CLASSICS OF SOVIET MATHEMATICS

L. S. PONTRYAGIN SELECTED WORKS

IN FOUR VOLUMES

VOLUME FOUR

The Mathematical Theory of Optimal Processes

CRC Press
Taylor & Francis Group

L. S. PONTRYAGIN
SELECTED WORKS

Volume 4

The Mathematical Theory of Optimal Processes

Classics of Soviet Mathematics

L. S. PONTRYAGIN SELECTED WORKS

Edited by R. V. Gamkrelidze

Volume 1: **Selected Research Papers**
Volume 2: **Topological Groups**
Volume 3: **Algebraic and Differential Topology**
Volume 4: **The Mathematical Theory of Optimal Processes**

ISSN 0743–9199
This book is part of a series. The publishers will accept continuation orders which may be cancelled at any time and which provide for automatic billing and shipping of each title in the series upon publication. Please write for details.

L. S. PONTRYAGIN SELECTED WORKS

Volume 4

The Mathematical Theory of Optimal Processes

L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko

Translated from the Russian
by K. N. Trirogoff
Aerospace Corporation,
El Segundo, California

English Edition Edited by
L. W. Neustadt
Aerospace Corporation
El Segundo, California

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Gordon and Breach Science Publishers

P.O. Box 786
Cooper Station
New York, NY 10276
United States of America

P.O. Box 197
London WC2E 9PX
England

58, rue Lhomond
75005 Paris
France

14-9 Okubo 3-chome
Shinjuku-ku, Tokyo 160
Japan

Originally published in Russian as Математическая теория оптимальных процессов by Izdatel'stvo Nauka, Moscow, 1961.
First published in English by Interscience Publishers, a division of John Wiley & Sons, Inc.

Third Printing, September 1965.
Reprinted, by arrangement, by Gordon and Breach Science Publishers S.A., 1986.
Reprinted from a copy in the collection of the Brooklyn Public Library.

Library of Congress Cataloging-in-Publication Data

Pontryagin, L. S. (Lev Semenovich), 1908–
The mathematical theory of optimal processes.

(L. S. Pontryagin selected works ; v. 4) (Classics of Soviet mathematics, ISSN 0743–9199)
Translation of: Matematicheskaia teoriia optimal'nykh protsessov.
Reprint. Originally published: New York : Interscience Publishers, 1962. With new introd.
Bibliography: p.
Includes index.
1. Mathematical optimization. I. Neustadt, Lucien W. II. Title. III. Title: Optimal processes.
IV. Series: Pontryagin. L. S. (Lev Semenovich), 1908– . Selections. Polyglot. 1985 ; v. 4.
V. Series: Classics of Soviet mathematics.
QA3.P76 1985 vol. 4 [QA402.5] 510 s [519] 86–4732

ISBN 2–88124–077–1 (Switzerland)

Volume 4: ISBN 2–88124–077–1; 4-volume set: ISBN 2–88124–134–4.

Printed in Great Britain by Bell and Bain Ltd., Glasgow.

Lev Semenovich Pontryagin

Contents

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

Editor's Preface

On 3 September 1983 Lev Semenovich Pontryagin was seventy-five. To mark this important event in the life of this outstanding contemporary mathematician we are beginning the publication of his scientific works in four volumes, according to a decision taken by the Mathematics Division of the USSR Academy of Sciences. The first volume contains the most important mathematical papers of L. S. Pontryagin and also includes a bibliography of his basic scientific works, the second is his well-known monograph *Topological Groups*, the third comprises two monographs, *Foundations of Algebraic Topology* and *Smooth Manifolds and Their Applications in Homotopy Theory*, and the fourth is a revised edition of *The Mathematical Theory of Optimal Processes* by L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko.

The scientific activity of Lev Semenovich Pontryagin has left a deep imprint on many crucial areas of modern mathematics, both pure and applied. His work has had a definitive influence on the development of topology and topological algebra, and because of him optimal control theory is one of the topical trends in present-day applied mathematics. In a brief review we can neither delve deeply into his important works nor describe the profound, multifaceted impact of these works on the advancement of the respective fields. This is, therefore, only a broad outline that may be of help in studying his works.

While still a second-year student at Moscow State University, Pontryagin began his scientific activity under the guidance of P. S. Aleksandrov. In this early period, his interests concentrated mainly on two central topics of algebraic (combinatorial) topology, namely, topological duality theorems and dimension theory, which Pontryagin looked upon as a local variant of duality theory.

The discovery of "Pontryagin duality," the culmination of his work in topological duality theorems, and the construction of the general theory of characters of a locally compact commutative group are Pontryagin's two most notable contributions and are undoubtedly among the finest achievements in modern mathematics.

We begin with a survey of his main works in duality theory and topological algebra. To assess the full value of the advances made by Pontryagin in this area, it is apt to recall here that at the time when Pontryagin had just begun his activity, homology groups were hardly used in topology; instead, Betti numbers with respect to different moduli and torsion coefficients were used, and the Alexander duality theorem was formulated as an equality of Betti numbers (modulo 2) of dimensions $n - r - 1$ and r of a polyhedron

$K \subset R^n$ and its complement $R^n \backslash K$,

$$p^r(R^n \backslash K) = p^{n-r-1}(K).$$

In his first published paper,[1*] Pontryagin improved this theorem by extending the duality between the Betti numbers of a polyhedron and its complement in R^n to the duality between the r- and $(n - r - 1)$-dimensional homology groups (modulo 2) of the polyhedra $(R^n \backslash K)$ and K. The full statement of this theorem follows. In $R^n \backslash K$ and K, two bases

$$z_1^r, \dots z_s^r \quad \text{and} \quad \zeta_1^{n-r-1}, \dots \zeta_s^{n-r-1}$$

of homology (mod 2) of dimensions r and $(n - r - 1)$, respectively, can be chosen, such that the square matrix of linking coefficients (mod 2)

$$\| (z_i^r, \zeta_j^{n-r-1}) \| (i, j = 1, \dots s)$$

is the identity matrix.

Thus, the duality between the homology groups (mod 2) established here by means of the linking coefficients led to a group isomorphism.

The next paper[2] deals with the same problem in modulus 2, but the polyhedron K is now imbedded in an arbitrary closed n-dimensional manifold M^n. Its solution demanded, probably for the first time in the history of topology, a study of homological properties of continuous mappings. That is, Pontryagin was led to the study of kernels and images of homomorphisms of homology groups (mod 2) for the inclusions $K \subset M^n$ and $M^n \backslash K \subset M^n$, and the duality theorem was formulated in terms of the ranks of the corresponding kernels. Later, the study of the homological properties of mappings acquired immense significance in topology and greatly influenced the creation of homological algebra.

This paper also contained a statement, known subsequently as the "Pontryagin cycle removal theorem," that asserted: If an r-dimensional cycle Z^r in M^r intersects every $(n - r)$-dimensional cycle in K with a zero intersection index, then the cycle Z^r can be "homologically removed" from K, i.e., there exists in $M^n \backslash K$ an r-dimensional cycle that is homologous to Z^r in M^n. This theorem found successful applications in the topological theory of variational problems; Pontryagin himself used it in estimating the category of a manifold.

