval. However, according to the assumption made at the point $t = \xi$, it changes sign, and therefore vanishes. Therefore, before that it must have been negative, which again contradicts equality (7.11). The impossibility of the existence of a control, which was initially equal to –1, and at some point, becomes equal to 1, is proved similarly. Thus, the system of optimality conditions has no solution at all.

Let us now try to establish the properties of the optimal control problem under consideration directly, without resorting to optimality conditions. Based on the form of the set of admissible controls, we establish the inequality $|u(t)| \geq 1$. Obviously, $|x(t)| \geq 0$. Then the integrand in the definition of the functional to be minimized turns out to be at least one. Thus, the condition $I(u) \geq 1/2$ is valid for all admissible controls.

Consider the sequence

$$u_k(t) = \begin{cases} 1, & \text{if } 2j/2k \leq t < (2j+1)/2k, \\ -1, & \text{if } (2j+1)/2k \leq t < (2j+2)/2k, \end{cases}$$

considered earlier in the study of Example 7.1. It was also shown there that for the corresponding sequence of solutions of the equations of state, the inequality

$$0 \leq x_k(t) \leq 1/2k, \ t \in (0,1), \ k = 1, 2, \dots .$$

Thus, $x_k(t) \to 0$ for all t. As a result, we conclude that $I(u_k) \to 1/2$, which implies the equality $\inf I(U) = 1/2$.

Let us now assume that for some admissible control the equality $I(u) = 1/2$ is true. This is possible only if the following two equalities are fulfilled simultaneously[46]

$$\int_0^1 x^2 dt = 0, \quad \int_0^1 u^2 dt = 1.$$

According to the first of them, the function x vanishes. However, this not only contradicts the second of the above equalities, but also corresponds to zero control, which is not admissible.

Based on the results obtained, we conclude that the optimal control problem under consideration has no solution, and the optimality conditions are necessary and sufficient[47]. We also note that, in this case, a bounded from below, continuous, and (which is especially important) convex functional is minimized. In this case, the set of admissible controls is closed and bounded. However, it is not convex[48], which explains the absence of optimal control[49].

7.2.4 Extension methods

We have some optimal control problem, which, as it turned out, has no solution. What can be done in such a situation? The most natural, it seemed, would be to correct the formulation of the problem in such a way that a solution would certainly exist in the new formulation. However, is it really a meaningless problem of optimal control that has no solution?

Let us go back to Example 7.1. As we have established, here there is no admissible control on which the lower bound of the functional is attained. However, the lower bound itself exists. This means that there is such an admissible control, the value of the functional on which is arbitrarily close to this lower bound. Indeed, the value of the functional on the control u_k defined by formula (7.6) for a sufficiently large number k differs from this lower bound by an arbitrarily small amount. Therefore, maybe it makes sense to consider this control as an approximate solution of the optimal control problem under consideration? In a real situation, we almost never can find an exact solution to the problem, as a result of which we have to be content with its approximate solution.

It should also be borne in mind that the problem, which is initially unsolvable, can turn into solvable if we expand the class of objects from which the solution of the problem is chosen. For example, the equation $2 + x = 1$ has no solution on the set of natural numbers, it becomes solvable on the set of integers. The equation $3x = 2$ has no solution on the set of integers, but it turns out to be solvable when passing to rational numbers. The equations of mathematical physics under certain conditions do not have a classical solution, which is understood as a sufficiently smooth function. However, it can have a ***generalized solution***, which is an element of a much wider Sobolev space[50]. Such reasoning makes sense not only for equations of one nature or another, but also for problems of extremum theory. In particular, the problem of minimizing the function $f(x) = x$ on the set of positive numbers has no solution, but it turns out to be solvable on a wider set of non-negative numbers. Methods for analyzing extremal problems based on these considerations are called ***extension methods***.

It is not our intention to describe specific extension methods[51]. We confine ourselves to presenting only the main idea of these methods. Let there be a problem of minimization of some functional I on the set U, which has no solution. First of all, let us consider a wider set V such that any of its elements can be approximated arbitrarily accurately by elements of the original set[52]. More precisely, there is an operator $A : U \to V$ such that for any element v of the set V there exists an element u from the set U such that the value Au will be arbitrarily close to v. Next, we consider some continuous functional J such that $J(Au) = I(u)$ for all $u \in U$, the problem of minimizing the functional J on the set V has a solution, besides the following equality holds

$$\min J(V) = \inf I(U).$$

Now we can solve the minimization problem for the functional J on the set V (extended problem), which has a solution v. By construction, there is such an element $u \in U$ that the value of Au will be arbitrarily close to v. Then, due to the continuity of the functional J, the number $J(Au)$ equal to $I(u)$ is arbitrarily close to $J(v)$ equal to $\min J(V)$, and hence also $\inf I(U)$. This defines a control u from the set U, the value of the functional I on which turns out to be arbitrarily close to its lower bound on this set. It can be understood as an approximate solution of the problem under conditions when the exact solution of the problem does not exist[53].

For the specified simplest example, as U we have a set of positive numbers, and V is a wider set of non-negative numbers. Obviously, any non-negative number is arbitrarily accurately approximated by a positive number. As an operator $A : U \to V$, a transformation is chosen that matches each positive number with its own, but understood as a non-negative number[54]. The functional I corresponds to the function $I(u) = u$, defined on the set U, and to the functional J, the function $J(v) = v$, defined on the set V, i.e., extension of the function I to the set U. The solution to the problem of minimizing the functional J on the set V is the number $v = 0$, which does not belong to the set U. However, there exists an arbitrarily close number u from this set, and the value of $I(u)$ turns out to be arbitrarily close to the unattainable lower bound of the functional I on the set U. Thus, in the absence of a minimal positive number, one can find an arbitrarily small positive number, which is understood as an approximate solution of the original problem.

7.2.5 Some features of non-linear boundary value problems

Chapter 5 considered an optimal control problem for which the corresponding system of optimality conditions had an infinite number of solutions. This system, in accordance with the method of elimination of unknowns, was reduced to a boundary value problem for a non-linear second-order differential equation with very unusual properties. Let us try to carry out similar transformations for the optimality conditions corresponding to Example 7.1.

We have the state system

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0,$$

the adjoint system

$$p'(t) = x(t),\ t \in (0,1);\ p(1) = 0$$

and the formula

$$u(t) = F(p(t)) = \begin{cases} 1, & \text{if } p(t) > 0, \\ -1, & \text{if } p(t) < 0. \end{cases}$$

Differentiating the adjoint equation using the equation of state and the formula for finding the control, we determine the boundary value problem

$$p''(t) = F(p(t)),\ t \in (0,1);\ p(1) = 0,\ p'(0) = 0. \tag{7.12}$$

It differs from the boundary value problem (5.16) only by the sign in the corresponding function F. However, being equivalent to the system of optimality conditions for Example 7.1, it has qualitatively different properties. If the boundary value problem (5.16) has an infinite set of solutions, then problem (7.12) has no solution at all due to the absence of solutions for the optimality conditions for the considered example[55].

As in Chapter 5, we will also consider the non-linear heat equation

$$\frac{\partial y(\tau,\xi)}{\partial \tau} = \frac{\partial^2 y(\tau,\xi)}{\partial \xi^2} - F(y(\tau,\xi)),\ \tau > 0,\ 0 < \xi < 1 \tag{7.13}$$

with the boundary conditions

$$\frac{\partial y(\tau,0)}{\partial \xi} = 0,\ y(\tau,1) = 0,\ \tau > 0 \tag{7.14}$$

and an initial condition. Obviously, the equilibrium position $z = z(\xi)$ for this system satisfies the equalities

$$z''(t) = F(z(t)),\ t \in (0,1);\ z(1) = 0,\ z'(0) = 0$$

coinciding with the boundary value problem (7.12). Since the latter has no solution, we conclude that the system characterized by equation (7.13) with the boundary condition (7.14) generally does not have an equilibrium position[56]. However, by changing the sign before the function F, we get the system considered in Chapter 5 with an infinite set of equilibria.

Additional conclusions

Based on the results obtained in Appendix, we can draw the following additional conclusions about the problem of solvability of optimal control problems.

- The existence of an optimal control is guaranteed if, in the corresponding theorem, the requirement that the set of admissible controls be bounded is replaced by the coercivity of the functional to be minimized.
- The optimal control problem from Example 3.4 has a solution in accordance with Theorem 7.2, although the conditions of Theorem 7.1 are violated for it.
- For Example 7.2, both the solvability of the optimal control problem and the sufficiency of optimality conditions are simultaneously absent.
- For Example 7.2, there is a singular control in the absence of a solution to the considered problem under.
- In Example 7.3, there is no optimal control, although a convex functional is minimized.
- The reason for the unsolvability of the problem from Example 7.3 is the absence of convexity of the set of admissible controls.

- In the absence of an optimal control, one can pose the problem of finding such an admissible control, the value of the functional on which is arbitrarily close to its lower bound. An analysis of an unsolvable optimal control problem can be carried out using extension methods.

- The system of optimality conditions for Example 7.1 is reduced to a boundary value problem for a non-linear second-order ordinary differential equation that has no solution.

- There is a boundary value problem for a non-linear equation of the heat conduction type, for which the equilibrium position does not exist.

Notes

1. Examples of the absence of a solution to problems in the theory of extremum, related to the calculus of variations, are given in [5], [67], [44], [60], [142], [200]; for optimal control problems described by ordinary differential equations; see [5], [95], [60], [170], [200]; for optimal control problems related to partial differential equations; see [36], [116], [121], [124], [136], [137].

2. See Examples 1.9, 1.10, and 1.11.

3. This system of optimality conditions exactly coincides with the system (6.12), (6.13), and (6.15) obtained in the analysis of Example 6.2 and differs only in the sign in the control formula from system (5.1), (5.2), and (5.4). We also note that the resulting system is invariant under sign change, just like the original control problem. We have already encountered this property many times before.

4. Here we can use the invariance of the system (7.1), (7.2), and (7.4) with respect to sign change.

5. As already noted, system (7.1), (7.2), and (7.4) differs from system (5.1), (5.2), and (5.4) considered in Chapter 5 only by the sign in the formula for determining the control. In both cases, the control can only take the values 1 or -1, with the possibility of switching from one of these values to another. If in that case, any choice of the initial approximation through iteration led to the same value, which corresponded to the convergence of the iterative process, then in this case, through iteration, a value with the opposite sign is obtained, which means the oscillation of the algorithm.

6. It is present with a constant multiplier, as a result of which it is impossible to realize the degeneration of the maximum principle.

7. In connection with the above reasoning, certain doubts may arise. We proved the absence of optimal control by establishing the unsolvability of the system of necessary optimality conditions. However, the derivation of the maximum principle began with the assumption of the existence of an optimal control. One might get the impression that we are dealing with a vicious circle. First, the maximum principle is derived under the assumption that an optimal control exists, and then a conclusion is made that there is no optimal control due to the unsolvability of the maximum principle, which itself is established under the assumption that a solution exists. Nevertheless, the conclusions made are quite justified. We can state that if an optimal control exists, then it would satisfy the maximum principle. However, since the maximum principle does not really have a solution, our initial assumption about the solvability of the optimization problem turned out to be false.

8. The sufficiency of the optimality conditions for the example under consideration can be established using Theorem 5.2 given in Chapter 5. Its conditions imply that the function f, which is the right side of the equation of state, can be represented as $f(t,u,x) = f_1(t,u)+cx$, the integrand g in the definition of the optimality criterion is equal to $g(t,u,x) = g_1(t,u)+g_2(t,x)$, where c is some constant and f_1, g_1 and g_2 are some functions, and the second derivatives with respect to x of g_2 and the function h, which determines the final state of the system in optimality criteria is not negative. For Example 7.1, we have $f = u$, $g = (x^2 - u^2)/2$, $h = 0$. Thus, all the conditions of Theorem 5.2 are satisfied, which means that the optimality conditions are indeed necessary and sufficient.

9. An iterative process for solving a system of necessary optimality conditions for an unsolvable optimal control problem can converge, but, of course, not to a solution of the problem, which is absent. This situation is possible in the case when the optimality conditions are necessary, but not sufficient, having some solution. In particular, the function $f(x) = x^3$ considered in Chapter 1 has no minimum points, but the corresponding stationary condition has a solution $x = 0$. If we try here to use approximate methods for solving the stationary condition, then we can quite get this point as a limit, although it does not minimize the function, but is its inflection point. The most unpleasant thing here is that in practice we often have to solve the problem formally, without knowing in advance whether the system of optimality conditions has a solution or not. Under these conditions, in the case of an observed divergence of the algorithm, we may not understand whether this circumstance is caused by the fundamental absence of a solution or by the negative properties of the algorithm itself, for example, an unsuccessful choice of the initial approximation. However, it is clear that the absence of a solution can undoubtedly be one of the reasons for the divergence of the algorithm.

10. Of interest are the properties of the sequence $\{x_k\}$. Obviously, we are dealing with a sequence of continuous functions converging pointwise to a function identically equal to zero. At the same time, each function x_k has no derivative at all those points at which the corresponding control u_k has a discontinuity. Thus, the set of points of non-differentiability grows indefinitely, and in an arbitrarily small neighborhood of any point from the domain of the considered functions, for sufficiently large k, there are points at which x_k has no derivative. Nevertheless, the limit function x, which is identically equal to zero, turns out to be infinitely differentiable. Consider now the sequence $\{l(x_k)\}$ of lengths of curves x_k on the unit interval. Obviously, the equality $l(x_k) = \sqrt{2}$ is true for any k. However, $l(x) = 1$. Thus, the length of the limit curve does not coincide with the limit of the sequence of lengths of curves.

11. Examples of unsolvable optimal control problems for systems with a fixed finite state are given in Chapter 11. Chapters 15 and 17 consider optimal control problems that have no solution, which differ from Example 7.1, respectively, only in the presence of an additional isoperimetric condition and the absence of the initial condition.

12. The absence of an optimal control is also surprising because we seem to have a sequence of admissible controls $\{u_k\}$ defined by relations (7.6). We know that it is minimizing, i.e., as the number k increases, the corresponding values of the functional converge to its lower bound on the set of admissible controls. What prevents us from considering the limit of this sequence as an optimal control? Indeed, if this sequence converges to a function u, then, due to the continuity of the functional, which will be proved in Subsection 7.2.1, it would be possible to establish the convergence of $I(u_k) \to I(u)$. However, as is already known, the sequence of values $\{I(u_k)\}$ converges to the lower bound of the functional. Then, on the basis of the established results, one could come to the conclusion that the function u turns out to be a solution to the same problem, the absence of which we have established, even in two different ways. All possible contradictions will be eliminated if the sequence $\{u_k\}$ defined above does not converge at all, i.e., has no limit in any reasonable class of functions. Indeed, as the number k increases, the number of discontinuities of the corresponding control also increases indefinitely. Moreover, for a sufficiently large number k, on any arbitrarily small interval, the function u_k

has an arbitrarily large number of discontinuity points. In this connection, we can conclude that the sequence $\{u_k\}$ has no limit. Naturally, one can speak about the convergence of a sequence, especially a functional one, only when it is indicated in what sense this convergence should be understood. The indicated phenomenon in control theory is called the sliding mode; see [58], [191]. Note, however, that the elements of this sequence with sufficiently large numbers can be chosen as weak approximate solutions of the problem under consideration, since they are admissible, and the values of the minimized functional corresponding to them are close enough to the lower bound of the latter; see Chapter 8.

13. In all the examples considered in Chapter 1, the absence of minimum points of a function was observed in problems for an unconditional extremum, i.e., in the case of unboundedness of the set on which the minimization problem is solved.

14. Here there is a complete analogy with the problem of determining the minimum number of an open interval $(0, 1)$. No matter how small a positive number we take, there is always a positive number with an even smaller value. Even though the interval (0,1) is limited, we will never be able to specify the smallest of its elements for the simple reason that the minimum positive number simply does not exist.

15. The ***functional boundedness*** means that there is a number that does not exceed the value of the functional on any admissible control. The important thing here is that we are working with a functional, and not with a general operator. As a result, the image of any set under its action turns out to be a numerical set, which means that the limitedness of this image has a natural meaning.

16. It is assumed that the number $\|u\|$ equals zero if and only if u is the zero element of the space, and for any elements u and v of the space and numbers λ the following relations hold $\|u + v\| \leq \|u\| + \|v\|$, $\|\lambda u\| = |\lambda|\|u\|$; see [94], [100], [106], [158].

17. The concept of boundedness can be defined not only for normed spaces, but also for a much wider class of metric spaces. However, metric properties are not enough to obtain further results, in particular, due to the absence of an analogue of the Bolzano–Weierstrass theorem for this class of spaces.

18. The Bolzano−Weierstrass theorem is valid only for finite-dimensional spaces, to which function spaces do not apply. In particular, the sequence $\{u_k\}$ defined earlier in accordance with equalities (7.6) is bounded in the space of integrable functions, but does not have a subsequence that converges in the natural sense (that is, in the sense of the corresponding norm).

19. For Hilbert spaces see [94], [100], [106], [158]. The ***dot product*** over the field of real numbers on the vector space V is such a transformation that assigns to any two elements u and v from V the number (u, v), and the equalities $(u, v) = (v, u)$, $(u, v + w) = (v, u) + (u, w)$ and $(u, \lambda v) = \lambda(u, v)$ for any number λ, and moreover, the scalar square (u, u) is non-negative and vanishes exclusively for zero space elements. Sometimes, in addition, they require that the space be infinite-dimensional. The completeness of a space means that any sequence in it whose elements converge indefinitely (i.e., the fundamental sequence) is convergent. Among the Hilbert spaces is the space L_2 of measurable functions that are Lebesgue integrable with a square in some domain, which we work with in a number of specific examples.

20. In fact, the Banach−Alaoglu theorem is valid not only for Hilbert spaces, but also for a wider class of reflexive Banach spaces, i.e., complete normed vector spaces (***Banach spaces***) whose second conjugation coincides with the original space, where weak convergence can also be defined; see [94], [100], [106], [158]. Moreover, there is an even wider class of spaces adjoint to Banach spaces, where one can also use this assertion with the weak convergence replaced by

an even weaker convergence, called *-weak. For reflexive spaces, weak and *-weak convergence coincide, and in the finite-dimensional case, weak convergence coincides with ***strong convergence*** when the norm of the difference between an arbitrary element of the sequence and its limit tends to zero. Thus, in the finite-dimensional case, the Banach–Alaoglu theorem actually reduces to the Bolzano–Weierstrass theorem. Reflexive Banach spaces include the space L_p of measurable functions Lebesgue integrable with degree $p > 1$ in some domain. The space L_∞ of essentially bounded functions is not reflexive, but is adjoint to the Banach space L_1 of Lebesgue integrable functions, as a result of which this statement remains valid. But L_1 itself is not adjoint to any Banach space, as a result of which it is not possible to use the Banach–Alaoglu theorem for it. On spaces L_p; see [94], [100], [106], [158].

21. The sequence of numbers $\{1/k\}$ from the interval (0,1) converges in the space of real numbers. However, the corresponding limit does not belong to this interval, since the latter is not a closed set.

22. The closedness of a set makes sense on a much wider class of ***topological spaces***, which include normed vector spaces and not only them; see [101].

23. For example, the sequence $\{\sin k\pi t\}$, which we will encounter in the next chapter, converges weakly, but not strongly.

24. This assertion follows from the ***Mazur's lemma***; see [60].

25. For the semicontinuity of a functional; see, for example, [44], [60], [95].

26. The ***lower limit*** of a sequence is the lower bound of the limits of all its subsequences.

27. Moreover, the entire sequence $\{I(u_k)\}$ converges to the lower bound of the functional.

28. For the existence of an optimal control, much more general objects can be used as the control space. We restrict ourselves to the use of Hilbert spaces solely because they are sufficient to study the considered examples. In Chapter 14, the solvability of one optimal control problem with a non-convex set of admissible controls will be established.

29. It was noted earlier that the norm can be equal to zero only on the zero element of the corresponding vector space. However, the integral is equal to zero not only for a function that is identically equal to zero, but also for functions that differ from zero only at individual points, and even for a ***Dirichlet function*** equal to zero at all irrational points and one for all rational points. Here, it should be borne in mind that we are dealing with ***measurable functions*** in the sense of Lebesgue. Moreover, two functions that differ only on a set of zero measure (the set of rational numbers just has zero measure) are understood in measure theory as one and the same object. This means that, in fact, the elements of the space $L_2(0,1)$ are not separate functions, but classes of functions that coincide almost everywhere (differ only on a set of zero measure), i.e., we are dealing with a quotient space. From this, in particular, it follows that in the definition of the set U, the validity of the inequality $|u(t)| \leq 1$ can be realized not for all, but for almost all values of $t \in (0,1)$.

30. This means that the set of points in a given interval for which the specified condition is not satisfied has zero Lebesgue measure; see [94], [100], [106], [158]. The indicated literature also provides a proof of the assertion under consideration.

31. In fact, we have established the validity of the inequality $|u(t)| \leq 1$ not for all, but for almost all values of $t \in (0,1)$. However, as already noted, the elements of the space $L_2(0,1)$ are defined up to a set of zero measure, as a result of which the result obtained is sufficient to justify the inclusion $u \in U$.

32. Even the strict convexity of the functional was proved earlier, although strict convexity is not needed in Theorem 7.1. In fact, this functional has even stronger properties; see next Chapter.

33. This property is included in the definition of the norm.

34. In Chapter 9 Theorem 7.1 will be used to prove the existence of an optimal control for a system with a fixed final state, and in Chapter 13 for a system with an isoperimetric condition.

35. The difference in signs in front of the control square does not affect the continuity properties.

36. It was already noted earlier that a change of sign in a convex function turns it into a concave function. Remember that for Example 7.1, a minimizing sequence $\{u_k\}$ was defined, which is a sequence of piecewise constant controls with an increasing number of discontinuity points; see Figure 7.1. Using the Banach–Alaoglu theorem, we establish that it is possible to extract from it a weakly convergent subsequence in the space $L_2(0,1)$. However, we have previously argued that this sequence does not converge. Thus, we are dealing with a sequence that converges weakly and does not converge strongly in the sense of the space $L_2(0,1)$. It can be seen that the weak limit of this sequence is a function identically equal to zero. It is certainly an element of the set of admissible controls (recall that this set is convex and closed, and therefore weakly closed, i.e., contains its own weak limits). The question arises why the weak limit of the minimizing sequence is not an optimal control? We know that the functional to be minimized is continuous. However, this property is not yet sufficient to derive the convergences of the sequence of functionals from the weak convergence of the sequence of controls. This requires weak continuity (or at least semicontinuity) of the functional. The absence of convexity of the functional did not allow us to derive its weak lower semicontinuity from its strong continuity. However, in this case we can assert that the functional to be minimized is not weakly lower semicontinuous at all, since otherwise, having weak convergence of controls, we could prove the optimality of the corresponding weak limit. Thus, the minimized functional for Example 7.1 is continuous, but not weakly lower semicontinuous.

37. We also note the solvability of the problem from Example 3.4 with an unbounded set of admissible controls.

38. General existence theorems for solutions to extremal problems are given in [5], [60], [95], [116], [194]. The solvability of the calculus of variations is considered in [5], [95], [200], [208]; for optimal control problems described by ordinary differential equations see [5], [44], [200]; for optimal control problems described by partial differential equations see [73], [116], [118], [171]. We also note the study of the solvability of optimal control problems for integral equations [20], for stochastic systems [87], for minimax problems [19], and for multi-extremal problems [209].

39. Chapter 14 will consider a solvable minimization problem for a lower bounded convex continuous coercive functional on a non-convex subset of a Hilbert space; see Example 14.2. However, the set under consideration is weakly closed. This property is quite sufficient to prove the existence of an optimal control.

40. In Chapter 14, Theorem 7.2 will be used to prove the existence of a solution to the optimal control problem in the presence of an isoperimetric condition and a fixed final state of the system. This statement will also be used in Chapter 16 to prove the existence of a solution to optimal control problems for systems with a free initial state, when the optimality criterion is minimized on a set of ”control-state” pairs.

41. The same control, of course, was also optimal in Example 3.3.

42. It is easy to see that for the example under consideration the conditions of Theorem 6.1 are realized, under which the existence of a singular control is guaranteed.

43. Let us establish the unsolvability of the optimal control problem under consideration directly. Let us define the sequence of controls $\{u_k\}$ in accordance with the equality $u_k(t) = k$. The corresponding state of the system is $x_k(t) = kt$. Let us find the values of the functional

$$I(u_k) = \frac{1}{2}\int_0^1 \Big[x^2t^2 + (1-t^2)k\Big]dt \geq \frac{1}{2}\int_0^1 \big(k^2t^2 + k\big)dt = \frac{k^3}{6} + \frac{k}{2}.$$

It follows that $I(u_k) \to \infty$ at $k \to \infty$. Thus, the value of the functional to be maximized can be arbitrarily large. Therefore, the problem under consideration does not really have a solution. Note the fundamental difference between Examples 7.1 and 7.2. In the first case, in the absence of minimum points of the functional on the set of admissible controls, there exists a corresponding infimum of the functional. In the second case, the functional to be maximized is not bounded from above, which means that its upper bound does not exist on the set of admissible controls.

44. In fact, this functional is even strongly convex; see the following chapter.

45. Actually, the last result is a consequence of the invariance of the problem with respect to the sign change.

46. We are dealing with the already familiar equalities (7.7).

47. Note that this sequence $\{u_k\}$ can be used to find a weak approximate solution to the problem.

48. In fact, this set is not even connected, being made up of two disjoint parts; see [101].

49. One should not think that the absence of convexity of the set characterizing the constraints on the system necessarily entails the unsolvability of the extremal problem. Particularly, in the Weierstrass theorem on the existence of an extremum of a function, the convexity of its domain of definition is not required. In this case, the problem of minimizing the considered functional on any non-convex set that includes the zero point (solution of the problem from Example 3.3) will be solvable. A solvable optimal control problem with a non-convex set of admissible controls is considered in Chapter 14.

50. On generalized solutions to problems of mathematical physics see [63], [116], [130], [172].

51. On extension methods for extremum theory problems see [95], [44], [60], [108], [166], [167], [200], [208].

52. This means that the set U is dense in V, that is, its closure coincides with V; see [94], [100], [106], [158]. Thus, the set of rational numbers is dense in the set of real numbers, i.e., any real number can be arbitrarily accurately approximated by rational numbers. Similarly, the set of polynomials is dense in the set of continuous functions on an interval according to the Weierstrass theorem, i.e., any continuous function can be approximated as accurately as polynomials.

53. A rigorous definition of an approximate solution to an optimal control problem will be given in a subsequent chapter.

54. Such an operator is called an embedding.

55. The boundary value problem for a second-order differential equation obtained from the system of optimality conditions for Example 7.3 has a similar property.

56. The boundary value problem for the heat equation associated with the system of optimality conditions for Example 7.3 has a similar property.

CHAPTER 8

Ill-posed optimal control problems

As is known, various kinds of difficulties can arise in solving optimal control problems. In particular, there may be no solution to the problem at all or be non-unique, the necessary optimality conditions may not be sufficient or degenerate, and the corresponding iterative processes may diverge. However, other side effects are also possible. We will consider an optimal control problem that has a unique solution, the corresponding necessary optimality conditions are sufficient. Nevertheless, it is possible to construct such a sequence of admissible controls, the values of the functional on which tend to its minimum, while this sequence itself does not converge to the optimal control. A similar situation is typical for optimization problems that are ill-posed in the sense of Tikhonov. The lecture establishes sufficient conditions for the well-posedness of the problem. In addition, an example of an optimal control problem is given for which there is no continuous dependence of the solution on the parameter, which corresponds to the absence of the Hadamard well-posedness. Appendix defines various types of approximate solutions to optimal control problems, establishes conditions that guarantee the well-posedness in the sense of Hadamard, and also describes some methods for studying ill-posed problems.

8.1 LECTURE

We consider a fairly simple optimal control problem for which there is a unique solution, and the optimality conditions are necessary and sufficient. An example of a minimizing sequence that does not converge to an optimal control is given. This situation corresponds to the absence of the Tikhonov well-posedness of the problem. Sufficient conditions for the Tikhonov well-posedness of the problem are given. For another example, the absence of a continuous dependence of the optimal control on some parameter is established, which is due to the Hadamard ill-posedness.

DOI: 10.1201/9781003398585-8

8.1.1 Tikhonov well-posedness

To solve the optimal control problems, we used optimality conditions in the form of the maximum principle and other methods. We know that various difficulties arise in their practical application. However, the list of these difficulties is far from being exhausted.

Let us return to Example 6.2. The optimal control problem is to find such a function $u = u(t)$ from the set

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,1)\},$$

which minimizes on this set the functional

$$I(u) = \frac{1}{2}\int_0^1 x(t)^2 dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{8.1}$$

In Chapter 6, the unique solution to this problem $u = 0$ was found. It would seem that a complete analysis has been carried out for the considered example. However, this optimal control problem has another surprising property.

Consider the sequence of controls defined by the formula; see Figure 8.1

$$u_k(t) = \sin k\pi t,\ k = 1, 2, \ldots . \tag{8.2}$$

Obviously, these functions are infinitely differentiable and do not exceed unity in absolute value. Thus, we are dealing with a sequence of admissible controls.

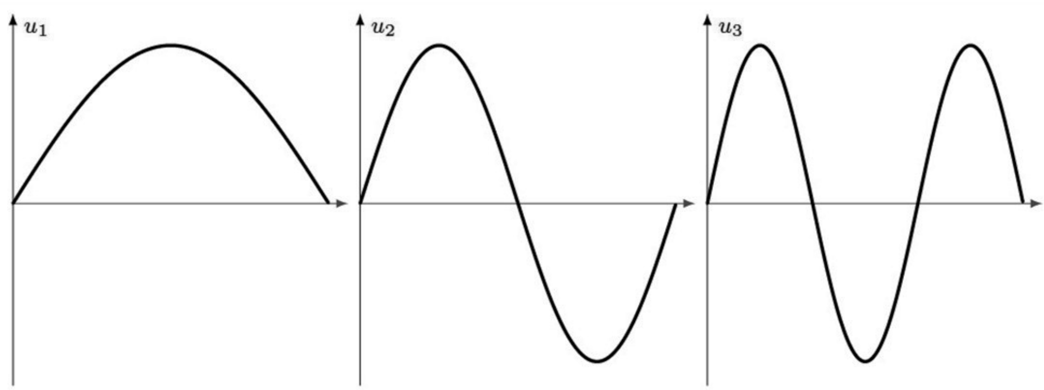

Figure 8.1 Minimizing sequence for Example 6.2.

The corresponding solutions to problem (8.1) are defined as follows; see Figure 8.2

$$x_k(t) = \int_0^t \sin k\pi\tau d\tau = \frac{1 - \cos k\pi t}{k\pi}. \tag{8.3}$$

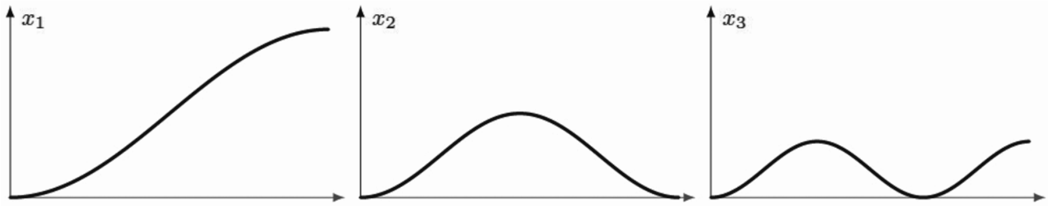

Figure 8.2 The sequence of states determined by the formula (8.3).

The following inequality holds

$$0 \leq x_k(t) \leq 2/k\pi, \; t \in (0,1), \; k = 1, 2, \ldots .$$

From this follows the relation

$$0 \leq I(u_k) = \int_0^1 x_k^2 dt \leq \frac{4}{(k\pi)^2}, \; k = 1, 2, \ldots .$$

Then the sequence of functionals corresponding to the controls $\{u_k\}$ converges to 0, i.e., to the minimum value of the optimality criterion on the set of admissible controls. Thus, the sequence of controls $\{u_k\}$ determined by formulas (8.2) is minimizing for this example.

A natural question arises: does the minimizing sequence converge to the optimal control u_0? The convergence of a sequence to its limit in a normed space means that the numerical sequence of the norms of the difference between the elements of the sequence and this limit tends to 0. In the space $L_2(0,1)$ the square of the norm of a function is equal to the integral of the square of this function. Let us estimate the value

$$\|u_k - u_0\|^2 = \int_0^1 \left|u_k(t) = u_0(t)\right|^2 dt = \int_0^1 \sin^2 k\pi t dt = \frac{1}{2}.$$

Thus, the optimal control u_0 is not the limit of the sequence $\{u_k\}$. Thus, the unique solution to the optimal control problem is not the limit of the minimizing sequence[1]

Obviously, this sequence does not converge at all. Indeed, if this sequence converged to some limit u^*, then, due to the continuity of the functional, the corresponding limit of the sequence of functionals would be equal to the value of the functional at the limit of this sequence, i.e., $I(u^*)$. However, this limit is equal to the minimum of the functional on the set of admissible controls, which means that the control u^* would be optimal. However, we already know that the only optimal control in this case is u_0, to which the minimizing sequence does not converge. Therefore, the sequence $\{u_k\}$ has no limit at all[2]. As a result, we conclude that the minimizing sequence for an optimal control problem with a continuous functional either converges to the optimal control or diverge altogether.

The divergence of the minimizing sequence, in principle, should not cause much surprise. We have already encountered a similar situation in the previous chapter when examining Example 7.1. In particular, the control sequence depicted in Figure 7.1 diverges, being minimizing for the corresponding optimal control problem[3]. However, optimal control did not exist there at all, i.e., the minimizing sequence simply had nowhere to converge. In this case, the optimal control exists, but the sequence of controls under consideration does not converge to it. This suggests that we are faced with a qualitatively different effect, consisting in the fact that the minimizing sequence does not always converge to the existing optimal control[4].

Definition 8.1 *The optimal control problem is called* ***Tikhonov well-posed***[5] *if any minimizing sequence for it converges to an optimal control. Otherwise, this is called* ***Tikhonov ill-posed****.*

Thus, the optimal control problem for Example 6.2 is Tikhonov ill-posed[6]. We are convinced that even if there is a unique optimal control and if the necessary optimality conditions are sufficient, we are not immune from certain surprises. In ill-posed problems, even having a control, the value of the minimized functional, on which is arbitrarily close to its lower bound, we cannot be sure that we will find the control with any predetermined degree of accuracy[7].

Obviously, an unsolvable optimization problem is not Tikhonov well-posed, since there is no optimal control for it at all. The question arises whether the situation is possible when the optimal control problem has not a unique solution, but is Tikhonov well-posed? Let us consider the situation when the optimal control problem has two different optimal controls u and v. Consider an arbitrary sequence of admissible controls $\{u_k\}$ and $\{v_k\}$ converging to u and v, respectively. Then, in the case of continuity of the minimized functional I, the conditions $I(u_k) \to \inf I(U)$ and $I(v_k) \to \inf I(U)$ are valid. Let us define a sequence $\{w_k\}$ whose elements take the values u_k and v_k alternately. Obviously, the sequence $\{I(w_k)\}$ converges to the value $\inf I(U)$, since all its subsequences have this property. Thus, the sequence $\{w_k\}$ turns out to be minimizing. At the same time, it does not converge because it has subsequences that converge to different limits. Hence, it follows that the considered problem is Tikhonov ill-posed.

Thus, we are convinced that the class of optimal control problems that are Tikhonov well-posed is narrower than the class of problems that have a unique solution. In particular, the optimal control problem for Example 6.2 is uniquely solvable, but not Tikhonov well-posed. In this regard, we can expect that in order to prove the well-posedness of the problem, stronger restrictions on the system will be required in comparison with those that we had in the study of the existence and uniqueness of the optimal control.

8.1.2 Justification of Tikhonov well-posedness

Our goal is to establish conditions that guarantee the Tikhonov well-posedness of optimal control problems[8]. Note that in the process of proving the existence, the convexity of the minimized functional was used; see Theorem 7.1 To ensure the uniqueness of the optimal control, a stronger condition was required that is the strict convexity; see Theorem 5.1. The Tikhonov well-posedness of the optimization problem is established under an even stronger constraint[9].

Definition 8.2 *A functional I on a convex subset U of a normed space is said to be* ***strongly convex*** *if for any elements $u, v \in U$ and a number $\alpha \in (0,1)$ the following inequality holds*

$$I[\alpha u + (1–\alpha)v] \le \alpha I(u) + (1–\alpha)I(v) – c\alpha(1–\alpha)\|u - v\|^2,$$

where c is a positive constant.

Obviously, any strongly convex functional is strictly convex, and even more so, convex.

The following assertion is true.

Theorem 8.1 *If the problem of minimizing a strongly convex functional on a convex subset of a Hilbert space has a solution, then it is Tikhonov well-posed*[10]

Proof. Let us assume that the function u is a solution to the problem of minimizing a strongly convex functional I on a convex set U. Consider an arbitrary minimizing sequence, i.e., such a sequence $\{u_k\}$ of elements of the set U that the convergence $I(u_k) \to \inf I(U)$ takes place. Then we have the inequality

$$I[\alpha u_k + (1–\alpha)u] \le \alpha I(u_k) + (1–\alpha)I(u) – c\,\alpha(1–\alpha)\|u_k – u\|^2.$$

From this follows the relation

$$I[\alpha u_k + (1–\alpha)u] - I(u) \le \alpha\big[I(u_k) - I(u)\big] – c\,\alpha(1–\alpha)\|u_k – u\|^2.$$

Due to the optimality of the control u, the value at the left side of the last inequality is not negative. As a result, after dividing by α, we have

$$c(1–\alpha)\|u_k – u\|^2 \le I(u_k) – I(u).$$

Taking into account the arbitrariness of the parameter α, we pass to the limit as $\alpha \to 0$. Thus, we establish the inequality

$$0 \le c\|u_k – u\|^2 \le I(u_k) – I(u).$$

Since $\{u_k\}$ is a minimizing sequence, the value on the right side tends to 0 here. Thus, the convergence $\|u_k – u\| \to 0$ takes place, which means that the sequence $\{u_k\}$ converges to the optimal control u. Hence, due to the arbitrariness of the minimizing sequence, it follows that the problem is Tikhonov well-posed. □

As an application, we turn first to Example 3.3. It required to choose such a control from the set[11]

$$U = \{u |\, |u(t)| \le 1,\ t \in (0,1)\},$$

which minimizes on this set the functional

$$I(u) = \frac{1}{2}\int_0^1 (u(t)^2 + x(t)^2)dt,$$

where the state of the system x is described by the equalities

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{8.4}$$

The existence of a solution to this problem was established in the previous chapter. Thus, to apply Theorem 8.1, it suffices to establish only the strong convexity of this functional.

Consider first the quadratic function $f(x) = x^2$. We have the equality

$$f[\alpha x + (1-\alpha)y] - \alpha f(x) - (1-\alpha)f(y) = [\alpha x + (1-\alpha)y]^2 =$$

$$= \alpha^2 x^2 + 2\alpha(1-\alpha)xy + (1-\alpha)^2 y^2 - \alpha x^2 - (1-\alpha)y^2 = -\alpha(1-\alpha)(x-y)^2$$

for all numbers x, y and $\alpha \in (0,1)$. Now we obtain

$$f[\alpha x + (1-\alpha)y] = \alpha f(x) + (1-\alpha)f(y) - \alpha(1-\alpha)(x-y)^2$$

As a result, we conclude that the function f is strongly convex, and $c = 1$.

Assuming in the previous formula $x = u(t)$ and $y = v(t)$ integrating over t, for all $u, v \in U$, and $\alpha \in (0,1)$ we get

$$\int_0^1 [\alpha u + (1-\alpha)v]^2 dt = \alpha \int_0^1 u^2 dt + (1-\alpha)\int_0^1 v^2 dt - \alpha(1-\alpha)\int_0^1 (u-v)^2 dt.$$

Hence, it follows that the functional J defined by the formula

$$J(u) = \int_0^1 u^2 dt,$$

is strongly convex such that the equality

$$J[\alpha u + (1-\alpha)v] = \alpha J(u) + (1-\alpha)J(v) - \alpha(1-\alpha)\|u-v\|^2.$$

In Chapter 5, the convexity of the functional was established

$$K(u) = \int_0^1 x[u]^2 dt,$$

where $x[u]$ is solution of the Cauchy problem defined above, corresponding to the control u. Considering that the functional to be minimized is equal to the half-sum of the functionals J and K, we establish the inequality[12]

$$I[\alpha u + (1-\alpha)v] \le \alpha I(u) + (1-\alpha)I(v) - \alpha(1-\alpha)\|u-v\|^2.$$

Thus, the minimized functional in Example 3.3 is strongly convex. Then, using Theorem 8.1, we conclude that the considered optimal control problem is Tikhonov well-posed.

Let us now turn to the consideration of Example 3.1, which differs from Example 3.3 only in the form of the functional to be minimized. In particular, here it has the form

$$I = \int_0^1 \Big(\frac{u^2}{2} - 3x\Big)dt.$$

It can be represented as the sum of the functional J defined above, divided in half, and the functional

$$M(u) = -3\int_0^1 x[u]dt,$$

where $x[u]$ has the same meaning as before. The functional J, as we already know, is strongly convex. Let us establish properties of the functional M.

Consider the equations of state (8.4) for some admissible controls u and v. We multiply the first of them by α, and the second by $1{-}\alpha$ and adding the results together, we have

$$\alpha x'[u] + (1{-}\alpha)x'[v] = \alpha u + (1{-}\alpha)v.$$

However, the control $\alpha u + (1{-}\alpha)v$ corresponds to the state $x[\alpha u + (1{-}\alpha)v]$. Taking into account that the initial state of the system is zero, we conclude

$$x[\alpha u + (1{-}\alpha)v] = \alpha x[u] + (1{-}\alpha)x[v].$$

Then after integration, we obtain the equality

$$M[\alpha u + (1{-}\alpha)v] = \alpha M[u] + (1{-}\alpha)M[v].$$

As a result, we conclude that the functional M is convex, and the convexity relation for it is realized in the form of the equality[13]

Thus, the functional I to be minimized for Example 3.1 is the sum of a strongly convex functional J and a convex functional M. Thus, the functional I turns out to be strongly convex, which means that the considered optimal control problem is Tikhonov well-posed.

Of the examples considered earlier, the Tikhonov well-posed problem was considered in Example 3.4, which differs from Example 3.3 only in the absence of restrictions on the control[14]. Problems from Examples 5.1, 5.3, 6.1, 6.6–6.8, 6.10, 6.12, and 6.14 turn out to be ill-posed, since the solution of these problems is not unique, as well as problems from Examples 6.9, 7.1, and 7.2, for which optimal control does not exist. The problem from Examples 4.2 and 4.3 is ill-posed, in view of the fact that even for the equation of state, there is no guaranteed unique solvability. Finally, we cannot use Theorem 8.1 to study the optimization problems from Examples 3.3 and 6.3 that have a unique solution due to the lack of convexity properties for the functional being minimized.

For Example 8.1, we can establish only strict but not strong convexity of the functional to be minimized. This is sufficient to prove the uniqueness of the optimal control; see Chapter 5. However, as we have already established, this problem is not Tikhonov well-posed. As a result, the corresponding functional here cannot be strongly convex.

8.1.3 Hadamard well-posedness

So far, we have considered the well-posedness of optimal control problems in the sense of Tikhonov. However, in the theory of equations of various nature, the concept of Hadamard well-posedness[15] is widely used, which implies the existence of a unique solution to the problem that continuously depends on its parameters. Let us try to investigate this problem for optimal control problems.

Example 8.1 *It is required to minimize the functional*

$$I_k(u) = \int_0^1 \left[(x - y_k)^2\right] dt$$

on the set

$$U = \{u \in L_2(0,1) \mid |u(t)| \leq 1,\ t \in (0,1)\},$$

where $y_k(t) = (k\pi)^{-1} \sin k\pi t$*, k is a numerical parameter*[16]*, and x is a solution of the Cauchy problem*

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{8.5}$$

The difference of this problem from all the previous ones is the presence of a certain parameter that can change. In fact, we have not one, but a family of optimization problems. We have not only to find the corresponding optimal control, but also to investigate its dependence on the parameter k. We have already done something similar in Chapter 2 when studying function minimization problems.

Let us explicitly find a solution to the problem posed. Obviously, the value of the minimized functional is not negative. Its equality to 0 is possible only under the condition $x(t) = y_k(t)$ for all t. Substituting this value into problem (8.5), we find the function

$$u_k(t) = y_k'(t) = \cos k\pi t.$$

Given that this control is admissible, we conclude that it is the unique optimal control[17] for Example 8.1.

Passing to the limit in the formula of the optimality criterion, we find

$$I_\infty = \lim_{k\to\infty} I_k(u) = \int_0^1 (x - y_\infty) dt = \int_0^1 x^2 dt,$$

where y_∞ is the limit of the sequence $\{y_k\}$, i.e., $y_\infty = 0$. Obviously, the problem of minimizing the functional I_∞ on the set U, which coincides with Example 6.2, has a unique solution $u_\infty = 0$. The question arises whether the sequence of solutions $\{u_k\}$ converges to the solution u_∞ of the limit problem?

Let us estimate the value

$$\|u_k - u_\infty\|^2 = \int\limits_0^1 |u_k(t) - u_\infty(t)|^2 dt = \frac{1}{2}.$$

Thus, the optimal control u_∞ for the limit problem is not the limit of the sequence $\{u_k\}$ of solutions to the original problem. Thus, for the considered problem, there is no continuous dependence of the solution on the parameter.

Definition 8.3 *An optimal control problem is called* ***Hadamard well-posed*** *if it has a unique solution that depends continuously on the parameters of the problem.*

Based on the analysis, we conclude that the optimal control problem for Example 8.1 is not Hadamard well-posed[18].

The meaning of the Hadamard ill-posedness of the problem it is as follows. When setting any applied problem of optimal control, there are always some characteristics that are determined experimentally, and therefore known with a certain degree of accuracy. For Hadamard well-posed problems, a small error in determining the system parameters does not cause a large error in optimal control. In the absence of well-posedness, the situation changes. Let us assume that the true value (that is, corresponding to the practical formulation of the problem) is the value y_∞ corresponding to the functional I_∞. However, during the measurement process, the value of y_k was obtained, which is quite close to y_∞ for large enough k. Thus, instead of the true functional I_∞, we actually minimize the functional I_k, which is sufficiently close to it. We, as if, have the right to expect that a small error in the experiment should cause a small error in the optimal control. However, instead of a true optimal control u_∞, we get a solution u_k of the problem under consideration that is very far from it. Thus, in the process of practical solution of a problem that is Hadamard ill-posed, strong distortions of the results are obtained[19].

RESULTS

Here is a list of questions in the field of well-posedness of optimal control problems, the main conclusions on this topic, as well as the problems that arise in this case.

Questions

It is required to answer questions concerning the well-posedness of optimal control problems according to Tikhonov and Hadamard.

1. What are the properties of the optimal control problem in Example 6.2?
2. What properties does the optimality criterion for Example 6.2 have?

3. What properties does the sequence of controls $\{u_k\}$ defined by formula (8.2) have?
4. What properties does the sequence of states $\{x_k\}$ defined by formula (8.3) have?
5. How can one experimentally detect the absence of Tikhonov well-posedness?
6. Why is the absence of Tikhonov well-posedness an undesirable phenomenon?
7. Can a well-posed Tikhonov problem have a non-unique solution?
8. Can a minimizing sequence in a Tikhonov ill-posed problem converge to an optimal control?
9. Why is the optimal control problem Tikhonov well-posed for Example 3.3, but ill-posed for Example 6.2?
10. Why is there an assumption about the existence of an optimal control under the conditions of Theorem 8.1, but there is no assumption about its uniqueness?
11. How do the concepts of convexity, strict convexity, and strong convexity relate to each other?
12. What do the optimal control problems from Examples 6.2 and 8.1 have in common, and how do they differ?
13. What is the difference between Tikhonov and Hadamard well-posedness?
14. What is the similarity between Hadamard well-posedness?
15. Why, when analyzing the Tikhonov and Hadamard well-posedness, it is necessary to take into account the functional properties of the control?
16. Why is the absence of Hadamard well-posedness an undesirable phenomenon?
17. What properties does the optimality criterion for Example 8.1 have?
18. Why does the optimal control problem in Example 8.1 have a unique solution?
19. What properties does the sequence of controls $\{u_k\}$ considered for Example 8.1 have?
20. What properties does the sequence of states $\{x_k\}$ considered for Example 8.1 have?

Conclusions

Based on the study of the well-posedness of optimal control problems, we can come to the following conclusions.

- Not every minimizing sequence converges to an optimal control, which is due to the absence of Tikhonov well-posedness of the problem.

- The optimal control problem from Example 6.2 is Tikhonov ill-posed.
- The minimizing sequence used in the analysis of Example 6.2 is weakly but not strongly convergent.
- For Tikhonov ill-posed optimal control problems, an admissible control on which the value of the minimized functional is sufficiently close to its minimum may not be close to the optimal control.
- The class of Tikhonov well-posed optimization problems is narrower than the class of uniquely solvable optimization problems.
- The Tikhonov well-posedness of the optimization problem can be established under the condition of strong convexity of the minimized functional.
- The optimality criterion for Example 6.2 is strictly convex, but not strongly convex.
- In optimal control problems, it is possible that the solution does not depend continuously on the parameters of the problem, which is due to the absence of the Hadamard well-posedness.
- The optimal control problem from Example 8.1 is Hadamard ill-posed.
- In the absence of the Hadamard well-posedness, a small error in determining the parameters of the system can cause significant distortion of the results.
- In the absence of the Hadamard well-posedness, various kinds of algorithmic errors can cause significant distortions of the results.

Problems

Based on the analysis of the well-posedness, the following additional problems arise.

1. **Meaning of the approximate solution**. The main conclusion in the analysis of Example 6.2 and the concept of Tikhonov well-posedness of the optimal control problem was that the meaning of the approximate solution is far from unambiguous. What should be understood by this approximate solution? Is it a control that is close enough to the optimal one or one for which the value of the minimized functional turns out to be sufficiently close to its lower bound? We also recall the penalty method considered in Chapters 2 and 3, which provided an approximate solution to problems for a conditional extremum. In contrast to the examples considered above, a slight violation of the given restrictions was also allowed there, which corresponds to a qualitatively different form of the approximate solution. Appendix gives various definitions of the approximate solution of optimal control problems.

2. **Justification of the Hadamard well-posedness**. Theorem 8.1 was presented in the Lecture, in which the restrictions are indicated that guarantee

the Tikhonov well-posedness of the problem. However, there is no similar statement in relation to Hadamard well-posedness. At the same time, in Chapter 2, properties were described that guarantee the Hadamard well-posedness of the problem of minimizing a function of one variable. In Appendix, this assertion is extended to optimal control problems.

3. **Relationship between Tikhonov and Hadamard well-posedness**. The two considered types of well-posedness of optimal control problems definitely differ. However, comparing Examples 6.2 and 8.1 of both types of ill-posed problems, one can note their undoubted closeness. In this regard, it can be assumed that these concepts are somehow related to each other. The connection between Tikhonov and Hadamard well-posedness is established in Appendix.

4. **Analysis of ill-posed problems**. If we have an ill-posed problem, then, in principle, we can simply try to correct the problem statement so that the new problem turns out to be well-posed. However, there are situations when the object of research is obviously not a well-posed problem, and it is necessary to solve it. It should also be borne in mind there are significantly more ill-posed problems than correct ones. In this regard, it seems extremely important to develop methods for solving ill-posed optimal control problems. Some results in this direction are given in Appendix.

8.2 APPENDIX

Below, we present additional results concerning the well-posedness of optimal control problems. In particular, the absence of Tikhonov well-posedness means that ensuring the closeness of the approximate solution to the exact one in the sense of controls and functionals can be implemented by different means. In this regard, it is of significant interest to clarify what is meant by an approximate solution of an optimal control problem. Subsection 8.2.1 is devoted to these issues. Next, we give a theorem that makes it possible to establish the Hadamard well-posedness of the problem. The final subsection deals with methods for the practical solution of ill-posed optimal control problems.

8.2.1 Types of approximate solution of the problem of finding an extremum

As has been repeatedly noted, the practical solving of optimal control problems, as a rule, is carried out approximately. In this case, it is desirable to clarify what exactly is meant by an approximate solution of the problem. This problem is of particular relevance when considering problems that are Tikhonov ill-posed. In Chapter 2, various types of approximate solutions to the function minimization problem were defined. Let us extend them to problems of functional minimization of a general form and problems of optimal control. Consider the first of them.

Problem 8.1 *It is required to minimize a functional I on a subset U of a normed space.*

By analogy with Definition 2.7, the following concepts are introduced.

Definition 8.4 *A control $u \in U$ is called a* ***strong approximate solution*** *of Problem* 8.1 *if it is close enough to the optimal control u_{opt}, i.e., the inequality $\|u - u_{opt}\| \le \varepsilon$ holds for a small enough positive number ε. The control $u \in U$ is called a* ***weak approximate solution*** *of the optimal control problem if the value of the functional I corresponding to it is close enough to its lower bound on the set U, i.e., the inequality $I(u) \le \inf I(U) + \delta$ holds*[20] *for a small enough positive number δ.*

In the case of continuity of the minimized functional, any strong approximate solution of the problem turns out to be its weak solution. If the optimal control problem is Tikhonov well-posed, then any weak approximate solution of the problem turns out to be its strong solution. However, in ill-posed problems, a weak approximate solution may not be strong. In particular, the function u_k defined by formula (8.2) for sufficiently large k is a weak but not strong approximate solution of the problem for Example 6.2.

For an unsolvable problem, the existence of an infimum of the functional on the set of admissible controls is possible; see Example 7.1. In this connection, the notion of a weak approximate solution makes sense for such a problem. In particular, the function u_k, defined by formula (7.6) for a large enough number k, turns out to be a weak approximate solution. It is the search for a weak approximate solution that the methods of extension of optimal control problems are oriented to; see Section 7. On the other hand, for the problem of maximizing the functional in Example 7.2, the upper bound of the functional does not exist. As a result, for such an example, a weak approximate solution does not make sense here, and the use of extension methods noted in the previous chapter is meaningless.

In Chapter 2, it was noted that in problems on the conditional extremum of a function, it makes sense to have such an approximate solution that satisfies the existing restrictions with a certain degree of accuracy. It can be naturally generalized to Problem 8.1.

Definition 8.5 *A control u is called a* ***strong conditional approximate solution*** *to Problem* 8.1 *if, for a small enough number $\delta > 0$, the inequality $\|u - u_{opt}\| \le \delta$ is true, where u_{opt} is the exact solution of this problem. The control u is called a* ***weak conditional approximate solution*** *of this problem if for small enough numbers $\varepsilon > 0$ and $\chi > 0$ the following inequalities hold $I(u) \le \inf I(U) + \chi$ and $\|u - v\| \le \chi$ for an element $v \in U$.*

A conditional approximate solution of the problem may not be an element of a given set, but it certainly turns out to be close enough to some element of a given set[21].

Let us now turn to optimal control problems.

Problem 8.2 *It is required to minimize the functional $I = I(u, x)$ on a subset U of some normed space, where x is the solution of the state equation*[22] *of $A(u, x) = 0$ defined by some operator A, corresponding to the control u.*

For Problem 8.2, we can introduce the concepts of strong and weak approximate solutions, as well as the corresponding conditional approximate solutions based on Definitions 8.1 and 8.2, assuming that the equation of state determines the implicit dependence $x = x[u]$. However, another type of approximate solution arises in connection with the use of the penalty method; see Chapter 3.

Definition 8.6 *A pair* (u, x) *is called a* ***strong approximate solution*** *of Problem* 8.2 *if, when under the inclusion* $u \in U$ *it is close enough to the optimal pair* (u_{opt}, x_{opt})*, i.e., the inequalities* $\|u - u_{opt}\| \leq \varepsilon_1$ *and* $\|x - x_{opt}\| \leq \varepsilon_2$ *hold for small enough positive numbers* ε_1 *and* ε_2*. A pair* (u, x) *is called a* ***weak approximate solution*** *of Problem* 8.2 *if under the inclusion* $u \in U$ *satisfy the inequalities* $I(u) \leq \inf I + \varepsilon$ *and* $\|A(u, x)\| \leq \chi$ *for sufficiently small positive numbers* ε *and* χ*, where* $\inf I$ *is the infimum of the mapping* I *on the set of all pairs* (v, y) *satisfying the equality* $A(v, y) = 0$ *and the inclusion* $v \leq U$*.*

According to Definition 8.6, the approximate solution (u, x) does not necessarily satisfy the equation of state $A(u, x) = 0$, but in some sense satisfies it approximately[23]. In particular, in Example 3.3, the relation $A(u, x) = 0$ was considered to correspond to an equation $x' = u$ with an initial condition. However, in the process of applying the penalty method, a pair $(u_\varepsilon, x_\varepsilon)$, was determined that satisfies an equation $x'_\varepsilon = u_\varepsilon + \varepsilon p_\varepsilon$ with a small parameter ε, where p_ε is the solution of the corresponding adjoint system. Similarly, in Example 4.2, the state equation was $x' = x^2 + u$. However, during the analysis, a pair $(u_\varepsilon, x_\varepsilon)$, was obtained, which satisfies the equality $x'_\varepsilon = x_\varepsilon^2 + u_\varepsilon + \varepsilon p_\varepsilon$.

8.2.2 Justification of Hadamard well-posedness

The concepts of the well-posedness of an optimization problem in the sense of Tikhonov and Hadamard differ significantly by definition. However, there is a lot in common between them. In particular, for each fixed value of the parameter k, the problem from Example 8.1, which is Hadamard ill-posed also turns out to be Tikhonov ill-posed. In both cases, the ill-posednes manifests itself as the absence of convergence of the corresponding sequence of controls. Taking into account the explicit connection between the concepts of well-posedness under consideration, we can expect that both the principle of proof of well-posedness and the methods for solving ill-posed problems of both types will be somewhat similar.

Suppose that there is some functional $I = I(\mu, u)$, where μ is a parameter that takes values from some set[24] M, and u is a control defined on the set U. The problem of minimizing the functional I is posed for a fixed value of the parameter $\mu \in M$ on the set U. The following statement[25], takes place, which is a natural generalization of Theorem 2.2 from Chapter 2.

Theorem 8.2 *Suppose that for any fixed value* $\mu \in M$ *the problem of minimizing the functional* $I = I(\mu, u)$ *on the set* U *is Tikhonov well-posed, and the mapping* $\mu \to I(\mu, u)$ *is continuous on* M *uniformly*[26] *in* $u \in U$*. Then the considered extremal problem is Hadamard well-posed.*

Proof. Consider an arbitrary sequence $\{\mu_k\}$ converging on M. Then there exists such an element $\mu \in M$ that convergence $\mu_k \to \mu$ takes place. Denote by J_k and J the functionals characterized by the equalities $J_k(u) = I(\mu_k, u)$ and $J(u) = I(\mu, u)$. As is well known, a Tikhonov well-posed problem has a unique solution. Denote by u_k and u the solutions of minimization problems on the set U of the functionals J_k and J, respectively. To prove the theorem, it suffices to establish that the sequence of solutions $\{u_k\}$ converges to the value u.

We have the equality

$$J(u_k)–J(u) = [J(u_k)–J_k(u_k)] + [J_k(u_k)–J_k(u)] + [J_k(u)–J(u)].$$

Considering that the quantities u_k and u are solutions of the corresponding extremal problems, we establish the inequalities

$$0 \le J(u_k)–J(u),\ J_k(u_k)–J_k(u) \le 0.$$

As a result, we get

$$0 \le J(u_k)–J(u) \ = \ [J(u_k)–J_k(u_k)] + [J_k(u_k)–J_k(u)] + [J_k(u)–J(u)] \le$$

$$2 \sup_{v \in u} |J_k(v) - J(v)|.$$

Since the mapping $\mu \to I(\mu, v)$ is continuous on M uniformly in $v \in U$, the value on the right side of the last inequality tends to zero. Thus, the convergence $J(u_k) \to J(u)$ takes place. This implies that the sequence $\{u_k\}$ is minimizing for the functional J.

By the hypothesis of the theorem, the problem of minimizing the functional J is Tikhonov well-posed. Then the minimizing sequence $\{u_k\}$ converges to the optimal value u. Therefore, the convergence of the parameters $\mu_k \to \mu$ implies the convergence of solutions of the corresponding extremal problems $u_k \to u$, which means that the considered problem is Hadamard well-posed. □

According to Theorem 8.2, proof of the Hadamard well-posedness of optimization problems is reduced to proving its correctness in the sense of Tikhonov for a fixed value of the parameter and justifying the uniform continuity of the functional. As an application, consider the following example.

Example 8.2 *It is required to minimize the functional*

$$I(\mu, u) \ = \int\limits_0^1 \big[(x - \mu)^2 + u^2\big] dt$$

for a fixed square-integrable function μ on the set

$$U \ = \ \big\{u \in L_2(0,1) \big| \ |u(t)| \le 1,\ t \in (0,1)\big\},$$

where x is a solution of the Cauchy problem

$$x'(t) \ = \ u(t),\ t \in (0,1);\ x(0) = 0.$$

The Tikhonov well-posedness of this problem for a fixed value of the parameter μ is proved using Theorem 8.1 in the same way as for a similar problem from Example 3.3. Let us establish that the functional under consideration is uniformly continuous.

For arbitrary values of the functions μ and ν, the following equality holds

$$\sup_{v\in U}\big|I(\mu,u)-I(\nu,u)\big| \le \Big|\int_0^1 \big[(x-\mu)^2-(x-\nu)^2\big]dt\Big| = \Big|\int_0^1 (2x-\mu-\nu)(\mu-\nu)dt\Big|.$$

This implies the inequality[27]

$$\sup_{v\in U}\big|I(\mu,u)-I(\nu,u)\big| \le \|2x-\mu-\nu\|\,\|\mu-\nu\| \le \big(2\|x\|+\|\mu\|+\|\nu\|\big)\|\mu-\nu\|. \quad (8.6)$$

From the state equation, we find the function

$$x(t) = \int_0^t u(\tau)d\tau.$$

Since any admissible control takes values from the interval $[-1,1]$, we conclude that the state of this system takes values from the interval $[-t,t]$, which means that $\|x\|^2 \le 1/3$. Thus, inequality (8.6) implies the uniform continuity of the considered functional. Now Theorem 8.2 implies Hadamard well-posedness of the considered optimal control problem[28].

The absence of Hadamard well-posedness according to for Example 8.1 is due to the Tikhonov ill-posedness of the corresponding problem for a fixed value of the parameter.

8.2.3 Regularization of optimal control problems

In the practical solution of ill-posed extremal problems, certain difficulties arise. In particular, algorithms that ensure the minimization of the functional may not guarantee finding the optimal control with the desired degree of accuracy. Various regularization methods are used to overcome the difficulties that arise[29].

Let us turn, in particular, to Example 8.1, in which we studied the optimal control problem that was Tikhonov ill-posed. The ***Tikhonov regularization method*** for this problem involves considering the functional

$$I_\varepsilon(u) = I(u)+\varepsilon\int_0^1 u^2dt,$$

where $\varepsilon > 0$ is a regularization parameter. Obviously, the problem of minimizing a regularized functional up to a constant factor coincides with the problem from Example 3.3, the well-posedness of which was established earlier. Thus, the solution of the regularized optimal control problem should not cause serious difficulties. Let the solution u_ε of the regularized problem be somehow found[30]. Then the solution of the original problem can be obtained as a result of passage to the limit as $\varepsilon \to 0$.

In practice, they usually proceed as follows. Initially, the value of ε is set relatively large. This is explained by the fact that the convergence of the iterative process for solving a regularized problem is the better and the less dependent on the choice of the initial approximation, the larger the regularization parameter. Therefore, despite the fact that the initial approximation for the iterative process is given arbitrarily, and therefore, generally speaking, far enough from the solution of the problem, the algorithm will converge relatively quickly. Then the parameter ε decreases, and as the initial approximation of the iterative process at the new step of the regularization method, the control obtained from the previous step of the iterative method is selected. The deterioration in the convergence of the algorithm due to the approximation to the original ill-posed problem is compensated to a certain extent by refining the initial approximation for the iterative process[31]. The specified procedure for splitting the regularization parameter is repeated many times until the desired degree of accuracy is achieved[32].

The described algorithm implies a nested iterative process, in which, at each step of the regularization method, the regularized optimal control problem is solved, for example, using the method of successive approximations described earlier. ***Iterative regularization*** often turns out to be more effective, in which at each step of the method of successive approximations, the parameter ε is gradually reduced[33]. An additional opportunity to improve the efficiency of the algorithm for solving the problem is related to the error in the numerical solution of the equation of state and the adjoint system[34].

In the previous subsection, the connection between the concepts of Tikhonov and Hadamard well-posedness was noted. In this regard, we can conclude that the use of regularization methods can also be effective for solving problems that are Hadamard ill-posed.

Let us note also that the maximum principle for Example 6.2 degenerates, i.e., the optimal control is singular. This circumstance significantly complicates the practical solution of the problem. At the same time, in a regularized problem, the maximum condition no longer degenerates[35], and its solution at a fixed step of the method of successive approximations is found explicitly[36]. Thus, regularization methods can also be used for practical finding of singular controls.

Additional conclusions

Based on the results presented in Appendix, we can draw some additional conclusions regarding the well-posedness of optimal control problems and methods for solving ill-posed problems.

- There are various types of approximate solutions to optimal control problems.
- A strong approximate solution of the problem is close enough to its exact solution.
- The value of the functional on a weak approximate solution is sufficiently close to its lower bound on the set of admissible controls.

- If the functional is continuous, a strong approximate solution is always a weak approximate solution.
- In Tikhonov well-posed problems, a weak approximate solution turns out to be a strong approximate solution.
- A weak approximate solution exists for unsolvable problems, provided that the lower bound of the functional exists.
- The conditional approximate solution allows minor violations of the state equation and given constraints.
- The Hadamard well-posedness of an optimization problem is related to its Tikhonov well-posedness.
- When justifying the Hadamard well-posedness of the problem, the uniform continuity of the functional with respect to the parameter is used.
- For the practical solution of ill-posed problems, one can use the Tikhonov regularization method.
- To improve the efficiency of the algorithm for solving an ill-posed problem, iterative regularization can be used.
- Regularization methods can also be used to find a singular control.

Notes

1. In fact, this is quite natural. It is enough to compare the elements of the indicated sequence, i.e., sinusoids with an infinitely increasing frequency of oscillation; see Figure 8.1 with optimal control, identically equal to zero.

2. Curiously, the sequence $\{u_k\}$ tends to u_0 weakly in the space $L_2(0,1)$. Indeed, weak convergence in a Hilbert space is convergence in the sense of the corresponding dot product, i.e., convergence to 0 of the numerical sequence $(u_k – u_0, \lambda)$ for any function λ from the space $L_2(0,1)$. Let us find the value

$$(u_k - u_0, \lambda) = \int_0^1 \big[u_k(t) - u_0(t)\big]\lambda(t)dt = \int_0^1 \sin k\pi t \lambda(t)dt.$$

The value on the right side of this equality, up to a constant set, is the coefficient of the expansion of the function λ in a ***Fourier series*** in terms of sines; see [94], [100], [106], [158]. Due to the convergence of the corresponding series for any element of the space $L_2(0,1)$, the corresponding Fourier coefficients must converge to zero, which implies the weak convergence of the sequence $\{u_k\}$ to u_0. By the way, for the sequence under consideration, the Banach–Alaoglu theorem on the existence of a weakly convergent subsequence of any bounded sequence is applicable. The boundedness of the minimizing sequence is realized in view of the obvious boundedness of the set of admissible controls.

3. In that case, with an increase in the number of the element of the sequence, the number of discontinuity points of the function increased indefinitely, while in this case, the oscillation

frequency of the periodic function increases indefinitely. Another form of minimizing control sequence divergence will be encountered in Chapter 11.

4. Naturally, in this case there also exist convergent minimizing sequences. Indeed, an arbitrary sequence of admissible controls descending to u_0 turns out to be minimizing due to the continuity of the functional.

5. On the Tikhonov well-posedness of optimization problems; see [60], [70], [122], [134], [187], [194], [211].

6. Examples of ill-posed optimal control problems are given in [67] and [194]. Tikhonov ill-posed optimal control problems are considered in Chapters 11 (Example 11.2) and 12 (Example 12.3) in the case of a pinned finite state, in Chapter 15 for a problem with an isoperimetric condition (Example 15.12), and in Chapter 17 for a system without the initial condition (Example 17.3).

7. The fact that the minimizing sequence under consideration weakly converges to the optimal control is a small consolation here, since the elements of the sequence, which are sinusoids with increasing frequency, do not in any way remind us of the optimal control, which is a function identically equal to zero. We could call the problem of optimal optimality weakly Tikhonov well-posed according to, but this concept can hardly be considered as meaningful.

8. Experimentally, it is possible to detect the absence of Tikhonov well-posedness correctness tracking the value of the optimality criterion at each step of the algorithm. If it turns out that the sequence of functionals converges and even decrease, while the sequence of controls corresponding to it does not converge, then this is a sure sign of the lack of Tikhonov ill-posedness.

9. In fact, to justify Tikhonov well-posedness, it is sufficient to have the property of ***strictly uniformly convexity*** of the functional; see [194]. A functional I on a convex set U is called strict uniform convex if there exists such a continuous function $\delta = \delta(\tau)$ satisfying the conditions $\delta(0) = 0$, $\delta(\tau) > 0$ for $\tau > 0$ and $\delta(\tau) \to \infty$ for $\tau \to \infty$, which for any elements $u, v \in U$ and $\alpha \in (0, 1)$ takes place the inequality

$$I[\alpha u + (1-\alpha)v] \leq \alpha I(u) + (1-\alpha)I(v) - \alpha(1-\alpha)\delta(\|u-v\|).$$

Obviously, any strictly uniformly convex functional is strictly convex, and even more so, convex. In this case, a strongly convex functional is strictly uniformly convex for the case when the function δ is quadratic.

10. Naturally, the conditions of the theorem implicitly contain restrictions that guarantee the solvability of the optimal control problem. We do not require here the uniqueness of the solution of the problem, since it is already guaranteed by the strong convexity of the functional. However, it is also possible to relax the requirement on the functional space, and also to use in the definition of strong convexity not the square of the norm, but any power greater than or equal to one. The choice of the square of the norm is due solely to the association with the definition of the Hilbert space norm.

11. In Chapter 3, we did not focus on the functional properties of the control. However, they would be indicated in Chapter 7 when substantiating the existence of an optimal control for this example.

12. In passing, we note that the sum of a convex and strongly convex functional is always strongly convex.

13. This means that this functional is affine.

14. The Tikhonov well-posed problem for a system with a fixed final state is considered in Chapter 12; see Example 12.1. In Chapter 13, we will establish the well-posedness in the sense of Tikhonov of optimal control problems with the isoperimetric condition.

15. Concerning the general concept of problem correctness; see [70], [99], [187].

16. Obviously, when $y_k = 0$ we have (up to a constant factor in front of the integral) the previously considered Example 6.2.

17. This problem has almost all the properties of the optimal control problem from the previous example. Using the technique described earlier, we can prove that this problem has a unique solution, the maximum principle for it is the necessary and sufficient condition for optimality, and the corresponding optimal control is singular.

18. An example of a Hadamard ill-posed optimal control problem for a system with a fixed final state is given in Chapter 12; see Example 12.4. Chapter 15 gives examples of optimal control problems with an isoperimetric condition that are Hadamard ill-posed; see Examples 15.7 and 15.13. In Chapter 17, we study the Hadamard ill-posed problem for a system with a free initial state; see Example 17.6.

19. The negative effects of ill-posed problems according to Hadamard appear even when all parameters of the problem are determined absolutely exactly. The fact is that in the process of practical calculation, individual procedures (numerical solution of differential equations, numerical integration, calculation of special functions, roots, etc.) are carried out approximately. In ill-posed problems, small computational errors can cause a large error in determining the solution to the problem. Chapter 12 will consider an optimal control problem with extremely unusual properties that can be interpreted as a strong form of Hadamard ill-posedness; see Examples 12.5 and 12.6.

20. It makes no sense to use the inequality $|I(u) - \min I(U)| \leq \delta$, since the value of the functional on an admissible control u cannot be less than its minimum on the set of all admissible controls.

21. We can also say that a weak conditional approximate solution u of the problem is close enough to the set U itself, where the distance $\rho(u, U)$ from the point u to the set U is determined by the formula $\rho(u, U) = \inf \|u - v\|$, and the infimum is taken over all elements v from the set U.

22. For the optimal control problems under consideration, the equality $A(u, x) = 0$ reduces to the corresponding Cauchy problem.

23. By analogy with Definition 8.5, here one could define a conditional approximate solution, for which the inclusion $u \in U$ is also implemented not exactly, but approximately.

24. In what follows, the convergence of the sequence of parameters will be determined, as a result of which it is assumed that the structure of the topological space is defined on the set M.

25. An analysis of Hadamard well-posedness for problems of the calculus of variations is carried out in [21], [60], [122], [145]; for optimal control problems for systems with lumped parameters in [197], [211] (including, in the presence of phase constraints in [128]); for systems with distributed parameters in [36], [60], [165], [127]; for systems described by variational inequalities in [115]; for minimax problems in [114]; for vector optimization problems in [51].

26. The ***uniform continuity*** of the mapping $\mu \to I(\mu, u)$ in $u \in U$ means that in the case of convergence $\mu_k \to \mu$ we have $\sup |I(\mu_k, u) - I(\mu, u)| \to 0$, where the upper bound is taken over all controls $u \in U$.

27. Here, we use the Schwartz inequality and the fact that the norm of a sum does not exceed the sum of norms.

28. An example of a well-posed optimal control problem for a system with a fixed final state is given in Chapter 12; see Example 12.2.

29. About regularization methods for ill-posed problems; see [70], [99], [187], [194].

30. For example, according to the iterative method described in Chapter 3.

31. To improve the efficiency of the algorithm, it is also desirable to match the rate of decrease in the regularization parameter with the criterion for terminating the iterative process.

32. Naturally, it is not always possible to find the optimal control using Tikhonov regularization method. However, for a fairly wide class of problems to be solved, in particular, for Example 8.1, the desired result is indeed achieved.

33. It would be desirable, of course, to reduce the regularization parameter quickly in order to implement a relatively small number of regularization method steps. However, reducing it too quickly can lead to the divergence of the algorithm.

34. Suppose we are considering some optimal control problem that is solved iteratively without regard to the well-posedness problem, as was done in Chapter 3. We will assume that the algorithm used is sufficiently efficient so that the minimized functional decreases from iteration to iteration. It seems that the longer we count, the more accurate the final result will be. However, in practice, the situation is often different. After a certain number of steps, the decrease in the functional stops. This is explained by the fact that the differential equations (the original and adjoint systems) usually solve themselves approximately, i.e., there is some error due to the approximation method used. Equations, as a rule, are solved relatively accurately, i.e., the approximation error is quite small. At the initial stage of the iterative process, we are far enough from the exact solution so that the iterative error is large enough, and the approximation error does not affect the process. However, as the algorithm converges, the iterative error decreases and at some stage becomes comparable with the approximation error. Further refinement of the algorithm becomes impossible, since now the approximation error becomes decisive, and the algorithm is not tuned to reduce it. If the regularization method is also used, then we have three types of error that are iterative, due to the initial approximation, regularization, associated with the regularization parameter, and approximation, determined by the steps with which the differential equations are solved. The efficiency of the algorithm is determined by the ways of matching the indicated types of error. In particular, at an early stage of the algorithm, it makes no sense to solve the equations accurately enough. In this regard, a fairly coarse grid can be used here. However, as the current approximation is refined (the iterative error decreases) and the regularization parameter decreases (the regularization error decreases), it is possible to switch to a finer grid using interpolation, thereby reducing the approximation error. The corresponding algorithm with corrective approximation can be effective in solving optimal control problems for rather complex equations of state, in particular, for partial differential equations; see [174].

35. Due to regularization, the conditions for the existence of a singular control are violated; see Theorem 6.1.

36. The description of this algorithm for a problem that practically does not differ from a regularized one was given earlier.

III

OPTIMAL CONTROL PROBLEMS FOR SYSTEMS WITH A FIXED FINAL STATE

Part II dealt with optimal control problems for systems described by ordinary differential equations. In this case, the initial state of the system was considered to be known. However, in practice, problems often arise with a known final state, i.e., we are talking about the transfer of the system from one given state to another when a certain result is achieved. This may be minimizing the cost of performing some action, maximizing the quality or quantity of output, minimizing the time to transfer the system to the desired state, etc. In this case, we have a system with a fixed final state, and the control is considered admissible not only if the given restrictions on its value are realized. It is also required that it guarantees the transfer of the system to the desired state. Problems of this nature form the third part of this book, which consists of four chapters. Chapter 9 establishes the maximum principle for systems with a fixed final state. Chapter 10 describes alternative methods for solving such problems. Chapter 11 gives examples of problems of this nature for which the existence or uniqueness of a solution is not realized, as well as cases where the maximum principle is not a sufficient condition for optimality or degenerates. Finally, Chapter 12 considers ill-posed optimal control problems with a fixed final state.

CHAPTER 9

Maximum principle for systems with a fixed final state

The problem of optimal control of a system described by an ordinary differential equation with given initial and final states is considered. In this case, certain restrictions are imposed on the control values, and it is required to minimize an integral functional. For this problem, the necessary optimality conditions are derived in the form of the maximum principle. In the simplest case, they have an analytical solution. In the general case, the problem is solved iteratively.

9.1 LECTURE

The subject of this lecture is optimal control problems for systems described by ordinary differential equations with given initial and final states. It is required to choose such a control, satisfying some additional restrictions, which delivers a minimum to an integral functional. This problem is solved by the maximum principle used earlier for solving optimal control problems with a free final state. The maximum condition here has the same form as in the problems described in Part II. However, in this case, there are no boundary conditions for the adjoint equation, while for the equation of state, on the contrary, there are conditions at both ends of the considered interval. Although this form of the system of optimality conditions turns out to be much more difficult, in the simplest case the optimal control can be found analytically. In the general case, the problem is solved iteratively using the shooting method.

9.1.1 Problem statement

We again consider the controlled system described by the Cauchy problem

$$x'(t) \;=\; f(t, u(t), x(t)),\; t \in (0,1);\; x(0) = x_0 \tag{9.1}$$

DOI: 10.1201/9781003398585-9

with a known function f and a number x_0. Here u is a control chosen from the set[1]

$$U = \{u|\ a(t) \leq u(t) \leq b(t),\ t \in (0,T)\},$$

where a and b are known functions. The functional

$$I(u) = \int_0^T g(t, u(t), x(t))dt$$

is chosen as an optimality criterion[2], where g is a known function. A feature of this problem is the presence of an additional condition

$$x(T) = x_T, \tag{9.2}$$

where the final state of the system x_T is known. Now we get the following ***optimal control problem with a fixed final state***[3].

Problem 9.1 *It is required to find a function u from the set U that ensures the fulfillment of condition* (9.2) *and minimizes the functional I on the subset of the function from U that guarantees the fulfillment of equalities* (9.2).

Thus, the solution of the optimal control problem is such a control function from a given set that transfers the system characterized by the given equation from one known state to another, while minimizing the optimality criterion. The presence of an additional condition (9.2) is the main feature of this problem, as a result of which, in order to solve it, it is necessary to modify the mathematical apparatus at our analysis.

9.1.2 Maximum principle

To determine the necessary optimality condition, we use the previously described method. Let us assume that a function u is an optimal control, so the following inequality holds

$$\Delta I = I(v,y) - I(u,x) \geq 0,$$

where x is the optimal state of the system, and v is an arbitrary control from the set U, for which the corresponding state y satisfies the equalities[4] (9.1) and (9.2).

