UNCLASSIFIED

AD 268 491

Reproduced
by the

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

268 491

268491

CATALOGED BY ASTIA
AS AD No.

MEMORANDUM
RM-2888-PR
DECEMBER 1961

VARIATIONAL METHODS IN PROBLEMS OF CONTROL AND PROGRAMMING

LEONARD D. BERKOVITZ

ASTIA

PREPARED FOR:

UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM

RM-2888-PR

December 1961

VARIATIONAL METHODS IN PROBLEMS OF CONTROL AND PROGRAMMING

Leonard D. Berkovitz

This research is sponsored by the United States Air Force under Project RAND— Contract No. AF 49(638)-700 — monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Technology, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force. Permission to quote from or reproduce portions of this Memorandum must be obtained from The RAND Corporation.

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

PREFACE

The modern control and programming problems studied in this memorandum are a common mathematical formulation of situations that arise in diverse areas. Common examples occur in the control of aircraft and missile regimes, control of reactors, and control of inventories. These problems have been studied at RAND over a period of years by various methods, and have been the subject of extensive study by Soviet mathematicians as well. The present memorandum studies these problems from the point of view of the calculus of variations. The relationship of this study to previous RAND work and some of the Soviet work is discussed.

This paper is scheduled to be published in The Journal of Mathematical Analysis and Applications.

SUMMARY

It is shown how a fairly general control problem, or programming problem, with constraints can be reduced to a special type of classical Bolza problem in the calculus of variations. Necessary conditions from the Bolza problem are translated into necessary conditions for optimal control. It is seen from these conditions that Pontryagin's maximum principle is a translation of the usual Weierstrass condition, and is applicable to a wider class of problems than that considered by Pontryagin. The differentiability and continuity properties of the value of the control are established under reasonable hypotheses on the synthesis, and it is shown that the value satisfied the Hamilton-Jacobi equation. As a consequence, a rigorous proof of a functional equation of Bellman is obtained and is shown to be valid for a much wider class of problems than heretofore considered. A sufficiency theorem for the synthesis of control is also given.

CONTENTS

PREFACE iii

SUMMARY v

Section

1. INTRODUCTION 1
2. NOTATION AND STATEMENT OF PROBLEM 5
3. THE EQUIVALENT BOLZA PROBLEMS 8
4. NECESSARY CONDITIONS FOR PROBLEM II 12
5. NECESSARY CONDITIONS FOR PROBLEM I 14
6. NECESSARY CONDITIONS FOR THE CONTROL PROBLEM 17
7. INTEGRABLE CONTROLS 21
8. NORMALITY 22
9. DISCONTINUOUS f, G, AND R 24
10. DEFINITION OF SYNTHESIS 25
11. THE FUNCTIONS λ, μ, AND L 32
12. THE VALUE AND THE HAMILTON-JACOBI EQUATION 35
13. AN EQUATION OF DYNAMIC PROGRAMMING 40
14. THE PROBLEM OF SYNTHESIS 41

REFERENCES 45

VARIATIONAL METHODS IN PROBLEMS OF CONTROL AND PROGRAMMING

1. INTRODUCTION

A controlled, or programmed, system is one for which the state at time t is represented by a real n-dimensional vector $x(t) = (x^1(t), \ldots, x^n(t))$ that is determined by a system of differential equations and initial conditions,

$$(1.1) \qquad \frac{dx^i}{dt} = G^i(t,x,u), \quad x^i(t_0) = x_0^i, \quad i = 1, \ldots, n,$$

where $u = (u^1(t), \ldots, u^m(t))$. The m-dimensional vector $u(t)$ is called the control function, or control, or the program for the system; it is usually required to satisfy constraints

$$(1.2) \qquad R^j(t,x,u) \geq 0, \qquad\qquad j = 1, \ldots, r.$$

The problem of optimal control, or the programming problem, is to choose the control $u(t)$ so as to bring the system from the given initial state to a terminal state (t_1,x_1), or one of a collection of terminal states $\{(t_1,x_1)\}$, in such a way as to minimize (or maximize) a functional

$$(1.3) \qquad J(u) = g(t_1,x_1) + \int_{t_0}^{t_1} f(t,x,u)dt,$$

where g is a function defined on the set of terminal states and the integral is evaluated along the solution of (1.1) corresponding to the choice of $u(t)$. A more complete and precise statement of the problem will be given in Sec. 2.