From the foregoing it is clear how far one of the central problems of algebraic topology of the late twenties had been advanced in two short papers of a 19-year old sophomore.

---

* Reference numbers refer to the bibliography of Pontryagin's publications, pp. 609–618.

The next work concerning duality theorems, his master's thesis,[6] was stimulated by the course in algebra given by E. Noether. It gives a profound analysis of the algebraic nature of topological duality theorems. Duality for an arbitrary modulus $m > 0$ obtained a final solution in the form of an isomorphism of the corresponding groups, in view of the fact, now well understood, that a finite cyclic group is the Pontryagin dual of itself (a concept which Pontryagin had not yet arrived at that time).

A particular corollary of the results of the paper is that, for any $m > 0$, the homology groups (mod m), $H_r^{(m)}(R^n \backslash K)$ and $H_{n-r-1}^{(m)}(K)$ of dimensions r and $n - r - 1$, respectively, are isomorphic, and, consequently, all homology groups (mod m) of the complement $R^n \backslash K$ are invariant, i.e., they depend only on the homology groups of the polyhedron K, but do not depend on the inclusion of K in R^n.

Duality theorems for full homology groups with integral coefficients cannot be formulated in terms of isomorphisms and, therefore, could not be fitted into the framework of the paper. For instance, the full r-dimensional integral homology group $H_r(R^n \backslash K)$ is neither isomorphic to the group $H_{n-r-1}(K)$ nor even determined by it. There exist only isomorphisms (also noted in the paper) separately between the r- and $(n - r - 1)$-dimensional weak homology groups and between the r- and $(n - r - 2)$-dimensional torsion groups of the sets K and $R^n \backslash K$, obviously implying the invariance of the full integral homology groups of the complement $(R^n \backslash K)$.

If, instead of a finite polyhedron K, an arbitrary compact set F is considered in R^n, then the corresponding integral and weak homology groups are, in general, no longer finitely generated, and a special investigation is needed to establish the invariance of the homology groups of the complement $R^n \backslash F$. Pontryagin also studied the duality for an arbitrary compact set $F \subset R^n$ and established the invariance of the groups $H_r^{(m)}(R^n \backslash F)$, $m > 0$, as well as the invariance of weak homology groups of $R^n \backslash F$, thereby significantly advancing the problem.

But the central question of the independence of the full group of integral homology $H_r(R^n \backslash F)$ of the inclusion of the compact set $F \subset R^n$ still remained unsolved. Its solution demanded the introduction of a new homological invariant of the set F, namely, a homology group related not to a discrete but to a compact coefficient group. This permitted him, while rejecting the narrow concept of duality as an isomorphism, to define "Pontryagin duality." In 1931–32, he made this decisive step and completely solved all problems relating to duality and also the longstanding problem of the proper definition of homology groups of compact metric spaces.

In constructing the homology group $H_r(F)$ of the set F, the coefficients are not taken from a discrete group of residues (modulo m) or from the group of integers, but are taken from a compact topological group of rotations of a circle. The group $H_r(F)$ is, in itself, a compact commutative topological group. The group $H_r(F)$ and the $(n - r - 1)$-dimensional integral homology group $H_{n-r-1}(R^n \backslash F)$ proved to be Pontryagin duals, i.e., each is the character group of the other (for a detailed exposition of the theory of characters, see reference 110 or the second volume of the *Selected Works*).

Generally, let Γ, G be a dual group pair, i.e., each is the character group of the other, and let Γ be compact and G discrete. Take Γ as the coefficient group for constructing the homology group $H_r^{\Gamma}(F)$. Then its dual (i.e., its character group) is the homology group of the complement $H_{n-r-1}^{G}(R^n \backslash F)$, which is constructed, using G, the dual of Γ, as the coefficient group. Duality is realized through linking coefficients.

The general duality theorem for a closed set $F \subset R^n$ was first reported as a short communication in the Proceedings of the International Mathematics Congress held in Zurich in 1932, while its full exposition is given in reference 18.

This paper actually marks the end of Pontryagin's research into topological duality theorems. These theorems, being a powerful tool for studying general homological problems in topology, resolved the crucial question in algebraic topology of the thirties. Particularly after Pontryagin's duality theorems, homology groups have gained a firm foothold in topology as the basic homological invariants in place of the Betti numbers and torsion coefficients, which had fully served the purpose of homology groups until the main circle of topological problems led to finitely generated groups.

Topological duality theorems for a (finite) polyhedron in an arbitrary closed n-dimensional manifold are given in their final formulation in reference 54.

A logical continuation of the duality theorems is the general theory of characters of locally compact commutative groups created by Pontryagin. The main result of this theory is the assertion that every compact commutative group is the character group of some discrete group. Its proof rests on the construction of the invariant measure introduced by Haar in 1933, which had played a key role in the development of topological algebra.

The general theory of characters had enabled Pontryagin to elucidate the structure of compact and locally compact groups, the results obtained for compact and locally compact commutative groups being final. A positive answer to Hilbert's fifth problem for a compact and locally compact commutative group follows directly from these results. (For a detailed exposition of the structure of compact and locally compact commutative groups, refer to the third edition of *Topological Groups*, Volume 2 of the *Selected Works*.) However, the significance of the theory of characters of locally compact topological groups does not end here. Its creation has indeed laid the foundation of topological algebra as an independent discipline, which has been primarily responsible for the development of general harmonic analysis on topological groups. Pontryagin's works in duality theory and character theory had a deep impact on algebraic-topological reasoning in the thirties and, in particular, made a great contribution to "functorial thinking" in mathematics.

His first publications on the general theory of characters of commutative topological groups, on the structure of compact groups, and on locally compact commutative groups are references 16, 17, and 19, respectively.

His remarkable theorem (see reference 10) that asserts that the field of real numbers, the field of complex numbers, and the division ring of quaternions are the only locally compact connected division rings should also be classified under topological algebra.

The methods developed here were later fully utilized by Pontryagin in elucidating the structure of locally compact commutative groups with the help of the theory of characters, as we have already pointed out.

The outcome of his studies in topological algebra was the famous monograph *Topological Groups*, first published in 1938, which has had several editions both in the USSR and in many other countries, in most of the major European languages. It became a classic that influenced many generations of mathematicians and that has not lost its value even today, forty-five years since its first publication, a rare event in mathematics. Its third English edition forms the second volume of the *Selected Works* of L. S. Pontryagin.

The early works of Pontryagin also deal with dimension theory. He constructed examples of compact metric spaces that have different dimensions in different moduli. He later used these examples (see reference 4) to construct the famous "dimensionally deficient" continuums, which disproved the longstanding hypothesis that the dimension of compact sets is additive under topological multiplication. He found two two-dimensional compact sets whose product is of dimension three, instead of four. His theorem that any n-dimensional compact set is homeomorphically mapped into R^{2n+1} (see reference 7) also fits into the category of dimension theory.

The homological dimension theory due to P. S. Aleksandrov owes much to Pontryagin's work in dimension theory. For Pontryagin himself, his studies in dimension theory had a far-reaching consequence — under their influence he began, in the mid-thirties, a systematic investigation of homotopic problems in topology.