As in Chapter 3, we use the ***Lagrange multiplier method*** to account for the state equation[5]. Introduce the Lagrange functional

$$L(u,x,p) = I(u) + \int_0^T p(t)\big[x'(t) - f(t, u(t), x(t))\big]dt,$$

where the function p is arbitrary. Then we obtain the inequality

$$\Delta L = L(v,y,p) - L(u,x,p) \geq 0 \ \ \forall v, p, \tag{9.3}$$

similar (3.5).

Combine all the terms that are under the integral in the functional L, which explicitly depend on the control, as was done in Chapter 3. We have the function

$$H(t,u,x,p) \;=\; pf(t,u,x) - g(t,u,x).$$

As a result, inequality (9.3) takes the form

$$\int_0^T (p\Delta x' - \Delta H)dt \ge 0 \;\; \forall v, p, \tag{9.4}$$

where $\Delta x = y - x$, $\Delta H = H(t,v,y,p) - H(t,u,x,p)$.

The first term under the integral in inequality (9.4) is transformed using the formula of integration by parts

$$\int_0^T p\Delta x' dt \;=\; p(T)\Delta x(T) - p(0)\Delta x(0) - \int_0^T p'\Delta x dt.$$

This takes into account that both the initial and final values of the state function are fixed, which means that they take the same values for any control, as a result of which the increment Δx at the initial and final time is equal to zero[6]. The ΔH increment is converted using the Taylor series expansion in the same way as in Chapter 3 using the notation adopted there. We have

$$\Delta H \;=\; \Delta_u H + H_x(t,u,x,p)\Delta x + \eta_1 + \eta_2,$$

where $\Delta_u H = H(t,v,x,p) - H(t,u,x,p)$, H_x is the partial derivative of the function H with respect to x, η_1 is a second order term with respect to Δx, obtained by expanding the value $H(t,u,x+\Delta x,p)$ in a Taylor series, $\eta_2 = [H_x(t,v,x,p) - H_x(t,u,x,p)]\Delta x$.

As a result, inequality (9.4) is reduced to the form

$$-\int_0^T \Delta_u H dt \;-\; \int_0^T [H_x(t,u,x,p) + p']\Delta x dt \;+\; \eta \ge 0 \;\; \forall v, p, \tag{9.5}$$

where the remainder term η is determined by the formula

$$\eta \;=\; -\int_0^T (\eta_1 + \eta_2) dt.$$

The resulting formulas differ from similar relations from Chapter 3 only by the absence of terminal term[7].

At the next stage, inequality (9.5) is simplified by choosing the solution of the adjoint system as an arbitrary function p. This was achieved in Chapter 3 by equalizing the value to zero, multiplied by the increment of the state function Δx under the integral and in the terminal term. As a result, a differential equation was obtained for the function p with the condition at the final moment of time, which corresponded to the adjoint system (3.8) and (3.9). In this case, since there is no terminal term in inequality (9.5), we can only obtain the adjoint equation

$$p'(t) = -H_x(t,u,x,p),\ t \in (0,T), \tag{9.6}$$

coinciding with (3.8), but without an additional condition at the final time. As a result of such a choice of the function p, inequality (9.5) takes the form

$$-\int_0^T \Delta_u H dt + \eta \geq 0 \ \ \forall v \in U.$$

The resulting value differs from the analogous inequality (3.10) only in the absence of a terminal component in the remainder term. By analogy with Theorem 3.1, we prove the following assertion[8].

Theorem 9.1 *In order for the control u to be a solution to Problem* 9.1, *it is necessary that it satisfies the maximum condition*

$$H\big[t,u(t),x(t),p(t)\big] = \max_{v\in[a(t),b(t)]} H\big[t,v,x(t),p(t)\big],\ t \in (0,T), \tag{9.7}$$

where x is the solution of problem (9.1) *corresponding to it, which also satisfies condition* (9.2), *and p is the solution of the conjugate equation* (9.6).

Thus, the system of optimality conditions is characterized by formulas (9.1), (9.2), (9.6), and (9.7), whence it is required to find three unknown functions u, x, p. At first glance, the resulting system is obviously not correct. On the one hand, the problem regarding the state function is overdetermined: for a first-order differential equation, two boundary conditions are set at once. On the other hand, for the adjoint equation, there are no additional conditions at all. However, in reality, we can find from the maximum condition the dependence of the control on the functions x and p and substitute it into the equation of state and the adjoint equation. The result is a system of two differential equations of the first order, for which there are two boundary conditions. In this connection, the given optimality conditions make sense.

Let us now turn to the practical solution of the optimality conditions.

9.1.3 Example of an analytical solving to a problem

In Chapter 3 it was shown that for fairly simple examples, the solution to the system of optimality conditions can be found analytically without using any iterative process. Let us show that a similar situation is also possible for systems with a fixed final state.

Example 9.1 *It is required to minimize the functional*

$$I(u) = \int_0^1 (u+x)dt$$

on a subset of such functions $u = u(t)$ from the set

$$U = \{u|\ 0 \leq u(t) \leq 2,\ t \in (0,1)\},$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0, \tag{9.8}$$

ensure the fulfillment of the condition

$$x(1) = 1. \tag{9.9}$$

We have Problem 9.1 with the following parameter values:

$$f(t,u,x) = u,\ T = 1,\ x_0 = 0,\ x_T = 1,\ a(t) = 0,\ b(t) = 2,\ g(t,u,x) = u + x.$$

In accordance with the method described above, we determine the function

$$H(t,u,x,p) = pf–g = pu–u–x.$$

Then the adjoint equation (9.6) has the form

$$p'(t) = 1,\ t \in (0,1). \tag{9.10}$$

According to Theorem 9.1, the optimal control satisfies the maximum condition

$$[p(t) - 1]u(t) - x(t) = \max_{v\in[0,2]} [p(t) - 1]v - x(t),\ t \in (0,1). \tag{9.11}$$

Thus, to find three unknown functions u, x, p there is a system (9.8) – (9.11).

Note that the adjoint equation (9.10) does not depend on other unknown functions. Then one can find its general solution $p(t) = t + c$, where c is an arbitrary constant. It can be found after substituting the value of the function p into the maximum condition (9.11) and considering the resulting relation together with equalities (9.8) and (9.9).

Let us now turn to the solution of the maximum principle (9.11). First of all, we check whether there is a singular control for this problem. Note that in this case there is the possibility of the control coefficient vanishing in the definition of the function H. This is realized when $p(t) = 1$. However, it was previously established that the function p is a variable. Thus, the maximum principle cannot degenerate.

Taking into account the linearity of the function H with respect to the control, we conclude that the solution of the maximum condition is reached on the boundary of the set of admissible controls. Thus, we find

$$u(t) = \begin{cases} 2, & \text{if } p(t) > 1, \\ 0, & \text{if } p(t) < 1. \end{cases}$$

Note that the function p is increasing. As a result, three variants of control behavior are possible. Either the function p at the initial time, and hence all the time, is greater than 1, and then the control is identically equal to 2; either p at the final moment of time, and hence all the time, is less than 1, and then the control is identically equal to 0; or at some point from the interval (0,1) the function p, increasing, reaches the value of 1, and then before this time the control is equal to 0, and after that this is equal to 2. Let us consider all three options.

If the control is identically equal to 2, then the corresponding solution to problem (9.8) is equal to $x(t) = 2t$. Then $x(1) = 2$, which contradicts condition (9.9). If the control is identically equal to 0, then the corresponding solution to problem (9.8) is equal to $x(t) = 0$, and hence $x(1) = 0$. This also contradicts condition (9.9).

Now let there be a point $\xi \in (0, 1)$ such that $u(t) = 0$ for $t < \xi$ and $u(t) = 2$ for $t > \xi$. Then the solution to problem (9.8) for $t < \xi$ is equal to $x(t) = 0$. In particular, $x(\xi) = 0$. This value is chosen as the initial condition when solving the equation of state for $t > \mu$. The corresponding solution has the form $x(t) = 2(t–\xi)$, and hence $x(1) = 2(1–\xi)$. Taking into account equality (9.9), we conclude that $\xi = 1/2$. Thus, the unique solution to the necessary optimality condition has the form[9]

$$u(t) = \begin{cases} 0, & \text{if } t < 1/2, \\ 2, & \text{if } t > 1/2. \end{cases}$$

It is the solution of the considered optimal control problem[10].

This Example is in a certain sense similar to Examples 1.1, 3.1, and 3.2, in which the necessary extremum condition had a unique solution, which was found analytically and turned out to be a solution to the problem posed[11]. Naturally, for more difficult systems, the optimal control can only be found using some iterative methods.

9.1.4 Approximate solving of a problem with a fixed final state

As seen in Example 9.1, the system of optimality conditions for optimal control problems with a fixed final state can be analyzed analytically. The possibility of finding an analytical solution in this case is due to the fact that both the equation of state and the optimality criterion are linear with respect to the function x, as a result of which the solution of the adjoint equation was found without resorting to the equation of state and the maximum condition[12]. For more difficult problems, it is not possible to explicitly find the optimal control, and the desired result is established using some iterative methods.

To find an approximate solution to Problem 9.1, one could try to use the method of successive approximations described in Chapter 3 and repeatedly used earlier. Indeed, from the maximum condition (9.7), in principle, one can find an explicit dependence of the control on the functions x and p, as was done in Part II. Next, one can specify some initial approximation of the control from the set U and substitute it into the Cauchy problem (9.1). However, it is extremely unlikely that the corresponding function x will satisfy the final condition (9.2). Moreover, even if we succeed, we are not be able to find a function p from equality (9.6) in view of the absence of an additional condition for this differential equation. Therefore, we cannot to move to the next iteration using the maximum condition. Thus, to find a solution to the problem, it is necessary to use a qualitatively different iterative method.

Difficulties in the practical solution of the optimality conditions in this case are due to the fact that we are actually dealing not with the Cauchy problem, but with a boundary value problem for differential equations, since the function x has conditions at both ends of the interval $(0, T)$. As a result, to find an approximate solution to the system of optimality conditions, we use the ***shooting method***, which allows us to reduce the boundary value problem to the Cauchy problem[13]

As already noted, we cannot proceed to the next iteration in accordance with the iterative process described in Chapter 3 due to the absence of an additional condition for the p function. We set this condition artificially

$$p(T) = \psi, \tag{9.12}$$

where ψ is unknown numeric parameter. Let us now assume that from the maximum condition (9.7) we have succeeded in determining the control as a function of x and p. Then, by setting some value of the number ψ and solving the Cauchy problem (9.1), (9.2), (9.6), and (9.12), one can find the functions x and p, which naturally depend on the choice of the number ψ. Then for each number ψ we can define the function $F = F(\psi)$ defined by the formula

$$F(\psi) = x(T) - x_T,$$

where $x(T)$ is the value of the function $x = x(t)$, corresponding to the given value of the parameter ψ, at the point $t = T$. For the validity of condition (9.2), it is required to choose the parameter ψ in such a way that the following equality holds

$$F(\psi) = 0. \tag{9.13}$$

Considering that the parameter ψ is numerical, equality (9.13) can be interpreted as a non-linear algebraic equation with respect to ψ. For its approximate solution, one can use any iterative method. As such, one can choose, for example, the following iterative method[14]

$$\psi_{k+1} = \psi_k - \beta_k F(\psi_k), \tag{9.14}$$

where β_k is an iterative parameter.

The practical implementation of the described algorithm is as follows[15].

1. The initial approximations of the control u_0, the final value ψ_0 of the function p and a sequence of algorithm parameters $\{\beta_k\}$ are chosen.

2. At the current k-th iteration, from the Cauchy problem (9.1), with a known value $u = u_k$, the corresponding value of the state function $x = x_k$ is determined.

3. With known values $u = u_k$, $x = x_k$ and $\psi = \psi_k$, the solution $p = p_k$, of the problem (9.6) and (9.12) is calculated.

4. With known values $x = x_k$ and $p = p_k$ from maximum condition (9.7) the next iteration of control u_{k+1} is determined.

5. A new approximation of the parameter $\psi_{k+1} = \psi_k - \beta_k[x_k(T) - x_T]$ is calculated.

In the case of convergence of the described algorithm, the result is a solution to the system of optimality conditions.

RESULTS

Here is a list of questions about problems of optimal control of systems with a fixed final state, the main conclusions on this topic, as well as the problems that arise in this case and require additional research.

Questions

It is required to answer questions concerning optimal control problems for systems with a fixed final state and the optimality condition for such problems.

1. Why are optimal control problems for systems with a fixed final state of practical interest?

2. In Problem 9.1, the state is described by a first-order ordinary differential equation with two boundary conditions, i.e., the system is redefined. Why does the equation of state underlying the optimal control problem make sense under these conditions?

3. In the general optimal control problem considered in Chapter 3, the optimality criterion consisted of integral and terminal terms. Why is there no terminal term in Problem 9.1?

4. Which control in Problem 9.1 is considered acceptable?

5. Does there always exist a control from a given set U that transfers the system from one given state to another in a specified time?

6. Is there any difference between the form of the remainder term in the functional increment formula for similar systems with a free and a fixed final state?

7. What is the fundamental difference between optimality conditions for systems with a free and a fixed final state?

8. What is the reason for the difference between the optimality conditions for systems with a free and with a fixed final state?

9. Is there a difference between the maximum condition for systems with a free and a fixed final state?

10. Why is there no boundary condition when deriving optimality conditions for an unknown function p?

11. Why do the above transformations not provide a complete justification of the maximum principle for Problem 9.1?

12. How many controls from a given set U take the system considered in Example 9.1 from a given initial state to a given final state?

13. Can the solution of the maximum principle for Example 9.1 be a singular control?

14. Can, in principle, the solution of the maximum principle for a system with a fixed final state be a singular control?

15. Why was it possible to find the optimal control for Example 9.1 without uniquely defining the solution of the adjoint equation?

16. Why was it possible to obtain the optimal control analytically for Example 9.1?

17. Why did the optimal control for Example 9.1 turn out to be a piecewise constant function?

18. How reasonable is the assertion that the formula for control obtained as a result of the analysis of the optimality conditions for Example 9.1 really gives a solution to this problem?

19. Why is it not possible to use the iterative algorithm used in solving optimal control problems for systems with a free final state for a system of optimality conditions for systems with a fixed final state?

20. What is the meaning of the shooting method?

21. In accordance with the shooting method for the adjoint equation, the missing condition (9.12) is added. What does it give if the parameter a included in it is still not known in advance?

22. Why is it not enough to set the control at the initial iteration in the considered iterative process?

23. Is there any confidence that in the case of convergence of the iterative process, we get a control that guarantees the fulfillment of the final condition (9.2)?

24. Is there any confidence that in the case of convergence of the iterative process, we really get the solution of the optimal control problem?

Conclusions

Based on the study of the optimal control problem for systems with a fixed final state, we can come to the following conclusions.

- The problems of optimal transfer of a system from one state to another are of great practical interest.
- For problems of optimal control of systems with a fixed final state, one can use the maximum principle.
- Optimality conditions for problems with a fixed final state include a maximum condition, a state equation with two boundary conditions, and an adjoint equation without boundary conditions.
- In the simplest case, the solution of optimality conditions for systems with a fixed final state can be found analytically.
- In the general case, the solution of optimality conditions for systems with a fixed final state can be found approximately.
- The previously described iterative method for solving optimality conditions in the form of the maximum principle turns out to be unsuitable for solving problems with a fixed final state due to the overdetermination of the equation of state and the underdetermination of the adjoint system.
- For a practical solving of the system of optimality conditions in this case, one can use the shooting method.
- The shooting method is based on the reduction of the boundary value problem for differential equations to the Cauchy problem.
- The shooting method involves introducing the missing boundary condition for the adjoint equation with an unknown value of the corresponding function and interpreting the final condition for the state function as an algebraic equation with respect to the specified unknown value.
- For the practical application of the shooting method, at the initial iteration, the control and the final state for the adjoint equation are set.

Problems

In the process of analyzing optimal control problems for systems with a fixed final state, additional problems arise that require additional research.

1. **Nontriviality of the problem statement**. In principle, we do not know in advance whether there is at least one control from a given set U that transfers the system from one given state to another. These questions constitute the problem of system ***controllability***, see Notes[16]

2. **Justification of the maximum principle**. Theorem 9.1 is given without a full justification. Regarding the substantiation of the maximum principle for various systems with a fixed final state, see Notes[17]

3. **Vector case**. The above results are naturally generalized to the case when both the state function and the control are vector quantities. Such problems are discussed in Chapter 10.

4. **Generalization of restrictions**. We consider here two types of constraints that are explicit constraints on control values and fixing the final state of the system. In Part IV, optimal control problems will be considered, in which some integral constraints, called isoperimetric constraints, are additionally assumed to be satisfied. In this case, both systems with a free final state and with a fixed final state will be considered.

5. **Qualitative analysis of Example 9.1**. Part II presents general results concerning the existence and uniqueness of solutions to optimal control problems, as well as the sufficiency of optimality conditions. It is of interest to extend them to systems with a fixed final state. In Appendix, these results are applied to the analysis of Example 9.1.

6. **Maximizing the functional from Example 9.1**. Previously, we have repeatedly encountered a situation where the minimization and maximization problems of the same functional have qualitatively different properties. In this regard, it is of interest to solve the problem of maximizing the functional from Example 9.1. Its solving is carried out in Appendix.

7. **Examples from Part II**. In Part II, we considered examples of optimal control problems for systems with a free final state that has an analytical solution. It would be interesting to consider similar examples for systems with a fixed final state. Appendix considers an analogue of Example 3.1.

8. **Decoupling method**. In Chapter 3, it was shown that for a linear system with a free final state with a quadratic functional in the absence of restrictions on controls, the system of optimality conditions can be solved by the decoupling method without using an iterative process. In Chapter 10, these results are extended to systems with a fixed final state.

9. **Variational inequality**. The Lecture described a method for solving the simplest optimal control problem with a fixed final state using the maximum principle. However, in Chapter 4, a variational inequality was used to solve the problem with a free final state. In Chapter 10 it is used to analyze optimal control problems for systems with a fixed final state.

10. **Penalty method**. To solve problems of finding an extremum with equality-type constraints, we have previously repeatedly used the penalty method; see Chapters 2 and 4. It is based on removing the existing constraint by introducing an additional parameter with a small parameter in the denominator into the

optimality criterion. It seems natural to use this idea to solve optimal control problems for systems with a fixed final state. In this case, the equality that specifies the final state for the system is interpreted as a restriction removed using the penalty method. The corresponding results are given in Chapter 10.

11. **Counterexamples for optimal control problems for systems with a fixed final state**. For Example 9.1, as well as for the examples studied in Appendix, there is a unique optimal control, and the maximum conditions are necessary and sufficient and do not degenerate. However, in Part II, examples of problems with a free final state were given, where these properties were violated. It is of interest to consider optimal control problems for a system with a fixed final state, for which these properties are not satisfied. Examples of such problems are given in Chapter 11. Examples of ill-posed optimal control problems for a system with a fixed final state are given in Chapter 12.

12. **Optimal control problems with a free initial state**. Moving from Part II to Part III, we imposed an additional condition on the system, considering the final state of the system to be known. However, the opposite situation is also possible, when the initial state of the system re-mains unknown. Part V is devoted to consideration of optimal control problems with a free initial state.

9.2 APPENDIX

Below, we present additional results in the field of optimal control theory for systems with a fixed final state. In particular, in Section 9.2.1, the existence and uniqueness of the optimal control for Example 9.1 are proved, as well as the sufficiency of the corresponding optimality conditions. Section 9.2.2 maximizes the functional from Example 9.1. In Section 9.2.3, problems of optimal control of systems with a fixed final state with a functional quadratic in control and linear in the state of the system are studied, which are analogues of Examples 3.1 and 3.2.

9.2.1 Qualitative analysis of Example 9.1

Chapter 5 provides theorems on the uniqueness of solutions to optimal control problems and on the sufficiency of the optimality condition in the form of the maximum principle. In Chapter 7, Theorem of the existence of an optimal control is proved. Let us apply these results to the analysis of Example 9.1.

Theorem 5.2 provides properties that guarantee the sufficiency of optimality conditions for a system with a free final state, based on the non-negativity of the remainder term in the functional increment formula. These properties make it possible to reverse the whole chain of reasoning, starting with the assumption that some control is optimal and ending with the conclusion that it satisfies the maximum condition. In this case, the presence or absence of a final condition for the state function affects only the absence or presence of a similar additional condition for the adjoint equation. Thus, the non-negativity of the remainder term guarantees the sufficiency of the optimality condition for the considered class of problems.

As noted earlier, the remainder term for Problem 9.1 is calculated by the formula

$$\eta = -\int_0^T (\eta_1 + \eta_2) dt,$$

where η_1 is determined by terms of the second order when expanding the value $H(t, u, x + \Delta x, p)$ in a series of Δx, and $\eta_2 = [H_x(t, v, x, p) - H_x(t, u, x, p)]\Delta x$. For Example 9.1, we have is $H = pu - u - x$. Considering that its derivative with respect to the state of the system is constant, we find $\eta = 0$, which guarantees the sufficiency of the optimality condition. It was previously established that the system of optimality conditions has a unique solution. If the optimality conditions are sufficient, both the existence and the uniqueness of the optimal control already follow from this. However, we will try to establish these properties based on the general results obtained in Part II.

According to Theorem 5.1, a strictly convex functional on a convex subset cannot have two minimum points. In this case, a control belonging to the set

$$U = \{u|\, 0 \le u(t) \le 2,\ t \in (0,1)\},$$

which ensures the fulfillment of condition (9.2), i.e., equalities $x(1) = 1$. From problem (9.2) we find

$$x(1) = \int_0^1 u(t) dt.$$

Thus, the optimality criterion is minimized at the intersection of the sets U and

$$V = \Big\{u\Big| \int_0^1 u(t) dt = 1\Big\}.$$

Let there be two functions u and v belonging to both these sets, i.e., satisfying inequalities

$$0 \le u(t) \le 2,\ \ 0 \le v(t) \le 2,\ \ t \in (0,1)$$

and equalities

$$\int_0^1 u(t) dt = 1,\ \ \int_0^1 v(t) dt = 1.$$

For an arbitrary number $\alpha \in (0,1)$ we multiply the first of the above inequalities by $1-\alpha$, the second one by α and add the resulting values. We have inequality

$$0 \le (1-\alpha)u(t) + \alpha v(t) \le 2,\ t \in (0,1),$$

whence follows the convexity of the set U. Now we multiply the first of the above integral equalities by $1-\alpha$, the second by α and add the results. We get

$$\int_0^1 \big[(1-\alpha)u(t) + \alpha v(t)\big] dt = 1.$$

Thus, the set V is convex, and hence its intersection with U is convex too.

It remains to investigate the properties of the optimality criterion. Taking into account the representation of the solution of the Cauchy problem (9.1), we define the functional

$$I(u) = \int_0^1 \big[u(t) + x(t)\big]dt = \int_0^1 \Big[u(t) + \int_0^t u(\tau)d\tau\Big]dt.$$

Obviously, this one is linear. As a result, we obtain the equality

$$I\big((1-\alpha)u + \alpha v\big) = (1-\alpha)I(u) + \alpha I(v).$$

Thus, the minimized functional is convex, but not strictly convex. Therefore, we are unable to use Theorem 5.1 to prove the uniqueness of the solution to the problem. Nevertheless, the optimal control is unique, since only one control satisfies the optimality condition[18].

Let us now try to use Theorem 7.1 to prove the existence of a solution to the problem under consideration. According to this statement, the problem of minimizing a convex continuous functional bounded below on a convex closed bounded subset of a Hilbert space has a solution.

We investigate the properties of the set of admissible controls for Example 9.1, which is the intersection $U \cap V$ of the sets defined above. First of all, we clarify that the control in this example can again be chosen from the Hilbert space $L_2(0,1)$ of square-integrable functions on the given interval (0,1). The boundedness of the set under consideration is obvious[19], and its convexity was established above. Let us show that this set is closed.

Let there be a sequence of functions $\{u_k\}$ of the set under consideration such that $u_k \to u$ in $L_2(0,1)$. It is required to establish that the limit u belongs to the sets U and V. As already noted in Chapter 7, the convergence of a sequence in the space $L_2(0,1)$ implies the existence of some of its subsequences converging almost everywhere. Thus, from $\{u_k\}$ one can single out a subsequence $\{u_s\}$ such that the numerical sequence $u_s(t) \to u(t)$ converges for almost values of t from the interval (0,1). The inclusion $u_s \in U$ implies the validity of the inequality $0 \leq u_s(t) \leq 2$. Hence, as a result of passing to the limit, we obtain the inequality[20] $0 \leq u(t) \leq 2$. Therefore, the inclusion $u \in U$ is true.

Let us now show that the indicated limit also belongs to the set V. It is known that the strong convergence of a sequence in a Hilbert space implies its weak convergence[21], i.e., condition $(u_k, v) \to (u, v)$ for all v from $L_2(0,1)$. We choose here as v a function that is identically equal to unity. As a result, we get the convergence

$$\int_0^1 u_k(t)dt \to \int_0^1 u(t)dt.$$

Given that the left side of this ratio is equal to one, we conclude that its right side is the same, and hence $u \in V$. Thus, the optimality criterion for Example 9.1 is minimized on the closed set.

Let us now turn to the properties of the functional to be minimized. Its convexity has been proven above. Let us establish that it is lower bounded. It follows from the definition of the set U that all values of the admissible control u are non-negative. Then all values of the corresponding state function

$$x(t) = \int_0^t u(\tau)d\tau$$

are also non-negative. This implies that the optimality criterion on any admissible control cannot take negative values, and hence is lower bounded.

Let us now establish the continuity of the optimality criterion. Let the convergence $u_k \to u$ again take place in $L_2(0,1)$. Denote by x_k and x the solutions to problem (9.1) corresponding to the controls u_k and u. We get the inequality

$$\big|x_k(t) - x(t)\big| = \Big|\int_0^t \big[u_k(\tau) - u(\tau)\big]d\tau\Big| \leq \int_0^t \big|u_k(\tau) - u(\tau)\big|d\tau \leq \int_0^1 \big|u_k(\tau) - u(\tau)\big|d\tau.$$

Here, the function under the integral on the right-hand side can be interpreted as the scalar product of the given expression and a function that is identically equal to one. Then, applying the Schwartz inequality, we obtain the estimate

$$\big|x_k(t) - x(t)\big| \leq \|u_k - u\|,$$

and hence $x_k(t) \to x(t)$. Taking into account the previously established convergence of the sequence of integrals of controls, we have

$$\int_0^1 \big[u_k(t) + x_k(t)\big]dt \to \int_0^1 \big[u(t) + x(t)\big]dt,$$

and hence $I(u_k) \to I(u)$. Thus, the minimized functional is continuous. Thus, all the conditions of Theorem 7.1 are satisfied, which means that the optimal control for Example 9.1 does indeed exist[22].

9.2.2 Maximizing the functional from Example 9.1

As already noted, the problems of minimizing and maximizing the same functional often have qualitatively different properties[23]. Let us now turn to the problem of maximizing the functional from Example 9.1.

Example 9.2 *It is required to maximize the functional*

$$I(u) = \int_0^1 (u + x)dt$$

on a subset of such functions u=u(t) from the set

$$U = \{u|\ 0 \le u(t) \le 2,\ t \in (0,1)\},$$

which for solutions of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

ensure the fulfillment of the condition

$$x(1) = 1.$$

As in Example 9.1, the function H is defined by the formula

$$H(t,u,x,p) = pu–u–x,$$

where p is a solution of the adjoint equation

$$p'(t) = 1,\ t \in (0,1),$$

coinciding with (9.10). However, instead of the maximum condition (9.11), the optimal control satisfies the corresponding minimum condition

$$[p(t) - 1]u(t) - x(t) = \min_{v \in [0,2]} [p(t) - 1]v - x(t),\ t \in (0,1).$$

The solution of the adjoint equation is again defined by the formula $p(t) = t + c$, where c is an arbitrary constant, i.e., the function p is increasing. However, the solution of the optimality condition now has the form

$$u(t) = \begin{cases} 0, & \text{if } p(t) > 1, \\ 2, & \text{if } p(t) < 1. \end{cases}$$

Compared to the similar formula from Example 9.1, here the boundaries are reversed. We again have three possible behaviors of the system. Here, either the function p at the initial time, and hence all the time, is greater than one, and then the control is identically equal to zero; either p at the final time, and hence all the time, is less than one, and then the control is identically equal to two; or at some point in time from the interval (0,1), the function p, increasing, reaches the value of one, and then before this time the control is equal to two, and after that it is equal to zero.

As noted in the study of Example 9.1, controls that are identically equal to zero or two do not guarantee the fulfillment of the equality $x(1) = 1$, which means that they are certainly not solutions to the problem. Thus, it suffices to consider only the third case. Now let there be a point $\xi \in (0,1)$ such that $u(t) = 2$ for $t < \xi$ and $u(t) = 0$ for $t > \xi$. Then the solution of the Cauchy problem for $t < \xi$ is $x(t) = 2t$. In particular, $x(\xi) = 2\xi$. Since the function x does not change in the next time interval (its derivative, i.e., the control is equal to zero), then the equality

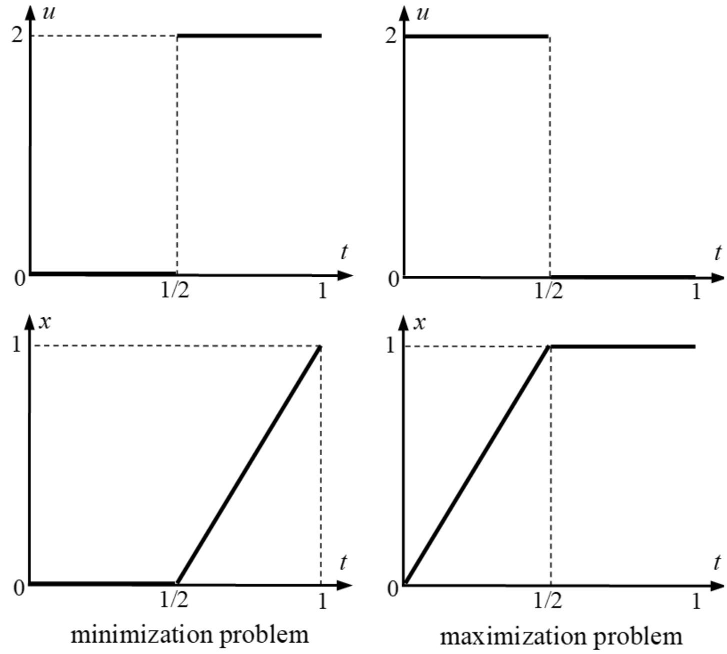

Figure 9.1 Solutions of problems from Examples 9.1 and 9.2.

$x(1) = 2\xi$ is satisfied. As a result, we find $\xi = 1/2$. Therefore, the unique solution to the optimality condition is

$$u(t) = \begin{cases} 2, & \text{if } t < 1/2, \\ 0, & \text{if } t > 1/2. \end{cases}$$

Compared to Example 9.1, the values that the control takes in the first and second half of the considered time interval have changed here; see Figure 9.1.

Since the remainder term in the minimization and maximization problem is the same, and in the preceding Subsection it is stated that it is equal to zero, we conclude that the optimality conditions for Example 9.2 are also necessary and sufficient[24].Thus, the above formula does indeed give a solution to the optimization problem[25].

9.2.3 Problem with a quadratic functional

Examples 9.1 and 9.2 considered earlier, the solution of which was found analytically, are characterized by a linear functional. However, in Chapter 3, the solution to the problem for Examples 3.1 and 3.2 was found analytically, although the optimality criterion was quadratic with respect to control. Let us show that similar results can be obtained for a system with a fixed final state.

Example 9.3 *It is required to minimize the functional*

$$I(u) = \int_0^1 \Big(\frac{1}{2}u^2 - 3x\Big)dt$$

on a subset of such functions $u = u(t)$ *from the set*

$$U = \big\{u \big| \, 1 \leq u(t) \leq 2, \ t \in (0,1)\big\},$$

where x *is a solution of the Cauchy problem*

$$x'(t) = u(t), \ t \in (0,1); \ x(0) = 0 \tag{9.15}$$

ensure the fulfillment of the condition

$$x(1) = 3/2. \tag{9.16}$$

This problem differs from the one considered in Example 3.1 only by the presence of an additional condition (9.16). To solve it, we use Theorem 9.1. Define a function

$$H = pu – u^2/2 + 3x.$$

The function p in this formula is a solution to the adjoint equation

$$p'(t) = -3, \ t \in (0,1). \tag{9.17}$$

From the condition of the maximum of the function H with respect to the control on the interval [1,2], we find the control

$$u(t) = \begin{cases} 1, & \text{if } p(t) < 1, \\ p(t), & \text{if } 1 \leq p(t) \leq 2, \\ 2, & \text{if } p(t) > 2. \end{cases} \tag{9.18}$$

The result obtained exactly coincides with what was established in the analysis of Example 3.1. However, in that case for the adjoint equation (9.17) there was an additional condition for $t = 1$. Thus, for the function p, the Cauchy problem was obtained, the solution of which was then substituted into formula (9.18). The resulting formula determined the optimal control for Example 3.1.

In this case, due to the presence of additional condition (9.16), there is no boundary condition for the adjoint equation. Thus, we cannot define the function p uniquely. We can only assert that equation (9.17) is satisfied by any function of the form $p(t) = c – 3t$, where c is an arbitrary constant. As a result, formula (9.18) takes the form

$$u(t) = \begin{cases} 2, & \text{if } c - 3t > 2, \\ c - 3t, & \text{if } 1 \leq c - 3t \leq 2, \\ 1, & \text{if } c - 3t < 1. \end{cases} \tag{9.19}$$

Further, in principle, one should substitute the found control into the equation of state (9.15). Naturally, the solution of the latter will depend on the choice of the constant c, which should be chosen in such a way that condition (9.16) is satisfied.

Note that, regardless of the constant c, the function p is decreasing. As a result, the following behaviors of the system are possible. Either $c > 5$, which means $c–3t > 2$ for all $t \in (0, 1)$, and then $u(t) = 2$; or $c < 1$, and hence $c–3t < 1$ for all $t \in (0, 1)$, and then $u(t) = 1$; or $1 \leq c \leq 5$, and then the control, in accordance with formula (9.19), has one or two discontinuity points on the interval (0,1).

For $u(t) = 2$, for any t, problem (9.15) has a solution $x(t) = 2t$, and hence $x(1) = 2$. However, this contradicts condition (9.16). For $u(t) = 2$, for all t, problem (9.15) has a solution $x(t) = t$, and hence $x(1) = 1$, which also contradicts equality (9.16). Consequently, the control is discontinuous. This means that there are points ξ and η such that $0 \leq \xi < \eta \leq 1$, and the following equalities hold

$$c - 3\xi = 2, \quad c - 3\eta = 1. \tag{9.20}$$

Note that for $\xi = 0$ and for $\eta = 1$ there is only one discontinuity point, and in other cases, the control takes the value 2 on the interval $(0, \xi)$, equals the function $c–3t$ on the interval (ξ, η), after which it takes value 1. Let us find the corresponding solution of problem (9.15).

Integrating the equality $x'(t) = 2$ from 0 to ξ, taking into account the initial condition $x(0) = 0$, we find the value $x(\xi) = 2\xi$. It is the initial condition for the equation $x'(t) = c - 3t$ on the interval (ξ, η). From equalities (9.20) we find the constants $c = 3\xi + 2$ and $\eta = \xi + 1/3$. Thus, there is an equation $x'(t) = 2 + 3\xi - 3t$ on the interval $(\xi, \xi + 1/3)$ with the initial condition $x(\xi) = 2\xi$. As a result of solving this Cauchy problem, the value $x(\xi + 1/3) = 2\xi + 1/2$ is determined. Finally, we obtain an equation $x'(t) = 1$ on the interval $(\xi + 1/3, 1)$ with the initial condition $x(\xi + 1/3) = 2\xi + 1/2$. Solving this problem, we determine $x(1) = \xi + 7/6$.

Comparing the result obtained with formula (9.16), we conclude that $\xi = 1/3$. Taking into account the connection between the numbers c, η and ξ indicated earlier, we conclude that $c = 3$ and $\eta = 2/3$. Substituting the found values into formula (9.19), we find the unique solution of the optimality conditions

$$u(t) = \begin{cases} 2, & \text{if } 0 < t < 1/3, \\ 3 - 3t, & \text{if } 1/3 \leq t \leq 2/3, \\ 1, & \text{if } 2/3 < t < 1. \end{cases}$$

The result obtained exactly coincides with the optimal control for Example 3.1 and is the solution of the optimal control problem for the considered example[26]. Thus, this problem has a unique solution, and the optimality conditions used for it are necessary and sufficient[27].

Let us consider one more example.

Example 9.4 *It is required to find the maximum of the functional from Example* 9.3 *on the same set.*

The only difference from Example 9.3 here is that the solution to the problem is determined from the minimum condition for the function H. Its only stationary point corresponds to the maximum of this function. Thus, its minimum can be reached exclusively at the boundary. Find the values

$$H(1) = p–1/2 + 3x, \quad H(2) = 2p–2 + 3x.$$

Now we obtain

$$u(t) = \begin{cases} 1, & \text{if } p(t) > 3/2, \\ 2, & \text{if } p(t) < 3/2. \end{cases}$$

The solution of the adjoint equation again has the form $p(t) = c–3t$, where c is an arbitrary constant, which means that the function p is decreasing. Therefore, there are three possible behaviors of the system. Either the function p is always greater than 3/2, and the control is equal to one; either p is always less than 3/2 and the control is two; or p is first greater than 3/2 and then less than this value, and the control has a jump from value 1 to 2. When examining Example 9.3, it was noted that the first two cases cannot ensure the condition $x(1) = 3/2$. Thus, only the last option remains.

Suppose that there is a point ξ from the interval (0,1), that $u(t) = 1$ for $t < \xi$ and $u(t) = 2$ for $t > \xi$. From the equation $x' = 1$ with the initial condition $x(0) = 0$ we find $x(\xi) = \xi$. Choosing this value as the initial condition for the equation $x' = 2$, we determine $x(1) = \xi + 2(1–\xi)$. Equating the result 3/2, we find the control switching point $\xi = 1/2$. Thus, the solution to the problem is determined by the formula

$$u(t) = \begin{cases} 1, & \text{if } 0 < t < 1/2, \\ 2, & \text{if } 1/2 < t < 1. \end{cases}$$

The resulting function exactly coincides with the optimal control for Example 3.4, which is quite natural, since the latter minimizes this functional on the entire set U, and not just on its subset, which ensures the equality $x(1) = 3/2$.

Additional conclusions

Based on the study of examples of optimal control problems for systems with a fixed final state, carried out in Appendix, we can draw the following additional conclusions.

- For problems of optimal control of systems with a fixed final state, one can use the previously obtained assertions about the existence and uniqueness of a solution and the sufficiency of optimality conditions.
- The set of controls from a given set that guarantees the transfer of the system to a given state from Example 9.1 is convex, closed and bounded.
- The solution of the optimal control problem from Example 9.1 is unique, although the minimized functional for it is convex, but not strictly convex.

- The maximum principle for Example 9.1 gives the necessary and sufficient conditions for optimality.
- For Example 9.2, which differs from Example 9.1 only in the type of extremum, there is also a unique solution that can be found analytically.
- The maximum principle for Example 9.2 gives the necessary and sufficient conditions for optimality.
- The optimal control problem from Example 9.3, which differs from the one considered in Example 3.1 only by the presence of a pinned final state condition, admits an analytical solution that coincides with the optimal control from Example 3.1.
- The optimal control problem from Example 9.4, which differs from that considered in Example 3.2 only by the presence of a pinned final state condition, admits an analytical solution that coincides with the optimal control from Example 3.2.

Notes

1. To substantiate the maximum principle, it is required to indicate the functional properties of admissible controls. However, in this case, our goal is not maximum rigor. We only want to establish the form of the necessary optimality conditions for a problem with a fixed final state.

2. Note that there is no terminal term in the definition of the optimality criterion. This is due to the fact that the final state of the system in this case is given, and therefore does not change with a change in control.

3. Optimal control problems for distributed systems with a fixed finite state are considered; for example, in [25], [169].

4. In this case, we ignore the question of how to find a control that is an element of the set U and transfers the system from one known state to another, and whether it exists at all. The goal is to obtain relations that are satisfied by the optimal control.

5. Naturally, the penalty method can also be applied here; see Chapter 10.

6. Chapter 3 dealt with an optimal control problem with a free final state. As a result, the increment of the state function at the final moment of time was, generally speaking, different from zero. It is this circumstance that predetermined the presence of a final condition for the adjoint equation for systems with a free final state and its absence for systems with a fixed final state.

7. This is a consequence of fixing the final state of the system, i.e., condition (9.2)

8. Naturally, the assertions of the theorem must be rigorously substantiated. To do this, we can use the same technique as in Chapter 3.

9. Chapter 10 will show that applying the variational inequality for this example leads to the same result.

10. For this conclusion to be true, we must also verify that the optimal control problem has a solution. Recall the problem of minimizing the function $f(x) = x^3$ considered in Chapter 1, for which the unique solution of the stationary condition $x = 0$ is not the minimum point of this function.

11. Chapter 10 will consider other optimal control problems for systems with a fixed final state, the solution of which is also found analytically; see Examples 10.2 and 10.3. It is characteristic that there the maximum principle has three solutions, which turn out to be special controls, and only two of them are optimal. Chapter 13 will also provide an analytical solution to the optimal control problem for a system with isoperimetric constraints; see Example 13.1.

12. A similar situation was observed for Example 3.1.

13. On the shooting method for the practical solution of boundary value problems for differential equations; see [46], [65], [149], [154], [183]. In Chapter 13, the shooting method is extended to problems with a fixed final state in the presence of an additional isoperimetric condition. In Chapter 16, optimal control problems of systems with a free initial state will be considered. In this case, on the contrary, there is no initial condition for the equation, and for the adjoint equation there are two boundary conditions. To approximately solve the optimality conditions in this case, you can also use the shooting method.

14. We already used this method in Chapter 2 to approximate the problem of minimizing a function of one variable.

15. In the described algorithm, the fourth and fifth steps can be interchanged.

16. The general theory of systems ***controllability*** is considered in [66], [98], [150], [182], [196]. On controllability of systems with distributed parameters, see [18], [64], [66], [89], [105], [116]; for integral equations, see [7]; for integro-differential equations, see [18]; for lagging systems, see [80]; for stochastic systems, see [199]; for differential inclusions, see [45]. Examples of non-controllable systems are given in [73], [89].

17. Regarding the substantiation of the maximum principle for a wide class of systems described by ordinary differential equations; see [5], [74], [95], [152], [193].

18. In Chapter 5, we encountered a situation where a convex but not strictly convex functional has an infinite set of solutions. At the same time, convex but not strictly convex $f(x) = |x|$ has a unique minimum point.

19. The boundedness of the set U is obvious. We recall here that the boundedness of a subset of a normed space means the uniform boundedness of the norms of the elements of this set. In our case, all values of the functions under consideration lie on the interval [0,2], which implies the corresponding boundedness of the norms.

20. We remind you that a function from the space $L_2(0,1)$ is defined up to a set of zero measure. As a result, the presence of convergence $u_s(t) \to u(t)$ for almost all t is sufficient to pass to the limit in the last inequality and obtain the desired result.

21. In fact, such a statement is true not only for Hilbert spaces; see [94], [100], [106], [158]. However, in Hilbert spaces it is easily established using the Schwartz inequality given in Chapter 6. In particular, for any function v from the space under consideration, we have

$$|(u_k, v)-(u, v)| = |(u_k-u, v)| \leq \|u_k-u\| \, \|v\|.$$

Then the convergence of $u_k \to u$ in the sense of the $L_2(0,1)$ norm implies that $(u_k, v) \to (u, v)$ for all v from $L_2(0,1)$, i.e., corresponding weak convergence.

22. However, we already know this, since earlier a unique solution of the optimality conditions was found and it was proved that they are sufficient. Similarly, the existence of an optimal control will be proved in Chapter 13 for one problem with the isoperimetric condition.

23. Thus, the problem of minimizing the function $f(x) = x^2$ on the entire real line has a solution, but the problem of maximizing the same function has no solution. The solution of the minimization problem for the same function on the interval $[-1, 1]$ is unique, and the corresponding maximization problem has two solutions. The functional minimization problem from Example 3.3 has a unique solution, and the corresponding necessary optimality conditions are sufficient for it. At the same time, Example 5.1, which differs from Example 3.3 only in the type of extremum, is characterized by the presence of two optimal controls and an essentially insufficient optimality condition.

24. In Chapter 10, to solve this example, the variational inequality will also be applied, which in this case is not a sufficient condition for optimality.

25. Examples 9.1 and 9.2 consider the functional

$$I = \int_0^1 (u + x)dt.$$

The solution of the equations of state at a specific point t is equal to the integral of the control from zero to t. Therefore, the condition $x(1) = 1$ is equivalent to the unity of the integral of the control on the interval (0,1). Thus, in the optimality criterion, in fact, only the integral of the function x can vary. This value is actually equal to the area of the figure bounded by the curve $x = x(t)$. As can be seen from Figure 9.1 for Example 9.1 this area is minimal, and for Example 9.2 it is maximal.

26. This should not be surprising, since condition (9.16) is realized on this control. The optimal control from Example 3.1 delivers a minimum to this functional on the set U. Naturally, it also minimizes this functional on the subset of functions from U that guarantee the fulfillment of condition (9.16). Note that, by changing the type of extremum, it is easy to establish that the solution of the corresponding problem will be the optimal control from Example 3.2, which maximizes the given functional on the entire set U, and not only on its specified subset.

27. These statements can be justified in the same way as for Example 9.1. Moreover, one can establish a strong convexity of the functional, as was done in Chapter 8, and prove the well-posedness of the optimal control problem in the sense of Tikhonov.

CHAPTER 10

Addition

The previous chapter considers the optimal control problems for systems described by an ordinary differential equations with given initial and final states. To solve this problem, the maximum principle was used. However, in Part II, other methods were used to solve optimal control problems for systems with a free final state. In this chapter, these results are extended to the class of problems under consideration. The decoupling method, variational inequality and the penalty method will be considered. Appendix establishes the applicability of the Bellman optimality principle, considers the vector problem of optimal control of a system with a fixed final state, and solves two problems of this class that have practical meaning.

10.1 LECTURE

The subject of this lecture is the optimal control problem posed in the previous section, which is described by ordinary differential equations with given initial and final states. It is required to choose such a control, subject to some additional restrictions, which delivers a minimum to the integral functional of a general form. To solve it, the decoupling method described in Part II, the variational inequality and the penalty method are used.

10.1.1 Decoupling method

In Examples 9.1 and 9.2, the solution of the optimal control problem for systems with a fixed final state was found without using an iterative process. This is explained by the fact that both the state equation and the optimality criterion there were linear with respect to the state and control functions. In Examples 9.3 and 9.4, a similar result was achieved in the case of a functional that is quadratic with respect to the control. In the general case, to find an approximate solution to the problem, one can use the previously described shooting method. However, in Chapter 3 it was shown that for the analysis of general linear systems with a quadratic functional without restrictions on the control, iterative methods can be dispensed with. Similar method is also applicable for systems with a fixed final state.

DOI: 10.1201/9781003398585-10

We consider a system described by the following Cauchy problem

$$x'(t) \;=\; a(t)x(t) + b(t)u(t) + f(t),\; t \in (0,T);\; x(0) = x_0, \qquad (10.1)$$

where the functions a, b, f, and the number x_0 are known. The final state of the system is fixed, i.e., the following equality holds

$$x(T) \;=\; x_T. \qquad (10.2)$$

Consider the functional

$$I(u) \;=\; \frac{1}{2}\int_0^T \Big\{\alpha\big[x(t) - z(t)\big]^2 + \beta\big[u(t)\big]\Big\}dt,$$

where the function z and the positive constants α and β are known, and x is the solution to problem (10.1). We pose the ***linear-quadratic optimal control problem with a fixed final state***.

Problem 10.1 *It is required to find a function u that minimizes the functional I on the set of controls that guarantee the fulfillment of the equality* (10.2).

This problem is a partial case of Problem 9.1 and differs from Problem 3.2 only in the presence of the additional condition (10.2). In accordance with Theorem 9.1, define the function

$$H \;=\; p(ax + bu + f) - \big[\alpha(x{-}z)^2 + \beta u^2\big]/2.$$

Then the adjoint equation (9.6) takes the form

$$p'(t) \;=\; a(t)[x(t) - x(t)] - b(t)p(t),\; t \in (0,T). \qquad (10.3)$$

In contrast to problem (3.24), there is no final condition here. The solution of the maximum condition for the function H with respect to the control is the function

$$u(t) \;=\; \beta^{-1}b(t)p(t),\; t \in (0,T). \qquad (10.4)$$

The result obtained coincides exactly[1] with formula (3.25).

Thus, with respect to three unknown functions u, x, the system (10.1)–(10.4) is obtained. It differs from the system (3.23)–(3.25) obtained in Chapter 3 only by the presence of the final condition (10.2) for the state function and the absence of a similar condition for the function p. Substituting the control from formula (10.4) into the first equality (10.1), we have the Cauchy problem

$$x'(t) \;=\; \beta^{-1}b(t)^2p(t) + a(t)x(t) + f(t),\; t \in (0,T);\; x(0) = x_0 \qquad (10.5)$$

exactly coinciding with (3.26). The result is a system of two linear differential equations (10.3) and (10.5) with two boundary conditions (10.2).

When solving Problem 3.2, two linear differential equations with two boundary conditions were also obtained. It was assumed that the functions x and p are connected between some kind of linear transformation. The search for such a transformation was the basis for using the decoupling method and solving the problem under consideration without using an iterative process. Let us show that a similar technique is applicable for Problem 10.1.

Let us define the system state function by the formula

$$x(t) \;=\; r(t)p(t) + q(t),\; t \in (0,T), \tag{10.6}$$

where the functions r and q are chosen in such a way that the result is a solution to system (10.2), (10.3), and (10.5). In contrast to the similar formula (3.27), here we determine x as a function of p, and not vice versa. This is explained by the fact that in this case the condition at the final moment of time is given by the function x, and not p. Such a representation will later allow us to obtain final conditions for the functions r and q.

Substituting the function x from equality (10.6) into equation (10.5), we obtain

$$r'p + rp' + q' \;=\; \beta^{-1}b^2 p + a(rp+q) + f.$$

Taking into account equation (10.3), we have

$$r'p + r\big[\alpha(x-z) - bp\big] + q' \;=\; \beta^{-1}b^2 p + a(rp+q) + f.$$

Substituting here the value of the function x from equality (10.6), we get

$$\big[r' + \alpha r^2 - (a+b)r - \beta^{-1}b^2\big]p \;+\; \big(q' + \alpha rq - aq - \alpha zr - f\big) \;=\; 0.$$

Determining in equality (10.6) $t = T$ and taking into account condition (10.2), we have

$$r(T)p(T) + q(T) \;=\; x_T.$$

The last two relations represent the equalities to zero of some linear functions with respect to the function p, respectively, at an arbitrary and final time. These equalities can be satisfied if both the coefficients before p and the free terms in them vanish. As a result, we arrive at the following problems for the functions r and q

$$r'(t) + \alpha r(t)^2 - \big[a(t)+b(t)\big]r(t) \;=\; \beta^{-1}b(t)^2,\; t\in(0,T);\; r(T)=0; \tag{10.7}$$

$$q'(t) + \alpha r(t)q(t) - a(t)q(t) - \alpha z(t)r(t) \;=\; f(t),\; t\in(0,T);\; q(T)=x_T. \tag{10.8}$$

Note that, as in Problem 3.2, the equation for the function r includes a quadratic non-linearity, i.e., is the ***Riccati equation***, and the equation for the function q is linear.

Based on the results obtained, we arrive at the following algorithm for solving the considered problem:

1. The Cauchy problem (10.7) is solved in the reverse direction of time for the Riccati differential equation with respect to the function r.

2. The Cauchy problem (10.8) is solved in the reverse direction of time for a linear differential equation with respect to the function q.

3. After substituting the function $p = r^{-1}(x - q)$ into problem (10.5), we find the function x from the Cauchy problem for the linear differential equation

$$x'(t) = [\beta^{-1}b(t)^2 r(t)^{-1} + a(t)]x(t) + [f(t) - \beta^{-1}b(t)^2 r(t)^{-1}q(t)], \; t \in (0,T); \; x(0) = x_0. \tag{10.9}$$

4. Using formulas (10.4) and (10.6), we find the control[2]

$$u(t) = \beta^{-1}b(t)r(t)^{-1}[x(t) - q(t)], \; t \in (0,T). \tag{10.10}$$

   Thus, the solution of the considered optimal control problem can be found explicitly without using the iterative process[3].

As an application, consider the following example.

Example 10.1 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t), \; t \in (0,1); \; x(0) = 0,$$

satisfying the condition $x(1) = 0$.

This example differs from Example 3.4 only in the presence of an additional condition at the end time. We have Problem 10.1 with the following parameter values

$$a = 0, \; b = 1, \; f = 0, \; x_0 = 0, \; z = 0, \; \alpha = 1, \; \beta = 1, \; T = 1.$$

Problems (10.7) and (10.8) here have the following form

$$r'(t) + r(t)^2 - r(t) = 1, \; t \in (0,1); \; r(1) = 0,$$

$$q'(t) + r(t)q(t) = 0, \; t \in (0,1); \; q(1) = 0.$$

Note that the last of them has a zero solution. Then problem (10.9) is transformed to the form

$$x'(t) = r(t)^{-1}x(t), \; t \in (0,1); \; x(0) = 0,$$

and also has a zero solution. Then it follows from formula (10.10) that $u = 0$. This corresponds to the optimal control in Example 3.4, which satisfies the condition $x(1) = 0$.

10.1.2 Variational inequality

In Part II, to study optimal control problems, along with the maximum principle, we also used the optimality condition in the form of a variational inequality. Let us show that, in principle, it can also be used to analyze systems with a fixed final state. For simplicity, we restrict ourselves to the examples described in the previous chapter.

In Example 9.1, the problem is to minimize the functional

$$I(u) = \int_0^1 (u+x)dt,$$

where x is a solution of Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0,$$

on the set of functions u satisfying the inequality $0 \le u(t) \le 2$, $t \in (0,1)$ and ensuring the fulfillment of the condition $x(1) = 1$.

Let u be a solution to this problem and x the corresponding solution to the Cauchy problem. Then the given functional I satisfies the inequality

$$I(v,y) - I(u,x) \ge 0,$$

where v is an arbitrary admissible control and y is the corresponding solution to the Cauchy problem. We define the Lagrange functional

$$L = \int_0^1 \big[(u+x) + p(x'-u)\big]dt$$

with arbitrary function p. Then the following inequality holds

$$\int_0^1 (1-p)(v-u)dt + \int_0^1 (\Delta x + p\Delta x')dt \ge 0,$$

where $\Delta x = y–x$. After integrating by parts, taking into account that the initial and final states of the system do not depend on the control, we obtain

$$\int_0^1 (1-p)(v-u)dt + \int_0^1 (1-p')\Delta x dt \ge 0.$$

Here, we choose the function p so that it satisfies the condition

$$p'(t) = 1,\ t \in (0,1),$$

which corresponds to the adjoint equation. Then the previous inequality takes the form

$$\int_0^1 (1-p)(v-u)dt \ge 0.$$

Choosing here as v the needle variation of the control, taking into account the constraints on the control, as was done in Chapter 3, we obtain the variational inequality

$$[1-p(t)]\,[w-u(t)] \;\geq\; 0 \;\; \forall w \in [0,2].$$

Obviously, it is equivalent to the maximum condition (9.11), and hence leads to the same result

$$u(t) \;=\; \begin{cases} 0, & \text{if } \; t < 1/2, \\ 2, & \text{if } \; t > 1/2. \end{cases}$$

Thus, the maximum principle and the variational inequality for Example 9.1 are equivalent[4].

Example 9.2 differs from the one considered above only in the type of extremum. Then the variational inequality for it is obtained by replacing the relation $\geq$ in the previous inequality with $\leq$. The corresponding optimality condition

$$[1-p(t)]\,[w-u(t)] \;\leq\; 0 \;\; \forall w \in [0,2]$$

is also identical to the maximum principle for this example and has a solution

$$u(t) \;=\; \begin{cases} 2, & \text{if } \; t < 1/2, \\ 0, & \text{if } \; t > 1/2. \end{cases}$$

Consider now Example 9.3, which consists in minimizing the functional

$$I(u) \;=\; \int\limits_0^1 \Big(\frac{u^2}{2} - 3x\Big)dt,$$

where x is a solution of Cauchy problem

$$x'(t) \;=\; u(t), \; t \in (0,1); \; x(0) = 0$$

on the set of functions u satisfying the inequality $1 \leq u(t) \leq 2$ for $0 < t < 1$ and guaranteeing the fulfillment of the condition $x(1) = 3/2$. Let u be a solution to this optimal control problem, and x be the corresponding state of the system. Then we get the inequality

$$I(v,y) - I(u,x) \;\geq\; 0,$$

where v is an arbitrary admissible control and y is the corresponding solution to the Cauchy problem. We define the Lagrange functional

$$L \;=\; \int\limits_0^1 \Big[\Big(\frac{u^2}{2} - 3x\Big) + p(x' - u)\Big]dt.$$

Then we get the inequality

$$\int\limits_0^1 \Big[\Big(\frac{v^2 - u^2}{2} - 3\Delta x\Big) + p\Delta x' - p(v-u)\Big]dt \;\geq\; 0.$$

We choose the function p so that it satisfies the condition

$$p'(t) = -3, \; t \in (0,1).$$

Then the previous inequality takes the form

$$\int_0^1 \Big[\frac{v^2-u^2}{2} - p(v-u)\Big]dt \geq 0.$$

It is easy to show that the set of controls that satisfy both given constraints is convex[5]. Then, using the convex variation of the control, i.e., $v = u+\sigma(s-u)$, where $\sigma \in (0,1)$, and s is an arbitrary control that satisfies the given constraints, we have

$$\sigma \int_0^1 \big[(u-p)(s-u) + \sigma(s-u)^2/2\big]dt \geq 0.$$

After dividing by σ and passing to the limit at $\sigma \to 0$, we obtain the variational inequality

$$\int_0^1 (u-p)(s-u)dt \geq 0 \;\; \forall s.$$

Choosing here as s the needle variation and passing to the limit taking into account the mean value theorem, as has already been done repeatedly, we obtain the variational inequality

$$[p(t)-u(t)]\,[w-u(t)] \,\forall \leq 0 \;\; \forall w \in [1,2],$$

exactly the same as corresponding condition obtained in Chapter 4 for Example 3.1. Its solution, as is already known, has the form

$$u(t) = \begin{cases} 2, & \text{if } 0 < t < 1/3, \\ 3-3t, & \text{if } 1/3 \leq t \leq 2/3, \\ 1, & \text{if } 2/3 < t < 1. \end{cases}$$

The resulting relation coincides with formula, established for the example under consideration using the maximum principle. Thus, the variational inequality for Example 9.3 again turns out to be equivalent to the maximum principle.

Let us now turn to Example 9.4, in which it is required to find the maximum of the functional from Example 9.3. The corresponding variational inequality for it is obtained from the previous one by replacing the sign $\leq$ with $\geq$, i.e., looks like coinciding with the similar inequality obtained in Chapter 4 for Example 3.2. It was shown that $u(t) = 1$ for $p(t) > 1$, $u(t) = 2$ for $p(t) < 2$, and $u(t) = p(t)$ for $p(t) \in [1,2]$. Thus, for $p(t) \in [1,2]$ the solution of this variational inequality can take three values at once 1, 2, and $p(t)$. Now the set of its solutions turns out to be wider than the set of solutions of the maximum principle, i.e., the latter is more efficient. Thus, the relationship between the considered optimality conditions for a system with a fixed final state remains the same as for a similar system with a free final state.

10.1.3 Penalty method

In Chapters 2 and 4, the penalty method was used to solve extremal problems with constraints in the form of equalities. It is quite natural to try to use it to solve problems with a fixed final state. In this case, it is the additional condition at the final moment of time that can be eliminated by introducing an additional term into the optimality criterion with a small parameter.

Let us return to the consideration of the general Problem 9.1. It consists in finding such a control u from the set

$$U = \{u \big| \; a(t) \leq u(t) \leq b(t), \; t \in (0,T)\},$$

guaranteeing the fulfillment of the condition

$$x(T) = x_T$$

for a function x that is a solution to the problem

$$x'(t) = f(t,u(t),x(t)), \; t \in (0,T), \; x(0) = x_0$$

while minimizing the functional

$$I = \int_0^T g(t,u(t),x(t))dt.$$

In accordance with the penalty method, the functional is define the functional

$$I_\varepsilon = \int_0^T g(t,u(t),x(t))dt + \frac{1}{2\varepsilon}\big[x(T) - x_T\big],$$

where ε is a small positive parameter. The problem of minimizing the functional I_ε on the set U is a partial case of the previously considered Problem 3.1. In this case, the second term on the right side of this functional corresponds to the terminal term of the optimality criterion. To solve the problem obtained, we use the method described earlier.

We define

$$H(t,u,x,p) = pf(t,u,x) – g(t,u,x).$$

The solution of the adjoint system is chosen as the function p

$$p'(t) = -H_x(t,u(t),x(t),p(t)), \; t \in (0,T), \; p(T) = \varepsilon^{-1}[x_T - x(T)].$$

The optimal control is now determined from the maximum condition

$$H(t,u(t),x(t),p(t)) = \max_{v \in [a(t),b(t)]} H(t,v,x(t),p(t)), \; t \in (0,T).$$

The resulting system of optimality conditions is solved iteratively, as was done in Part II. The result for small enough values of the parameter ξ is chosen as an approximate solution of Problem 9.1.

In particular, for Example 9.1, the corresponding adjoint system has the form

$$p'(t) \;=\; 1,\; t \in (0,1),\; p(1) = \varepsilon^{-1}[1 - x(1)].$$

The function H here is equal to pu–u–x, i.e., is defined in the same way as in Chapter 3. Accordingly, the solution of the maximum condition is again calculated by the formula

$$u(t) \;=\; \begin{cases} 2, & \text{if } p(t) > 1, \\ 0, & \text{if } p(t) < 1. \end{cases}$$

Find a solution to the adjoint system

$$p(t) \;=\; t - 1 + \varepsilon^{-1}[1 - x(1)].$$

Note that the function p is monotonically increasing. Therefore, in principle, three cases are possible. Either the function p is always greater than 1, or it is always less than 1, or it is first less than and then greater than 1. Let us consider these three cases.

Assume that $p(t)$ is always greater than 1. Then we have a control $u(t) = 2$ for all t. Therefore, $x(t) = 2t$, and hence $x(1) = 2$. In this case, the solution of the adjoint system is $p(t) = t$–1–ε^{-1}. For sufficiently small ε, the function p certainly takes a value less than 1, which contradicts the assumption made, which means that this case is not realized at all.

Suppose now that $p(t)$ is always less than 1. Then we have the control $u(t) = 0$ for all t. Therefore, $x(t) = 0$, and hence $x(1) = 0$. In this case, the solution of the adjoint system $p(t) = t$–$1+\varepsilon^{-1}$. For sufficiently small ε, the function p certainly takes on a value greater than 1, which again contradicts the assumption made. Therefore, this case is also not implemented.

Now let there be a point ξ from the interval (0,1) at which the increasing function p is equal to one. Then the control $u(t)$ is equal to 0 for $t < \xi$ and equal to 2 for $t > \xi$. We find the state function $x(t) = 0$ for $t < \xi$, which means $x(\xi) = 0$. For $t > \xi$ we have $x(t) = 2(t$–$\xi)$, and then $x(1) = 2(1$–$\xi)$. As a result, we find $p(t) = t$–$1 + \varepsilon^{-1}(2\xi$–$1)$. According to the assumption made, the equality $p(\xi) = 1$ should hold. Having defined $t = \xi$, in the previous formula, we have $p(\xi) = \xi$–$1 + \varepsilon^{-1}(2\xi$–$1) = 1$. From here, we find the point

$$\xi \;=\; \frac{2+\varepsilon^{-1}}{1+2\varepsilon^{-1}} \;=\; \frac{1+2\varepsilon}{2+\varepsilon}.$$

For small enough values of ε, the point ξ really lies on the interval (0,1), which means that the function

$$u_\varepsilon(t) \;=\; \begin{cases} 0, & \text{if } t < \frac{1+2\varepsilon}{2+\varepsilon}, \\ 2, & \text{if } t > \frac{1+2\varepsilon}{2+\varepsilon} \end{cases}$$

is the unique solution to the system of optimality conditions for the problem of minimizing the functional I_ε. It is the minimum point of this functional[6].

Passing to the limit in the last equality as $\varepsilon \to 0$, we find the function

$$u_\varepsilon(t) = \begin{cases} 0, & \text{if } t < 1/2, \\ 2, & \text{if } t > 1/2 \end{cases}$$

which we know is the optimal control for Example 9.1. Thus, using the penalty method, we can find not only an approximate, but also an exact solution to this problem[7].

RESULTS

Here is a list of questions in the field of various methods for solving problems of optimal control of systems with a fixed final state, the main conclusions on this topic, as well as the problems that arise in this case and require additional research.

Questions

It is required to answer questions concerning optimal control problems for systems with a fixed final state and the optimality condition for such problems.

1. Why did the decoupling method turn out to be applicable for the linear-quadratic problem of optimal control of systems with a fixed final state?
2. What is the difference between the decoupling method for the systems with free and fixed final state?
3. Why is the function p expressed in terms of x when using the decoupling method for systems with a free final state, and vice versa for systems with a fixed final state?
4. What is the difference between the algorithm for finding the optimal control in accordance with the decoupling method for systems with a free and a fixed final state?
5. Can the decoupling method be used to solve the problem in Example 9.1?
6. What is the reason for the possibility of applying the variational inequality in the study of Examples 9.1–9.4?
7. Whence follows the convexity of the set of controls that satisfy all the given constraints for the considered examples?
8. Why, for the considered examples, the variational inequality is sometimes a sufficient condition for optimality, and sometimes it is not?
9. Why, for the considered examples, is the variational inequality sometimes equivalent to the maximum principle, and sometimes not equivalent?

10. Can the variational inequality be applied to the analysis of Example 10.1?
11. What is the point of using the penalty method for solving optimal control problems for systems with a fixed final state?
12. Why, unlike the problems considered earlier, for Example 9.1, using the penalty method, not only an approximate, but also an exact solution of the problem is found?
13. Can the penalty method be applied to the analysis of Examples 9.2 and 10.1?
14. Which of the methods of analysis of Example 9.1 seems to be more effective?

Conclusions

Based on the study of the optimal control problem for systems with a fixed final state, we can come to the following conclusions.

- The decoupling method is applicable for solving the linear-quadratic optimal problem with a fixed final state.
- When using the decoupling method for a linear-quadratic optimal problem with a fixed final state, the state function is determined through the solution of the adjoint equation, and not vice versa, as for systems with a free final state.
- The decoupling method allows solving the synthesis problem.
- As a result of applying the decoupling method, the Riccati equation is obtained.
- The use of variational inequalities for solving the optimal problem with a fixed final state is possible if the set of controls that satisfy the given constraints and guarantee the transfer of the system to the given final state is convex.
- The variational inequalities for Examples 9.1, 9.2, and 9.3 are equivalent to the maximum principle and are necessary and sufficient conditions for optimality.
- The variational inequality for Example 9.4 is not equivalent to the maximum principle and are not sufficient optimality conditions due to the non-convexity of the optimality criterion.
- The penalty method allows one to reduce a problem with a fixed final state to a problem with a free final state.
- The penalty method allows one to find an approximate solution to an optimal control problem with a fixed final state.
- The penalty method for Example 9.1 allows you to find the exact value of the optimal control by passing to the limit in the formula for solving the problem of minimizing the penalty functional.

Problems

In the process of analyzing optimal control problems for systems with a fixed final state, additional problems arise that require additional research.

1. **Bellman principle**. In Chapter 4, to solve the problem of optimal control with a free final state, dynamic programming was used, which is based on the Bellman principle. At the same time, the applicability of this statement for a concrete example was proved. In Appendix, a similar result will be obtained for one system with a fixed final state.

2. **Penalty method and alternative methods**. The penalty method made it possible to reduce a problem with a fixed final state to some problem with a free final state. For the latter one can use not only the maximum principle, but also other methods. These results are obtained in Appendix.

3. **Vector case**. The above results are naturally generalized to the case when both the state function and the control are vector quantities. In this case, in the state equation, the adjoint equation, and the maximum condition, u, x, and p are vector quantities. Such a problem is considered in Appendix.

4. **Practical applications**. Interest in optimal control problems for systems with a fixed final state is largely due to their practical significance. In this regard, these problems with real practical meaning are of interest. In Appendix, two such problems are solved, one of which relates to geometry, and in mechanics.

5. **Insufficiency of the maximum condition for systems with a fixed final state**. The maximum principle for the considered examples gives a necessary and sufficient optimality condition. However, in Part II, examples of problems with a free final state with an insufficient optimality condition were given. It is of interest to consider the problem of optimal control of a system with a fixed final state, for which the maximum principle turned out to be a necessary but not sufficient optimality condition. Such an example is given in Chapter 11.

6. **Non-uniqueness of the solution in optimal control problems for systems with a fixed final state**. The optimal control for the considered examples turned out to be unique. However, in Part II, examples of problems with a free final state with a non-unique solution were given. It is of interest to consider the problem of optimal control of a system with a fixed final state, the solution of which would not be unique. Such an example is given in Chapter 11.

7. **Absence of a solution in optimal control problems for systems with a fixed final state**. The solution of optimal control problems in all considered examples exists. However, in Part II, examples of unsolvable problems with a free final state were given. It is of interest to consider the problem of optimal control of a system with a fixed final state, the solution of which does not exist. Such an example is given in Chapter 11.

8. **Optimal control problems for systems with a fixed final state with singular controls**. The maximum principle does not degenerate for all the considered examples. However, in Part II, examples of problems with a free final state were given, for which the solutions of the maximum condition turned out to be special controls. It is of interest to consider the problem of optimal control of a system with a fixed final state, in which the maximum principle degenerates. Such an example is given in Chapter 11.

9. **Well-posedness of optimal control problems for systems with a fixed final state**. When studying optimal control problems for systems with a free final state, we encountered various manifestations of their ill-posedness. Examples of ill-posed optimal control problems for a system with a fixed final state are given in Chapter 12.

10.2 APPENDIX

Below, we present some additional results in the field of optimal control theory for systems with a fixed final state. In particular, Section 10.2.1 shows the applicability of Bellman optimality principle to an example of an optimal control problem for a system with a fixed final state. Section 10.2.2 solves one geometric problem of the type under consideration. Section 10.2.3 shows that the penalty method can be used in conjunction with various derivative-based methods. In Section 10.2.4, the results obtained earlier are extended to the case when both the control and the state of the system are vector quantities. As an application, Section 10.2.5 investigates the time optimal problem.

10.2.1 Applicability of Bellman optimality principle

The analysis of optimal control problems for systems with a free final state in Part II was also carried out using dynamic programming. It is based on the Bellman optimality principle. It was shown in Chapter 4 that it holds for one particular example. Let us verify the validity of the Bellman optimality principle for an optimal control problem for a system with a fixed final state, in particular, for Example 9.1.

In this example, it was required to minimize the functional

$$I(u) = \int_0^1 (u+x)dt,$$

where x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

on the set of functions u satisfying the inequality $0 \le u(t) \le 2$ for $t \in (0,1)$ and ensuring the equality $x(1) = 1$. The solution to this problem, as is already known, has the form

$$u(t) = \begin{cases} 0, & \text{if } t < 1/2, \\ 2, & \text{if } t > 1/2. \end{cases} \tag{10.11}$$

The corresponding state of the system is determined by the formula

$$x(t) = \begin{cases} 0, & \text{if } t < 1/2, \\ 2t-1, & \text{if } t > 1/2. \end{cases} \tag{10.12}$$

Consider a family of problems that differ from the one above only in the choice of the initial time ξ and the initial state x. In this case, it is required to minimize the functional

$$I_\xi^x(v) = \int_0^1 (v+y)dt,$$

where y is a solution of the Cauchy problem

$$y'(t) = v(t),\ t \in (\xi, 1);\ y(0) = x$$

on the set of functions v satisfying the inequality $0 \leq v(t) \leq 2$ for $\xi < t < 1$ and ensuring the equality $y(1) = 1$.

In accordance with the maximum principle, we introduce the function

$$H = pv–(v+y),$$

where p satisfies the adjoint equation

$$p'(t) = 1,\ t \in (\xi, 1).$$

In this case, the corresponding optimal control is found from the condition of the maximum of the function H with respect to the control. The result is the formula

$$v(t) = \begin{cases} 2, & \text{if } p(t) > 1, \\ 0, & \text{if } p(t) < 1, \end{cases}$$

similar to the one that was obtained in Chapter 9 when examining this example.

The adjoint equation has a solution $p(t) = t + c$, where c is an arbitrary constant. Thus, the function p is increasing. There are three possible options here. Either the function p in is always greater than 1, in which case the control v is identically equal to 2; or p is always less than 1, in which case the control is identically equal to 0; or at some point in time from the interval (0,1), the function p, increasing, reaches unity, and then before this time the control is equal to 0, and after that it is equal to 2.

For $v = 0$, the considered Cauchy problem has a solution $y(t) = x$. The condition $y(1) = 1$ is possible only for the initial state $x = 1$. However, as can be seen from formula (10.12), the optimal state of the system for Example 9.1 reaches the value 1 only at the final time. Therefore, this case is of no interest.

For $v = 2$, the Cauchy problem has a solution $y(t) = x + 2(t - \xi)$. The equality $y(1) = 1$ is admissible if the initial time and the initial state of the system are related by the equality $x = 2\xi–1$. We turn to formula (10.12), which characterizes the optimal state of the system under study. We see that $x(t) = 2t–1$ for $t > 1/2$, which corresponds to the previously obtained equality. Thus, when $\xi > 1/2$, as an optimal control on the interval $(\xi, 1)$, one can choose a function that is identically equal to 2, which is consistent with formula (10.11).

Now let there be a point η from the interval $(\xi, 1)$ such that $v(t) = 0$ for $t < \eta$ and $v(t) = 2$ for $t > \eta$. Then $y(\eta) = x$ and $y(1) = x+2(1–\eta)$. Equating the last value to one, we find $\eta = (x+1)/2$. Thus, if at time ξ the system is in state x, then at $t < (x+1)/2$ the control takes the value 0, and at $t > (x+1)/2$ it takes the value 2. Here, we should distinguish between two cases: whether the starting point ξ is before or after the switching point 1/2 of the optimal control for Example 9.1; see formula (10.11).

For $\xi < 1/2$, the state of the system at this point in time, according to formula (10.12), is equal to $x = 0$. Then from the formula $\eta = (x + 1)/2$ it follows that the control switching point is equal to $\eta = 1/2$. Therefore, the optimal control v on the section $(\xi, 1)$ coincides with the final part of the optimal control determined by the formula (10.11).

For $\xi > 1/2$, the state of the system at this point in time, according to formula (10.12), is equal to $x = 2\xi–1$. Then the point η starting from which the control takes the value 2 is equal to $\eta = (x + 1)/2 = \xi$. Hence, in this case, $v(t) = 2$ for $t > \xi$, which is consistent with the optimal control for Example 9.1.

Thus, the optimal control for this example always remains optimal when considering the problem on any finite interval $(\xi, 1)$, i.e., the Bellman optimality principle is indeed satisfied[8].

10.2.2 Shortest Curve

Consider now a geometric problem, which is a special case of Problem 9.1.

Example 10.2 *It is required to find the curve $x = x(t)$ passing through two given points and having the smallest length.*

We select the coordinate system in such a way that one of these points is at the origin, which corresponds to the initial condition $x(0) = 0$. Denote by (T, X) the coordinates of the second given point, which corresponds to the final condition $x(T) = X$. For a complete statement of the problem, it is required to obtain a formula characterizing the length of the curve.

Consider a part of the curve $x = x(t)$, bounded by points M with coordinates (t, x) and N with coordinates $(t+\Delta t, x+\Delta x)$; see Figure 10.1. For sufficiently small values of Δt, the length of the arc of the curve connecting these points is close enough to the length of the segment Δs connecting them. The latter is calculated using the Pythagorean theorem

$$\Delta s = \sqrt{\Delta t^2 + \Delta x^2} = \sqrt{1 + \Big(\frac{\Delta x}{\Delta t}\Big)^2}\Delta t.$$

This implies the formula for an infinitely small element of length[9]

$$ds = \sqrt{1 + x'(t)^2}dt.$$

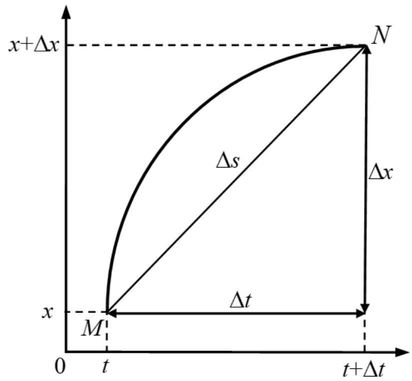

Figure 10.1 Arc length calculation.

To calculate the length of the arc of the curve connecting two initially given points, one should integrate this equality[10]. We get the equality

$$S = \int_0^T \sqrt{1 + x'(t)^2}dt.$$

To reduce the problem of minimizing the arc length of a curve connecting two given points to Problem 9.1, we denote the derivative of the function x by u. As a result, we obtain the problem of minimizing the functional[11]

$$I(u) = \int_0^T \sqrt{1 + u(t)^2}dt.$$

on the set of all functions satisfying the condition $x(T) = X$, where $x = x(t)$ is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0, T);\ x(0) = 0.$$

To solve the problem under consideration, we use the previously described method[12]. Define a function

$$H = pu - \sqrt{1 + u^2}.$$

Considering that it does not depend on x, we conclude that the adjoint equation (9.6) here has the form $p' = 0$. This implies that p is a constant c.

Let us now turn to the maximum condition (9.7). We draw attention to the absence of explicit restrictions on control. As a result, we have a problem for the unconditional maximum of the function H. We vanish the derivative of this function. We get

$$H_u = c - \frac{u}{\sqrt{1 + u^2}} = 0.$$

After squaring, we find

$$u^2 = \frac{c^2}{1-c^2}.$$

The value on the right side is some constant. Then the modulo of the function u is equal to a constant. However, it follows from the preceding equality that u and c have the same sign. Therefore, the function u cannot change sign, and hence it is some constant c_1. It is easy to verify that the second derivative of the function H is negative, which means that the function $u(t) = c_1$ is indeed the solution of the maximum[13].

To find the unknown constant, we turn to the equation of state. Obviously, the solution to the existing Cauchy problem for a given control is determined by the formula $x(t) = c_1 t$. This implies that $x(T) = c_1 T$. However, according to the condition of the problem, this value is equal to X. As a result, we find a constant and determine the optimal control $u(t) = X/T$. It corresponds to the function $x(t) = Xt/T$. Thus, the curve of least length connecting two fixed points is a straight line passing through these points, which is quite natural.

The considered example is connected in some way with the most amazing mathematical object that is the Cantor function[14]. For simplicity, we confine ourselves to considering the case $X = T = 1$. On the segment [0,1], we construct some set M using the following procedure. At the first step, we include in it the middle third of the segment under consideration – the interval (1/3,2/3). At the second step, we add to M the middle thirds of the remaining segments – the intervals (1/9,2/9) and (7/9,8/9). Next, we complement M with the middle thirds of the remaining segments. Continuing this process ad infinitum, we obtain the desired set M. The ***Cantor set*** C is the complement in [0,1] to the resulting set; see Figure 10.2. This set is surprising in that it has the cardinality of the continuum[15] and zero measure[16].