It is generally recognized that in the absence of the constraints (1.2), control problems, as usually formulated, are special cases of the problem of Bolza in the calculus of variations. In attacking problems in which constraints of the form (1.2) are present, as well as constraints of the form

$$(1.4) \qquad \int_{t_0}^{t_1} \varphi^k(t,x,u)dt \leq C^k, \qquad k = 1, \ldots, K,$$

several avenues have been explored. One is the "maximum principle" developed by Pontryagin and his collaborators Boltyanski and Gamkrelidze [13] for problems of the following type. The constraints are independent of x and require u to lie in a closed set, the function g is absent, the terminal state x_1 is a prescribed vector, and the terminal time is arbitrary. An extension of the maximum principle to problems in which the time of termination t_1 is fixed, x_1 is free, and g is a linear function of the coordinates was given by Rozonoèr [14].

Another approach, which is formal and heuristic in character, is the dynamic-programming argument of Bellman [1], who presents a functional equation that the value of the minimum as a function of initial position must satisfy. The terminal condition in this class of problems has t_1 fixed and x_1 free. Rozonoèr [14] has rigorously established the validity of the functional equation presented by Bellman for those problems in this class in which

$$g = \sum_{i=1}^{n} c^i x_1^i(t_1).$$

A different set of techniques has been used in dealing with linear systems (1.1). The problem of determining a control $u(t)$, subject to the constraints $|u^i(t)| \leq 1$, $i = 1, \ldots, m$, that brings $x(t)$ to 0 in minimum time was studied for systems with

$$G^i = \sum_{j=1}^{n} a_{ij}x^j + \sum_{j=1}^{m} b_{ij}u^j$$

by Bushaw [5], Bellman, Glicksberg, and Gross [2], and Gamkrelidze [6]. The problem of determining u so as to minimize the time required for $x(t)$ to hit a moving particle $z(t)$ for linear systems in which a_{ij} and b_{ij} are functions of time was studied by Krasovskii [8] and LaSalle [10]. The paper by LaSalle gives a brief survey of the other papers cited in this paragraph. Krasovskii [9] has considered the last problem for systems (1.1) of the form

$$G^i = f^i(t,x) + b^i(t)u.$$

In the first part of this paper we shall show how a fairly general control problem with constraints can be reduced to a special type of classical Bolza problem. Necessary conditions from the Bolza problem will be translated into necessary conditions for optimal control. These conditions give more information than the necessary conditions presented

by the authors cited, and are applicable to wider classes of problems. For example, it will be seen that the maximum principle is a restatement of the Weierstrass condition in the calculus of variations and is applicable to more general problems than those considered in [13] and [14]. Results on "bang-bang" control can be derived from Corollaries 1 and 2 of Theorem 2, but we shall not develop this topic here.

Theorem 2 of the present paper, which is the main theorem concerning the necessary conditions, was stated in slightly different form by Hestenes [7] in connection with aircraft climb problems, but was never published by him. Because of the relative unavailability of [7], we shall present the proof of Theorem 2. The constraint conditions of the present paper are slightly different from those of Hestenes. We also consider the case of discontinuous f, G^1, and R^1, and give simple criteria for normality in a special class of problems.