His studies in homotopic topology likewise reached their climax (at the beginning of the forties) in the discovery of methods that basically paved the way for a new field in modern mathematics, differential topology. Here we have in mind his discovery of characteristic classes and his contributions to the theory of fiber bundles.

Prior to taking up the "homotopic period," mention should be made of his outstanding topological paper written in 1935,[21] a full exposition of which is given in reference 21. It gives the solution to the Cartan problem of calculating the homology groups of compact group manifolds for the four main series of compact Lie groups. Historically, in this paper, the homological invariants were first found for a large and extremely important class of manifolds defined, not by triangulation, but by analytical (in this case, by algebraic) relations. To solve this problem, Pontryagin used, instead of Cartan's method based on the algebra of exterior invariant forms on a group (R. Brauer applied this method later), Morse's method of defining a smooth function on a manifold with isolated critical points and constructing trajectories orthogonal to level surfaces of the function. He refined this method further — the critical points were no longer "isolated," but formed "critical manifolds."

The methods developed in this paper were fruitfully used by H. Hopf and others to advance further the topology of group manifolds and homogeneous spaces, and later by Pontryagin himself to solve certain auxiliary problems in homotopy theory, and, in particular, to calculate the homology groups of Grassmann manifolds.

A direct consequence of this work is an elegant result obtained by Pontryagin many years later.[39] The point is that, for all compact simple Lie groups, the Betti numbers are equal to the corresponding Betti numbers of the direct products of spheres of different dimensions. The question therefore naturally arose: is a compact simple Lie group homeomorphic to the product of spheres of appropriate dimensions? Through the use of homotopic techniques, he found the answer to be negative. The special unitary group of third-order matrices has the same Betti numbers as the product of a 3-dimensional sphere and a 5-dimensional sphere, but the group itself is not homeomorphic to the product of the spheres: this was established through the use of the classification of the mappings of S^4 into S^3.

We shall now outline the homotopic works of L. S. Pontryagin. The topical problem in homotopic topology in the early stages of its development centered around the homotopic classification of the mappings of a sphere into a sphere of lesser dimension. Pontryagin encountered this problem while making fruitless attempts at giving a local characterization of the dimension of a compact set in R^n in terms of the homological characteristics of its complement.

In the beginning, he tried to solve the homotopic classification problem of the mappings of the sphere S^{n+k} into S^n using homological methods. But, shortly after learning about Hopf's work on the classes of mappings of S^3 onto S^2, he came to fully appreciate the situation; that was the beginning of a fifteen-year period during which Pontryagin was completely engaged in homotopic topology.

First, he demonstrated that the Hopf invariant is unique and, consequently, that Hopf's construction gives all the classes of the mappings of S^3 into S^2; thus, he obtained the full classification of the mappings of S^3 into S^2. Soon after, in 1936, he discovered an amazing result: the number of classes of mappings of S^{n+1} into S^n, for $n \geqslant 3$, is two (see reference 28). A mistake was made, however, in classifying the mappings of S^{n+2} into S^n, which led to an erroneous result. It was noticed and corrected by Pontryagin in 1950 (see reference 63). For these mappings, too, the number of classes was found to be two.

The initial proofs of these theorems were incredibly cumbersome. Only later, after the discovery of the method of framed manifolds (see below), could they be greatly simplified.

Then followed the solution to a series of problems in the homotopic classification of mappings of polyhedra into spheres and vice versa. Of these papers we mention here only two, reference 40 and 43. These papers introduced such basic concepts in homotopy theory as "obstructions" and "difference cochains" and a new cohomological operation — the Pontryagin square, the predecessor of Steenrod's cohomological operations.

But the major problem, the classification of the mappings of S^{n+k} into S^n for $k \geqslant 3$, still defied solution. This is exactly the problem that led Pontryagin to discover the so-called "framed manifold method," to define new invariants of smooth manifolds — characteristic classes known as "Pontryagin classes," and to create the theory of fiber bundles, i.e., to create a new and very important field in modern mathematics, differential topology.

Among the pioneers in this field, besides L. S. Pontryagin, we should name H. Hopf, E. Stiefel, H. Whitney, and C. S. Chern.

The framed manifold technique was designed to study the homotopic properties of mappings with the help of the information available about the differential-topological structure of a manifold. It was only fruitful in classifying the mappings of S^{n+k} into S^n for $k \leqslant 3$ (as had already been noted at the beginning of the fifties by Pontryagin for $k = 1, 2$, and by Rokhlin for $k = 3$), because, for $k > 3$, information was needed about smooth manifolds of dimensions > 3, which could not be obtained by the methods available in the early fifties. However, the framed manifold technique is equally effective for the opposite purpose, studying smooth manifolds when we have homotopic information at our disposal, which can be more successfully derived with the help of Leray's algebraic (spectral sequence) method. This reversal of the method, known as bordism theory, is due to R. Thom. Most of the far-reaching results in the modern theory of smooth manifolds have been obtained precisely through a combination of the Pontryagin–Thom differential-topological method and Leray's algebraic method.

Today, characteristic classes constitute the central topic not only in differential topology, but also in modern differential geometry as a whole; fiber bundle theory has long since become a common research tool in topology, geometry, and analysis.

The theory of characteristic classes and the closely related theory of singularities of vector fields are presented in three large papers.[56,57,61] The results of these papers were reported in earlier preliminary works.[45,48–50] Reference 49 also reports briefly on the theory of classifying spaces, which subsequently played an important role in the development of fiber bundle theory.

The framed manifold method and a full classification of the mappings of S^{n+k} into S^n for $k = 0$, 1, and 2 are presented in reference 69 (see also Volume 3 of the *Selected Works*), which was the original exposition in the literature of the fundamentals of differential topology.

The "topological period" in the activity of L. S. Pontryagin ends with reference 69; from the early fifties on, he switched over exclusively to applied fields. Up until this time he had turned his attention to applied and nontopological topics only occasionally, but with great success.

We begin the survey of his earlier nontopological works with the famous paper written in collaboration with A. A. Andronov,[29] in which the concept of the structural stability of a dynamical system in a plane was first introduced, using the term "rough system," and the roughness condition was formulated.

In a broad context there are two motives behind the idea of roughness: physical and mathematical. The physical motive arose in connection with Andronov's investigations into auto-oscillations and consists of the following: if a dynamic system describing a physical phenomenon is known only approximately, then the qualitative portrait of the system's phase plane can reflect the phenomenon only if this portrait does not change under small perturbations of the dynamic system. The mathematical motive is related to the idea of "typicality," or "general position," which is not at all specific to differential equations and which is widely used in different fields of mathematics, including some topological works of L. S. Pontryagin. For the "general position" case, the phase portrait should be expected to be simpler than in exceptional cases; thus, the "general position" case deserves the utmost attention.

In this paper, smooth flow (of class C^1) in a domain $O \subset R^2$ bounded by a smooth closed curve everywhere transversal to the trajectories is called rough, if, for any flow sufficiently C^1-close to the initial flow, there exists a homeomorphism of the domain O onto itself, C^0 close to the identity, that sends the trajectories of one flow into the trajectories of another, preserving the direction of motion along these trajectories.