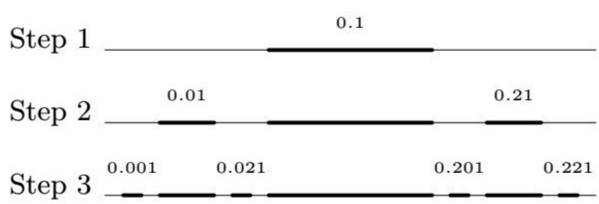

Figure 10.2 Construction of the complement to the Cantor set.

Let us now set the function $x = x(t)$, setting it equal to 1/2 on the interval (1/3,2/3), 1/4 on the interval (1/9,2/9), 3/4 on the interval (7/9,8/9), 1/8 on the interval (1/27,2/27), etc.; see Figure 10.3. This defines a certain function on the previously indicated set M. Extending it by continuity to the Cantor set, we obtain the ***Cantor function***, which turns out to be continuous and non-decreasing on the unit interval.

Note that the Cantor function is constant on each interval that makes up the set S, i.e., outside the Cantor set, which has zero measure. Consequently, its derivative

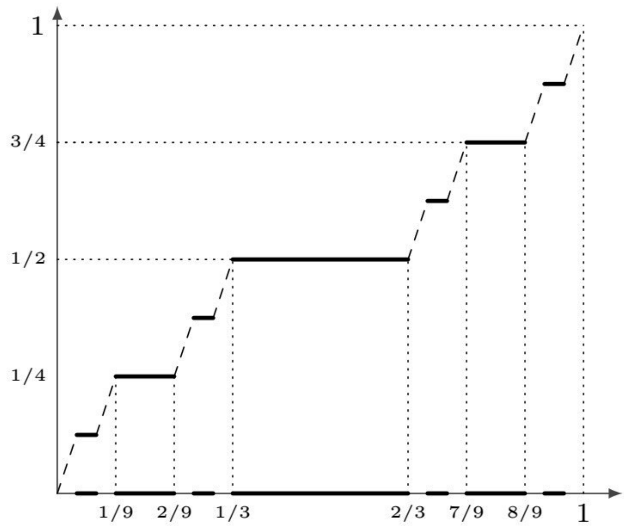

Figure 10.3 Cantor function.

$v = x'$ vanishes almost everywhere on the unit interval. In addition, this function satisfies the boundary conditions from Example 10.2 at $X = T = 1$. In this regard, there is a temptation to pass it off as a solution to the corresponding problem, considering $I(v) = 1$, while the value of the functional on the previously found optimal control is clearly greater and equal to $\sqrt{2}$ for $X = T = 1$. However, the derivative of the Cantor function is not defined on the entire interval [0,1], as a result of which we cannot substitute it into the optimality criterion.

10.2.3 Penalty method with functional differentiation

It was shown in the Lecture that, using the penalty method, Problem 9.1 can be reduced to an optimal control problem with a free final state. In particular, the problem of minimizing the functional

$$I_\varepsilon(u) = \int_0^T g(t, u(t), x(t))dt + \frac{1}{2\varepsilon}\left[x(T) - X\right]^2$$

with a small positive parameter ε on the set

$$U = \{u \mid a(t) \le u(t) \le b(t),\ t \in (0, T)\},$$

where x is a solution of the Cauchy problem

$$x'(t) = f(t, u(t), x(t)),\ t \in (0, T),\ x(0) = x_0.$$

To solve it, the maximum principle described in Part II was applied. However, as was shown in Chapter 4, similar problems can be solved by other methods.

We note, in particular, methods related to the differentiation of functionals. As is known, if u is a minimum point of a differentiable functional I on a convex set U, then the variational inequality

$$I'(u)(v–u) \geq 0 \;\; \forall v \in U,$$

where $I'(u)$ is the derivative of this functional at the considered point. For an approximate solution of this problem, one can use the gradient projection method

$$u_{k+1} \;=\; P\big[u_k – \beta_k I'(u_k)\big],$$

where k is the iteration number, the positive constant β_k is the algorithm parameter, and P is the projection operator onto the set of admissible controls. Note that the direct application of these results for Problem 9.1 in the general case is not possible, since the set of admissible controls there is characterized not only by explicit restrictions on the values of the control, but also by fixing the final state of the system. It is possible to establish the convexity of this set only in exceptional cases[17]. At the same time, the minimization of the penalty I_ε obtained as a result of applying the method is carried out directly on a given set U, which is convex.

As an illustration, let us return to the consideration of Example 10.2, in which it is required to minimize the functional

$$I(u) \;=\; \int\limits_0^T \sqrt{1+u^2}dt$$

on the set of all functions ensuring the fulfillment of the condition $x(T) = X$, where $x = x(t)$ is a solution of the Cauchy problem

$$x'(t) \;=\; u(t),\; t \in (0,T),\; x(0) = 0.$$

The corresponding functional I_ε is characterized by the equality

$$I_\varepsilon(u) \;=\; \int\limits_0^T \sqrt{1+u^2}dt \;+\; \frac{1}{2\varepsilon}\big[x(T) - X\big]^2.$$

Because of the absence of explicit restrictions on the control, the variational inequality here reduces to the stationary condition (equality to zero of the derivative of the functional at its minimum point), and the gradient projection method to the usual gradient method (the projection operator is a unitary operator).

Let us find the derivative of the considered functional. We recall that the Gateaux derivative of a functional J defined on a normed vector space V at a point u is a linear continuous functional $J'(u)$ that satisfies the relation

$$\lim_{\sigma\to 0}\frac{J(u+\sigma h) - J(u)}{\sigma} \;=\; J'(u)h \;\; \forall h \in V.$$

It is easy to establish that the derivative of the functional I_ε at the point u is equal to[18]

$$I'_\varepsilon(u) = \frac{u}{\sqrt{1+u^2}} + p,$$

where p is a solution to the problem

$$p'(t) = 0,\ t \in (0,T),\ p(T) = \frac{1}{\varepsilon}\big[y(T) - X\big].$$

Because of the absence of explicit restrictions on the control, we use the stationary condition, equating the derivative of the functional to zero. We get

$$\frac{u}{\sqrt{1+u^2}} + p = 0.$$

Let us compare these results with those obtained in the previous subsection. The function x in both cases is a solution to the same Cauchy problem. The last equality exactly coincides with what was established earlier when using the maximum principle. The adjoint equation is the same in both cases. The only difference has to do with the end condition. Previously, it was set for the function x and had the form $x(T) = X$. In this case, there is a final condition for the function p. However, it can be written as $\varepsilon p(T) = x(T)–X$. Note that for sufficiently small values of ε, the term on the left side of the last equality will be arbitrarily close to zero, which means that this relation can be interpreted as an approximate final condition for the state function. Thus, there is reason to believe that as a result of applying the penalty method, we obtain an approximate solution of this problem.

Indeed, since the stationary condition and the adjoint equation have the same form as in the previous subsection, we can again conclude that the control is constant, i.e., $u(t) = c_1$. The equation of state also has the same form as before, which means that the function x is linear, i.e., $x(t) = c_1 t + c_2$. Taking into account the zero initial condition, we conclude that $c_2 = 0$. Another constant was previously found from the end condition for the function x. In this case, there is a similar condition for the function p. This function itself is connected with the control of the stationary condition. We get the equality

$$p = -\frac{c_1}{\sqrt{1+c_1^2}}.$$

Thus, the constant c_1 is found from the equality

$$-\frac{c_1}{\sqrt{1+c_1^2}}\varepsilon = c_1 T - X.$$

Obviously, for small ε the left side of the equality is arbitrarily small, and hence $c_1 \approx X/T$, which is consistent with the result obtained earlier. Naturally, for $\varepsilon \to 0$, the equality $c_1 = X/T$ is true, which means that the desired curve is characterized by the formula $x(t) = tX/T$.

10.2.4 Vector optimal control problem with a fixed finite state

When analyzing optimal control problems for systems with a free final state, we actually limited ourselves to considering the scalar case, when there was one control and one state function. However, in the final part of Chapter 3, a general scheme for solving problems with many controls and system states was described. We proceed in the same way for systems with a fixed final state.

We assume that the control and the state functions are vector functions

$$u = (u_1, u_2, \dots, u_r), \;\; x = (x_1, x_2, \dots, x_n).$$

In this case the state system

$$x'(t) = f(t, u(t), x(t)), \; t \in (0, T); \; x(0) = x_0$$

is a Cauchy problem for a system of differential equations, where f is an nth order vector function of $r+n+1$ variables, and x_0 is some n-dimensional vector. The control u belongs to the set

$$U = \{u | \; u(t) \in G(t), \; t \in (0, T)\},$$

where $G(t)$ is some subset of the r-dimensional Euclidean space. The final state of the system is fixed, i.e., we have the equality

$$x(T) = x_T.$$

The optimality criterion is determined by the formula

$$I = \int_0^T g(t, u(t), x(t))dt,$$

where g is a function of $r + n + 1$ variables. We have the following ***vector optimal control problem with a fixed final state***[19].

Problem 10.2 *Find a vector function u that minimizes the functional I on sets U for a fixed final state of the system.*

To solve this problem, a function of $r + 2n + 1$ variables is defined

$$H(t, u, x, p) = \langle p, f(t, u, x)\rangle - g(t, u, x),$$

where on the right side of the equality is the dot product of the corresponding vectors. In accordance with the ***maximum principle***, the optimal control satisfies the condition

$$H(t, u(t), x(t), p(t)) = \max_{v \in G(t)} H(t, v, x(t), p(t)), \; t \in (0, T),$$

where p is a solution if the ***adjoint equation***

$$p'(t) = -H_x(t, u(t), x(t), p(t)), \; t \in (0, T).$$

Here the maximum condition is a conditional extremum problem for the function H with respect to r variables. The adjoint equation actually includes n differential equations, and H_x is a vector whose elements are partial derivatives of H with respect to the variables x_i, $i = 1, \dots, n$.

In a concrete case, to find the optimal control, one should first find the dependence of the control on the state of the system and the solution of the adjoint equation from the maximum condition, which corresponds to Problem 2.3. The result is substituted into the equation of state and the adjoint equation, resulting in a system of $2n$ differential equations with $2n$ boundary conditions for the state of the system. Substituting the solution of this problem into the formula for the control from the previous step, we find the solution to the optimality conditions.

One example of a vector optimal control problem with a fixed final state that admits an analytical solution is considered in the next subsection.

10.2.5 Time optimal problem

Let us consider one more applied optimal control problem for a system with a fixed final state, which arises in mechanics.

Example 10.3 *There is some body that makes a rectilinear movement under the action of a force. It is required to select this force, acting within the given limits, in such a way as to transfer the body from one state to another in the minimum time.*

This corresponds to the ***time optimal problem***[20]. Let us give a mathematical formulation of this problem. First of all, we note that the rectilinear motion of a body is described by Newton's second law

$$mx''(t) = F(t),$$

where t is the time, x is the coordinate of the body in the direction of motion, m is the mass of the body, and F is the acting force. The body mass is considered known. After dividing the last equality by it, we obtain the equation

$$x''(t) = u(t), \tag{10.13}$$

where u is the acceleration, which is chosen as a control. The initial state of the system is characterized by its initial position a and initial velocity v, which are considered to be known. Choosing the moment of time $t = 0$ as the starting point, we obtain the initial conditions

$$x(0) = a, \;\; x'(0) = v. \tag{10.14}$$

We choose the coordinate system in such a way that the final state of the system is zero. Thus, it is required that at the final time T the system should move to the origin of coordinates and at the same time have zero velocity[21]. As a result, the final state of the system is characterized by the equalities

$$x(T) = 0, \;\; x'(T) = 0. \tag{10.15}$$

By the condition of the problem, the force acting on the system, and hence the corresponding acceleration, are limited. Thus, there are some restrictions on the control values at an arbitrary point in time. For definiteness, we assume that the set of admissible controls has the form

$$U = \{u|\ |u(t)| \leq 1,\ t \in (0,T)\}.$$

Thus, the problem of optimal control is to find such a function u from the set U that transfers the system described by equation (10.13) from the initial state (10.14) to the final state (10.15) in the minimum time T.

Now we need to convert this problem to the standard form. First of all, we note that equation (10.13) is of the second order. However, it can be reduced to a system of two first-order equations by introducing the unknowns $x_1 = x$, $x_2 = x'$. As a result, we obtain a system of equations

$$x_1'(t) = x_2(t),\ \ x_2'(t) = u(t),\ t \in (0,T). \tag{10.16}$$

Accordingly, the initial conditions (10.14) and the final conditions (10.15) are written as

$$x_1(0) = a,\ \ x_2(0) = v; \tag{10.17}$$

$$x_1(T) = 0,\ \ x_2(T) = 0. \tag{10.18}$$

To obtain a special case of Problem 10.2, it remains to bring the minimized functional (in this case, the time of movement T) to the standard integral form. Obviously, it can be written as follows

$$I(u) = \int_0^T dt,$$

which corresponds to the optimality criterion in Problem 10.2 with the subintegral function $g = 1$. Thus, it is required to choose such a control u from the set U that transfers the system described by equations (10.16) from the initial state (10.17) to the final state (10.18), while minimizing the functional I.

Let us now turn to the solution of the problem. First of all, the function H is introduced. Previously, it was defined by the formula

$$H(t,u,x,p) = \langle p, f(t,u,x)\rangle - g(t,u,x).$$

In this case, x is a second-order vector function whose components are the functions x_1 and x_2 defined above (the coordinate and velocity of the body). Then f also turns out to be a vector quantity whose components are the values on the right-hand sides of the system (10.16), i.e., x_2 and u. The function p also turns out to be a vector function with some components p_1 and p_2. In this case, $\langle p, f\rangle$ is a dot product, i.e., the sum of the products of the components of the corresponding vectors. Thus, the function H in this case has the form

$$H = p_1 f_1 + p_2 f_2 - g = p_1 x_2 + p_2 u - 1.$$

The adjoint equation, which has the form $p' = -H_x$, also turns out to be a vector equation, i.e., a system of differential equations. Here H_x is the vector of partial derivatives of the function H with respect to the variables x_1 and x_2. As a result, we obtain the adjoint system of equations

$$p_1'(t) = 0, \;\; p_2'(t) = -p_1(t), \; t \in (0, T). \tag{10.19}$$

Finally, the maximum condition implies the maximization of the control function H defined above on the previously specified set U. Taking into account the linearity of the dependence of H on u, we conclude that its maximum can be achieved only at the ends of the given segment $[-1, 1]$, depending on the sign of the function p_2. As a result, we find

$$u(t) = \begin{cases} 1, & \text{if } p_2(t) > 0, \\ -1, & \text{if } p_2(t) < 0. \end{cases} \tag{10.20}$$

Thus, the optimal control is determined from the equalities (10.16)–(10.20).

Let us find a solution to the resulting system. Note that the adjoint system (10.19) does not depend on other unknown functions. Find its solution $p_1(t) = c_1$ and $p_2(t) = -c_1 t + c_2$, where c_1 and c_2 are constants. Their arbitrariness should not be embarrassing, since as a result of substituting the control with the indicated function p_2 into equations (10.16), a system of two differential equations with four boundary conditions (10.17) and (10.18) is obtained.

According to formula (10.20), the control depends on the sign of the function p_2, which is linear. A linear function can change sign on the interval [0,1] at most once. If it does not change sign at all, then it is either everywhere positive, in which case $u(t) = 1$, or it is negative everywhere, in which case $u(t) = -1$ for any t. If the function p_2 changes sign once, then either $u(t)$ first takes the value 1 and then –1 if p_2 is decreasing, or vice versa if p_2 is increasing. Thus, four scenarios are possible.

For $u(t) = 1$ for all t, the solution to the Cauchy problem (10.16) and (10.17) has the form $x_1(t) = a + vt + t^2/2$, $x_2(t) = v + t$. Eliminating the parameter t from these equalities, we establish the following connection between the functions x_1 and x_2

$$x_1(t) = \left(a - \frac{v^2}{2}\right) + \frac{[x_2(t)]^2}{2}.$$

In the x_1 and x_2 plane, called the ***phase plane***, the set of points with coordinates $x_1(t), x_2(t)$ for all possible values of t forms a certain curve, called the ***phase curve***[22], which is a parabola. Naturally, any pair of initial states a, v has its own parabola; see Figure 10.4, and the arrows here indicate the direction of movement of the point in the phase plane with increasing time t, due to the fact that at $u = 1$ the function x_2 increases. Note that there is a single parabola, moving along which one can eventually get to the origin, which corresponds to the fulfillment of the final conditions (10.18). It corresponds to positive values of the parameter a and $v = -\sqrt{2a}$. Thus, for the indicated sets of parameters a and v, the solution to the problem is the function u, which is identically equal to 1.

For $u(t) = -1$ for all t, the solution to the Cauchy problem (10.16) and (10.17) has the form $x_1(t) = a + vt - t^2/2$, $x_2(t) = v - t$. Once again excluding the parameter t

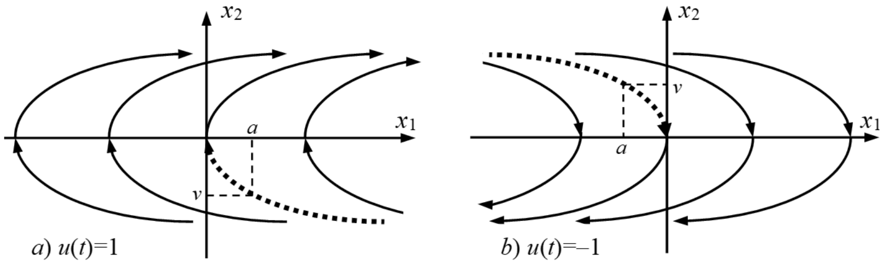

Figure 10.4 Phase curves for system (10.16) with constant controls.

from these equalities, we establish the following connection between the functions x_1 and x_2

$$x_1(t) = \Big(a + \frac{v^2}{2}\Big) - \frac{[x_2(t)]^2}{2}.$$

The set of points with coordinates $x_1(t)$, $x_2(t)$ for all possible values of t again forms a parabola defined by the parameters a and v. The direction of change of functions is due to the fact that at $u = -1$ the function x_2 decreases. Here, too, there is a single parabola, moving along which you can get to the origin, thereby ensuring the fulfillment of conditions (10.18). It corresponds to negative values of the parameter a and $v = \sqrt{-2a}$; see Figure 10.4. Thus, for the specified set of parameters a and v, the solution to the problem is the function u, which is identically equal to –1.

Let us now assume that at the initial moment of time the system is at some point A of the phase plane with coordinates a and v that do not satisfy the above conditions. Let it lie above a curve made up of two halves of parabolas, moving along which you can get to the origin; see Figure 10.5. This corresponds to the initial states $a > 0$, $v > -\sqrt{2a}$ or $a < 0$, $v < \sqrt{-2a}$. In this case, the control is initially assumed to be $u(t) = -1$. The movement along the corresponding parabola continues until reaching point B, which lies on the parabola leading to the origin. This happens at the moment of time τ, when the following conditions holds $x_1(\tau) > 0$, $x_2(\tau) = -\sqrt{2x_1(\tau)}$. For $t > \tau$, we have $u(t) = 1$.

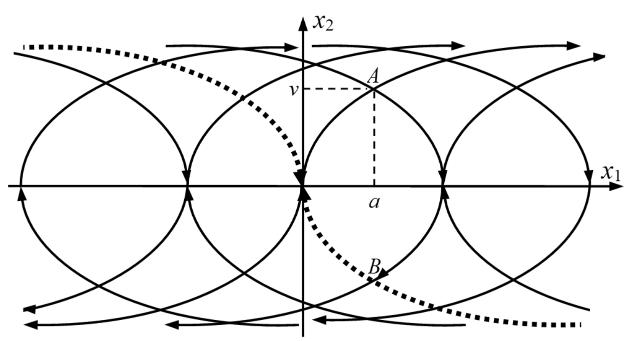

Figure 10.5 Solution of the time optimal problem.

If the initial states of the system satisfy the conditions $a < 0$, $v < -\sqrt{2a}$ or $a > 0$, $v > \sqrt{-2a}$, then the control is first assumed to be equal to $u(t) = 1$, which corresponds to motion in the phase plane until reaching the parabola leading to the origin. This happens at the moment of time τ, when $x_1(\tau) < 0$ and $x_2(\tau) = \sqrt{-2x_1(\tau)}$. For $t > \tau$, we have $u(t) = -1$.

Thus, the solution of the problem of optimal speed is found for any values of the initial position a and initial velocity v of the considered body.

Additional conclusions

Based on the additional analysis of optimal control problems for systems with a fixed final state, carried out in Appendix, we can draw the following additional conclusions.

- The Bellman optimality principle for the problem of optimal control of a system with a fixed final state, considered in Example 9.1, is satisfied.
- The problem of finding the curve of the smallest length passing through two given points is reduced to the problem of optimal control of a system with a fixed final state. The curve of least length is a straight line.
- The penalty method can be used in conjunction with various methods related to calculating the derivative of a functional, particularly, with variational inequalities.
- The maximum principle naturally extends to the vector problem of optimal control of a system with a fixed final state.
- The time optimal problem is reduced to a vector problem of optimal control of a system with a fixed final state.
- The time optimal problem admits an analytical solution for any initial position and initial velocity of the body.

Notes

1. This is quite natural, since the function H is the same in similar problems with a free and a fixed final state.

2. The obtained analytical dependence of the optimal control on the state function, as in the case of the system with a free final state considered in Chapter 3, gives a solution to the problem of the ***synthesis problem***.

3. In Appendix, these results are extended to the vector case.

4. This is explained by the linearity of the equation of state and the optimality criterion both in terms of control and state. We have already met with the equivalence of the indicated optimality conditions in Examples 3.1, 3.2, and 3.3. However, for Example 5.1, the maximum principle turned out to be more efficient, since the set of its solutions was essentially narrower than the set of solutions of the corresponding variational inequality.

5. This is proved in the same way as the analogous property for Example 9.1 in Section 9.2.1. We also note that for non-linear systems, it is rather difficult to establish the convexity of the set of controls that guarantee the transition of the system to a given final state. As a result, the use of variational inequalities turns out to be difficult.

6. One can, for example, prove the sufficiency of the optimality conditions using Theorem 5.2.

7. It is clear that it is possible to substantiate the convergence of the penalty method and find with its help the exact solution of the problem only in exceptional cases.

8. We will deal with the violation of Bellman optimality principle of in Chapter 15.

9. This object is called the ***differential***.

10. We recall that the integral by its nature is, roughly speaking, an infinitely large sum of infinitely small quantities. As applied to this case, the length of the entire arc of the curve is the sum of arbitrarily small lengths (differentials) ds, which corresponds to the integration procedure.

11. Minimization problems for integral functionals are the subject of the ***calculus of variations***; see [37], [61], [208]. In principle, the problem of minimizing the curve length in the original formulation, i.e., without introducing control u, is a particular case of the classical ***Lagrange problem*** of the calculus of variations. We are talking about minimizing an integral functional that explicitly depends on the desired function and its first derivative (in this case, only on the derivative) on the set of functions that takes known values at the boundaries of the integration interval. The necessary condition for the extremum here is the ***Euler equation***, which is some second-order differential equation, which is solved with two given boundary conditions. The system of optimality conditions for the considered example is reduced to the same boundary value problem as a result of applying the previously considered method of eliminating unknowns.

12. Let us pay attention to the fact that the optimality criterion in this case does not depend on the function x, i.e., we have its explicit dependence on the required function u. In this regard, it seems that we can simply solve the problem of minimizing this functional without resorting to the equation of state. Obviously, the value of the optimality criterion for any control is not less than 1. Moreover, $I(u) = 1$ only for $u(t) = 0$ for all t. However, in this case $x(t) = 0$ for all t, which means that the equality $x(T) = X$ will not hold. Thus, we cannot ignore the equation of state, since it implicitly characterizes the set of admissible controls.

13. Indeed, using the formula for the derivative of the quotient of two functions, we find

$$H_{uu} = -\frac{\sqrt{1+u^2} - \frac{u^2}{\sqrt{1+u^2}}}{1+u^2} = -\frac{1}{(1+u^2)^{3/2}}.$$

14. About Cantor function; see [77].

15. Any number from the segment [0,1] can be represented as a ternary fraction. Obviously, all numbers that fall into M at the first step of the described process have the digit 1 in ternary notation; see Figure 10.2. At the second step, M includes numbers whose second significant digit is 1, and at the third step, numbers with one as the third decimal place, and so on. Thus, the Cantor set includes all those and only those numbers from the segment [0,1] that do not have units in the ternary notation. Therefore, this set is characterized by the equality

$$C = \{0, 0.2, 0.02, 0.22, 0.002, 0.022, 0.202, 0.222, \dots\}.$$

Any element x of the Cantor set can be uniquely associated with the number $x/2$ from the set

$$C' = \{0, 0.1, 0.01, 0.11, 0.001, 0.011, 0.101, 0.111, \dots\}.$$

The latter can be associated with the set C'', each element of which has the same notation as the corresponding element C', but only in the binary system (for example, the second element 0.1 of the set C' is a fraction $1/3$, and the corresponding element 0.1 of C'' is equal to $1/2$). Obviously, the set C'' consists of all possible numbers of the interval [0,1] written in binary. Therefore, up to the notation, the set C coincides with the segment [0,1]. Then the Cantor set, which is equivalent to it, has the cardinality of the continuum. On the concept of cardinality; see, for example, [100], [106].

16. The set M is the union of open intervals. Therefore, its measure is equal to the sum of the measures, i.e., the lengths of its constituent intervals. At the kth step of the set construction process, we have $2k$ intervals of length $1/3k$. As a result, we find the measure of the set M

$$\mu(M) = \sum_{k=1}^{\infty} \frac{2^{k+1}}{3^k} = \frac{1}{3}\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^k.$$

Using the formula for the sum of members of a geometric progression, we establish that $\mu(M) = 1$. Since the sets C and M form a partition of the segment [0,1] (they do not intersect, and their union makes up the entire segment), the sum of their measures is equal to the measure of their union, i.e., the length of the considered segment, equal to one. From here, it follows that the Cantor set has zero measure. About the concept of measure; see [94], [100], [106], [158].

17. In particular, for Example 9.1 considered in Section 10.1.2, this result is established due to the linearity of the equation of state of the system both in control and in the state of the system.

18. Indeed, we find the value

$$I_\varepsilon(u + \sigma h) = \int_0^T \sqrt{1 + (u + \sigma h)^2}dt + \frac{1}{2\varepsilon}\big[y(T) - X\big]^2,$$

where y is a solution of the problem

$$y'(t) = u(t) + \sigma h(t),\ t \in (0, T),\ y(0) = 0.$$

Using the Taylor series expansion, we obtain the equality

$$\sqrt{1 + (u + \sigma h)^2} = \sqrt{1 + u^2} + \frac{uh}{\sqrt{1 + u^2}}\sigma + o(\sigma),$$

where $o(\sigma)/\sigma \to 0$ as $\sigma \to 0$. Besides, we get

$$\big[y(T) - X\big]^2 = [X(T)]^2 + 2\big[y(T) - X\big]z(T) + [z(T)]^2,$$

where $z = y – x$. Now we have

$$I_\varepsilon(u + \sigma h) - I_\varepsilon(u) = \int_0^T \Big[\sigma\frac{uh}{\sqrt{1 + u^2}} + o(\sigma)\Big]dt + \frac{1}{\varepsilon}\big[y(T) - X\big]z(T) + \frac{1}{2\varepsilon}[z(T)]^2.$$

The function z here is a solution to the problem

$$z'(t) = \sigma h(t),\ t \in (0, T),\ z(0) = 0,$$

and hence directly proportional to the number σ. Multiplying the last equation by an arbitrary function p and integrating over t, after integrating by parts, taking into account the initial condition, we obtain

$$-\int_0^T p'(t)z(t)dt + p(T)z(T) = \sigma \int_0^T p(t)h(t)dt.$$

We choose here as p the solution of the problem

$$p'(t) = 0,\ t \in (0,T),\ p(T) = \frac{1}{\varepsilon}\big[y(T) - X\big].$$

Then the following equality holds

$$I_\varepsilon(u + \sigma h) - I_\varepsilon(u) = \int_0^T \Big[\sigma\Big(\frac{u}{\sqrt{1+u^2}} + p\Big)h + o(\sigma)\Big]dt + \frac{1}{2\varepsilon}[z(T)]^2.$$

Dividing this equality by σ and passing to the limit at $\sigma \to 0$, we obtain

$$I'_\varepsilon(u)h = \int_0^T \Big(\frac{u}{\sqrt{1+u^2}} + p\Big)hdt,$$

whence it follows

$$I'_\varepsilon(u) = \frac{u}{\sqrt{1+u^2}} + p.$$

19. It is also possible that only a part of the final states of the system are fixed, for example, the conditions $x_i(T) = x_{Ti}$, $i = 1, \dots, l$, where $l < n$ are specified. In this case, for the vector function p at the final moment of time, the conditions corresponding to the indices $i = l+1, \dots, n$.

20. On problems of optimal performance for systems with lumped parameters; see [5], [76], [61], [62], [67], [90], [140], [152], [182], and for systems with distributed parameters; see [3], [23], [104], [111], [116], [129], [157].

21. This means that the body not only gets to the origin of coordinates at the final moment of time, but also remains there afterwards.

22. The phase plane and phase curve are fundamental concepts of the qualitative theory of differential equations; see [10], [86].

CHAPTER 11

Counterexamples of optimal control problems with a fixed final state

The previous two chapters dealt with optimal control problems for systems with a fixed final state. The effectiveness of the respective optimization methods was illustrated by specific examples. However, in the practical solving of such problems, certain difficulties may arise. They are the subject of this chapter.

11.1 LECTURE

We will consider rather simple examples of optimal control problems for systems with a fixed final state, for which the application of the standard technique runs into various troubles. In particular, the Sections 11.1.1, 11.1.2, 11.1.3, and 11.1.4 describe problems, respectively, with insufficient optimality conditions, with a degeneration of the maximum principle, with a non-unique optimal control and in the absence of an optimal control. The presence of a fixed final state complicates the analysis, but is not an obstacle to its implementation.

11.1.1 Insufficiency of optimality conditions

In the examples of optimal control problems with a fixed finite state considered earlier, the necessary optimality condition was always sufficient. Let us give an example of the insufficiency of the optimality conditions.

Example 11.1 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt$$

on a subset of such functions $u = u(t)$ from the set

$$U = \{u \mid |u(t)| \leq 1,\ t \in (0,1)\}$$

DOI: 10.1201/9781003398585-11

that guaranty the equality

$$x(1) = 1/5, \tag{11.1}$$

where x is a solution of the Cauchy problem

$$x'(t) \;=\; u(t),\; t \in (0,1);\; x(0) = 0. \tag{11.2}$$

This problem differs from the one considered in Example 5.1 only by the presence of condition (11.1). As a result, the system of optimality conditions here differs from the one used in that example only by the presence of a final condition for the state function and the absence of such a condition for the adjoint equation. In particular, the function H is defined by the formula

$$H \;=\; pu–(u^2 + x^2)/2,$$

where p satisfies the adjoint equation

$$p'(t) \;=\; x(t),\; t \in (0,1). \tag{11.3}$$

Using minimum condition for the function H find

$$u(t) \;=\; \begin{cases} \;\;1, & \text{if } p(t) < 0, \\ -1, & \text{if } p(t) > 0. \end{cases} \tag{11.4}$$

Thus, the system of optimality conditions includes formulas[1] (11.1)–(11.4).

Note, first of all, that controls that are identically equal to 1 or −1 do not guarantee the realization of condition (11.1), and therefore are not admissible. Thus, the solution to the problem can be an exclusively discontinuous function[2]. Suppose there exists a point ξ from the interval (0,1) such that the function p is negative if $t < \xi$ and this is positive if $t > \xi$. Then the control is 1 on the interval $(0, \xi)$ and –1 on the interval $(\xi, 1)$. The corresponding solutions to problem (11.2) are determined by the formulas $x(t) = t$ and $x(t) = 2\xi–t$. In this case, the condition (11.1) holds that is

$$x(1) \;=\; 2\xi–1 \;=\; 1/5.$$

Find $\xi = 3/5$. The equality (11.3) on the interval (0,3/5) is $p' = t$. By the positiveness of the derivative, the function p is strictly increasing. However, at the point ξ it must vanish. Therefore, up to this point it was negative, which is consistent with the accepted assumption. We have the adjoint equation $p' = 6/5–t$ for $t > 3/5$. This derivative is positive, which means that the function p is still increasing. Taking into account the equality $p(3/5) = 0$, we conclude that the considered function is positive in this interval, which also agrees with the existing assumption. Thus, the control that takes the value 1 on the interval $(0, 3/5)$ and –1 on the interval $(3/5, 1)$ satisfies the system of optimality conditions (11.1)–(11.4); see Figure 11.1.

Let now, on the contrary, the control is equal to –1 on the interval $(0, \xi)$ and 1 on the interval $(\xi, 1)$, which means, according to equality (11.4), the function p is positive on the first set, and negative on the second.

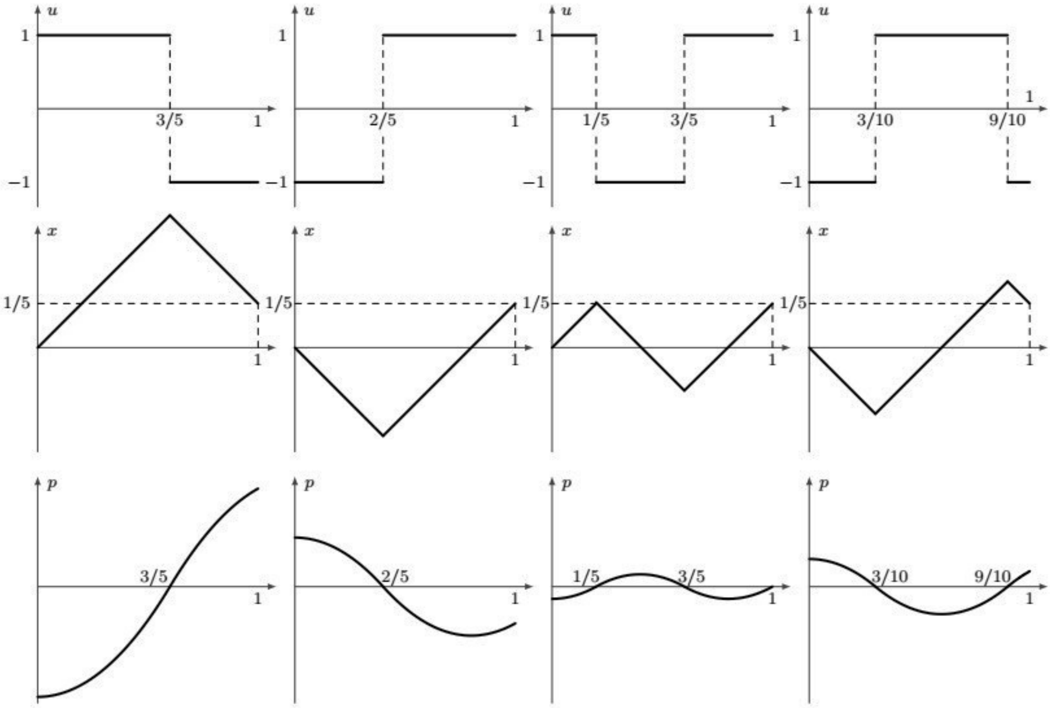

Figure 11.1 The first four solutions of the optimality conditions for Example 11.1.

The corresponding solution to problem (11.2) is $x(t) = -t$ on the interval $(0, \xi)$ and $x(t) = t–2\xi$ on the interval $(\xi, 1)$. Then equality (11.1) implies

$$x(1) = 1–2\xi = 1/5,$$

which means, $\xi = 2/5$. The adjoint equation (11.3) on the interval (0,2/5) is characterized by the equality $p' = -t$. From here, it follows that the function p decreases in the initial segment. However, at its right end it must vanish. Therefore, before that it was positive, as expected. On the interval (2/5,1) the adjoint equation has the form $p' = t–4/5$. Thus, the function p continues to decrease, which means it will become negative. At point 4/5, the function has a minimum. After that, it increases and takes the value $p(1) = –3/50$. Thus, this function remains negative throughout the interval (2/5,1), as it should be. Thus, the control that takes the value −1 on the interval (0,2/5), and 1 on the interval (2/5,1) also satisfies the system of optimality conditions; see Figure 11.1. Therefore, the system has two solutions with one discontinuity point[3].

Now let the control have two discontinuity points ξ and η, and it is equal to 1 on the interval $(0, \xi)$, −1 on (ξ, η), and 1 on $(\eta, 1)$. Then, according to equality (11.4), the function p must be negative on the first and third intervals and positive on the second. The solution x of problem (11.2) is equal to t, $2\xi–t$ and $t+2\xi–2\eta$, respectively, on the considered intervals. From equality (11.1), we find

$$x(1) = 1 + 2\xi–2\eta = 1/5.$$

From here, we find $\eta = \xi + 2/5$. Consider the adjoint equation $p' = t$ on the interval $(0, \xi)$. Obviously, its solution increases here, reaching zero at the point ξ. Therefore, this function initially takes negative values, as expected. For the next interval $(\xi, \xi + \ 2/5)$, we have the adjoint equation $p' = 2\xi – t$, besides the solution is equal to zero at the ends of the interval. Integrating the last equality by this interval, find $\xi = 1/5$, so $\eta = 3/5$. Then from the equality $p' = 2/5 – t$, it follows that this function at $t > 1/5$ increases to point 2/5 and then decreases to zero. Thus, the considered function is positive on the second interval. Finally, on interval (3/5,1), we have the equation $p' = t – 4/5$. Its solution here has a minimum at the point $t = 4/5$, with $p(1) = 0$. The results obtained are consistent with formula (11.4), which means that control equal to 1 at interval $(0, 1/5)$ and $(3/5, 1)$ and –1 at interval $(1/5, 3/5)$ also satisfies the system optimality conditions[4]; see Figure 11.1.

Now suppose the control is –1 on intervals $(0, \xi)$ and $(\eta, 1)$ and 1 on (ξ, η). Then the function p must be positive on the first and third intervals and negative on the second. The state x is, respectively, $-t$, $t - 2\xi$, and $-t - 2\xi + 2\eta$ successively at the indicated intervals. From equality (11.1), we find

$$x(1) \;=\; –1 – 2\xi + 2\eta \;=\; 1/5,$$

so $\eta = \xi + 3/5$. Consider the adjoint equation $p' = -t$ on interval $(0, \xi)$. The function p decreases here, reaching zero at the point ξ. Then it is positive on first interval. The adjoint equation on second interval $(\xi, \xi + 3/5)$ is $p' = t – 2\xi$, besides the solution is zero at its ends. Integrating the last equality by this interval, determine $\xi = 3/10$, so $\eta = 9/10$. From the equality $p' = t – 3/5$, it follows that function p at $t > 3/10$ decreases to point 3/5, where it reaches its minimum, and then increases to zero. Thus, this function is negative on the second interval. Finally, on the third interval (9/10,1), we have the equation $p' = -t + 6/5$. Thus, the function p is increasing here and becomes positive. Therefore, the control equal with value –1 on interval (0,3/10) and (9/10,1) and 1 on interval (3/10,9/10) is also a solution of optimality conditions; see Figure 11.1.

Suppose the control has three break points ξ, η, and ζ, besides this function is equal to 1 on interval $(0, \xi)$ and (η, ζ) and to –1 on (ξ, η) and $(\zeta, 1)$. The function x is t, $2\xi – t$, $t + 2\xi – 2\eta$ and $2\xi – 2\eta + 2\zeta – t$ sequentially on the considered intervals. From equality (11.1), we get

$$x(1) \;=\; 2\xi – 2\eta + 2\zeta – 1 \;=\; 1/5.$$

Therefore, $\zeta = 3/5 + \eta - \xi$. We have the adjoint equation $p' = t$ on first interval. His solution rises to zero at the point ξ. Therefore, it is negative on this interval, which is consistent with equality (11.4). The function p satisfies the equation $p' = 2\xi – t$ on interval (ξ, η), besides it is zero at its ends. Integrating this equality over the specified interval, we find $\eta = 3\xi$, so $\zeta = 3/5 + 2\xi$. Then the function p satisfies the equation $p' = t – 4\xi$ on interval $(3\xi, 3/5 + 2\xi)$, and this is equal to zero at the ends. Integrating this equality, determine $\xi = 1/5$. Therefore, $\eta = 3/5$ and $\zeta = 1$. Thus, the last point turns out to be boundary, which means that the previously accepted assumption is not realized at all[5].

Suppose, on the contrary, that control is –1 on intervals $(0, \xi)$ and (η, ζ) and this is 1 on (ξ, η) and $(\zeta, 1)$. The function x is equal to $-t$, t–2ξ, $-t$–$2\xi + 2\eta$ and t–$2\xi + 2\eta$–2ζ sequentially on the considered intervals. Using equality (11.1), we get

$$x(1) \;=\; 1 - 2\xi + 2\eta - 2\zeta \;=\; 1/5.$$

Find $\zeta = 2/5 + \eta$–ξ. We have the adjoint equation $p' = t$ on first interval. Its solution increases to zero at the point ξ. Therefore, this is negative here, which is consistent with equality (11.4). The function p satisfies the equation $p' = t$–2ξ on (ξ, η), besides this is zero at the ends. Integrating this equality, we find again $\eta = 3\xi$ and $\zeta = 2/5 + 2\xi$. Then on the interval $(3\xi, 2/5 + 2\xi)$ the function p satisfies the equation the equation $p' = -t + 4\xi$, and it is equal to zero on its boundaries. Integrating this equality over this interval, we determine $\xi = 2/15$. It follows that $\eta = 2/5$ and $\zeta = 2/3$. It is easy to verify that the corresponding function p has properties consistent with relation (11.4).

Continuing this process, we can establish that there is a unique solution of optimality conditions with each number of discontinuity points greater than two[6], and for an odd number of discontinuity points on the first segment of continuity, control takes the value –1, and for an even number it takes the value 1. It is easy to make sure[7], that the maximal value of the functional is realized for the control that is equal to 1 for $t < 3/5$ and to –1 for $t > 3/5$. Thus, the problem has the unique solution with infinite set of solutions of optimality conditions[8].

11.1.2 Singular control

In the considered examples of optimal control problems for systems with fixed final state, the maximum principle did not degenerate. However, the presence of a singular control is not related to the behavior of the system at the final time. As a result, we can expect the appearance of singular control for this class of problems as well.

Example 11.2 *Consider the system described by the Cauchy problem*

$$x'(t) \;=\; u(t),\; t \in (0,1);\; x(0) = 0. \tag{11.5}$$

The control u is chosen from the set

$$U \;=\; \{u \,|\, |u(t)| \le \pi,\; t \in (0,1)\}.$$

Besides it guaranties the final condition

$$x(1) = 0. \tag{11.6}$$

It is required to minimize on this set the functional

$$I(u) \;=\; \frac{1}{2}\int\limits_0^1 (x^2 - 2x \sin \pi t)dt.$$

For solving this problem, we use the standard method. Determine the function

$$H = pu - (x^2 - 2x \sin \pi t)/2,$$

where p satisfies the adjoint equation

$$p'(t) = x(t) - \sin \pi t, \ t \in (0,1). \tag{11.7}$$

Due to the linearity of the function H with respect to control, two variants of the solution maximum condition are possible: either its maximum is reached on the boundary of the set of admissible controls, or control is singular.

The regular solution is determined by the equality

$$u(t) = \begin{cases} \pi, & \text{if } p(t) > 0, \\ -\pi, & \text{if } p(t) < 0. \end{cases} \tag{11.8}$$

The control here can have only two values that are π and $-\pi$. Suppose $u(t) = \pi$ everywhere. Then the solution of the problem (11.5) is $x(t) = \pi t$, so $x(1) = \pi$. However, this contradicts condition (11.6). If $u(t) = -\pi$ everywhere, then $x(t) = -\pi t$, and $x(1) = -\pi$. This result is also inconsistent with equality (11.6). Therefore, the control determined by the formula (11.8) can be discontinuous only or that is equivalent, the function p changes the sign on interval (0,1).

Denote by ξ the first (and possibly not the only) point from the origin in interval (0,1) where equality $p(\xi) = 0$ holds. For definiteness, suppose that $p(t) > 0$ for $0 < t < \xi$. Then on this interval $u(t) = \pi$, and hence $x(t) = \pi t$. Substituting this value into the adjoint equation (11.7), we find the derivative

$$p'(t) = \pi t - \sin \pi t, \ t \in (0, \xi).$$

Obviously, $\pi t > \sin \pi t$, so p' is positive on interval $(0, \xi)$. Therefore, is function p increases monotonically. Thus, we have a function that, on the considered interval, takes exclusively positive values and increases. Under these conditions equality $p(\xi) = 0$ is impossible. This means that the previously accepted assumption about the behavior of the function p turned out to be wrong. The impossibility of this function being negative on the first interval of continuity is proved[9].

Thus, the function p cannot be positive or negative. Therefore, the formula (11.8) is impossible, and the system (11.5)–(11.8) has no solution. This does not exclude the case $p(t) = 0$ for all $t \in (0,1)$, which corresponds to the degeneration of the maximum condition. Therefore, if its solution exists, then it is singular[10].

For $p = 0$, from the adjoint equation it follows that $x(t) = \sin \pi t$. Note that function x satisfies both boundary conditions. It turns out to be the solution equation of the state for $u(t) = \pi \cos \pi t$. This satisfies the existing restrictions on the values of control, which means that it is a solution of the optimality conditions.

In Chapter 6, the Kelley condition was given, which is a necessary condition for the optimality of the singular control[11]. By this result, optimal singular control satisfies the inequality

$$\frac{\partial}{\partial u} \frac{d^2}{dt^2} \frac{\partial H}{\partial u} \geq 0.$$

Thus, if a singular control satisfies the Kelley condition, then it can be optimal. However, this is not optimal if it does not satisfy this condition.

Check the validity of the Kelley condition for this singular control. We find

$$\frac{\partial H}{\partial u} = p.$$

Using the adjoint equation, determine

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' = x(t) - \sin \pi t.$$

By the state equation, we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = x' - \pi \cos \pi t = u - \pi \cos \pi t.$$

Now, we have

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 1.$$

Thus, the Kelley condition is true; hence, the considered singular control can be optimal.

In reality, the considered functional is

$$I = \frac{1}{2}\int\limits_0^1 \left(x - \sin \pi t\right)^2 dt - \frac{1}{2}\int\limits_0^1 \sin^2 \pi t.$$

The second term here does not depend on control, and the first one takes non-negative values, and its equality to zero is possible only for $x(t) = \sin \pi t$. This corresponds to singular control $u(t) = \pi \cos \pi t$, which is optimal. Thus, the optimal control problem under consideration has a unique solution, and the optimality conditions are necessary and sufficient[12].

11.1.3 Non-uniqueness of the optimal control

In all the considered problems with fixed final state, the optimal control was the only one. Consider another example.

Example 11.3 *We have the system described by the Cauchy problem*

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{11.9}$$

The control u is chosen from the set

$$U = \{u \,|\, |u(t)| \leq 1,\ t \in (0,1)\}.$$

Besides, it guaranties the final condition

$$x(1) = 0. \tag{11.10}$$

It is required to minimize on this set the functional

$$I(u) = \frac{1}{4}\int_0^1 (x^4 - 2x^2 \sin^2 \pi t)dt.$$

Determine the function

$$H(t,u,x,p) = pu - \frac{1}{4}(x^4 - 2x^2 \sin^2 \pi t).$$

Then we have the adjoint equation

$$p'(t) = x(t)^3 - x(t)\sin^2 \pi t,\ t \in (0,1). \tag{11.11}$$

Optimal control satisfies here the maximum condition

$$p(t)u(t) - \frac{1}{4}\Big[x(t)^4 - 2x(t)^2 \sin^2 \pi t\Big] = \max_{|v|\le 1} p(t)v - \frac{1}{4}\Big[x(t)^4 - 2x(t)^2 \sin^2 \pi t\Big],\ t \in (0,1). \tag{11.12}$$

Thus, we have the system (11.9)–(11.11) for finding the unknown functions u, x, p.

The function H is linear with respect to control. Therefore, its maximum is reached on the boundary of the set of admissible controls, or the solution of the maximum condition is a singular. The absence of a regular solution is set in the same way as in the previous example[13]. Then the solution of optimality conditions is singular[14]. This is possible only if $p(t) = 0$ for all $t \in (0,1)$, which corresponds to the vanishing of the coefficient at control in the definition of the function H. Determining $p(t) = 0$ at the adjoint equation (11.11), we get

$$0 = p'(t) = x(t)^3 - x(t)\sin^2 \pi t,\ t \in (0,\xi).$$

We have a cubic equation for $x(t)$. It has three solutions

$$x_1(t) = 0,\ x_2(t) = \sin \pi t,\ x_3(t) = -\sin \pi t.$$

Note that these functions satisfy the given boundary conditions. These states correspond to controls

$$u_1(t) = 0,\ u_2(t) = -\cos \pi t,\ u_3(t) = \cos \pi t.$$

Therefore, system of optimality conditions has three solutions, besides these controls are singular. The question arises, are they (or at least some of them) optimal?

Check the validity of the Kelley condition. Determine the derivative

$$\frac{\partial H}{\partial u} = p.$$

Using the adjoint equation (11.11), we find

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' = x^3 - \sin^2 \pi t.$$

Using the state equation (11.9), we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = (3x^2 - \sin^2 \pi t)x' - \pi x \sin 2\pi t = (3x^2 - \sin^2 \pi t)u - \pi x \sin 2\pi t.$$

Now, we have

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 3x^2 - \sin^2 \pi t.$$

Substitute in this equality the values of the state functions defined earlier, corresponding to the singular control. We determine

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=u_1} = 3x_1^2 - \sin^2 \pi t = -\sin^2 \pi t,$$

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=u_2} = 3x_2^2 - \sin^2 \pi t = 2\sin^2 \pi t,$$

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=u_3} = 3x_3^2 - \sin^2 \pi t = 2\sin^2 \pi t.$$

Obviously, first of these values is negative, and the second and third are positive for all $t \in (0,1)$. From Kelley condition, it follows that the singular controls control $u_2 = -\cos \pi t$ and $u_3 = \cos \pi t$ can be optimal, and the control $u_1(t) = 0$ is not optimal.

It remains to be seen whether control u_2 and u_3 (or at least one of them) is actually optimal. Let us return to the consideration of the optimality criterion

$$I = \frac{1}{4}\int_0^1 (x^4 - 2x^2 \sin^2 \pi t)dt.$$

It can be transformed to

$$I = \frac{1}{4}\int_0^1 (x^2 - \sin^2 \pi t)^2 dt - \frac{1}{4}\int_0^1 sin^4 \pi t dt.$$

The second integral here does not depend from control. The first one is not negative; besides it equals to zero only for $x^2 = \sin^2 \pi t$. This is possible for $x(t) = \sin \pi t$ or for $x(t) = -\sin \pi t$, which corresponds to determined early states x_2 and x_3. Therefore, the controls u_2 and u_3 are in reality optimal[15]. Thus, the considered optimal control problem has two solutions[16]; and the maximum principle is a necessary, but non-sufficient condition[17].

11.1.4 Unsolvable optimal control problem

Consider another optimal control problem for a system with a fixed final state[18].

Example 11.4 *We have the system described by the Cauchy problem*

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{11.13}$$

Control u is determined from such functions that ensure the fulfillment of the conditions

$$x(1) = 0. \tag{11.14}$$

It is required to minimize on this set functional[19]

$$I(u) = \int_0^1 \sqrt[4]{1+u^2}dt.$$

Determine the function

$$H = pu - \sqrt[4]{1+u^2}.$$

The optimal control satisfies the maximum condition

$$pu - \sqrt[4]{1+u^2} = \max_v \Big(pv - \sqrt[4]{1+v^2}\Big), \tag{11.15}$$

where p is the solution of adjoint equation

$$p'(t) = -H_x = 0, \ t \in (0,1). \tag{11.16}$$

Thus, we have the system of optimality conditions (11.13)–(11.16).

The formula (11.15) is the non-conditional extremum problem. Equaling to zero the derivative of the function H with respect to control, we get

$$\frac{\partial H}{\partial u} = p - \frac{u}{2(1+u^2)^{3/4}} = 0.$$

From equation (11.16) it follows that the function p is constant. From the last equality, it follows that

$$\frac{u}{2(1+u^2)^{3/4}} = c, \tag{11.17}$$

where c is a constant.

We have a nonlinear algebraic equation. Its solution, as it were, can only be a constant c_0 (perhaps not the only one), because the argument t does not appear at all in this equality[20].

Substituting this control to the problem (11.13), we find the function $x(t) = c_0 t$. Determine here $t = 1$. Using equality (11.14), we get $c_0 = 1$. Thus, there exists the unique solution of optimality conditions that is the function $u_0 = 1$. We conclude that there exists the unique constant control u_0, for which the derivative of the function H vanishes when condition (11.14) is satisfied. It is easy to see that the second derivative of the function H at the point u_0 is negative, i.e., we are really having the point of maximum for this function[21].

From the problem (11.13) with $u = u_0$ it follows $x_0(t) = t$. Using equality (11.16), determine the solution of the adjoint equation $p_0(t) = 2^{-7/4}$. Thus, the functions

$$u_0(t) = 1, \ x_0(t) = t, \ p_0(t) = 2^{-7/4}$$

are unique solution of the system of optimality conditions[22] (11.13)–(11.16). Let us try to prove that control u_0 is optimal for this example.

Let us turn to a direct analysis of the problem. Obviously, the integrand of the given functional is in no way less than unity. As a result, we obtain a lower bound for it $I(u) \geq 1$ for any control u. Let us define the sequence of controls[23] (see Figure 11.2)

$$u_k(t) = \begin{cases} 0, & \text{if } 0 < t < (k-1)/k, \\ k, & \text{if } (k-1)/k < t < 1, \end{cases} \quad k = 1, 2, \ldots .$$

The corresponding state sequence is (see Figure 11.2)

$$x_k(t) = \begin{cases} 0, & \text{if } 0 < t < (k-1)/k, \\ kt - k + 1, & \text{if } (k-1)/k < t < 1, \end{cases} \quad k = 1, 2, \ldots .$$

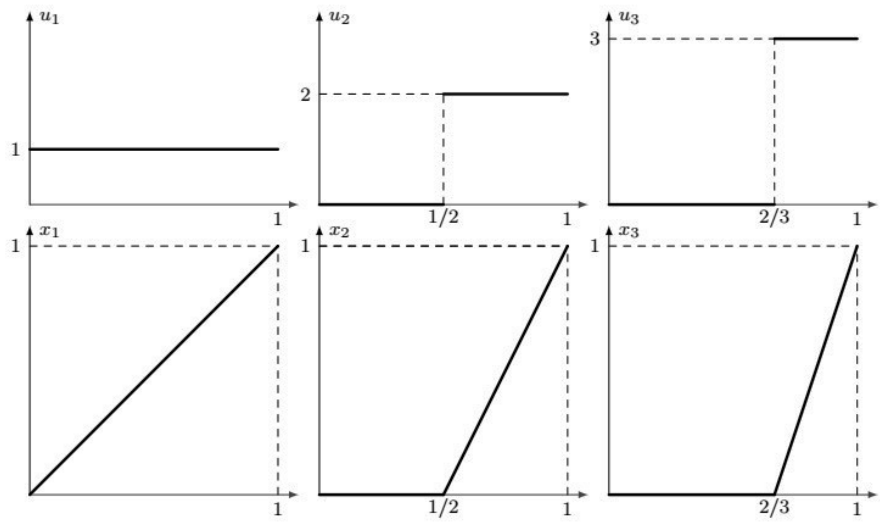

Figure 11.2 Sequences of controls and state for Example 11.4.

Obviously, we have the equality $x_k(1) = 1$, so the final condition (11.14) is true. Thus, the control u_k is admissible. Find the value of the functional

$$I(u_k) = \int_0^1 \sqrt[4]{1 + u_k^2}\, dt = \int_0^{(k-1)/k} dt + \int_{(k-1)/k}^1 \sqrt[4]{1 + k^2}\, dt = \frac{k-1}{k} + \frac{\sqrt[4]{1+k^2}}{k}.$$

Now we have the convergence $I(u_k) \to 1$ as $k \to \infty$.

Therefore, there is no control on which the given functional takes a value less than one, but there is a sequence of controls that ensure the transfer of the system to a given final state, on which the sequence of functionals converges to 1. Hence, it follows that the lower bound of the minimized functional on the set of controls that ensure the transfer of the system to a given final state is 1.

Previously, we defined the unique solution u_0 of the maximum principle. It is natural to assume that it is at this point that the control functional reaches its minimum equal to 1. However, the following equality holds

$$I(u_0) = \int_0^1 \sqrt[4]{1+u_0^2}dt = \int_0^1 \sqrt[4]{2}dt = \sqrt[4]{2}.$$

We are forced to make a seemingly paradoxical conclusion: the value of the minimized functional on the only found solution u_0 of the optimality conditions is greater than its lower bound on the set of admissible controls.

The possible non-optimality of a concrete solution of the maximum principle is not surprising. We have encountered a similar situation many times in the case of insufficient optimality conditions. Not so surprising is the absence of solutions of the maximum principle in the case of unsolvability of the optimization problem. The paradox of this situation lies in the fact that it turns out that the unique solution of the maximum principle is not optimal[24].

By definition, a necessary condition is such an optimality condition that any optimal control necessarily satisfies. The non-optimality of the unique solution of the maximum principle is possible only in the case when the optimization problem has no solution at all. Indeed, if optimal control existed, then it would necessarily satisfy the maximum principle. However, its unique solution u_0 is not optimal. This is possible only in the case when the considered optimal control problem does not have a solution[25].

The absence of optimal control can also be proven directly. Obviously, if the lower bound of the given functional (i.e., 1) is reached, then control must necessarily be identically equal to zero. However, the corresponding solution of problem (11.13) will also be equal to zero, which means that it cannot satisfy the boundary condition (11.14) in any way. Thus, the only control on which the functional reaches its lower bound cannot transfer the considered system to the given final state, and therefore is not admissible. Thus, this optimal control problem does not really have a solution[26].

RESULTS

Below, there is a list of questions in the field of optimal control problems for systems with fixed final state, related to the considered examples, the main conclusions from the results of the analysis, as well as the problems that arise in this case and need additional research.

Questions

It is required to answer questions about examples of optimal control problems by systems with fixed final state, discussed in the lecture.

1. Why in Example 11.1, unlike Example 5.1, which is close to it, cannot there be continuous solutions to the optimality conditions?

2. Why cannot there be continuous solutions to the optimality conditions in Example 11.1?
3. Why in Example 5.1 each number of discontinuity points in control corresponds to two solutions of optimality conditions, while in Example 11.1 this is true only for the case of one or two breaks?
4. Why does Example 11.1 have a unique solution, while the optimal control problem from Example 5.1 with the same equation of state, optimality criterion, and set U has two solutions?
5. Why some controls from Table 11.1 (see Notes), cannot be solutions of optimality conditions?
6. Why Example 11.2 cannot have regular solutions of the maximum principle?
7. Is it possible to use the method of proving the absence of regular solutions of the maximum principle from Example 11.2 to obtain a similar result for Example 6.2, which considers a fairly close problem with a free final state?
8. Formula (11.8) does not exclude the presence of several discontinuity points of control. Why, while studying it, did we limit ourselves to considering only one point of break?
9. What would change if Example 11.2 did not require control to belong to U?
10. Whence does it follow that the maximum principle for Example 11.2 is a necessary and sufficient optimality condition?
11. What conclusion can be drawn as a result of checking the validity of the Kelley condition for Example 11.2?
12. What happens if we use the optimal control uniqueness theorem from Chapter 5 for Example 11.2?
13. What happens if for Example 11.2 we use the theorem on the sufficiency of the maximum condition from Chapter 5?
14. What happens if we use the optimal control existence theorem from Chapter 7 for Example 11.2?
15. What is the fundamental difference between the properties of Examples 11.2 and 11.3?
16. What is the reason for the difference between the properties of Examples 11.2 and 11.3?
17. Why was there one singular control in Example 11.2, but three in Example 11.3?

18. What conclusion can be drawn as a result of checking the validity of the Kelley condition for Example 11.3?

19. What happens if we use the optimal control uniqueness theorem from Chapter 5 for Example 11.3?

20. What happens if for Example 11.3 we use the theorem on the sufficiency of the maximum condition from Chapter 5?

21. What happens if we use the optimal control existence theorem from Chapter 7 for Example 11.3?

22. Why are there no singular controls in Example 11.4, unlike in previous examples?

23. Why is the optimal control problem in Example 11.4 so different from previous examples?

24. On what basis is it concluded that the control u_0 for Example 11.4 corresponds to the maximum of the function H?

25. What happens if for Example 11.4 we use the theorem on the sufficiency of the maximum condition from Chapter 5?

26. Why is the only control u_0 determined from the maximum condition not optimal?

27. Whence follows the absence of optimal control for Example 11.4?

28. Is the set of admissible controls for Example 11.4 bounded?

29. Why cannot we use the optimal control existence theorems for Example 11.4?

Conclusions

Based on the study of the considered problems of optimal control systems with fixed final state, we can come to the following conclusions.

- The optimality conditions for Example 11.1 have an infinite number of solutions.
- All solutions of the optimality conditions for Example 11.1 are discontinuous.
- For Example 11.1, there are two solutions with one and two break points, and one each with any number of breakpoints greater than two.
- The optimality conditions for Example 11.1 are not sufficient.
- The optimal control problem for Example 11.1 has a unique solution.

- For Example 11.2, the function H included in the optimality conditions is linear with respect to control, as a result of which optimal control can be singular or take values on the boundary of the set of admissible controls.
- The presence of a fixed final state for Example 11.2 allows us to exclude the regular solution maximum conditions from consideration.
- For Example 11.2, there is a unique singular control.
- The singular control for Example 11.2 satisfies the Kelley condition.
- The singular control for Example 11.2 is optimal.
- The optimal control problem from Example 11.2 has a unique solution.
- The maximum principle for Example 11.2 is a necessary and sufficient condition for optimality.
- For Example 11.3, the function H included in the optimality conditions is linear with respect to control, as a result of which optimal control can be singular or take values on the boundary of the set of admissible controls.
- The presence of a fixed final state makes it possible to exclude the regular solution maximum conditions from consideration in Example 11.3.
- For Example 11.3, there are three singular controls.
- One of the singular controls for Example 11.3 does not satisfy the Kelley condition, but two do.
- Two singular controls for Example 11.3 are optimal.
- The optimal control problem for Example 11.3 has two solutions.
- The maximum principle for Example 11.3 is not sufficient optimality condition.
- The function H for Example 11.4 has a unique stationary point, where the second derivative of H is negative.
- For Example 11.4, there exists a minimizing control sequence.
- For Example 11.4, a minimizing sequence of equations is constructed, which is not limited.
- The value of the optimality criterion in the found stationary point of the function H from Example 11.4 is greater than the lower bound of functional on the set of admissible controls.
- The optimal control problem from Example 11.4 has no solution.

Problems

In the process of analyzing the considered problems of optimal control for systems with fixed final state, additional problems arise that need to be studied.

1. **Singular control for maximization problems with fixed final state**. In Part II, it was noted that in the problems of minimization and maximization of the same functional, the sets of singular controls coincide. However, when changing the type of extremum, the properties of the problem can change qualitatively. It would be interesting to analyze the problems considered in the Lecture with a change in the type of extremum. Such results are given in Appendix.

2. **Optimal control problem with fixed final state in the presence of both singular and regular solutions of the maximum condition**. In all considered problems of optimal control with fixed final state, either singular controls did not exist at all, or only singular controls exist. However, for problems with free final states, we encountered a situation where the maximum condition had both singular and regular solutions at the same time. Appendix provides an example of a similar problem with fixed final state.

3. **Optimal control with fixed final state with infinite set of solution maximum principle and non-unique optimal control**. In the Lecture, the problem of optimal control with a fixed final state with an infinite set of solutions of the maximum condition under the uniqueness of the optimal control was considered. However, for problems with free finite states, we encountered a situation where the maximum condition had an infinite number of solutions, and the optimal control was not unique. Appendix gives an example of a similar problem with fixed final state.

4. **Optimal control problem with fixed final state with non-optimality of all singular control**. In all considered optimal control problems with fixed final state, when the maximum condition degenerates, at least one singular control was optimal. However, for problems with free final states, we have encountered a situation where all singular controls are not optimal. Appendix gives an example of a similar problem with fixed final state.

5. **Optimal control problem with fixed final state, in which the singular control that satisfies the Kelley condition is not optimal**. In all considered problems of optimal control with fixed final state, singular controls satisfying the Kelley condition turned out to be optimal. However, earlier we encountered a situation where the validity of the Kelley condition did not guarantee the optimality of the singular control. Appendix gives an example of a similar problem with fixed final state.

6. **Complete analysis of maximum conditions for Example 11.4**. The analysis of the optimal control problem from Example 11.4 was in fact incomplete and not entirely accurate. Appendix completes the analysis of this example.

7. **Optimal control with fixed final state that has an infinite set of solutions**. Previously, we have already encountered a situation where the optimal control problem with a free final state has an infinite number of solutions. It would be interesting to give a similar example for a system with fixed final state. Such an example is given in Appendix.

11.2 APPENDIX

We continue to consider examples of optimal control problems with fixed final state. Sections 11.2.1 and 11.2.2 deal with maximization problems for functionals with one and three singular controls. Section 11.2.3 completes the analysis of Example 11.4. Sections 11.2.4 and 11.2.5 provide examples of the considered class of optimal control problems with an infinite set of solutions and degeneration of the Kelley condition.

11.2.1 Maximization of a functional with unique singular control

Consider a problem similar to Example 11.2, but related to the maximization of a functional.

Example 11.5 *Consider the system described by the Cauchy problem*

$$x'(t) \;=\; u(t),\; t \in (0,1);\; x(0) = 0. \tag{11.18}$$

The control u is chosen from the set

$$U \;=\; \{u \,|\, |u(t)| \le 2,\; t \in (0,1)\}.$$

Besides, it guaranties the final condition

$$x(1) = 1. \tag{11.19}$$

It is required to maximize on this set the functional

$$I(u) \;=\; \frac{1}{2}\int_0^1 (x^2 - 2xt)dt.$$

Determine the function

$$H \;=\; pu–(x^2–2xt)/2,$$

where p satisfies the adjoint equation

$$p'(t) \;=\; x(t) - t,\; t \in (0,1). \tag{11.20}$$

The optimal control here can be found from the minimization of the function H. Because of its linearity with respect to the control, the solution of optimality conditions can be singular or is reached at the boundary of the set U. Singular control here corresponds to the equality $p(t) = 0$, that is realized for $x(t) = t$. This state satisfies the equality (11.19) and corresponds to the admissible control $u(t) = 1$. However, this is the point of minimum of the given functional[27].

The regular solution of minimization problem for the function H is determined by the formula

$$u(t) = \begin{cases} 2, & \text{if } p(t) < 0, \\ 0, & \text{if } p(t) > 0. \end{cases} \tag{11.21}$$

If $u(t) = 2$ everywhere, then the state function is $x(t) = 2t$, hence, $x(1) = 2$ that contradicts equality (11.19). For $u(t) = 0$, we get $x(t) = 0$, so $x(0) = 0$ that contradicts equality (11.19) too. Therefore, has function u at least one discontinuity point, besides the function p change the sign there.

Let ξ be a unique discontinuity point of control, besides $u(t) = 2$ and the function p is negative for $t < \xi$ and $u(t) = 0$ and p is positive for $t > \xi$. From equation (11.20) it follows that $x(t) = 2t$ for $t < \xi$. Therefore, $x(\xi) = 2\xi$. This value does not change later, which means, $x(1) = 2\xi$. Using the condition (11.20), we find $\xi = 1/2$. Thus, $x(t) = 1–t$ for $t > 1/2$.

The adjoint equation for $t < 1/2$ is $p'(t) = x(t) - t = t$. Obviously, this derivative is positive, so the function p increases if $t < 1/2$. At $t = 1/2$, it vanishes. This means that it was initially negative, which is consistent with the choice of control. Similarly, for $t > 1/2$ we have the adjoint equation $p'(t) = x(t) - t = 1 - t$ This derivative is again positive, hence the function p increases if $t > 1/2$. Then it became positive there, which is consistent with formula (11.21). Thus, the control, which equal to 2 for $t < 1/2$ and to 0 for $t > 1/2$ satisfies, in reality, the optimality conditions.

Suppose now $u(t) = 0$ and $p(t) > 0$ for $t < \xi$, and $u(t) = 2$ and $p(t) < 0$ for $t > \xi$. Then $x(t) = 0$ for $t < \xi$. Therefore, $x(\xi) = 0$. Solving the state equation with this initial condition, we find $x(t) = 2(t–\xi)$. Particularly, $x(1) = 2(1–\xi)$. Using equality (11.20), we find $\xi = 1/2$. Thus, $x(t) = 2t–1$ for $t > 1/2$. The adjoint equation is $p'(t) = x(t) - t = -t$ for $t < 1/2$. The function p decreases on first interval with zero value in the middle of this interval. Therefore, initially this function is negative. For $t > 1/2$, the adjoint equation is $p'(t) = x(t) - t = t - 1$. This derivative is negative, and the function p decreases if $t > 1/2$. However, it is vanishing at the point $t = 1/2$. Therefore, in what follows, this function is negative, which is consistent with the choice of control. Hence, the control, which is equal to 0 for $t < 1/2$ and 2 for $t > 1/2$ also satisfies the system of optimality conditions.

Let us denote the first of the found solutions by u_1^2, and the second one by u_1^0, and the corresponding states of the system by x_1^2 and x_1^0. In order to choose the best of these controls, we find the corresponding values of functional. Note the equality

$$I = \frac{1}{2}\int_0^1 (x-t)^2 dt - \frac{1}{2}\int_0^1 t^2 dt = \frac{1}{2}\int_0^1 (x-t)^2 dt - \frac{1}{6}.$$

Determine the value of the integral at the right-hand side of this equality, which is denoted by J. We have

$$J(u_1^2) = \int_0^1 (x_1^2 - t)^2 dt = \int_0^{1/2} t^2 dt + \int_{1/2}^1 (1-t)^2 dt = \frac{1}{12},$$

$$J(u_1^0) = \int_0^1 (x_1^0 - t)^2 dt = \int_0^{1/2} (-t)^2 dt + \int_{1/2}^1 (t-1)^2 dt = \frac{1}{12}.$$

Thus, found two solutions are equivalent.

Suppose now there exists two discontinuity points ξ and η of control. Suppose $u(t) = 2$ for $0 < t < \xi$ and $\eta < t < 1$ and $u(t) = 0$ for $\xi < t < \eta$. Then function x successively takes the values $2t$, 2ξ, and $2\xi – 2\eta + 2t$ on the three considered intervals. Using condition (11.20), we have $x(1) = 2\xi – 2\eta + 2 = 1$, which implies $\eta = \xi + 1/2$. Thus, there are an infinite number of options for two points of discontinuity[28]. However, one should check the signs of the function p at each of the corresponding intervals.

Suppose $u(t) = 2$ for $0 < t < \xi$. Then $x(t) = 2t$, and the function p satisfies the adjoint equation $p'(t) = t$. Therefore, its derivative is positive, and p increases. Using the equality $p(\xi) = 0$, we conclude that this function is negative for $0 < t < \xi$. This is consistent with equality (11.21).

Then we have $u(t) = 0$ for $\xi < t < \xi + 1/2$. We get $x(t) = x(\xi) = 2\xi$, and the adjoint equation is $p'(t) = 2\xi - t$. Integrate this equality by the considered interval. Considering that function p vanishes on its boundaries, we have

$$0 = \int_{\xi}^{\xi+1/2} (2\xi - t) dt = \Big(2\xi t - \frac{t^2}{2}\Big)\Big|_{\xi}^{\xi+1/2} = \xi - \Big(\xi + \frac{1}{4}\Big)\frac{1}{2} = \frac{\xi}{2} - \frac{1}{8}.$$

Now we find $\xi = 1/4$. Therefore, for $1/4 < t < 3/4$ we get the adjoint equation $p'(t) = 1/2 - t$. This derivative is positive, hence the function p increases from zero, and this is positive. The derivative is equal to zero at the point $t = 1/2$, where the function p has. Then it decreases and reaches zero at $t = 3/4$. Thus, it is positive throughout this interval, which also agrees with equality (11.21).

If $t > 3/4$, then the control is $u(t) = 2$, and $x(t) = 2t – 1$. Then we have the adjoint equation $p'(t) = t - 1$. Therefore, function p decreases from zero, and this is negative. This result again agrees with equality (11.21). Thus, the third solution of the optimality conditions is found.

Determine the value of the functional J at this control u_2^2, to which the state defined above corresponds x_2^2. We have

$$J(u_2^2) = \int_0^1 (x_2^2 - t)^2 dt = \int_0^{1/4} t^2 dt + \int_{1/4}^{3/4} \Big(\frac{1}{2} - t\Big)^2 dt + \int_{3/4}^1 (t-1)^2 dt = \frac{1}{48}.$$

The resulting value is less than those that were set earlier, which means that the control optimal is not optimal. We can also find the solution u_2^0, which is equal to 0 on intervals $(0, 1/4)$ and $(3/4, 1)$ to 2 on (1/4,3/4). This is not optimal too, and the corresponding value of the functional J is 1/48.

The existence of solution u_k^2 for optimality conditions for is established in a similar way. It is defined as follows. The segment [0,1] is divided into $2k$ equal parts. On the first of them, control is assumed to be 2, on the next two this is equal 0, on the next two this is 2, etc., Finally, on the last this is 2. Solution u_k^0 is obtained from the previous one by swapping the values 2 and 0. The corresponding system states are shown in Figure 11.3. In this case, we determine the values of functional

$$J(u_k^2) = J(u_k^0) = \frac{1}{3}\frac{1}{(2k)^2}, \; k = 1, 2, \dots .$$

Thus, the values of this functional decrease with increasing number of discontinuity points[29].

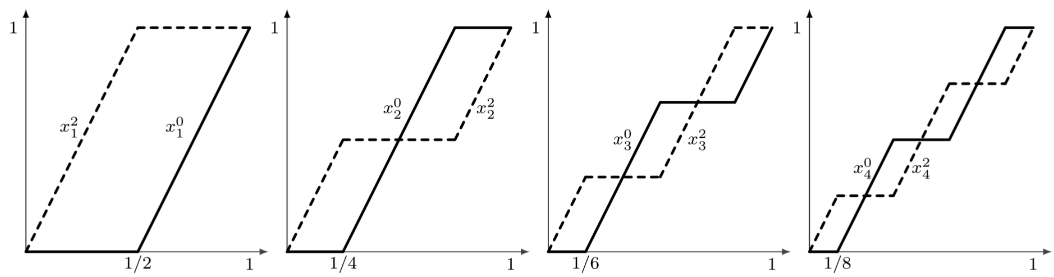

Figure 11.3 State sequence for Example 11.4.

Thus, for Example 11.4, the maximum principle gives a necessary, but not sufficient optimality condition. In this case, there is a unique singular control, which is not optimal, and an infinite set of regular solutions of the optimality conditions that differ in the number of discontinuity points[30] (two for each number of jump). Two of them are optimal, having a unique discontinuity point.

11.2.2 Maximization of a functional with three singular controls

Consider the maximization problem for the functional from Example 11.3. For simplicity, we confine ourselves to considering the case when there are no restrictions on the values of control.

Example 11.6 *Consider the system described by the Cauchy problem*

$$x'(t) = u(t), \; t \in (0, 1); \; x(0) = 0.$$

The control u is chosen from the set of functions that enforce the condition

$$x(1) = 0.$$

It is required to maximize on this set the functional

$$I(u) = \frac{1}{4}\int_0^1 (x^4 - 2x^2 \sin^2 \pi t)dt.$$

The optimal control is the point of minimum of the function

$$H(t,u,x,p) = pu - \frac{1}{4}(x^4 - 2x^2 \sin^2 \pi t),$$

where p is the solution of the adjoint equation

$$p'(t) = x(t)^3 - x(t)\sin^2 \pi t,\ t \in (0,1).$$

By linearity of the function H and the absence of the direct limitation for the controls, the solution of the optimality conditions can be only singular[31].

As we already know, the singular controls in maximization and minimization problems are the same. Thus, have again three singular controls

$$u_1(t) = 0,\ u_2(t) = -\pi \cos \pi t,\ u_3(t) = \pi \cos \pi t.$$

For the maximization problem, the sign in the Kelley condition should be changed. If in Example 11.3 the first singular control did not satisfy the Kelley condition, but the second and third ones did and turned out to be the result of optimal control, then in this case the situation is opposite. Thus, only control u_1 satisfies the Kelley condition. Being the only singular solution of the maximum principle that satisfies the Kelley condition, it could have optimal control. However, this will only be the case if the optimal control problem is solvable.

As noted in Section 11.1.3, the optimality criterion can be written as

$$I = \frac{1}{4}\int_0^1 (x^2 - \sin^2 \pi t)^2 dt - \frac{1}{4}\int_0^1 \sin^4 \pi t dt.$$

Consider the sequence

$$x_k(t) = k \sin \pi t,\ k = 1, 2, \dots .$$

There are the states of the system for the controls $u_k(t) = -k\pi \cos \pi t$; they satisfy also the final condition $x_k(1) = 0$. Determine the corresponding value of the functional

$$I_k = \frac{(k^2-1)^2 - 1}{4}\int_0^1 \sin^4 \pi t dt.$$

Obviously, this value can be arbitrarily large for sufficiently large. Thus, the considered functional is upper unbounded on the set of admissible controls, which transform the system to the given final state. Therefore, the problem of its maximization has no solution[32].

11.2.3 Completion of the analysis of Example 11.4

Return to the analysis of Example 11.4. This is the minimization of the functional

$$I = \int_0^1 \sqrt[4]{1+u^2}dt$$

on the set of functions $u = u(t)$ such that the solution x of the problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

satisfies the equality $x(1) = 1$.

We considered the function

$$H = pu - \sqrt[4]{1+u^2}.$$

From equality to zero its derivative was found the only control $u_0 = 1$, under which the given final condition is guaranteed, and the second derivative of H at this point is negative. However, this problem does not have a solution, as a result of which it was concluded that the optimality condition is not sufficient.

Let us now try to check the sufficiency of the maximum condition directly, as was done in Chapter 5. To do this, consider the remainder term in the formula for the increment functional. In accordance with formula (5.12), it is equal to

$$\eta = \eta_3 - \int_0^1 (\eta_1 + \eta_2)dt.$$

Here η_3 corresponds to the second order term obtained as a result of transforming the part of the optimality criterion that characterizes the terminal term functional, η_1 is associated with the second order terms when expanding the value $H(t, u, x + \Delta x, p)$ into a series in Δx, and $\eta_2 = [H_x(t, v, x, p) - H_x(t, u, x, p)]\Delta x$. For this example, there is no terminal term in the optimality criterion, and function H does not depend on x. This means that $\eta = 0$. However, in this case, the optimality conditions will certainly be necessary and sufficient.

It was previously established that the considered optimal control problem has no solution, which means that the set of its solutions is empty. If the optimality condition is necessary and sufficient, then the set of its solutions exactly coincides with the set of optimal controls, and therefore is certainly empty. The results obtained force us to reconsider some of the conclusions that we made in the process of studying the optimality conditions for the problem posed.