In the second part of the paper we study the function $W(t,x)$, which is defined as the value of the minimum (or maximum) of (1.3) as a function of initial position. We determine the differentiability properties of W under reasonable assumptions on the synthesis of control, and show that in its regions of differentiability the function W satisfies the Hamilton-Jacobi equation. By combining this equation with the Weierstrass condition (or maximum principle), we can rigorously establish the functional equation of Bellman [1] and obtain a statement about its regions of validity for a very general class of problems.

Our last theorem is a sufficiency theorem that is useful in synthesizing the control. This theorem is a variant of the standard sufficiency theorem in the calculus of variations. A similar theorem was stated by Breakwell [4]; his statement, however, needs an additional hypothesis to be valid, and his proof is formal.

We conclude our introductory remarks with the observation that problems in which constraints of the form (1.4) are present can be reduced to problems without these constraints by the introduction of new state variables and associated initial and terminal conditions as follows:

$$\frac{dx^{n+k}}{dt} = \varphi^k(t,x,u), \quad x^{n+k}(t_0) = 0, \quad x^{n+k}(t_1) \leq c^k, \qquad k = 1, \ldots, K.$$

2. NOTATION AND STATEMENT OF PROBLEM

Vector matrix notation will generally be used. Vectors and matrices will be denoted by single letters. Superscripts will be used to denote the components of a vector; subscripts will be used to distinguish vectors. Vectors will be written as matrices consisting of either one row or one column. We shall not use a transpose symbol to distinguish between the two usages, as it will be clear from the context how the vector is to be considered. If A is a matrix of m rows and n columns, x is an m-dimensional vector, and y is an n-dimensional vector, then in the product xA, x must be a

row matrix, and in the product Ay, y must be a column matrix. Thus we write the inner product of two vectors x and y simply as xy; a quadratic form with matrix A we write as xAx.

The operator (d/dt) will generally be denoted by a prime. Thus, the system (1.1) will be written as

$$(1.1) \qquad x' = G(t,x,u), \quad x(t_0) = x_1,$$

and the constraints (1.2) as $R(t,x,u) \geq 0$. (A vector is nonnegative if and only if every component is nonnegative.) If $Z(t,x,u)$ is a vector-valued function that is differentiable on a region $\mathcal{S}$ of (t,x,u) space, we denote the matrix of partial derivatives $(\partial Z^i/\partial x^j)$ by Z_x; the symbol Z_u has similar meaning. For real-valued functions $Z(t,x,u)$, the symbols Z_x and Z_u represent vectors of partial derivatives. We denote the determinant of a square matrix A by $||A||$.

Let $\mathcal{D}$ be a bounded region of $(n + 1)$-dimensional (t,x) space and let $\mathcal{U}$ be a region of m-dimensional u space. Let $\mathcal{S} = \mathcal{D} \times \mathcal{U}$. Let $\mathcal{T}$ be a manifold of class C'', of dimension $p \leq n$, lying in $\mathcal{D}$, and given parametrically by equations

$$(2.1) \qquad t = t_1(\sigma), \quad x = x_1(\sigma),$$

where $\sigma = (\sigma^1, \ldots, \sigma^p)$ ranges over an open cube $\mathcal{K}$ in p-dimensional space. Points of $\mathcal{T}$ will henceforth be denoted as (t_1,x_1); we shall call $\mathcal{T}$ the terminal manifold. Let

$f(t,x,u)$ be a real-valued function of class C'' on $\mathcal{D}$, let $g(\sigma)$ be a real-valued function of class C'' on $\mathcal{K}$, and let the vector-valued functions $G(t,x,u) = (G^1, \ldots, G^n)$ of (1.1) and $R(t,x,u) = (R^1, \ldots, R^r)$ of (1.2) be of class C'' on $\mathcal{S}$. Furthermore, let the constraint vector R satisfy the following <u>constraint conditions</u>:

(2.2)

(i) If $r > m$, then at each point of $\mathcal{S}$ at most m components of R can vanish.