After giving this definition, the authors show that the rough systems on a plane are typical (they form an everywhere dense open set) and that their qualitative portrait is quite simple. Here the three ideas, "simplicity", "roughness", and "typicality", merge together (the corresponding classes of the systems coincide). This merger is specific to the small dimension of the phase space and fails for higher dimensions. But these three ideas are themselves of great interest for higher-dimensional systems also, and the questions of the behavior of trajectories for the corresponding class of systems and of the mutual relations between these classes have dominated the study of dynamic systems through the past twenty or twenty-five years, and go back, in the final analysis, to reference 29.

Still earlier, reference 29 had influenced the development of the two-dimensional qualitative theory of differential equations. First, it outlines the role of "singular" (orbitally unstable) trajectories, subdividing the phase plane into "cells" filled with trajectories of identical behavior. Second, the solution of the problem concerning rough systems on a plane paved the way for studies of "typical" bifurcations of a parameter-dependent dynamic system in the two-dimensional case.

Of his early works on dynamic systems, mention should be made of one more paper,[13] which gives simple conditions, conveniently applied, for the birth of a cycle from a closed trajectory of a plane nonlinear Hamiltonian system under small autonomous (nonconservative) perturbations.

Among the early nontopological works of Pontryagin, reference 47 also deserves special mention, and had a considerable impact on the development of functional analysis on spaces with an indefinite metric. It was written during World War II at Kazan in connection with a purely applied problem of stability in ballistics. Its main result is that any Hermitian operator in a Hilbert space with an indefinite metric of index k has a k-dimensional invariant subspace on which all eigenvalues of the operator have nonnegative imaginary parts, and the main (indefinite) form of the space is nonnegative.

One more work completed during wartime at Kazan concerns stability theory. It formulates the conditions that must be fulfilled for a quasipolynomial to have roots with negative real parts (see reference 42). These conditions were later extended to functions of the type f/g having no poles, where f is a quasi-polynomial and g a polynomial (see reference 66).

We shall now take up the period that dates approximately from the beginning of the fifties, when Pontryagin was basically devoting himself to problems in applied mathematics.

Here, too, he displays with great strength his exceptional talent to perceive amidst the primal chaos in each new problem the main path, which leads to the goal via the shortest route. He forges ahead on this pathway, overcoming technical difficulties that seem, at times, to be insurmountable.

To study new topics, Pontryagin founded a special seminar in oscillation and control theory in 1952 at the Steklov Mathematics Institute. He believed that, to gain success in any applied field of mathematics, one should not confine oneself to the existing mathematical models, but start the study with technical problems, not only to gain a deeper insight into the existing models, but also to formulate new mathematical problems that have a pure mathematical interest as well as a technical interest.

Soon, as a result of this seminar, two basic advances emerged: the theory of relaxation (discontinuous) oscillations and the optimal control theory, which later Pontryagin began to elaborate on with great success jointly with his younger collaborators V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko.

Relaxation oscillations are encountered in physical, and, in particular, in radio engineering systems described by differential equations with a small parameter ε attached to higher derivatives. Mathematically, relaxation oscillations can be defined as the periodic solutions of differential equations (or a system of differential equations) with a small parameter attached to higher derivatives that contain "slow motion" sections traversed by a phase point in a finite time, as well as "junction points" where the "fast motion" sections start and which are traversed in infinitely small time as $\varepsilon \to 0$. A classical example of these oscillations is the Van der Pohl equation. The study of the asymptotic behavior of these oscillations in relation to ε is a very difficult mathematical problem and was only partially solved in some simplest cases. Pontryagin's studies have made much headway with this problem for general systems and are of fundamental value.

Of great help to Pontryagin in these investigations was his phenomenal ability to do long mental calculations and to memorize complicated expressions.

Pontryagin's works on relaxation oscillations are listed in that part of the bibliography which comprises papers published in 1955–1963.

In the mid-fifties, he discovered the famous "Pontryagin maximum principle," which, though universal, is easily formulated and is an effective tool in solving a broad range of optimization problems from purely applied questions in diverse engineering fields to complicated theoretical questions. The maximum principle includes the first-order theory of the classical calculus of variations, which had proved futile in tackling many new technical problems, the analysis of which has led to the discovery of the maximum principle.

The maximum principle is simple to formulate and we state it for the important time-optimal case.

A process is called controlled if it can be described by an n-dimensional vector differential equation

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

where $x \in R^n$ is the phase point and u is an r-dimensional vector control parameter that takes values from some given subset $U \subset R^r$, which is, as a rule, a closed domain. The problem then is to choose a control $u(t) \in U$, as a function of time t, such that the corresponding trajectory $x(t)$ of the equation

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

is shifted from a given point x_0 to some other given point x_1 in minimum time. This control and its corresponding trajectory are called optimal. Let us introduce the following scalar function

$$H(x, \psi, u) = \psi f(x, u),$$

where $\psi f(x, u)$ is the scalar product of an n-dimensional vector ψ and f, and write the canonical system of equations

$$\dot{x} = f = \frac{\partial H}{\partial \psi}, \qquad \dot{\psi} = -\psi \frac{\partial f}{\partial x} = -\frac{\partial H}{\partial x}.$$

The Pontryagin maximum principle asserts that, for a control $u(t)$, $t_0 \leqslant t \leqslant t_1$, and the corresponding trajectory $x(t)$ to be optimal, it is necessary that there exist a nonzero variable vector $\psi(t)$ such that $u(t)$, $x(t)$, and $\psi(t)$ satisfy the above canonical system of equations and "Pontryagin's maximum condition":

$$H(x(t), \psi(t), u(t)) = \max_{u \in U} H(x(t), \psi(t), u), \ \forall t \in [t_0, t_1].$$

The discovery of the maximum principle proved a startling event that soon gave birth to a new advance, the optimal control theory, which, at present, is a vital and flourishing area in applied mathematics — and the stream of papers brought forth by this theory is truly immense.

Among the works of Pontryagin that have greatly influenced the development of optimal control theory, we may mention his Plenary Address to the International Congress of Mathematicians held in Edinburgh in 1958[89] and his monograph "The Mathematical Theory of Optimal Processes" written jointly with V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko (see Volume 4 of the *Selected Works*).

A natural development of optimal control theory proposed by Pontryagin himself is differential game theory, which he is presently pursuing. A review of this theory is outlined in his Plenary Address to the International Congress of Mathematicians held at Nice in 1970.[126] A full exposition of his theory of linear differential games is given in references 127 and 143, which are also included in this volume.

Since 1934, L. S. Pontryagin has been working at the Steklov Mathematics Institute of the USSR Academy of Sciences; he was made a full-time member of the Institute and given the position of Head of the Topology Division in 1939. From 1961 to the present, he has held the position of Head of the Division of the Theory of Ordinary Differential Equations and Control Theory. At the same time, he has always attached great importance to the teaching of mathematics and has devoted much time to giving lectures at Moscow State University. Being an excellent teacher, he always prepared his lectures with utmost care, even designing notation to the minutest detail. Four of his books, *Topological Groups* (Volume 2 of the *Selected Works*), *Combinatorial Topology*, *Algebraic and Differential Topology* (Volume 3 of the *Selected Works*), and *Ordinary Differential Equations* (English edition[100]), which have been translated into many languages, were based on his lecture courses at Moscow State University; they have greatly influenced the education of many generations of mathematicians all over the world.