Reducing the derivative of the function H to zero, we determined the equality

$$p - \frac{u}{2(1+u^2)^{3/4}} = 0,$$

where p is constant. Note that the zero values of the functions u and p satisfying the last equality cannot be the solution of the optimality conditions, since the zero control corresponds to the zero state of the system, which contradicts the condition $x(1) = 1$.

The resulting relation can be written as

$$cu = f(u), \tag{11.22}$$

where $c = 1/2p$, $f(u) = (1 + u^2)^{3/4}$. Depending on the values of the constant c, the algebraic equation (11.22) may have one, two, or no solutions; see Figure 11.4. The constant c itself is determined by the solution of the adjoint system, which is constant, and must be such that the control corresponding to it ensures the fulfillment of the condition $x(1) = 1$.

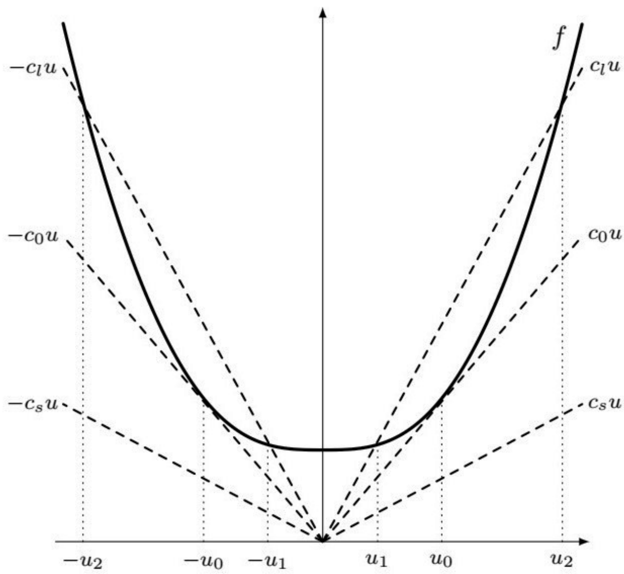

Figure 11.4 The equation (11.22) has one, two or no solution.

If the value of c is sufficiently small in absolute value (in Figure 11.4, this case corresponds to one of the equalities $c = c_s$ or $c = -c_s$, then equation (11.22), and hence the system of optimality conditions, does not have a solution at all. The only solution equation (11.22) is realized for two values of the constant c that are a positive number c_0 and the negative one $-c_0$. A positive constant gives us some constant control u_0 (see Figure 11.4), which corresponds to the solution $x(t) = u_0 t$ of the state equation, which takes the value $x(1) = u_0$ at the point $t = 1$. Then the equality $x(1) = 1$ is possible only for $u_0 = 1$. Therefore, we have obtained the previously defined control u_0. If equality $c = -c_0$ is true, then the corresponding control is negative, which means function x is decreasing, the specified equality will not be able to be performed.

Consider now the case where the constant c is so large in modulus that equation (11.22) has two solutions. Obviously, if one of the found constants is chosen as control, then as a result of consideration of the condition $x(1) = 1$, we either do not get a result at all (control is negative), or we get a known value u_0 (control is positive). However, it is possible that control is piecewise constant. Moreover, it can take such values u_1 and u_2 that correspond to two solutions of equation (11.22), but to the same solution of the adjoint equation. As Figure 11.4 shows, both of these values must have the same sign. Since control, which takes only negative values, ensures that the function x decreases, and therefore cannot lead to the validity of condition (11.22), we focus exclusively on positive values of u_1 and u_2.

Suppose there exists a point $\xi \in (0,1)$ such that

$$u(t) = \begin{cases} u_1, & \text{if } t < \xi, \\ u_2, & \text{if } t > \xi. \end{cases} \tag{11.23}$$

The corresponding state function is

$$x(t) = \begin{cases} u_1 t, & \text{if } t < \xi, \\ (u_1 - u_2)\xi + u_2 t, & \text{if } t > \xi. \end{cases} \tag{11.24}$$

Using now the final condition

$$x(1) = (u_1 - u_2)\xi + u_2 = 1,$$

we get the value

$$\xi = \frac{u_2 - 1}{u_2 - u_1},$$

which must belong to the interval (0,1). Therefore, we have the inequality $u_1 < 1 < u_2$ for $u_1 < u_2$ and the condition $u_2 < 1 < u_1$ if $u_2 < u_1$.

For definiteness, we denote by u_1 the smallest of the values of u_1 and u_2. Then the existence of a unique point ξ, for which control, determined by the formula (11.23) ensures the validity of the final condition, is possible only when the inequality $u_1 < 1 < u_2$ is satisfied. As we can see from Figure 11.4, this relationship does indeed hold, since equality $u_0 = 1$ holds true. Thus, the triple of functions, including control and the state, determined by formulas (11.23) and (11.24) with the above value ξ, as well as the corresponding function p, really give a solution to the system, which includes the equation of the state with two given boundary conditions, adjoint equation and equality (11.22); see Figure 11.5.

The values included in formula (11.23) are uniquely determined from equation (11.22) for a given value of the parameter c. Obviously, for any value $c > c_0$ there is a unique pair u_1, u_2 that satisfies the necessary requirements. The constant c_0 can be determined from equality (11.22) if $u = u_0 = 1$. As a result, we find the value $c_0 = 2^{3/4}$. Thus, each of the values $c > 2^{3/4}$ corresponds to a unique pair of positive numbers u_1 and u_2, which determine the solution of the specified system in accordance with formulas (11.23) and (11.24); see Figure 11.5. Thus, this system has an infinite and even uncountable set of solutions.

Assuming that control, which takes only the values u_1 and u_2, has two discontinuity points, we get a new solution of the system. Characteristically, for each constant c exceeding c_0, one can find its own pair of u_1 and u_2, and hence a new solution of the

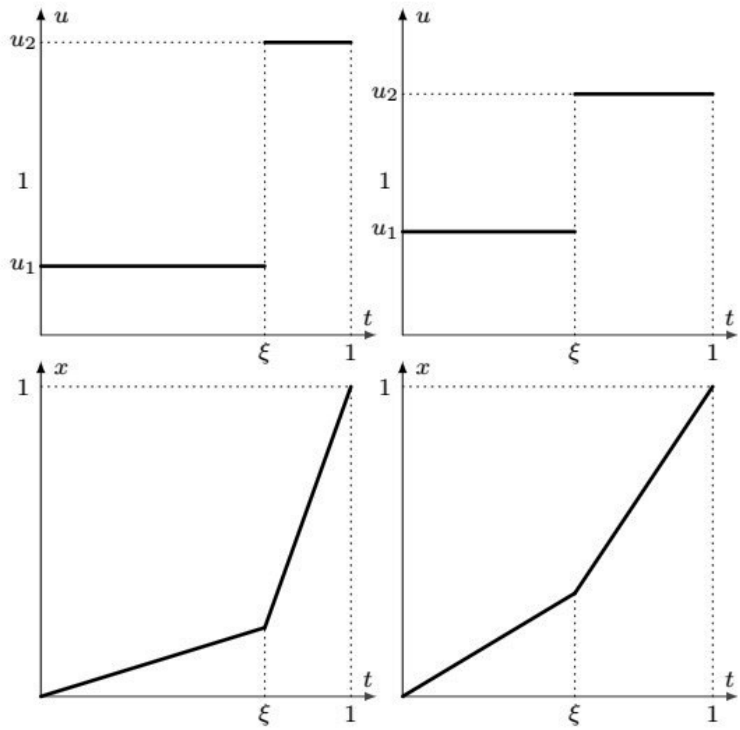

Figure 11.5 Possible controls with unique break point and the corresponding state functions.

specified system. Moreover, having one such solution, one can shift the discontinuity point ξ and η. Thus, so that the distance between them remains unchanged, and they themselves belong to interval (0,1). The result is a new control with two breakpoints, also satisfying the specified system; see Figure 11.6. Finally, the option of three or more check breaks is not ruled out.

It should be noted that so far, we have been talking about a system that includes a state equation with two boundary conditions, adjoint equation and equality (11.22), which, as it turned out, has an infinite number of solutions. However, it is far from obvious that the solutions found by her really correspond to the maximum principle for Example 11.4.

Obviously, the function of one variable $H = pu - \sqrt[4]{1+u^2}$ for all fixed p can take on an arbitrarily large value, which means that point does not have a global maximum at all. Thus, the condition maximum does not have a solution, which means that it is indeed a necessary and sufficient optimality condition for the considered example.

11.2.4 Problem with infinite set of solutions

Consider the following optimal control problem.

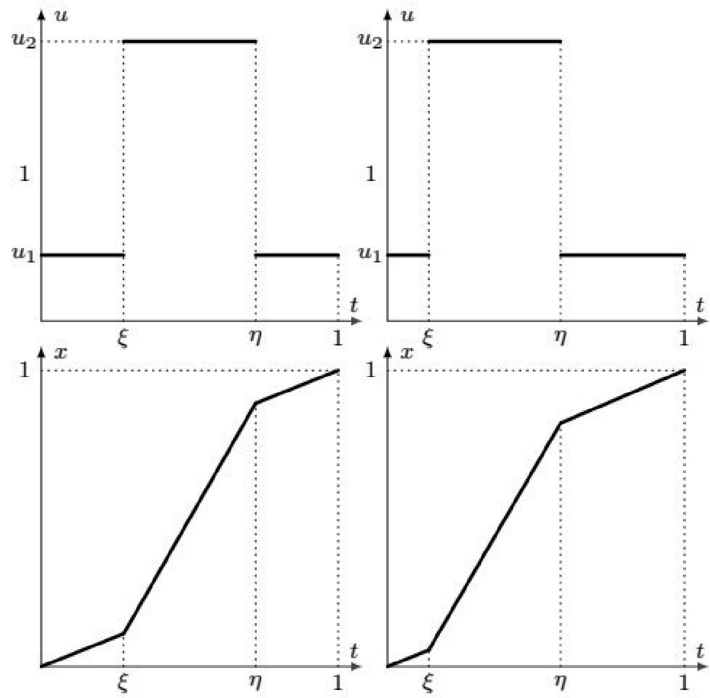

Figure 11.6 Possible controls with two break points and the corresponding state functions.

Example 11.7 *We have the system described by the Cauchy problem*

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{11.25}$$

The control u is chosen from the set

$$U = \{u \mid |u(t)| \leq 1,\ t \in (0,1)\}.$$

Besides, it guaranties the final condition

$$x(1) = 0. \tag{11.26}$$

It is required to maximize on this set the functional

$$I(u) = \int_0^1 u(t)x(t)dt.$$

This problem differs from the one considered in Example 6.1 only by the presence of a fixed final state. Denote by V is such a subset of controls from U that the state of the system corresponding to them satisfies condition (11.26). The given example implies the problem of minimizing functional I on a subset V, in contrast to Example

6.1, where the problem of minimizing the same functional on the entire set U was studied without fixing the final state of the system. Naturally, we have the inequality[33]

$$\min I(U) \leq \min I(V).$$

Thus, if a solution of first problem (Example 6.1) belongs to the set V, then it is the solution of this example.

When analyzing Example 6.1, it was previously shown that the optimal controls for it are all functions from the set U that satisfy the equality

$$\int_0^1 u(t)dt = 0. \tag{11.27}$$

Solving the problem (11.25), we get

$$x(1) = \int_0^1 u(t)dt.$$

Thus, condition (11.25) will be satisfied by those and only those controls that satisfy equality (11.27). Thus, all optimal controls for Example 6.1 belong to the set V, which means that there will be optimal controls for Example 11.7 as well. Therefore, the considered problem has an infinite and even uncountable set of solutions[34].

11.2.5 Problems with degeneracy of the Kelley condition

In Chapter 6, we considered a situation where not only the maximum principle, but also the Kelley condition degenerated. Let us give a similar example for a system with fixed final state.

Example 11.8 *We have the system described by the Cauchy problem*

$$x_1'(t) = x_2(t),\ x_2'(t) = u(t),\ t \in (0,1);\ x_1(0) = x_2(0) = 0. \tag{11.28}$$

The control u is chosen from the set

$$U = \{u \,|\, |u(t)| \leq 1,\ t \in (0,1)\}.$$

Besides it guaranties the final condition

$$x_1(1) = 0. \tag{11.29}$$

It is required to maximize on this set the functional

$$I(u) = \frac{1}{2}\int_0^1 (x_1^4 - x_1^2 \sin^2 \pi t)dt.$$

Consider also the corresponding maximization problem.

Example 11.9 *It is required to maximize the functional from Example* 11.8 *on the same set of controls.*

Determine the function

$$H = p_1 x_2 + p_2 u - (x_1^4 - x_1^2 \sin^2 \pi t)/2,$$

where p_1 and p_2 satisfy the equalities

$$p_1'(t) = 2x_1(t)^3 - 2x_1(t)\sin^2 \pi t, \;\; p_2'(t) = -p_1(t), \; t \in (0,1); \; p_2(1) = 0. \qquad (11.30)$$

The degeneration of the maximum principle here is possible for $p_2 = 0$. Then it follows from the second equation (11.30) that $p_1 = 0$, and from the first one that $x_1^3 - x_1 \sin^2 \pi t = 0$. The resulting cubic equation has three solutions: $x_1 = 0$, $x_1 = \sin \pi t$, and $x_1 = \sin \pi t$. Substituting these values into the first equation (11.28), we find the corresponding functions $x_2 = 0$, $x_2 = \pi \cos \pi t$, and $x_2 = \pi \cos \pi t$. Now it follows from the first equation (11.28) that for the considered example there are three special controls $u = 0$, $u = -\pi^2 \sin \pi t$, and $u = -\pi^2 \sin \pi t$. Note that they all belong to the set U and ensure the fulfillment of condition (11.29).

Let us verify the validity of the Kelley condition for the found singular controls. We have

$$\frac{\partial H}{\partial u} = p_2.$$

Using the adjoint equation, we get

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p_2' = -p_1.$$

After differentiation with using the first adjoint equation, we find

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = -p_1' = -2x_1(t)^3 + 2x_1(t)\sin^2 \pi t.$$

Now we have

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 0,$$

a.e. the Kelley condition degenerates.

To verify the validity of the Kopp–Moyer condition, we continue the differentiation

$$\frac{d^3}{dt^3}\frac{\partial H}{\partial u} = -6x_1^2 x_1' + 2x_1' \sin^2 \pi t + x_1 \pi \sin 2\pi t = -6x_1^2 x_2 + 2x_2 \sin^2 \pi t + x_1 2\pi \sin 2\pi t$$

because of the first equality (11.30). Using the second equality (11.30), determine

$$\frac{d^4}{dt^4}\frac{\partial H}{\partial u} = -12x_1 x_2^2 - 6x_1^2 u + 2u \sin^2 \pi t + 4\pi x_2 \sin 2\pi t + 4\pi^2 x_1 \cos 2\pi t.$$

Finally, we find the value

$$\frac{\partial}{\partial u}\frac{d^4}{dt^4}\frac{\partial H}{\partial u} = -6x_1^2 + 2\sin^2 \pi t,$$

that is not zero. Put here the known singular controls, we get

$$\frac{\partial}{\partial u}\frac{d^4}{dt^4}\frac{\partial H}{\partial u}\Big|_{u=0} = 2\sin^2 \pi t,\ \frac{\partial}{\partial u}\frac{d^4}{dt^4}\frac{\partial H}{\partial u}\Big|_{u=\pm\pi^2\sin^2 \pi t} = -(6\pi^2 - 2)\sin^2 \pi t.$$

Thus, for Example 11.7 the Copp–Moyer condition holds for the singular controls $u = -\pi^2 \sin \pi t$, and $u = -\pi^2 \sin \pi t$, and fails for control $u = 0$, and vice versa for Example 11.8. Thus, in both cases, the optimality condition is not sufficient, but in Example 11.7 optimal can be control $u = -\pi^2 \sin \pi t$, and $u = -\pi^2 \sin \pi t$, and in Example 11.8 the control $u = 0$ can be optimal. It is not difficult to make sure that these special controls are indeed optimal in the first case[35], and in the second case, either a zero singular control or a regular control equal to 1 for negative values of p_2 and –1 for positive values of this function is optimal[36].

Additional conclusions

Based on the analysis of the above examples of optimal control problems by systems with fixed final state, we have the following additional conclusions.

- The maximum principle for Example 11.4 can degenerate, but the corresponding unique singular control is not optimal.
- The maximum principle for Example 11.4 has an infinite number of solutions, differing in the number of points of break.
- The maximum principle for Example 11.4 is not a sufficient optimality condition.
- The optimal control problem from Example 11.4 has two solutions, each with one discontinuity point.
- The maximum principle for Example 11.5 has three solutions that are singular controls.
- Among the three singular controls for Example 11.5, only one satisfies the Kelley condition.
- The maximum principle for Example 11.5 is not a sufficient optimality condition.
- The singular control for Example 11.5 satisfying the Kelley condition is not optimal.
- The optimal control problem for Example 11.5 has no solution.

- The remainder term in the formula of functional increment for Example 11.5 is zero.
- The maximum principle for Example 11.3 is sufficient optimality condition.
- There exists an uncountable set of stationary points of the function H for Example 11.3, which transform the system to the final state, but they are not optimal.
- The function H for Example 11.3 is not upper bounded.
- For optimal control problems with fixed final state, there may be an infinite set of singular controls.
- For optimal control problems with fixed final state, there may be a degeneration of the Kelley condition; and to check the optimality of special controls, one can use the Kopp–Moyer condition.
- The maximum principle for Examples 11.7 and 11.8 is not a sufficient optimality condition.
- The optimal control problem for Example 11.7 has two solutions, which are non-zero singular controls.

Notes

1. In contrast to the similar system (5.1)–(4.4), there is no sign change invariance here, as a result of which several other properties can be expected from this system. There is no invariance under the change of sign in the formulation of the problem itself. Particularly, if some function is a solution to the problem, then for a function taken with the opposite sign, condition (11.1) will certainly be get broken.

2. Note that for Example 5.1 these two continuous (constant) functions were optimal controls.

3. We encountered a similar result in Example 5.1, but there the two solutions of the optimality conditions differed only in sign.

4. Curiously, this control is exactly the same as function u_2^+ from Example 5.1. This is not surprising, because this is the only solution of the system of optimality conditions for the specified example that satisfies equality (11.2). Recall that the optimality conditions for Examples 5.1 and 11.1 differ only in the presence of this equality and the absence of the final condition for the adjoint equation.

5. We got here the previously discussed solution with two control break points and a value of 1 on the first interval. There is nothing surprising in this, since function p in the indicated previous variant vanished not only at points of the control break, but also at $t = 1$.

6. Table 11.1 shows the distribution of control discontinuity points that are the solution and the resulting system of optimality conditions for the considered example. In this case, the second column shows the coordinates of the discontinuity points control with a value of 1 in the first interval of continuity, and in the third column shows it with a value of –1 in this interval.

TABLE 11.1 Distribution of control discontinuity points.

Numbers of discontinuity	Coordinates of breakpoints of controls with value 1 on first interval	Coordinates of breakpoints of controls with value –1 on first interval
1	3/5	2/5
2	2/10, 6/10	3/10, 9/10
3	3/15, 9/15, **15/15**	2/15, 6/15, 10/15
4	2/20, 6/20, 10/20, 14/20	3/20, 9/20, 15/20, **21/20**
5	3/25, 9/25, 15/25, 21/25, **27/25**	2/25, 6/25, 10/25, 14/25, 18/25
6	2/30, 6/30, 10/30, 14/30, 18/30, 22/30	3/30, 9/30, 15/30, 21/30, 27/30, **33/30**

It is easy to see that in the case of n breakpoints, control, which takes the value 1 on first interval, has jumps in points

$$\frac{5+(-1)^{n+1}}{10n}(2k-1),\ k=1,2,...,n.$$

If it is equal to –1 there, then the points of discontinuity are

$$\frac{5+(-1)^{n}}{10n}(2k-1),\ k=1,2,...,n.$$

Note that in the resulting formulas, the last point of the break can have a coordinate greater than or equal to 1 (these values are highlighted in the table), which indicates that such a variant of control is inadmissible. Thus, there are two solutions of optimality conditions each with one and two discontinuity points, and one solution each with more than three discontinuity points.

7. Obviously, the square of control included in the optimality criterion for all solutions of the optimality conditions is equal to 1; however, it is for the specified control that function x reaches its maximum value. Something similar was observed in Example 5.1. However, there a continuous function was admissible and turned out to be the very solution of the problem. Moreover, due to the invariance of the system with respect to the sign change, there were two, and not one, optimal controls.

8. Characteristically, the uniqueness of the solution of the problem is realized when the uniqueness theorem presented in Chapter 5 is violated. In particular, in this case we are minimizing a functional that is not convex. The results obtained are significantly different from the properties of the related problem of optimal control with a free final state from Example 5.1, characterized by the invariance of the problem under the change of sign of control. There, each number of discontinuity points corresponds to two solutions of optimality conditions, which differ in signs, and solutions turn out to be two continuous controls. In Example 12.7, considered in the following section, we analyze the problem of optimal control, in which the corresponding homogeneous condition is used instead of the final condition (11.1). In this regard, the problem statement and optimality conditions become invariant under sign change. As a result, each number of discontinuity points corresponds to two solutions of optimality conditions, which differ in signs, and optimal are two controls with one discontinuity point.

9. Indeed, suppose now that $p(t) < 0$ for $0 < t < \xi$, where ξ is the first point in interval (0,1) such that $p(\xi) = 0$. Then on the considered interval $u(t) = -\pi$, and hence $x(t) = -2\pi t$. Substituting this value into the adjoint equation (11.7), we find the derivative

$$p'(t) = \pi t - \sin \pi t,\ t \in (0, \xi).$$

Thus, $p'(t)$ is negative on interval $(0, \xi)$. Therefore, function p decreases monotonically. Given that it is negative at this interval, we conclude that equality $p(\xi) = 0$ is impossible.

10. Existence of singular control for this example can be proved by Theorem 6.1. Note that for easier Example 6.2, where the analogical system with free final state was be considered,

we cannot so easily prove the absence of a regular solution of the maximum condition. The presence of additional condition (11.1) makes it possible to significantly simplify the analysis at this stage of the study. Chapter 15 gives an example of an optimal control problem with an isoperimetric condition for which there is a singular control.

11. In Chapter 6, Kelley condition was formulated for the optimal control problem by a system with a free final state. However, the degeneration of the condition maximum is determined by the form of the function H and is not related to the presence or absence of an additional condition at the final time.

12. For the considered example, we can prove the existence of optimal control, its uniqueness, and sufficiency of optimality conditions using Theorems 7.1, 5.1, and 5.2, respectively. To do this, it is enough to carry out the same transformations as in the analysis of Example 6.2. Note also that the obtained results apply to the problem of determining such a control u from the set

$$U = \Big\{u \,\Big|\, a(t) \leq u(t) \leq b(t),\ t \in (0,T)\Big\},$$

that transforms the system described by Cauchy problem

$$x'(t) \;=\; u(t),\ t \in (0,T):\ x(0) = 0.$$

to the state $x(T) = x_T$ with minimization of the functional

$$I \;=\; \frac{1}{2}\int\limits_0^T \big[x - z(t)\big]^2 dt$$

under following conditions

$$z(0) = x_0,\ z(T) = x_T;\ a(t) \leq z'(t) \leq b(t),\ t \in (0,T).$$

In this case, the problem has a unique solution, the optimality conditions are necessary and sufficient, and the optimal control is singular and equal to z'. We will return to consideration of Example 11.2 in Chapter 12.

13. Regular solution of the optimality conditions (if it exists) is determined by the formula $u(t) = 1$ for positive values $p(t)$ and $u(t) = -1$ for its negative values. If $u(t) = 1$ everywhere, then the solution of the problem (11.9) is $x(t) = t$, hence $x(1) = 1$. This contradicts the equality (11.10). If $u = -1$ everywhere, then $x(t) = -t$, so $x(1) = -1$. This result is also inconsistent with equality (11.10). Therefore, the regular solution of the maximum condition, if it exists, is discontinuous, and the function p changes the sign. Denote by ξ the first point from the origin in interval (0,1) where equality $p(\xi) = 0$ is satisfied. Suppose $p(t) > 0$ for $0 < t < \xi$. Then $u(t) = 1$ here, and $x(t) = t$. Putting it the adjoint equation (11.11), we find

$$p'(t) \;=\; x(t)^3 - x(t)\sin^2 \pi t, \;=\; t(t^2 - \sin^2 \pi t),\ t \in (0,\xi).$$

Obviously, $t^2 > \sin^2 \pi t$; hence, $p'(t)$ is positive on interval $(0,\xi)$. Therefore, the function p increases. Thus, this function has only positive values. In this case, the equality $p(\xi) = 0$ is impossible. It means that the supposition that the function p is positive turned out to be wrong. Suppose now, that $p(t) < 0$ for $0 < t < \xi$, where ξ is the first point from the origin in interval (0,1) such that $p(\xi) = 0$. Therefore, $u(t) = -1$ here and $x(t) = -t$. Substantiating this value to the adjoint equation, we find

$$p'(t) \;=\; x(t)^3 - x(t)\sin^2 \pi t, \;=\; -t(t^2 - \sin^2 \pi t),\ t \in (0,\xi).$$

From inequality $t^2 > \sin^2 \pi t$ it follows that the derivative $p'(t)$ is negative on interval $(0,\xi)$. Therefore, the function p decreases. Because it is negative on this, we conclude that the equality $p(\xi) = 0$ is impossible. Thus, the regular solution of optimality conditions is absent.

14. We can use Theorem 6.1 for proving the existence of singular control. By this result, it is sufficient so that the right side of the state equation looks like $f_1(t,x)\varphi(u) + f_2(t,x)$, the integrand was represented as $g_1(t,x)\varphi(u) + g_2(t,x)$, and there exists an admissible control such that $f_1(t,x(t))p(t) – g_1(t,x(t)) = 0$. For the considered example, $f_1(t,x) = 1$, $\varphi(u) = u$, $f_2(t,x) = 0$, $g_1(t,x) = 0$, $g_2(t,x) = x^4 – 2x^2 \sin^2 \pi t$. The last condition of theorem corresponds the equality $p(t) = 0$, which can be true for three admissible controls because of the equality (11.7).

15. Note that both the problem statement itself and the condition maximum are sign-invariant. Therefore, if some control u is optimal, or at least a solution of the optimality condition, then control $–u$, which differs from it, will also be so. Therefore, the number of optimal controls and solutions of the optimality condition must necessarily be even (if it is finite), if their number does not include control, which is identically equal to zero. The latter does not change when the sign is changed. In particular, a third (non-optimal) singular control is added to the two optimal controls, which is identically equal to zero.

16. Examples of optimal control problems with isoperimetric conditions and non-unique solutions will be considered in Chapters 14 and 15.

17. It is easy to see that the conditions of Theorem 5.2 on the sufficiency of the maximum condition are not realized in this case. We also note that it is possible to construct an example of the optimal control problem with m optimal singular controls. Let the system be characterized by the equalities

$$x'(t) = u(t),\ t \in (0,T); x(0) = x_0,\ x(T) = x_T.$$

Consider arbitrary differentiable functions $z_1,..., z_m$, satisfying the conditions

$$z_i(0) = x_0,\ z_i(T) = x_T,\ i = 1,...,m.$$

Then the minimization problem for the functional

$$I = \int_0^T \prod_{i=1}^m \big[x(t) - z_i(t)\big]^2 dt$$

on the set of controls for which the state of the system satisfies the above relations, has the solutions $u_i = z_i'$, $i = 1,...,m$ that are singular controls. They are also the solutions of the problem with additional condition $u \in U$ if $z_i' \in U$, $i = 1,...,m$. Corresponding examples for more difficult equations can be determined similarly.

18. A similar example is considered in [142] from the point of view of the calculus of variations; see also [170].

19. This problem differs from the one considered in Example 10.2 only in that here the integral is the root of the fourth, and not the second degree.

20. This assertion will be refined later.

21. Obviously, the second derivative is

$$\frac{\partial^2 H}{\partial u^2} = -\frac{d}{du}\left[\frac{u}{2\big(1+u^2\big)^{3/4}}\right] = -\frac{2-u^2}{\big(1+u^2\big)^{7/4}}.$$

22. In fact, this statement is erroneous. We will see this in Section 11.2.3.

23. By the properties of this sequence (see Figure 11.2), we conclude that the set of controls that transform the system to the given final state is unbounded. This means that does not exist a positive constant, which is greater than the norm of arbitrary element of this set. Particularly, the norm of element u_k of considered sequence with respect to the space $L_p(0,1)$ of integrable functions with degree p is

$$\|u_k\| = \Big(\int\limits_{(k-1)/k}^{1} k^p dt \Big)^{1/p} = k^{(p-1)/p}.$$

This is arbitrarily large when $p > 1$. Therefore, the sequence $\{u_k\}$ is not bounded. It is possible to prove the boundedness of this sequence in the space $L_1(0,1)$ of integrable functions. However, the Banach–Alaoglu theorem, which is used in the optimal control existence theorems, does not hold for it. The unboundedness of the set of admissible controls means that when studying the solvability of this problem, we cannot use Theorem 7.1. However, there is also Theorem 7.2, in which, instead of the boundedness of the set of admissible controls, the coercivity of the functional being minimized is used. This means that for $\|v_k\| \to \infty$ the corresponding sequence of functionals must increase indefinitely. However, this is not the case, at least for the considered minimizing sequence, since the corresponding sequence of functionals converges to 1. Thus, the coercivity of functional is not satisfied. Moreover, Theorem 7.2 contains the convexity property of the functional being minimized. However, the given functional is not convex, because the integrand function g, determined by the formula $g(u) = \sqrt[4]{1+u^2}$ is not convex. Thus, we also cannot use Theorem 7.2, which is quite natural, since this optimal control problem, indeed, does not have a solution. Chapter 14 will prove the existence of an optimal control for a particular example, not only in the absence of a bounded set of admissible controls, but even in the absence of its convexity.

24. Appendix will show that this is not entirely true.

25. Actually, we have already met with such a situation in Chapter 1 when minimizing the function $f(x) = x^3$. The only stationary point $x = 0$ here does not minimize the considered function.

26. In Chapter 15, we consider an example of an insolvable optimal control problem for the system with isoperimetric condition.

27. The considered functional is equal to

$$I(u) = \frac{1}{2}\int\limits_0^1 (x-t)^2 dt - \frac{1}{2}\int\limits_0^1 t^2 dt.$$

The second integral here does not depend from control, and the first one is non-negative. It is equal to zero for the state $x(t) = t$, which corresponds to the singular control $u(t) = 1$. Therefore, this control actually delivers the minimum, not the maximum, of the given functional. The non-optimality of this singular control can also be proved using the Kelley condition.

28. The result obtained indicates that at half of the specified interval (0,1) control should take the value 0, and at half this is 2. A similar pattern can be traced for solutions of the optimality condition with a large number of point of discontinuity.

29. In a certain sense, this optimal control problem resembles the one considered in Example 5.1.

30. In Part II, we turned to the method of elimination in order to exclude two of their three unknown functions u, x, and p from the system of optimality conditions and obtain a boundary

value problem for a differential equation with respect to one of the unknowns. One can do the same for problem (11.18)–(11.21). It can be reduced to the boundary value problem

$$p''(t) \;=\; F(p(t)) - 1,\; t \in (0,1),\; p(0) = 0,\; p'(1) = 0,$$

where $F(p)$ denotes the value on the right side of equality (11.21). The resulting second boundary value problem for a second-order differential equation has an infinite set of solutions that are discontinuous (piecewise constant) functions.

31. If we have extra the set of admissible controls $U = \{u|\;|u(t)| \leq 1,\; t \in (0,1)\}$, then in addition to the three special controls, there is an infinite number of regular solutions to the optimality condition, similar to those in Example 11.4. They differ in the number of control breakpoints and in which of the values 1 or –1 is realized at the initial moment of time.

32. Naturally, if control is required to belong to a bounded set, then the defined control u_k will no longer be valid. The optimality criterion will turn out to be bounded from above, and regular optimality conditions appear.

33. The minimum of the functional on the whole set does not exceed its minimum on the subset.

34. Another example of an optimal control problem with an infinite set of solutions is given in Chapter 15. In this case, some isoperimetric condition is additionally imposed on the system.

35. The given optimality criterion can be written as

$$I(u) \;=\; \frac{1}{2}\int_0^1 \big[x_1(t) - \sin \pi t\big]^2 \big[x_1(t) + \sin \pi t\big]^2 dt \;-\; \frac{1}{2}\int_0^1 \sin^4 \pi t.$$

Here, the second integral does not depend on control, and the first one is non-negative. It can only vanish on the specified non-zero singular controls.

36. In this case, it does not matter which control is optimal in Example 11.8. The main thing is that we are dealing with problems in which there are three singular controls, and for testing them for optimality, it turns out that not the Kelley condition, but the Kopp–Moyer condition is effective.

CHAPTER 12

Ill-posed optimal control problems with a fixed final state

We continue to consider optimal control problems for systems described by ordinary differential equations with given initial and final states. In this case, it was required to minimize some integral functional in the presence of certain restrictions on the control values. This chapter deals with a range of issues related to the concept of well-posedness. Examples of such problems are given that are ill-posed in the sense of Tikhonov and Hadamard, and also in the presence of an extremal bifurcation.

12.1 LECTURE

The subject of this lecture is the study of questions of the well-posedness of optimal control problems for systems with a fixed final state. In particular, Section 12.1.1 gives examples of such problems that are Tikhonov and Hadamard well-posed. Subsections 12.1.2 and 12.1.3 explore the corresponding ill-posed problems. Subsections 12.1.4 and 12.1.5 describe examples of optimal control problems for systems with a fixed final state in the presence of parameters that are characterized by bifurcation of extremals.

12.1.1 Well-posed optimal control problems with a fixed final state

For problems of optimal control of systems with a free final state in Chapter 8, the concept of Tikhonov well-posedness was defined, which implies the convergence of any minimizing sequence to the optimal control. Consider an analog of this problem for a system with a fixed final state.

Example 12.1 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt$$

DOI: 10.1201/9781003398585-12

on a subset of such functions $u = u(t)$ *from the set*

$$U = \{u \,|\, |u(t)| \leq 1,\ t \in (0,1)\},$$

which for solutions of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

guarantee the realization of the condition

$$x(1) = 0.$$

This problem differs from the one considered in Example 3.3 only by the presence of a final condition. In Chapter 8, sufficient conditions were given for the Tikhonov well-posedness of the optimization problem. In particular, according to Theorem 8.1, if the problem of minimizing a strongly convex functional on a convex subset of a Hilbert space has a solution, then it is Tikhonov well-posed.

It is clear that the solution of the considered problem under is a control that is equal to zero, i.e., optimal control exists. The space $L_2(0,1)$ is Hilbert. The optimality criterion in this case is minimized on a subset of U, guaranteeing the realization of the final condition. The convexity of this subset is set in the same way as the similar result[1] for Example 9.1. The strong convexity of the functional to be minimized was proved in Chapter 8. Thus, this optimal control problem turns out to be well-posed in the sense of Tikhonov by virtue of Theorem 8.1.

Along with Tikhonov well-posedness, Chapter 8 also considered the Hadamard well-posedness of optimization problems, which implies the existence of a unique solution that continuously depends on the parameter. In particular, the problem from Example 8.1 was Hadamard well-posed. Consider an analog of this problem for a system with a fixed final state.

Example 12.2 *It is required to minimize the functional*

$$I = I(\mu, u) = \frac{1}{2}\int_0^1 [u^2 + (x-\mu)^2]\,dt$$

for a fixed value of the parameter μ *on a subset of such functions* $u = u(t)$ *from the set*

$$U = \{u \,|\, |u(t)| \leq 1,\ t \in (0,1)\},$$

which for solutions of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

guarantee the realization of the condition

$$x(1) = 0.$$

According to Theorem 8.2, if for any fixed value μ from a set M, the problem of minimizing the functional $I = I(\mu, u)$ on the set U is Tikhonov well-posed, and the mapping $\mu \to I(\mu, u)$ is continuous on M uniformly in $u \in U$, then this problem is Hadamard well-posed. The Tikhonov well-posedness of this problem for a fixed value of the parameter μ is established in the same way as the similar result for Example 12.1, and the uniform continuity of the functional with respect to the parameter was established when analyzing Example 8.1. Thus, the optimization problem from Example 12.2 is Hadamard well-posed.

12.1.2 Tikhonov ill-posed problem

Let us now turn to ill-posed optimal control problems. In particular, the optimal control problem from Example 6.2 turned out to be Tikhonov ill-posed. Consider a similar example for a system with a fixed final state.

Example 12.3 *It is required to minimize the functional*

$$I(u) = \frac{1}{2}\int_0^1 x^2 dt$$

on a subset of such functions $u = u(t)$ from the set

$$U = \{u \in L_2(0,1) \mid |u(t)| \le 1,\ t \in (0,1)\},$$

which for solutions of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0 \tag{12.1}$$

guarantee the realization of the condition

$$x(1) = 0. \tag{12.2}$$

This problem differs from the one considered in Example 6.2 only by the presence of an additional condition (12.2). For its analysis, we use the standard method. Here the function H is defined by the formula

$$H = pu - x^2/2,$$

where p is a solution of the equation

$$p'(t) = x(t),\ t \in (0,1). \tag{12.3}$$

As in the examples in the preceding chapter, the solution to the maximum condition is either reached at the boundary of the set U or is a singular control.

The regular solution of the maximum condition is characterized by the equality

$$u(t) = \begin{cases} 1, & \text{if } p(t) > 0, \\ -1, & \text{if } p(t) < 0. \end{cases} \tag{12.4}$$

Assume that $u(t) = 1$ over the entire interval (0,1). Then the solution of the Cauchy problem (12.1) is $x(t) = t$, and hence $x(1) = 1$. However, this contradicts condition (12.2). On the other hand, for $u(t) = -1$ we find $x(t) = -t$, and hence $x(1) = -1$. This result also disagrees with equality (12.2). Consequently, the control characterized by formula (12.4) can only be discontinuous, which is possible when the function p changes sign at least at one point in the interval (0,1).

Denote by ξ the first point from the origin of coordinates from the interval (0,1), where the equality $p(\xi) = 0$ is fulfilled. For definiteness, we assume that $p(t) > 0$ for $0 < t < \xi$. Then on this interval $u(t) = 1$, and hence $x(t) = t$. Substituting this value into the adjoint equation (12.3), we find the derivative $p'(t) = t$ for $0 < t < \xi$. This implies that the function p is monotonically increasing on the interval under consideration. Considering that it vanishes at the point ξ, we conclude that for $t < \xi$ it is negative. However, according to equality (12.4), the function p must be positive at all points t, where $u(t) = 1$. If, on the contrary, $p(t) < 0$ for $0 < t < \xi$, then $u(t) = -1$, and hence $x(t) = -t$. Then the adjoint equation has the form $p'(t) = -t$. Thus, the function p on the interval $(0, \xi)$ and reaches zero at the point ξ. This is possible if it is positive, which again contradicts condition (12.4). Thus, the function p cannot be either positive or negative on any initial subinterval from (0,1). Thus, the solution of the maximum principle cannot be represented in the form (12.4), i.e., can not be regular[2].

A singular control is implemented if the coefficient before the control in the definition of the function H is equal to zero. This corresponds to the equality $p(t) = 0$. Then equality (12.3) implies $x(t) = 0$. Note that this function x also satisfies the final condition (12.2). Substituting this function into problem (12.1), we find the control $u_0(t) = 0$. It belongs to the set U and is the optimal control for Example 12.3. The value of the optimality criterion corresponding to it is equal to $I(u_0) = 0$.

Consider now the sequence of controls characterized by the equalities

$$u_k(t) = \cos k\pi, \; k = 1, 2, \dots .$$

The corresponding solutions to problem (12.1) have the form

$$x_k(t) = \int_0^t \cos k\tau) d\tau = \frac{\sin \pi k t}{\pi k}, \; k = 1, 2, \dots .$$

Note that they all satisfy condition (12.2), and hence the sequence $\{u_k\}$ is admissible. We find the corresponding values of the minimized functional

$$I(u_k) = \frac{1}{2} \int_0^1 x_k^2 dt = \frac{1}{2k^2\pi^2} \int_0^1 \sin^2 \pi k t dt = \frac{1}{4k^2\pi^2}.$$

Obviously, $I(u_k) \to 0$ as $k \to \infty$. Thus the sequence $\{u_k\}$ is a minimizing.

Let us check the convergence of this sequence to the optimal control. Calculate the value

$$\|u_k - u_0\|^2 = \int_0^1 |u_k(t) - u_0(t)|^2 dt = \int_0^1 \cos^2 \pi k t dt = \frac{1}{2}.$$

Therefore, the minimizing sequence does not converge to the optimal control, which means that the considered problem is not Tikhonov well-posed[3].

12.1.3 Hadamard ill-posed problem

Let us give an example of a Hadamard ill-posed optimal control problem for a system with a fixed final state.

Example 12.4 *It is required to minimize the functional*

$$I_k(u) = \int_0^1 (x - y_k)^2 dt,$$

where $y_k(t) = (k\pi)^{-1} \sin k\pi t$*,* k *is a numerical parameter (natural number) on a subset of such functions* $u = u(t)$ *from the set*

$$U = \{u \,|\, |u(t)| \le 1,\ t \in (0,1)\},$$

which for solutions of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

guarantee the realization of the condition

$$x(1) = 0.$$

Using the technique described in the previous chapter, we establish that this problem has a unique solution, the maximum principle gives the necessary and sufficient optimality condition, and the solution here is a singular control, in which the functional vanishes. This corresponds to the function

$$u_k(t) = y_k'(t) = \cos k\pi t.$$

Obviously, it belongs to the set U and guaranties the equality $x(1) = 0$.

Passing to the limit in the formula defining the optimality criterion, we find the value

$$I_\infty = \lim_{k\to\infty} I_k = \int_0^1 x^2 dt.$$

The problem of minimizing this functional on a subset U of controls that ensure the fulfillment of a given final condition corresponds to Example 12.3 and has a solution $u_\infty = 0$.

Find the value

$$\|u_k - u_\infty\|^2 = \int_0^1 \left|u_k(t) - u_\infty(t)\right|^2 dt = \int_0^1 \cos^2 k\pi t dt = \frac{1}{2}.$$

Thus, the optimal control u_∞ for the limit problem is not the limit of the sequence $\{u_k\}$ of solutions to the original problem. Therefore, the solution of the problem under consideration is not continuous in the parameter, which means that the problem is not Hadamard well-posed[4].

12.1.4 Bifurcation of extremals

When analyzing in Chapter 2 problems of minimizing a function depending on a parameter, we encountered a situation where, with a small change in the parameter, the solution of the problem does not just change by a considerable amount, which corresponds to the absence of Hadamard well-posedness. In the process of changing the parameter of the problem, the number of stationary points of the function under study may change. This phenomenon, called bifurcation, also arises in problems of minimizing functionals. Consider an optimal control problem with a fixed final state and a functional depending on the parameter[5].

Example 12.5 *It is required to minimize the functional*

$$I(u) = \int_0^\pi \Big[\frac{u^2}{2} - \mu(1-\cos x)\Big]dt$$

on a set of such functions $u = u(t)$, which transform the system characterized by the equalities

$$x'(t) = u(t),\ t \in (0,\pi);\ x(0) = 0 \tag{12.5}$$

to the state

$$x(\pi) = 0, \tag{12.6}$$

where μ is a positive constant that is a problem parameter.

To obtain optimality conditions, we use the standard method[6]. Define a function

$$H(u,x,p) = up – u^2/2 + \mu(1-\cos x).$$

Then the optimal control satisfies the maximum condition

$$H(u,x,p) = \max_v H(v,x,p), \tag{12.7}$$

where p is a solution of the adjoint equation

$$p' = -H_x = -\mu \sin x. \tag{12.8}$$

Therefore, the system of optimality conditions includes relations (12.5) – (12.8) with respect to three unknown functions u, x, p. To study it, we use the elimination method in order to obtain a certain problem with respect to one of the unknown functions. Turning to zero the derivative of the function H with respect to control, we find the value $u = p$. Note that the corresponding second derivative is negative, i.e., the maximum point of the function H is found. As a result of substituting this value into the equations of state after differentiation, taking into account equation (12.8), we obtain a second-order nonlinear differential equation[7]

$$x''(t) + \mu \sin x(t) = 0,\ \ t \in (0,\pi). \tag{12.9}$$

Equalities (12.5) and (12.6) imply the boundary conditions

$$x(0) = 0, \;\; x(\pi) = 0. \tag{12.10}$$

Obviously, the boundary value problem (12.9) and (12.10) for any values of the parameter μ has a zero solution. To establish its additional properties, we need one result from the theory of boundary value problems for nonlinear differential equations. Consider the following problem

$$x'' + f(t,x) = 0, \;\; t \in (0,T); \; x(0) = 0, \; x(T) = 0. \tag{12.11}$$

Definition 12.1 *The function y is called a* ***lower solution*** *of problem* (12.11) *if the following relations hold*

$$y'' + f(t,y) \geq 0, \;\; t \in (0,T); \; y(0) \leq 0, \; y(T) \leq 0.$$

The function z is called an ***upper solution*** *of problem* (12.11) *if the following relations hold*

$$z'' + f(t,z) \leq 0, \;\; t \in (0,T); \; z(0) \geq 0, \; z(T) \geq 0.$$

Obviously, the usual solution to problem (12.11) is always both the lower and the upper solution of this problem. In this case, all the above relations are true in the form of equality. Naturally, the converse statement is not true in general.

The following assertion is true[8]:

Theorem 12.1 *If, with sufficient smoothness of the function f, there are lower and upper solutions y and z, respectively, of problem* (12.11)*, and the inequality $y(t) \leq z(t)$ is true for all $t \in (0,T)$, then this problem has a solution x such that*

$$y(t) \leq x(t) \leq z(t), \; t \in (0,T).$$

We use Theorem 12.1 to analyze the boundary value problem (12.9), (12.10). Let us define the function $y(t) = \varepsilon \sin t$, where ε is some non-negative number. We have

$$y''(t) + \mu \sin y(t) = -\varepsilon \sin t + \mu \sin(\varepsilon \sin t).$$

It is known that the sine of a small enough angle is sufficiently close to the value of the angle itself[9]. Then, for $\mu > 1$, the parameter ε can be chosen so small that the term on the right side of the last equality would be non-negative. Taking into account that the function y vanishes at the ends of the considered interval, we conclude that it is a lower solution to problem (12.9) and (12.10). Obviously, the constant function $z(t) = \pi$ satisfies equation (12.9) and takes positive values at the boundaries of this interval. Therefore, it is the upper solution of this boundary value problem. In this case, the inequality $\varepsilon \sin t \leq \pi$ is valid. Thus, the conditions of Theorem 12.1 are satisfied. Then for $\mu > 1$, there exists a solution x of problem (12.9) and (12.10) satisfying the condition

$$\varepsilon \sin t \leq x(t) \leq \pi, \; t \in (0,\pi).$$

Thus, for any value $\mu > 1$, the boundary value problem (12.9) and (12.10) has a positive solution. It was noted earlier that it always has a zero solution. It can be established[10] that for $\mu \leq 1$ there is an exclusively zero solution to this problem. Thus, the value $\mu = 1$ is its ***bifurcation point***. It can be seen[11] that for $\mu \leq 1$ the set optimal control problem is a function that is identically equal to zero, and for $\mu > 1$ it turns out to be the corresponding positive solution to the boundary value problem (12.9) and (12.10). This effect is called ***bifurcation of extremals***[12]. As the parameter increases from the bifurcation point, there is a sharp change in the optimal control, which means the absence of Hadamard well-posedness[13].

12.1.5 Chafee–Infante problem

Let us now consider one optimal control problem for a system with a fixed final state and a functional depending on two parameters[14].

Example 12.6 *It is required to minimize the functional*

$$I(u) = \frac{1}{4}\int_0^{\pi} (u^2 + \nu x^4 - 2\mu x^2)dt$$

on a set of such functions $u = u(t)$, which transform the system characterized by the equalities

$$x'(t) = u(t),\ t \in (0,\pi);\ x(0) = 0 \tag{12.12}$$

to the state

$$x(\pi) = 0, \tag{12.13}$$

where μ and ν are positive constants that are system parameters.

To obtain optimality conditions, we use the standard method[15]. Define a function

$$H(u,x,p) = up – 2u^2 – \nu x^4 + 2\mu x^2.$$

Then the optimal control satisfies the maximum condition

$$H(u,x,p) = \max_v H(v,x,p), \tag{12.14}$$

where p is a solution of the adjoint equation

$$p' = -H_x = 4\nu x^3 - 2\mu x. \tag{12.15}$$

Thus, to find the optimal control, we have the system (12.12)–(12.15). Reducing to zero the derivative of the function H with respect to the control, we find its only stationary point

$$u = p/4. \tag{12.16}$$

Obviously, the second derivative of the function H with respect to control is negative. Thus, the control determined by the formula (12.16) really corresponds to the maximum of the function H.

As in the previous example, we reduce the system of optimality conditions to a problem with respect to a single unknown quantity. Differentiating the state equation by t and taking into account relations (12.15) and (12.16), we have

$$x'' = u' = p'/4 = \nu x^3 - \mu x.$$

As a result, we obtain the equation[16]

$$x''(t) + \mu x(t) - \nu x(t)^3 = 0, \ t \in (0, \pi) \tag{12.17}$$

with boundary conditions

$$x(0) = 0, \ \ x(\pi) = 0. \tag{12.18}$$

Boundary value problem (12.17) and (12.18) is called the ***Chafee–Infante problem***.

If the solution to problem (12.17) and (12.18) is the optimal state of the system, then, in accordance with equation (12.12), the optimal control turns out to be a derivative of the solution under consideration. Thus, to find a solution to the optimization problem, it suffices to find solutions to the Chafee–Infante problem, and then check whether they correspond to the minimum of the considered functional.

The presence of a trivial (zero) solution to the Chafee–Infante problem is obvious. To prove the existence of a non-trivial solution to the system (12.17) and (12.18), we again use Theorem 12.1. As in the case of Example 12.5, we define the function $y(t) = \varepsilon \sin t$, where ε is a positive constant. We have

$$y''(t) + \mu y(t) - \nu y(t)^3 = (\mu - 1)\varepsilon \sin t - \nu \varepsilon^3 \sin^3 t = \varepsilon \sin t(\mu - 1 - \varepsilon^2 \nu \sin^2 t).$$

For $\mu > 1$ and small enough ε, the value on the right side of the last equality is non-negative. Then, taking into account the obvious equalities $y(0) = 0$ and $y(\pi) = 0$, we establish that the function y is the lower solution to problem (12.17) and (12.18). Suppose now that the function z takes an exclusively constant value c. Then the following equality holds

$$z''(t) + \mu z(t) - \nu z(t)^3 = c(\mu - \nu c^2).$$

This value is non-positive for $c \geq \sqrt{\mu/\nu}$. Hence, the function $z(t) = c$ is the upper solution of the Chafee–Infante problem.

Choosing the constant

$$c = \max\left\{\varepsilon, \sqrt{\mu/\nu)}\right\}.$$

we establish the validity of the inequality $y(t) \leq z(t)$ for all values of t. Using Theorem 12.1, we establish that the problem (12.17) and (12.18) has a solution x satisfying the inequality

$$y(t) \leq x(t) \leq z(t), \ t \in (0, \pi).$$

Thus, for any values of $\mu > 1$, $\nu > 0$ the Chafee–Infante problem has a positive solution. In fact, it can be shown[17], that under these conditions the positive solution of the problem is unique, and for $\mu \leq 1$ the Chafee–Infante problem has only zero solution. Thus, the value $\mu = 1$ turns out to be the bifurcation point of this problem, and we again encounter the phenomenon of bifurcation of extremals. The corresponding control is the derivative of a positive solution to the Chafee–Infante problem.

Note that if the function x is a solution to problem (12.17), (12.18), then the function y characterized by the equality $y(t) = -x(t)$ for all t is also a solution to this problem[18]. Therefore, changing the sign of the solution is certainly lead to a new solution of the Chafee–Infante problem, and due to the presence of even powers of the functions under the integral in the definition of the optimality criterion, the value of the optimality criterion in both cases will be the same. Naturally, changing the sign in the zero solution does not change anything. However, according to the positive solution of the Chafee–Infante problem, a negative solution that differs from it in sign can be determined. Thus, when $\mu > 1$, there are three solutions (zero, positive, and negative) to this problem[19].

RESULTS

Here is a list of questions in the field of well-posedness of optimal control problems for systems with a fixed final state, the main conclusions on this topic, as well as the problems that arise in this.

Questions

It is required to answer questions about examples of optimal control problems by systems with fixed final state, discussed in the lecture.

1. Why is the optimal control for Example 3.3 still optimal for Example 12.1?
2. Why is Tikhonov well-posedness substantiation technique, described earlier in relation to systems with a free final state, also applicable to systems with a fixed final state?
3. What stage of Tikhonov well-posedness justification for systems with a fixed final state turns out to be much more difficult than the corresponding stage for systems with a free final state?
4. What about the Hadamard well-posedness of the problem in Example 12.1?
5. Why is the Tikhonov well-posedness proof for Example 12.1 easily extended to Example 12.2 with an arbitrary parameter μ?
6. Why does Theorem 8.2 turn out to be applicable to justify Hadamard well-posedness of a problem with a fixed final state?
7. Why in Example 12.3 the solution of the optimality conditions can be either singular or be achieved on the boundary of the set of admissible controls?

8. Why cosines were chosen as the minimizing sequence in Example 12.3, and sines in the analogous Example 6.2?
9. What happens to the sequence $\{u_k\}$ from Example 12.3 when the number k increases without limit?
10. Will the main properties of the optimal control problem from Example 12.3 still hold if the control is chosen from the class of continuous rather than square-integrable functions?
11. Why for Example 12.4 it is not possible to use Theorem 8.2 on the Hadamard well-posedness of the optimization problem?
12. What properties differ between the optimization problems from Examples 12.2 and 12.4?
13. Is it possible for Example 12.4 to choose control from the space of continuous functions?
14. Is the optimization problem from Example 12.5 Hadamard well-posed and why?
15. Is it possible to use Theorem 12.1 to analyze the boundary value problem (12.9) and (12.10) for an arbitrary value of the parameter μ?
16. What happens to the properties of the problem (12.9) and (12.10) in the neighborhood of the parameter value $\mu = 1$?
17. What is the qualitative difference between Examples 12.5 and 12.6?
18. Why does the transformation $z(t) = x(\pi - t)$, which converts one solution of the optimality conditions for Example 12.6 to another, not give a new solution of the optimality conditions?
19. Why does it follow that the positive solution of the optimality conditions for Example 12.6 has mirror symmetry?
20. Why is the study of Example 12.6 not completed?

Conclusions

Based on the study of the considered problems of optimal control systems with fixed final state, we can come to the following conclusions.

- To substantiate the Tikhonov well-posedness of optimal control problems for systems with a fixed final state, the technique described for systems with a free final state is applicable.
- The optimal control problem from Example 12.1 is Tikhonov well-posed.

- To substantiate Hadamard well-posedness of optimal control problems for systems with a fixed final state, the technique described for systems with a free final state is applicable.
- The optimal control problem from Example 12.2 is Hadamard well-posed.
- The optimal control problem from Example 12.3 is not Tikhonov well-posed.
- The optimal control problem from Example 12.4 is not Hadamard well-posed.
- A situation is possible when the boundary value problem for a second-order nonlinear differential equation has a different number of solutions depending on the value of the problem parameter, which corresponds to the bifurcation phenomenon.
- It is possible that the system of optimality conditions has a different number of solutions depending on the value of the problem parameter, which corresponds to the bifurcation of extremals.
- For Example 12.5, the value of the parameter $\mu = 1$ is the bifurcation point.
- For Example 12.5, with $\mu \leq 1$, there is an exclusively zero solution of the optimality conditions, which is optimal, and with $\mu > 1$, an additional positive solution appears, which turns out to be optimal.
- As the parameter in Example 12.5 increases from the bifurcation point, the optimal control changes abruptly, which means that there is Hadamard ill-posed.
- For Example 12.6, there is a bifurcation of extremals, and the value of the parameter $\mu = 1$ is the bifurcation point.
- For Example 12.6, with $\mu \leq 1$, there is an exclusively zero solution of the optimality conditions, which is optimal, and with $\mu > 1$, two additional solutions appear, which are positive and negative.
- The positive solution of the system of optimality conditions for Example 12.6 is unique.
- The change of sign transforms any solution of the system of optimality conditions for Example 12.6 into a solution of the same system.
- The transformation, which consists of a mirror image about the middle of a given interval, translates any solution of the system of optimality conditions for Example 12.6 into a solution of the same system.
- A positive solution to the system of optimality conditions for Example 12.6 is symmetrical with respect to the middle of the given interval.

Problems

In the process of analyzing the considered problems of optimal control for systems with fixed final state, additional problems arise that need to be studied.

1. **Maximization problems**. When studying optimal control problems for systems with a free final state, we paid attention to the fact that the problems of finding the minimum and maximum of the same functional may have qualitatively different properties. Appendix considers the problem of maximizing the functional from Example 12.1.

2. **Ill-posedness of problems with unique singular control**. In Part II, we encountered a situation where a uniquely solvable optimal control problem whose solution is a singular control turned out to be Tikhonov ill-posed. In Chapter 11, there was an example of a similar problem for a system with a fixed final state, but questions of well-posedness were not considered in this case. Appendix will show that it is also Tikhonov ill-posed.

3. **General analysis of the bifurcation phenomenon**. For questions related to the phenomenon of bifurcation, see Notes[20].

4. **Completion of the analysis of the Chafee–Infante problem**. When analyzing Example 12.6, the phenomenon of bifurcation of extremals was discovered. At the same time, it was shown that for small values of the problem parameter, there is a unique solution to the Chafee–Infante problem, and for sufficiently large values, two new solutions appear. However, this does not guarantee the absence of other solutions. It will be shown in Appendix that a multiple bifurcation is observed for this problem.

12.2 APPENDIX

Below, we present some additional results related to the well-posedness of optimal control problems for systems with a fixed final state. In particular, Section 12.2.1 considers the functional maximization problem from Example 12.1, which has properties that differ significantly from those of the corresponding minimization problem. Section 12.2.2 provides further analysis of what was dis-cussed in the previous Chapter Example 11.1. It is shown that the corresponding optimal control problem is Tikhonov ill-posed. Finally, Section 12.2.3 completes the analysis of Example 12.6, where there is a multiple bifurcation of extremals.

12.2.1 Maximization of the functional from Example 12.1

Example 12.1 is a natural analog of Example 3.3 for a system with a fixed final state. Both of these problems have ideal properties, being Tikhonov well-posed of and differing by the sufficiency of the optimality condition in the form of the maximum principle. In Chapter 5 it was shown that changing the type of extremum in Example 3.3 leads to a qualitative change in the properties of the optimal control problem; see Example 5.1. The optimality conditions here turn out to be essentially insufficient and have an infinite number of solutions, and these solutions are arranged in pairs of functions that differ in signs. There are continuous, with one discontinuity point, with two discontinuity points, etc. In this case, two continuous solutions turn out to be optimal, i.e., uniqueness is also violated. Consider the functional maximization problem from Example 12.1, which is a natural analog of Example 5.1 for a system with a fixed final state[21].

Example 12.7 *It is required to maximize the functional*

$$I(u) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt$$

on a subset of such functions $u = u(t)$ *from the set*

$$U = \{u \,|\, |u(t)| \leq 1, \; t \in (0,1)\},$$

which for solutions of the Cauchy problem

$$x'(t) = u(t), \; t \in (0,1); \; x(0) = 0$$

guarantee the realization of the condition

$$x(1) = 0.$$

In accordance with the maximum principle, the optimal control here is determined from the minimum condition on a given set of function

$$H = pu - (u^2 + x^2)/2,$$

where p is a solution of the adjoint equation

$$p'(t) = x(t), \; t \in (0,1).$$

We find the function

$$u(t) = \begin{cases} 1, & \text{if } p(t) < 0, \\ -1, & \text{if } p(t) > 0. \end{cases} \tag{12.19}$$

The unique difference from the system of optimality conditions (5.1), (5.2), and (5.4) for Example 5.1 is that for $t = 1$, here the condition (also homogeneous) is specified for the function x, and not for p. However, in view of the presence of a fixed final state, the study of this system will be carried out differently than it was done in Chapter 5, i.e., using an iterative process, but using the methodology described in the previous chapter.

As can be seen from formula (12.19), the control can take only boundary values. When $u(t) = 1$ on the entire interval (0,1), then the equation of state has a solution $x(t) = t$, and hence $x(1) = 1$, which contradicts the given final condition. For $u(t) = -1$, we have $x(t) = -t$, and hence $x(0) = -1$, which also contradicts the existing restrictions. Therefore, the function u can only be discontinuous.

Let ξ be the unique discontinuity point of the control, with $u(t) = 1$ for $t < \xi$ and $u(t) = -1$ for $t > \xi$. Then $x(t) = t$ for $t < \xi$, and hence $x(\xi) = \xi$. For $t > \xi$, the state function is determined by the formula $x(t) = 2\xi - t$. Then from the equality $x(1) = 0$ we find the only possible break point $\xi = 1/2$. Thus $x(t) = 1 - t$ for $t > 1/2$. The adjoint equation for $t < 1/2$ has the form $p' = t$. Since the derivative is positive, the function p increases for $t < 1/2$. However, at $t = 1/2$ it vanishes. This means that it was initially negative, which corresponds to formula (12.19) for $u(t) = 1$. Similarly, for $t > 1/2$ we have the adjoint equation $p' = 1 - t$. This derivative is again positive, which means that the function p increases for $t > 1/2$. However, at $t = 1/2$ it vanishes. Therefore, in what follows, this function is positive, which is consistent with the control $u(t) = -1$.

Thus, the control equal to 1 for $t < 1/2$ and –1 for $t > 1/2$ really satisfies the system of optimality conditions. It is easy to see that, as a result of the change of sign, we obtain the second solution of the optimality conditions with one discontinuity point. This is a consequence of the invariance of this system, as well as of the optimal control problem itself with respect to sign change.

Let now there are two control discontinuity points ξ and η such that $0 < \xi < \eta < 1$. Suppose that the control takes the value 1 on the intervals $(0, \xi)$ and $(\eta, 1)$ and the value –1 on the interval (ξ, η). Then the state function takes the values t in the first part, $2\xi - t$ in the second part and $t + 2\xi - 2\eta$ in the third part. Then we obtain the equality

$$x(1) = 1 + 2\xi - 2\eta = 0.$$

Hence, we find $\eta = \xi + 1/2$, that is the distance between the break points is equal to half of the considered interval[22].

Thus, the equality $x(t) = t$ is true for $0 < t < \xi$. Then the adjoint equation has the form $p' = t$, which means that the function p increases. Considering that $p(\xi) = 0$, we conclude that this function is negative in the first interval, and therefore equality (12.19) is satisfied. Then we have $x(t) = 2\xi - t$ for $\xi < t < \xi + 1/2$. Hence, the adjoint equation has the form $p' = 2\xi - t$. We integrate this equality over the second interval, taking into account that the function p vanishes on its boundaries. We have

$$0 = \int\limits_{\xi}^{\xi+1/2} (2\xi - t)dt = \Big(2\xi t - \frac{t^2}{2}\Big)\Big|_{\xi}^{\xi+1/2} = \frac{\xi}{2} - \frac{1}{8}.$$

As a result, we determine $\xi = 1/4$. Therefore, when $1/4 < t < 3/4$, the adjoint system takes the form $p' = 1/2 - t$. For $t > 1/4$, this derivative is positive, which means that the function p increases from zero and is positive. At $t = 1/2$, the derivative vanishes, after which the function p decreases and reaches zero at $t = 3/4$. Thus, it is positive on the second interval, which also agrees with equality (12.19). Finally, for $t > 3/4$, we have $x(t) = t - 1$. In this case, we have the adjoint equation $p' = t - 1$. Therefore, the function p decreases from zero, and hence is negative. This result again agrees with equality (12.19). Thus, the third solution of the optimality conditions is found. The fourth solution is obtained from the third one by changing the sign.

Repeating the reasoning from the analysis of Examples 11.1 and 11.5, one can establish the existence of an infinite set of solutions to the system of optimality conditions. There exist two solutions for each number of control discontinuity points. At the same time, in contrast to the above, each solution with the same number of discontinuity points differs only in signs. To determine solutions with k discontinuity points, the segment [0,1] is divided into $2k$ equal parts. On the first of them, the control u_k^+ is assumed to be equal to 1, on the next two this is −1, on the next two the control is equal to 1, and so on, and finally, on the last this is 1. The second control u_k^+ with k discontinuity points differs from the first one only in sign. On Figure 12.1 one shows the corresponding states[23] x_k^+ and x_k^-.

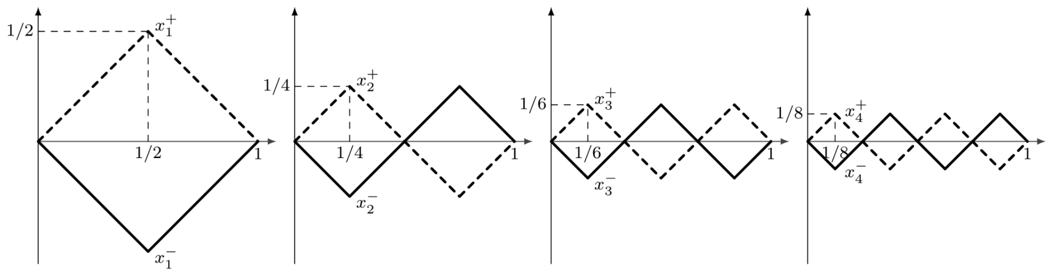

Figure 12.1 States satisfying the optimality conditions.

Obviously, the values of the functional on both controls with the same number of discontinuity points are the same. Calculate

$$I(u_k^+) = I(u_k^-) = \frac{1}{2}\int_0^1 [(u_k^+)^2 + (x_k^+)^2]dt = \frac{1}{2}\Big(1 + 2k\int_0^{1/2k} t^2 dt\Big) = \frac{1}{2} + \frac{1}{24k^2}.$$

Hence, it follows that the controls with one discontinuity point u_1^+ and u_1^- are optimal[24].

12.2.2 Ill-posedness of the problem from Example 11.2

Let us go back to Example 11.2. For simplicity, we restrict ourselves to the case when there are no restrictions on the control values. Thus, we consider the problem of minimizing the functional

$$I(u) = \frac{1}{2}\int_0^1 (x^2 - 2x\sin\pi t)dt.$$

on a set of functions $u = u(t)$, providing a translation of the system characterized by the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

to the state $x(1) = 0$.

This problem has a unique solution $u_0(t) = \pi \cos \pi t$, besides the minimum value of the functional is

$$\min I = -\frac{1}{2}\int_0^1 \sin^2 \pi t dt = -\frac{1}{4}.$$

Determine the control sequence

$$u_k(t) = \pi\big(\cos \pi t + \cos k\pi t\big),\ k = 1, 2, \ldots.$$

The corresponding sequence of states is characterized by the equality

$$x_k(t) = \pi \int_0^t \big(\cos \pi\tau + \cos k\pi\tau\big) d\tau = \sin \pi t + \frac{\sin k\pi t}{k}.$$

Note that $x_k(1) = 0$; so the $\{u_k\}$ control sequence is admissible.

Find the value

$$I(u_k) = \frac{1}{2}\int_0^1 \big(x_k - \sin \pi t\big)^2 dt - \frac{1}{4}.$$

The following inequality holds

$$0 \le \big|x_k - \sin \pi t\big| \le \frac{1}{k}.$$

Now we have

$$-\frac{1}{4} \le I(u_k) \le \frac{\pi}{2k^2} - \frac{1}{4}.$$

Thus, $I(u_k) \to -1/4$ as $k \to \infty$, so the sequence $\{u_k\}$ is minimizing.

Determine the value

$$\big\|u_k - u_0\big\|^2 = \int_0^1 \big\|u_k(t) - u_0(t)\big\|^2 dt = \int_0^1 \cos^2 k\pi t dt = \frac{1}{2}.$$

Thus, the optimal control u_0 is not the limit of the sequence $\{u_k\}$, i.e., the minimizing sequence does not converge to the optimal control. Therefore, the optimal control considered problem is Tikhonov ill-posed.

12.2.3 Analysis of the Chafee–Infante problem

We return to the analysis of the ***Chafee–Infante problem***, which includes the equation

$$x''(t) + \mu x(t) - \nu x(t)^3 = 0,\ t \in (0, \pi) \tag{12.20}$$

with boundary conditions

$$x(0) = 0,\ \ x(\pi) = 0, \tag{12.21}$$

where μ and ν are positive constants. It was noted earlier that it always has a zero solution x_0, which turns out to be unique for $\mu \le 1$. When $\mu > 1$, there is also a positive solution, which we denote by x_{+1}, as well as a negative solution $x_{-1} = -x_{+1}$. Note that if a function is a solution to problem (12.20), (12.21), then a function that differs from it only in sign is a solution to the same problem[25].

Let us try to establish the possibility of the existence of additional solutions to problem (12.20) and (12.21). Consider the equation

$$y''(t) + \frac{\mu}{4}y(t) - \frac{\nu}{4}y(t)^3 \;=\; 0,\; t \in (0,\pi)$$

with homogeneous boundary conditions. Using the arguments from Section 12.1.5 and Theorem 12.1, we establish that for $\mu > 4$ this problem has a positive solution, which we denote by y_+. Define the function

$$x_{+2}(t) \;=\; \begin{cases} y_+(2t), & \text{if } 0 < t < \pi/2, \\ -y_+(2\pi - 2t), & \text{if } \pi/2 < t < \pi. \end{cases}$$

Obviously, for $0 < t < \pi/2$ we get

$$x''_{+2}(t) \;=\; 4y''_+(2t) \;=\; -\mu y_+(2t) + \nu[y_+(2t)]^3 \;=\; -\mu x_{+2}(t) + \nu[x_{+2}(t)]^3.$$

Analogically, for $\pi/2 < t < \pi$ we obtain

$$x''_{+2}(t) \;=\; -4y''_+(2\pi - 2t) \;=\; \mu y_+(2t) - \nu[y_+(\pi - 2t)]^3 \;=\; -\mu x_{+2}(t) + \nu[x_{+2}(t)]^3.$$

Therefore, the function defined above satisfies relations (12.20) and (12.21). Thus, for $\mu > 4$, the Chafee–Infante problem has at least five solutions: zero x_0, positive x_{+1}, negative x_{-1}, and two new solutions x_{+2} and $x_{-2} = -x_{+2}$, which change sign exactly once. Thus, for $1 < \mu \leq 4$, the Chafee–Infante problem has exactly three solutions, and for $\mu > 4$, at least five solutions.

Similarly, the homogeneous boundary value problem for the equation

$$z''(t) + \frac{\mu}{9}z(t) - \frac{\nu}{9}z(t)^3 \;=\; 0,\; t \in (0,\pi)$$

for $\mu > 9$ has a positive solution, which we denote by z_+. Define the function

$$x_{+3}(t) \;=\; \begin{cases} z_+(3t), & \text{if } 0 < t < \pi/3, \\ -z_+(2\pi - 3t), & \text{if } \pi/3 < t < 2\pi/3, \\ z_+(3t - 2\pi), & \text{if } 2\pi/3 < t < \pi. \end{cases}$$

For $0 < t < \pi/3$ we get

$$x''_{+3}(t) \;=\; 9z''_+(3t) \;=\; -\mu z_+(3t) + \nu[z_+(3t)]^3 \;=\; -\mu x_{+3}(t) + \nu[x_{+3}(t)]^3.$$

Then for $\pi/3 < t < 2\pi/3$ we have

$$x''_{+3}(t) \;=\; -9z''_+(2\pi - 3t) \;=\; \mu z_+(2\pi - 3t) - \nu[z_+(2\pi - 3t)]^3 \;=\; -\mu x_{+3}(t) + \nu[x_{+3}(t)]^3.$$

Finally, for $2\pi/3 < t < \pi$ we obtain

$$x''_{+3}(t) \;=\; 9z''_+(3t - 2\pi) \;=\; -\mu z_+(3t - 2\pi) + \nu[z_+(3t - 2\pi)]^3 \;=\; -\mu x_{+3}(t) + \nu[x_{+3}(t)]^3.$$

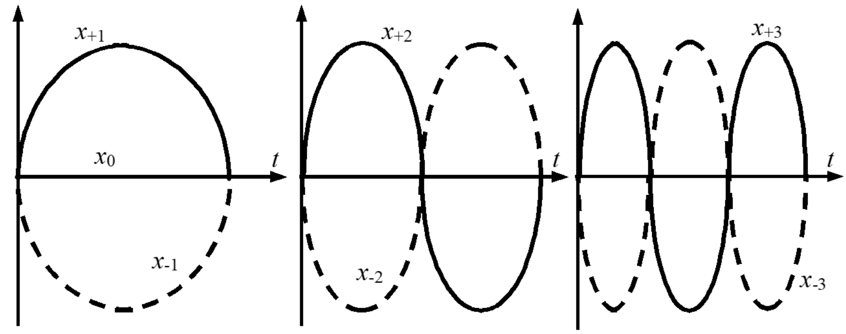

Figure 12.2 Solutions of the Chafee–Infante problem for $4 < \mu \leq 9$.

Thus, for $\mu > 9$, the problem has two additional solutions x_{+3} and $x_{-3} = -x_{+3}$, which change sign twice. Figure 12.2 shows the established solutions to the Chafee–Infante problem[26]. Therefore, for $4 < \mu \leq 9$, the Chafee–Infante problem has exactly five solutions, and for $\mu > 9$, at least seven solutions.

In the general case, for $\mu > k^2$, the equation

$$v''(t) + \frac{\mu}{k^2}v(t) - \frac{\nu}{k^2}v(t)^3 = 0, \; t \in (0, \pi)$$

with homogeneous boundary conditions has a positive solution, which we denote by v_+. Define the function

$$x_{+k}(t) = \begin{cases} z_+(kt), & \text{if } 0 < t < \pi/k, \\ -z_+(2\pi - kt), & \text{if } \pi/k < t < 2\pi/k, \\ z_+(kt - 2\pi), & \text{if } 2\pi/k < t < 3\pi/k, \\ -z_+(4\pi - kt), & \text{if } 3\pi/k < t < 4\pi/k, \\ \quad . \quad . \quad . & \quad . \quad . \quad . \quad . \end{cases}$$

By direct substitution, we make sure that it also satisfies relations (12.20) and (12.21) and changes sign k–1 times. Thus, for $(k–1)^2 < \mu \leq k^2$, the Chafee–Infante problem has exactly $2k$–1 solutions, and for $\mu > k^2$, at least $2k + 1$ solutions. Hence, in this case, not only is there no continuous dependence of the solution of the problem on the parameter μ. In this case, the values $\mu_k = k^2$, $k = 1, 2, ...$ are the bifurcation points for the Chafee–Infante problem. Thus, as in Example 12.5, we are dealing with an extremal bifurcation, but with an infinite number of bifurcation points.

The general behavior of the solution of the problem with a change in the positive parameter μ can be described as follows. For small values of this coefficient, the problem has a unique (zero) solution x_0. A change (increase) in the parameter μ does not affect the solution of the problem in any way, as a result of which there is reason to believe that there is a continuous dependence of the solution on the parameter. However, when μ passes through the critical value equal to 1, the problem acquires two more solutions x_{+1} and x_{-1}, i.e., the properties of the system have changed abruptly. Further growth of the parameter μ leads to a gradual change in these solutions, although their general structure remains unchanged. This continues until the parameter μ passes through the second bifurcation point equal to 4, after which two new solutions x_{+2} and x_{-2} appear from somewhere. Until the variable coefficient reaches the value $\mu = 9$, all non-trivial problems do not change qualitatively, but exceeding this value leads to the appearance of another pair of solutions x_{+3} and x_{-3}. In the general case, as long as the parameter μ changes on the interval between the squares of two natural numbers, the number of solutions to the Chafee–Infante problem remains unchanged. However, when passing through the values $\mu = k^2$, two new solutions x_{+k} and x_{-k} appear. Such a striking nature of the influence of the coefficient μ with the lowest linear term of the equation cannot but cause surprise and admiration[27].

The process described above is shown schematically in Figure 12.3, called the ***bifurcation diagram***. Here the abscissa axis corresponds to the numerical parameter μ, and the ordinate axis corresponds to the functional space X of the problem solutions. The bifurcation diagram clearly shows that at the bifurcation points, new ones branch off from the old solutions.

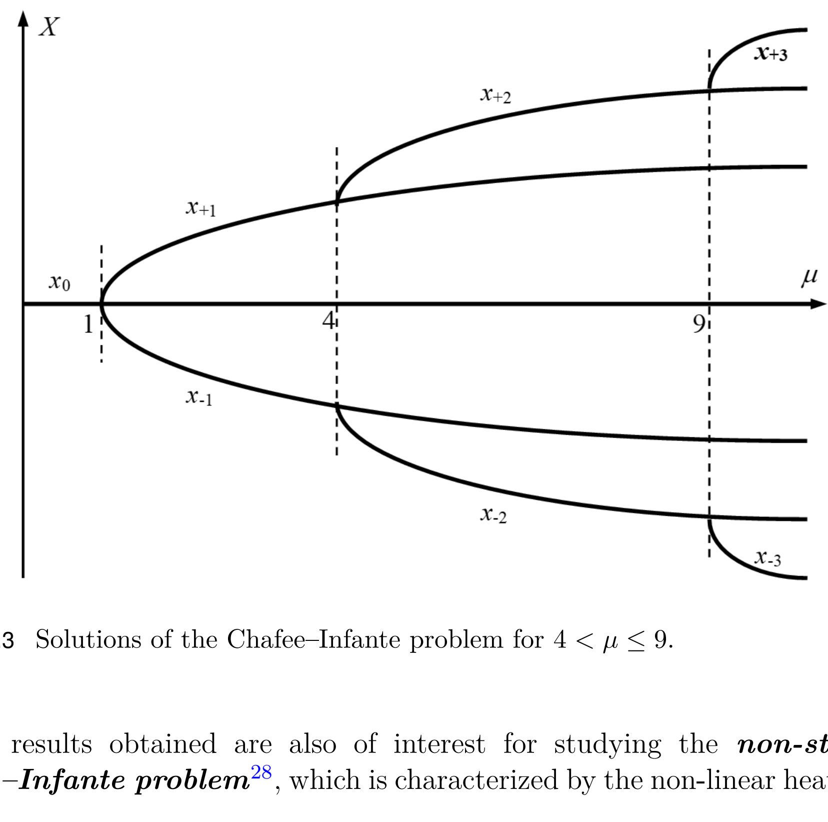

Figure 12.3 Solutions of the Chafee–Infante problem for $4 < \mu \leq 9$.

The results obtained are also of interest for studying the ***non-stationary Chafee–Infante problem***[28], which is characterized by the non-linear heat equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial \xi^2} + \mu u - \nu u^3,\ \tau > 0,\ 0 < \xi < \pi$$

with the boundary conditions

$$u(0,\tau) = 0,\ u(\pi,\tau) = 0,\ \tau > 0$$

and the initial condition

$$u(\xi,0) = u_0(\xi),\ 0<\xi<\pi.$$

Obviously, the equilibrium position of the non-stationary Chafee–Infante problem is necessarily the solution of the boundary value problem (12.20), (12.21). Then, based on the analysis done earlier, we can conclude that for $(k–1)^2 < \mu \leq k^2$ the non-stationary Chafee–Infante problem has exactly $2k–1$ equilibria[29].

We will not clarify the question of which of the solutions of the boundary value problem under consideration (optimality conditions) provides the minimum of the functional, since this is associated with certain technical difficulties. The main thing for us in this example is to obtain a boundary value problem that meets the optimality conditions and has amazing properties.