(ii) At each point of $\mathcal{S}$ the matrix $(\partial R^i/\partial u^j)$, where i ranges over those indices such that $R^i(t,x,u) = 0$ and $j = 1, \ldots, m$, has maximum rank.

Consider the class of all functions $u = u(t)$ that are piecewise C' (i.e., each component u^i of u is piecewise continuous and has piecewise continuous first and second derivatives) on the closure of the projection of $\mathcal{D}$ on the t axis, and have range contained in $\mathcal{U}$. For each such u we can obtain a continuous solution of (1.1) that defines a curve K, with possible corners, in $\mathcal{D}$. Let $\mathcal{A}$ be the subclass of this class of functions u with the following properties:

(i) The curve K is defined and is interior to $\mathcal{D}$ for $t_0 \leq t \leq t_1$, where $(t_1,x_1) = (t_1,x(t_1))$ is a point of $\mathcal{T}$, and K does not intersect $\mathcal{T}$ for any $t_0 \leq t < t_1$.

(ii) Along K, the constraints (1.2) are satisfied; i.e., $R(t,x(t),u(t)) \geq 0$.

The class $\mathcal{A}$, which depends on (t_0, x_0), is called the class of admissible controls. For a given (t_0, x_0) it may be void.

The problem of optimal control is to find an element $u^* \epsilon \mathcal{A}$ that minimizes (or maximizes) the functional

$$(2.3) \qquad J(u) = g(\sigma) + \int_{t_0}^{t_1} f(t,x,u)dt$$

over all $u \epsilon \mathcal{A}$, where the integral is taken along the curve K corresponding to u, and σ is the parameter value associated with $(t_1, x_1) = (t_1, x(t_1))$. For definiteness we shall henceforth assume that (2.3) is to be minimized.

We note that the problem of optimal control as presented here is equivalent to the problem in which $g \equiv 0$ or the problem in which $f \equiv 0$. The equivalence of these problems can be shown by making transformations similar to those used to show the equivalence of the problems of Bolza, Lagrange, and Mayer in the calculus of variations ([3], pp. 189–190).

3. THE EQUIVALENT BOLZA PROBLEMS

Let $y = (y^1, \ldots, y^m)$ be an m-dimensional vector. To the system (1.1) adjoin the system of differential equations

$$(3.1) \qquad y' = u, \quad y(t_0) = 0.$$

The following problem of Bolza in $(n + m + 1)$-dimensional (t,x,y) space with differential inequalities as added side conditions is clearly equivalent to the problem of optimal control posed in Sec. 2.

Problem I. Find an arc $(x(t),y(t))$ that minimizes

$$(3.2) \qquad g(\sigma) + \int_{t_0}^{t_1} f(t,x,y')dt$$

in the class of arcs that are piecewise C'', that satisfy the differential equations

$$(3.3) \qquad G(t,x,y') - x' = 0,$$

the differential inequalities

$$(3.4) \qquad R(t,x,y') \geq 0,$$

and the end conditions

$$(3.5) \qquad \begin{aligned} &x(t_0) = x_0, \quad y(t_0) = y_0 = 0, \\ &t_1 = t_1(\sigma), \quad x_1 = x_1(\sigma), \\ &\qquad y_1 = y(t_1) = \eta, \end{aligned}$$

where $\eta = (\eta^1, \ldots, \eta^m)$.

By means of a device used by Valentine in [15], we obtain the following problem of Bolza, which has no inequality side conditions and is equivalent to Problem I.

Problem II. Find an arc $(x(t),y(t),\xi(t))$, where $\xi = (\xi^1, \ldots, \xi^r)$, that minimizes (3.2) in the class of arcs that are piecewise C'', that satisfy the differential equations

$$(3.6)\qquad \begin{aligned} &G(t,x,y') - x' = 0, \\ &R(t,x,y') - (\xi')^2 = 0, \end{aligned}$$

and the end conditions (3.5) and

$$(3.7)\qquad \xi(t_0) = \xi_0 = 0, \quad \xi(t_1) = \xi_1 = \tau,$$

where $\tau = (\tau^1, \ldots, \tau^r)$ and $(\xi')^2 = ((\xi^{1'})^2, \ldots, (\xi^{r'})^2)$.