R. V. GAMKRELIDZE

Preface to the English Translation

This monograph makes available the powerful results in optimal control theory obtained by the group of mathematicians led by Academician Pontryagin at the Steklov Mathematical Institute in Moscow. Until now, the material described in this book has been available only in Russian mathematical periodicals and in scattered English translations.

It is hoped that the publication of this English edition will stimulate research in this budding field of applied mathematics.

The translation has been as faithful as is possible within the constraints of good English usage and in keeping with the differences in the technical terminologies in English and Russian. Further, the redundancy so common in Russian mathematical writing was oftentimes reduced so as to confirm with the more succinct style of western mathematics. The references to works in Russian were changed to corresponding ones in English whenever possible. As noted, some references were added in the present edition; there is no pretence of completeness in this list.

Typographical errors (where found) were corrected. Less obvious corrections made by the editor were always described in translators' footnotes. Most of these were discussed with Professors Gamkrelidze and Mishchenko during their visit to the United States in the spring of 1962 and met with their approval.

The following individuals contributed to the English edition: Professor A. V. Balakrishnan of UCLA who made constructive criticisms of the translation of Chapter VII; Professor G. Leitman of the University of California (Berkeley) and L. Berkovitz of the Rand Corporation, who called attention to the material described in the footnote on page 240. Further acknowledgements are due to the following: Misses Stella Allabashi and Geraldine Matlick for expertly typing the manuscript; the Aerospace Corporation for its aid and cooperation in the preparation of the manuscript; and finally, the Interscience Division of John Wiley & Sons, Inc., for the rapidity with which they executed the publication. In fact, the English translation will have appeared within 18 months of the time that the preparation of the Russian manuscript was begun.

L. W. NEUSTADT
K. N. TRIROGOFF
Los Angeles, California
June, 1962

INTRODUCTION

The physical processes which take place in technology are, as a rule, controllable, i.e., they can be realized by various means depending on the will of man. In this connection, there arises the question of finding the very best (in one sense or another) or, as is said, the *optimal* control of the process. For example, one can speak about optimality in the sense of rapidity of action, i.e., about achieving the aim of the process in the shortest time; about achieving this aim with a minimum expenditure of energy, etc. Mathematically formulated, these are problems in the calculus of variations, which in fact owes its origin to these problems. However, the solution of a whole range of variational problems, which are important in contemporary technology, is outside the classical calculus of variations. The solution presented here of a considerable number of such non-classical variational problems is due to the collective authors of this book. In its essential features, this solution is unified in one general mathematical method, which we call the maximum principle. It should be noted that all the fundamental necessary conditions in the classical calculus of variations (with ordinary derivatives) follow from the maximum principle (see Chapter V).

Here we shall consider control processes which can be described by a system of ordinary differential equations:

$$\frac{dx^i}{dt} = f^i(x^1, \cdots, x^n, u^1, \cdots, u^r), \qquad i = 1, 2, \cdots, n; \tag{1}$$

where $x^1, \cdots, x^n$ are the variables which characterize the process, i.e., the phase coordinates of the controlled object which define its state at each instant of time t; $u^1, \cdots, u^r$ are the control parameters which determine the course of the process; and t is time. In order to determine the course of the control process (1) in a certain time interval $t_0 \leq t \leq t_1$, it is sufficient to give the control parameters $u^1, \cdots, u^r$:

$$u^j = u^j(t), \qquad j = 1, \cdots, r, \tag{2}$$

as functions of time on this time interval.

Then, for the given initial values

$$x^i(t_0) = x_0^i, \qquad i = 1, \cdots, n, \tag{3}$$

the solution of system (1) is uniquely determined. The variational problem to be solved, which is related to the control process (1), consists of the following. We consider the integral functional

$$J = \int_{t_0}^{t_1} f^0(x^1, \cdots, x^n, u^1, \cdots, u^r)\, dt, \tag{4}$$

where $f^0(x^1, \cdots, x^n, u^1, \cdots, u^r)$ is a given function. For each control (2), given on a certain interval $t_0 \leq t \leq t_1$, the course of the control process is uniquely determined, and the integral (3) takes on a definite value. Let us assume that there exists a control (2) which transfers the controlled object from a given initial phase state (3) to a prescribed terminal phase state

$$x^i(t_1) = x_1^i, \qquad i = 1, \cdots, n. \tag{5}$$

It is required to find a control

$$\bar{u}^j(t), \qquad j = 1, \cdots, r, \tag{6}$$

which realizes the transfer of the controlled object from state (3) to state (5) in such a manner that the functional (4) has a minimum value. The times t_0 and t_1, in this statement of the problem, are not fixed. We only require that the object should be in state (3) at the initial time, and in state (5) at the final time, and that the functional (4) should achieve a minimum. (The case where the times t_0 and t_1 are fixed is also of interest; this case easily reduces to problems referred to in this introduction. See §8.) In the special case where the function $f^0(x^1, \cdots, x^n, u^1, \cdots, u^r)$, which defines the functional (4), is identically equal to unity, the functional (4) has the value $t_1 - t_0$, and our variational problem becomes the time-optimal problem.

In technical problems where the control parameters $u^1, \cdots, u^r$, for example, determine the position of a machine's controllers, the u^j cannot assume arbitrary values, but are subject to certain restrictions. Because of the arrangement of the mechanism described by system (1), the parameter u^1 can, let us say, only assume values which satisfy the condition

$$|u^1| \leq 1. \tag{7}$$

Or, for example, if the parameters u^1 and u^2 characterize a vector in a plane, where the length of the vector does not exceed unity but its direction is arbitrary, these parameters are subject to the condition

$$(u^1)^2 + (u^2)^2 \leq 1. \tag{8}$$

Generally, it is necessary to assume that the point $(u^1, \cdots, u^r)$ must belong to a certain set U of the space with coordinates $u^1, \cdots, u^r$. Moreover, the choice of this set U reflects specific features of the object (1). In the mathematical statement of the problem the set U (the "control region") is considered as arbitrary; but in technical problems the case where U is a closed set [compare with inequalities (7) and (8)] is particularly important and characteristic. This condition signifies that the extreme positions of the controller [the values $u^1 = \pm 1$ in inequality (7), or boundary points of the circle (8)], which may, in particular, yield an optimal control, are also admissible. In fact, this circumstance makes the variational problem under consideration non-classical, since, in the classical calculus of variations, the variable parameters cannot satisfy inequalities of the type (7) or (8), when these also include equalities.

The non-classical character of our variational problem is brought out particularly clearly by the time-optimal problem for system (1), where the right-hand sides are linear functions of the variables $x^1, \cdots, x^n, u^1, \cdots, u^r$ with constant coefficients, and where the set U is a closed convex polyhedron. For example, U may be the *cube* defined by the inequalities:

$$|u^j| \leq 1, \qquad j = 1, \cdots, r.$$

In this case it turns out that the optimal control (6) is realized by the point $(u^1(t), \cdots, u^r(t))$ which is located, in turns, at various vertices of U. The rules according to which the control point jumps from one vertex to another also give the optimal control law. This linear variational problem, which has important practical applications, is solved in Chapter III on the basis of general methods. Classical methods are completely inapplicable to the solution of such a problem.