Additional conclusions

Based on the analysis of the above examples of optimal control problems by systems with fixed final state, we have the following additional conclusions.

- For Example 12.6, there are an infinite number of solutions to the optimality conditions, which are all piecewise constant functions.
- Each number of discontinuity points corresponds to two solutions of the optimality conditions for Example 12.6, which differ in signs.
- The maximum principle for Example 12.6 is a necessary but not sufficient condition for optimality.
- The optimal control problem from Example 12.6 has two solutions, which are functions with a unique breakpoint.
- The optimal control problem from Example 11.1 is Tikhonov ill-posed.
- For Example 12.6, there is a bifurcation of extremals.
- The Chafee–Infante problem has an infinite number of bifurcation points.
- If the parameter μ for Example 12.6 satisfies the inequality $(k–1)^2 < \mu \leq k^2$, then the Chafee–Infante problem has exactly $2k–1$ solutions.
- Depending on the values of the parameter μ, the Chafee–Infante problem has one zero solution and two solutions, differing in signs, corresponding to each number of sign changes of the function.
- For the non-stationary Chafee–Infante problem, the number of equilibrium positions depends on the value of the parameter μ in Example 12.6.
- The equilibria of the non-stationary Chafee–Infante problem are solutions to the stationary Chafee–Infante problem.

Notes

1. In Example 9.1, the interval [0,2] was considered as a set of control values, and in this case we are dealing with the interval [–1,1]. It is clear that this difference does not affect the properties of the set under study. However, if the equation is non-linear, proving the convexity of the set would present serious difficulties.

2. We recall that the absence of regular solutions of the maximum principle for Example 6.2 was proved on the basis of an unsuccessful attempt to apply an iterative process to solve a system of optimality conditions. The presence of an additional end condition allows you to get this result directly.

3. The principle of constructing an example of a Tikhonov ill-posed problem for a system with a fixed final state is quite simple. There is an example of the Tikhonov ill-posed optimal problem for a system with a free final state. In this case, it is Example 6.2. An additional condition at the final moment of time is determined by the value of the optimal state of the system at this moment of time, which in this case gives condition (12.2). Then the optimal control from the old example will also be optimal for the new example, in which the final state is fixed. It remains only to choose a divergent minimizing sequence of controls so that it ensures the transfer of the system to the chosen final state. In particular, in this case, the functions $u_k = \cos k\pi t$ were considered, while in Example 6.2 the functions $u_k = \sin k\pi t$ were used, which do not guarantee the transfer of the system to the desired state. In Chapter 15, an example of a Tikhonov ill-posed optimal control problem is given, which differs from the one considered in Example 12.3 only by the presence of an additional isoperimetric condition.

4. The method of constructing an example of an ill-posed Hadamard problem for a system with a fixed finite state is quite natural. A Hadamard ill-posed problem with a free final state is chosen. In this case, this is Example 8.1. As the end value of the system state of the new example, the end value of the optimal state from the old example is chosen. Thus, the optimal controls in the new and old examples are the same, which makes it possible to obtain the desired result. Chapter 15 gives an example of a Hadamard ill-posed optimal control problem with the isoperimetric condition.

5. The considered example is related to the ***Euler problem on an elastic rod***; see [88], [188]. It considers an elastic rod standing vertically on a flat horizontal surface. A downward force acts on the upper end of the rod. The profile of the rod is determined from the condition of minimum potential energy corresponding to the considered functional I. In Chapter 15, we will consider the problem of optimal control of the system in the presence of an isoperimetric condition, which is equivalent to this example.

6. Naturally, one can substitute the derivative of the state of the system into the optimality criterion instead of the control. The result is the Lagrange problem, which can be analyzed by means of the calculus of variations.

7. Relation (12.9) is the Euler equation for the corresponding problem of the calculus of variations.

8. Theorem 12.1 is given in [94].

9. Indeed, the sine expands into the Taylor series

$$\sin\varphi = \sum_{k=0}^{\infty}(-1)^k \frac{\varphi^{2k+1}}{(2k+1)!}.$$

For sufficiently small φ, one can neglect the terms of a higher order of smallness and establish the relation $\sin\varphi \approx \varphi$.

10. For proof of this statement; see [88].

11. At small values of the parameter μ, which corresponds to a small force acting on the elastic rod, the potential energy reaches a minimum when the rod is in a strictly vertical position, which corresponds to a zero solution. At the same time, for large values of this parameter, corresponding to a large force, the potential energy minimum is achieved when the rod is bent, which corresponds to a positive solution to the boundary value problem; see [88].

12. An ***extremal*** is a solution to the Euler equation in the calculus of variations; see [37], [61], [208]. As already noted, relation (12.9) exactly coincides with the Euler equation for the considered example, if it is interpreted from the standpoint of the calculus of variations. On bifurcations of extremals see [52].

13. In addition, in accordance with the method of establishment, it is possible to interpret the solutions of the boundary value problem (12.9) and (12.10) as the equilibrium positions of some non-stationary system with very unusual properties. However, even more unusual properties appear in the following problem.

14. Chapter 15 will consider the problem of optimal control of the system in the presence of an isoperimetric condition, which is equivalent to this example.

15. This example also reduces to a Lagrange problem, for which the corresponding Euler equation can be obtained.

16. Relation (12.17) is the Euler equation for the considered example.

17. For a proof of this statement; see [88].

18. The existence of a negative solution can also be proved using Theorem 12.1, choosing the function $y(t) = -c$ as the lower solution, and the function $z(t) = -\varepsilon \sin t$ as the upper solution

19. It is easy to find one more transformation that takes one solution of the Chafee–Infante problem to another. It is easy to see that if the function x is a solution to problem (12.17) and (12.18), then the function z, characterized by the equality $z(t) = x(\pi - t)$ for all t, is also a solution to this problem. This action corresponds to a mirror transformation with respect to the middle of the considered interval $(0, \pi)$. Thus, one gets the impression that for $\mu > 1$ there is a fourth solution to the Chafee–Infante problem obtained from its positive solution using the specified mirror image. Then we can expect the presence of the fifth solution, as a result of a similar transformation of the negative solution. However, since the positive solution of the problem is unique, we can conclude that it has mirror symmetry, i.e., the equality $x(t) = x(\pi - t)$ is fulfilled at $t \in (0, \pi)$. It is clear that the negative solution of the problem has the same property.

20. On the general theory of bifurcations; see, [96], [107], [184], [201]. On bifurcations of extremals; see [52].

21. Another analogue of Example 5.1 is the one discussed in the previous Chapter Example 11.1. It differs from Example 12.7 in that the corresponding end condition is non-uniform. In this case, the invariance of the problem with respect to sign change is violated, as a result of which the properties of its solution change significantly.

22. Note that the same result was obtained for Example 12.4, where we considered the maximization of a functional without a quadratic term with respect to the control. This led to the existence of a singular control. In addition, in that problem, the control changed on the interval

[0,2], and the final state was the value 1, not 0, as in this case. However, in both cases, the final state was achieved with a constant control equal to the average value from a given interval.

23. The results obtained are quite close, but do not fully coincide with those that were established in the analysis of Examples 5.1 and 11.4; see, in particular, Figures 5.3 and 11.3.

24. Recall that in Example 11.4, controls with one break point were also optimal, and in Example 5.1, continuous controls. In all these problems, it was required to find the maximum of the quadratic functional on a limited set of admissible controls.

25. From this, in particular, it follows that the number of solutions to the Chafee–Infante problem will certainly be odd, since a change of sign for the function x_0 does not give a new solution.

26. Naturally, the solutions of the Chafee–Infante problem depend on the specific values of the parameters μ and ν. In this regard, the functions depicted in Figure 12.2, only schematically depict the corresponding solutions.

27. We encountered a similar phenomenon in Chapter 2 when analyzing the logistic equation.

28. A comprehensive analysis of the non-stationary Chafee–Infante problem is carried out in [88].

29. As in previous examples (see Chapters 3 and 5), a non-stationary problem can be considered as a means for practically finding a non-trivial solution to the corresponding stationary problem in accordance with the method of establishing.

IV

OPTIMAL CONTROL PROBLEMS FOR SYSTEMS WITH ISOPERIMETRIC CONDITIONS

The next part of the book deals with the study of optimal control problems in the presence of an additional integral constraint in the form of equality, called the isoperimetric condition. In this case, it is possible, but not necessarily, given the final state of the system. The part consists of three chapters. The first of them describes the general principles of analysis of the considered problem class, as well as some examples illustrating the application of the described theory. In the next chapter, examples of the violation of the uniqueness of the solution of such problems, as well as the sufficiency of the corresponding necessary optimality conditions, are given. Finally, in the final chapter, examples of problems with isoperimetric conditions are described, for which some effects are observed similar to those that we encountered in the previous parts of the book.

CHAPTER 13

Optimization of systems with isoperimetric conditions

The subject of this chapter is the problem of optimal control of systems with additional integral constraints. In the Lecture, the corresponding optimality conditions are derived and examples of problems of this class are given, for which an analytical solution can be established in this way. Appendix presents methods for finding an approximate solution to such problems, conducts a qualitative analysis of the considered examples, solves one geometric problem with a similar constraint, describes some alternative methods and the corresponding vector problems.

13.1 LECTURE

In this lecture, we study problems of optimal control of systems with integral constraints in the form of equality, called isoperimetric conditions. Section 13.1.1 gives a general statement of such a free finite state problem and derives necessary optimality conditions in the form of the maximum principle. Section 13.1.2 gives an example of a problem with an isoperimetric condition that admits an analytical solution. Section 13.1.3 investigates an optimal control problem with an isoperimetric condition in the presence of a fixed final state, a particular case of which admits an analytical solution; see Section 13.1.4.

13.1.1 Optimal control problem with isoperimetric condition

We again consider the controlled system described by the Cauchy problem

$$x'(t) \;=\; f(t,u(t),x(t)),\; t\in(0,T);\; x(0)=x_0 \tag{13.1}$$

with known function f and numbers x_0 and T. The control u is chosen from the set

$$U \;=\; \{u|\; a(t)\le u(t)\le b(t),\; t\in(0,T)\},$$

DOI: 10.1201/9781003398585-13

where a and b are known functions. In contrast to the problems considered earlier, it is additionally assumed that the condition

$$\int_0^T q(t, u(t), x(t))dt = 0, \tag{13.2}$$

where q is a known function of its arguments, and x is a corresponding solution of problem (13.1), i.e., the state function.

Definition 13.1 *Equality* (13.2) *is called the* ***isoperimetric condition***[1].

The standard functional is chosen as the optimality criterion

$$I(u) = \int_0^T g(t, u(t), x(t))dt + h(x(T)),$$

where g and h are known functions, x is a solution of Cauchy problem (13.1) for the control u. The following optimal control problem is posed.

Problem 13.1 *It is required to find a function u that minimizes the functional I on a subset of functions from U that guarantees the equality* (13.2).

The presence of an additional condition (13.2) is the main feature of this problem, which requires some adjustment of the mathematical apparatus at our disposal[2].

Let a function u be an optimal control. Then the following inequality holds

$$\Delta I = I(v, y) - I(u, x) \geq 0,$$

where x is the optimal state, and v is an arbitrary element of the set U such that the corresponding state y satisfies[3] the equality (13.2).

Using the ***Lagrange multiplier method***, determine the functional

$$L(u, x, p) = I(u) + \int_0^T p(t)[x'(t) - f(t, u(t), x(t))]dt + \lambda \int_0^T q(t, u(t), x(t))dt,$$

where the function p and the number λ (Lagrange multipliers) are arbitrary. Then we get[4]

$$\Delta L = L(v, y, p) - L(u, x, p) \geq 0 \;\; \forall v \in U, \; \forall p, \; \forall \lambda. \tag{13.3}$$

We combine all the terms that are under the integral in the functional L, which explicitly depend on the control, as was done in Chapters 3 and 9. We get the function

$$H(t, u, x, p, \lambda) = pf(t, u, x) + \lambda q(t, u, x) - g(t, u, x).$$

Now the inequality (13.3) takes the form

$$\int_0^T (p\Delta x' - \Delta H)dt + \Delta h \geq 0 \;\; \forall v \in U, \; \forall p, \; \forall \lambda, \tag{13.4}$$

where we use the standard denotations

$$\Delta x = y – x,\ \Delta H = H(t,v,y,p,\lambda) – H(t,u,x,p,\lambda),\ \Delta h = h(y(T)) – h(x(T)).$$

Let us transform the terms in inequality (13.4) in the usual way. We have

$$\int_0^T p\Delta x' dt \ =\ p(T)\Delta x(T)\ -\ \int_0^T p'\Delta x dt,$$

$$\Delta H\ =\ \Delta_u H\ +\ H_x(t,u,x,p)\Delta x\ +\ \eta_1\ +\ \eta_2,$$
$$\Delta h\ =\ h_x[x(T)]\Delta x(T)\ +\ \eta_3$$

with the preservation of the previously accepted notation, where η_1 is a value of a higher order with respect to the increment Δx, $\eta_2 = \big[H_x(t,v,x,p,\lambda) – H_x(t,u,x,p,\lambda)\big]\ \Delta x$, η_3 is of the second order with respect to $\Delta x(T)$. As a result, inequality (13.4) is reduced to the form

$$-\int_0^T \Delta_u H dt - \int_0^T \big[H_x(t,u,x,p,\lambda)+p'\big]\Delta x dt + h_x(x(T))\Delta x(T) + \eta \geq 0\ \forall v \in U,\ \forall p,\ \forall \lambda, \tag{13.5}$$

where the remainder term η is determined by the well-known formula

$$\eta\ =\ \eta_1\ -\ \int_0^T (\eta_2+\eta_3)dt.$$

The resulting relation (13.5) is similar to inequalities (3.7) and (9.5).

We define, as usual, the adjoint system

$$p'(t)\ =\ -H_x(t,u,x,p,\lambda),\ t\in(0,T);\ \ p(T)\ =\ -h_x(x(T)). \tag{13.6}$$

Then inequality (13.5) takes the form

$$-\int_0^T \Delta_u H dt\ +\ \eta \geq 0\ \ \forall v \in U,\ \forall \lambda.$$

The resulting formula differs from the analogous inequality (3.10) only in the form of the function H, which includes an additional term. By analogy with Theorems 3.1 and 9.1, we have the following assertion[5].

Theorem 13.1 *In order for the control u to be a solution to Problem* 13.1*, it is necessary that it satisfies the maximum condition*

$$H(t,u(t),x(t),p(t),\lambda)\ =\ \max_{v\in[a(t),b(t)]} H(t,v,x(t),p(t),\lambda),\ t\in(0,T), \tag{13.7}$$

where x is the corresponding solution to system (13.1)*, which also satisfies condition* (13.2)*, and p is the solution of the adjoint system* (13.6).

Thus, the system of optimality conditions is characterized by relations (13.1), (13.2), (13.6), and (13.7), whence it is required to find three unknown functions u, x, p, and the constant λ. Let us give an example of the simplest optimal control problem with an isoperimetric constraint that admits an analytical solution.

13.1.2 Analytical solving of the problem with isoperimetric condition

Consider the following optimization problem.

Example 13.1 *We have a system characterized by the Cauchy problem*

$$x'(t) \;=\; u(t),\; t \in (0,1);\; x(0) \;=\; 0$$

on a subset of such functions $u = u(t)$ from the set

$$U \;=\; \{u|\, |u(t)| \le 1,\; t \in (0,1)\},$$

under the isoperimetric condition

$$\int_0^1 x(t)dt \;=\; 0,$$

which minimizes the functional

$$I \;=\; \frac{1}{2}\int_0^1 u(t)^2 dt.$$

We have Problem 13.1 with following values of the parameters

$$f(t,u,x) = u,\; T = 1,\; x_0 = 0,\; a(t) = -1,\; b(t) = 1,$$

$$g(t,u,x) = u^2/2,\; h(x) = 0,\; q(t,u,x) = x.$$

In accordance with the described method, we determine the function

$$H(t,u,x,p,\lambda) \;=\; pf + \lambda q - g \;=\; pu + \lambda x - u^2/2.$$

Then the adjoint equation (13.6) takes the form

$$p'(t) \;=\; -\lambda,\; t \in (0,1);\; p(1) = 0.$$

The optimal control is determined from the maximum condition for the function H on a given set. Obviously, its unique stationary point $u = p$ delivers the maximum of this function because of the negativity of its second derivative. Then the solution of the maximum condition is

$$u(t) \;=\; \begin{cases} -1, & \text{if } p(t) < -1, \\ p(t), & \text{if } -1 \le p(t) \le 1, \\ -1, & \text{if } p(t) > 1. \end{cases} \tag{13.8}$$

The solution of the adjoint system has the form $p(t) = \lambda(1–t)$. Thus, the function p is linear, its initial value is λ, and the final value is zero. There are three possibilities here; see Figure 13.1. For $\lambda > 1$ on the interval (0,1), there is a point ξ such that $p(t) > 1$ for $t < \xi$ and $|p(t)| \leq 1$ for $t > \xi$. For $\lambda < -1$ on the interval (0,1), there exists a point η such that $p(t) < -1$ for $t < \eta$ and $|p(t)| < 1$ for $t > \eta$. Finally, for $|\lambda| \leq 1$ we have $|p(t)| \leq 1$ for all t. Let us take a closer look at all three cases.

If $\lambda > 1$, then from formula (13.8) it follows that $u(t) = 1$ for $t < \xi$ and $u(t) = p(t)$ for $t > \xi$ (see Figure 13.1, dashed line). Thus, control is always positive. Then it follows from the state equation $x'(t) = u(t)$ that the function x is increasing. However, it is equal to zero at $t = 0$. Therefore, this function takes exclusively positive values, which means that the integral of it is also positive. Thus, the isoperimetric condition is not satisfied.

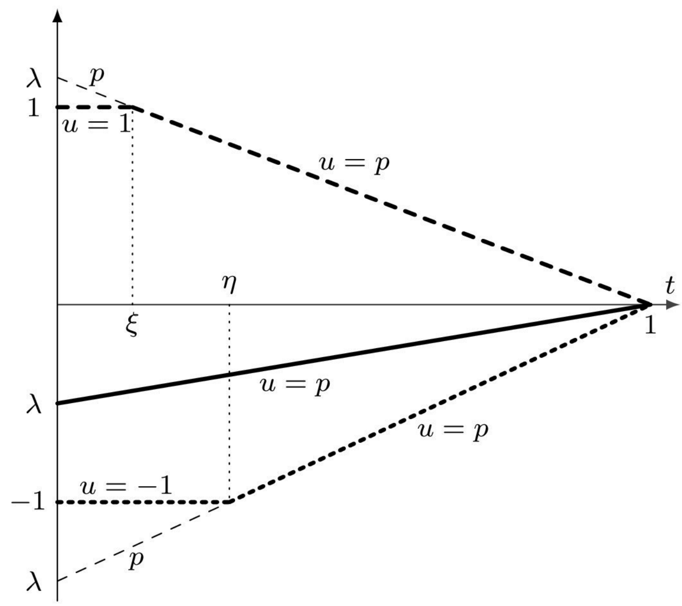

Figure 13.1 Possible control behavior for different λ.

Now suppose $\lambda < -1$. From equality (13.8), it follows that $u(t) = -1$ for $t < \xi$ and $u(t) = p(t)$ for $t > \xi$ (see Figure 13.1, dotted line). Therefore, the control takes exclusively negative values. In this case, it follows from the equation of state that the function x decreases. Given the zero initial condition, we conclude that all its values are negative. Then the integral of it is negative, which again contradicts the isoperimetric condition.

Now let the inequality $|\lambda| \leq 1$ be true, and hence $u(t) = p(t) = \lambda(1–t)$. The solution of the Cauchy problem

$$x'(t) = \lambda(1-t),\ t \in (0,1);\ x(0) = 0$$

is determined by the formula

$$x(t) = \lambda(t – t^2/2).$$

Find the integral[6]

$$\int_0^1 x dt = \lambda \int_0^1 \Big(t - \frac{t^2}{2}\Big) dt.$$

Turning the resulting value to zero, we find $\lambda = 0$, whence it follows that $u = 0$. Thus, the optimality conditions have a unique solution $u = 0$.

Indeed, the minimized functional takes non-negative values, and equality to zero is possible only when $u = 0$. The corresponding solution of the state equation is $x = 0$. The integral of this function is equal to zero, which means that the isoperimetric condition is satisfied. Therefore, we have found a unique solution to the optimal control problem, and the maximum principle gives a necessary and sufficient optimality condition[7].

13.1.3 Problem with isoperimetric condition and fixed final state

The results obtained can be extended to optimal control problems for systems with an isoperimetric condition and a fixed final state. The controlled system described by the Cauchy problem is again considered

$$x'(t) = f(t, u(t), x(t)), \ t \in (0, T); \ x(0) = x_0 \tag{13.9}$$

with given function f and numbers x_0, T. The control u belongs to the set

$$U = \{u| \ a(t) \le u(t) \le b(t), \ t \in (0, T)\},$$

where a and b are known functions. It is also assumed that the final state of the system is fixed, i.e., the equality holds

$$x(T) = x_T, \tag{13.10}$$

where the final state x_T is given. Besides, the following isoperimetric condition is true

$$\int_0^T q(t, u(t), x(t)) dt = 0, \tag{13.11}$$

where the function q is known. We have the optimality criterion

$$I(u) = \int_0^T g(t, u(t), x(t)) dt$$

with known function g. One poses the following optimal control problem.

Problem 13.2 *It is required to find such a function u from the set U that ensures the fulfillment of conditions* (13.10) *and* (13.10) *and minimizes the functional I on the set of controls that satisfy all given constraints.*

This is the extension of Problems 9.1 and 13.1. The problem is solved in a standard way. Let the function u be the optimal control, and x be the corresponding state system. Then we obtain the inequality

$$\Delta I = I(v,y) - I(u,x) \geq 0$$

for all admissible control v with corresponding state y. Define again the functional

$$L(u,x,p) = I(u) + \int_0^T p(t)\big[x'(t) - f(t,u(t),x(t))\big]dt + \lambda \int_0^T q(t,u(t),x(t))dt,$$

where the function p and the number λ are arbitrary. We have the inequality

$$\Delta L = L(v,y,p) - L(u,x,p) \geq 0 \;\; \forall v, \; \forall p, \; \forall \lambda,$$

which is the analog of (13.3). Determine the function

$$H(t,u,x,p,\lambda) = pf(t,u,x) + \lambda q(t,u,x) - g(t,u,x).$$

As a result, the preceding inequality takes the form

$$\int_0^T (p\Delta x' - \Delta H)dt \geq 0 \;\; \forall v, \; \forall p, \; \forall \lambda,$$

while retaining here and below the previously adopted notation. After integrating by parts, taking into account the existing boundary conditions, we establish

$$\int_0^T p\Delta x' dt = -\int_0^T p'\Delta x dt.$$

Using the Taylor series expansion, we get

$$\Delta H = \Delta_u H + H_x(t,u,x,p,\lambda)\Delta x + \eta_1 + \eta_2,$$

where the terms η_1 and η_2 have the same form as in Section 13.1.2. As a result, we obtain the inequality

$$\int_0^T \Delta_u H dt - \int_0^T \big[H_x(t,u,x,p,\lambda) + p'\big]\Delta x dt + \eta \geq 0 \;\; \forall v, \; \forall p, \; \forall \lambda$$

with the remainder term

$$\eta = -\int_0^T (\eta_1 + \eta_2)dt.$$

Determine the adjoint equation

$$p'(t) \;=\; -H_x(t,u,x,p,\lambda),\; t\in(0,T). \tag{13.12}$$

Then we get the inequality

$$\int_0^T \Delta_u H dt \,+\, \eta \ge 0 \;\; \forall v,\; \forall\lambda,$$

which is transformed in the standard way. The following assertion is true[8].

Theorem 13.2 *In order for the control u to be a solution to Problem* 13.2*, it is necessary that it satisfies the maximum condition*

$$H(t,u(t),x(t),p(t),\lambda) \;=\; \max_{v\in[a(t),b(t)]} H(t,v,x(t),p(t),\lambda),\; t\in(0,T). \tag{13.13}$$

Thus, the system of optimality conditions is characterized by formulas (13.9)–(13.13). An analytical solution here can be established only in exceptionally simple cases.

13.1.4 Analytical solving of a problem with an isoperimetric condition and a fixed final state

Consider the following example.

Example 13.2 *We have a system characterized by the Cauchy problem*

$$x'(t) \;=\; u(t),\; t\in(0,1);\; x(0) \;=\; 0.$$

The problem of optimal control is to find such a function $u = u(t)$ that ensures the fulfillment of the final condition

$$x(1) \;=\; 1$$

and the isoperimetric condition

$$\int_0^1 x(t)dt \;=\; \frac{1}{2},$$

minimizing under these constraints the functional

$$I \;=\; \frac{1}{2}\int_0^1 u(t)^2 dt.$$

Determine the function

$$H \;=\; pu + \lambda(x–1/2) - u^2/2.$$

Obviously, the unconditional maximum of the control function H is achieved at $u = p$. In this case, the adjoint equation has the form $p' = -\lambda$. It has a general solution $p(t) = -\lambda t + c$, where c is an arbitrary constant. Thus, we find the control $u(t) = -\lambda t + c$. Then the function x turns out to be a solution to the Cauchy problem

$$x'(t) \;=\; -\lambda t + c,\; t \in (0,1);\; x(0) \;=\; 0.$$

Find the solution

$$x(t) \;=\; -\frac{\lambda t^2}{2} + ct.$$

The right side of the obtained equality includes two unknown constants λ and c. To find them, there is a finite condition for the state of the system and an isoperimetric condition. We substitute the value of the function x into these relations. We have

$$x(1) \;=\; -\lambda/2 + c \;=\; 1,$$

$$\int_0^1 x(t)dt \;=\; \int_0^1 \Big(-\frac{\lambda t^2}{2} + ct\Big)dt \;=\; -\frac{\lambda}{6} + \frac{c}{2} \;=\; \frac{1}{2}.$$

As a result, we find $\lambda = 0$ and $c = 1$. Thus, we determine the unique solution of the system of optimality conditions $u(t) = 1$. It is easy to verify that it is the solution of the given optimal control problem[9].

RESULTS

Here is a list of questions in the field of optimal control problems with isoperimetric constraints, the main conclusions on this topic, as well as the problems that arise in this case and require further analysis.

Questions

It is required to answer questions concerning the analysis of optimal control problems with isoperimetric constraints.

1. What is the peculiarity of applying the Lagrange multiplier method for problems with an isoperimetric constraint?
2. What is the peculiarity of applying the Lagrange multiplier method for problems with an isoperimetric constraint?
3. Why, when defining the Lagrange functional, is the Lagrange multiplier p a function, and the other multiplier λ a constant?
4. How will the Lagrange functional be defined if there are not one, but two isoperimetric constraints?

5. How will the Lagrange functional be defined if there are not one, but two equations for two state functions?

6. How will the Lagrange functional be defined if there are not one, but two controls?

7. Why does Theorem 13.1 remain unfounded?

8. How many unknown values does the system of optimality conditions for Problem 13.1 include?

9. To what class of equations does the isoperimetric condition belong if it is interpreted as the problem of finding the parameter λ ?

10. Why is the isoperimetric condition used to find the Lagrange multiplier λ, although the parameter λ itself is not included in this condition?

11. What explains the possibility of finding an analytical solution to the optimal control problem from Example 13.1?

12. What are the functional properties in the general case, does the control defined by formula (13.7) have?

13. Why does the parameter λ in Example 13.1 belong to the interval $[-1, 1]$?

14. Why is the optimality condition for Example 13.1 necessary and sufficient?

15. Why is the Lagrange functional for Problems 13.1 and 13.2 defined in the same way?

16. What is the difference between the systems of optimality conditions for Problems 13.1 and 13.2?

17. What explains the possibility of finding an analytical solution to the optimal control problem from Example 13.2?

18. How does it follow that the control found in the process of analyzing the optimality conditions for Example 13.2 is indeed optimal?

Conclusions

Based on the study of optimal control problems with isoperimetric constraints, we can come to the following conclusions.

- The isoperimetric condition is an additional constraint in the form of an equality, taken into account in accordance with the Lagrange multiplier method.
- The maximum principle can also be established for optimization problems with isoperimetric constraints.

- The optimality conditions for a problem with an isoperimetric constraint is a system that includes a state equation and an adjoint equation with the corresponding boundary conditions, a maximum condition and the isoperimetric condition itself with respect to three unknown functions and a numerical Lagrange multiplier.
- The optimal control for Example 13.1 can be found analytically in view of the linearity of the equation and the isoperimetric condition with respect to the state function and the absence of an explicit dependence on it of the optimality criterion.
- The optimal control problem from Example 13.1 has a unique solution, and the corresponding optimality condition is both necessary and sufficient.
- The results obtained are extended to optimal control problems with isoperimetric and fixed final state of the system.
- The optimal control for Example 13.2 can be found analytically in view of the linearity of the equation and the isoperimetric condition with respect to the state function, the absence of an explicit dependence on it of the optimality criterion and the absence of explicit restrictions on the control values.
- The optimal control problem from Example 13.2 has a unique solution, and the corresponding optimality condition is both necessary and sufficient.

Problems

As a result of the analysis of the optimal control problems for the systems with isoperimetric conditions, the following additional problems arise.

1. **Justification of optimality conditions**. Theorems 13.1 and 13.2 are not justified. For the rationale for these results, see Notes[10].
2. **Practical solving of optimality conditions**. The system of optimality conditions for the considered examples was solved analytically. Naturally, this is possible only in exceptionally simple cases. In fact, iterative methods are usually used to analyze optimality conditions; see Appendix.
3. **Qualitative analysis of the considered examples**. For Examples 13.1 and 13.2, an analytical solution of the optimality conditions was found. However, the fact that exactly the optimal control was found needs to be substantiated. This result can be established with the help of a qualitative analysis of the tasks set, which is carried out in Appendix.
4. **Alternative methods**. In addition to the maximum principle, based on the method of Lagrange multipliers, other approaches can be used for the considered class of problems. In Appendix for case studies, the penalty method and variational inequality are applied.

5. **Counterexamples**. Of interest are optimal control problems with isoperimetric conditions, for which the existence and uniqueness of a solution, the sufficiency of optimality conditions, etc., are violated. The following chapters are devoted to these issues.

6. **Application**. Optimization problems with isoperimetric conditions naturally arise in practical situations. One such problem with geometric meaning is studied in Appendix.

13.2 APPENDIX

Below, we present some additional results for optimal control problems with isoperimetric conditions. In particular, Section 13.2.1 describes an algorithm for an approximate solution of the obtained optimality conditions. In Section 13.2.2, for Examples 13.1 and 13.2, the existence and uniqueness of an optimal control, Tikhonov well-posedness, and sufficiency of optimality conditions are proved based on the corresponding theorems from Part II. Section 13.2.3 considers an optimal control problem of a system with an isoperimetric condition, which has a natural geometric meaning. Section 13.2.4 uses the penalty method to obtain a problem without an isoperimetric condition, which is analyzed using a variational inequality. Finally, in Section 13.2.5, the results obtained earlier are extended to the vector case, when there are an arbitrary number of controls, state functions, and isoperimetric conditions.

13.2.1 Approximate solving of the problem with isoperimetric condition

Naturally, the possibility of an analytical solving of the system of optimality conditions is an exception, typical for fairly simple problems. As a rule, the solution of such problems is found approximately on the basis of iterative processes.

The system of optimality conditions for Problem 13.1 includes the state system (13.1), the adjoint system (13.6), the maximum condition (13.7), and the isoperimetric condition (13.2) for three unknown functions u, x, p, and the constant λ. This system is solved iteratively. In this case, as in the previous cases, the control is determined from the maximum condition, the state of the system is determined from the Cauchy problem (13.1), and the function p is determined from the adjoint system. As a result, with respect to the Lagrange multiplier λ, we have an isoperimetric condition, which can be interpreted as an algebraic equation with respect to this parameter. Its approximate solution can be found in the same way as the unknown final value for the function p in the shooting method; see Chapter 9. As a result, we have the following algorithm[11].

1. Initial approximations of the control u_0 and the Lagrange multiplier λ_0 are given. The sequence of algorithm parameters $\{\gamma_k\}$ is chosen, which will be used when finding the next approximation λ.

2. At the kth iteration, with a known control value u_k, the corresponding state function x_k is found from the Cauchy problem

$$x_k'(t) = f(t, u_k(t), x_k(t)),\ t \in (0, T),\ x_k(0) = x_0.$$

3. The value of p_k is determined with the known functions u_k, x_k, and the parameter λ_k from the previous iteration from the Cauchy problem

$$p_k'(t) = -H_x(t, u_k(t), x_k(t), p_k(t), \lambda_k),\ t \in (0,T),\ p_k(T) = -h_x(x_k(T)).$$

4. With known values of x_k, p_k, and λ_k, a new control approximation is found from the problem to the conditional extremum of the function

$$H(t, u_{k+1}(t), x_k(t), p_k(t), \lambda_k) = \max_{v \in [a(t),b(t)]} H(t, v, x_k(t), p_k(t), \lambda_k),\ t \in (0,T).$$

5. A new approximation of the parameter λ is found by the formula

$$\lambda_{k+1} = \lambda_k - \gamma_k \int_0^T q(t, u_{k+1}(t), x_k(t))dt.$$

In the case of convergence of the described algorithm, the result is a solution to the system of optimality conditions[12].

Let us now turn to the optimality conditions for Problem 13.2. The corresponding iterative process combines the properties of the algorithm described above with the shooting method from Chapter 9. In this case, the adjoint equation is supplemented by the final condition

$$p(T) = \psi, \tag{13.14}$$

where ψ is an unknown numerical parameter. Now, for three unknown functions u, x, p and two numerical parameters λ and ψ, there are two first-order differential equations with two boundary conditions (13.9), (13.12), and (13.14), the problem for the conditional maximum of the function (13.13) and two additional equalities (13.10) and (13.11). To solve this system, one can use the following algorithm.

1. Initial approximations of the control u_0 and the Lagrange multiplier λ_0, the number ψ_0 of the equality (13.14), and the sequence of algorithm parameters $\{\gamma_k\}$ and $\{\beta_k\}$ are given.

2. At the kth iteration, with a known control value u_k, the corresponding state function x_k is found from the Cauchy problem

$$x_k'(t) = f(t, u_k(t), x_k(t)),\ t \in (0,T),\ x_k(0) = x_0.$$

3. The value of p_k is determined with the known functions u_k, x_k, and the parameter λ_k from the previous iteration from the Cauchy problem

$$p_k'(t) = -H_x(t, u_k(t), x_k(t), p_k(t), \lambda_k),\ t \in (0,T),\ p_k(T) = \psi_k.$$

4. With known values of x_k, p_k, and λ_k, a new control approximation is found from the problem to the conditional extremum of the function

$$H(t, u_{k+1}(t), x_k(t), p_k(t), \lambda_k) = \max_{v \in [a(t),b(t)]} H(t, v, x_k(t), p_k(t), \lambda_k),\ t \in (0,T).$$

5. A new approximation of the parameter λ is found by the formula

$$\lambda_{k+1} = \lambda_k - \gamma_k \int\limits_0^T q(t, u_{k+1}(t), x_k(t))dt.$$

6. A new approximation of the parameter ψ is found by the formula

$$\psi_{k+1} = \psi_k - \beta_k[x_k(T) - x_1].$$

13.2.2 Qualitative analysis of the considered examples

In the study of Examples 13.1 and 13.2, a solution was found to the system of optimality conditions. In both cases, due to the uniqueness of the obtained solution, a conclusion was made about the optimality of the resulting control. However, as we already know, the unique solution to the optimality condition may not be the optimal control. This is possible in the case when the problem has no solution at all, and the extremum condition is not sufficient[13]. Let us check the sufficiency of optimality conditions, the existence of a solution for the examples under study, and at the same time other general properties of extremal problems that we encountered earlier.

Let us return to the consideration of Example 13.1. We minimize the functional

$$I = \frac{1}{2}\int\limits_0^1 u^2 dt$$

on the set U of functions $u = u(t)$ satisfying the inequality $|u(t)| \leq 1$ for all $t \in (0,1)$ under the following equality

$$\int\limits_0^1 x dt = 0,$$

where x is a solution to the problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Let us first check the sufficiency of the optimality conditions, as was done in Chapter 5. To do this, we will show that the remainder term η in the functional increment formula is non-negative. We have

$$\eta = \eta_3 - \int\limits_0^1 (\eta_1 + \eta_2)dt,$$

where η_3 is a value of a higher order with respect to the increment $\Delta x(T)$, determined by the terminal term of the optimality criterion, η_1 is a second-order term with respect to Δx, obtained by expanding the function H into a Taylor series in x, and η_2 is determined by the formula

$$\eta_2 = \left[H_x(t, v, x, p, \lambda) - H_x(t, u, x, p, \lambda)\right]\Delta x.$$

For Example 13.1, there is no terminal term in the optimality criterion, and the function H has the form $H = pu + \lambda x – u^2/2$. It is linear with respect to x, and the coefficient at x does not depend on the control. Then $\eta = 0$, which guarantees the sufficiency of optimality conditions. Thus, the set of solutions of the original problem coincides with the set of solutions of the optimality conditions. The latter consists of a unique element, which thus turns out to be an optimal control.

Another way to justify this result is related to the proof of the existence of an optimal control. For this purpose, we use Theorem 7.1, according to which the problem of minimizing a convex continuous functional bounded below on a convex closed bounded subset of a Hilbert space has a solution. The boundedness from below, the convexity and continuity of the quadratic integral functional were established in Chapter 3. Thus, it remains to check the desired properties of the set of admissible controls, which is the intersection of a given set U and a set V of controls that guarantee the isoperimetric condition.

First of all, we again choose $L_2(0, 1)$ as the control space. In Chapter 3, the convexity, closeness, and boundedness of the set U were established. From this, the boundedness of the intersection $U \cap V$ already follows[14]. Thus, it remains to prove the convexity and closeness of the set V.

Let u and v be elements of the set V, i.e., we have

$$\int_0^1 x dt = 0, \ \int_0^1 y dt = 0, \tag{13.15}$$

where x and y are the corresponding states. Then we obtain

$$x'(t) = u(t), \ t \in (0, 1); \ x(0) = 0; \quad y'(t) = v(t), \ t \in (0, 1); \ y(0) = 0.$$

For any $\alpha \in (0, 1)$ the control $w = \alpha u + (1–\alpha)v$ corresponds to the state z that is a solution of the problem

$$z'(t) = w(t), \ t \in (0, 1); \ z(0) = 0.$$

As a result, we obtain the equality $z = \alpha x + (1–\alpha)y$. Multiplying the first equality (13.15) by α, and the second one by $(1–\alpha)$ and summing up the results, we have

$$\int_0^1 z dt = 0.$$

This implies the inclusion $z \in V$, which means that the set V is convex.

Now let $\{u_k\}$ be a sequence of elements of the set V such that $u_k \to u$ in $L_2(0, 1)$. By virtue of the inclusions $u_k \in V$, the following equalities hold

$$\int_0^1 x_k dt = 0, \ k = 1, 2, \dots, \tag{13.16}$$

where x_k is a solution of the problem

$$x_k'(t) = u_k(t),\ t \in (0,1);\ x_k(0) = 0.$$

In Chapter 9, when analyzing Example 9.1 with the same state equation, $x_k(t) \to x(t)$ was found to converge for any t, where x is the system state corresponding to the control limit u. Then, as a result of the passing to the limit in equality (13.16), we obtain

$$\int_0^1 x dt = 0,$$

whence follows the inclusion $u \in V$. Thus, the set V is closed. Thus, all the conditions of Theorem 7.1 are satisfied, which means that the optimal control problem considered in Example 13.1 does indeed have a solution. Naturally, it must satisfy the optimality condition. Since the latter has a unique solution, we can again conclude that the previously found control u=0 is indeed the optimal control.

Let us now turn to Theorem 5.1 on the uniqueness of the optimal control. According to this result, a strictly convex functional on a convex subset cannot have two minimum points. The convexity of the corresponding set was established above, and the strict convexity of the quadratic integral functional was proved in Chapter 5. Thus, the uniqueness of the solution of the considered problem follows from Theorem 5.1.

Further, according to Theorem 8.1, if the problem of minimizing a strongly convex functional on a convex subset of a Hilbert space has a solution, then it is Tikhonov well-posed. The strong convexity of the quadratic integral functional was proved in Chapter 8. Then the optimal control problem from Example 13.1 is Tikhonov well-posed.

Example 13.2 considers the same state equation with the same optimality criterion. However, there are no explicit restrictions on the control, and the final condition $x(1) = 1$ is given in addition to the isoperimetric condition

$$\int_0^1 x dt = \frac{1}{2}.$$

Thus, the set of admissible controls in this case is the intersection of the sets V and W, where the set V is determined by the isoperimetric condition, and W is determined by the fixed final state. The convexity and closure of the first of them can be proved in the same way as for the previous example. The same properties for the set W were established in Chapter 9.

Because of the absence of restrictions on the values of the control, here we cannot establish the boundedness of the set of admissible controls and prove the solvability of the optimization problem using Theorem 7.1. However, it is obvious that the quadratic integral functional is coercive. Therefore, the existence of an optimal control for Example 13.2 follows from Theorem 7.2. With the existing properties of the optimality criterion and the set of admissible controls, the uniqueness of the optimal control, the well-posedness of the problem according to Tikhonov, and the sufficiency of the optimality conditions are established in the same way as for the previous example.

13.2.3 Dido problem

Chapter 2 dealt with the problem of finding a square with a given perimeter and maximum area. Consider its generalization, which consists of finding a curve $x = x(t)$ of a given length with given ends, for which the area of the corresponding curvilinear trapezoid bounded by this curve, a segment of the coordinate axis t and straight lines perpendicular to this axis and passing through the ends of the curve[15]; see Figure 13.2. This problem is called ***Dido problem***[16].

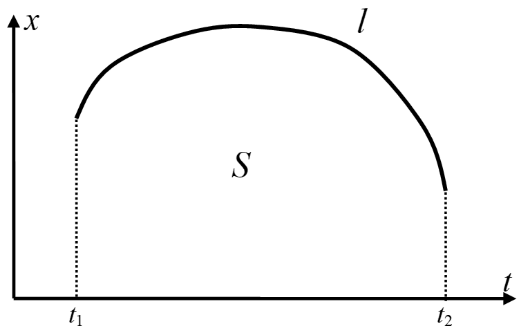

Figure 13.2 Curve of length l enveloping a curvilinear trapezoid of area S.

Let us give a mathematical formulation of this problem. The area of a curvilinear trapezoid bounded by the curve $x = x(t)$ on the interval (t_1, t_2) is equal to the integral

$$S = \int_{t_1}^{t_2} x(t)dt.$$

The length of the corresponding arc of the curve is

$$l = \int_{t_1}^{t_2} \sqrt{1 + [x'(t)]^2}dt.$$

The known coordinates of the ends of the curve arc are characterized by the equalities

$$x(t_1) = x_1, \; x(t_2) = x_2.$$

We now transform this problem to the standard form[17] of Problem 13.2, setting $t_1 = 0$, $t_2 = T$.

Example 13.3 *The optimal control problem is to find a function $u = u(t)$ that minimizes the functional*

$$I(u) = \int_0^T x(t)dt,$$

where x is a solution to the problem

$$x'(t) = u(t), \ t \in (0,T); \ x(0) = x_1$$

on the set of such functions u that guaranty the fulfillment of the final condition

$$x(T) = x_2$$

and isoperimetric condition

$$\int_0^T \sqrt{1+[u(t)]^2}dt = l.$$

The parameters T, x_1, x_2, and l are known.

Let us establish the optimality conditions for the considered example. Obviously, the last equality can be written as

$$\int_0^T \Big[\sqrt{1+[u(t)]^2} - \frac{l}{T}\Big]dt = 0.$$

Determine the function

$$H = pu + \lambda\Big(\sqrt{1+u^2} - \frac{l}{T}\Big) - x.$$

The function p here is a solution of the adjoint equation

$$p'(t) = 1, \ t \in (0,T).$$

The control is found from the minimum condition for the function H. Equaling its derivative with respect to control to zero, we have

$$\frac{\partial H}{\partial u} = p + \lambda\frac{u}{\sqrt{1+u^2}} = 0.$$

Differentiating the resulting equality, taking into account the adjoint equation, we establish the relation

$$1 + \lambda\frac{d}{dt}\frac{x'}{\sqrt{1+x'^2}} = 0.$$

In order to simplify the resulting second-order differential equation, calculate the derivative[18]

$$\frac{d}{dt}\Big(x - \lambda\sqrt{1+x'^2} + x'\frac{\lambda x'}{\sqrt{1+x'^2}}\Big) = x' - \frac{\lambda x'}{\sqrt{1+x'^2}}x'' +$$

$$+x''\frac{\lambda x'}{\sqrt{1+x'^2}} + x'\frac{d}{dt}\frac{\lambda x'}{\sqrt{1+x'^2}} = x'\Big(1+\frac{d}{dt}\frac{\lambda x'}{\sqrt{1+x'^2}}\Big) = 0$$

because of the previous equality. As a result, we obtain the first-order differential equation

$$x - \lambda\sqrt{1+x'^2} + x'\frac{\lambda x'}{\sqrt{1+x'^2}} = c_1,$$

where c_1 is an arbitrary constant. Thus, we get the differential equation

$$x - \frac{\lambda}{\sqrt{1+x'^2}} = c_1.$$

Determine a function φ such that $x' = \tan\varphi$. Then the previous equality takes the form

$$x - \lambda\cos\varphi = c_1. \tag{13.17}$$

After differentiation by φ, we obtain

$$\frac{dx}{d\varphi} + \lambda\sin\varphi = 0.$$

Now we have

$$dt = \frac{dx}{\tan\varphi} = -\frac{\lambda\sin\varphi d\varphi}{\tan\varphi} = -\lambda\cos\varphi d\varphi.$$

Integrating this equality, we get

$$t + \lambda\sin\varphi = c_2, \tag{13.18}$$

where c_2 is an arbitrary constant.

Equalities (13.17) and (13.18) imply[19]

$$(x-c_1)^2 + (t-c_2)^2 = \lambda^2(\cos^2\varphi + \sin^2\varphi) = \lambda^2. \tag{13.19}$$

Thus, the desired curve is a part of a circle[20] in the plane (t,x) of radius λ centered at the point with coordinates (c_2, c_1). To find concrete values of the radius and coordinates of the center of the circle (three unknowns), there are two boundary and one isoperimetric conditions[21].

13.2.4 Penalty method and variational inequality

The derivation of necessary optimality conditions for problems with isoperimetric conditions previously relied on the Lagrange multiplier method. However, the problems of finding an extremum in the presence of constraints in the form of equality were also studied using the penalty method[22]. One can try to use this method for solving the problem with the isoperimetric condition, in particular, for Problem 13.1.

Thus, we have the problem of minimizing the functional

$$I(u) = \int_0^T g(t,u(t),x(t))dt + h(x(T))$$

on the set of functions $u = u(t)$ from

$$U = \{u|\ a(t) \le u(t) \le b(t),\ t \in (0,T)\}$$

that guaranty the equality

$$\int_0^T q(t,u(t),x(t))dt = 0,$$

where x is a solution to the problem

$$x'(t) = f(t,u(t),x(t)),\ t \in (0,T);\ x(0) = x_0.$$

Using the ***penalty method***, determine the functional

$$I_\varepsilon(u) = I(u) + \frac{1}{2\varepsilon}\Big[\int_0^T q(t,u(t),x(t))dt\Big]^2,$$

where ε is a small enough positive number. The minimization problem for the functional I_ε can be solved by the standard optimization method. Let us illustrate this approach with a concrete example.

Consider, in particular, Example 13.1, which poses the problem of minimizing the functional

$$I = \int_0^1 u^2 dt$$

on the set of functions $u = u(t)$ from $U = \{u|\ |u(t)| \le 1,\ t \in (0,1)\}$ such that

$$\int_0^1 x dt = 0,$$

where x is a solution to the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Determine the functional

$$I_\varepsilon(u) = \int_0^1 u^2 dt + \frac{1}{2\varepsilon}\Big(\int_0^1 x dt\Big)^2.$$

We have obtained the problem of minimizing the functional I_ε on a convex set U, which can be understood as a subset of the Hilbert space $L_2(0,1)$. Try to solve this problem naturally with the help of variational inequalities, as was done in Chapters 4 and 10.

If the functional I_ε has a Gateaux derivative $I'_\varepsilon(u)$ at its minimum point u on the set U, then the variational inequality holds

$$I'_\varepsilon(u)(v-u) \ \geq\ 0 \ \ \forall v \in U. \tag{13.20}$$

However, it is necessary to find a functional derivative.

Find the difference

$$I_\varepsilon(u+\sigma h) \ - \ I_\varepsilon(u) \ = \ \frac{1}{2}\int\limits_0^1 \big[(u+\sigma h)^2 - u^2\big]dt \ + \ \frac{1}{2\varepsilon}\Big[\Big(\int\limits_0^1 ydt\Big)^2 - \Big(\int\limits_0^1 xdt\Big)^2\Big],$$

where σ is a number, h is an arbitrary element of the space $L_2(0,1)$, and y is a solution to the problem

$$y'(t) \ = \ u(t) + \sigma h(t), \ t \in (0,1); \ y(0) \ = \ 0.$$

We have the equalities

$$(u+\sigma h)^2 - u^2 \ = \ 2\sigma uh \ + \ \sigma^2 h^2,$$

$$\Big(\int\limits_0^1 ydt\Big)^2 - \Big(\int\limits_0^1 xdt\Big)^2 \ = \ \int\limits_0^1 (y+x)dt \int\limits_0^1 (y-x)dt \ = \ \int\limits_0^1 xdt \int\limits_0^1 zdt \ + \ \frac{1}{2}\Big(\int\limits_0^1 zdt\Big)^2,$$

where the function $z = y–x$ is a solution to the problem

$$z'(t) \ = \ \sigma h(t), \ t \in (0,1); \ z(0) \ = \ 0.$$

Now we have

$$z(t) \ = \ \sigma \int\limits_0^1 h(\tau)d\tau,$$

so the function z is proportional to the number σ.

Thus, we have the equality

$$I_\varepsilon(u+\sigma h) - I_\varepsilon(u) \ = \ \sigma \int\limits_0^1 \Big(uh + \frac{\sigma h^2}{2}\Big)dt + \frac{1}{\varepsilon}\Big[\int\limits_0^1 xdt \int\limits_0^1 zdt \ + \ \frac{1}{2}\Big(\int\limits_0^1 zdt\Big)^2\Big]. \tag{13.21}$$

Multiply the equation for z by an arbitrary function p and integrate the result. After integrating by parts, taking into account the initial condition, we obtain

$$-\int\limits_0^1 p'(t)z(t)dt \ + \ p(1)z(1) \ = \ \sigma \int\limits_0^1 p(t)h(t)dt.$$

We choose here as p the solution to the problem

$$p'(t) \ = \ \frac{1}{\varepsilon}\int\limits_0^1 xdt, \ t \in (0,1); \ p(1) = 0. \tag{13.22}$$

Then the equality (13.21) takes the form

$$I_\varepsilon(u+\sigma h) - I_\varepsilon(u) \;=\; \sigma\int\limits_0^1 (u-p)h dt + \frac{\sigma^2}{2}\int\limits_0^1 h^2 dt + \frac{1}{2\varepsilon}\Big(\int\limits_0^1 z dt\Big)^2.$$

Dividing this formula by σ and passing to the limit as $\sigma \to 0$ with using the linear dependence of z from σ, we get

$$I'_\varepsilon(u)h \;=\; \int\limits_0^1 (u-p)h dt \;\; \forall h.$$

As a result, the variational inequality (13.22) takes the form

$$\int\limits_0^1 (u-p)(v-u)dt \;\geq 0 \;\; \forall v \in U. \tag{13.23}$$

The optimality condition (13.23) was obtained earlier in Chapter 4 by analyzing Example 3.3. Its solution is determined by the formula

$$u(t) \;=\; \begin{cases} -1, & \text{if } \; p(t) < -1, \\ p(t), & \text{if } -1 \leq p(t) \leq 1, \\ 1, & \text{if } \; p(t) > 1. \end{cases} \tag{13.24}$$

Earlier, for this example, optimality conditions were obtained, including the same equation of state and the same formula for control. The difference is only on the right side of the adjoint equation, where previously there was a Lagrange multiplier with the opposite sign $-\lambda$, and in this case the integral of the state function divided by ε is found. Denoting this integral as $-\lambda$, we can reproduce the previous analysis and, in particular, exclude the case when this integral exceeds unity in absolute value. Since the derivative of p is constant, this function itself is linear and for $|\lambda| \leq 1$ also does not exceed unity in absolute value, having the form $p(t) = \lambda(1–t)$. Under these conditions, it follows from formula (13.24) that $u(t) = p(t)$, and hence $u(t) = \lambda(1–t)$. Substituting this value into the equation of state, we have

$$x'(t) \;=\; \lambda(1-t) \;=\; \frac{1}{\varepsilon}\int\limits_0^1 x(\tau)d\tau(t-1).$$

This is the ***integro-differential equation***.

Integrating this equality with a homogeneous initial condition, we find

$$x(t) \;=\; \frac{1}{\varepsilon}\int\limits_0^1 x(\tau)d\tau\Big(\frac{t^2}{2} - t\Big).$$

After integration, we obtain

$$\varepsilon\int\limits_0^1 x(\tau)d\tau \;=\; \int\limits_0^1 x(\tau)d\tau\int\limits_0^1 \Big(\frac{t^2}{2} - t\Big)dt \;=\; -\frac{1}{3}\int\limits_0^1 x(\tau)d\tau.$$

This equality can only hold if the integral of the function x is equal to zero. Then it follows from the previous equality that $x = 0$, and hence $u = 0$. The result obtained exactly coincides with what was established in the Lecture[23].

13.2.5 Vector problem with isoperimetric conditions

The results obtained can be easily extended to the vector case. Let the control and the state function be vector functions

$$u = (u_1, u_2, ..., u_r), \;\; x = (x_1, x_2, ..., x_n).$$

Now the state equation

$$x'(t) = f(t, u(t), x(t)), \; t \in (0, T), \; x(0) = x_0$$

is the Cauchy problem for the system of differential equations, besides f is n order vector function of $r + n + 1$ variables, and x_0 is an n-dimensional vector. The vector control u belongs to the set

$$U = \{u| \; u(t) \in G(t), \; t \in (0, T)\},$$

where $G(t)$ is a subset of the r-dimensional Euclidean space. In addition, some number s of isoperimetric conditions are given

$$\int_0^T q(t, u(t), x(t))dt = 0,$$

where q is a vector function of order s in $r+n+1$ variables. The optimality criterion is determined by the formula

$$I = \int_0^T g(t, u(t), x(t))dt + h(x(T)),$$

where g is a function of $r + n + 1$ variables and h is a function of n variables. We have the following ***vector optimal control problem***.

Problem 13.3 *Find a control u that minimizes on a subset U of functions satisfying the given isoperimetric conditions the functional I whose definition includes the function x, which is a solution to the system of state equations for a given control u.*

Define the function of $r + 2n + s + 1$ variables

$$H(t, u, x, p, \lambda) = \langle p, f(t, u, x)\rangle + \langle \lambda, q(t, u, x)\rangle - g(t, u, x),$$

where λ is a vector of s order, the first scalar product is understood in the sense n-dimensional Euclidean space, and the second in the sense of an s-dimensional space.

In accordance with the ***maximum principle***, the optimal control satisfies the condition

$$H\big(t, u(t), x(t), p(t), \lambda\big) \;=\; \max_{v \in G(t)} H\big(t, v, x(t), p(t), \lambda\big),\; t \in (0, T),$$

where p is a solution to the ***adjoint system***

$$p'(t) \;=\; -H_x(t, u, x, p, \lambda),\; t \in (0, T),\; p(T) \;=\; h_x(x(T)).$$

As a result, a system of optimality conditions is obtained, which includes as unknown r controls u, n states x and functions p, as well as s numerical Lagrange multipliers λ. In this case, the maximum condition is a problem for the conditional extremum of the function H with respect to r variables (controls). The equations of state and the adjoint system include n differential equations each with the corresponding boundary conditions, and the vectors H_x and h_x include n components each, which are partial derivatives with respect to the variables x_i, $i = 1, ..., n$. Finally, there are s isoperimetric conditions that ultimately allow us to determine the vector λ.

One can also consider a vector optimal control problem with isoperimetric conditions and fixed final states[24]. In this case, the equations of state have the same form as in the previous case. However, in addition to the set U and the isoperimetric conditions, the final conditions are given

$$x(T) \;=\; x_T,$$

where x_T is a vector of n order. Determine an optimality criterion

$$I \;=\; \int_0^T g(t, u(t), x(t))dt.$$

Problem 13.4 *Find a control u that minimizes the functional I, whose definition includes the function x, which is a system of equations of state for a given control u, on a subset U of functions that satisfy the given isoperimetric and finite conditions.*

To solve this problem, the function H is introduced in the same way as before. The optimal control here will satisfy the maximum condition given above. However, the function p will satisfy the adjoint equation

$$p'(t) \;=\; -H_x(t, u, x, p, \lambda),\; t \in (0, T)$$

without boundary conditions. The same optimality conditions are obtained with respect to the same set of unknown functions. However, there are $2n$ boundary conditions for n equations of state, and there are no boundary conditions for n adjoint equations. The resulting system can, in principle, be solved approximately using the shooting method.

Additional conclusions

Based on the properties of optimal control problems with isoperimetric conditions established in Appendix, we can draw the following conclusions.

- The optimality conditions for a problem with an isoperimetric constraint are usually solved iteratively.
- The optimality conditions for a problem with an isoperimetric constraint and a fixed final state are usually solved iteratively using the shooting method.
- The theorems given in Part II on the sufficiency of the maximum condition, the existence and uniqueness of the optimal control, and the well-posedness of optimal control problems according to Tikhonov are also applicable to the analysis of problems with isoperimetric conditions.
- The optimal control problem from Example 13.1 has a unique solution and is Tikhonov well-posed, and the corresponding optimality conditions are necessary and sufficient.
- The optimal control problem from Example 13.2 has a unique solution and is Tikhonov well-posed, and the corresponding optimality conditions are necessary and sufficient.
- Dido problem belongs to the class of optimal control problems for systems with an isoperimetric condition and a fixed final state.
- The solution to Dido problem is an arc of a circle.
- A problem with an isoperimetric condition can be reduced to a problem without an isoperimetric condition using the penalty method.
- Applying the penalty method to the analysis in Example 13.1 leads to the same result as applying the maximum principle.
- Methods for studying optimal control problems for systems with an isoperimetric constraint, both with a free and a fixed final state, are extended to the vector case.

Notes

1. Chapter 2 dealt with the problem of finding a rectangle with the maximum area of a given perimeter. A summary of this problem is given in Appendix. Fixing the perimeter is a kind of restriction in the form of equality. It is in this connection that the term “isoperimetric condition” appeared. In the calculus of variations, integral constraints in the form of equality are usually called ***isoperimetric conditions***; see [37], [61], [208].

2. To solve Problem 13.1, we can also use the following idea. We define a new state function $y = y(t)$ as a solution to the following Cauchy problem

$$y'(t) \,=\, q(t, u(t), x(t))\; t \in (0, T);\;\; y(0) = 0.$$

Then the isoperimetric condition (13.2) is equivalent to the condition $y(T) = 0$. Thus, Problem 13.1 is reduced to an optimal control problem for a system of two differential equations with a fixed final state.

3. We again ignore the question of how to find a control that is an element of the set U and guarantees the fulfillment of condition (13.2), and whether it exists at all.

4. This is an analogue of the inequalities (3.5) and (9.3).

5. Naturally, the assertions of the theorem must be rigorously substantiated. To do this, one can use the same technique as in Chapter 3.

6. In principle, one can immediately conclude that for negative λ the function x is everywhere negative, and for positive values, it is everywhere positive. In both cases, the integral of x does not vanish in any way, which means that the isoperimetric condition cannot be satisfied. There remains the case $\lambda = 0$, which leads to the optimal control. However, we have chosen just such a way of research in order to emphasize that in practice the isoperimetric condition is used to find the Lagrange multiplier λ.

7. This example is in a certain sense similar to Examples 3.1, 3.2, and 9.1, in which the extremum necessary condition also had a unique solution, which was found analytically and turned out to be a solution to the problem. Note that here we solved an optimal control problem with a linear equation and an optimality criterion that does not depend on the function x at all. As a result, the function H turned out to be linear with respect to x. Then the adjoint system does not depend on x at all, and its solution can be found without connection with the state equation. This is precisely what determines the possibility of finding an analytical solution to the problem.

8. We again refrain from substantiating this assertion.

9. To do this, it is enough to prove the existence of an optimal control or the sufficiency of optimality conditions; see Appendix.

10. Regarding the substantiation of optimization methods in problems with difficult constraints [5], [56], [95], [140].

11. In the described algorithm, the actions at the fourth and fifth steps can be swapped. However, in this case, when calculating the parameter λ, the control is selected at the current iteration, and when calculating the control, its newly found value at the next iteration is chosen as λ.

12. Naturally, this does not guarantee that the found limit value of the control will be optimal, since the optimality condition is not sufficient for the general case.

13. In particular, for the function $f(x) = x^3$, the stationary condition has a unique solution $x = 0$, which is not the minimum point of this function.

14. Indeed, the intersection $U \cap V$ is part of a bounded set U, and hence is also bounded.

15. This problem is considered, for example, in [61].

16. According to legend, the Phoenician princess Dido, having landed on the African coast of the Mediterranean Sea, received permission from a local tribe to settle for a piece of land that can be covered with the skin of a bull. Dido cut the skin into narrow straps, wove a rope out of them, and covered a fairly large part of the coast with it. It is believed that the city of Carthage was founded in this way.

17. Naturally, here one can use the methods of the calculus of variations; see [37], [61], [208].

18. In the calculus of variations, the value in brackets on the left side of this equality is called the ***first integral***, [37], [61], [208].

19. It is enough in each of the equalities to transfer the trigonometric term to the right side of the equality, the term from the right side to the left, square it, and then add the resulting equalities.

20. Argument t belongs to the interval $(0, T)$.

21. Let us consider, for example, the parameters $T = 1$, $x_1 = 0$, $x_2 = 0$, $l = \pi/2$. Then there are three equations for three unknown parameters λ, c_1, and c_2. Using two boundary conditions, from equality (13.19) we obtain two equations

$$(c_1)^2 + (c_2)^2 = \lambda^2, \ (1-c_1)^2 + (1-c_2)^2 = \lambda^2.$$

Because of the equation

$$x - \frac{\lambda}{\sqrt{1+x'^2}} = c_1$$

the isoperimetric condition takes the form

$$\int_0^1 \frac{\lambda dt}{x - c_1} = \frac{\pi}{2}.$$

Taking into account equality (13.19), we establish the third equation for unknown parameters

$$\frac{\pi}{2\lambda} = \int_0^1 \frac{dt}{\sqrt{\lambda^2 - (t-c_1)^2}} = \arcsin\frac{1-c_1}{\lambda} + \arcsin\frac{c_1}{\lambda}.$$

It is easy to verify that the resulting system of equations has a solution $\lambda = 1$, $c_1 = 0$, $c_2 = 1$. Thus, the desired curve turns out to be an arc of a circle

$$x^2 + (t-1)^2 = 1, \ t \in (0, 1).$$

22. In Chapter 2, the penalty method was used for function minimization, in Chapter 4 for the free finite state optimal control problem, and in Chapter 10 for the fixed final state problem.

23. The question arises, what happens if we try to use the variational inequality for Example 13.1 directly, i.e., without using the penalty method. Indeed, we are dealing with the minimization of the original functional I, which has an extremely simple form, at the intersection of a given set U and a set V, which guarantee the fulfillment of the isoperimetric condition. In the Lecture it was shown that the intersection $U \cap V$ is convex. This circumstance allows us to use the variational inequality to solve this problem. Obviously, the derivative of the functional I at the point u is equal to $I'(u) = u$. Then the corresponding variational inequality has the form

$$\int_0^1 u(t)[v(t) - u(t)]dt \geq 0 \ \forall v \in (U \cap V).$$

This inequality is close enough to condition (13.23). However, the solution of the latter was associated with the use of the needle variation; see Chapter 4. Unfortunately, the needle variation does not guarantee the isoperimetric condition, so the solution of the obtained variational inequality is not obvious.

24. Naturally, an intermediate variant is also possible, when a part of the system states is specified at the final moment of time. In this case, for the adjoint equation, a part of the conditions at the final moment of time will be set.

CHAPTER 14

Absence of sufficiency and uniqueness in problems with isoperimetric conditions

The fourth part of the book is devoted to problems of optimal control of systems with isoperimetric conditions. This and the following chapters provide examples of these problems, for which the effects that we encountered in the previous parts are realized. This includes the absence of existence or uniqueness of a solution, insufficiency and degeneration of optimality conditions, etc. Below are examples of problems of this class, for which the solution is not unique, and the optimality conditions are not sufficient. In this case, systems with both free and fixed final states are considered.

14.1 LECTURE

In the previous chapter, optimality conditions were given for optimal control problems in the presence of isoperimetric conditions. The effectiveness of the research method was illustrated with relevant examples. However, in the past, when solving extreme problems, we encountered different difficulties. Similar effects are observed in problems with isoperimetric conditions. This lecture will give examples of problems of the considered class, for which there are no sufficiency of optimality conditions and the uniqueness of the solution to the problem. The first example corresponds to a linear system of optimality conditions, while in the second case this system turns out to be non-linear.

14.1.1 Linear system of optimality conditions for a fixed final state

We continue to study examples of optimal control problems with isoperimetric conditions.

Example 14.1 *There is a system described by the equalities*

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0. \tag{14.1}$$

DOI: 10.1201/9781003398585-14

The optimal control problem is to find a function $u = u(t)$ that minimizes the functional

$$I(u) = \int_0^1 u(t)^2 dt$$

under the final condition

$$x(1) = 0 \tag{14.2}$$

and the isoperimetric condition[1]

$$\int_0^1 x(t)^2 dt = 1. \tag{14.3}$$

To solve this problem[2], we will use the technique described in the previous chapter[3]. We write the last equality in the form

$$\int_0^1 (x^2 - 1) dt = 0.$$

We have Problem 13.2 with following parameters

$$f(t,u,x) = u,\ T = 1,\ x_0 = 0,\ x_T = 0,\ a(t) = -N,\ b(t) = N,$$

$$g(t,u,x) = u^2,\ h(x) = 0,\ q(t,u,x) = (x^2 – 1).$$

Determine the function

$$H(t,u,x,p,\lambda) = pf + \lambda q – g = p + (x^2 – 1) – u^2.$$

Then the adjoint equation takes the form

$$p'(t) = -2\lambda x(t),\ t \in (0,1). \tag{14.4}$$

The control is found from the maximum condition of the function H. As a result, we determine

$$u(t) = p(t)/2. \tag{14.5}$$

Thus, with respect to the three unknown functions u, x, p, and the number λ, a system of optimality conditions (14.1) – (14.5) is obtained, and for a fixed value of λ the resulting system is linear[4]. In accordance with the method of eliminating unknowns, we will try to reduce the existing system to a problem with respect to the state function[5]. We substitute the control from formula (14.5) into the state equation. After its differentiating, taking into account the adjoint equation, we have

$$x''(t) = u'(t) = p'(t)/2 = -\lambda x(t).$$

Thus, the function x satisfies the second-order differential equation[6]

$$x''(t) + \lambda x(t) = 0,\ t \in (0,1) \tag{14.6}$$

with initially specified boundary conditions

$$x(0) = 0, \;\; x(1) = 0. \tag{14.7}$$

The number λ included in equation (14.6) is unknown. However, in addition to the resulting boundary value problem, there is also an isoperimetric condition (14.3). Note that the solution to problem (14.6) and (14.7) is certainly a function x that is identically equal to zero. However, in this case equality (14.3) is not hold, and therefore we will be interested in non-trivial solutions to this problem. As a result, we give the extremely important concept of the theory of differential equations.

Definition 14.1 *The **Sturm–Liouville problem**[7] consists of finding non-trivial solutions to the boundary value problem* (14.6) *and* (14.7) *and the values of λ for which such a solution exists. These solutions are called the **eigenfunctions** of a given problem, and the corresponding numbers λ are called its **eigenvalues**.*

Multiplying equality (14.6) by x and integrating the resulting relation taking into account condition (14.3), we have

$$\int_0^1 x'' x dt \, + \, \lambda \int_0^1 x^2 dt \; = \; 0.$$

Integrating the first integral by parts, taking into account boundary conditions (14.7) and using equality (14.3), we find

$$\lambda \; = \; \int_0^1 x'^2 dt.$$

Substituting the found value into equality (14.6), we obtain

$$x''(t) \, + \, x(t) \int_0^1 x'^2 dt \; = \; 0, \; t \in (0, 1). \tag{14.8}$$

Thus, the system of optimality conditions is reduced to a unique integro-differential equation (14.8) with boundary conditions (14.7).

Let us return to consideration of equation (14.6). From the formula obtained earlier it follows that the parameter λ cannot be negative. When $\lambda = 0$, the general solution of the equation has the form $x(t) = c_1 + c_2 t$. As a result of substitution into equalities (14.7), we determine $c_1 = 0$ and $c_2 = 0$, and therefore $x = 0$. Thus, the eigenvalues are positive, which means that the general solution to equation (14.6) is

$$x(t) \; = \; c_1 \sin \sqrt{\lambda} t + c_2 \cos \sqrt{\lambda} t,$$

where c_1, c_2 are arbitrary constants. Using the first of conditions (14.7), we find

$$x(0) \; = \; c_2 \; = \; 0.$$

Using the second equality (14.7), we obtain

$$x(1) = c_1 \sin\sqrt{\lambda} = 0.$$

Taking into account that if the constant c_1 is equal to zero, the solution to the boundary value problem again turns out to be zero, we have the equality

$$\sin\sqrt{\lambda} = 0.$$

Now we find

$$\lambda = \lambda_k = (k\pi)^2,\ k = 1, 2, \dots . \tag{14.9}$$

Thus, there is an infinite number of values of the Lagrange multiplier λ, i.e., eigenvalues[8]. The corresponding non-zero solutions to the boundary value problem (14.6), (14.7) are determined by the formulas

$$x_k(t) = c_k \sin k\pi t,\ k = 1, 2, \dots , \tag{14.10}$$

where c_k are arbitrary constants. The equalities (14.9) and (14.10) give a complete set of solutions to the Sturm–Liouville problem, i.e., corresponding eigenvalues and eigenfunctions.

We choose the constants c_k in such a way as to ensure the fulfillment of the isoperimetric condition (14.3). We have

$$\int_0^1 [x_k(t)]^2 dt = c_k^2 \int_0^1 \sin^2 k\pi t dt = 1.$$

Therefore, $c_k = \pm\sqrt{2}$. As a result, we define the functions[9]

$$x_k^+(t) = \sqrt{2}\sin k\pi t,\ \ x_k^-(t) = -\sqrt{2}\sin k\pi t,\ k = 1, 2, \dots .$$

The corresponding controls are equal

$$u_k^+(t) = \sqrt{2}k\pi\cos k\pi t,\ \ u_k^-(t) = -\sqrt{2}k\pi\cos k\pi t,\ k = 1, 2, \dots .$$

Thus, the system of optimality conditions (14.1)–(14.5) has the infinite set of solutions.

In order to choose the best of them, calculate the corresponding values of the optimality criterion

$$I(u_k^+) = I(u_k^-) = 2(k\pi)^2 \int_0^1 \cos^2 k\pi t dt = k^2\pi^2,\ \ k = 1, 2, \dots, .$$

The minimum of these values, corresponding to $k = 1$, is the minimum of functionality for Example 14.1. Thus, the solutions to this problem turn out to be controls

$$u_1^+(t) = \sqrt{2}\pi\cos\pi t,\ \ u_1^-(t) = -\sqrt{2}\pi\cos\pi t.$$

Therefore, the optimal control problem under consideration has two solutions, and the corresponding optimality conditions are necessary, but not sufficient[10]

14.1.2 Non-linear system of optimality conditions for a fixed final state

The possibility of an analytical solution to the optimal control problem posed in Example 14.1 is due to the fact that the state equation is linear both in control and in the state of the system, and the optimality criterion and the isoperimetric condition are quadratic. As a result, with respect to the three functions u, x, and p, we obtained a system that reduces to a linear boundary value problem (14.6) and (14.7), admitting an analytical solution. The situation becomes significantly more complicated with a slight change in any of the components of the problem statement[11].

Example 14.2 *There is a system described by the equalities*

$$x'(t) = u(t),\ t \in (0,1);\ \ x(0) = 0. \tag{14.11}$$

The optimal control problem is to find a function $u = u(t)$ that minimizes the functional

$$I(u) = \int_0^1 u(t)^2 dt$$

under the final condition

$$x(1) = 0 \tag{14.12}$$

and the isoperimetric condition[12]

$$\int_0^1 x(t)^4 dt = 1. \tag{14.13}$$

One can verify that optimal control for Example 14.2 exists[13]. Define the function

$$H(u, x, p, \lambda) = up + \lambda(x^4 - 1) - u^2/2.$$

The function p satisfies the adjoint equation

$$p'(t) = -4\lambda x(t)^3,\ t \in (0,1). \tag{14.14}$$

Control is found from the maximum condition

$$H(u, x, p, \lambda) = \max H(v, x, p, \lambda). \tag{14.15}$$

Thus, to find three unknown functions u, x, p, and the number λ, we have a system[14] that includes two first-order differential equations with two boundary conditions (14.11), (14.12), and (14.14), the problem for the unconditional extremum (14.15) and the equality (14.13)

To study the resulting system, as in the previous problem, we use the method of eliminating unknowns. By turning the derivative of the function H to zero, we find the control $u = p$, which actually delivers the maximum of this function. Then, differentiating the state equation (14.11) and taking into account the adjoint equation (14.14), we have

$$x'' = u' = p' = -4\lambda x^3. \tag{14.16}$$

The result is similar to equation (14.6) established for the previous example. However, unlike the latter, equation (14.16) turns out to be nonlinear, which is a consequence of the existing isoperimetric condition.

Multiplying the resulting equality by the function x and integrating the result taking into account the existing boundary conditions, we establish the relation

$$4\lambda \int_0^1 x^4 dt = -\int_0^1 x'' x dt = \int_0^1 x'^2 dt.$$

Taking into account the isoperimetric condition (14.13), we reduce the last equality to the form

$$4\lambda = \int_0^1 x'^2 dt.$$

Substituting this value into equation (14.16), we obtain the equality

$$x'' + x^3 \int_0^1 x'^2 dt = 0. \tag{14.17}$$

Thus, the optimal state of the system satisfies the ***integro-differential equation*** (14.17), similar to (14.8) with homogeneous boundary conditions

$$x(0) = 0, \ \ x(1) = 0. \tag{14.18}$$

Determine the value[15]

$$\|x\| = \sqrt{\int_0^1 x'^2 dt}.$$

Let us set the function

$$y(t) = \|x\|^2 x(t), \ t \in (0,1). \tag{14.19}$$

Using equation (14.17), we establish the equality

$$y'' = \|x\| x'' = \|x\|^3 x^3 = -y^3.$$

Thus, the function y satisfies the nonlinear ordinary differential equation of the second order[16]

$$y''(t) + y(t)^3 = 0, \ \ t \in (0,1) \tag{14.20}$$

with boundary conditions

$$y(0) = 0, \ \ y(1) = 0. \tag{14.21}$$

Let us assume that somehow a solution to problem (14.20) and (14.21) has been found. Condition (14.19) implies the equality

$$\|y\| = \|x\|^2.$$

Now we find

$$x(t) = \|x\|^{-1}y(t) = \|y\|^{-1/2}y(t),\ t \in (0,1). \tag{14.22}$$

Thus, having a solution to problem (14.20) and (14.21), we can determine the state function for Example 14.2 using formula (14.22). The required control is a derivative of this function.

Let us now consider the non-linear boundary value problem (14.20) and (14.21). Obviously, a function identically equal to zero is its solution. At the same time, if the function x is the optimal state for Example 14.2, then, due to the isoperimetric condition, it is different from zero, which was also true for the previous example. Then the function y determined by formula (14.19) is also non-zero and is a solution to the boundary value problem (14.20) and (14.21). Consequently, this boundary value problem does not have a unique solution. Along with the trivial one, there is also a solution that is different from zero and associated with the optimal state of the system under study[17].

A natural question arises: is the set of solutions to the boundary value problem (14.20) and (14.21) not limited to two elements? If this turned out to be the case, then the problem under study would have a unique solution. However, it is obvious that in the case when the function x is the optimal state of the system, the function $-x$ also satisfies the isoperimetric condition, and the values of the functional I coincide in both cases. Thus, the solution to the optimal control problem is certainly not unique: if control u is optimal, then so is control $-u$. Solutions of equations (14.17) and (14.20) are also invariant with respect to changing the sign: if a certain function is a solution to the given equations with homogeneous boundary conditions, then by changing the sign, in a non-trivial case we obtain a new solution to the same problems. Thus, the boundary value problem (14.20) and (14.21) has at least three solutions, one of which is trivial, and the other two correspond to the solutions of the studied optimal control problem[18].

Naturally, we do not yet know whether there are other solutions to the considered problems. However, the method of proving the existence of the last solution (the second for the optimality conditions and the third for the boundary value problem) suggests the direction of further research. Try to find a transformation that converts one solution to the problem into another.

By direct verification, we are convinced that in the case when the function y is a solution to the boundary value problem (14.20) and (14.21), then the functions

$$z_1(t) = y(1-t),\ \ z_2(t) = -y(1-t),\ \ t \in (0,1) \tag{14.23}$$

are also solutions to the same problem. In this case, two options are possible: either the function y satisfies one of the following conditions of symmetry with respect to the middle of the time interval (see Figure 14.1)

$$y(t) = y(1-t),\ t \in (0,1), \tag{14.24}$$

$$y(t) = -y(1-t),\ t \in (0,1), \tag{14.25}$$

and equalities (14.23) do not give new solutions, or the solution is not symmetric and then we obtain new solutions to the boundary value problem. If a non-trivial solution to problem (14.20) and (14.21) does not satisfy the symmetry conditions (14.24) or (14.25), then this problem has at least five solutions, and four non-trivial solutions determine the solutions to the optimal control problem[19].

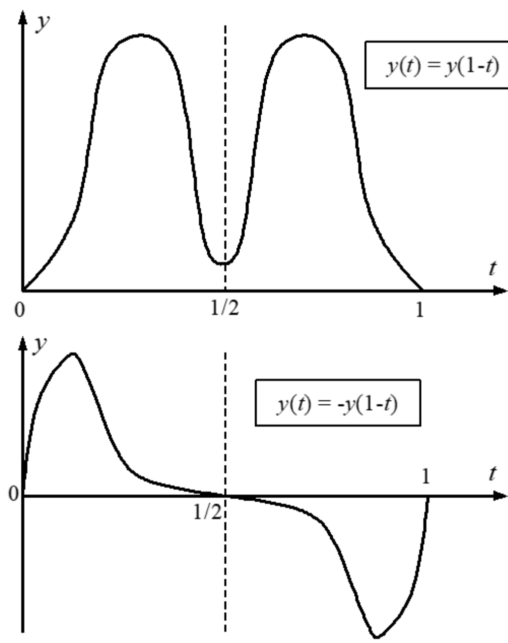

Figure 14.1 Form of possible symmetric solutions to the problem (14.20) and (14.21).

Thus, the boundary value problem (14.20) and (14.21) has at least three solutions (in the case of symmetry of the non-trivial solution), and possibly even five solutions (in the absence of symmetry). However, it is possible that the problem has other solutions.

Let us consider a boundary value problem of type (14.20) and (14.21) on an arbitrary interval

$$z''(t) + z(t)^3 = 0, \ t \in (0, a), \tag{14.26}$$

$$z(0) = 0, \ z(a) = 0, \tag{14.27}$$

where a is a positive number.

Suppose y is a non-trivial solution of problem (14.20) and (14.21). Define the function

$$z(t) = a^{-1}y(t/a), \ t \in (0, a). \tag{14.28}$$

The following equality holds

$$z''(t) + [z(t)]^3 = a^{-3}y''(t/a) + a^{-3}[y(t/a)]^3 = a^{-3}\{y''(t/a) + [y(t/a)]^3\} = 0.$$

Thus, we obtain a solution to problem (14.26) and (14.27).

For each value of the parameter a, transformation (14.28) associates a new solution to equation (14.20), which also satisfies the first of the boundary conditions (14.21). However, the need to ensure the second boundary condition significantly limits the choice of acceptable specified parameters.

Denote by y_1 a non-zero solution to problem (14.20) and (14.21), the existence of which is known. Then the function

$$z_2(t) = 2y_1(2t),\ t \in (0, 1/2)$$

is a solution to problem (14.26) and (14.27) for $a = 1/2$.

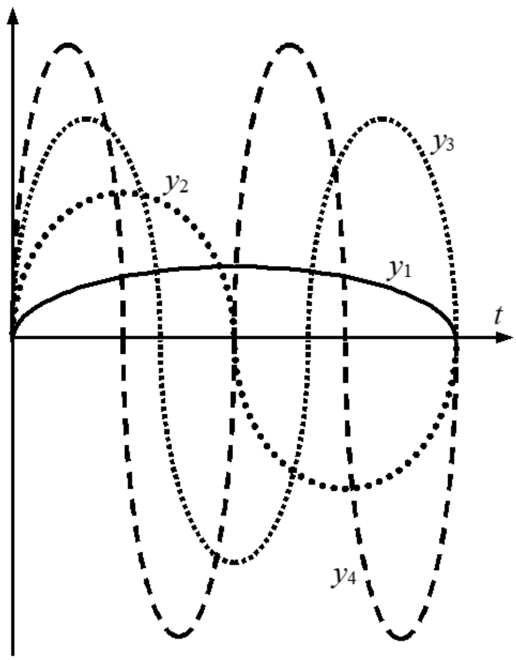

Figure 14.2 Solutions to problem (14.20) and (14.21).

Consider the function (see Figure 14.2)

$$y_2(t) = \begin{cases} z_2(t), & \text{if } 0 < t < 1/2, \\ -z_2(1-t), & \text{if } 1/2 < t < 1. \end{cases}$$

Obviously, the boundary condition (14.21) is satisfied for it. On the interval (0,1/2) it satisfies equation (14.20), since it is $z = z_2(t)$ a solution to equation (14.26) for $0 < t < 1/2$. At the same time, the function $z = -z_2(1-t)$ is a solution to the same equation for $1/2 < t < 1$. Note that the function $y2$ defined above at the point $t = 1/2$ is continuously differentiable (if, of course, the function y_1 is such) by construction. Considering that, according to equation (14.20), its second derivative is equal to the cube of the function taken with the opposite sign, we conclude that y_2 at the point $t = 1/2$ is twice continuously differentiable. As a result, we establish that function y_2

turns out to be a solution to problem (14.20) and (14.21), different from all previous ones. Another solution would be a function that differs from y_2 in sign[20].

Now we have an algorithm for constructing new solutions to the boundary value problem under study. Obviously the function

$$z_3(t) = 3y_1(3t),\ t \in (0, 1/3)$$

is the solution to problem (14.26) and (14.27) for $a = 1/3$. Then the function (see Figure 14.2)

$$y_3(t) = \begin{cases} z_3(t), & \text{if } 0 < t < 1/3, \\ -z_3(2/3 - t), & \text{if } 1/3 < t < 2/3, \\ z_3(t - 2/3), & \text{if } 2/3 < t < 1 \end{cases}$$

is also the solution to the boundary value problem under consideration.

In general, for any natural number k the function

$$z_k(t) = ky_1(kt), t \in (0, 1/k)$$

is the solution to problem (14.26) and (14.27) for $a = 1/k$. Repeating the above reasoning, we establish that the function

$$y_k(t) = \begin{cases} z_k(t), & \text{if } 0 < t < 1/k, \\ -z_k(2/k - t), & \text{if } 1/k < t < 2/k, \\ z_k(t - 2/k), & \text{if } 2/k < t < 3/k, \\ -z_k(4/k - t), & \text{if } 3/k < t < 4/k, \\ \cdots & \cdots \end{cases}$$

is also the solution our boundary value problem. One or three more solutions can be obtained by acting on this function with the above transformations[21]. Thus, the system of optimality conditions for Example 14.2 has an infinite number of solutions[22]. We encountered a similar situation in the previous example.

Now we can return to the study of the original optimization problem. First of all, in accordance with formula (14.22) using the known solution y_k of problem (14.20), (14.21) we find the solution

$$x_k(t) = \|y_k\|^{-1/2} y_k(t),\ t \in (0, 1)$$

of the integro-differential equation (14.17) with homogeneous boundary conditions. The corresponding control is As noted earlier, any non-zero solution to this problem satisfies the system of conditions for Example 14.2. Let us estimate the value of the functional

$$I(u_k) = \int_0^1 |x_k'(t)|^2 dt = \int_0^1 \left[\|y_k\|^{-1/2} y_k'(t)\right]^2 dt$$

$$= \|y_k\|^{-1} \int_0^1 |y_k'(t)|^2 dt = \|y_k\|^{-1} \|y_k\|^2 = \|y_k\|.$$

Taking into account the definition of the function y_k (see Figure 14.2), we find the integral

$$\int_0^1 |y_k'(t)|^2 dt = \|y_k\|^2 = \sum_{j=1}^{k} \int_{(j-1)/k}^{j/k} |y_k'(t)|^2 dt = k \int_0^{1/k} |y_k'(t)|^2 dt.$$

From the equality

$$y_k(t) = k y_1(kt), \; 0 < t < 1/k$$

it follows

$$\int_0^{1/k} |y_k'(t)|^2 dt = k^4 \int_0^{1/k} |y_1'(t)|^2 dt = k^3 \int_0^1 |y_1'(t)|^2 dt = k^3 \|y_1\|^2.$$

Now we get

$$I(u_k) = k^2 \|y_1\|^2, \; k = 1, 2, \dots .$$

Thus, among all the previously found solutions to the necessary conditions for an extremum, only that control that corresponds to the value y_1 (and the functions resulting from it after the action of previously determined transformations) delivers a minimum to the functional under consideration. To find the optimal state of the system, you should now use formula (14.22). By differentiating the resulting function, we find the optimal control[23]. Thus, this example largely has the same properties as the previous example[24].

RESULTS

Here is a list of questions based on the results of the lecture, the main conclusions on this topic, as well as problems arising in this case, which are solved partly in Appendix, partly in the subsequent chapter.

Questions

It is required to answer questions related to the previously given lecture material.

1. Obviously, the minimized functional for Example 14.1 takes exclusively non-negative values. Why is control $u = 0$, on which the functional is equal to zero, not a solution to the optimal control problem under consideration?

2. Why the system of optimality conditions (14.1)–(14.5) for a fixed value of λ turned out to be linear?

3. Why are only non-trivial solutions considered for the boundary value problem (14.6) and (14.7)?

4. Why is the general solution of equation (14.6) not represented in terms of exponentials?

5. Why was it necessary to establish the dependence of the parameter λ through x, if the equation (14.8) obtained on its basis is not used in further studies?

6. Are the controls u_k obtained by differentiating the functions x_k determined by formula (14.10) solutions to the system of optimality conditions for Example 14.1?

7. How was the isoperimetric condition used to find a solution to the system of optimality conditions for Example 14.1?

8. Which of the conditions of Theorem 5.1 on the uniqueness of optimal control is violated for Example 14.1?

9. What is the fundamental difference between the problem statements in Examples 14.1 and 14.2?

10. What is the fundamental difference between the system of optimality conditions for Examples 14.1 and 14.2 and what is the reason for this difference?

11. Why is the transition from equation (14.17) to equation (14.20) made?

12. Where do the boundary conditions for equation (14.20) come from?

13. Why, when using equality (14.28) to construct new solutions to the optimality conditions for Example 14.2, are exclusively the values $a = 1/k$ used, although the properties of equation (14.26) do not depend on this value?

14. Why cannot the analysis of Example 14.2 be considered complete?

15. How many optimal controls are there for Example 14.2?

16. Why can the properties of Example 14.2 be extended to the case where the isoperimetric condition contains an arbitrary even power greater than two?

17. What happens if the isoperimetric condition for Example 14.2 has an odd degree?

18. Is it possible to extend the results obtained in the study of Example 14.2 to the case when under the integral in the isoperimetric condition there is the value $|x|^r$, where $r > 1$?

Conclusions

Based on the analysis, we come to the following conclusions.

- The system of optimality conditions for Example 14.1 with a fixed value of the parameter λ is linear, which predetermined the possibility of a complete analysis of the problem.

- The linearity of the system of optimality conditions for Example 14.1 at a fixed value of the parameter λ is a consequence of the linearity of the equation of state and the quadraticity of the optimality criterion and the isoperimetric condition.
- The system of optimality conditions for Example 14.1 is reduced to a boundary value problem for an integro-differential equation.
- The optimality conditions for Example 14.1 reduce to the Sturm–Liouville problem.
- The optimality conditions for Example 14.1 have an infinite set of solutions.
- The solutions to the optimality conditions for Example 14.1 form an orthonormal family.
- The optimality conditions for Example 14.1 are not sufficient.
- In Example 14.1, there are two optimal controls that differ in sign.
- Optimal control problems with an isoperimetric condition are characterized by the absence of convexity of the set of admissible controls.
- The system of optimality conditions for Example 14.2 at a fixed value of the parameter λ is non-linear, which is due to the presence of the fourth degree in the isoperimetric condition.
- The system of optimality conditions for Example 14.2 is reduced to a boundary value problem for a non-linear integro-differential equation.
- Using a special substitution, the integro-differential equation considered in Example 14.2 can be reduced to a differential equation.
- A boundary value problem for a second-order differential equation with cubic non-linearity has an infinite number of solutions that differ significantly in their properties.
- The optimality conditions for Example 14.2 are not sufficient.
- In conditions of ambiguous solvability of the problem, you can try to find transformations that allow you to construct new solutions to the problem based on existing ones.
- The optimal control problem for Example 14.2 has at least two solutions.
- The results of the analysis of Example 14.2 were obtained under the assumption of the existence of optimal control.

Problems

Based on the results obtained above, we come to the following problems.

1. **Systems with a free final state**. Lecture examined optimal control problems with an isoperimetric condition in the case of a fixed final state. Problems with a free final state in the presence of an isoperimetric condition are also of interest. Examples of such problems are discussed in Appendix and in the subsequent chapter.

2. **Systems with constraints on control values**. In the examples considered, the set of admissible controls was determined exclusively by the isoperimetric condition and fixation of the final state of the system, as a result of which the means of classical calculus of variations could be used to solve such problems. It would be interesting to consider obtaining the described effects in the case of optimal control problems in the presence of isoperimetric conditions and explicit restrictions on the control values. Examples of such problems are discussed in Appendix and in the subsequent chapter.

3. **Sufficiency of optimality conditions in the absence of uniqueness of optimal control**. In the examples considered, both the sufficiency of optimality conditions and the uniqueness of optimal control were simultaneously violated. It would be interesting to give an example of an optimal control problem of the class under consideration, which would have a non-unique solution if the optimal control conditions are insufficient. One such example is given in Appendix.

4. **Uniqueness of optimal control is in the absence of sufficiency of optimality conditions**. In the examples in this lecture, the solution to the problem was not unique, and the optimality conditions were not sufficient. It would be natural to consider an example with a unique optimal control when the optimality condition is not sufficient. Such an example is given in Chapter 15.

5. **Optimal control problem with an infinite set of solutions**. In the previous sections, optimal control problems that admit of an infinite set of solutions were considered. One would like to get an example of an optimal control problem with an isoperimetric condition that has a similar property. Such an example is given in Appendix.

6. **Solvability of the optimal control problem for Example 14.2**. In Example 14.2, a non-linear system of optimality conditions with very non-trivial properties was obtained. However, the assumption of the existence of optimal control was used. The proof of the solvability of the problem under consideration is given in Appendix.

7. **Unsolvable optimization problems with isoperimetric conditions**. Previously, various extremal problems without solutions were considered. Similar problems are of interest in the presence of isoperimetric constraints. Such examples are given in Chapter 15.

8. **Singular controls in problems with isoperimetric conditions**. In the previous parts of the book, examples of various optimal control problems were given for which the maximum principle degenerates. It would be interesting to consider problems with isoperimetric conditions in which singular controls exist. Different problems of this type are given in Chapter 15.

9. **Ill-posed optimal control problems with isoperimetric conditions**. Previously, different ill-posed optimal control problems were considered. It would be interesting to consider ill-posed problems with isoperimetric conditions. Examples of such problems are given in Chapter 15.

10. **Bifurcation of extremals in optimal control problems with isoperimetric conditions**. Chapter 12 gave examples of optimal control problems with bifurcation of extremals. We would like to establish a similar effect for problems with isoperimetric conditions. Examples of such problems are given in Chapter 15.

11. **Optimal control problems with general phase constraints**. We limited ourselves to considering optimal control problems in which only restrictions on control values, fixation of the final state of the system and integral restrictions in the form of equality were allowed. However, in practice, problems arise with integral restrictions in the form of inequalities or restrictions on the values of the state function at individual points, including boundary ones (phase restrictions). For methods for solving similar problems, see Notes[25].

14.2 APPENDIX

We continue to study the problems of sufficiency of optimality conditions and uniqueness of solutions for optimal control problems with isoperimetric conditions. In particular, Section 14.2.1 considers an analog of Example 14.1 in the absence of a fixed final state. Subsection 14.2.2 provides an example of a problem of the class under consideration with constraints on control values, for which there are quite a lot of solutions. Section 14.2.3 proves the existence of an optimal control for Example 14.2.

14.2.1 Non-uniqueness and insufficiency for a system with a free final state

In the Lecture, optimal control problems with an isoperimetric condition were studied for the case when the final state of the system is fixed. Let us consider a similar problem for a system with a free final state.

Example 14.3 *There is a system described by the equalities*

$$x'(t) = u(t),\ t \in (0,1);\ \ x(0) = 0.$$

The optimal control problem is to find a function $u = u(t)$ that minimizes the functional

$$I(u) = \int_0^1 u(t)^2 dt$$

under the isoperimetric condition

$$\int_0^1 x(t)^2 dt = 1.$$

This problem differs from the one considered in Example 14.1 only in the absence of a fixed final state. Define the function

$$H(t,u,x,p,\lambda) = pf + \lambda q - g = pu + \lambda(x^2 - 1) - u^2$$

and the adjoint system

$$p'(t) = -2\lambda x(t),\ t \in (0,1);\ p(1) = 0.$$

From the maximum condition, we find

$$u(t) = p(t)/2.$$

The resulting system of optimality conditions differs from a similar problem (14.1)–(14.5) for Example 14.1 only in the presence of a final condition for the function p and its absence for the state function[26]. To analyze this system, we will use the same method as for the above example. Using the equation of state and the adjoint equation, we obtain the equality

$$x''(t) = u'(t) = p'(t)/2 = -\lambda x(t).$$

Thus, the function x is the solution of the equation

$$x''(t) + \lambda x(t) = 0,\ t \in (0,1), \tag{14.29}$$

which coincides with (14.6). However, instead of the second boundary value in a problem with a fixed final state, the following equality from the final condition of the conjugate system is used. Thus, equation (14.29) is considered with boundary conditions

$$x(0) = 0,\ x'(1) = 0. \tag{14.30}$$

As a result of the relatively unknown function x and parameter λ, the boundary value problem (14.29) and (14.30) is obtained, supplemented by the isoperimetric condition.

Note that the trivial (zero) solution of the boundary value problem does not satisfy the isoperimetric condition. Therefore, such a solution is not considered. We are again dealing with the ***Sturm–Liouville problem*** for the same equation, but with boundary conditions (14.30) instead of (14.7).

Multiplying equality (14.29) by x and integrating the result taking into account equalities (14.30) and the isoperimetric condition in the same way as was done for Example 14.1, we again obtain the equality

$$\lambda = \int_0^1 x'^2 dt,$$

whence the non-negativity of λ follows. When $\lambda = 0$, the general solution to equation (14.29) is determined by the formula $x(t) = c_1 + c_2 t$. As a result of substitution into equalities (14.30), we find $c_1 = c_2 = 0$, and therefore $x = 0$. Thus, the value of λ can only be positive, and the general solution of the equation has the form

$$x(t) = c_1 \sin \sqrt{\lambda} t + c_2 \cos \sqrt{\lambda} t,$$

where c_1 and c_2 are arbitrary constants. Using the first of the boundary conditions, we find $x(0) = c_2 = 0$. Using the second equality (14.30), we obtain

$$x'(1) = c_1 \sqrt{\lambda} \cos \sqrt{\lambda} = 0.$$

Taking into account that if the constant c_1 is equal to zero, the solution to the boundary value problem again turns out to be zero, we get $\cos \sqrt{\lambda} = 0$. From here we find the eigenvalues

$$\lambda = \lambda_k = \Big(\frac{\pi}{2} + k\pi\Big)^2, \ k = 1, 2, \dots .$$

Thus, problem (14.29) and (14.30) has non-trivial solutions

$$x_k(t) = c_k \sin \Big(\frac{1}{2} + k\Big)\pi t, \ k = 1, 2, \dots ,$$

where the constants c_k are arbitrary. To find them, we use the isoperimetric condition

$$\int_0^1 \big[x_k'(t)\big]^2 dt = (c_k)^2 \int_0^1 \sin^2 \Big(\frac{1}{2} + k\Big)\pi t dt = 1.$$

Thus, $c_k = \pm\sqrt{2}$, so for all k there exist two solutions

$$x_k^+ = \sqrt{2} \sin \Big(\frac{1}{2} + k\Big)\pi t, \ x_k^- = -\sqrt{2} \sin \Big(\frac{1}{2} + k\Big)\pi t, \ k = 1, 2, \dots .$$

The corresponding controls are defined as follows

$$u_k^+ = \sqrt{2}\Big(\frac{1}{2} + k\Big) \cos \Big(\frac{1}{2} + k\Big)\pi t, \ u_k^- = -\sqrt{2}\Big(\frac{1}{2} + k\Big) \cos \Big(\frac{1}{2} + k\Big)\pi t, \ k = 1, 2, \dots .$$

Thus, the system of optimality conditions for the considered example has an infinite set of solutions. The optimal control is the one that corresponds to the smallest value of the optimality criterion. We calculate

$$I(u_k^+) = I(u_k^-) = \Big(\frac{1}{2} + k\Big)^2 \pi^2 \int_0^1 \cos^2 \Big(\frac{1}{2} + k\Big)\pi t dt = \frac{1}{2}\Big(\frac{1}{2} + k\Big)^2 \pi^2, \ k = 1, 2, \dots .$$

The smallest of these values corresponds to the value $k = 1$. Thus, our problem has two solutions

$$u_1^+ = \frac{3}{\sqrt{2}}\pi \cos \frac{3\pi t}{2}, \ u_1^- = -\frac{3}{\sqrt{2}}\pi \cos \frac{3\pi t}{2}.$$

Thus, the optimal control problems with fixed and free final states considered in Examples 14.1 and 14.3 have similar properties[27].

14.2.2 System with constraints on control values

In the previous examples there were no restrictions on the control values, as a result of which these problems could be solved using the calculus of variations. Let us now consider a problem with explicit restrictions on control in a free final state.

Example 14.4 *The system characterized by the Cauchy problem*

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

The optimal control problem is to find a function $u = u(t)$ from the set

$$U = \{u |\, |u(t)| \leq 1,\ 0 < t < 1\}$$

under the isoperimetric condition

$$\int_0^1 x(t)dt = 0, \tag{14.31}$$

that minimizes the functional

$$I(u) = \int_0^1 u(t)^2 dt.$$

To obtain a minimization problem, it is enough to change the sign in the definition of the optimality criterion. Thus, we have Problem 13.1 with the following parameter values:

$$f(t,u,x) = u,\ T = 1,\ x_0 = 0,\ a(t) = -1,\ b(t) = 1,$$
$$g(t,u,x) = -u^2,\ h(x) = 0,\ q(t,u,x) = x.$$

Define the function

$$H(t,u,x,p,\lambda) = pu + \lambda x + u^2$$

and the adjoint system

$$p'(t) = -\lambda,\ t \in (0,1);\ p(1) = 0.$$

The control is determined from the condition for the maximum of the function H on a given set. Note that its unique stationary point $u = -p$ provides the minimum of this function due to the positivity of its second derivative. Consequently, the solution to the optimality condition is achieved on the boundary of the set of admissible controls. Find the corresponding boundary values

$$H|_{u=-1} = -p + \lambda x + 1/2,\ H|_{u=1} = p + \lambda x + 1/2.$$

We get

$$u(t) = \begin{cases} -1, & \text{if } p(t) < 0, \\ 1, & \text{if } p(t) > 0. \end{cases}$$

The solution to the adjoint system is $p(t) = \lambda(1–t)$. Considering that t takes values from the unit interval, we conclude that the function p is of constant sign. For $\lambda > 0$, the function p is positive everywhere, which means $u(t) = 1$. Then $x(t) = t$, and therefore condition (14.31) is violated. For $\lambda < 0$, the function p is negative everywhere, which means $u(t) = –1$. Then $x(t) = –t$, and therefore condition (14.31) is again not satisfied. There remains the case $\lambda = 0$, in which $p(t) = 0$. Then the function H takes the same value on both boundaries of the set of admissible controls. Consequently, the control at each point can equally take on the values 1 and –1. However, in this case equality (14.31) must certainly be satisfied. Let us try to determine under what conditions this is possible.

Let us note, first of all, that both the original formulation of the problem and the system of optimality conditions are again invariant with respect to a change of sign. Thus, if the function u satisfies the optimality conditions, then the function $–u$ also has similar properties. It was noted earlier that control cannot take the values 1 or –1 everywhere, i.e., optimality conditions can only be satisfied by a discontinuous function.

Let us assume that there is a single control discontinuity point $\xi \in (0, 1)$. Suppose $u(t) = 1$ for $t < \xi$ and $u(t) = –1$ for $t > \xi$. Then from the equation of state, we find $x(t) = t$ for $t < \xi$ and $x(t) = 2\xi–t$ for $t > \xi$. Find the integral

$$\int_0^1 x(t)dt \;=\; \int_0^\xi tdt \;+\; \int_\xi^1 (2\xi - t)dt \;=\; -\xi^2 + 2\xi - \frac{1}{2}.$$

Equating this value to zero, we establish a quadratic equation. It has two solutions $\xi = 1 + \sqrt{2}/2$ and $\xi = 1 - \sqrt{2}/2$. The first of them lies outside the interval (0,1). Thus, there is a single control discontinuity point at which the corresponding state function satisfies the isoperimetric condition. Thus, a control equal to 1 for t less than $1 - \sqrt{2}/2$ and –1 for t greater than this value is a solution to the system of optimality conditions. Naturally, a function that differs in sign from the specified control will also be a solution to the optimality conditions.

Let now there be two control discontinuity points ξ and η such that $0 < \xi < \eta < 1$. Suppose that the control is equal to 1 on the intervals $(0, \xi)$ and $(\eta, 1)$ and equal to –1 on the interval (ξ, η). The corresponding state function is equal to t on the interval $(0, \xi)$, $2\xi–t$ on (ξ, η) and $2\xi–2\eta + t$ on $(\eta, 1)$. We calculate the integral

$$\int_0^1 x(t)dt \;=\; \int_0^\xi tdt \;+\; \int_\xi^\eta (2\xi - t)dt \;+\; \int_\eta^1 (2\xi - 2\eta + t)dt \;=\; \xi^2 + \eta^2 + 2\xi - 2\eta + \frac{1}{2}.$$

We equate the result to zero in accordance with equality (14.31) and consider the resulting formula as an equation for η. It has two solutions

$$\eta_1 \;=\; 1 + \sqrt{1/2 - \xi^2 - 2\xi},\; \eta_2 \;=\; 1 - \sqrt{1/2 - \xi^2 - 2\xi}.$$

These equalities make sense if $\xi^2 + 2\xi - 1/2 < 0$, which, due to the positivity of ξ, holds when $\xi < \sqrt{3/2} - 1$. Here, point η_1 does not belong to the interval (0,1). Thus, there is an infinite and not even countable set of pairs of points ξ and η at which the control has a discontinuity (switching from value 1 to –1 or vice versa) so that condition (14.31) is satisfied. Each such pair corresponds to two controls that differ in signs and satisfy the system of optimality conditions.

It is easy to see that there is also an infinite number of solutions with three, four, etc. break-points (see in particular Figure 14.3). The question arises: which of these solutions to optimality conditions are optimal? Obviously, for any control from the set U, the value of the maximized functional does not exceed 1, and the equality $I(u) = 1$ is possible only in the case when $|u(t)| = 1$ for all[28] t. Thus, the control can only take on the values 1 or –1. However, in this case the isoperimetric condition (14.31) must be satisfied. All solutions to optimality conditions have precisely these properties. Thus, all solutions of the maximum principle are optimal, i.e., the problem under consideration has an infinite and not even countable set of solutions, and the optimality condition is necessary and sufficient[29].

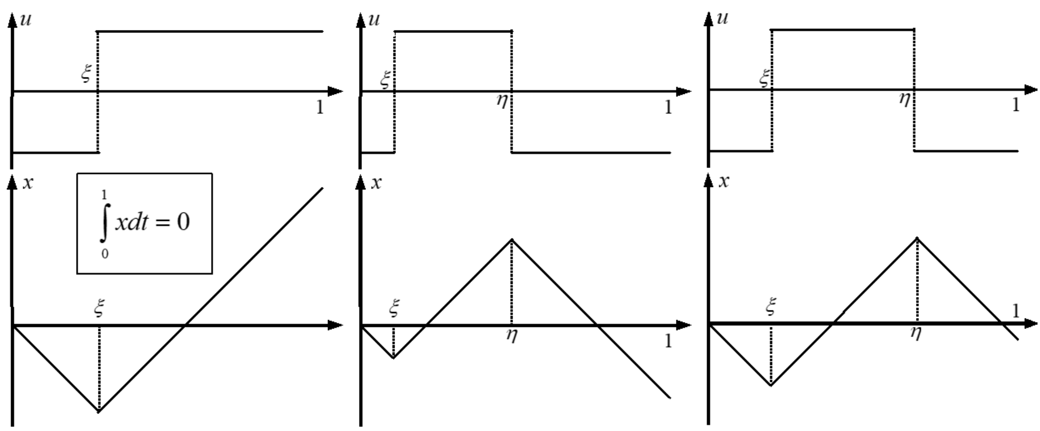

Figure 14.3 Optimal controls and states for Example 14.3.

14.2.3 Existence of optimal control for Example 14.2

We return to Example 14.2. It is about minimizing functional

$$I(u) = \int_0^1 u^2 dt$$

on a set of controls that guarantee the fulfillment of the condition

$$\int_0^1 x^4 dt = 1,$$

where x satisfies the equalities

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Justification of the solvability of this problem encounters significant difficulties due to the rather difficult structure of the set of admissible controls.

Let us define functional spaces that allow us to write the problem statement in a more concise form. The ***space*** $L_4(0,1)$ is the set of functions $x = x(t)$ that are Lebesgue integrable with the fourth degree[30] on the interval (0,1), i.e., satisfying the condition

$$\int_0^1 x(t)^4 dt < \infty.$$

Sobolev space $H_0^1(0,1)$[31] is a set of functions $x = x(t)$, equal to zero at points $t = 0$ and $t = 1$ and square-integrable along with its first generalized derivative on the interval (0,1), i.e., satisfying the inequalities

$$\int_0^1 x(t)^2 dt < \infty,\quad \int_0^1 x'(t)^2 dt < \infty.$$

The space $L_4(0,1)$ is Banach[32] with the norm

$$\|x\|_4 = \sqrt[4]{\int_0^1 x(t)^4 dt},$$

and the space $H_0^1(0,1)$ is Hilbert with the norm

$$\|x\| = \sqrt{\int_0^1 x'(t)^2 dt}.$$

The sequence $\{x_k\}$ is weakly convergent to an element x in the space $H_0^1(0,1)$ if we have

$$\int_0^1 x_k'(t)\lambda'(t)dt \to \int_0^1 x'(t)\lambda'(t)dt\ \ \forall \lambda \in H_0^1(0,1).$$

Define the space $X = H_0^1(0,1)$, the set

$$V = \Big\{x \in X \Big|\ \|x\|_4 = 1\Big\}$$

and the functional

$$J(x) = \|x\|^2.$$

Then Example 14.2 reduces to the problem of minimizing the functional J on a subset V of the space X. In this formulation, the proof of the existence of a solution can be carried out much simpler.

Since the functional J is bounded below (non-negative), its infimum exists on the set V. This means that there is a sequence $\{x_k\}$ from this set such that $J(x_k) \to \inf J(V)$. If the sequence $\{x_k\}$ turns out to be unbounded, i.e., for $k \to \infty$ we have $\|x_k\| \to \infty$, then from the definition of the minimized functional it would follow that $J(x_k) \to \infty$, which contradicts the statement that the sequence $\{x_k\}$ is minimizing. Thus, this sequence is limited. Using the Banach–Alaoglu theorem, we select from it a subsequence, which for simplicity, we also denote by $\{x_k\}$, such that $x_k \to x$ converges weakly in X.

In the theory of Sobolev spaces, the ***Rellich–Kondrashov theorem***[33] is known, according to which for $x_k \to x$ weakly in X, $x_k \to x$ converges strongly in $L_4(0,1)$, and therefore $\|x_k – x\|_4 \to 0$. Consider the difference

$$\int_0^1 x_k(t)^4 dt - \int_0^1 x(t)^4 dt = \int_0^1 \Big[|x_k(t)|^2 - |x(t)|^2\Big]\Big[|x_k(t)|^2 + |x(t)|^2\Big] dt.$$

Obviously, if a function belongs to the space $L_4(0,1)$, then its square is an element of the space $L_2(0,1)$. Using Schwarz inequality, we get

$$\Big|\int_0^1 x_k(t)^4 dt - \int_0^1 x(t)^4 dt\Big| \le \|x_k^2 - x^2\|_2 \|x_k^2 + x^2\|_2. \tag{14.32}$$

By the ***triangle inequality***[34], the norm of the sum of elements does not exceed the sum of their norms. Then we can estimate the second factor on the right side of inequality (14.32). We have

$$\|x_k^2 + x^2\|_2 \le \|x_k^2\|_2 + \|x^2\|_2 = \sqrt{\int_0^1 (x_k^2)^2 dt} + \sqrt{\int_0^1 (x^2)^2 dt} = \|x_k\|_4^2 + \|x\|_4^2.$$

Due to the convergence of $x_k \to x$ in $L_4(0,1)$, the sequence of norms $\{\|x_k\|_4\}$ is bounded. Then the last inequality implies the existence of a positive constant c_1 such that the estimate holds

$$\|x_k^2 + x^2\|_2 \le c_1.$$

We have

$$\|x_k^2 - x^2\|_2 = \int_0^1 \big(x_k^2 - x^2\big)^2 dt = \int_0^1 (x_k - x)^2 (x_k + x)^2 dt \le \|(x_k - x)^2\|_2 \|(x_k + x)^2\|_2$$

because of Schwartz inequality. Determine

$$\|(x_k + x)^2\|_2 = \|x_k + x\|_4^2 \le \Big(\|x_k\|_4 + \|x\|_4\Big)^2 \le c_2,$$

where c_2 is a positive constant.

Finally, from the equality

$$\|(x_k - x)^2\|_2 = \|x_k - x\|_4^2$$

and the convergence $x_k \to x$ in $L_4(0,1)$, it follows that the quantity on the left side of the last equality tends to zero. As a result, inequality (14.32) implies convergence

$$\int_0^1 |x_k(t)|^4 dt \to \int_0^1 |x(t)|^4 dt.$$

Then, passing to the limit in the equality $\|x_k\|_4 = 1$, following from the inclusion $x_k \in V$, we conclude that $x \in V$.

Obviously, the functional to be minimized is convex and continuous[35]. Then it is weakly lower semicontinuous (see Chapter 7), which means that the inequality

$$J(x) \le \inf \lim J(x_k).$$

Since the sequence $\{x_k\}$ minimizes on the set V, the value on the right side of the last inequality is $\inf J(V)$. Thus, on the set V there is an element x that satisfies the inequality $J(x) \le \inf J(V)$. However, x is an element of the set V, which means that the last relation can only be satisfied in the form of equality. Consequently, the infimum of the functional J on the set V is achievable, which means that the optimal control problem initially considered has a solution. It is characteristic that this result was obtained in the absence of convexity of the set of admissible controls[36], i.e., when the conditions of Theorem 7.1, and Theorem 7.1 are violated.

Additional conclusions

Based on the analysis of optimal control problems for systems with isoperimetric conditions presented in Appendix, the following additional conclusions can be drawn.

- The optimal control problem for a linear system with a quadratic optimality criterion and an isoperimetric condition, considered in Example 14.3, like a similar problem with a fixed final state, is reduced to a linear system of optimality conditions and then to the corresponding Sturm–Liouville problem.
- The system of optimality conditions for Example 14.3 has an infinite set of solutions, of which only two are optimal, which differ in sign.
- The optimality conditions for Example 14.4 are necessary and sufficient.
- The optimal control problem from Example 14.4 has an infinite and not even countable set of solutions.
- All optimal controls for Example 14.4 are discontinuous.
- An infinite and even non-countable set of solutions to an optimization problem can consist of non-singular controls.
- The existence of an optimal control can be proven even in the absence of convexity of the set of admissible controls, if it can be established that it is weakly closed.

Notes

1. It is curious that the set of admissible control values for the problem under consideration is certainly not convex, since the half-sum of two controls that differ in sign gives a function identically equal to zero, for which the isoperimetric condition is not satisfied. The isoperimetric condition characterizes the set of functions lying on the surface of the sphere. In particular, the isoperimetric condition reduces to the equality $\|x\|^2 = 1$, where the norm is understood in the sense of the space $L_2(0,1)$. This corresponds to the equation of a spherical surface of unit radius centered at zero; see [94], [106], [158]. Naturally, the center itself, i.e., a function identically equal to zero lies outside the sphere.

2. This problem has a natural physical meaning. Let us consider the process of oscillation of a spring on a time interval $[0,T]$. This is characterized by the function $x = x(t)$, which is the deviation of the spring from the equilibrium position at a given time. Let us estimate the kinetic and potential energies of the spring. Kinetic energy is determined by the formula $K = mv^2/2$, where m is the mass of the spring, and v is its velocity. Velocity is a derivative of the x coordinate, i.e., $v = x'$. As a result, we obtain the formula $K = mx'^2/2$. If the velocity were constant, then the kinetic energy of the system at a given time interval would be equal to the product of the value of K and the length of the time interval T. In the case of a variable velocity, the total kinetic energy is obtained as a result of integrating the function K over time, which gives the value

$$I = \frac{m}{2}\int_0^1 x'(t)^2 dt.$$

Potential energy is equal to the product of force and the distance traveled, i.e., $U = Fx$. In accordance with ***Hooke's law***, the elastic force acting on a spring is proportional to the deviation of the spring from the state of equilibrium, i.e., $F = kx$, where k is the elasticity coefficient. As a result, we obtain the formula $U = kx^2$. If the position of the spring did not change over time, then the potential energy of the spring at a given time interval would be equal to the product UT. When the position of the spring changes, the time function U is integrated over time. As a result, we get the value

$$J = k\int_0^1 x(t)^2 dt.$$

Let the initial and final states be given $x(0) = x_0$, $x(T) = x_T$. One can set the problem of finding a law of spring motion that, for a given value of potential energy J, ensures a minimum of its kinetic energy. The example considered earlier is a special case of this problem with $T = 1$, $x_0 = 0$, $x_T = 0$, $k = 1$, $m = 1$, and $J = 2$.

3. Due to the absence of explicit restrictions on control, the problem under consideration can be studied using the calculus of variations; see [37], [61], [208]. To do this, it is enough to replace the control in the optimality criterion with the function derivative.

4. For linear systems with a quadratic functional, we previously used the decoupling method; see Chapters 3 and 8. In this case, the optimality conditions were reduced to the Riccati equations. However, we cannot use a similar technique to study Example 14.1 due to the presence of an isoperimetric condition. Note also that on the right side of equation (14.4) in front of the unknown function x there is a coefficient λ, which itself is unknown. Thus, in general, the system of optimality conditions is nonlinear, and it is linear only for a fixed value of λ.

5. We have already used this technique many times, starting with Chapter 3.

6. Formula 14.6 is the Euler equation for the corresponding problem of the calculus of variations.

7. On the Sturm–Liouville problem; see, for example, [86]. It is closely related to the spectral theory of operators, see Dunford and is widely used in the theory of equations of mathematical physics; see [130].

8. The set of all eigenvalues forms the ***spectrum*** of the operator corresponding to a given boundary value problem.

9. It is curious that the found family of functions is ***orthonormal*** in the space $L_2(0,1)$, i.e., the scalar product of any two different functions of this family is equal to zero, and any two identical functions are equal to one.

10. Note that the given problem, like many studied earlier, is invariant with respect to changes in the sign (see Examples 3.1 and 9.1.). In particular, any two controls that differ only in sign correspond to the same value of the minimized functional. From here (due to the inadmissibility of a control that is identically equal to zero), it immediately follows that the problem cannot have a unique solution.

11. This problem differs from the one considered in Example 14.1 only in the type of isoperimetric condition. In particular, in that problem the integral was not the fourth, but the second power of the function x, and the integral itself was equal to two, although the latter circumstance, unlike the first, does not play a significant role.

12. The isoperimetric condition, as in the previous example, characterizes a spherical surface of unit radius with a center at zero. However, in this case we consider the space $L_4(0,1)$ of functions that are integrable with the fourth power on the unit interval; see Appendix for more details, where the study of this example will be continued. Naturally, we again encounter a non-convex set of admissible controls.

13. A proof of the problem solvability for Example 14.2 is given in Appendix.

14. Optimality conditions here can also be obtained using the calculus of variations; see [37], [61], [208].

15. The number $\|x\|$ is the norm of x in the Sobolev space $H_0^1(0,1)$ of functions that are square-integrable by Lebesgue together with their derivative on the interval (0,1) and take zero values on the boundary of this interval; see Section 14.2.3 for more details.

16. [118] considers a more general boundary value problem for the case of a multidimensional equation. In particular, instead of an ordinary second-order differential equation, the corresponding equation of elliptic type is considered.

17. The trivial solution cannot be an optimal control, since it does not satisfy the isoperimetric condition.

18. Naturally, this is true provided that optimal control exists. This fact will be established in Appendix.

19. Let us denote by Y the set of non-trivial solutions to the boundary value problem. As we already know, at least three transformations are defined on it, which carry out the transition from one element of the set Y to another (see Figure 14.4):

$$A_1y(t) = -y(t),\ A_2y(t) = y(1-t),\ A_3y(t) = -y(1-t),\ t \in (0,1).$$

We can also consider the identity transformation on Y, which we denote by A_0. Obviously,

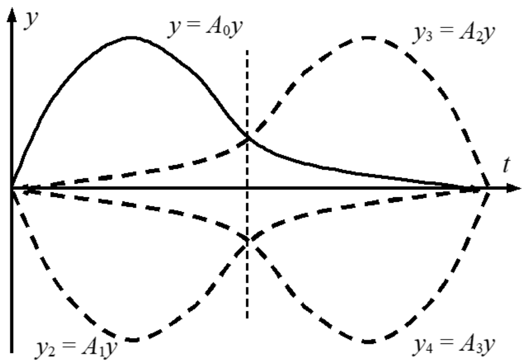

Figure 14.4 Possible asymmetric solutions to the problem (14.20) and (14.21).

the composition of transformations acting on the set Y also turns out to be a transformation of the set Y into itself. In this regard, one would expect that by performing the composition of the indicated transformations, we would be able to find new solutions to the problem. By direct verification, you can make sure that the composition does not take the indicated transformations beyond the class Y, i.e., is an operation (more precisely, a group operation) on this set, see the table called the ***Cayley table***; see Table 14.1.

TABLE 14.1 Cayley table.

	A_0	A_1	A_2	A_3
A_0	A_0	A_1	A_2	A_3
A_1	A_1	A_0	A_3	A_2
A_2	A_2	A_3	A_0	A_1
A_3	A_3	A_2	A_1	A_0

Moreover, the set of transformations Y with the indicated composition forms an ***Abelian group*** (the composition of transformations is associative, commutative, there is an identity transformation, and each transformation is invertible; see [192]).

20. If the function y_1 does not satisfy the symmetry conditions (14.24) or (14.25), then using the previously defined transformations A_2 and A_3, two more new solutions to the boundary value problem can be found.

21. As already noted, to construct solutions to equation (14.22), one can choose any value of the parameter a. However, the fulfillment of the second boundary condition (14.23) can be achieved only when $a = 1/k$, where $k = 1, 2, ...$ It should be noted that we are not completely sure that there are no other transformations that determine new solutions to the problem.

22. The results obtained also make it possible to establish the amazing properties of one non-linear equation of parabolic type associated with equation (14.22). Let us consider the non-linear ***heat equation***

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial \xi^2} + v^3, \ \xi \in (0,1), \ \tau > 0$$

with boundary conditions

$$v(\xi, 0) = v_0(\xi), \ \xi \in (0,1),$$
$$v(0,\tau) = 0, \ v(1,\tau) = 0, \ \tau > 0$$

It is required to establish the behavior of this system with an unlimited increase in the parameter τ. Obviously, if at $\tau \to \infty$ the convergence $v(\xi, \tau) \to y(\xi)$ takes place, then the function y, which is the ***equilibrium position*** of the system under consideration, is a solution to the differential equation $y'' + y^3 = 0$ with homogeneous boundary conditions. Thus, the equilibrium position for the equation under study turns out to be a solution to the boundary value problem (14.20) and (14.21). Then, based on the above analysis, we can come to the following conclusion that the considered non-stationary system under has an infinite number of equilibrium positions. The implementation of one or another equilibrium position is determined by the choice of a specific value of the initial state of the system v_0. Note also that the considered non-stationary system can be used to practically find non-trivial solutions to problem (14.20) and (14.21).

23. Naturally, we never found the function y_1 itself, but only proved its existence with the corresponding set of properties. To finally solve the problem, a practical solution to problem (14.20) and (14.21) is necessary. Due to the non-linearity of the equation, this will require some approximate method for solving boundary value problems for differential equations. However, due to the significant ambiguity of the solution, a direct search for function y_1 may be associated with significant difficulties. Note also that we do not know whether the function y_1 satisfies the symmetry conditions (14.22) or (14.23), and therefore whether the studied optimization problem has two or four solutions. This question can be answered after practically finding the function y_1. Strictly speaking, we are not even sure that we have found all the solutions to the boundary value problem (14.20) and (14.21), and therefore the necessary condition for the extremum (14.17). If there is at least one more such function, then using the above method one can determine another infinite set of solutions to the existing boundary value problem.

24. Similar results can be obtained by replacing the fourth power in the isoperimetric condition with an arbitrary odd power greater than two. Indeed, if in relation (14.13) instead of 4 an arbitrary power $2m$ is given, where m is a natural number greater than one, then instead of (14.15) we obtain the integro-differential equation

$$x'' + x^{2m-1}\|x\|^2 = 0.$$

As a result of substitution $y(t) = \|x\|^{1/(m-1)}x(t)$, the last formula is reduced to the non-linear differential equation

$$y'' + y^{2m-1} = 0,$$

possessing properties similar to those of the considered equation (14.20).

25. Optimal control problems with general constraints are considered; for example, in [5], [56], [95], [140], [193].

26. Naturally, the presence or absence of a fixed final state does not in any way affect the definition of the function H, and therefore, the form of the conjugate system and the maximum condition.

27. In Example 14.3, as in Example 14.1, the problem statement is invariant with respect to the change of control sign, and zero control is unacceptable due to the isoperimetric condition. Naturally, if in Example 14.1 the final state is non-zero, then the first property is violated, and the result may turn out to be qualitatively different.

28. More precisely, for almost everyone.

29. In Chapter 5, it was noted that the sufficiency of optimality conditions is guaranteed in the case of non-negativity of the remainder term in the formula for the increment of the functional.

In general, it is determined by the formula

$$\eta = \eta_3 - \int_0^T (\eta_1 + \eta_2) dt.$$

where η_3 corresponds to the second-order term obtained as a result of transforming the part of the optimality criterion characterizing the state of the system at the final moment of time, η_1 is associated with the second-order terms when expanding the value $H(t,u,x+\Delta x,p)$ into a series in terms of Δx, and $\eta_2 = [H_x(t,v,x,p)–H_x(t,u,x,p)]\Delta x$. For the example under consideration, there is no terminal term, and the function H is defined by the formula $H = pu + \lambda x + u^2$. Then $\eta = 0$, as a result of which the established sufficiency of optimality conditions is quite natural.

30. In general, the ***space*** $L_p(\Omega)$ is the set of Lebesgue measurable functions in the domain Ω of arbitrary dimension and integrable there with degree $p \geq 1$, i.e., such for which the Lebesgue integral of the modulus of a function raised to the power p is bounded. About spaces $L_p(\Omega)$; see, for example, [94], [106], [158].

31. In general, ***Sobolev space*** $W_p^m(\Omega)$ is the set of functions belonging to the space $L_p(\Omega)$ together with their generalized derivatives up to order m. Space $W_p^0(\Omega)$ is identified with $L_p(\Omega)$. For $p = 2$, the Sobolev space is Hilbert and is denoted by $H^m(\Omega)$. The space $H_0^1(0,1)$ is a subspace of functions from $H^1(\Omega)$, that take zero values on the boundary of the region Ω. On Sobolev spaces; see [1], [181].

32. Banach spaces are complete normed spaces; see [94], [106], [158]. All spaces $L_p(\Omega)$ are Banach, but only $L_2(\Omega)$ is Hilbert.

33. According to the ***Rellich–Kondrashov theorem***, in the case of a bounded domain Ω, from the weak convergence of $x_k \to x$ in the Sobolev space $W_p^m(\Omega)$ follows its strong convergence in the space $W_q^l(\Omega)$ where $1 < p < \infty$, $1 < q < \infty$, $n(1/p–1/q) < (m–l)$, n is the dimension of the set Ω. In the our case $\Omega = (0,1)$, $n = 1$, $p = 2$, $m = 1$, $q = 4$, $l = 0$. About the Rellich–Kondrashov theorem; see, for example, [1].

34. The triangle inequality is included in the definition of norm; see [94], [106], [158].

35. The convexity and continuity of the integral of the square of a function has been used many times before. The presence under the integral not of the function itself, but of its derivative does not change the properties of the functional, since the differentiation operator is linear.

36. In fact, the convexity condition for the set of admissible controls in the existence theorems from Chapter 7 was used solely to justify the weak closedness of this set. If this property can be established in the absence of its convexity, then the solvability of the problem can also be established.

CHAPTER 15

Different counterexamples for optimization problems with isoperimetric conditions

In the previous two chapters, problems of optimal control of systems with isoperimetric conditions were considered. In the first of them, general theoretical results were described, and in the second one we analyzed examples of problems with violation of the uniqueness of optimal control and the sufficiency of optimality conditions. This chapter provides examples of similar problems, both with free and fixed final states, for which other effects are implemented. We have previously encountered them in previous parts of the book. The Lecture describes unsolvable problems, problems with singular controls, and ill-posed problems of the indicated class. Appendix discusses some additional properties, in particular, bifurcation of extremals and violation of Bellman optimality principle.

15.1 LECTURE

In Chapter 13, optimality conditions were given for optimal control problems in the presence of isoperimetric conditions. The effectiveness of research methods was illustrated with relevant examples. In Chapter 14, examples of similar problems were considered in the absence of sufficiency of optimality conditions and uniqueness of optimal control. However, when studying extremal problems of other classes, we previously encountered other difficulties. This lecture gives examples of problems of the specified class, for which there is no optimal control, there are singular controls and the effects of ill-posedness of problems appear.

15.1.1 Insolvability of a problem with a free final state

Consider the following problem.

DOI: 10.1201/9781003398585-15

Example 15.1 *Find a function $u = u(t)$, which minimizes the functional*

$$I(u) = \int_0^1 [x(t)^2 - u(t)^2]dt$$

on the subset of functions

$$U = \{u \,|\, |u(t)| \leq 1, \; t \in (0,1)\}$$

satisfying the isoperimetric condition

$$\int_0^1 u(t)^2 dt = 1, \tag{15.1}$$

where the function x is determined by the equalities

$$x'(t) = u(t), \; t \in (0,1); \; x(0) = 0. \tag{15.2}$$

This optimal control problem differs from the one considered in Example 7.1 only by the presence of the additional isoperimetric condition (15.1). Let us estimate from below the values of the functional, as was done in Chapter 7. Taking into account the definition of the set U, we obtain the inequality

$$I = \frac{1}{2}\int_0^1 x^2 dt - \frac{1}{2}\int_0^1 u^2 dt \geq \frac{1}{2}.$$

Consider the control sequence determined by the equalities[1]

$$u_k(t) = \begin{cases} 1, & \text{if } \frac{2j}{2k} \leq t < \frac{2j+1}{2k}, \\ -1, & \text{if } \frac{2j+1}{2k} \leq t < \frac{2j+2}{2k}, \end{cases} \tag{15.3}$$

where $j = 0, 1, ..., k$–1. All its elements belong to the set U and satisfy the equality (15.1). The corresponding sequence of the solutions to problem (15.2) takes the form[2]

$$x_k(t) = \begin{cases} t - \frac{2j}{2k}, & \text{if } \frac{2j}{2k} \leq t < \frac{2j+1}{2k}, \\ \frac{2j+2}{2k} - t, & \text{if } \frac{2j+1}{2k} \leq t < \frac{2j+2}{2k}, \end{cases} \tag{15.4}$$

where $j = 0, 1, ..., k$–1. Obviously, we have the inequality

$$0 \leq x_k(t) \leq 1/2k, \; t \in (0,1), \; k = 1, 2, \ldots. \tag{15.5}$$

Then for all t, we have the convergence $x_k(t) \to 0$ as $k \to \infty$. Now we get $I(u_k) \to$ –1/2. Thus, the 1/2 point is the lower bound of the functional, and the sequence $\{u_k\}$ is minimizing.

Let us now assume that on some admissible control u the lower bound of the functional is achieved. This is only possible at $x = 0$. According to equation (15.2), this means that $u = 0$. However, this control is not admissible, since it does not satisfy condition (15.1). Thus, the infimum of the functional on the subset of controls from U that satisfy the isoperimetric condition is not achieved. Thus, the optimal control problem under consideration has no solution[3].

Previously, we have already encountered a situation where, in the absence of convexity of the functional, a solution to the problem of its minimization does not exist[4]. Let us now give an example of a problem with a convex functional for which optimal control also does not exist.

Example 15.2 *Find a function $u = u(t)$, which minimizes the functional*

$$I(u) = \int_0^1 x(t)^2 dt$$

on the subset of functions

$$U = \{u \,|\, |u(t)| \leq 1, \ t \in (0,1)\}$$

satisfying the isoperimetric condition

$$\int_0^1 u(t)^2 dt = 1,$$

where the function x is determined by the equalities

$$x'(t) = u(t), \ t \in (0,1); \ x(0) = 0.$$

This problem differs from the previous one only in the absence of a control square in the optimality criterion, as a result of which this functional turns out to be convex[5].

Note that the functional takes only non-negative values. Consider the same sequence (15.3) as in the previous example. Obviously, all its elements are admissible. The corresponding sequence of states $\{x_k\}$ is again characterized by equalities (15.4), which means that inequalities (15.5) and the convergence of $x_k(t) \to 0$ as $k \to \infty$ are still satisfied. It follows that $I(u_k) \to 0$. Thus, the point 0 is the lower bound of the functional, and the sequence $\{u_k\}$ is minimizing.

Let us now assume that on some admissible control u the lower bound of the functional is achieved. This is only possible at $x = 0$, which corresponds to control $u = 0$. However, this control is not admissible, since it contradicts the isoperimetric condition. Thus, the lower bound of the functional on the set of admissible controls is not achieved, i.e., the considered optimal control problem has no solution[6].

15.1.2 Insolvability of problem with a fixed final state

We have given examples of optimal control problems for systems with an isoperimetric condition in the presence of restrictions on control values with a free final state. Now we consider the case when, on the contrary, there are no restrictions on the control values, and the final state of the system is fixed. Example 14.1 considered an optimal control problem with an optimality criterion that is quadratic with respect to the control and an isoperimetric condition that is quadratic with respect to the system state function. Let us consider the opposite situation.

Example 15.3 *We have the following system*

$$x'(t) \;=\; u(t),\; t \in (0,1);\;\; x(0) \;=\; 0. \tag{15.6}$$

The optimal control problem is to find a function $u = u(t)$ that minimizes the functional

$$I(u) \;=\; \int_0^1 x(t)^2 dt$$

under the final condition

$$x(1) = 0 \tag{15.7}$$

and the isoperimetric condition

$$\int_0^1 u(t)^2 dt \;=\; 1. \tag{15.8}$$

Determine the function

$$H(t,u,x,p,\lambda) = pu + \lambda(u^2 – 1) - x^2.$$

Then the adjoint equation takes the form

$$p'(t) \;=\; 2x(t),\; t \in (0,1). \tag{15.9}$$

Equaling the derivative of H with respect to the control to zero, we get

$$u(t) = p(t)/2\lambda. \tag{15.10}$$

However, this value corresponds to the maximum of this function only for negative values of the parameter λ. For its other values, the maximum condition has no solution.

The resulting system of optimality conditions differs from the similar system (14.1)–(14.5) for Example 14.1 in that the parameter λ is present not in the adjoint equation, but in the control formula. However, the analysis of optimality conditions in this case is carried out in the same way as in Example 14.1. To eliminate the functions u and p, we substitute the control from formula (15.10) into the equation of state. Then after differentiation, taking into account the adjoint equation, we find

$$x''(t) \;=\; u'(t) \;=\; \lambda^{-1} x(t).$$

Thus, the function x satisfies the second order differential equation

$$x''(t) - \lambda^{-1}x(t) = 0, \; t \in (0,1). \tag{15.11}$$

with boundary conditions

$$x(0) = 0, \quad x(1) = 0. \tag{15.12}$$

The boundary value problem (15.11) and (15.12) includes the unknown parameter λ and is supplemented by the isoperimetric condition (15.8). Note that the trivial solution of the boundary value problem in the form of equality (15.6) corresponds to a zero control, which does not satisfy formula (15.8). Thus, only a non-trivial solution makes sense, which means we again obtain the Sturm–Liouville problem. Equation (15.11) differs from (14.6) only in that instead of the value λ, there is the value $-\lambda^{-1}$.

Since the solution of optimality conditions is possible only for negative λ, the general solution to equation (15.11) has the form

$$x(t) = c_1 \sin \frac{t}{\sqrt{-\lambda}} + c_2 \cos \frac{t}{\sqrt{-\lambda}},$$

where c_1, c_2 are arbitrary constants. Using the first equality (15.12), we get $x(0) = c_2 = 0$. From the second equality (15.12), it follows

$$x(1) = c_1 \sin \frac{1}{\sqrt{-\lambda}} = 0.$$

The constant c_1 is not equal to zero[7], so we get

$$\lambda = \lambda_k = -\frac{1}{(k\pi)^2}, \; k = 1, 2, \dots . \tag{15.13}$$

Thus, there exists an infinite set of Lagrange multipliers λ such that the boundary problem (15.11) and (15.12) has a non-zero solution[8]

$$x_k(t) = c_k \sin k\pi t, \; k = 1, 2, \dots ,$$

where c_k is a constant. Find the corresponding control

$$u_k(t) = k\pi c_k \cos k\pi t, \; k = 1, 2, \dots .$$

This function must satisfy the isoperimetric condition (15.8). Calculate the integral

$$\int_0^1 u_k^2 dt = (k\pi c_k)^2 \int_0^1 \cos^2 k\pi t = 1.$$

Now we obtain

$$c_k = \pm \frac{\sqrt{2}}{k\pi}.$$

Thus, the system (15.8), (15.11), and (15.12) has an infinite set of solutions

$$x_k^+(t) = \frac{\sqrt{2}}{k\pi} \sin k\pi t, \;\; x_k^-(t) = -\frac{\sqrt{2}}{k\pi} \sin k\pi t, \; k = 1, 2, \dots .$$

The corresponding controls are

$$u_k^+(t) = \sqrt{2}\cos k\pi t, \;\; u_k^-(t) = -\sqrt{2}\cos k\pi t, \;\; k = 1, 2, \dots .$$

Calculate the value of the optimality criterion for choosing the best of them

$$I(u_k^+) = I(u_k^-) = \frac{2}{k^2\pi^2}\int_0^1 \sin^2 k\pi t dt = \frac{1}{k^2\pi^2}, \;\; k = 1, 2, \dots .$$

Obviously, the sequences $\{I(u_k^+)\}$ and $\{I(u_k^-)\}$ are decreasing and tend to zero. By definition, the functional to be minimized is non-negative. Note that $\{u_k^+\}$ and $\{u_k^-\}$ are sequences of admissible controls, since for any of their elements both finite and isoperimetric conditions are satisfied. Thus, they turn out to be minimizing sequences. However, the optimality criterion can vanish only at $x = 0$. This state function is a solution to problem (15.6) under control $u = 0$, which does not satisfy the isoperimetric condition (15.8). Therefore, the minimum of the functional is not achieved on the entire set of admissible controls[9].

Thus, the considered optimal control problem has no solution. The maximum principle here is not a sufficient condition for optimality. The system of optimality conditions has an infinite set of solutions, which, surprisingly, form two minimizing sequences for the studied problem[10]

15.1.3 Singular controls

The following examples are related to the degeneration of the maximum principle for optimal control problems with isoperimetric constraints.

Example 15.4 *The optimization control problem is to find a function $u = u(t)$, which minimizes the functional*[11]

$$I(u) = \int_0^1 x(t)^2 dt$$

on the subset of functions

$$U = \{u \,|\, |u(t)| \leq 1, \; t \in (0, 1)\}$$

satisfying the isoperimetric condition

$$\int_0^1 u(t)dt = 0,$$

where the function x is the solution to the problem

$$x'(t) = u(t), \; t \in (0, 1); \;\; x(0) = 0.$$

Example 15.5 *Consider the maximization problem for the functional from Example* 15.4 *on the same set.*

The function H for both problem is determined by the same formula

$$H = pu + \lambda u - x^2/2,$$

where p is the solution to the problem

$$p'(t) \; = \; x(t), \; t \in (0,1); \;\; p(1) = 0.$$

Here, degeneration of the maximum principle is possible if the equality $p + \lambda = 0$ is satisfied, which implies $p(t) = -\lambda$. Thus, the function p is constant, which means its derivative is zero. Then from the adjoint equation, we obtain that $x = 0$. Substituting this value into the state equation, we conclude that $u = 0$. Therefore, in Examples 15.4 and 15.5 there is a unique singular control that is zero.

Now we check the validity of Kelley condition. Finding the derivative

$$\frac{\partial H}{\partial u} \; = \; p(t) + \lambda.$$

Differentiate this equality by t using the adjoint equation. We obtain

$$\frac{d}{dt}\frac{\partial H}{\partial u} \; = \; p'(t) \; = \; 2x(t).$$

After the next differentiation with using the state equation, we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} \; = \; 2x'(t) \; = \; 2u(t).$$

Now we find

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} \; = \; 2.$$

For Example 15.4 (minimization problem), the term on the left side of the resulting equality must not be negative, and for Example 15.5 (maximization problem), it must not be positive. Thus, the Kelley condition is satisfied for the first example, but not for the second one. Therefore, the found singular control $u = 0$ for Example 15.4 may be optimal, but for Example 15.5, it cannot be.

Obviously, the minimized functional is non-negative, and equality to zero is possible only for $x = 0$. From the state equation, it follows that this is implemented for control $u = 0$. It belongs to the set U and satisfies the isoperimetric condition, which means it is admissible. Therefore, it is the unique optimal control for Example 15.4 and cannot be such for Example 15.5.

In the examples considered, singular control was the only one. Here is an example of a problem with many singular controls.

Example 15.6 *Consider the system, which is described by the Cauchy problem*

$$x'(t) = u(t),\ t \in (0,1);\ \ x(0) = 0.$$

The control $u = u(t)$ is chosen from the set

$$U = \{u \,|\, |u(t)| \leq 1,\ t \in (0,1)\}$$

with isoperimetric condition

$$\int_0^1 (ut - x)dt = 0.$$

The optimal control problem is to minimize the functional

$$I(u) = \int_0^1 \Big(\frac{x^3}{3} - \frac{x^2 t}{2}\Big)dt.$$

Determine the function

$$H = pu + \lambda(ut - x) - \frac{x^3}{3} + \frac{x^2 t}{2},$$

where p is the solution of the adjoint system

$$p'(t) = x(t)^2 - tx(t) + \lambda,\ t \in (0,1);\ \ p(1) = 0.$$

Based on the form of the function H, we conclude that the degeneration of the maximum principle is possible only for $p(t) = -\lambda t$. This implies the equality $p'(t) = -\lambda$. Comparing the resulting equality with the adjoint equation, we conclude that $x(t)^2 - tx(t) + 2\lambda = 0$. This is possible at least for $\lambda = 0$ in two cases: for $x(t) = 0$ and for $x(t) = t$. The resulting states of the system are implemented, respectively, under the controls $u(t) = 0$ and $u(t) = 1$. Obviously, both functions found belong to the set U and ensure the fulfillment of the isoperimetric condition[12].

Thus, for our example, there are at least two singular controls[13]. We check the validity of Kelley condition for them. The derivative of the function H with respect to control is equal to $p + \lambda t$. Taking into account the form of the adjoint system, we find

$$\frac{d}{dt}\frac{\partial H}{\partial u} = x(t)^2 - tx(t).$$

Using the state equation, we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = (2x - t)x' - x = (2x - t)u - x.$$

Finally, we calculate

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 2x - t.$$

Let us establish the sign of this value on singular controls. When $u = 0$, we have $x = 0$, which means $2x–t = −t$, which is negative. Thus, the Kelley condition is not satisfied, which means that the singular control $u = 0$ is not optimal. When $u = 1$, we have $x(t) = t$, which means $2x–t = t$, which is positive. Thus, the singular control $u = 1$ can be optimal. Therefore, the optimality condition in this case is necessary, but not sufficient[14].

By analogy with Example 15.6, we can consider optimal control problems in the presence of an isoperimetric condition with an arbitrary number of singular controls[15].

15.1.4 Ill-posed problems

Chapter 6 provided examples of ill-posed optimal control problems in the absence of isoperimetric conditions. Similar problems occur in the presence of similar restrictions. Let us return, in particular, to Example 15.4, which considers the problem of minimizing the functional

$$I(u) = \frac{1}{2}\int_0^1 x^2 dt$$

on the subset of functions

$$U = \{u \,|\, |u(t)| \le 1,\ t \in (0,1)\},$$

satisfying the isoperimetric condition

$$\int_0^1 u dt = 0,$$

where the function x is described by the equalities

$$x'(t) = u(t),\ t \in (0,1);\ \ x(0) = 0.$$

Obviously, the minimized functional is non-negative, and equality to zero here is possible only for $x = 0$. According to the given equation, this is possible for control identically equal to zero. The latter is an element of the set U and satisfies the isoperimetric condition. Thus, this optimal control problem has a unique solution $u = 0$. The corresponding minimum of the functional is equal to zero.

Consider the sequence

$$u_k(t) = \cos k\pi t,\ k = 1, 2, \dots .$$

It belongs to the set U, besides the following equality holds

$$\int_0^1 u_k(t)dt = \int_0^1 \cos k\pi t dt = \frac{1}{k\pi}\sin k\pi t\Big|_0^1 = 0,\ \ k = 1, 2, \dots .$$

Thus, the isoperimetric condition is also satisfied. Therefore, the considered sequence of controls is admissible. The corresponding states of the system are determined by the formula

$$x_k(t) = \frac{\sin k\pi t}{k\pi}, \ k = 1, 2, \dots .$$

It follows that $x_k(t) \to 0$ at $k \to \infty$ for any t. Thus, the convergence $I(u_k) \to 0$ takes place. As a result, we conclude that the sequence $\{u_k\}$ is minimizing. However, it does not converge to optimal control in any reasonable function space. Thus, the considered optimal control problem is not Tikhonov well-posed.

Consider now the following example[16].

Example 15.7 *The optimization control problem is to find the function $u = u(t)$, which minimize the functional*

$$I_k = \int_0^1 (x - y_k)^2 dt$$

on the subset of functions

$$U = \{u \, | \, |u(t)| \leq 1, \ t \in (0,1)\}$$

satisfying the isoperimetric condition

$$\int_0^1 u dt = 0,$$

where $y_k(t) = (k\pi)^{-1} \sin k\pi t$, k is a numerical parameter (natural number), and the function x is described by the Cauchy problem

$$x'(t) = u(t), \ t \in (0,1); \ \ x(0) = 0.$$

Obviously, the optimality criterion is not negative in this case. Moreover, its equality to zero is possible only for $x(t) = y_k(t)$, which corresponds to the control

$$u_k(t) = y_k'(t) = \cos k\pi t.$$

This function belongs to the set U and satisfies the isoperimetric condition, which means it is the unique optimal control.

If $k \to \infty$, then the considered optimal control problem is reduced to the previous example, which means it has a unique solution $u_\infty(t) = 0$. However, the sequence $\{u_k\}$ does not converge to u_∞ in any natural function space. Consequently, the problem under study is Hadamard ill-posed[17].

RESULTS

Here is a list of questions based on the results of the lecture, the main conclusions on this topic, as well as problems arising from this, partially solved in Appendix.

Questions

It is required to answer questions related to the previously given lecture material.

1. What does the system of optimality conditions look like for Example 15.1?
2. What is the infimum of the functional on the set of admissible controls in Example 15.1?
3. What happens to the sequence $\{u_k\}$ from Example 15.1 when $k \to \infty$?
4. On what basis is it concluded that there is no optimal control for Example 15.1?
5. What made the optimal control problem in Example 15.1 insolvable?
6. How to prove that the set of admissible controls is closed for Example 15.1?
7. Is the set of admissible controls from Example 15.1 convex?
8. What properties differ between the problems from Example 15.1 and 15.2?
9. What does the system of optimality conditions look like for Example 15.2?
10. Is it possible to establish the sufficiency of the optimality conditions for Example 15.2 in accordance with Theorem 5.2?
11. How is Example 15.3 different from previous examples?
12. What is the difference between Example 15.3 and the fairly close Example 14.1?
13. What happens to the sequences $\{x_k^+\}$ and $\{x_k^-\}$ from Example 15.3 at $k \to \infty$?
14. What happens to the corresponding control sequences at $k \to \infty$?
15. What properties do the solutions to the optimality conditions for Example 15.3 have?
16. How can an approximate solution be found for Example 15.3?
17. Why is the optimal control for Example 15.4 necessarily singular?
18. Is it possible to establish the sufficiency of the optimality conditions for Example 15.4 in accordance with Theorem 5.2?
19. Can we use Theorem 5.1 to prove the uniqueness of the optimal control for Example 15.4?
20. What is the fundamental difference between the properties of Example 15.4 and 15.5, and what causes it?
21. Based on this analysis, can we conclude that there are only two singular controls for Example 15.6?

22. Based on the analysis, can we conclude that for Example 15.6 the solution to the maximum principle is necessarily a singular control?

23. What happens to the sequence $\{u_k\}$ from Example 15.7 when $k \to \infty$?

24. What happens to the sequence $\{x_k\}$ from Example 15.7 when $k \to \infty$?

25. In what sense can the u_k control from Example 15.7, for sufficiently large k, be chosen as an approximate solution to the problem?

26. Is the set of feasible controls from Example 15.7 convex?

27. Is the minimized functional from Example 15.7 convex?

Conclusions

Based on the analysis, we come to the following conclusions.

- For optimal control problems with isoperimetric conditions, both with a free and a fixed final state, there may be no optimal control.
- For optimal control problems with isoperimetric conditions, degeneration of the maximum principle is possible, and the number of singular controls can be arbitrary.
- The optimal control problem for Example 15.1 has no solution.
- The insolvability of the problem from Example 15.1 is due to the absence of convexity of both the set of admissible controls and the optimality criterion.
- The optimal control problem for Example 15.2 has no solution.
- The insolvability of the problem in Example 15.2 is due to the absence of convexity of the set of admissible controls, while the optimality criterion is convex.
- There is an infinite set of solutions to the optimality conditions, for Example 15.3.
- The optimality conditions for Example 15.3 are not sufficient.
- The optimal control problem for Example 15.3 has no solution.
- The solutions to the optimality conditions for Example 15.3 form a minimizing sequence.
- For optimal control problems with isoperimetric conditions, degeneration of the maximum principle is possible, and the number of singular controls can be arbitrary.

- Singular controls in problems with isoperimetric conditions can be either optimal or non-optimal, and the Kelley condition can be used to identify non-optimal singular controls.
- The optimal control problem for Example 15.4 has a unique solution.
- The unique solution to the maximum principle for Example 15.4 is singular control.
- The singular control for Example 15.4 satisfies the Kelley condition.
- The singular control for Example 15.4 is optimal.
- The maximum principle for Example 15.4 is a sufficient condition for optimality.
- There is only one singular control for Example 15.5.
- The singular control for Example 15.5 does not satisfy the Kelley condition.
- The singular controls for Example 15.5 are not optimal.
- The optimality condition for Example 15.5 is not sufficient.
- For the optimal control problem for Example 15.6, there are at least two singular controls.
- One of the singular controls for Example 15.6 satisfies the Kelley condition, but the other does not.
- The optimality condition for Example 15.6 is not sufficient.
- Optimal control problems with isoperimetric conditions may be Tikhonov and Hadamard ill-posed.
- The optimal control problem for Example 15.4 is Tikhonov ill-posed.
- The optimal control problem for Example 15.7 has a unique solution.
- The optimal control problem for Example 15.7 is Hadamard ill-posed.

Problems

Based on the results obtained above, we come to the following problems.

1. **Infinity of the set of singular controls**. In the examples discussed in the Lecture, the set of singular controls was finite. However, we have previously encountered optimal control problems in which this set is infinite. It would be interesting to give an example of problems with isoperimetric conditions with an infinite number of singular controls. Such examples are available in Appendix.

2. **Singular control for problems with fixed final state**. Previously, an example of optimal control problems with degeneracy of the maximum principle was given in the case when there is either only an isoperimetric condition, or only a fixed final state, or both of these conditions are absent. It would be interesting to consider problems with singular controls when both of these conditions are present simultaneously. Such an example is given in Appendix.

3. **Ill-posed problem with isoperimetric conditions and fixed final state**. Previously, various ill-posed optimal control problems were considered. In the case when there is either only an isoperimetric condition, or only a fixed final state, or both of these conditions are absent. It would be interesting to consider problems with singular controls when both of these conditions are present simultaneously. Such an example is given in Appendix.

4. **Bifurcation of extremals in optimal control problems with isoperimetric conditions**. Chapter 12 gave examples of optimal control problems with extremal bifurcation. I would like to establish a similar effect for problems with isoperimetric conditions. Examples of such tasks are given in Appendix.

5. **Bellman principle in optimal control problems with isoperimetric conditions**. In the previous parts of the book, examples of optimal control problems were given in which the Bellman principle was fulfilled. It would be interesting to check the applicability of Bellman principle for problems with isoperimetric conditions.

15.2 APPENDIX

Here are examples of optimal control problems with isoperimetric conditions with some additional features. In particular, Section 15.2.1 considers problems of the indicated class with an infinite set of singular controls. Sections 15.2.2 and 15.2.3 describe, respectively, problems with singular controls and ill-posed problems with a fixed final state. Section 15.2.4 provides an example of the considered class of problems with bifurcation of extremals, and in Section 15.2.5, we give an example with a violation of the Bellman optimality principle.

15.2.1 Problems with an infinite set of singular controls

The Lecture considered optimal control problems in the presence of an isoperimetric condition with degeneracy of the maximum principle. Moreover, the set of singular control was finite. However, in Chapter 6 examples of problems with an infinite set of singular controls were given. Let us make sure that such a situation is also possible for problems with isoperimetric constraints.

Example 15.8 *The optimal control problem is to find a function $u = u(t)$ that minimizes the functional*

$$I(u) = \int_0^1 x(t)u(t)^2 dt$$

on the subset of functions

$$U = \{u \mid |u(t)| \le 1,\ t \in (0,1)\}$$

satisfying the isoperimetric condition

$$\int_0^{1/2} u(t)dt = 0, \tag{15.14}$$

where the function x satisfies the equalities

$$x'(t) = u(t),\ t \in (0,1);\ \ x(0) = 0. \tag{15.15}$$

Define the function

$$H = pu + \lambda\varphi u - xu,$$

where $\varphi(t) = 1$ for $t < 1/2$ and $\varphi(t) = 0$ for $t > 1/2$. The function p is here the solution of the adjoint system

$$p'(t) = u(t),\ t \in (0,1);\ \ p(1) = 0. \tag{15.16}$$

The function H is linear with respect to the control. Therefore, we can have two different variants. Maybe we have the maximum on the boundary on the given set, but it can be a singular control. For the first (regular) case, the control is determined by the formula.

$$u(t) = \begin{cases} 1, & \text{if } p(t) - x(t) + \lambda\varphi(t) > 0, \\ -1, & \text{if } p(t) - x(t) + \lambda\varphi(t) < 0. \end{cases} \tag{15.17}$$

Solving problems (15.15) and (15.16), we find

$$p(t) - x(t) + \lambda\varphi(t) = -\Big[\int_t^1 u(\tau)d\tau + \int_0^t u(\tau)d\tau\Big] + \lambda\varphi(t) = -\int_0^1 u(\tau)d\tau + \lambda\varphi(t).$$

The first term on the right side of the resulting equality does not depend on t, and the second one is equal to λ for $t < 1/2$ and zero for $t > 1/2$. Then, in accordance with formula (15.17), the control is either a constant, i.e., is identically equal to 1 or −1, or has a discontinuity at the point $t = 1/2$, passing there from the value 1 to −1 or vice versa. Obviously, all four of these functions do not satisfy the isoperimetric condition (15.14), and therefore cannot be solutions to the problem. Thus, if optimal control for the considered problem exists, then it is certainly singular.

Singular control can be implemented only in the case when the control coefficient in the definition of the function H is equal to zero. Then, comparing the obtained result with condition (15.14), from the last equality, we deduce

$$-\int_{1/2}^1 u(\tau)d\tau = 0,\ t < 1/2;\ \ -\int_{1/2}^1 u(\tau)d\tau + \lambda = 0,\ t > 1/2.$$

As a result, we conclude that $\lambda = 0$, and all admissible controls (i.e., belonging to the set U and satisfying the isoperimetric condition) for which the equality

$$\int_{1/2}^{1} u(\tau)d\tau = 0. \tag{15.18}$$

is true are singular. Naturally, there is an infinite and not even countable set of such functions[18].

Check the Kelley condition. Find the derivative

$$\frac{\partial H}{\partial u} = p(t) - x(t) + \lambda.$$

After differentiation, we get

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p'(t) - x'(t) = u - u = 0$$

because of equalities (15.15) and (15.16). It follows that for any singular control the Kelley condition is realized (in the form of equality), which means that all of them can be optimal, although, generally speaking, they will not necessarily be so.

To answer the question, which of the singular controls is optimal, we find the corresponding values of the optimality criterion. Taking into account the state equation and the isoperimetric condition, we obtain

$$I = \int_0^1 xudt = \int_0^1 xx'dt = \frac{1}{2}\int_0^1 \frac{d}{dt}x^2 dt = \frac{x(1)^2}{2} = \frac{1}{2}\Big[\int_0^1 u(\tau)d\tau\Big]^2 = \frac{1}{2}\Big[\int_{1/2}^1 u(\tau)d\tau\Big]^2.$$

It follows that for an arbitrary control the optimality criterion is non-negative, and equality to zero is possible only on those controls that satisfy equality (15.18), and therefore are singular. Thus, all singular controls turn out to be optimal, which means that the problem has an infinite and not even countable set of solutions[19]. It is characteristic that the optimality condition turns out to be necessary and sufficient.

In the considered example, the singular controls certainly turned out to be optimal. A natural example of the non-optimality of singular controls is provided by the problem of maximizing this functional[20].

Example 15.9 *It is necessary to maximize the functional from Example* 15.8 *on the same set.*

As already noted, the singular controls in minimization and maximization problems are the same. This property was established earlier for problems without an isoperimetric condition. However, the degeneracy of the maximum principle is associated exclusively with the properties of the function H. In the presence of the isoperimetric condition, this function is supplemented with one more term, which does not change the situation. Therefore, the singular controls for Example 15.9 will be all functions from the set U defined in the previous example that satisfy the isoperimetric condition (15.14). As was established in the previous chapter, they all minimize the considered functional, which means they cannot maximize it. Thus, the singular controls for Example 15.9 are not optimal.

It is also interesting to check the validity of Kelley condition for this example, in which the corresponding value must (unlike the previous case) be non-positive, since here we are looking for the maximum, not the minimum of the functional. Since the function H for Examples 15.8 and 15.9 is the same, we can again obtain the relation (15.18). It follows that for any singular control the Kelley condition are again satisfied in the form of equality. However, this circumstance does not contradict the previously established result about the non-optimality of singular controls[21].

In the considered examples, the isoperimetric condition (15.14) is an integral of the control over the first half of the interval (0,1) on which the system is considered. The problem in which in equality (15.14) the integral is taken over any part of this interval has similar properties. A special role is played by the case when the integral over the entire given set is considered.

Example 15.10 *We have the minimization problem from Example* 15.8 *with changing the isoperimetric condition* (15.14) *by the equality*

$$\int_0^1 u(\tau)d\tau = 0. \tag{15.19}$$

Now we have $H = pu + \lambda u – xu$. The maximum principle is degenerate if the coefficient before the control is equal to zero. Thus, we obtain

$$p(t) - x(t) + \lambda = -\int_0^1 u(\tau)d\tau + \lambda = 0.$$

Using equality (15.19), we conclude $\lambda = 0$. Then any admissible control is singular. By analogy with Example 15.8, we can determine that all these functions satisfy the Kelley condition.

We know that the optimality criterion can be determined by the equality

$$I = \frac{1}{2}\Big[\int_0^1 u(\tau)d\tau\Big]^2.$$

Obviously, this integral vanishes for any control that satisfies condition (15.19). Thus, all admissible controls turn out to be not only singular, but also optimal. Thus, this optimal control problem is trivial, but has an infinite and even uncountable set of solutions, and the maximum principle provides a necessary and sufficient condition for optimality[22].

15.2.2 Singular controls for problems with a fixed final state

In all previously considered optimal control problems with isoperimetric conditions, which allowed the degeneration of the maximum principle, there were no restrictions on the final state of the system. However, it is also possible to consider systems with a fixed final state in the presence of isoperimetric constraints.