From what has been said about the jumps of the optimal control point from vertex to vertex of U, it follows that the class of admissible controls (2) cannot be considered to consist of continuous functions. We usually suppose that it consists of *piecewise continuous* functions. The phase coordinates $x^1, \cdots, x^n$ are assumed to be continuous and piecewise differentiable functions of time. Under these assumptions the necessary conditions for optimality are formulated in the form of a *maximum principle* (see Chapter I), which is proved in Chapter II.

If the object under consideration is a mechanical system, the part $x^1, \cdots, x^k$ of the phase coordinates describes the geometrical state of the system, and the part $x^{k+1}, \cdots, x^{2k}(2k = n)$ describes its velocity. In certain problems the goal of the control process may not be to have the object get to a definite point $(x_1^1, \cdots, x_1^n)$ in phase space, but to have the mechanical system attain a definite spatial position $(x_1^1, \cdots, x_1^k)$, with arbitrary velocities. Thus, we have here the variational problem of an optimal transition of the object from a definite initial point $x_0^1, \cdots, x_0^n$ in phase space, to an arbitrary point on the k-dimensional plane defined by the equations

$$x^1 = x_1^1, \cdots, x^k = x_1^k.$$

We can see that the optimal problem formulated earlier does not include a number of important problems. Because of this fact, the problem of optimally transferring an object from an initial manifold M_0 of points in phase space, to a terminal manifold M_1, where the dimensions of M_0 and M_1 are arbitrary (in particular, when they are both zero we obtain the first described problem), is examined in §6 of Chapter I.

It is quite clear that because of the very character of the technical problem, not only the object's control parameters, but also its phase coordinates must sometimes be subject to certain restrictions. For example, if we discuss the motion of an aircraft, and if x^1 denotes its altitude above the ground, the inequality $x^1 \geq h > 0$, where h is the minimum allowed flight altitude, must be satisfied. The inequality $x^1 \geq h$ follows neither from the properties of the system of eqs. (1), nor from the inequalities imposed on the control parameters; but is completely independent. The problem of optimally controlling an object, when the point in phase space which represents it must remain in a certain closed region G of the space at all times, is solved in Chapter VI. It is therein assumed that G has a piecewise smooth boundary.

Under these conditions, the object's motion takes place partially in the interior of G, where it is subject to the usual maximum principle, and partially along the boundary of G, where it is subject to a complicated form of the maximum principle. The transitions from trajectory segments which lie in the interior of G to segments which lie on the boundary of G are subject to peculiar rules which recall the laws of the refraction of light and which, in a certain sense, generalize them.

Up to now we have spoken about an optimal control which brought the object to a given point or onto a given sub-manifold of the phase space. However, the optimal control problem may consist of optimally getting to a moving point in phase space. Let us assume that there exists a moving point

$$x^i = \theta^i(t), \qquad i = 1, \cdots, n, \tag{9}$$

in phase space. Then, there arises the problem of optimally bringing the object (1) in coincidence with the moving point (9). This problem is easily reduced to the one considered above. It is sufficient to introduce new variables by setting

$$y^i = x^i - \theta^i(t), \qquad i = 1, \cdots, n.$$

As a result of this transformation, the control system (1) becomes a new system, which, it is true, is no longer autonomous; and the goal of the control process becomes that of bringing the new object $(y^1, \cdots, y^n)$ to the stationary point $(0, \cdots, 0)$ in phase space. Since our basic results are easily extended to non-autonomous control processes (see §7) the problem may be considered solved.

We assumed here that the motion of the pursued point (9) is predetermined on the entire time interval under consideration. A completely new problem, which is important in practice, arises when the motion of the pursued object is not known beforehand, and information on its motion is received in the course of time. In order to solve such a problem about a pursued object it is necessary to have some data on its behavior. Extremely important is the case when the pursued object is itself controlled, so that its motion is described by a system of equations

$$\frac{dz^i}{dt} = g^i(z^1, \cdots, z^n, v^1, \cdots, v^s), \qquad i = 1, \cdots, n, \tag{10}$$

and when the motion takes place in the same phase space as the motion of the pursuing object (1). The problem consists of the following: knowing the technical capabilities of the pursued object [i.e., the system of eqs. (10)] and its position at a given instant of time, determine the control for the pursuing object at the same instant of time in such a way that the pursuit is carried out in an optimal fashion. In this formulation the problem is still unsolved. In Chapter VII another pursuit problem is solved. It is assumed that the position of the pursued object is known at the initial time, and that its subsequent behavior is described in a probabilistic manner; namely, the process of its motion is assumed to be Markovian. Under these assumptions we seek a control for the pursuing object (1) for which an encounter of a small neighborhood of the object (1) with the pursued object is the most probable.

Initially we tried to find the optimal control (6) for fixed initial (3) and terminal (5) positions of the object. However, it is often necessary to find not only the optimal control (6), but the general solution of the problem with arbitrary positions (3) and (5). We shall assume, for the sake of definiteness, that the object's terminal position (5) is fixed, but that its initial position (3) is an arbitrary point in the space. Then, the desired optimal control (6) becomes a function not only of time, but also of the initial point

$$x_0 = (x_0^1, \cdots, x_0^n),$$

so that we have the optimal control

$$\bar{u}^1(t, x_0), \cdots, \bar{u}^r(t, x_0). \tag{11}$$

Let us set

$$\bar{u}^j(t_0, x_0) = \bar{u}^j(x_0).$$

If $x(t)$ is the position of the controlled object at the time t with the control (11), the obvious identity

$$\bar{u}^j(t, x_0) = \bar{u}^j(x(t)), \qquad j = 1, \cdots, r,$$

holds, expressing the fact that at each time t one must control the object in an optimal manner. Therefore, instead of the functions (11) of $n + 1$ variables we may consider the functions

$$\bar{u}^1(x), \cdots, \bar{u}^r(x) \tag{12}$$

of n variables. These functions yield the so-called *synthesis* of the optimal control. The question of the very existence of a synthesizing control (12) is rather complicated, but has a positive answer for linear systems under certain additional assumptions of an extremely general character (see Chapter III). The synthesizing control (12) is also constructed for some specific linear systems (see §5 and §20).

Starting from the assumption that the synthesizing control (12) does exist, and that the corresponding functional (4), which is now a function of the point x:

$$J = J(x) = J(x^1, \cdots, x^n) \tag{13}$$

is a continuously differentiable function of the variables $x^1, \cdots, x^n$, the American mathematician R. Bellman constructed a partial differential equation for the functional (13). This equation of Bellman's gives rise to another approach to the solution of the optimal control problem (see §9). It is different from the one given in this book, but is closely related to it. It must be noted that the assumption on the continuous differentiability of the functional (13) does not hold in the simplest cases. Thus, Bellman's considerations yield a good heuristic method, rather than a mathematical solution of the problem. The maximum principle, in addition to its complete mathematical validity, also has the advantage that it results in a system of ordinary differential equations, whereas Bellman's approach requires the solution of a partial differential equation.

Finally, a few words should be said about the heretofore unmentioned Chapter IV. Here, some generalizations of the considered fundamental problems, as well as some applications, are compiled. In particular, we consider the case where some additional numerical parameters (which may be chosen prior to beginning the process, but cannot be changed once the process is under way) enter into eq. (1), as well as the case where the object's equation of motion is complicated by the presence of the effect of delay. In addition, the application of the maximum principle to one problem in the theory of the approximation of functions is presented. The solution of one (rather special) pursuit problem is given in the last paragraph of Chapter IV.

CHAPTER I

THE MAXIMUM PRINCIPLE

§1. ADMISSIBLE CONTROLS

We shall consider the behavior of an object whose state at any instant of time is characterized by n real numbers $x^1, x^2, \cdots, x^n$ (for example, these may be coordinates and velocities). The vector space X of the vector variable $x = (x^1, \cdots, x^n)$ is the phase space of the object under consideration. The behavior (motion) of the object (from a mathematical viewpoint) consists of the fact that the variables $x^1, \cdots, x^n$ change with time. It is assumed that the object's motion can be controlled; i.e., that the object is equipped with certain "controllers" on whose position the motion of the object depends. The positions of the "controllers" are characterized by points u, of a certain *control region* U, which may be any set in some r-dimensional Euclidean space E_r. Giving a point

$$u = (u^1, u^2, \cdots, u^r) \in U$$

is equivalent to giving a numerical system of parameters u^1, u^2, $\cdots$, u^r. In applications, the case where U is a closed region in E_r is important. In particular, the control region U may be a *cube* in the r-dimensional space of the variables $u^1, u^2, \cdots, u^r$:

$$|u^j| \leq 1, \qquad j = 1, 2, \cdots, r, \tag{1}$$

or any other closed and bounded set in this r-dimensional space. The physical meaning of considering a closed and bounded control region U (in the space of the variables $u^1, u^2, \cdots, u^r$) is clear. The quantity of fuel being supplied to a motor, temperature, current, voltage, etc., which cannot take on arbitrarily large values, may serve as the control parameters $u^1, u^2, \cdots, u^r$. In addition, because of the physical construction of the object's control portion, relations given by one, or several equations of the form $\phi(u^1, u^2, \cdots, u^r) = 0$, may exist among the control parameters. In this case, U may have a more or less complicated geometric character. If, for example, there are two control parameters u^1 and u^2, which, because of the object's construction, have the form $u^1 = \cos\phi$ and $u^2 = \sin\phi$, where ϕ is an (arbitrarily given) angle, the control region is a circle

$$(u^1)^2 + (u^2)^2 = 1. \tag{2}$$

Henceforth, we shall simply speak of the *control region* U, and of its points $u \in U$. We shall think of U as a certain set in the space of the variables $u^1, u^2, \cdots, u^r$, considering the system of control parameters $u = (u^1, u^2, \cdots, u^r)$, located arbitrarily in U, as its "point" u [see, for example, (1) or (2)].

We shall call every function $u = u(t)$, defined on some time interval $t_0 \leq t \leq t_1$, with range in U, a *control*. Since U is a set in the space of the control parameters $u^1, u^2, \cdots, u^r$, each control

$$u(t) = (u^1(t), u^2(t), \cdots, u^r(t))$$

is a vector function (given for $t_0 \leq t \leq t_1$) whose range is in U. Henceforth, depending on the character of the stated problem, we shall impose various conditions (piecewise continuity, piecewise differentiability, etc.) on $u(t)$. The controls which satisfy these conditions will be called admissible controls. *In this chapter we shall consider the admissible controls to be arbitrary piecewise continuous controls* (with range in U); i.e., controls $u = u(t)$ which are continuous for all t under consideration, with the exception of only a finite number of t, at which $u(t)$ may have discontinuities of the first kind. To avoid any misunderstanding, let us note that, from the definition of discontinuities of the first kind, we assume the existence of the finite limits

$$u(\tau - 0) = \lim_{\substack{t\to\tau \\ t<\tau}} u(t), \qquad u(\tau + 0) = \lim_{\substack{t\to\tau \\ t>\tau}} u(t),$$

at a point of discontinuity, τ. In particular, it therefore follows that every control $u(t)$ is bounded (even if U is not).

The value of a piecewise continuous control $u(t)$ at a point of discontinuity does not play an essential role in what follows; however, for the sake of definiteness, it is convenient to assume that at each point of discontinuity, τ, the value of $u(t)$ is equal to its left-hand limit:

$$u(\tau) = u(\tau - 0), \tag{3}$$

and that each control $u(t)$ under consideration is continuous at the endpoints of the interval $t_0 \leq t \leq t_1$, on which it is given.

Thus, in this chapter, we shall agree to call any piecewise continuous function $u(t)$, $t_0 \leq t \leq t_1$, whose range is in U, which satisfies condition (3) at its points of discontinuity, and which is continuous at the endpoints of the interval $t_0 \leq t \leq t_1$ on which it is given, an *admissible control*. Piecewise continuous controls correspond to the assumption of "inertialess" controllers, since the values of the function $u(t)$ may jump (at an instant of discontinuity) instantaneously from one point of the control region to another. This class of admissible controls seems to be the most interesting for the practical applications of the theory developed here.

§2. STATEMENT OF THE FUNDAMENTAL PROBLEM

We shall assume that the object's law of motion (also the law for the effect of the "controllers" on this motion) can be written in the form of a system of differential equations

$$\frac{dx^i}{dt} = f^i(x^1, x^2, \cdots, x^n, u^1, \cdots, u^r) = f^i(x, u), \quad i = 1, 2, \cdots, n, \tag{4}$$

or in vector form,

$$\frac{dx}{dt} = f(x, u), \tag{5}$$

where $f(x, u)$ is the vector with coordinates

$$f^1(x, u), f^2(x, u), \cdots, f^n(x, u).$$

The functions f^i are defined for $x \in X$ and for $u \in U$. They are assumed to be continuous in the variables $x^1, x^2, \cdots, x^n, u$, and continuously differentiable with respect to $x^1, x^2, \cdots, x^n$. In other words, the functions

$$f^i(x^1, x^2, \cdots, x^n, u) \quad \text{and} \quad \frac{\partial f^i(x^1, x^2, \cdots, x^n, u)}{\partial x^j}; \quad i, j = 1, 2, \cdots, n,$$

are defined and continuous on the direct product $X \times U$.

Let us note that system (4) is *autonomous*, i.e., its right-hand sides do not depend explicitly on the time, t. We shall consider the case wherein the right-hand sides do depend on t below (see §7).