Example 15.11 *The optimization control problem is to find the function $u = u(t)$, which minimizes the functional*

$$I(u) = \frac{1}{2}\int_0^1 x(t)^2 dt$$

on the subset of functions

$$U = \{u|\, |u(t)| \le 1,\ 0 < t < 1\},$$

which transform the system described by the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0$$

to the final state

$$x(1) = 1/2 \tag{15.20}$$

and guaranty the isoperimetric condition

$$\int_0^1 x(t)tdt = 1/6. \tag{15.21}$$

Define the function

$$H = pu + \lambda(xt–1/6) - x^2/2,$$

where p satisfies the adjoint equation

$$p'(t) = x(t) - \lambda t,\ t \in (0,1).$$

The degeneration of the maximum principle is possible $p(t) = 0$. We have the equality $x(t) = \lambda t$. Putting this result to the isoperimetric condition, we get

$$\int_0^1 x(t)tdt = \lambda \int_0^1 t^2 dt = \frac{\lambda}{3} = \frac{1}{6}.$$

Now we find $\lambda = 1/2$; so, $x(t) = t/2$. From the state equation, it follows $u(t) = 1/2$.

Check the Kelley condition. We find

$$\frac{\partial H}{\partial u} = p(t).$$

Determine the derivative

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p'(t) = x(t) - \frac{t}{2}.$$

After differentiation we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = x'(t) - \frac{1}{2} = u(t) - \frac{1}{2}.$$

Finally, we find

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 1.$$

Therefore, the singular control satisfies the Kelley condition; so, it can be optimal[23].

15.2.3 Ill-posed problem with a fixed final state

Previously, an example of ill-posed optimal control problems was given, when there was no isoperimetric condition, or fixation of the final state, or when none of these restrictions were specified. Let us consider problems in which both of these restrictions are present.

Example 15.12 *The optimization control problem is to find the function $u = u(t)$, which minimizes the functional*

$$I(u) = \frac{1}{2}\int_0^1 x(t)^2 dt$$

on the subset of functions

$$U = \{u|\,|u(t)| \le 1,\ 0 < t < 1\}$$

satisfying the final condition

$$x(1) = 0 \tag{15.22}$$

and isoperimetric condition[24]

$$\int_0^{1/2} u(t)dt = 0, \tag{15.23}$$

where the function x is a solution of the Cauchy problem

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0.$$

Obviously, the minimized functional is non-negative, and equality to zero here is possible only for $x = 0$. For a given equation, this is possible only for a control that is identically equal to zero. This is an element of the set U and guarantees the fulfillment of both additional restrictions. Thus, this optimal control problem has a unique solution $u = 0$. The corresponding minimum of the functional is equal to zero.

Consider now the control sequence $u_k(t) = \cos 2k\pi t$, where $k = 1, 2, \dots$. Calculate the integral

$$\int_0^{1/2} u_k(t)dt = \int_0^{1/2} \cos 2k\pi t dt = \frac{1}{2k\pi} \sin 2k\pi t \Big|_0^{1/2} = 0, \quad k = 1, 2, \dots.$$

Thus, the isoperimetric condition (15.23) is true. Calculate the state function

$$x_k(t) = \int_0^t \cos 2k\pi t dt = \frac{\sin 2k\pi t}{2k\pi}.$$

Then $x_k(1) = 0$, which means that condition (15.22) is also satisfied. Thus, the considered control sequence is admissible. In addition, for all t the convergence $x_k(t) \to 0$ takes place, which implies that $I(u_k) \to 0$. Thus, the sequence $\{u_k\}$ is minimizing. However, it does not converge to optimal control, which means that this problem is Tikhonov ill-posed.

Now we consider another class of well-posedness.

Example 15.13 *The optimization control problem is to find the function $u = u(t)$, which minimizes the functional*[25]

$$I_k(u) = \frac{1}{2} \int_0^1 (x - y_k)^2 dt,$$

where $y_k(t) = (2k\pi)^{-1} \sin 2k\pi t$, k is a numerical parameter (natural number). The functional is minimized on the such subset of functions

$$U = \{u \,|\, |u(t)| \leq 1, \ 0 < t < 1\}$$

under the final condition

$$x(1) = 0$$

and isoperimetric condition

$$\int_0^{1/2} u(t)dt = 0,$$

where the function x is described by the equalities

$$x'(t) = u(t), \ t \in (0, 1); \ x(0) = 0.$$

Obviously, the optimality criterion is non-negative. Besides, it can be equal to zero only for $x(t) = y_k(t)$, i.e., we have

$$u_k(t) = y_k(t) = \cos 2k\pi t.$$

This function belongs to the set U and ensures that both additional conditions are satisfied. Therefore, it is the unique optimal control.

For $k \to \infty$, the considered optimal control problem is reduced to the previous example, which means it has a unique solution $u_\infty(t) = 0$. However, the sequence $\{u_k\}$ does not converge to u_∞. Therefore, this problem is Hadamard ill-posed.

15.2.4 Extremal bifurcation for problems with isoperimetric condition

We considered before optimization control problems with the extremal bifurcation; see Chapter 12. This property is possible for the problems with isoperimetric conditions too.

Example 15.14 *The optimization control problem is to find the function $u = u(t)$, which minimizes the functional*

$$I_(u) = \int_0^\pi \Big[\frac{u^2}{2} - \mu(1 - \cos x)\Big] dt$$

under the isoperimetric condition

$$\int_0^\pi u(t)dt = 0, \tag{15.24}$$

where μ is a positive constant (problem parameter), and x is a solution to the problem

$$x'(t) = u(t), \; t \in (0, \pi); \; x(0) = 0.$$

This problem is quite close to the one discussed in Example 12.5. The unique difference is that instead of the isoperimetric condition, the set of admissible controls was characterized by the presence of a fixed final state $x(\pi) = 0$. However, from the state equation it follows that

$$x(\pi) = \int_0^\pi u(t)dt. \tag{15.25}$$

It follows that the equality $x(\pi) = 0$ is equivalent to condition (15.24), which means that Example 15.14 actually coincides with Example 12.5. In Chapter 12, it was shown that for the latter there is a bifurcation of extremals, and the value $\mu = 1$ is the corresponding bifurcation point. Thus, a similar effect is observed for Example 15.14 with the isoperimetric condition (15.24).

In Chapter 12, another example of a problem with bifurcation of extremals was considered, which could also be reduced to an optimal control problem with an isoperimetric constraint.

Example 15.15 *The optimization control problem is to find the function $u = u(t)$, which minimizes the functional*

$$I(u) = \frac{1}{4} \int_0^\pi (2u^2 + \nu x^4 - 2\mu x^2) dt$$

on the set of function satisfies the equality

$$\int_0^\pi u(t)dt = 0,$$

where μ and ν are positive constants (problem parameters), and x is a solution to the problem

$$x'(t) = u(t),\ t \in (0,\pi);\ x(0) = 0.$$

The above problem differs from the one considered in Example 12.6 only in that instead of the isoperimetric condition, the set of admissible controls there was characterized by the equality $x(\pi) = 0$. However, due to the coincidence of the equations of state, in this case formula (15.25) is again valid, which means that Examples 12.6 and 15.15 are equivalent. Then, in accordance with the results from Chapter 12 for Example 15.15, there is an infinite set of bifurcation points $\mu_k = k^2$, $k = 1, 2, ...,$ and for $(k–1)^2 < \mu \leq k^2$ the corresponding system of optimality conditions has exactly $2k–1$ solutions, and for $\mu > k^2$ at least $2k + 1$ solutions exist.

15.2.5 Applicability of the Bellman principle for problems with isoperimetric conditions

In the previous parts of the book, the Bellman optimality principle was considered, according to which optimal control does not depend on the history of the system and is determined by its state at a given moment. Thus, if $u = u(t)$ is the optimal control of the system on the time interval $(0, T)$, then for any point ξ of this interval the same function[26] is the optimal control of the system on the interval (ξ, T), i.e., any finite part of the optimal control of the system is itself optimal. Let us check the feasibility of Bellman principle for a fairly simple optimal control problem with an isoperimetric condition

Example 15.16 *Consider a system described by the Cauchy problem*

$$x'(t) = u(t),\ t \in (\xi, 1);\ x(0) = 0,$$

where $\xi < 1$. The optimization control problem is to find the function $u = u(t)$, which minimize the functional $I(u) = [x(1)]^2$ on the set of functions satisfying the equality

$$\int_{\xi}^{1} u(t)^2 dt = 1.$$

The isoperimetric condition can be transformed to the equality

$$\int_{\xi}^{1} \Big[u(t)^2 - \frac{1}{1-\xi}\Big] dt = 0.$$

Determine the function

$$H = up + \lambda\Big(u^2 - \frac{1}{1-\xi}\Big).$$

The function p here is a solution to the adjoint system

$$p'(t) = 0,\ t \in (0,\pi);\ p(1) = -2x(1).$$

Equaling to zero the derivative of H with respect to the control, we find

$$u(t) = -p(t)/2\lambda.$$

The solution of the adjoint system is

$$p(t) = -2x(1);$$

so, we get

$$u(t) = x(1)/\lambda.$$

Thus, the control is a constant. Then the solution of the state system at the arbitrary point t is

$$x(t) = 1 + u(t-\xi) = 1 + (t-\xi)x(1)/\lambda.$$

Determine $t = 1$, we obtain

$$x(1) = 1 + (1-\xi)x(1)/\lambda.$$

As a result, we find

$$x(1) = \frac{\lambda}{\lambda - 1 + \xi},$$

so, we have

$$u = \frac{1}{\lambda - 1 + \xi}.$$

Put this value to the isoperimetric condition for finding the constant λ. We get

$$\frac{1-\xi}{(\lambda - 1 + \xi)^2} = 1.$$

Now, we have

$$\frac{1}{(\lambda - 1 + \xi)} = \frac{1}{\sqrt{1-\xi}},$$

so the optimal control is

$$u(t) = \frac{1}{\sqrt{1-\xi}}, \ t \in (\xi, 1).$$

Thus, each value ξ (origin) corresponds to its own optimal control, which is a constant. Thus, any final part of the optimal trajectory is not optimal. Consequently, the Bellman principle for the considered optimal control problem with the isoperimetric condition is not implemented.

Consider now an example of optimization control problem with an isoperimetric condition with respect to the state function.

Example 15.17 *We have the problem of minimization of the functional*

$$I(u) = \frac{1}{2}\int_{\xi}^{1} u(t)^2 dt$$

on the set of functions $u = u(t)$ such that the following equality holds

$$\int_{\xi}^{1} x(t)dt = 1,$$

where $\xi < 1$ and x is a solution to the problem

$$x'(t) = u(t),\ t \in (\xi, 1);\ x(0) = 0.$$

The isoperimetric condition can be transformed to the equality

$$\int_{\xi}^{1} \Big[x(t) - \frac{1}{1-\xi}\Big]dt = 0.$$

Then the function H takes the form

$$H = up + \lambda\Big(x - \frac{1}{1-\xi}\Big) = \frac{u^2}{2}.$$

The function p is a solution to the problem

$$p'(t) = -\lambda,\ t \in (\xi, 1);\ p(1) = 0.$$

Obviously, the control $u = p$ maximizes the function H. Then the solution to the adjoint system is $p(t) = \lambda(1–t)$, so the control is determined by the formula $u(t) = \lambda(1–t)$. Now we find the state function

$$x(t) = \lambda\Big(t - \xi - \frac{t^2 - \xi^2}{2}\Big).$$

Put the result to the isoperimetric condition. We have the equality

$$\lambda \int_{\xi}^{1} \Big(t - \xi - \frac{t^2 - \xi^2}{2}\Big)dt = 1.$$

Calculate the integral at the final formula, which is noted by J. We get

$$J = \int_{0}^{1-\xi} \Big[\tau - \frac{\tau(\tau + 2\xi)}{2}\Big]d\tau = \frac{(1-\xi)^3}{3}.$$

Now we find the parameter

$$\lambda = \frac{1}{J} = \frac{3}{(1-\xi)^3}.$$

Putting this value for the formula for the control, we find the optimal control

$$u(t) = \lambda(1-t) = \frac{3(1-t)}{(1-\xi)^3}.$$

Thus, for any value ξ there exists an optimal control that is the linear function. Therefore, the any final part of the optimal curve is not optimal. We conclude that the Bellman principle is not true for the considered example.

Additional conclusions

Based on the analysis of optimal control problems for systems with isoperimetric conditions carried out in Appendix, the following additional conclusions can be drawn.

- An optimal control problem with an isoperimetric condition can have an infinite set of singular controls, which can be both optimal and non-optimal.
- For optimal control problems with an isoperimetric condition and a fixed final state, the existence of singular controls is possible.
- Optimal control problems with an isoperimetric condition and a fixed final state may turn out to be ill-posed according to both Tikhonov and Hadamard.
- In optimal control problems with an isoperimetric condition, bifurcation of extremals is possible.
- In optimal control problems with an isoperimetric condition, a violation of the Bellman optimality principle is possible.

Notes

1. This sequence was considered before; see Chapter 7, Figure 7.1.

2. This sequence was considered before in Chapter 7; see Figure 7.2.

3. One may wonder what properties of the considered problem were the reason for its insolvability? According to Theorem 7.1, the problem of minimizing a convex continuous functional bounded from below on a convex closed bounded subset of a Hilbert space has a solution. We choose $L_2(0, 1)$ as the control space, which is Hilbert. Lower boundedness and continuity of the optimality criterion were established earlier. However, as noted in Example 7.1, the considered functional is non-convex. As a result, the conditions of Theorem 7.1 are violated, and the absence of optimal control is quite natural.

4. This applies, in particular, to Examples 7.1 and 11.4. Let us also note the functions minimization problems discussed in Chapter 1 $f(x) = -x^2$ and $f(x) = x^3$.

5. A complete proof of the convexity of this functional is given in Chapter 3.

6. One can again ask the question, what properties of this problem were the reason for its absence of a solution? As already noted, the optimality criterion here is convex, as a result of which the obtained result can be explained exclusively by the properties of the set of admissible controls. In Theorem 7.1, this set is required to be convex, closed, and bounded. The boundedness of the set of admissible controls is beyond doubt. In Chapter 7, its closeness was proven. Then the insolvability of the optimal control problem can only be explained by the lack of convexity of the set of admissible controls. Indeed, if some control u belongs to the set U, then the control $-u$ has also a similar property. However, their half-sum gives a zero function for which the isoperimetric condition does not hold. Thus, the property of convexity of the set of admissible controls used in Theorem 7.1 is violated, which makes it possible that there is no optimal control.

7. Otherwise, we get a trivial solution that contradicts the isoperimetric condition.

8. It is curious that the resulting formula coincides exactly with the one that was established in the analysis of the boundary value problem (14.6) and (14.7).

9. However, if optimal control existed, then it would satisfy the system of optimality conditions. However, all solutions of this system are exhausted by families $\{u_k^+\}$ and $\{u_k^-\}$, the elements of which are not optimal, since the corresponding values of the minimized functional are positive.

10. This means that for a large enough number k, the values of the minimized functional on admissible controls u_k^+ and u_k^- are arbitrarily close to its lower bound on the entire set of admissible controls, i.e., these controls can be chosen as weak approximate solutions to the problem.

11. The considered optimal control problem is transformed from Example 6.2 by adding an isoperimetric condition, which is certainly satisfied by the optimal control for the specified example.

12. Note that both singular controls are on the boundary of this set.

13. In this case, it was important for us to obtain an example in which for an optimal control problem with an isoperimetric condition the singular control is not unique. In this regard, the presence or absence of other singular controls is not significant.

14. In this case, it does not matter to us whether the control $u = 1$ is optimal and whether there are non-singular solutions to the maximum principle. Our goal is to get an example with a non-unique singular control and show that not all singular controls are optimal.

15. Using the methodology described in Chapter 6, we can give examples of optimal control problems with an arbitrary number of singular controls, some of which may be optimal, and some of which may not. For this purpose, you can, in particular, use the examples from the indicated chapter, adding isoperimetric conditions to the corresponding formulations of the problem so that the solutions to these problems satisfy them.

16. This optimal control problem is derived from the one considered in Example 8.1 by adding an isoperimetric condition that is likely to be satisfied by the corresponding optimal control.

17. Obviously, the optimal control for Example 15.7 is singular for all k.

18. For example, any function, which is equal to zero on the interval (0,1/2), to number a on the interval (1/2,3/4), and to $-a$ on the interval (3/4,1) for all $a \in [-1, 1]$ is the singular control.

19. In fact, the considered problem differs from Example 6.1 solely in the presence of an isoperimetric condition. Naturally, any solution to the problem from Example 15.8 is a solution to the problem from Example 6.1, but not vice versa.

20. In fact, we are dealing with Example 6.6, supplemented by the isoperimetric condition.

21. We recall that the validity of Kelley condition for singular control means only about its possible optimality, but nothing more. Only the violation of this condition allows us to conclude that this singular control is not optimal.

22. Naturally, the problem of maximizing this functional also has the same property, since it takes the same zero value on any admissible control.

23. In this case, the very fact of the existence of a singular control in a problem with an isoperimetric condition and a fixed final state is important for us. Therefore, it is not important whether non-singular solutions of the maximum principle exist and whether they can be optimal. However, it is obvious that a regular solution of the maximum principle, if it really exists, satisfies the equality

$$u(t) = \begin{cases} 1, & \text{if } p(t) > 0, \\ -1, & \text{if } p(t) < 0. \end{cases}$$

In this case, the function u cannot take the values 1 or –1 everywhere, since in this case both conditions (15.20) and (15.21) are certainly violated. Let us assume that there is a unique control break point. Finding the corresponding solution to the equation of state and substituting it into condition (15.20), we obtain a linear equation with respect to this point. However, as a result of substituting the result into the left side of equality (15.21), the value 1/6 is not obtained. Thus, if there are regular solutions of the maximum principle, then they have at least two discontinuity points.

24. If in the isoperimetric condition, we define the integration from 0 to 1, then this constraint is equivalent to the final condition $x(1) = 0$. However, this requires an example in which both constraints are present in a non-trivial way.

25. As in the previous example, the isoperimetric condition, which consists of the equality to zero of the integral of the function u from 0 to 1, is equivalent to the final condition $x(1) = 0$.

26. More precisely, its narrowing to the corresponding finite subinterval of a given time interval.

V

OPTIMAL CONTROL PROBLEMS WITH A FREE INITIAL STATE

Previously, we considered optimal control problems for which either only the initial state of the system, or both the initial and final states are known. The final part of the book analyzes systems in which the initial and final states are not specified at all. There are two chapters here, the first of which describes the general theory of studying such problems, and the second one considers some examples.

CHAPTER 16

Optimal control systems with a free initial state

The purpose of this chapter is to consider optimal control problems for which the initial state of the system is not known. In this case, each admissible control corresponds not to a concrete state of the system, but to a whole class of states corresponding to the general solution of the equation of state, determined up to arbitrary constants. This circumstance does not turn out to be an obstacle to considering the corresponding optimization problems if we assume that the optimality criterion here is minimized on the set of control–state pairs related by this equation. The Lecture deduces the necessary optimality condition in the form of the maximum principle for the considered optimal control problem, provides examples of its analytical solutions, and describes an algorithm for approximate solving of the problem in the general case. Appendix provides a qualitative analysis of the considered examples based on the general theoretical results given earlier, and also describes other methods for solving this problem.

16.1 LECTURE

One can imagine a situation, where the state equation of the system is given in the absence of any boundary conditions. Then we cannot associate each control with a concrete system state function. It corresponds to an infinite set of solutions to a given differential equation, forming its general solution. We give below a rigorous formulation of the optimal control problem for such a case and present the corresponding necessary optimality conditions. They differ from those discussed in Part II in the absence of boundary conditions for the state equation and the presence of two boundary conditions for the adjoint equation. In addition, fairly simple examples of problems of the indicated class are given, for which, using optimality conditions, it is possible to explicitly find a solution to the problem, and an algorithm for solving the obtained optimality conditions, based on the shooting method, is also described.

DOI: 10.1201/9781003398585-16

16.1.1 Optimal control problem for a system with a free initial state

We again consider a controlled system described by the differential equation

$$x'(t) = f\big(t, u(t), x(t)\big),\ t \in (0,T), \tag{16.1}$$

where u is a control, x is a state function, and f is a known function. In Part II, this equation was supplemented with an initial condition, i.e., the Cauchy problem was considered. In Part III, in addition, a condition was specified at the final moment of time, i.e., the object of study was a system with a fixed final state. In this case, equation (16.1) is considered without any additional conditions, i.e., we have systems with a free initial state. The control, as before, is chosen from a set

$$U = \big\{u \big|\ a(t) \leq u(t) \leq b(t),\ t \in (0,T)\big\},$$

where a and b are known functions.

Now we determine the optimality criterion. Let us note that, due to the absence of an initial condition, the optimality criterion can depend on the initial state of the system. Thus, in the general case, the minimized functional includes an integral depending on the control and state of the system over the entire given interval $(0,T)$, as well as terms associated with both the final and initial state of the system. Thus, we obtain the integral

$$I = \int_0^T g\big(t, u(t), x(t)\big)dt + h\big(x(T)\big) + l\big(x(0)\big),$$

where g, h and l are known functions.

There is one more important circumstance, characteristic specifically for this class of problems. In all previous cases, the equation of state was supplemented with an initial condition, as a result of which the control included in this equation uniquely determines the state of the system[1]. Thus, this function, which is included in the definition of the optimality criterion, is not an independent object. As a result, the argument of the minimized functional was exclusively control. In this case, due to the absence of an initial condition for equation (16.1), the control determines only the ***general solution*** of the equation, determined up to an arbitrary constant[2]. Thus, the dependence of the state function on control is not unique[3]. As a result, the functional I is defined by the pair[4] (u,x) related by equality (16.1). Define the following concept.

Definition 16.1 ***Admissible pair*** *for the considered system is a pair* (u,x)*, where the function* u *belongs to the set* U*, and* x *satisfies equality* (16.1) *for a this function* u*.*

Let us denote the set of all admissible pairs of the system by W. Now we can formulate the ***optimal control problem with a free initial state***[5].

Problem 16.1 *Find an admissible pair* (u,x) *that minimizes the functional* I *on the set* W*.*

To find the optimal pair, we will try to use the method described earlier, adapting it to this formulation of the problem. In particular, we will establish optimality conditions for it in the form of the maximum principle.

16.1.2 Maximum principle for a system with a free initial state

Let the pair (u, x) be optimal for Problem 16.1. Then the following inequality holds

$$\Delta I = I(v, y) - I(u, x) \geq 0 \ \ \forall (v, y) \in W. \tag{16.2}$$

Let us define the Lagrange functional, as was done earlier

$$L(u, x, p) = I(u, x) + \int_0^T p(t)\big[x'(t) - f(t, u(t), x(t))\big]dt.$$

Obviously, for any admissible pair (u, x) and arbitrary function p, the value $L(u, x, p)$ coincides with $I(u, x)$. Then from inequality (16.2) it follows

$$\Delta L = L(v, y, p) - L(u, x, p) \geq 0 \ \ \forall (v, y) \in W, \ \forall p. \tag{16.3}$$

Determine the function

$$H(t, u, x, p) = pf(t, u, x) - g(t, u, x),$$

characterizing the explicit dependence of the integrand in the definition of the functional L on the control. Then the previous inequality takes the form

$$\Delta L = \int_0^T p(t)\Delta x'(t)dt - \int_0^T \Delta H dt + \Delta h + \Delta l,$$

where

$$\Delta x = y - x, \ \ \Delta H = H(t, v, y, p) - H(t, u, x, p),$$
$$\Delta h = h(y(T)) - h(x(T)), \ \ \Delta l = l(y(0)) - l(x(0)).$$

Let us use standard transformations. We get

$$H(t, v, y, p) - H(t, u, x, p) = H_x(t, v, x, p)\Delta x + \eta_1 = H_x(t, u, x, p)\Delta x + \eta_1 + \eta_2,$$

$$h(y(T)) = h(x(T)) + h_x(x(T))\Delta x(T) + \eta_3,$$
$$l(y(0)) = l(x(0)) + h_x(x(0))\Delta x(0) + \eta_4,$$

where $\eta_2 = \big[H_x(t, v, x, p) - H_x(t, u, x, p)\big]\Delta x$, η_1, η_3, and η_4 are second order quantities relative to increments Δx, $\Delta x(T)$, and $\Delta x(0)$, and H_x, h_x, and l_x are derivatives of the corresponding functions. Integrating by parts, we get

$$\int_0^T p(t)\Delta x'(t)dt = p(T)\Delta x(T) - p(0)\Delta x(0) - \int_0^T p'(t)\Delta x(t)dt.$$

As result, inequality (16.3) takes the form

$$-\int_0^T \Delta_u H dt - \int_0^T (H_x + p')\Delta x dt + [h_x + p(T)]\Delta x(T) + [l_x + p(0)]\Delta x(0) + \eta \geq 0 \quad (16.4)$$

for all controls $v \in U$ and arbitrary function p, where

$$\Delta_u H = H_x(t,v,x,p) - H_x(t,u,x,p),\ H_x = H_x(t,u,x,p),\ h_x = h_x(x(T)),\ l_x = l_x(x(0)),$$

and the remainder term is determined by the equality

$$\eta = \eta_3 + \eta_4 - \int_0^T (\eta_1 + \eta_2) dt.$$

Formula (16.4) differs from a similar inequality (3.7), obtained in the analysis of the optimal control problem with a given initial state, only by the presence of an additional term on its left side, as well as the value η_4 in the formula for the remainder term. In a similar inequality (9.5), obtained in the analysis of a system with a fixed final state, for obvious reasons, there were no terms associated with the increment of the state function at the ends of a given interval, and the remainder term included only the integral summand.

In all cases considered, at this stage of the study, the adjoint equation was determined by choosing the function p in such a way that the integrand of the resulting inequality, multiplied by the increment Δx, equals zero. As a result, we obtain the equality

$$p'(t) = H_x(t,u,x,p),\ t \in (0,T) \quad (16.5)$$

having the same form as in both previous cases.

For Chapter 3, in the process of transforming inequality (3.7), we also had the opportunity to choose the final condition for the resulting equation by setting the coefficient at $\Delta x(T)$ to zero. As a result, each of the existing first-order differential equations has one boundary condition, and for the state equation this condition is the initial one, and for the adjoint equation it is the final one. For Chapter 9, when considering inequality (9.5), we do not have the opportunity to supplement the adjoint equation with any boundary condition. However, the system of optimality conditions obtained later makes sense, since for the existing two interconnected differential equations there are exactly two boundary conditions in view of the fact that the state of the system is considered given not only at the initial, but also at the final moment of time.

In this case, the formula for the increment of the functional (16.4) contains not only a term containing the increment $\Delta x(T)$, but also a summand including $\Delta x(0)$. By setting the factors before both specified increments to zero, we know the value of the function p at both ends of the given interval $(0,T)$. This procedure may seem incorrect, since for one first-order differential equation (16.5) we obtain thereby two boundary conditions at once[6]

$$p(0) = l_x(x(0)),\ p(T) = -h_x(x(T)). \quad (16.6)$$

However, this equation is actually considered together with equation (16.1), for which there are no boundary conditions at all. Thus, as in the two previous cases, we obtain a system of two first-order differential equations with two boundary conditions.

Thus, in inequality (16.4) we choose as the function p the solution of the adjoint system (16.5) and (16.6). As a result, we obtain the following inequality

$$-\int_0^T \Delta_u H dt + \eta \geq 0 \ \forall v \in U.$$

It coincides exactly[7] with formula (3.10). The derivation from here of the optimality condition in the form of the ***maximum principle*** is carried out in the same way as in Chapters 3 and 9. Thus, the following statement is true[8].

Theorem 16.1 *In order for the control u to be a solution to Problem* 16.1*, it must satisfy the equality*

$$H\left[t, u(t), x(t), p(t)\right] = \max_{v \in [a(t), b(t)]} H\left[t, v, x(t), p(t)\right], \quad t \in [0, T], \tag{16.7}$$

where x satisfies equality (16.1)*, and p satisfies equalities* (16.5) *and* (16.6)*.*

Thus, the system of optimality conditions for Problem 16.1 includes differential equations (16.1) and (16.5) with boundary conditions (16.6) and the maximum condition (16.7) for three unknown functions u, x, and p.

16.1.3 Analytical solution of the problem in the absence of control restrictions

In the previous chapters, we began the direct study of other types of optimal control problems with the analysis of examples that can be solved analytically. Let us present one extremely simple special case of Problem 16.1, for which the optimal pair can be easily determined directly from the optimality conditions.

Example 16.1 *Required to minimize functional*

$$I(u, x) = \frac{1}{2} \int_0^1 (x^2 - xt^2 + u^2 - 2ut) dt,$$

where the functions x and u satisfy the state equation

$$x'(t) = u(t), \ t \in (0, 1). \tag{16.8}$$

We have Problem 16.1 for the following values of its constituent quantities:

$$T = 1, \ f = u, \ g = (x^2 - xt^2 + u^2 - 2ut)/2, \ a = -\infty, \ b = \infty, \ l = 0, \ h = 0.$$

Determine the function

$$H = pu - \frac{1}{2}(x^2 - xt^2 + u^2 - 2ut).$$

Then the adjoint equation (16.5) takes the form

$$p'(t) = x(t) - t^2/2, \ t \in (0,1), \tag{16.9}$$

and the boundary conditions (16.6) are written as follows:

$$p(0) = 0, \ p(1) = 0. \tag{16.10}$$

Setting the derivative of the function H with respect to the control to zero, we find

$$u(t) = p(t) + t. \tag{16.11}$$

Taking into account the negativity of the corresponding second derivative and the absence of restrictions on control, we conclude that this function is indeed a solution to the maximum condition (16.7).

Thus, to find three unknown functions u, x and p, we have equalities (16.8)–(16.11). We find the solution to the resulting problem using the method of eliminating unknowns. In particular, differentiating equality (16.9) and taking into account equation (16.8), we obtain

$$p''(t) = x'(t) - t = u(t) - t = p(t).$$

Now we have the following boundary value problem with respect to the function p.

$$p''(t) - p(t) = 0, \ t \in (0,1), \ p(0) = 0, \ p(1) = 0.$$

Obviously, it has the solution[9] $p(t) = 0$. Then from formula (16.11) it follows $u(t) = t$. Now from equality (16.9) it follows that $x(t) = t^2/2$. One can prove that this is in reality the optimal control for the considered problem[10].

16.1.4 Analytical solution of the problem with control constraint

We considered before the optimal control problem with a free initial state without any control constraints and find its solution directly. Let us now give an example of a similar problem with a control constraint[11].

Example 16.2 *Minimize the functional*

$$I(u,x) = \frac{1}{2}\int_0^1 (x^2 + u^2)dt,$$

where the functions x and u satisfy the state equation

$$x'(t) = u(t), \ t \in (0,1), \tag{16.12}$$

besides, the control belongs to the set

$$U = \{u \,|\, |u(t)| \le 1, \ t \in (0,1)\}.$$

Determine the function

$$H = pu - (x^2 + u^2)/2.$$

Then the adjoint system takes the form

$$p'(t) = x(t) - t^2/2,\ t \in (0,1);\ \ p(0) = 0,\ \ p(1) = 0. \tag{16.13}$$

Find the solution to the maximum principle

$$u(t) = \begin{cases} 1, & \text{if } p(t) < -1, \\ p(t), & \text{if } -1 \le p(t) \le -1, \\ -1, & \text{if } p(t) > 1. \end{cases} \tag{16.14}$$

The resulting system of optimality conditions (16.12)–(16.14), in accordance with the method of eliminating unknowns, can be reduced to a boundary value problem for a second order differential equation with respect to the function p, as was done for Example 16.1. However, in view of the non-linear relationship between the functions u and p according to equality (16.14), it is very difficult to find an analytical solution to the problem in a similar way.

Let us consider an auxiliary optimal control problem, which differs from the one given above only in the absence of restrictions on control. The optimality conditions for it include equalities (16.12) and (16.13), as well as the formula $u(t) = p(t)$ instead of condition (16.14). This system is no longer difficult to solve using the method of eliminating unknowns. In particular, differentiating the adjoint equation and using equality (16.12) and the last formula, we find $p'' = x' = u = p$. Then the function p turns out to be a solution to the boundary value problem

$$p''(t) - p(t) = 0,\ t \in (0,1);\ \ p(0) = 0,\ \ p(1) = 0.$$

Obviously, it has a zero solution. As a result, it follows from equality (16.13) that $x = 0$, which means $u = 0$ due to equality (16.12). It is clear that the minimum of the functional in the absence of any restrictions does not exceed the minimum of the functionality in the presence of restrictions. Then, taking into account that the found control belongs to the set U, we conclude that the resulting pair of functions (u, x) also turns out to be a solution to the problem considered in Example 16.2[12].

16.1.5 Algorithm of solving the optimality conditions

It is clear that it is possible to find an analytical solution to the problem only in exceptional cases for extremely simple situations. In the general case, optimal control can be found exclusively approximately using some iterative methods. We return again to problem (16.1), (16.5)–(16.7).

As before, we can find the control from the maximum condition (16.7) if the functions x and p are known. However, unlike all previous problems, knowledge of the control is not enough to determine the function x from the equation of state due to the absence of an initial condition. On the other hand, for the adjoint equation (16.5) there are two boundary conditions at once. We encountered a similar situation in Part III, with the only difference that the equation of state was overdetermined there, and the adjoint equation was underdetermined. To approximately solve the optimality conditions in Chapter 9, we used the shooting method, introducing the missing final condition for the adjoint equation in the form $p(T) = \psi$ with an unknown parameter ψ. In this case, the equation of state was considered with given initial conditions, and the additionally available final condition for the function x was interpreted as an equation for the parameter ψ.

To solve the existing system of optimality conditions (16.1), (16.5)–(16.7), we will also use the ***shooting method***. However, now we add to the equation of state (16.1) the missing initial condition

$$x(0) = \psi, \tag{16.15}$$

where the number ψ is unknown. Now, regarding the four unknowns u, x, p and ψ, we have system (16.1), (16.5)–(16.7), and (16.15). In this case, equation (16.1) is considered together with the initial condition (16.15), equation (16.5) is solved together with the second boundary condition (16.6), and the first condition (16.6) is understood as an algebraic equation for ψ. Its solution can be found iteratively, for example, using the algorithm

$$\psi_{k+1} = \psi_k - \beta_k \big[p_k(0) - l_x(x_k(0))\big], \; k = 0, 1, \dots , \tag{16.16}$$

where x_k and p_k are solution of the state equation and the adjoint system at the k iteration, and β_k is an iterative parameter. As result, we have the following algorithm.

1. The initial approximations of the control u_0 and the initial value ψ_0 of the function x are specified. The sequence of algorithm parameters $\{\beta_k\}$ is chosen.

2. At the current k-th iteration from the Cauchy problem (16.1) and (16.15) with the previously determined values $u = u_k$ and $\psi = \psi_k$ the corresponding value of the state function $x = x_k$ is determined.

3. Solving the adjoint equation (16.5) for known values $u = u_k$ and $x = x_k$ with the final condition $p_k(T) = -h_x(x_k(T))$ we find the function p_k.

4. For known values $x = x_k$ and $p = p_k$, a new control approximation u_{k+1} is determined from the maximum condition (16.7).

5. A new approximation of the parameter ψ_{k+1} is calculated using the formula (16.16).

If the described algorithm converges, the result is a solution to the system of optimality conditions.

RESULTS

Here is a list of questions in the field of optimal control problems for systems with a free initial state, the main conclusions on this topic, as well as problems arising in this case that require additional research.

Questions

It is required to answer questions related to the previously given lecture material.

1. Why does the problem of an optimal system described by a differential equation in the absence of an initial condition make sense?
2. What exactly can be found from the state equation in the absence of an initial condition?
3. What is the sense of the dependence of the system state function on control in Problem 16.1?
4. Why does the optimality criterion for Problem 16.1 alone include an additional term that is absent in the previously considered optimal control problems?
5. Why in the optimal control problems discussed earlier was it necessary to find a control that minimizes this functional, and in Problem 16.1 the question is raised about finding the optimal control–state pair?
6. Why, when deriving the optimality condition for Problem 16.1, does the remainder term in the formula for the increment of the functional include an additional term that is absent in the optimal control problems considered earlier?
7. What is the difference between the adjoint system for Problem 16.1 and similar relations for other types of optimal control problems?
8. Why is it possible to specify two boundary conditions (16.6) for the adjoint equation (16.5), which is a first-order differential equation?
9. How does the final inequality for the increment of the minimized functional, differ from similar relations for other types of optimal control problems?
10. How does the system of optimality conditions obtained in accordance with Theorem 16.1 differ from similar systems for the previously considered optimal control problems?
11. As a result of what properties could the solution to the optimal control problem for Example 16.1 be found analytically?
12. How does it follow that Example 16.1 obtained during the analysis process is really a solution to the problem?
13. Why were we able to find a solution to the problem directly from the system of optimality conditions for Example 16.1, but for Example 16.2 we cannot obtain a similar result?
14. How was the solution to the optimal control problem for Example 16.2 found?
15. Do we have confidence that the functions found in the analysis of Example 16.2 actually form solutions to the corresponding optimal control problem?

16. Why is it not possible to use the method of successive approximations described in Chapter 3 to find the solution of the system of optimality conditions obtained in accordance with Theorem 16.1?

17. Why is it possible to use the shooting method for solving of the system of optimality conditions obtained in accordance with Theorem 16.1?

18. How does the shooting method used in this case differ from the one discussed in Chapter 9?

19. When using the shooting method for the state equation, is it possible to add a fictitious final condition rather than an initial condition?

20. Is it possible, when using the shooting method, to consider the adjoint equation together with the initial condition available for it, and use the corresponding final condition when specifying the missing boundary condition for the state equation at the subsequent iteration?

21. Can we be sure that the shooting method is convergent?

22. Can we be sure that if the shooting method converges, we will find a solution to the existing optimal control problem?

23. Can we be sure that if the shooting method converges, we will find a solution to the resulting system of optimality conditions?

Conclusions

Based on the study of the problem of optimal control of systems with a free initial state, we can come to the following conclusions.

- In the absence of an initial condition, each admissible control determines the general solution of the state equation, so the state of the system according to a given control is not determined definitely.

- The optimal control problem for a system in the absence of initial conditions makes sense.

- The absence of initial conditions for the equation of state provides additional opportunities for solving the optimal control problem, since in addition to the freedom to choose control, it becomes possible to select one of the partial solutions of the state equation that provides the lowest value of the optimality criterion.

- For an optimal control problem with a free initial state, the optimality criterion is minimized on a set of control–state pairs related by a given equation.

- For an optimal control problem with a free initial state, a general scheme for deriving the necessary optimality conditions in the form of the maximum principle is applicable.

- The system of optimality conditions for the considered problem includes a state equation without any boundary conditions, an adjoint equation with two boundary conditions, and a maximum condition, which has the same form as for other previously considered problems.
- In the simplest case, the solution to the optimality conditions for problems with a free initial state can be found analytically.
- The optimality conditions for Example 16.1 allow for a direct analytical solution.
- The functions found in the process of analyzing the optimality conditions for Example 16.1 are indeed a solution to the corresponding optimal control problem.
- The solution to the optimal control problem for Example 16.2 is found by passing to a similar problem in the presence of an initial state.
- The functions determined during the analysis of Example 16.2 are indeed the solution to the corresponding optimal control problem.
- The solution to a system of optimality conditions for problems with a free initial state is, as a rule, determined approximately.
- To practically solve a system of optimality conditions for problems with a free initial state, one can use the shooting method.
- When using the shooting method to solve a system of optimality conditions in problems in this case, the state equation is supplemented with an initial condition with an unknown value on its right side, the adjoint equation is solved with one of the existing boundary conditions, and the second boundary condition is interpreted as an equation with respect to the unknown initial state of the system.

Problems

In the process of analyzing optimal control problems for systems with a free initial state, additional problems arise that require additional research.

1. **Qualitative analysis of the considered examples**. In the Lecture, two problems of optimal control of systems in the absence of initial conditions were solved. It would be interesting to apply the general theorems described in Part II for them, in particular to prove the existence and uniqueness of a solution to the problem, as well as the sufficiency of optimality conditions. These results are presented in Appendix.
2. **Linear-quadratic optimal control problems for systems with a free initial state**. In Chapters 3 and 10, for linear systems with quadratic functionals in the absence of explicit restrictions on control using the decoupling

method, a non-iterative algorithm for solving optimal control problems was described. One would like to establish similar results for systems with a free initial state. The corresponding result is obtained in Appendix.

3. **Penalty method in optimal control problems for systems with a free initial state**. In Problem 16.1, the optimality criterion is minimized on the set of feasible control-state pairs related by the state equation. This interpretation is also typical for the penalty method described in Chapter 4. In this regard, it seems natural to use the penalty method directly to solve the problem. In Appendix, the penalty method will be used to solve specific problems of optimal control of systems with a free initial state.

4. **Optimal control of singular systems with a free initial state**. Obviously, in the absence of an initial condition, any control corresponds to far from a single state function. Chapter 4 considered an optimal control problem for which the solution to the state equation has an infinite number of solutions given the initial condition. It was interesting to consider the optimization problem for such an equation in the absence of an initial condition. One such example is given in Appendix.

5. **Non-uniqueness of the solutions and non-sufficiency of optimality conditions for systems with a free initial state**. In the considered examples of optimal control problems in the absence of initial conditions, the solution to the problem was unique, and the optimality conditions were sufficient. It would be interesting to construct examples of such problems for which these properties are violated. Such examples are given in the next chapter.

6. **Singular controls for systems with a free initial state**. Until now, we have considered only regular solutions of the maximum condition for the considered systems. Chapter 17 gives examples of problems of this type with singular controls that may or may not be optimal.

7. **Non-solvable optimal control problems for systems with a free initial state**. Previously, for other classes of extremal problems, we were faced with the absence of a solution to the problem. Chapter 17 provides examples of optimal control problems without initial conditions that have no solution.

8. **Ill-posed optimal control problems for systems with a free initial state**. For other types of optimization problems, there were cases when the problems under study were ill-posed according to Tikhonov or Hadamard. Similar tasks for this class of systems are given in Chapter 17.

9. **Optimal control problems for systems with a free initial state under the isoperimetric conditions**. In the previous part, optimal control problems with a free or fixed final state in the presence of isoperimetric constraints were considered. However, such problems also make sense for systems in the absence of initial conditions. Examples of such problems are given in Chapter 17.

16.2 APPENDIX

Below is some additional information about optimal control problems for systems with a free initial state. In particular, Section 16.2.1 provides a qualitative analysis (existence and uniqueness of a solution, sufficiency of optimality conditions) of the examples discussed in the Lecture based on the general results given in Part II. In Section 16.2.2, the linear-quadratic problem of optimal control of a system with a free initial state is solved using the decoupling method. Section 16.2.3 uses one specific example to describe the application of the penalty method for the class of problems under consideration. Finally, in Section 16.2.4, the penalty method is used to solve the optimal control problem for a singular differential equation in the absence of an initial condition.

16.2.1 Qualitative analysis of examples

We return to the analysis of the examples discussed in the Lecture. Let us consider in particular at Example 16.1. Let us try to apply the general statements given earlier to study it. The first question of interest to us is the existence of optimal control. Chapter 7 provides two theorems regarding the solvability of the optimization problem. In this case, we do not have the opportunity to use Theorem 7.17.1 due to the absence of explicit restrictions on control, and therefore, the set of admissible controls is unlimited. However, we still have the second statement at our disposal.

According to Theorem 7.2, the problem of minimizing a convex continuous coercive functional bounded from below on a convex closed subset of a Hilbert space has a solution. The first step in using this result is to select the space on which the optimality criterion is defined. In this case, the argument of the functional I is the pair (u, x). Considering the presence of the control square under the integral, it is natural to assume that $u \in L_2(0,1)$. With the same reason we can assume that the function x must belong to the same space. However, equality (16.8) indicates that the derivative of this function, equal to the control, also belongs to the space $L_2(0,1)$. The fact that both the function of one variable itself and its first derivative[13] belong to the class of square-integrable functions corresponds to the ***Sobolev space*** $H^1(0,1)$. In this regard, it is natural to assume that the considered functional is defined on the ***space product***[14] $Z = L_2(0,1) \times H^1(0,1)$. Obviously, it is minimized on the set

$$W = \{(u,x) \in Z \mid x'(t) = u(t),\ t \in (0,1)\}.$$

Prove the convexity of this set.

Let (u, x) and (v, y) be arbitrary elements of the set W, and σ is a number from the unit interval. We obtain the equalities

$$x'(t) = u(t),\ y'(t) = v(t),\ t \in (0,1).$$

Multiplying the first equality by σ, and the second one by $1-\sigma$, we have

$$[\sigma x + (1-\sigma)y]'(t) = [\sigma u + (1-\sigma)v](t),\ t \in (0,1).$$

Thus, the pair $(\sigma u + (1-\sigma)v, \sigma x + (1-\sigma)y)$ belongs to the set W, so this set is convex.

Consider a pair sequence $\{(u_k, x_k)\}$ with elements of the set W such that the following convergence holds[15] $(u_k, x_k) \to (u, x)$ in Z. Prove that the limit pair (u, x) belongs to the set W. We have the equality

$$x_k'(t) = u_k(t), \ t \in (0,1), \tag{16.17}$$

and the convergence $u_k \to u$ in $L_2(0,1)$ and $x_k \to x$ in $H^1(0,1)$. The last condition means that in the space $L_2(0,1)$ not only the sequence of functions $\{x_k\}$ converges, but also the sequence of their derivatives. Thus, we have the convergence of $x_k' \to x'$ in $L_2(0,1)$. In Chapter 7, we noted that if a certain sequence converges in the space $L_2(0,1)$, then from it we can extract a subsequence that converges for almost all t in the unit interval[16]. Then, passing to the limit in equality (16.17), we conclude that $x'(t) = u(t)$, which implies that the set W is closed[17].

Consider now the minimized functional. It takes the form

$$I(u,x) = \frac{1}{2}\int_0^1 \Big[\Big(x - \frac{t^2}{2}\Big)^2 + (u-t)^2\Big]dt - \frac{1}{2}\int_0^1 \Big(\frac{t^4}{4} + t^2\Big)dt.$$

Since the second integral is a constant, the properties of the functional I are completely determined by the properties of the first integral. Considering that under the integral is the sum of two squares, and quadratic functions have been repeatedly studied previously[18], we conclude that we are indeed dealing with a convex continuous coercive functional.

Thus, all the conditions of Theorem 7.2 are satisfied, which means that the optimal control problem under consideration has a solution. Moreover, in view of the strict convexity of quadratic functions, this solution is unique by virtue of Theorem 5.1.

Now we prove the sufficiency of the optimality conditions. As we know (see Chapter 5), this is follow from the non-negativity of the remainder term

$$\eta = \eta_3 + \eta_4 - \int_0^1 (\eta_1 + \eta_2)dt. \tag{16.18}$$

Here, the first two terms on the right side of this equality are obtained by expanding into a Taylor series the terms in the formula for the increment of the functional, characterizing the function x at the final and initial times. In our case, the optimality criterion does not contain such terms, as a result of which the value η is determined exclusively by the integral in equality (16.18). Now we find the value $\eta_2 = [H_x(t,v,x,p) - H_x(t,u,x,p)]\Delta x$. For the given function H, we have $H_x = t^2/2 - x$. Since this value does not depend on the control, we conclude that $\eta_2 = 0$. Finally, the value η_1 is the second derivative of the function H with respect to x, equal to -1, multiplied by the square of the increment of the state function. As a result, we conclude that the remainder term η, determined by formula (16.18), is indeed non-negative, which means that the maximum principle for the example under consideration provides a necessary and sufficient condition for optimality.

Let us now turn to the analysis of Example 16.2. Here, it would seem, due to the presence of explicit restrictions on control (the formulation of the problem includes inclusion $u \in U$), we have Theorem 7.1 at our disposal to prove the existence of a solution to the problem. However, note that the functional is minimized on a set of control-state pairs. Therefore, if the control actually takes a value from a limited area, then the state function is not limited due to the absence of initial conditions. Thus, we again use Theorem 7.2 to prove the existence of an optimal pair.

In this case, we consider again the functional I in the space Z defined above for the same reasons as for Example 16.1. In this case, we define the set of admissible pairs as follows:

$$Y = \big\{(u, x) \in W \mid u \in U\big\},$$

where the set W was defined above. Obviously, from the previously established convexity and closedness of the sets W and U, the validity of similar properties of the set of admissible pairs follows[19]. The optimality criterion in this case is the sum of two quadratic functionals, for which coercivity, convexity and continuity are already known. This implies similar properties of this functional I. As a result, the existence of a solution for Example 16.2 follows from Theorem 7.2. Moreover, in view of the strict convexity of the quadratic functional, the functional I also turns out to be so. Consequently, the solution to this problem is unique.

To justify the sufficiency of optimality conditions, we define the remainder term in accordance with formula (16.18). As for the previous example, the equalities $\eta_3 = 0$ and $\eta_4 = 0$ are valid here due to the absence in the definition of the optimality criterion of terms characterizing the boundary values of the state function. The derivative of the function H in this case is determined by the formula $H_x = -x$. It follows that $\eta_2 = 0$, and the second derivative of the function H is negative. Then the remainder term η turns out to be non-negative, which means that the optimality conditions are necessary and sufficient.

16.2.2 Decoupling method

Let us now consider the linear-quadratic problem of optimal control of a system with a free initial state, which is an analogue of Problems 3.2 and 10.1. Given a system described by the following Cauchy problem

$$x'(t) = a(t)x(t) + b(t)u(t) + f(t), \ t \in (0, T), \tag{16.19}$$

where the functions a, b, and f are known. Determine the functional

$$I(u, x) = \frac{1}{2} \int_0^T \Big\{\alpha\big[x(t) - z(t)\big]^2 + \beta\big[u(t)\big]^2\Big\} dt,$$

where the function z and the constants α and β are known, and the relation between u and x is given by the equality (16.19).

Problem 16.2 *The **linear-quadratic optimal control problem with a free initial state** consists of finding a pair* (u, x) *that minimizes the functional* I *on the set of all pairs related by the equality* (16.19).

Using the known method define the function

$$H = p(ax + bu + f) - [\alpha(x - z)^2 + \beta u^2]/2.$$

Then the adjoint system takes the form

$$p'(t) = \alpha[x(t) - z(t)] - a(t)p(t),\ t \in (0, T);\ p(0) = 0,\ p(T) = 0. \quad (16.20)$$

From the maximum condition, we find the control

$$u(t) = \beta^{-1} b(t)p(t),\ t \in (0, T). \quad (16.21)$$

Thus, we have the linear system of optimality conditions (16.19)–(16.21) that is an analogue of problems (3.23)–(3.25) and (10.1)–(10.4).

To find a solution to the problem, we will use the decoupling method. In all three cases of the considered linear-quadratic problem under, we have the same state equations, adjoint equations and formulas for finding the control. However, in Chapter 3 there was an initial condition for the state function and a final condition for the function p, and in Chapter 10 there were two boundary conditions for the function x and no boundary conditions for the function p. In this case, the situation is the opposite, i.e., the equation of state is considered without boundary conditions, and the adjoint system includes two boundary conditions.

Using the linearity of system (16.19)–(16.21) with respect to all three unknown functions suppose again (as previous cases) the existence of linear relation between the functions x and p. This is the equality

$$p(t) = r(t)x(t) + q(t),\ t \in (0, T), \quad (16.22)$$

where r and q are unknown functions. After differentiation with using the first equality (16.20) we obtain

$$r'x + rx' + q' = \alpha(x - z) - a(rx + q).$$

Substantiate the derivative of the function x from the equality (16.19) using formula (16.21), we get

$$\big(r' + \beta^{-1} b^2 r^2 + 2ar - \alpha\big)x + \big(q' + \beta^{-1} b^2 rq + aq + fr + \alpha z\big) = 0.$$

Having defined $t = 0$ in formula (16.22) and taking into account the initial condition for the function p, we establish

$$r(0)x(0) + q(0) = 0.$$

Equating the coefficient of x and the quantity independent of x to zero in the two equalities obtained, we get the following problems regarding the functions r and q.

$$r'(t) + \beta^{-1} b(t)^2 r(t)^2 + 2a(t)r(t) = \alpha,\ t \in (0, T),\ r(0) = 0. \quad (16.23)$$

$$q'(t) + \beta^{-1} b(t)^2 r(t)q(t) + a(t)q(t) + f(t)r(t) + \alpha z(t) = 0,\ t \in (0, T),\ q(0) = 0. \quad (16.24)$$

Note that the equations themselves have the same form as before, but the boundary conditions turn out to be different. As a result, we arrive at the following algorithm for solving the problem under consideration:

1. Solve the Cauchy problem (16.23) in the forward time direction for an ordinary differential equation with quadratic non-linearity (***Riccati equation***) with respect to the function r.

2. Solve the Cauchy problem (16.24) in the forward time direction for a linear ordinary differential equation with respect to the function q.

3. From formula (16.22) with the already known functions r and q, the explicit dependence of the function x on p is determined

$$x(t) = r(t)^{-1}[q(t)-p(t)],\ t \in (0,T), \tag{16.25}$$

which is substituted into the adjoint equation.

4. The adjoint equation is solved in the opposite direction with a known homogeneous condition at $t = T$ with respect to the function p.

5. Using formulas (16.21) and (16.25) the required pair (u, x) is calculated.

Thus, the completely analysis of Problem 16.2 can be realized directly without using any iterative algorithm[20].

16.2.3 Penalty method

A special feature of problems with a free initial state is the fact that here control and state are considered as a single whole, the optimality criterion is minimized on a set of control-state pairs, and the state equation is understood as an equality connecting the elements of these pairs. A similar idea was used in Chapter 4, where the penalty method was used instead of the Lagrange multiplier method to solve the problem. In this regard, it is quite natural to try to use this method to study optimal control problems with a free initial state.

We consider only Example 16.2. Using penalty method determine the functional

$$J_\varepsilon(u,x) = \frac{1}{2}\int_0^1 (u^2 + x^2)dt + \frac{1}{2\varepsilon}\int_0^1 (x' - u)^2 dt,$$

where ε is a small positive parameter. We consider the problem of its minimization on the set of such pairs (u, x), the first element of which is defined on the set U of functions whose value at any point does not exceed one. Due to the presence of control restrictions, we use the variational inequality to solve the problem. Let $(u_\varepsilon, x_\varepsilon)$ be a solution to the given problem. Then the following inequality holds

$$J_\varepsilon\big(u_\varepsilon + \sigma(s-u_\varepsilon), x_\varepsilon\big) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0 \ \ \forall s \in U,\ \forall \sigma \in (0,1).$$

As result, we get

$$\frac{\sigma}{2}\int_0^1 \big[2u_\varepsilon(s-u_\varepsilon)+\sigma(s-u_\varepsilon)^2\big]dt + \frac{\sigma}{2\varepsilon}\int_0^1 \big[-2(x_\varepsilon'-u_\varepsilon)(s-u_\varepsilon)+\sigma(s-u_\varepsilon)^2\big]dt \geq 0.$$

Dividing this inequality by σ and passing to the limit as $\sigma \to 0$, we have

$$\int_0^1 \Big(u_\varepsilon - \frac{x_\varepsilon' - u_\varepsilon}{\varepsilon}\Big)(s-u_\varepsilon)dt \geq 0 \;\; \forall s \in U.$$

Defining the function

$$p_\varepsilon' \;=\; \varepsilon^{-1}(x_\varepsilon' - u_\varepsilon), \tag{16.26}$$

we obtain

$$\int_0^1 (u_\varepsilon - p_\varepsilon)(s-u_\varepsilon)dt \geq 0 \;\; \forall s \in U.$$

We analyzed this variational inequality in Chapter 4. It has the solution

$$u_\varepsilon(t) \;=\; \begin{cases} -1, & \text{if } \; p_\varepsilon(t) < -1, \\ p_\varepsilon(t), & \text{if } \; -1 \leq p_\varepsilon(t) \leq 1, \\ 1, & \text{if } \; p_\varepsilon(t) > 1. \end{cases} \tag{16.27}$$

For any continuously differentiable function h and positive number σ the following inequality holds

$$J_\varepsilon(u_\varepsilon, x_\varepsilon + \sigma h) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \;\geq\; 0.$$

Therefore, we get

$$\frac{\sigma}{2}\int_0^1 (2x_\varepsilon h + \sigma h^2)dt + \frac{\sigma}{2\varepsilon}\int_0^1 \big[2(x_\varepsilon' - u_\varepsilon)h' + \sigma h'^2\big]dt \;\geq\; 0.$$

Dividing by σ and passing to the limit as $\sigma \to 0$ using equality (16.26) we obtain

$$\int_0^1 (x_\varepsilon h + p_\varepsilon h')hdt \;\geq\; 0.$$

Use the formula of integration by parts

$$\int_0^1 p_\varepsilon h' dt \;=\; -\int_0^1 p_\varepsilon' h dt + p_\varepsilon(1)h(1) - p_\varepsilon(0)h(0) \;\geq\; 0.$$

As result, the previous inequality takes the form[21]

$$\int_0^1 (x_\varepsilon - p_\varepsilon)hdt + p_\varepsilon(1)h(1) - p_\varepsilon(0)h(0) \;\geq\; 0. \tag{16.28}$$

Formula (16.28) is true for all functions h, including those that vanish at the boundary of a given interval. For them this inequality is written as follows

$$\int_0^1 (x_\varepsilon - p_\varepsilon) h dt \geq 0.$$

We considered this inequality in Chapter 4. Using the arbitrariness of the function h we get the equality $p'_\varepsilon = x_\varepsilon$. Then from inequality (16.28) it follows the inequality $p_\varepsilon(1)h(1) - p_\varepsilon(0)h(0) \geq 0$. The value $h(1)$ is arbitrary here, so it the arbitrariness of can be equal to zero. As result, we have $p_\varepsilon(0)h(0) \geq 0$, so $p_\varepsilon(0) = 0$ because of the arbitrariness of $h(0)$. The equality $p_\varepsilon(1) = 0$ can be obtained analogically. Thus, the function p_ε is a solution to the problem

$$p'_\varepsilon(t) = x_\varepsilon(t),\ t \in (0,1);\ p_\varepsilon(0) = 0,\ p_\varepsilon(1) = 0. \tag{16.29}$$

From formula (16.26), it follows the equality

$$x'_\varepsilon(t) = u_\varepsilon(t) + \varepsilon p_\varepsilon(t),\ t \in (0,T). \tag{16.30}$$

Now we have the system (16.27), (16.29), (16.30) with respect to the unknown functions u_ε, p_ε, and x_ε. It differs from the optimality conditions (16.12)–(16.14) obtained in the Lecture only by equation (16.30), which for sufficiently small ε can be interpreted as an approximation of the state equation (16.12). The practical finding of a solution to the resulting system can be carried out using the iterative process described in Section 16.1.5.

16.2.4 Optimal control problem for a singular system with a free initial state

Chapter 4 analyzed examples of optimal control problems for singular systems. In them, the equation of state for a specific admissible control could have no solution at all or have an infinite number of solutions. Similar problems can be posed in the absence of initial conditions. Let us consider a similar example[22].

Example 16.3 *Minimize the functional*

$$I(u,x) = \frac{1}{2}\int_0^1 (x^2 + u^2) dt$$

on the set of pairs (u,x) *satisfying the equality*

$$x'(t) = \sqrt{x(t)} + u(t),\ t \in (0,1), \tag{16.31}$$

where the function u *belongs to the set*

$$U = \{u |\, 0 \leq u(t) \leq 1,\, t \in (0,1)\}.$$

A special feature of equation (16.31) is the fact that even when supplemented with an initial condition, the corresponding Cauchy problem may have significantly more than one solution[23], which, however, is not an obstacle to the analysis of the posed optimal control problem. In accordance with the penalty method described above, we define the functional

$$J_\varepsilon(u,x) = \frac{1}{2}\int_0^1 (x^2+u^2)dt + \frac{1}{2\varepsilon}\int_0^1 (x' - \sqrt{x} - u)^2 dt,$$

where ε is a small positive number. We minimize this functional on the set of all pairs (u,x) such that the first element belongs to the set U.

Let $(u_\varepsilon, x_\varepsilon)$ be the solution to the problem. The following inequality holds

$$J_\varepsilon\big(u_\varepsilon + \sigma(s - u_\varepsilon), x_\varepsilon\big) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0 \;\; \forall s \in U,\; \sigma \in (0,1).$$

It takes the form

$$\sigma \int_0^1 u_\varepsilon(s - u_\varepsilon)dt + \frac{\sigma}{\varepsilon}\int_0^1 (x_\varepsilon' - \sqrt{x_\varepsilon} - u_\varepsilon)(s - u_\varepsilon)dt + o(\sigma) \geq 0,$$

where $o(\sigma)/\sigma \to 0$ as $\sigma \to 0$. Define the function

$$p_\varepsilon = \varepsilon^{-1}\big(x_\varepsilon' - \sqrt{x_\varepsilon} - u_\varepsilon\big). \tag{16.32}$$

Dividing the previous inequality by σ and passing to the limit as $\sigma \to 0$ we get the variational inequality

$$\int_0^1 (u_\varepsilon - p_\varepsilon)(s - u_\varepsilon)dt \geq 0 \;\; \forall s \in U.$$

By analogy with formula (16.27), we find its solution

$$u_\varepsilon(t) = \begin{cases} 0, & \text{if } p_\varepsilon(t) < 0, \\ p_\varepsilon(t), & \text{if } 0 \leq p_\varepsilon(t) \leq 1, \\ 1, & \text{if } p_\varepsilon(t) > 1. \end{cases} \tag{16.33}$$

Then for any smooth enough function h and the arbitrary constant σ the following inequality holds

$$J_\varepsilon(u_\varepsilon, x_\varepsilon + \sigma h) - J_\varepsilon(u_\varepsilon, x_\varepsilon) \geq 0.$$

Dividing by σ and passing to the limit as $\sigma \to 0$ with using equality (16.32) after easy transformation[24] we get

$$\int_0^1 x_\varepsilon h dt + \int_0^1 p_\varepsilon\Big(h' - \frac{h}{2\sqrt{x_\varepsilon}}\Big)dt \geq 0.$$

Transform this inequality using arbitrariness of the function h as was done in the previous subsection, we establish the adjoint system

$$p_\varepsilon'(t) + \frac{1}{2\sqrt{x_\varepsilon}} p_\varepsilon(t) = x_\varepsilon(t),\ t \in (0,1);\ p_\varepsilon(0) = 0,\ p_\varepsilon(1) = 0. \tag{16.34}$$

Now from the formula (16.32) it follows the approximate state equation

$$x_\varepsilon'(t) = \sqrt{x_\varepsilon(t)} + u_\varepsilon(t) + \varepsilon p_\varepsilon(t),\ t \in (0,1). \tag{16.35}$$

Thus, we have the system (16.33)–(16.35) with respect to three unknown functions u_ε, x_ε, and p_ε. Similar systems of optimality conditions were obtained when studying other examples in this section. By solving this system for sufficiently small values of ε, it is possible, in principle, to find an approximate solution to the posed optimal control problem[25].

Additional conclusions

Based on the results obtained above, the following additional conclusions can be drawn.

- For Example 16.1, the existence of a unique pair and the sufficiency of the optimality conditions are proved in a standard way.
- For Example 16.2, the existence of a unique pair and the sufficiency of the optimality conditions are proved in a standard way.
- The optimal control problem for a linear-quadratic system with a free initial state can be solved without using an iterative process.
- To analyze a linear system with a free initial state in the case of quadratic optimality criterion and the absence of restrictions on controls, one can use the decoupling method.
- Optimal control problems for systems with a free initial state can be solved using the penalty method.
- For Example 16.2, an approximate system of optimality conditions is obtained using the penalty method.
- The penalty method is applicable to the analysis of optimal control problems for singular systems with a free initial state.
- Using the penalty method, an approximate system of optimality conditions for one singular system with a free initial state is obtained.

Notes

1. However, in Chapter 4, we considered an example of a controlled system for which the Cauchy problem for the corresponding differential equation had more than one solution; see also Example 16.3.

2. When considering an nth order differential equation or a system of n first-order equations, the general solution includes n arbitrary constants; see, for example [10], [86].

3. For controllable systems characterized by the Cauchy problem, the dependence of the state function on the control is characterized by an operator defined by this problem. In the absence of an initial condition, this dependence is a ***multivalued mapping*** that associates each control with a general solution to the equation, which is determined ambiguously. On multivalued mappings; see [40].

4. In Chapter 4, we already encountered the problem of minimizing a functional on a set of control-state pairs due to the fact that the existence of a solution to the state equation for an arbitrary admissible control was not guaranteed. Thus, the functional was minimized on the set of those pairs for which, under a given control, there is a corresponding state of the system. In this case, the situation is the opposite, since for any fixed control the solution to the equation of state is determined ambiguously, which means that there is an additional possibility, compared to varying the control, of minimizing the functional by choosing the best of the system states for a given control.

5. There is an alternative approach to formulating optimal control of a system with a free initial state (naturally, here we assume the unique solvability of the corresponding Cauchy problem, which is not always the case; see Examples 4.2 and 4.3). Indeed, we can supplement equation (16.1) with a fictitious initial condition $x(0) = x_0$, considering the unknown value x_0 as an additional control. Now the pair of controls (u, x_0) together with equation (16.1) uniquely determines the state. Thus, we obtain a standard optimal control problem of the type of Problem 3.1 (more precisely, a special case of a vector optimal control problem with a free final state, i.e., Problem 3.3). One could also consider a system in which, in the absence of an initial condition, the final state of the system is specified. However, this case, by changing the variables $\tau = T – t$, is reduced to the standard optimal control problem considered in Chapter 3.

6. Let us consider a special case of Problem 16.1, when the equation of state has the form $x' = u$, and there are no restrictions on control. In this case, we can exclude the control from consideration by substituting the derivative x' into the optimality criterion instead. As a result, we obtain the ***Bolza problem***, well known in the calculus of variations, and equalities (16.5) correspond to the ***transversality conditions***; see [37], [61], [208].

7. Naturally, the coincidence here makes sense up to the type of remainder term.

8. Naturally, this result requires strict justification.

9. Indeed, the general solution of the adjoint equation is $p(t) = c_1 e^t + c_2 e^{-t}$. Putting this value to the given boundary condition we get $c_1 = c_2 = 0$, so $p(t) = 0$.

10. Using the equality $p = 0$, from formula (16.9), we get $x(t) = p'(t) + t^2/2 = t^2/2$. Note that the integrand of the optimality criterion is

$$x^2 – xt^2 + u^2 – 2ut \ = \ (x – t^2/2)^2 – t^4/4 + (u – t)^2 – t^2.$$

As result, the minimized functional takes the form

$$I(u) = \frac{1}{2}\int_0^1 \Big[\Big(x - \frac{t^2}{2}\Big)^2 + (u-t)^2\Big]dt - \frac{1}{2}\int_0^1 \Big(\frac{t^4}{4} + t^2\Big)dt.$$

Obviously, the first integral here takes exclusively non-negative values, and the second does not depend on the control. Then the minimum value of the optimality criterion is achieved in the case when the first integral is equal to zero. This is possible only for $u(t) = t$ and $x(t) = t^2/2$, which corresponds to the control found in the process of analyzing the optimality conditions.

11. This problem differs from the one considered in Example 3.3 only in the absence of an initial condition for the state equation.

12. The solution to the problem can be found in another way. Obviously, the minimum of a functional in the presence of restrictions in the absence of initial conditions certainly does not exceed the minimum of the same functional with the same restrictions, but in the presence of a fixed initial condition, for example, zero. In the latter case, we get the already mentioned Example 3.3, the optimal control for which is a function identically equal to zero. The corresponding minimum value of this optimality criterion is zero. Therefore, the minimum of the functional for Example 16.2 does not exceed zero. However, this functional does not accept negative values. Therefore, its minimum is zero, which corresponds to zero control and state values. Then the zero solution is also a solution to this problem. Note that this technique does not work in the functional maximization problem. Indeed, the problem of maximizing a given functional with a zero initial condition corresponds to Example 5.1, which has its control solutions identically equal to 1 and –1, with a maximum of the functional equal to 1/2. It is clear that the maximum of the functional for Example 16.2 should be no less than this value. However, this functional is not limited from above due to the arbitrariness of the initial state of the system. As a result, the well-known solution to the maximization problem from Example 5.1 is not suitable for analyzing the corresponding problem without initial conditions.

13. We consider functions of one variable. For functions of many variables, the Sobolev space H^1 is characterized by functions that are square integrable along with all their first-order partial derivatives. Note also that since we are dealing with square-integrable functions, the corresponding derivatives are understood in a generalized sense, i.e., in the sense of distribution theory.

14. The set product $A \times B$ is a set of all pairs (a, b) such that the following inclusions hold $a \in A$, $b \in B$. If both considered sets are Hilbert spaces, then this is true for their product too; see, for example, [94], [100], [106], [158].

15. ***Convergence of pair sequence*** $(a_k, b_k) \to (a, b)$ is the space product $A \times B$ is the convergence $a_k \to a$ in A and $b_k \to b$ in B; see, for example, [94], [100], [106], [158].

16. In reality, this result is true for any space $L_p(\Omega)$ of functions, which are integrable with degree p in the set Ω with arbitrary dimension.

17. As in other examples of justification for the closedness of sets (see Chapter 7), we have proven the validity of limit relations (in this case the equality $x'(t) = u(t)$ not for all, but for almost all t from the unit interval. However, the elements of the set $L_2(0, 1)$, being measurable functions, are defined up to a set of measure zero, as a result of which the equality fulfilled almost everywhere is sufficient to justify the closedness of the set under consideration.

18. The fact that we are working here not with an ordinary square $f(z) = z^2$, but with functions of the type $f(z) = (z–a)^2$, does not play a fundamental role.

19. Indeed, consider two elements (u_1, x_1) and (u_2, x_2) of the set Y and the number $\sigma \in (0,1)$. Due to the convexity of the sets W and U, we have inclusions $\sigma(u_1, x_1) + (1–\sigma)(u_2, x_2) \in W$ and $\sigma u_1 + (1–\sigma)u_2 \in U$, and therefore $\sigma(u_1, x_1) + (1–\sigma)(u_2, x_2) \in Y$. Now let the sequence of elements $\{(u_k, x_k)\}$ of the set Y converges to some pair (u, x) in the space Z. This means the convergence of $u_k \to u$ in $L_2(0,1)$ and $x_k \to x$ in $H^1(0,1)$. Then, since the sets W and U are closed, the inclusions $(u, x) \in W$ and $u \in U$ are valid. Thus, the set Y is convex and closed. Note that the convexity of the set U was established in Chapter 5, and its closedness in Chapter 7.

20. It would be possible to consider a more general problem when the optimality criterion depends quadratically on the boundary values of the state function.

21. The resulting inequality differs from a similar formula from Chapter 4 solely in the presence of the last term, which is a consequence of the absence of an initial condition for the state equation.

22. This example differs from Example 4.3 only in the absence of an initial condition for the state equation.

23. In particular, for a control identically equal to zero, equation (16.31) with a zero initial condition is solved by a function equal to zero at $0 \le t \le \xi$ and the value $(t–\xi)^2/4$ at $t > \xi$ for any non-negative number ξ; see Chapter 4.

24. The corresponding transformations are carried out in Chapter 4 when studying Example 4.3

25. Naturally, from the statement of the problem itself it is clear that it has a unique solution equal to zero.

CHAPTER 17

Different optimal control problems for systems with a free initial state

In the previous chapter, optimality conditions were given for optimal control problems of systems in the absence of initial conditions. The purpose of this chapter is to study problems of this class for which optimal control does not exist or turns out to be non-unique, optimality conditions degenerate or are insufficient, and the optimization problems themselves are ill-posed. In some cases, additional isoperimetric conditions are imposed on the system.

17.1 LECTURE

Previously, optimality conditions for optimal control problems with a free initial state were considered, with the help of which some examples were studied. For this class of problems, the effects described in other parts of the book are possible. Below, we give examples of optimal control problems for systems with a free initial state, for which the solution does not exist or is not unique, the optimality conditions are not sufficient or degenerate, and the ill-posedness of the problems according to Tikhonov and Hadamard are analyzed.

17.1.1 Non-uniqueness and non-sufficiency

The first unpleasant effects that we encountered when studying optimal control problems were the absence of uniqueness of optimal control and sufficiency of optimality conditions. It is natural to assume that for the class of problems under consideration one may encounter similar situations.

DOI: 10.1201/9781003398585-17

Example 17.1 *Minimize the functional*

$$I(u,x) = \frac{1}{4}\int_0^1 (x^4 - 2x^2 + 2u^2)dt,$$

where the functions x and u are related by the state equation $x'(t) = u(t)$, $t \in (0,1)$.

We have Problem 16.1 with the following values of the parameters

$$T = 1,\ f(t,u,x) = u,\ g(t,u,x) = (x^4 - 2x^2 + 2u^2)/4,$$

$$a = -\infty,\ b = \infty,\ l(x) = 0,\ q(x) = 0.$$

Determine the function

$$H = pu - \frac{1}{4}(x^4 - 2x^2 + 2u^2).$$

Then the adjoint system takes the form

$$p'(t) = x(t)^3 - x(t),\ t \in (0,1);\ p(0) = 0,\ p(0) = 0.$$

Equaling to zero the derivative of the function H, we find the control $u(t) = p(t)$, which is the solution of the maximum principle.

After differentiation of the adjoint equation using the state equation and the found control, we get

$$p'' = (3x^2–1)x' = (3x^2–1)u = (3x^2–1)p.$$

Thus, the function p is a solution of the boundary problem

$$p''(t) = \left[3x(t)^2–1\right]p(t),\ t \in (0,1);\ p(0) = 0,\ \ p(1) = 0. \tag{17.1}$$

Obviously, the function p that is equal to zero is its solution. Using the obtained formula for the control we find $u_0 = 0$. From the adjoint equation it follows that three constant function $x_1 = 0$, $x_2 = 1$, $x_3 = –1$ correspond to this function p. There are particular solutions of the state equation for the control u_0. Thus, we have three pairs (u_0, x_1), (u_0, x_2), and (u_0, x_3), which are the solutions of the optimality conditions and can be optimal[1].

Find the values of the given functional for these pairs. We obtain

$$I(u_0, x_1) = 0,\ I(u_0, x_2) = I(u_0, x_3) = –1/4.$$

As a result, we conclude that the first pair is non-optimal, so the maximum principle is not a sufficient optimality condition.

In order to find out whether the second and third pairs are really solutions to this optimal control problem, we present the optimality criterion in the following form

$$I(u,x) = \frac{1}{4}\int_0^1 (x^2 - 1)^2 dt + \frac{1}{2}\int_0^1 u^2 dt - \frac{1}{4}.$$

Obviously, two integrals here are non-negative. Therefore, the value of the given functional at the any admissible pair is not less than $-1/4$. Its equality to this value is possible only for $u(t) = 0$ and $x(t)^2 = 1$ for all t. Thus, the function x at the arbitrary point t is equal to 1 or -1. However, for zero control the corresponding state function is constant. We conclude that the considered optimal control problem has two solutions[2] that are the pairs (u_0, x_2) and (u_0, x_3). Thus, we have the absence of the uniqueness of the problem solution and the sufficiency of the optimality conditions[3].

Consider another example with similar properties.

Example 17.2 *Minimize the functional*

$$I(u,x) = \int_0^1 \sin x(t)dt + \frac{1}{2}\int_0^1 u(t)^2 dt,$$

where the functions x and u are related by the state equation $x'(t) = u(t)$, $t \in (0,1)$.

Determine the function

$$H = pu - (\sin x + u^2/2).$$

The function p is a solution of the equation $p' = \cos x$ with homogeneous boundary conditions. From the maximum principle, it follows the equality $u = p$.

As in the previous example, we try to find a solution to the optimality conditions corresponding to the values $p(t) = 0$ for all t. According to the two previous equalities, this is possible for $u(t) = 0$ and $\cos x(t) = 0$ for any t. The last equality holds for the values $x(t) = \pi/2 + k\pi$ for all $t \in (0,1)$, where k is an arbitrary integer. Note that any constant function satisfies the state equation for control u_0, which is identically equal to zero. Thus, at least all pairs (u_0, x_k), where x_k is the kth of the above constant functions, satisfy the optimality conditions.

Let us now turn to the existing optimality criterion. The first integral in its definition is estimated from below by the value -1, and the second is non-negative, which means $I(u,x) \geq -1$ for any values of the argument of the functional. Obviously, the equality $I(u,x) = -1$ is possible only for $u(t) = 0$ and $x(t) = 3\pi/2 + 2k\pi$ for any $t \in (0,1)$ for all integer k. Naturally, all such pairs of functions are related by the state equation, i.e., are acceptable. It follows that among all solutions to the optimality conditions there are only pairs (u_0, x_k) for odd values of k. Thus, in this case, the optimal control problem has an infinite set of solutions, the optimality conditions are necessary, but not sufficient, and there is also an infinite set of non-optimal solutions to the optimality conditions.

17.1.2 Singular controls

For the previously considered classes of optimal control problems, we encountered the degeneration of the maximum principle. Let us show that a similar situation is possible for problems with a free initial state.

Example 17.3 *Minimize the functional*[4]

$$I(u,x) = \frac{1}{2}\int_0^1 x(t)^2 dt,$$

where the functions x and u are related by the state equation $x'(t) = u(t)$, $t \in (0,1)$, besides u belongs to the set U of functions satisfying the inequality $|u(t)| \le 1$ for all $t \in (0,1)$.

Determine the function $H = pu – x^2/2$. Then the adjoint system takes the form

$$p'(t) = x(t),\ t \in (0,1);\ \ p(0) = 0,\ p(1) = 0.$$

The corresponding maximum condition is

$$p(t)u(t) = \max_{|v|\le 1} p(t)v,\ t \in (0,1).$$

Obviously, it degenerates if $p = 0$. Now from the adjoint equation, it follows $x = 0$. Using the state equation, we get $u = 0$. The optimality criterion is non-negative. It can be equal to zero only for $x = 0$. Therefore, the optimization control problem has a unique solution that is the found singular control[5].

The optimality condition for the considered example is necessary and sufficient. Of course, another situation can be possible.

Example 17.4 *Minimize the functional*

$$I(u,x) = \frac{1}{4}\int_0^1 (x^4 - 2x^2t^2)dt$$

on the set of pairs (u,x), satisfying the state equation $x'(t) = u(t)$, $t \in (0,1)$, besides u belongs to the set U of functions satisfying the inequality $|u(t)| \le 1$ for all $t \in (0,1)$.

The function H here is determined by the formula

$$H = pu - (x^4 – 2x^2t^2)/4.$$

Now the adjoint system takes the form

$$p'(t) = x(t)^3 - x(t)t^2,\ t \in (0,1);\ \ p(0) = 0,\ p(1) = 0.$$

We can have a singular control if $p(t) = 0$ for all t. This is possible, if the function x has one of these three values t, $–t$, or 0 at the arbitrary point t. However, from the state equation, it follows the equality

$$x(t) = \int_0^t u(\tau)d\tau + c,$$

where c is an arbitrary constant. Then the function x is continuous. Thus, a singular control can exist for three state functions $x_1(t) = t$, $x_2(t) = -t$, and $x_3(t) = 0$ for all t, i.e., the points of jump are impossible here. By the state equation, these state functions correspond to the constant controls $u_1 = 1$, $u_2 = -1$, and $u_3 = 0$. There are admissible, so we found the singular solutions of the maximum principle. Check the validity of the Kelley condition.

Calculate the derivative

$$\frac{\partial H}{\partial u} = p.$$

Using the adjoint equation, we find

$$\frac{d}{dt}\frac{\partial H}{\partial u} = p' = x^3 - xt^2.$$

By the state equation, we get

$$\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 3x^2x' - x't^2 - 2xt = u(3x^2 - t^2) - 2xt.$$

Now we have

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} = 3x^2 - t^2.$$

Check the sign of this value for our singular controls. We obtain

$$\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=\pm 1} = 3t^2 - t^2 = 2t^2, \quad \frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u}\Big|_{u=0} = -t^2.$$

Thus, the Kelley condition is true for first two singular controls, and this gets broken for the third one. Therefore, the controls u_1 and u_2 can be optimal, but the control u_3 is non-optimal. It is easy to see[6] that this problem really has two solutions, which are the pairs (u_1, x_1) and (u_2, x_2).

17.1.3 Insolvability of an optimal control problem

In the process of analyzing other types of optimal control problems, we encountered another unpleasant situation. Consider the following example.

Example 17.5 *Minimize the functional*[7]

$$I(u,x) = \frac{1}{2}\int_0^1 (x^2 - u^2)dt$$

on the set of pairs (u,x), satisfying the state equation $x'(t) = u(t)$, $t \in (0,1)$, besides u belongs to the set U of functions satisfying the inequality $|u(t)| \leq 1$ for all $t \in (0,1)$.

For this problem, we could write the optimality conditions in the form of the maximum principle, similar to what was done in the previous examples. However, we can carry out here the same analysis as in Chapter 7. From the inequalities $x(t)^2 \geq 0$

and $u(t)^2 \leq 1$ for all $t \in (0,1)$, it follows a lower estimate for the value of the minimized functional

$$I = \frac{1}{2}\int_0^1 x^2 dt - \frac{1}{2}\int_0^1 u^2 dt \geq \frac{1}{2}.$$

Consider now the pair sequence $\{(u_k, x_k)\}$, determined by the equalities[8]

$$u_k(t) = \begin{cases} 1, & \text{if } \frac{2j}{2k} \leq t < \frac{2j+1}{2k}, \\ -1, & \text{if } \frac{2j+1}{2k} \leq t < \frac{2j+2}{2k}, \end{cases} \quad x_k(t) = \begin{cases} t - \frac{2j}{2k}, & \text{if } \frac{2j}{2k} \leq t < \frac{2j+1}{2k}, \\ \frac{2j+2}{2k} - t, & \text{if } \frac{2j+1}{2k} \leq t < \frac{2j+2}{2k}, \end{cases}$$

where $j = 0, 1, ..., k–1$; $k = 1, 2, ...$ Obviously, for any k, we have the inclusion $u_k \in U$ and the equality $x_k'(t) = u_k(t)$ for all $t \in (0,1)$, i.e., all pairs (u_k, x_k) are admissible. Using the conditions

$$|u_k(t)| = 1, \ 0 \leq x_k(t) \leq 1/2k, \ t \in (0,1), \ k = 1, 2, ...,$$

determine the inequality

$$-\frac{1}{2} \leq I(u_k, x_k) = \frac{1}{2}\int_0^1 (x_k^2 - u_k^2)dt \leq \frac{1}{8k^2} - \frac{1}{2}, \ k = 1, 2, ...$$

Passing to the limit as $k \to \infty$, we conclude that $I(u_k, x_k) \to –1/2$. Thus, the number $–1/2$ is the lower bound of the given functional on the set of admissible pairs, and the determined pair sequence is minimizing.

If there exists an admissible pair such that the corresponding value of the functional is $–1/2$, then two equalities hold

$$\int_0^1 x^2 dt = 0, \quad \int_0^1 u^2 dt = 1.$$

From first of them, the function x is zero. Then from the state equation, it follows that the control is zero too. However, this contradicts the second equality. Consequently, both of these equalities cannot be satisfied simultaneously, which means that the lower bound of the functional is not achieved. Thus, the considered optimal control problem has no solution[9].

17.1.4 Ill-posed optimal control problems

In the previous chapters, we encountered ill-posed optimal control problems. Naturally, similar problems arise in the absence of initial conditions for the state equation. We return to Example 17.3, where we consider the minimization of the functional

$$I(u,x) = \int_0^1 x^2 dt$$

on the set of pairs (u, x), related by the equation of state $x'(t) = u(t)$ for all $t \in (0, 1)$, besides u belongs to the set U of functions satisfying the inequality $|u(t)| \le 1$ for all $t \in (0, 1)$.

When studying the problem of well-posedness of a problem, one should indicate the functional space on which the optimality criterion is defined. As such, we choose the product $L_2(0, 1) \times H^1(0, 1)$, discussed in the previous chapter.

It was noted earlier that the solution to the problem posed is a pair of functions (u_0, x_0), identically equal to zero. Consider a sequence of pairs of functions $\{(u_k, x_k)\}$, where for an arbitrary number k the equalities $u_k(t) = \cos k\pi t$, $x_k(t) = \sin k\pi t / k\pi$ hold. This sequence is admissible, and we have the convergence $I(u_k, x_k) \to 0$. Thus, this sequence is minimizing. At the same time, the convergence of $u_k \to u_0$ in $L_2(0, 1)$ does not take place[10], which means that this optimal control problem is Tikhonov ill-posed.

Now consider the following example.

Example 17.6 *Minimize the functional*[11]

$$I_k(u, x) = \int_0^1 (x - y_k)^2 dt$$

on the set of pairs (u, x)*, satisfying the state equation* $x'(t) = u(t)$*,* $t \in (0, 1)$*, besides* u *belongs to the set* U *of functions satisfying the inequality* $|u(t)| \le 1$ *for all* $t \in (0, 1)$*, where* $y_k(t) = (k\pi)^{-1} \sin k\pi t$*,* k *is a number.*

Obviously, this functional has non-negative values. It can be equal to zero only for $x(t) = y_k(t)$ for all t. Then from the state equation, it follows $u_k(t) = y_k'(t) = \cos k\pi t$. This control belongs to the set U, so the pair (u_k, y_k) the unique solution of the considered optimal control problems.

Passing to the limit at the formula of the optimality criterion, we find the limit value

$$I_\infty = \lim_{k\to\infty} I_k(u, x) = \int_0^1 x^2 dt.$$

The problem of minimizing it on the set of admissible pairs, which actually coincides with Example 17.3, has a zero solution (u_0, x_0). However, due to the absence of convergence $u_k \to u_0$ in $L_2(0, 1)$, the sequence of optimal pairs $\{(u_k, y_k)\}$ for Example 17.6 does not converge to the limit optimal pair (u_0, x_0). Thus, the considered optimal control problem is Hadamard ill-posed.

RESULTS

Here is a list of questions in the area of considered examples of optimal control problems for systems with a free initial state, the main conclusions on this topic, as well as problems arising in this case that require additional research.

Questions

It is required to answer questions concerning optimal control problems for systems with a free initial state.

1. Is it possible to say that the solution to the optimality conditions for Example 17.1 is necessarily realized when the equality $p(t) = 0$ for all t is satisfied?

2. On what basis was it concluded that the optimality conditions for Example 17.1 are satisfied by the control that is identically equal to zero?

3. Why do several system states correspond to zero control?

4. Why do only three system states corresponding to zero control determine solutions to the optimality conditions for Example 17.1?

5. The identical equality to zero of the function p for Example 17.1 is realized if at an arbitrary point t the function x takes one of the three indicated values. Can this function take on different values at different points from among those available and why?

6. On what basis is it concluded that the optimality conditions for Example 17.1 are insufficient?

7. On what basis is it concluded that the two solutions to the optimality conditions for Example 17.1 are optimal?

8. What influenced the absence of uniqueness of the solution and sufficiency of optimality conditions for Example 17.1?

9. How are Examples 17.1 and 17.2 similar and how are they different?

10. Can we say that we have found all solutions to the optimality conditions for Example 17.2?

11. Can we say that we have found all the optimal pairs for Example 17.2?

12. On what basis is it concluded that certain solutions to the optimality conditions for Example 17.2 are optimal?

13. What influenced the absence of uniqueness of the solution and sufficiency of optimality conditions for Example 17.2?

14. In Examples 6.2 and 17.3, the same functional is minimized on the same set of admissible controls with the same equation of state, and in both cases the zero control turns out to be optimal. For the first of these examples, optimal control corresponds to a unique state of the system, and in the second, a whole family of solutions to the equation of state corresponds to the absence of an initial condition. Why do we only take one of them into account?

15. How does the absence of an initial condition affect the presence or absence of singular control?

16. How does the absence of an initial condition affect the optimality of a singular control?

17. Are there regular solutions to the maximum condition for Example 17.3?

18. Is it possible to prove the uniqueness of the optimal control for Example 17.3 using Theorem 5.1?

19. Is it possible to prove the sufficiency of the optimality condition for Example 17.3 using Theorem 5.2?

20. Why is the singular control optimal in Example 17.3, but there is a non-optimal singular control in Example 17.4?

21. Are there regular solutions to the maximum condition for Example 17.4?

22. Why is there not a unique solution to the problem in Example 17.4?

23. In the process of analyzing Example 17.4, it turned out that the system state function at any point t takes only one of three values. Why does it follow that she cannot switch from one of these values to another?

24. How does it follow that one of the three found singular controls is not optimal for Example 17.4?

25. Whence does it follow that two of the three found singular controls are optimal for Example 17.4?

26. Example 17.5 differs from Example 7.1 only in the absence of an initial state, and both of these optimal control problems have no solution. Is it, in principle, possible for a situation where, by eliminating the initial condition, an unsolvable problem becomes solvable?

27. How does the analysis of Examples 17.5 and 7.1 differ?

28. In Example 17.5, what would change if no restrictions were placed on the control values?

29. How does the considered sequence of controls behave for Example 17.5?

30. How does the considered sequence of states behave for Example 17.5?

31. Can the sequence of pairs considered for Example 17.5 be used to find an approximate solution to the problem?

32. How does the analysis of Examples 17.6 and 6.2 differ?

33. How does the considered sequence of controls behave for Example 17.6?

34. How does the considered sequence of states behave for Example 17.6?
35. What will change if, in Example 17.6, no restrictions are imposed on the control values?

Conclusions

Based on the study of specific problems of optimal control of systems with a free initial state, we can come to the following conclusions.

- In problems of optimal control of systems with a free initial state, a non-unique solution is possible.
- In optimal control problems for systems with a free initial state, optimality conditions may be insufficient.
- The optimality conditions for Example 17.1 have at least three solutions.
- The optimal control problem for Example 17.1 has two solutions.
- The optimality conditions for Example 17.2 have an infinite set of solutions.
- Among the solutions to the optimality conditions for Example 17.2, there are an infinite set of optimal pairs and an infinite set of non-optimal pairs.
- In problems of optimal control of systems with a free initial state, the maximum condition may degenerate.
- For the optimal control problem in Example 17.3, there is a unique singular control that is optimal.
- For the optimal control problem in Example 17.4, there are three singular controls.
- Two singular controls from Example 17.4 satisfy the Kelley condition, but one does not.
- The singular controls from Example 17.4 that satisfy the Kelley condition are optimal.
- In problems of optimal control of systems with a free initial state, there may be no solution.
- The optimal control problem in Example 17.5 has no solution.
- Optimal control problems for systems with a free initial state may turn out to be ill-posed according to both Tikhonov and Hadamard.
- The optimal control problem from Example 17.5 has a unique solution, but is Tikhonov ill-posed.
- The optimal control problem from Example 17.6 has a unique solution, but is Hadamard ill-posed.

Problems

In the process of analyzing optimal control problems for systems with a free initial state, additional problems arise that require additional research.

1. **Non-uniqueness of optimal pairs and insufficiency of optimality conditions for the case when solutions to optimality conditions differ not only in states, but also in controls**. In the examples discussed in the Lecture, the optimality conditions for systems with a free initial state sometimes had a non-unique solution, and different solutions turned out to be both optimal and non-optimal. However, the corresponding control-state pairs differed solely by state. Appendix considers a problem of the indicated class, for which there are several pairs of solutions to optimality conditions that differ in both states and controls.

2. **Special properties of the singular control**. When analyzing singular controls in Chapter 6, we encountered a situation where a certain control, rejected in the process of searching for regular solutions of the maximum principle, turns out to be its singular solution, and the optimal one. Appendix gives an example of an optimal control problem for a system with a free initial state, for which a similar property is observed.

3. **Infinity of the set of singular controls**. We have already encountered an infinite number of singular controls, both optimal and non-optimal, when studying systems with a given initial state. In particular, for the minimization problem from Example 6.1 all singular controls turned out to be optimal, but in the corresponding maximization problem they were not optimal. One would like to find out what happens if the initial condition is eliminated for this system. This problem is explored in Appendix.

4. **Optimal control problems with a free initial state under isoperimetric conditions**. In the previous part, optimal control problems with free and fixed final states with given initial conditions were considered. It would be interesting to consider similar problems in the absence of initial conditions. Examples of problems in this class are given in Appendix.

17.2 APPENDIX

We will consider below examples of optimal control problems for systems with a free initial state, different from those considered earlier. In particular, Section 17.2.1 analyzes a problem in which the non-uniqueness of the solution and the insufficiency of the optimality conditions are realized in a situation where pairs of solutions to the optimality conditions differ not only in the states of the system, but also in the controls. In Section 17.2.2, the control that was previously rejected in the study of regular solutions of the maximum principle turns out to be optimal. In Section 17.2.3, an infinite number of singular controls turn out to be non-optimal, although for a similar problem with an initial condition they were optimal.

Finally, Sections 17.2.4 and 17.2.5 consider examples of the class of systems under study in the presence of isoperimetric constraints.

17.2.1 Non-uniqueness and insufficiency under different controls

In Examples 17.1 and 17.2, the optimality conditions were satisfied by many control-state pairs, some of which were optimal and some of which were non-optimal. However, in both examples, the solutions to the optimality conditions differed only in the state of the system, while the control for all found pairs of each example was the same. Naturally, such a situation occurs only in exceptional cases.

Example 17.7 *Minimize the functional*[12]

$$I(u,x) = \frac{1}{4}\int_0^1 (x^4 - 2x^2t^2 + u^4 - 2u^2)\,dt,$$

where the functions x and u are related by the state equation $x'(t) = u(t)$, $t \in (0,1)$.

The function H is determined by the formula

$$H = pu - \frac{1}{4}(x^4 - 2x^2t^2 + u^4 - 2u^2).$$

Then the adjoint system takes the form

$$p'(t) = x(t)^3 - x(t)t^2,\ t \in (0,1);\ p(0) = 0,\ p(1) = 0.$$

From the maximum condition, it follows the equality

$$p(t) = u(t)^3 - u(t),\ t \in (0,1).$$

One can try to find the pairs (u,x) satisfying the system of optimality conditions if $p(t) = 0$ for all t. In accordance with the previous equality, this is possible only in the case when the function u for an arbitrary t takes one of three values 0, 1, or −1. On the other hand, from the adjoint equation for the indicated function p it follows that for an arbitrary t the function x takes one of three values 0, t, or $-t$. In this case, the functions u and x must be related by the state equation. It is easy to verify that only three pairs (u_1,x_1), (u_2,x_2), and (u_3,x_3) have the indicated properties, where $u_1(t) = 0$, $x_1(t) = 0$, $u_2(t) = 1$, $x_2(t) = t$, $u_3(t) = -1$, $x_3(t) = -t$ for all $t \in (0,1)$. Thus, these three function pairs satisfy the system of optimality conditions[13].

For finding a best pair, transform the integrand of the optimality criterion.

$$x^4 - 2x^2t^2 + u^4 - 2u^2 = (x^2 - t^2)^2 - t^4 + (u^2 - 1)^2 - 1.$$

The second and fourth terms on the right side here are concrete quantities, and the first and third take exclusively non-negative values. They can be identically equal to zero only in the case when, for an arbitrary t, the function u takes one of two values

1 or –1, and the function x takes the values t or $-t$. Considering that these functions are related by the equation of state, we conclude that the problem being solved has exactly two optimal pairs (u_2, x_2) and (u_3, x_3). Thus, we are again dealing with the absence of uniqueness of optimal control and sufficiency of optimality conditions, but now the solutions to optimality conditions differ not only in the functions of the state of the system, but also in the controls.

17.2.2 Special property of the singular control

Consider a non-standard relation between regular and singular solutions of the maximum principle.

Example 17.8 *Minimize the functional*[14]

$$I(u,x) = \frac{1}{4}\int_0^1 (x^2 - 2xt)dt$$

on the set of pairs (u, x) *related by the state equation* $x'(t) = u(t)$ *for all* $t \in (0,1)$ *satisfying the inequality* $|u(t)| \leq 1$ *for all* $t \in (0,1)$.

Determine the function

$$H = pu - (x^2 - 2xt)/2.$$

Then the adjoint system takes the form

$$p' = x - t,\ t \in (0,1);\ p(0) = 0,\ p(1) = 0. \tag{17.2}$$

In accordance with the maximum principle, there are two options here. In the regular case, the maximum of the function H reaches at the boundary of the set of admissible controls, and in the singular case a singular control is implemented.

Consider, at first, the regular case. Find the control from the maximum principle

$$u_\varepsilon(t) = \begin{cases} -1, & \text{if } p_\varepsilon(t) > 0, \\ 1, & \text{if } p_\varepsilon(t) < 0. \end{cases} \tag{17.3}$$

Suppose $u(t) = 1$ for all t that is possible only for the positive values of the function p because of the equality (17.3). Then we find $x(t) = t + c$ with arbitrary constant c using the state equation. Integrating the adjoint equation (17.2) with using the boundary conditions, we get

$$\int_0^1 p'dt = \int_0^1 cdt = c = 0.$$

Therefore, $x(t) = t$. Now the adjoint equation takes the form $p' = 0$. Using the corresponding boundary conditions, we conclude that $p(t) = 0$ for all t, which contradicts the formula (17.3). Thus, this control cannot determine the solution of the optimality conditions[15].

Suppose now $u(t) = –1$ for all t that is possible for negative p. Then the general solution of the state equation is $x(t) = –t + c$, where the constant c is arbitrary. After integration of the adjoint equation with using the boundary conditions, we obtain

$$\int_0^1 p' dt = \int_0^1 (c - 2t)dt = c - 1 = 0.$$

We find $c = 1$, so $x(t) = 1–t$. Therefore, the adjoint equation can be written as $p' = 1–2t$. Solving this equation with a zero initial condition, we find $p(t) = t–t^2$. This function is positive on a given interval, which contradicts equality (17.3). To summarize, we conclude that control cannot be continuous.

Let us assume that ξ is the first control discontinuity point, and initially $u(t) = 1$, which is possible if the function p is positive on the interval $(0, \xi)$. In this section $x(t) = t + c$. Then $p' = c$, i.e., the function p is monotonic. However, it goes to zero both at the beginning of this interval (initial condition) and at the end (when passing through point ξ, it changes sign). Consequently, it is equal to zero, which again contradicts condition (17.3)

Let us assume now that ξ is the first control discontinuity point, and initially $u(t) = –1$, which is possible if the function p is positive on the interval $(0, \xi)$. Then $x(t) = –t + c$. The adjoint equation on this interval is $p' = c–2t$. Integrating this equality over the interval $(0, \xi)$, we find $c = 1$. Then $p' = 1–2t$. Solving it with a zero initial condition, we find $p = t–t^2$. This function is positive on the unit interval, which contradicts equality (17.3).

Thus, there are no regular solutions to the maximum principle. However, the existence of singular control is possible. This is realized at $p = 0$, which corresponds to the equality $x(t) = t$. Therefore, we obtain $u = 1$. Previously, we rejected this control in the process of studying the possibility of the existence of regular solutions to the maximum principle. However, it turned out to be a singular control. It is easy to verify that it is optimal[16].

17.2.3 Infinite set of singular controls

Many examples considered in this part are analogs of previously studied optimal control problems with a given initial state of the system. Let us now consider the corresponding analog of Example 6.1.

Example 17.9 *Minimize the functional*

$$I(u, x) = \int_0^1 x(t)u(t)dt$$

on the set of pairs (u, x) related by the state equation $x'(t) = u(t)$ for all $t \in (0, 1)$ satisfying the inequality $|u(t)| \leq 1$ for all $t \in (0, 1)$.

Determine the function $H = pu–xu$. Then the adjoint equation takes the form

$$p'(t) = u(t),\ t \in (0,1);\ \ p(0) = 0,\ \ p(1) = 0.$$

Integrating the adjoint equation with the boundary conditions, we get

$$\int_0^1 u(t)dt = 0. \tag{17.4}$$

Obviously, degeneration of the maximum principle here is possible if the equality[17] $x(t) = p(t)$ is true for all $t \in (0,1)$. Thus, the function x satisfies the same conditions as p, so the following equalities hold

$$x'(t) = u(t),\ t \in (0,1);\ x(0) = 0,\ \ x(1) = 0. \tag{17.5}$$

Note that the second boundary condition (17.5) is a consequence of integrating the equation of state taking into account the first boundary condition and equality (17.4).

Thus, in this case, any element of the set U that satisfies equality (17.4) has special control. Naturally, such controls constitute an infinite and not even countable set. The result obtained exactly coincides with what was established during the study of Example 6.1. Thus, the solution to the system of optimality conditions for this example is any pair (u,x), where u is a singular control, and x is a solution to a given state equation with a zero initial condition.

Let us now turn to the direct study of the optimality criterion. We have

$$I(u,x) = \int_0^1 xudt = \int_0^1 xx'dt = \frac{1}{2}\int_0^1 \frac{d}{dt}(x^2)dt = \frac{x(1)^2}{2} - \frac{x(0)^2}{2}.$$

For a problem with a fixed (zero) state of the system, the second term on the right side of the last equality becomes zero, and the first is non-negative. Thus, the functional can be equal to zero only when $x(1) = 0$, which corresponds to equality (17.4). As a result, it turned out that any singular control for Example 6.1 turns out to be optimal, which means that the problem has an infinite number of solutions.

The initial state is not fixed for our case. Transform the previous equality using the state equation. We obtain

$$x(1) = x(0) + \int_0^1 u(t)dt.$$

As a result, the formula for the optimality criterion takes the form

$$I(u,x) = \frac{x(1)^2}{2} - \frac{x(0)^2}{2} = x(0)\int_0^1 u(t)dt = \frac{1}{2}\Big[\int_0^1 u(t)dt\Big]^2.$$

Based on the set of admissible controls, we conclude that the integral on the right side of this equality takes values from the interval [–1,1]. However, there are no restrictions on the value of $x(0)$. As a result, the functional under consideration turns out to be unbounded from below on the set of admissible pairs (u, x). Thus, this optimal control problem has no solution, and the optimality conditions are necessary, but not sufficient[18].

17.2.4 Problem with an isoperimetric condition with respect to control

For systems with free and fixed final states, optimal control problems with isoperimetric conditions were previously considered. Systems in the absence of initial conditions can be studied in a similar way. In this case, in accordance with the Lagrange multiplier method[19], the function H is introduced in the same way as for the previous problems, i.e., depends additionally on the numerical Lagrange multiplier associated with the isoperimetric condition. Regarding the optimal control, the maximum condition[20] for this function H is valid, the definition of which includes a function x that satisfies the equation of state without boundary conditions, and a function p that satisfies the adjoint equation with two boundary conditions.

Example 17.10 *Minimize the functional*

$$I(u, x) = \frac{1}{2}\int_0^1 (x^2 - xt^2 + u^2 - 2ut)dt$$

on the set of pairs (u, x) *related by the state equation* $x'(t) = u(t)$ *for all* $t \in (0, 1)$ *under the isoperimetric condition*

$$\int_0^1 udt = \frac{1}{2}.$$

In accordance with the described methodology, the following function is determined

$$H = pu + \lambda(u{-}1/2) - (x^2 - xt^2 + u^2 - 2ut)/2.$$

The function p here is the solution of the adjoint system

$$p'(t) = x(t) - t^2/2,\ t \in (0, 1);\ p(0) = 0,\ \ p(1) = 0. \tag{17.6}$$

Find the control from the maximum condition

$$u(t) = p(t) + t + \lambda. \tag{17.7}$$

Now with respect to the three unknown functions u, x, p and the number λ, there is a system that includes the control formula (17.7), the state equation and the adjoint equation with two boundary conditions (17.6), as well as the isoperimetric condition.

Differentiating the adjoint equation, we get

$$p'' = x' - t = u - t = p + \lambda.$$

Integrating the equality (17.7) using the isoperimetric condition, we obtain

$$\int_0^1 u dt = \int_0^1 p dt + \frac{1}{2} + \lambda = \frac{1}{2}.$$

Now we find

$$\lambda = -\int_0^1 p dt. \tag{17.8}$$

Then we have the following boundary value problem for the integro-differential equation

$$p''(t) - p(t) = \int_0^1 p(\tau)d\tau, \ y \in (0,1); \ p(0) = 0, \ p(1) = 0.$$

After multiplication by p and integration, we get

$$\int_0^1 p''(\tau)p(\tau)d\tau - \int_0^1 p(\tau)^2 d\tau = \Big[\int_0^1 p(\tau)d\tau\Big]^2.$$

Transforming the first integral here using the integration by parts formula and taking into account the existing boundary conditions, we have

$$\int_0^1 p'(\tau)^2 d\tau + \int_0^1 p(\tau)^2 d\tau + \Big[\int_0^1 p(\tau)d\tau\Big]^2.$$

Obviously, the resulting equality is true only for the function p identically equal to zero. Then from formula (17.8), it follows that $\lambda = 0$. As a result, from formula (17.7), we find the control $u(t) = t$, and from the first equality (17.6), we find the state function $x(t) = t^2/2$. Thus, the system of optimality conditions for the considered example has a unique solution that is the found pair (u, x). It is easy to verify that this is the optimal pair for this example[21].

17.2.5 Problem with an isoperimetric condition with respect to state

Example 17.10 considered an optimal control problem with a free initial state in the presence of an isoperimetric condition that includes control. Let us now consider a problem that differs from the previous one only in that the isoperimetric condition includes the system state function.

Example 17.11 *Minimize the functional*

$$I(u,x) = \frac{1}{2}\int_0^1 (x^2 - xt^2 + u^2 - 2ut)dt$$

on the set of pairs (u,x) *related by the state equation* $x'(t) = u(t)$ *for all* $t \in (0,1)$ *under the isoperimetric condition*

$$\int_0^1 xdt = \frac{1}{6}.$$

The function H now is determined by the formula

$$H = pu + \lambda(x - 1/6) - (x^2 - xt^2 + u^2 - 2ut)/2.$$

The corresponding adjoint system takes the form

$$p'(t) = -\lambda + x(t) - t^2/2,\ t \in (0,1);\ p(0) = 0,\ p(1) = 0.$$

From the maximum condition, it follows

$$u(t) = p(t) + t.$$

Let us find a solution to the system of optimality conditions. Integrating the adjoint equation taking into account the boundary conditions available for it, as well as the isoperimetric condition, we have

$$\int_0^1 p'dt = -\lambda + \int_0^1 xdt - \frac{1}{2}\int_0^1 t^2 dt = -\lambda = 0.$$

Then, after differentiating the adjoint equation, taking into account the equation of state and the previously obtained formula for control, we obtain

$$p'' = x' - t = u - t = p.$$

The resulting second-order equation with homogeneous boundary conditions has a solution that is identically equal to zero. Then from the control formula obtained above we find $u(t) = t$, and from the adjoint equation, taking into account the found value of λ, it follows that $x(t) = t^2/2$. Obviously, the resulting pair (u,x) is the only solution to the optimal control problem under consideration[22].

Additional conclusions

Based on the results of the study of the previously discussed examples, the following conclusions can be drawn.

- For Example 17.7, the system of optimality conditions is satisfied by three pairs of functions that differ in both the states they contain and the controls.
- The optimality conditions for Example 17.7 are necessary but not sufficient.
- The optimal control problem from Example 17.7 has two solutions that differ in sign. Example 17.8 explores the possibility of the existence of regular and singular solutions to the maximum principle, with a control rejected as a possible regular solution being a singular control.
- The singular control from Example 17.8 is optimal.
- The optimality conditions for Example 17.8 are necessary and sufficient.
- For Example 17.9, there is an infinite set of singular controls, coinciding with a similar set for Example 6.1, which differs from the one considered only in the presence of the initial condition.
- All singular controls for Example 17.9 satisfy the Kelley condition and its generalizations.
- All the singular controls for Example 17.9 are non-optimal, although for Example 6.1 they are all optimal.
- The optimal control problem for Example 17.9 has no solution.
- The optimality conditions for Example 17.9 are not sufficient.
- The optimal control problem for Example 17.10 with the isoperimetric condition has a unique solution.
- The optimality conditions for Example 17.10 with the isoperimetric condition are sufficient.
- The optimal control problem for Example 17.11 with the isoperimetric condition has a unique solution.
- The optimality conditions for Example 17.11 with the isoperimetric condition are sufficient.

Notes

1. In principle, it is possible that the boundary value problem (17.1) has other solutions.

2. Note that this optimal control problem is invariant under sign changes. In this regard, the presence of two optimal pairs, differing only in signs, is quite natural.

3. One may wonder why for this example the optimality conditions turned out to be insufficient, and the optimal control was not unique. If we turn to the general theorems on uniqueness and sufficiency from Chapter 5, we can pay attention to the requirement that the optimality

criterion be convex. It is easy to verify that the functional under consideration does not have this property.

4. This problem differs from the one considered in Example 6.2 only in the absence of an initial condition.

5. Naturally, to solve this problem it was possible to exclude control from consideration altogether, since the functional depends on the function x. It is clear that the quadratic functional reaches its minimum when the corresponding integrand is equal to zero.

6. Indeed, the integrand of the minimized functional can be written as $(x^4 – 2x^2t^2) = (x^2 – t^2)^2 + t^4$. Here the second term on the right side does not depend on the choice of control and state, and the first, being non-negative, can become zero only in the case when at any point t the function x takes the values t or $–t$. In the absence of an equation of state, we could conclude that any function x taking the value t on one part of the interval (0,1) and $–t$ on another part of it provides a minimum to the functional. However, the presence of an equation of state with restrictions on controls excludes the possibility of the existence of a discontinuous function x. Thus, solutions to the problem can only be functions (u_1, x_1) and (u_2, x_2).

7. This problem differs from the one considered in Example 7.1 only in the absence of an initial condition.

8. In Chapter 7, we considered the same control sequence $\{u_k\}$; see Figure 7.1. After solving the corresponding state equation with the given initial condition, we determined the considered state sequence $\{x_k\}$; see Figure 7.2. Now we can not to determine it after solving the state equation because of the absence of the initial condition. In this regard, we immediately set the corresponding sequence of pairs.

9. Obviously, the given minimizing sequence of pairs allows us to determine a weak approximate solution to the problem.

10. The rationale for this statement is given in Chapter 8, where an example is studied that differs from the one considered only in the absence of an initial condition; see Example 6.2. The analysis carried out in this case differs from that done in Chapter 8 by considering not a minimizing sequence of controls, but a minimizing sequence of control-state pairs.

11. This problem differs from the one considered in Example 8.1 only in the absence of an initial condition.

12. This problem differs from the one considered in Example 7.1 only in the absence of an initial condition.

13. In this case, it does not matter whether there are solutions to the optimality conditions with non-zero values of the function p. The results obtained are already sufficient to justify both the non-uniqueness of the solution to the problem and the insufficiency of optimality conditions.

14. This problem differs from the one considered in Example 6.5 only in the absence of an initial condition.

15. In reality, it cannot be a regular solution, but, as we will soon see, it is a singular control.

16. To do this, it is enough to present the integrand in the optimality criterion in the form $(x – t)^2 – t^2$. Does the second term here not depend on the choice of pair? And the first, being non-negative, vanishes at $x(t) = t$ at all points.

17. In this case, it does not matter to us whether there are regular solutions to the maximum principle.

18. As already noted, Example 17.9 differs from Example 6.1 solely in the absence of an initial condition. For the latter, the validity of Kelley condition and its generalizations was established in Chapter 6, albeit in a degenerate form. Moreover, all singular controls were optimal. Obviously, the presence or absence of an initial condition does not affect the validity of these statements, i.e., for Example 17.9 they are still executed. However, in this case, singular controls cannot be optimal due to the insolvability of the problem. This does not contradict the validity of Kelley condition and its generalizations, since they provide necessary, but in the general case not sufficient conditions for the optimality of a singular control.

19. As before, to study problems with an isoperimetric condition, one can use the penalty method.

20. This result is justified in the same way as similar statements for other types of systems.

21. Obviously, our integrand can be written as

$$x^2 – xt^2 + u^2 – 2ut = (x – t^2/2)^2 – t^4/4 + (u – t)^2 – t^2 = 0.$$

Here, the second and fourth terms on the right side do not depend on the choice of pair (u, x), and the first and third are non-negative. They can vanish exclusively for a pair characterized by the equalities $u(t) = t$ and $x(t) = t^2/2$, which corresponds to the found solution of the optimality conditions, which thereby turns out to be optimal. Note also that in this case it is possible to establish the existence and uniqueness of a solution to the problem and the sufficiency of optimality conditions based on the theorems given in Part II.

22. In fact, Examples 17.10 and 17.11 consider the same optimal control problem. In the absence of an isoperimetric condition, its solution is characterized by the functions $u(t) = t$ and $x(t) = t^2/2$, the first of which satisfies the isoperimetric condition from Example 17.10, and the second satisfies the isoperimetric condition from Example 17.11. In this regard, adding isoperimetric conditions does not change the result.

List of examples

Here is a list of examples discussed in this book, with the example's own number in brackets, followed by a description of its most important properties.

PART I. FUNCTION MINIMIZATION

1. (1.1). The stationarity condition has a unique solution and is a necessary and sufficient condition for a minimum. The problem has a unique solution and is Tikhonov well-posed.

2. (1.2). The stationary condition has three solutions and is a necessary but not sufficient condition for a minimum. The problem has only one solution.

3. (1.3). The stationary condition has three solutions and is a necessary but not sufficient condition for a minimum. The problem has two solutions.

4. (1.4). The stationary condition has five solutions and is a necessary but not sufficient condition for a maximum. The problem has three solutions.

5. (1.5). The stationary condition has a countable set of solutions and is a necessary but not sufficient condition for a minimum. The problem has a countably many solutions. The countable set of solutions to the stationary condition are not the minimum points of the function.

6. (1.6). The stationary condition has a continuum of solutions and is a necessary and sufficient condition for a minimum. The problem has a continuum of solutions.

7. (1.7). The stationary condition has three solutions and is a necessary but not sufficient condition for a minimum. The problem has a unique solution and is Tikhonov ill-posed.

8. (1.8). The stationary condition has no solution and is a necessary and sufficient condition for a minimum. The problem has no solution.

9. (1.9). The stationary condition has a unique solution and is a necessary but not sufficient condition for a minimum. The problem has no solution.

10. (1.10). The stationary condition has two solutions and is a necessary but not sufficient condition for a minimum. The problem has no solution.

11. (1.11). The stationary condition is not applicable due to the non-smoothness of the function. The problem has only one solution. The solution is found using the subgradient of the function.
12. (1.12). The stationary condition is not applicable due to limitations. The problem has only one solution. The solution is found using the variational inequality.
13. (2.1). The variational inequality has a unique solution and is a necessary and sufficient condition for a minimum. The problem has only one solution.
14. (2.2). Depending on the values of the parameters, the variational inequality can have one, two, or three solutions. The problem has only one solution.
15. (2.3). The problem has a unique solution and is Hadamard well-posed.
16. (2.4). The problem has a unique solution, but is Hadamard ill-posed.
17. (2.5). The problem is Hadamard ill-posed. There is a bifurcation of stationary points.
18. (2.6). The problem has a unique solution and is solved iteratively, and the algorithm converges for any initial approximation.
19. (2.7). The problem has a unique solution and is solved iteratively, and the algorithm diverges for any initial approximation.
20. (2.8). The problem has a unique solution and is solved iteratively, and the algorithm converges for some initial approximations and diverges for others.
21. (2.9). The problem has only one solution. The problem is solved iteratively, and the algorithm diverges for any initial approximation, but becomes convergent after modifying the algorithm.
22. (2.10). The problem has two solutions and is solved iteratively. Depending on the initial approximation, the solution algorithm converges to one or another solution.
23. (2.11). For the problem, multiple bifurcation of stationary points is observed. The problem is solved iteratively, and depending on the initial approximation, the algorithm either converges to one of the stationary points or diverges.
24. (2.12). Problem with equality type constraint. The problem is solved by the Lagrange multiplier method and the penalty method. The problem has only one solution.
25. (2.13). A problem in which a weak approximate solution is not strong.
26. (2.14). A problem in which a strong approximate solution is not weak.
27. (2.15). The problem is unsolvable, but a weak approximate solution exists.

PART II. OPTIMAL CONTROL PROBLEMS WITH A FREE FINAL STATE

28. (3.1). The maximum principle has a unique solution and is a necessary and sufficient condition for the minimum. The problem has a unique solution and is Tikhonov well-posed. The variational inequality is equivalent to the maximum principle. Bellman optimality principle is valid.

29. (3.2). The maximum principle has a unique solution and is a necessary and sufficient condition for the minimum. The problem has a unique solution. The variational inequality is not equivalent to the maximum principle.

30. (3.3). The maximum principle has a unique solution and is a necessary and sufficient condition for the minimum. The problem has a unique solution and is Tikhonov well-posed. The variational inequality is equivalent to the maximum principle. The problem is solved iteratively. The algorithm converges for any initial approximation. The problem is solved using the penalty method.

31. (3.4). The maximum principle has a unique solution and is a necessary and sufficient condition for the minimum. The problem has a unique solution and is Tikhonov well-posed. The problem is solved using the decoupling method, iteratively, using dynamic programming and the penalty method.

32. (4.1). The functionality is not smooth. Standard iterative methods are not applicable. The problem is solved by non-smooth optimization methods.

33. (4.2). The state equation is not always solvable. The problem is solved using the penalty method.

34. (4.3). The state equation may have a non-unique solution. The problem is solved using the penalty method.

35. (5.1). The maximum principle has an infinite number of solutions and is a necessary but not sufficient condition for the minimum. Iterative methods for solving the maximum condition converge to different limits for different initial approximations. The problem has two solutions. The variational inequality is less effective than the maximum principle.

36. (5.2). The maximum principle has a unique solution and provides a necessary and sufficient condition for optimality when the conditions of the theorem on the sufficiency of the optimality condition are violated. The problem has a unique solution if the conditions of the uniqueness theorem are violated.

37. (5.3). The problem has three solutions, which are found using the maximum principle.

38. (6.1). The optimality criterion is convex, but not strictly convex. The maximum principle gives an uncountable set of singular controls and is a necessary and sufficient condition for optimality. The variational inequality is equivalent to the maximum principle. All singular controls satisfy the Kelley condition and

the generalized Kopp–Moyer condition of any order in degenerate form. The problem has countless solutions.

39. (6.2). The maximum principle has a unique solution that is singular and is a necessary and sufficient condition for optimality. The singular control satisfies Kelley condition. The problem has a unique solution, but is not Tikhonov well-posed.

40. (6.3). The maximum principle has two solutions that are singular and is a necessary, but not sufficient, condition for optimality. One singular control satisfies the Kelley condition, but the other does not.

41. (6.4). The maximum principle has a unique solution and is a necessary and sufficient condition for optimality. A control suspected of being a regular solution to the maximum principle turns out to be optimal, being a special control. The problem has a unique solution.

42. (6.5). The maximum principle has three solutions, of which one turns out to be singular, is, and provides a necessary, but not sufficient condition for optimality. Singular control does not satisfy Kelley condition. The problem has two solutions, which are regular solutions of the maximum principle.

43. (6.6). The maximum principle has an infinite set of solutions, of which only two are regular, and is a necessary but not sufficient condition for optimality. All singular controls satisfy the Kelley condition and the generalized Kopp–Moyer condition of any order in degenerate form. The problem has two solutions.

44. (6.7). The maximum principle gives five singular controls and is a necessary, but not sufficient condition for optimality. Three singular controls satisfy the Kelley condition, but two do not. The problem has three solutions.

45. (6.8). The maximum principle gives five singular controls and is a necessary, but not sufficient condition for optimality. Two singular controls satisfy the Kelley condition, but three do not. The problem has no solution.

46. (6.9). The maximum principle has a unique solution that is special and is a necessary and sufficient condition for optimality. The singular control satisfies the Kelley condition in degenerate form, as well as the Kopp–Moyer condition. The problem has only one solution.

47. (6.10). The maximum principle gives three solutions, one of which is singular and two are regular. The singular control satisfies the Kelley condition in degenerate form, but does not satisfy the Kopp–Moyer condition. The problem has two solutions.

48. (6.11). The maximum principle gives a unique solution that is singular. The singular control satisfies the Kelley and Kopp–Moyer conditions in degenerate form, as well as the generalized third-order Kopp–Moyer condition. The problem has a unique solution, which is a singular control.

49. (6.12). The maximum principle has three solutions, one of which is singular and two are regular. Singular controls satisfy the Kelley and Kopp–Moyer conditions in degenerate form and do not satisfy the generalized third-order Kopp–Moyer condition. The problem has two solutions that are regular.

50. (6.13). The maximum principle has a unique solution that is singular. The singular control satisfies the Kelley condition and the generalized Kopp–Moyer condition of order $r-1$ in degenerate form, as well as the generalized Kopp–Moyer condition of order. The problem has only one solution.

51. (6.14). The maximum principle gives three solutions, one of which is singular and two are regular. The singular control satisfies the Kelley condition and the generalized Kopp–Moyer condition of order $r-1$ in degenerate form and does not satisfy the generalized Kopp–Moyer condition of order r. The problem has two solutions that are regular.

52. (7.1). The functional to be minimized is non-convex. The maximum principle does not have any solutions, but is a necessary and sufficient condition for the minimum. The problem has no solution. There is a weak approximate solution to the problem.

53. (7.2). The maximum principle has a unique solution that is singular. Singular control does not satisfy Kelley condition. The problem has no solution. There is no weak approximate solution to the problem.

54. (7.3). The functional to be minimized is convex. The maximum principle does not have any solutions, but is a necessary and sufficient condition for the minimum. The problem has no solution. There is a weak approximate solution to the problem.

55. (8.1). The maximum principle has a unique solution that is singular. The problem has a unique solution, but is not Hadamard well-posed. For a fixed value of the parameter, the problem is not Tikhonov well-posed. An approximate solution is found by the regularization method.

56. (8.2). For a fixed value of the parameter, the problem is Tikhonov well-posed. The problem Hadamard well-posed.

PART III. OPTIMAL CONTROL PROBLEM WITH A FIXED FINAL STATE

57. (9.1). The maximum principle has a unique solution. The variational inequality is equivalent to the maximum principle. The penalty method gives an exact solution to the problem. Bellman optimality principle is valid. The solution to the problem is unique, and discontinuous. Changing the type of extremum does not change the general properties of the problem.

58. (9.2). The maximum principle gives a unique solution. The variational inequality is equivalent to the maximum principle. The problem has a unique solution,

although the functional is not strictly convex. The solution to the problem is discontinuous.

59. (9.3). The maximum principle has a unique solution. The variational inequality is equivalent to the maximum principle. The problem has only one solution. Changing the type of extremum changes the general properties of the problem.

60. (9.4). The maximum principle gives a unique solution. The variational inequality is not equivalent to the maximum principle. The problem has a unique solution.

61. (10.1). The problem is solved using the decoupling method. The problem has a unique solution.

62. (10.2). The problem has a geometric meaning. The problem has a unique solution. An approximate solution to the problem is found using the penalty method.

63. (10.3). The problem is vector and has a physical meaning. The problem has a unique solution.

64. (11.1). Optimality conditions have an infinite set of solutions, moreover, two with one and two discontinuities and one with each larger number of discontinuity points. Optimality conditions are not sufficient. The problem has a unique solution, although the optimality criterion is non-convex.

65. (11.2). The maximum principle has a unique solution that is singular. The problem is Tikhonov ill-posed.

66. (11.3). The maximum principle has three solutions and is an insufficient condition for optimality. The problem has two solutions.

67. (11.4). There is an infinite and not even countable set of stationary points of the function H that are not optimal. The problem has no solution.

68. (11.5). The maximum principle has an infinite set of solutions, of which one is singular, and all solutions except the singular one are discontinuous. The optimality condition is not sufficient. Singular control is not optimal. The problem has two solutions.

69. (11.6). The maximum principle has three solutions that are singular. One singular control satisfies the Kelley condition, but is not optimal. The problem has no solution.

70. (11.7). The maximum principle has an infinite set of solutions and is a sufficient condition for optimality. The problem has an infinite and not even countable set of solutions.

71. (11.8). The maximum principle has three solutions that are singular. The Kelley condition is satisfied for all singular controls in degenerate form. Two singular controls satisfy the Kopp–Moyer condition, but one does not. The problem has two solutions.

72. (11.9). The maximum principle has three solutions that are singular. The Kelley condition is satisfied for all singular controls in degenerate form. Two singular controls do not satisfy the Kopp–Moyer condition, but one does. The problem has a unique solution.

73. (12.1). The maximum principle has a unique solution. The problem has a unique solution and is Tikhonov well-posed.

74. (12.2). The maximum principle has a unique solution. The problem has a unique solution and is Hadamard well-posed.

75. (12.3). The maximum principle has a unique solution that is singular. The problem has a unique solution and is Tikhonov ill-posed.

76. (12.4). The maximum principle has a unique solution that is singular. The problem has a unique solution and is Hadamard ill-posed.

77. (12.5). The problem has a unique solution and is Hadamard ill-posed. There is a bifurcation of extremals. For some values of the problem parameter, the maximum principle is a necessary and sufficient condition for optimality, but for others it is not.

78. (12.6). The problem is Hadamard ill-posed. There is a bifurcation of extremals with an infinite set of bifurcation points.

79. (12.7). There is an infinite set of solutions to optimality conditions. The problem has two solutions that are discontinuous.

PART IV. OPTIMAL CONTROL PROBLEMS WITH ISOPERIMETRIC CONDITIONS

80. (13.1). The problem with a free final state has a unique solution and is Tikhonov well-posed. The problem is solved using the maximum principle, penalty method and variational inequality.

81. (13.2). Problem with a fixed final state has a unique solution and is to well-posed Tikhonov.

82. (13.3). A problem with geometric meaning. The problem has a unique solution.

83. (14.1). A problem with a fixed final state and a non-convex set of admissible controls. The system of optimality conditions is linear. The maximum principle has an infinite set of solutions. Solutions of optimality conditions form an orthonormal family. The problem has two solutions.

84. (14.2). The set of admissible controls is not convex. The system of optimality conditions is nonlinear. Optimality conditions have an infinite set of solutions. There is a solution to the problem, but it is not unique.

85. (14.3). Problem with a free final state. The set of admissible controls is not convex. The system of optimality conditions is linear. The maximum principle has an infinite set of solutions. The problem has two solutions.

86. (14.4). The problem has a continuum of solutions that are not singular controls. All optimal controls are discontinuous. Optimality conditions are necessary and sufficient.

87. (15.1). Problem with a fixed final state. The set of admissible controls is not convex. The minimized functional is non-convex. The problem has no solution.

88. (15.2). Problem with a fixed final state. The set of admissible controls is not convex. The minimized functional is convex. The problem has no solution.

89. (15.3). System with a free final state. The set of admissible controls is not convex. There is an infinite set of solutions to optimality conditions. The problem has no solution. Solutions to the optimality conditions form a minimizing sequence and can be chosen as weak approximate solutions to a given problem.

90. (15.4). System with a free final state. The solution of the optimality conditions is a singular control. The singular control satisfies Kelley condition and is optimal. The problem has a unique solution, but is not Tikhonov well-posed.

91. (15.5). System with a free final state. There exists a unique singular control that satisfies the Kelley condition, but is not optimal.

92. (15.6). System with a free final state. There are two singular controls on the boundary of the set of admissible controls. First of them satisfies the Kelley condition, but the second does not.

93. (15.7). System with a free final state. The problem has a unique solution, but is Hadamard ill-posed.

94. (15.8). The problem has an infinite and not even countable set of solutions. All solutions of the maximum principle are singular controls. The Kelley condition is satisfied for all singular controls.

95. (15.9). There is an infinite and not even countable set of singular controls. The Kelley condition is satisfied for all singular controls, but all singular controls are not optimal.

96. (15.10). There is an infinite and not even countable set of singular controls. All admissible controls are singular. The Kelley condition is satisfied for all singular controls. The problem has an infinite and not even countable set of solutions.

97. (15.11). System with a fixed final state. There a unique control that satisfies Kelley condition.

98. (15.12). System with a fixed final state. There is a unique optimal control, but the problem Tikhonov ill-posed.

99. (15.13). System with a fixed final state. There is a unique optimal control, but the problem is Hadamard ill-posed.

100. (15.14). The problem has a point of extremal bifurcation.

101. (15.15). The problem has an infinite set of points of extremal bifurcation.

102. (15.16). A problem with control included in the optimality criterion and a state included in the isoperimetric condition. The problem has a unique solution. Bellman principle does not valid.

103. (5.17). A problem with control included in the isoperimetric condition and the state included in the optimality criterion. The problem has a unique solution. Bellman principle is not valid.

PART V. OPTIMAL CONTROL PROBLEMS WITH A FREE INITIAL STATE

104. (16.1). A problem without control restrictions has a unique solution. The maximum principle is a necessary and sufficient condition for optimality.

105. (16.2). A problem with control constraints has a unique solution. The maximum principle is a necessary and sufficient condition for optimality. An approximate solution is found using the penalty method.

106. (16.3). Problem with a singular state equation of. An approximate solution is found using the penalty method.

107. (17.1). The optimal control problem has two optimal pairs with the same control. The maximum principle is a necessary but not sufficient condition for optimality.

108. (17.2). The optimal control problem has an infinite set of optimal pairs with the same control. The maximum principle is a necessary but not sufficient condition for optimality. There is an infinite set of solutions to optimality conditions that are not optimal.

109. (17.3). The problem has a unique solution, which is singular. The maximum principle is a necessary and sufficient condition for optimality. The problem is Tikhonov ill-posed.

110. (17.4). The maximum principle has three singular solutions, one of which does not satisfy the Kelley condition. The optimal control problem has two optimal pairs with the same control. The maximum principle is a necessary but not sufficient condition for optimality.

111. (17.5). The problem has no solution. A weak approximate solution to the problem is determined.

112. (17.6). The problem has a unique solution, but is Hadamard ill-posed.

113. (17.7). The optimal control problem has two optimal pairs that differ in both control and state. The maximum principle is a necessary but not sufficient condition for optimality.

114. (17.8). The problem has a unique solution that is singular, and the corresponding pair was rejected in the process of searching for regular solutions of the maximum principle.

115. (17.9). The problem has an infinite set of singular controls, all of which satisfy the Kelley condition and its generalizations. All of them are not optimal, although when adding an initial condition, they become optimal. The problem has no solution.

116. (17.10). Problem with an isoperimetric condition with respect to control. The problem has only one solution. The conditions for maximum optimality are necessary and sufficient.

117. (17.11). Problem with an isoperimetric condition relative to the state. The problem has a unique solution. The maximum principle is a necessary and sufficient optimality conditions.

Bibliography

[1] R.A. Adams and J.J.F. Fournier. *Sobolev Spaces.* Academic Press, 2003.

[2] K.L. Ahmed and N.U. Teo. On the Optimal Control Systems Governed by Quasilinear Integro Partial Differential Equations of Parabolic Type. *J. Math. Anal. Appl.*, 59(1):33–59, 1977.

[3] N.U. Ahmed. On the Maximum Principle for Time-Optimal Controls for a Class of Distributed-Boundary Control Problems. *J. Optim. Theory Appl.*, 45(1):147–157, 1985.

[4] M.A. Aizerman. *Theory of Automatic Control.* Elsevier, 1963.

[5] V.M. Alekseev, V.M. Tikhomirov, and S.V. Fomin. *Optimal control.* Nauka, Moscow, 1979.

[6] V.V. Alsevich. Necessary Optimality Conditions for Minimax Optimization Problems. *Differ. Equat.*, 12(8):1384–1391, 1976.

[7] T.S. Angell. The Controllability Problem for Nonlinear Volterra Systems. *J. Optim. Theory Appl.*, 41(1):9–35, 1983.

[8] M. Aoki. *Optimization of Stochastic Systems.* Academic Press, New York – London, 1967.

[9] J.-L. Armand. *Application of the theory of optimal control of distributed-parameter system to structural optimization. Ph. D. Thesis.* Stanford univ. Stanford, 1971.

[10] V.I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations.* Springer, 1988.

[11] V.I. Arnold, A.N. Varchenko, and S.M. Gusein-Zade. *Singularities of Differentiable Mappings.* MTSNMO, Moscow, 2009.

[12] L.T. Ashchepkov and O.V. Vasiliev. On the Optimality of Singular Controls in Goursat–Darboux Systems. *Zh. Calc. Math. and Mat. Phys.*, 15(5):1157–1167, 1975.

[13] M. Athans and P.L. Falb. *Optimal Control: An Introduction to the Theory and its Applications.* Dover Publ., 2006.

[14] K.A. Atkinson. *An Introduction to Numerical Analysis.* New York, John Wiley and Sons, 1989.

[15] J.P. Aubin and I. Ekeland. *Applied Nonlinear Analysis.* John Wiley and Sons, New York, etc., 1984.

[16] V.I. Averbukh and O.G. Smolyanov. Different Definitions of Derivatives in Linear Topological Spaces. *Uspekhi Matem. Nauk*, 23(4):67–116, 1968.

[17] N.S. Bakhvalov, N.P. Zhidkov, and G.G. Kobelkov. *Numerical Methods.* Basic Knowledge Laboratory, Moscow, 2000.

[18] K. Balachandran and J.P. Dauer. Controllability of Nonlinear Systems in Banach Spaces: A Survey. *J. Optim. Theory Appl.*, 115(2):7–28, 2002.

[19] A.V. Balakrishnan. *Introduction to Optimization Theory in a Hilbert Space.* Springer-Verlag, 1974.

[20] E.J. Balder. An existence problem for the optimal control of certain nonlinear integral equation of Urysohn type. *J. Optim. Theory Appl.*, 42(3):447–463, 1984.

[21] J. Baranger and R. Temam. Non Convex Optimal Problems Depending on a Parameter. *SIAM J. Contr.*, 13:146–152, 1975.

[22] V. Barbu. Optimal Control of Variational Inequalities. In *Research Notes in Mathematics*, volume 100. Pitman, Boston – London – Melbourne, 1984.

[23] V. Barbu. The Dynamic Programming Equation for the Time-Optimal Control Problem in Infinite Dimensions. *SIAM J. Contr. Optim.*, 29(2):445–456, 1991.

[24] V. Barbu. Optimal Control Approach to Nonlinear Diffusion Equations Driven by Wiener Noise. *J. Optim. Theory Appl.*, 153(1):1–26, 2012.

[25] V. Barbu and G. Wang. Internal Stabilization of Semilinear Parabolic Systems. *J. Math. Anal. Appl.*, 285(2):387–407, 2003.

[26] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming. Theory and Algorithms.* Wiley, New York, 1993.

[27] R. Bellman. *Stability Theory of Differential Equations.* Dover Publications Inc., Mineola, 1953.

[28] R. Bellman. *Dynamic Programming.* Princeton University Press, 1957.

[29] R. Bellman and R. Calaba. *Dynamic programming and modern control theory.* Academic Press, 1966.

[30] E.A. Bender. *An Introduction to Mathematical Modeling.* New York, Dover, 2000.

[31] A. Bensoussan and J.L. Lions. Sur les Temps d'Arrêt Optimaux et les Inéquations Variationnelles d'évolution. *C. R. Acad. Sci. Paris*, 280:A989–A992, 1975.

[32] L.D. Berkovitz. A penalty function proof of the maximum principle. *Appl. Math. Optimiz.*, 2(4):291–303, 1976.

[33] D.P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods.* Academic Press, New York, 1982.

[34] D.P. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, 2005.

[35] J. Betts. *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming.* SIAM Press, Philadelphia, PA, 2010.

[36] M.F. Bidaut. Existence Theorems for Usual and Approximate Solutions of Optimizations Problems. *J. Optim. Theory and Appl.*, 15(4):397–411, 1975.

[37] G. Bliss. *Calculus of Variations.* Open Court Pub. Co., Illinois, 1944.

[38] V.G. Boltyansky. *Optimal Control of Discrete Systems.* Nauka, Moscow, 1973.

[39] I.M. Bomze, T. Csendes, R. Horst, and P.M. Pardalos (eds). *Developments in Global Optimization.* Kluwer, Dordrecht, 1997.

[40] Y.G. Borisovich, B.D. Gel'man, A.D. Myshkis, and V.V. Obukhovskii. Multivalued mappings . *J. Math. Sci.*, 24:719–791, 1984.

[41] J.P. Boyd. *Solving Transcendental Equations: The Chebyshev Polynomial Proxy and Other Numerical Rootfinders, Perturbation Series, and Oracles.* Society for Industrial and Applied Mathematics, New York, 2014.

[42] A. Bryson and Ho Yu-Chi. *Applied Optimal Control: Optimization, Estimation, and Control.* Routledge, 1975.

[43] A.L. Bukhgeim. *Introduction to the Theory of Inverse Problems.* De Gruyter, Utrecht, Boston, 1999.

[44] L. Cesari. *Optimization. Theory and Applications. Problems with Ordinary Differential Equations.* Springer Verlag, New York, 1983.

[45] Y.K. Chang and W.T. Li. Controllability of Second-Order Differential and Integro-Differential Inclusions in Banach Spaces. *J. Optim. Theory Appl.*, 129(1):77–87, 2006.

[46] F.L. Chernous'ko and V.V. Kolmanovsky. Computational and Approximate Methods of Optimization. In *Math. Analysis. Results of Science and Technology*, volume 14, pages 101–166. VINITI, Moscow, 1977.

[47] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983.

[48] F.H. Clarke. *Methods of Dynamic and Nonsmooth Optimization.* SIAM, Philodelphia, 1989.

[49] J. Céa. *Optimisation. Théorie et algorithmes.* Dunod, Paris, 1971.

[50] D. Cowles. An existence theorem for optimal problem involving integral equation. *SIAM J. Control*, 11(4):595–606, 1973.

[51] G.P. Crespi, M. Papalia, and M. Rocca. Extended Well-Posedness of Quasiconvex Vector Optimization Problems. *J. Optim. Theory Appl.*, 141(2):285–297, 2009.

[52] B.M. Darinsky, Yu.I. Sapronov, and S.L. Tsarev. Bifurcation of extremals of Fredholm Functionals. *Modern Mathematics*, 12:3–140, 2004.

[53] C. De La Vega. Necessary Conditions for Optimal Terminal Time Control Problems Governed by a Volterra Integral Equation. *J. Optim. Theory Appl.*, 130(1):79–93, 2006.

[54] V.F. Demyanov, F. Giannessi, and V.V. Karelin. Optimal Control Problems via Exact Penalty Functions. *J. Global Optim.*, 12(3):215–223, 1998.

[55] V.F. Demyanov and A.M. Rubinov. *Constructive Nonsmooth Analysis.* Lang, Frankfurt am Main, 1995.

[56] A.V. Dmitruk, A.A. Milutin, and N.P. Osmolovsky. Lusternik Theorem and Extremum Theory. *Uspehi Math. Nauk*, 36(6):11–46, 1980.

[57] G. Duvaut and J.L. Lions. *Inequalities in Mechanics and Physics.* Springer Verlag, Berlin, Heidelberg, New York, 1976.

[58] C. Edwards and S. Spurgeon. *Sliding Mode Control. Theory and Applications.* London, Taylor and Francis, 1998.

[59] A.I. Egorov. *Fundamentals of Control Theory.* Fizmatlit, Moscow, 2001.

[60] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems.* Nord Holland, 1976.

[61] L.E. Elsgolts. *Differential Equations and the Calculus of Variations.* University Press of the Pacific, 2003.

[62] L. Evans. *An Introduction to Mathematical Optimal Control Theory.* University of California, Berkeley, 2010.

[63] S. Farlow. *Partial Differential Equations for Scientists and Engineers.* Dover, New York, 1993.

[64] H.O. Fattorini. Some Remarks on Complete Controllability. *SIAM J. Control.*, (5):391–402, 1966.

[65] R.P. Fedorenko. *Approximate Solution of Optimal Control Problems.* Nauka, Moscow, 1978.

[66] A.A. Feldbaum and A.G. Butkovsky. *Methods of the Theory of Automatic Control.* Moscow, Nauka, 1975.

[67] W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control.* Springer, 1975.

[68] R. Fletcher. *Practical Optimization Methods.* John Wiley and Son, Chichester, 1987.

[69] C.A. Floudas. *Deterministic Global Optimization: Theory, Algorithms and Applications.* Kluwer Academic Publishers, 2000.

[70] C.A. Floudas and P. Pardalos (Eds.). *Encyclopedia of Optimization.* Springer, 2008.

[71] A. Frölicher and W. Bucher. *Calculus in Vector Spaces Without Norm.* Springer-Verlag Berlin etc., 1966.

[72] A.V. Fursikov. Lagrange Principle for Problems of Optimal Control of Ill-Posed or Singular Distributed Systems. *J. Math. Pure Appl.*, 71:139–195, 1992.

[73] A.V. Fursikov. *Optimal control of distributed systems. Theory and applications.* Amer. Math. Soc. Providence, 1999.

[74] R. Gabasov and F.M. Kirillova. *Maximum Principle in the Optimal Control Theory.* Science and Technology, Minsk, 1974.

[75] R. Gabasov and F.M. Kirillova. *Singular Optimal Controls.* Nauka, Moscow, 2018.

[76] R.V. Gamkrelidze. Time-Optimal Processes with Limited Phase Coordinates. *DAN USSR*, 125(3):475–478, 1959.

[77] B. Gelbaum and J. Olmsted. *Counterexamples in Analysis.* Holden Day, 1964.

[78] B. Gelbaum and J. Olmsted. *Theorems and Counterexamples in Mathematics.* Springer, 2012.

[79] P.E. Gill, W. Murrey, and M.H. Wright. *Practical Optimization.* Academic Press London, 1981.

[80] V.Y. Glizer. Novel Controllability Conditions for a Class of Singularly-Perturbed Systems with Small State Delays. *J. Optim. Theory Appl.*, 137(1):135–156, 2008.

[81] A. Granas, R. Guenther, and J. Lee. *Nonlinear Boundary Value Problems for Ordinary Differential Equations.* Państwowe Wydawn. Nauk, Warszawa, 1985.

[82] W.A. Gruver and E. Sachs. *Algorithmic Methods in Optimal Control.* Pitman Res. Math. 47, Pitman, London, 1981.

[83] O. Hajek. *Dynamical Systems in the Plane.* Academic Press, 1968.

[84] G. Hall and J.M. Watt (Eds). *Modern Numerical Methods for Ordinary Differential Equations.* Oxford University Press, 1976.

[85] M.A. Hanson. A Generalization of the Kuhn-Tucker Sufficiency Conditions. *J. Math. Anal. Appl.*, 184(1):146–155, 1994.

[86] P. Hartman. *Ordinary Differential Equations.* SIAM, Providence, 2002.

[87] U.G. Haussmann and J.P. Lepeltier. On the Existence of Optimal Controls. *SIAM J. Contr. Optim.*, 28(4):851–902, 1990.

[88] D. Henry. *Geometric Theory of Semilinear Parabolic Equations.* Springer-Verlag Berlin etc., 1981.

[89] J. Henry. Controllability of Some Nonlinear Parabolic Equations. *Lect. Not. Contr. Inf. Sci.*, (1):287–297, 1978.

[90] H. Hermes and J.P. La Salle. *Functional Analysis and Time Optimal Control.* Academic Press, New York, London, 1969.

[91] J.-B. Hiriart-Urruty and C. Lemarechal. *Fundamentals of Convex Analysis.* Berlin, Springer-Verlag, 2001.

[92] R. Horst and P.M. Pardalos (Eds). *Handbook of Global Optimization.* Kluwer, Dordrecht, 1995.

[93] D.G. Hull. *Optimal Control Theory for Applications.* Springer, New York, 2003.

[94] V. Hutson, J.S. Pym, and M.J. Cloud. *Applications of Functional Analysis and Operator Theory.* Elsevier Science, 2005.

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

[96] G. Iooss and D. Joseph. *Elementary Stability and Bifurcation Theory.* Springer-Verlag, New York, Heidelberg, Berlin, 1980.

[97] R. Isaacs. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization.* Courier Corporation, 1999.

[98] A. Isidori. *Nonlinear Control Systems.* Springer, 1989.

[99] S.I. Kabanikhin. *Inverse and Ill-posed Problems: Theory and Applications.* De Gruyter, 2012.

[100] L.V. Kantorovich and G.P. Akilov. *Functional Analysis.* Nauka, Moscow, 1977.

[101] J. Kelley. *General topology.* D. Van Nostrand Company, Inc., Toronto, etc., 1957.

[102] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and their Applications.* Academic Press, New York, 1980.

[103] D.E. Kirk. *Optimal Control Theory: An Introduction.* New Jersey, Englewood Cliffs, 2004.

[104] G. Knowles. Finite Element Approximation of Parabolic Time Optimal Control Problems. *SIAM J. Contr. Optim.*, 20(3):414–427, 1982.

[105] T. Kobayashi. Some Remarks on Controllability for Distributed Parameter Systems. *SIAM J. Control*, 16(15):733–742, 1978.

[106] A.N. Kolmogorov and S.V. Fomin. *Elements of Function Theory and Functional Analysis.* Dover Publ., 1999.

[107] M.A. Krasnosel'skii. *Topological Methods in the Theory of Nonlinear Integral Equations.* Pergamon, 1964.

[108] V.F. Krotov. *Global Methods in Optimal Control Theory.* Marcel Dekker, New York, 1996.

[109] A.G. Kusraev and S.S. Kutateladze. *Subdifferential Calculus: Theory and Applications.* Nauka, Moscow, 2007.

[110] L.D. Landau and E.M. Lifshits. *Mechanics.* Elsevier, 1982.

[111] I. Lasiecka. Ritz–Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions. *SIAM J. Contr. Optim.*, 22(3):477–500, 1984.

[112] J. Lee. *A First Course in Combinatorial Optimization.* Cambridge University Press, 2004.

[113] W.S. Levine (ed.). *The Control Handbook.* New York, CRC Press, 1996.

[114] M.B. Lignola and J. Morgan. Topological Existence and Stability for Min Sup Problems. *J. Math. Anal. Appl.*, 151(1):164–180, 1990.

[115] M.B. Lignola and J. Morgan. Well-posedness for Optimization Problems with Constraints Defined by Variational Inequalities Having a Unique Solution. *J. Global Optim.*, 16(1):57–67, 2000.

[116] J.L. Lions. *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles.* Dunod, Gauthier-Villars, Paris, 1968.

[117] J.L. Lions. *Quelques méthods de résolution des problèmes aux limites non linéaires.* Dunod, Gauthier-Villars, Paris, 1969.

[118] J.L. Lions. *Contrôle de systèmes distribués singuliers.* Gauthier-Villars, Paris, 1983.

[119] V.G. Litvinov. *Optimization in Elliptic Boundary Value Problems.* Nauka, Moscow, 1987.

[120] A. Lorenzi. *An introduction to identification problems via functional analysis.* De Gruyter, Utrecht – Boston, 2001.

[121] S. Luan. Nonexistence and Existence of an Optimal Control Problem Governed by a Class of Semilinear Elliptic Equations. *J. Optim. Theory Appl.*, 158(1):1–10, 2013.

[122] R. Lucchetti and F. Patrone. Hadamard and Tyhonov well-posedness of a certain class of convex functions. *J. Math. Anal. Appl.*, 88:204–215, 1982.

[123] D.G. Luenberger. *Optimization by Vector Space Methods.* Wiley, 1997.

[124] K.A. Lurie. *Optimal Control in Mathematical Physics.* Nauka, Moscow, 1975.

[125] L.A. Lusternik. On Conditional Extremum of Functionals. *Math. Sbornik*, (3):390–401, 1934.

[126] E.N. Mahmudov. Sufficient Conditions for Optimality for Differential Inclusions of Parabolic Type and Duality. *J. Global Optim.*, 41(1):31–42, 2008.

[127] M. Majewski. Stability Analysis of an Optimal Control Problem for a Hyperbolic Equation. *J. Optim. Theory Appl.*, 141(1):127–146, 2009.

[128] K. Malanowski. Stability Analysis for Nonlinear Optimal Control Problems Subject to State Constraints. *SIAM J. Optim.*, 18(3):926–945, 2007.

[129] H. Maurer and N.P. Osmolovskii. Second Order Sufficient Conditions for Time-Optimal Bang-Bang Control. *SIAM. J. Control Optim.*, 42(6):2239–2263, 2004.

[130] S.G. Mikhlin. *Linear Partial Differential Equations.* Vysshaya Shkola, Moscow, 1977.

[131] J. Milnor. *Morse theory.* Princeton University Press, 1963.

[132] M. Mitchell. *An Introduction to Genetic Algorithms.* Cambridge, MA: MIT Press, 1996.

[133] M. Mäkelä and P. Neittaanmäki. *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control.* WSPC, 1992.

[134] J. Morgan and V. Scalzo. Discontinuous but Well-Posed Optimization Problems. *SIAM J. Optim.*, 17(3):861–870, 2006.

[135] S.F. Morozov and V.I. Sumin. On Speed Problems in the Theory of Optimal Control of Transport Processes. *Differ. Equ.*, 11(4):727–740, 1976.

[136] F. Murat. Thèoremes de non existènce pour des problèmes de contrôle dans les coefficients. *C. R. Acad. Sci. Paris*, (5):A395—A398, 1972.

[137] F. Murat. Contre-examples pour divers problèmes ou le contrôle intervient dans les coefficients. *Ann. di Math. Pura et Appl.*, 112:46–68, 1977.

[138] P. Neittaanmäki and D. Tiba. On the finite element approximation of the boundary control for two-phase Stefan problem. *Lect. Not. Contr. Inf. Sci.*, 62:356–370, 1984.

[139] J.A. Nelder and Mead R. A Simplex Method for Function Minimization. *Computer Journal*, 7(4):308–313, 1965.

[140] L.W. Neustadt. *Optimization*. Princeton University Press, Princeton, 1976.

[141] J. Nocedal and S.J. Wright. *Numerical Optimization*. Berlin, Springer, 2006.

[142] C. Olech. Existence theory and optimal control theory. *Topol. Funct. Anal., Vienna*, 1:291–328, 1976.

[143] B.G. Pachpatte. *Inequalities for Differential and Integral Equations*. Academic Press, San Diego, 1998.

[144] J.Jr. Palis and W. de Melo. *Geometric Theory of Dynamical Systems*. Springer, 2012.

[145] L.-H. Peng, C. Li, and J.-C. Yao. Well-posedness of a Class of Perturbed Optimization Problems in Banach Spaces. *J. Math. Anal. Appl.*, 346(2):384–394, 2008.

[146] D. Peterson and J.H. Zalkind. A Review of Direct Sufficient Conditions in Optimal Control Theory. *Int. J. Control*, 28(4):589–610, 1978.

[147] J.D. Pintér. *Global Optimization in Action, Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications*. Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.

[148] V.I. Plotnikov. Necessary and Sufficient Optimality Conditions and Uniqueness Conditions for Optimizing Functions for Control Systems of General Form. *Izv. USSR Acad. Sci., ser. math.*, 26(3):652–679, 1972.

[149] E. Polak. *Computational Methods in Optimization: A Unified Approach*. Academic Press, New York, 1975.

[150] J. Polderman and J. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach*. Springer Verlag, New York, 1998.

[151] A. Polyanin and A. Manzhirov. *Handbook of Integral Equations*. CRC Press, Boca Raton, 1998.

[152] L.S. Pontryagin, V.G. Boltyansky, R.V. Gamkrelidze, and E.F. Mishchenko. *Mathematical Theory of Optimal Processes.* Nauka, Moscow, 1974.

[153] M.M. Postnikov. *Introduction to Morse Theory.* Nauka, Moscow, 1971.

[154] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes. The Art of Scientific Computing.* Cambridge University Press, 2007.

[155] U. Raitums. *Optimal Control Problems for Elliptic Equations.* Zinatne, Riga, 1989.

[156] J.P. Raymond. Existence and Uniqueness Results for Minimization Problems with Nonconvex Functionals. *J. Optim. Theory Appl.*, 82(3):571–592, 1994.

[157] J.P. Raymond and H. Zidani. Pontryagin's Principle for Time-Optimal Problems. *J. Optim. Theory Appl.*, 101(2):375–402, 1999.

[158] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. I. Functional Analysis.* Academic Press, New York, London, 1972.

[159] W.T. Reid. *Riccati Differential Equations.* New York, Academic Press, 1972.

[160] H.M. Robbins. A Generalized Legendre–Clebsch Condition for the Singular Cases of Optimal Control. *IBM J. Res. Develop.*, 11(4):361–372, 1967.

[161] R.T. Rockafellar. *Convex Analysis.* Princeton, NJ: Princeton University Press, 1997.

[162] I. Ross. *A Primer on Pontryagin's Principle in Optimal Control.* San Francisco, Collegiate Publishers, 2015.

[163] J. Rotman. *An Introduction to the Theory of Groups.* New York: Springer-Verlag, 1994.

[164] T. Roubiček. Optimal Control of Nonlinear Fredholm Integral Equation. *J. Optim. Theory Appl.*, 97(3):707–729, 1998.

[165] T. Seidman and H.X. Zhou. Existence and Uniqueness of Optimal Control for a Quasilinear Parabolic Equations. *SIAM J. Contr. Optim.*, 20(6):747–762, 1982.

[166] S. Serovajsky. Lower Completion and Extension of Extremum Problems. *Izv. vuzov. Math.*, (5):30–41, 2003.

[167] S. Serovajsky. Sequential extension in the coefficients control problems for elliptic type equations. *J. Inverse Ill-Posed Probl.*, 11(5):523–536, 2003.

[168] S. Serovajsky. Optimal Control for Singular Equation with Nonsmooth Nonlinearity. *J. Inverse Ill-Posed Probl.*, 14(6):621–631, 2006.

[169] S. Serovajsky. Optimal Control of a Singular Evolutional Equation with Nonsmooth Operator and Fixed Terminal State. *Differ. Equ.*, 43(2):259–266, 2007.

[170] S. Serovajsky. *Counterexamples in the Optimal Control Theory.* Walter de Gruyter, 2011.

[171] S. Serovajsky. *Optimization and Differentiation.* CRC Press, Taylor & Francis Group, 2017.

[172] S. Serovajsky. *Sequential Models of Mathematical Physics.* CRC Press, 2019.

[173] S. Serovajsky. *Mathematical Modelling.* Chapman and Hall/CRC, 2021.

[174] S. Serovajsky and I. Adylova. Numerical Solution of Coefficient Inverse Problems with Corrective Approximation. *Inzhenerno-fizicheskij zhurnal*, (5):858–859, 1988.

[175] S. Serovajsky, M. Sigalovsky, and A. Azimov. Non-smooth Optimization Methods in the Geometric Inverse Gravimetry Problem. *Adv. Math. Models Appl.*, 7(1):5–15, 2022.

[176] C. Simionescu. Optimal Control Problem in Sobolev Spaces with Weight. *SIAM J. Control Optim.*, 14(1):137–143, 1976.

[177] T.K. Sirazetdinov. *Optimization of Systems with Distributed Parameters.* Nauka, Moscow, 1977.

[178] T. Slawing. Shape Optimization for Semi-linear Elliptic Equations Based on a Embedding Domain Method. *Appl. Math. Optim.*, 49(2):183–199, 2004.

[179] M. Sniedovich. *Dynamic Programming: Foundations and Principles.* Taylor and Francis, 2010.

[180] J.A. Snyman and D.N. Wilke. *Practical Mathematical Optimization: Basic Optimization Theory and Gradient-Based Algorithms.* Springer, Berlin, 2018.

[181] S.L. Sobolev. *Some Applications of Functional Analysis in Mathematical Physics.* Amer. Math. Soc., 1963.

[182] E.D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems.* Springer, 1998.

[183] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer-Verlag, New York, 1980.

[184] S. Strogatz. *Nonlinear Dynamics and Chaos.* Perseus Publishing, 2000.

[185] D. Tiba. Optimal Control of Nonsmooth Distributed Parameter Systems. In *Lecture Notes in Mathematics*, volume 1459. WJ Springer-Verlag, Berlin, 1990.

[186] D. Tiba. New Results in Space Optimization. *Math. Repts.*, 5(4):389–398, 2003.

[187] A.N. Tikhonov and V.Ya. Arsenin. *Solutions of Ill-Posed Problems.* Halsted Press, 1977.

[188] S.P. Timoshenko and J.M. Gere. *Theory of Ellastic Stability.* McGraw-Hill, 2017.

[189] J.F. Traub. *Iterative Methods for the Solution of Equations.* American Mathematical Soc., 1982.

[190] A.I. Tretyak. On Necessary Conditions for Optimality of Odd Order in a Time-Optimal Problem for Systems That Are Linear in Control. *Math. Sbornik*, 181(5):625–641, 1990.

[191] V. Utkin. *Sliding Modes in Control and Optimization.* Springer-Verlag, 1992.

[192] L. van der Waerden. *Algebra. Vol. I.* Springer, 2003.

[193] F.P. Vasilyev. *Optimization Methods. Vol. 1.* MTSNMO, Moscow, 2011.

[194] F.P. Vasilyev. *Optimization Methods. Vol. 2.* MTSNMO, Moscow, 2011.

[195] F.P. Vasilyev, M.M. Potapov, B.A. Budak, and L.A. Artemyeva. *Optimization Methods.* Yurayt, Moscow, 2016.

[196] A.A. Voronov. *Stability, Controllability, Observability.* Nauka, Moscow, 1979.

[197] S. Walczak. Well-posed and Ill-posed Optimal Control Problems. *J. Optim. Theory Appl.*, 109(1):245–262, 2001.

[198] G. Wang, Y. Zhao, and W. Li. Some Optimal Control Problems Governed by Elliptic Variational Inequalities with Control and State Constraint on the Boundary. *J. Optim. Theory Appl.*, 106(3):627–655, 2000.

[199] L.W. Wang. Approximate Controllability for Integrodifferential Equations with Multiple Delays. *J. Optim. Theory Appl.*, 143(1):185–206, 2009.

[200] J. Warga. *Optimal Control of Differential and Functional Equations.* Academic Press, New York, 1972.

[201] M.M. Weinberg and V.A. Trenogin. *Theory of Branching of Solutions of Nonlinear Equations.* Nauka, Moscow, 1969.

[202] Z.Y. Wu. Sufficient Global Optimality Conditions for Weakly Convex Minimization Problems. *J. Global Optim.*, 39(3):427–440, 2007.

[203] A.I. Yablonsky. *Mathematical Models in the Study of Science.* Nauka, Moscow, 1986.

[204] Y. Ye and Q. Chen. Optimal Control of the Obstacle in a Quasilinear Elliptic Variational Inequality. *J. Math. Anal. Appl.*, 294(1):258–272, 2004.

[205] B.P. Yeo. A Modified Quasilinearization Algorithm for the Computation of Optimal Singular Control. *Int. J. Control*, 32(4):723–730, 1980.

[206] A. Yezza. Relaxed Optimal Control Problems Governed by Integral Equations. *J. Math. Anal. Appl.*, 175(1):126–142, 1993.

[207] K. Yokoyama. ε-Optimality Criteria for Vector Minimization Problems via Exact Penalty Functions. *J. Math. Anal. Appl.*, 187(1):296–305, 1994.

[208] L. Young. *Lectures on the Calculus of Variations and Optimal Control Theory.* W.B. Saunders Co., Philadelphia, etc., 1969.

[209] A.J. Zaslavski. Solutions for a Class of Optimal Control Problems with Time Delay, Part 2. *J. Optim. Theory Appl.*, 91(2):455–490, 1996.

[210] M.I. Zelikin. *Homogeneous Spaces and the Riccati Equation in the Calculus of Variations.* Berlin, Springer-Verlag, 2000.

[211] T. Zolezzi. A Characterizations of Well-posed Optimal Control Systems. *SIAM J. Control Optim.*, 19(5):604–616, 1981.

[212] V.I. Zubov. *Lectures on Control Theory.* Lan, SPb, 2009.

Index