If the control law is given, i.e., if a certain admissible control $u = u(t)$ is chosen, eq. (5) takes the form

$$\frac{dx}{dt} = f(x, u(t)), \tag{6}$$

from which [for any initial conditions $x(t_0) = x_0$] the motion of the object $x = x(t)$ is uniquely determined; i.e., the solution of eq.(6) is defined for a certain time interval. Namely, if $u(t)$ is given for $t_0 \leq t \leq t_1$, and $\theta_1, \theta_2, \cdots, \theta_k$ are its points of discontinuity (of the first kind), where $t_0 < \theta_1 < \theta_2 \cdots < \theta_k < t_1$, we shall first consider eq. (6) on the interval $t_0 \leq t \leq \theta_1$ where its right-hand side is continuous. We shall denote the solution of this equation, with initial condition $x(t_0) = x_0$, by $x(t)$. If this solution is defined on the entire interval $t_0 \leq t \leq \theta_1$, and has the value $x(\theta_1)$ at the point θ_1, we can consider eq. (6) on the interval $\theta_1 \leq t \leq \theta_2$, using $x(\theta_1)$ as the initial value. This solution will also be denoted by $x(t)$. Thus, the constructed solution $x(t)$ is continuous at all points at which it is defined, and, in particular, at the "junction point" θ_1. Now, if $x(t)$ is defined on the entire interval $t_0 \leq t \leq \theta_2$, and has the value $x(\theta_2)$ at θ_2, we can consider eq. (6) on the interval $\theta_2 \leq t \leq \theta_3$, using $x(\theta_2)$ as the initial value; etc. The thus obtained solution $x(t)$ of eq. (6) is continuous and piecewise differentiable; namely, at all points, except $\theta_1, \theta_2, \cdots, \theta_k$, $x(t)$ (where it is defined) is continuously differentiable. The solution $x(t)$ will be called the solution of system (4) [or of eq. (5)], *corresponding* to the control $u(t)$ for the initial condition $x(t_0) = x_0$. This solution may not be defined on the entire interval $t_0 \leq t \leq t_1$ on which $u(t)$ is given (it may run off to infinity).

We shall say that the admissible control $u(t)$, $t_0 \leq t \leq t_1$, *transfers* the phase point from the position x_0 to the position x_1 if the corresponding solution $x(t)$ of eq. (5) [or, what is the same, (6)], satisfying the initial condition $x(t_0) = x_0$, is defined for all t, $t_0 \leq t \leq t_1$, and passes through the point x_1 at the time t_1; i.e., it also satisfies the boundary condition $x(t_1) = x_1$.

Let us now suppose that we are given an additional function $f^0(x^1, x^2, \cdots, x^n, u) = f^0(x, u)$ which is defined and is continuous together with its partial derivatives $\partial f^0/\partial x^i$, $i = 1, 2, \cdots, n$, on all of $X \times U$. Then, the fundamental problem (finding the optimal controls) can be formulated as follows.

In the phase space X, two points x_0 and x_1 are given. Among all the admissible controls $u = u(t)$ which transfer the phase point from the position x_0 to the position x_1 (if such controls exist), find one for which the functional

$$J = \int_{t_0}^{t_1} f^0(x(t), u(t))\, dt \tag{7}$$

takes on the least possible value. Here, $x(t)$ is the solution of eq. (5) with initial condition $x(t_0) = x_0$ corresponding to the control $u(t)$, and t_1 is the time at which this solution passes through x_1.

Let us note that (for fixed x_0 and x_1) the upper and lower limits, t_0 and t_1, in the integral (7) are not fixed numbers, but depend on the choice of the control $u(t)$ which transfers the phase point from x_0 to x_1 (these limits are determined by the relations $x(t_0) = x_0$ and $x(t_1) = x_1$). We shall discuss the solution of the problem for the case of fixed limits below (see §8).

The control $u(t)$ which yields the solution of the problem cited above is called *an optimal control corresponding to a transition from x_0 to x_1*. The corresponding trajectory $x(t)$ is called an *optimal trajectory*. Thus, the fundamental problem consists of finding the optimal controls (and the corresponding optimal trajectories).

An important special case of the above cited optimal problem is the one where $f^0(x, u) \equiv 1$. In this case, the functional (7) takes the form

$$J = t_1 - t_0, \tag{8}$$

and the optimality of the control $u(t)$ signifies *minimality of the transition time from x_0 to x_1*. The problem of finding the optimal controls (and trajectories) in this case will be called the time-optimal problem.

In order to formulate and prove the necessary optimality condition it will be convenient to reformulate our problem. Namely, let us adjoin a new coordinate x^0 to the phase coordinates $x^1, x^2, \cdots, x^n$, which vary according to (4). Let x^0 vary according to the law

$$\frac{dx^0}{dt} = f^0(x^1, x^2, \cdots, x^n, u),$$

where f^0 is the function which appears in the definition of J [see (7)]. In other words, we shall consider the system of differential equations

$$\frac{dx^i}{dt} = f^i(x^1, x^2, \cdots, x^n, u^1, \cdots, u^r) = f^i(x, u), \quad i = 0, 1, 2, \cdots, n, \quad (9)$$

whose right-hand sides do not depend on x^0. Introducing the vector

$$\mathbf{x} = (x^0, x^1, x^2, \cdots, x^n) = (x^0, x)$$

in the $(n + 1)$-dimensional vector space $\mathbf{X}$, we may rewrite system (9) in vector form

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(x, u), \quad (10)$$

where $\mathbf{f}(x, u)$ is the vector in $\mathbf{X}$ with coordinates $f^0(x, u), \cdots, f^n(x, u)$. Note that $\mathbf{f}(x, u)$ does not depend on the coordinate x^0 of the vector $\mathbf{x}$.

Now let $u(t)$ be an admissible control transferring x_0 to x_1, and let $x = x(t)$ be the corresponding solution of eq. (5) with initial condition $x(t_0) = x_0$. Let us denote the point $(0, x_0)$ by $\mathbf{x}_0$; i.e., $\mathbf{x}_0$ is the point of $\mathbf{X}$ whose coordinates are $0, x_0^1, \cdots, x_0^n$, where $x_0^1, \cdots, x_0^n$ are the coordinates of x_0 in X. Then, it is clear that the solution of eq. (10) with initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$, corresponding to the control $u(t)$, is defined on the entire interval $t_0 \leq t \leq t_1$, and has the form

$$x^0 = \int_{t_0}^{t} f^0(x(t'), u(t'))\, dt',$$

$$x = x(t).$$

In particular, when $t = t_1$

$$x^0 = \int_{t_0}^{t_1} f^0(x(t), u(t))\, dt = J, \quad x = x_1,$$

i.e., the solution $\mathbf{x}(t)$ of eq. (10) with initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ passes through the point $\mathbf{x} = (J, x_1)$ at $t = t_1$. In other words, if we let Π be the line in $\mathbf{X}$ passing through the point $\mathbf{x} = (0, x_1)$ and parallel to the x^0 axis (this line is made up of all the points (ξ, x_1) where the number ξ is arbitrary; Fig. 1), we can say that $\mathbf{x}(t)$ passes through a point on Π, with coordinate $x^0 = J$, at the time $t = t_1$. Conversely, suppose that $u(t)$ is an admissible control such that the corresponding solution $\mathbf{x}(t)$ of eq. (10) with initial condition $\mathbf{x}(t_0) = \mathbf{x}_0 = (0, x_0)$, at some time t_1 passes through a point $\mathbf{x}_1 \in \Pi$, with coordinate $x^0 = J$.