Let $u^* \epsilon \mathcal{Q}$ be an optimal control, let K^* be the corresponding curve, and let $x^*(t)$ be the function defining K^*, for $t_0 \leq t \leq t_1$. Let $y^*(t)$ denote the solution of (3.1) for $u = u^*$. It follows from the preceding discussion that $(x^*(t), y^*(t))$ satisfies Eqs. (3.3) – (3.5) and minimizes (3.2). Hence the arc defined by $(x^*(t), y^*(t), \xi^*(t))$, where

$$(\xi^*(t)')^2 = R(t, x^*, y^{*\prime}), \quad \xi^*(t_0) = 0,$$

furnishes a minimum for Problem II. We denote this arc by K_2^*. We assert that at every element $(x^*, y^*, \xi^*, x^{*\prime}, y^{*\prime}, \xi^{*\prime})$ of K_2^*, the equations (3.6) are independent; that is, the matrix

$$(3.8)\qquad \begin{pmatrix} G_{y'} & -I & 0 \\ R_{y'} & 0 & -2\Xi' \end{pmatrix}$$

has rank $(n + r)$ along K_2^*, where I is the n-dimensional

identity matrix and $2\,\Xi'$ is an $r \times r$ diagonal matrix with entries $2(\xi^i)'$ on the diagonal, $i = 1, \ldots, r$. In order to prove the assertion we first suppose that the first r_1 rows, $0 \leq r_1 \leq r$, of the submatrix $(R_{y'}\ \ 0\ \ -2\ \ \Xi')$ have elements $2\,\xi^i \neq 0$, and the remaining rows have elements $2\,\xi^i = 0$. This can always be achieved by permuting rows and relabeling. The matrix (3.8) now has the form

$$\begin{pmatrix} A_1 & D \\ A_2 & 0 \end{pmatrix},$$

where D is an $(n + r_1)$ by $(n + r_2)$ diagonal matrix with nonzero entries on the diagonal and 0 is a zero matrix. The matrix A_2 consists of the last $r - r_1$ rows of the matrix $R_{y'}$. For each of these rows, we have $(\xi^i)' = 0$. Consequently, we have $R^i(t,x^*,y^{*\prime}) = 0$, $i = r_1 + 1, \ldots, r$. From (3.1) we see that this is equivalent to $R^i(t,x^*,u^*) = 0$, for $i = r_1 + 1, \ldots, r$. From the constraint conditions (2.2) it follows that $r - r_1 \leq m$ and that the matrix with elements $\partial R^i/\partial u^j$, $i = r_1 + 1, \ldots, r$, $j = 1, \ldots, m$, has rank $r - r_1$ for (t,x^*,u^*) along K_2^*. Hence, it follows that A_2 has rank $r - r_1$ and (3.8) has rank $(n + r_1) + (r - r_1) = n + r$, as required.

The above argument actually is not restricted to K_2^*; it shows that (3.8) has rank $(n + r)$ at all elements (t,x,y,ξ,x',y',ξ') for which (3.6) holds.

4. NECESSARY CONDITIONS FOR PROBLEM II

Since K_2^* furnishes a minimum for Problem II and the matrix (3.8) has rank $(n + r)$ wherever Eqs. (3.6) hold, it follows (Bliss [3], McShane [11]) that the following necessary conditions hold along K_2^*:

Theorem 1. *There exist a constant* $\lambda_0 \geq 0$, *an* n-*dimensional vector* $\lambda(t)$, *and an* r-*dimensional vector* $\mu(t)$, *defined on the interval* $t_0 \leq t \leq t_1$, *such that* $(\lambda_0, \lambda(t), \mu(t))$ *is never zero and such that* $\lambda(t)$ *and* $\mu(t)$ *are continuous, except perhaps at values of* t *corresponding to corners of* K_2^*, *where they possess unique right-hand and left-hand limits. Moreover, the function*

$$(4.1) \quad F(t,x,y,\xi,x',y',\xi',\lambda_0,\lambda,\mu) = \lambda_0 f + \lambda(G - x') + \mu(R - \xi'^2)$$

satisfies the following conditions along K_2^*:

(i) (Euler-Lagrange equations) *Between corners of* K_2^*, *we have*

$$(4.2) \quad \frac{dF_{x'}}{dt} = F_x, \quad \frac{dF_{y'}}{dt} = F_y, \quad \frac{dF_{\xi'}}{dt} = F_\xi.$$

At a corner, these equations hold for the unique one-sided limits.

(ia) (Weierstrass-Erdmann) *At a corner of* K_2^*, $F_{x'}$, $F_{y'}$, $F_{\xi'}$, *and* $(F - x' F_{x'} - y' F_{y'} - \xi' F_{\xi'})$ *have well-defined, one-sided limits that are equal.*

(ii) (Transversality) <u>At the end point</u> $(t_1, x_1^*(t), y_1^*(t), \xi_1^*(t))$ <u>of</u> K_2^*, <u>we have</u>

$$(4.3)\quad (F - x' F_{x'} - y' F_{y'} - \xi' F_{\xi'})t_{1\sigma} + F_{x'}x_{1\sigma} + \lambda_0 g_\sigma = 0,$$

$$F_{y'}\, y_{1\eta} = 0, \quad F_{\xi'}\, \xi_{1\tau} = 0.$$

(iii) (Weierstrass) <u>For all</u> $(t, x, \xi, y, X', Y', \zeta') \neq (t, x, \xi, y, x', y', \xi')$ <u>and satisfying</u> (3.6), <u>the inequality</u>

$$(4.4)\quad \mathcal{E}(t, x, y, \xi, x', y', \xi', X', Y', \zeta', \lambda_0, \lambda, \mu) \geq 0$$

<u>holds, where</u>

$$\mathcal{E} = F(t, x, y, \xi, X', Y', \zeta') - F(t, x, y, \xi, x', y', \xi')$$

$$- (X' - x')F_{x'} - (Y' - y')F_{y'} - (\zeta' - \xi')F_{\xi'},$$

<u>the functions</u> $F_{x'}$ <u>and</u> $F_{\xi'}$ <u>being evaluated at</u> $(t, x, y, \xi, x', y', \xi', \lambda_0, \lambda, \mu)$, <u>and the arguments</u> $(\lambda_0, \lambda, \mu)$ <u>being omitted throughout</u>.

(iv) (Clebsch) <u>For every vector</u> $(\pi, \rho, \kappa) \neq 0$, <u>where</u> $\pi = (\pi^1, \ldots, \pi^n)$, $\rho = (\rho^1, \ldots, \rho^m)$, <u>and</u> $\kappa = (\kappa^1, \ldots, \kappa^r)$, <u>that is a solution of the linear system</u>

$$(4.5)\quad G_{y'}\rho - I\pi = 0,$$

$$R_{y'}\rho - 2\Xi'\kappa = 0,$$

the following inequality holds:

$$(4.6) \qquad \pi F_{x'x'} \pi + \rho F_{y'y'} \rho - 2 \sum_{i=1}^{r} \mu^i (\kappa^i)^2 \geq 0.$$

5. NECESSARY CONDITIONS FOR PROBLEM I

We now follow Valentine [15] and translate the necessary conditions for Problem II into necessary conditions for Problem I. We first consider the Euler equations. From (4.1) we get

$$(5.1) \qquad F_\xi = 0, \quad F_{\xi^{i'}} = -2\mu^i \xi^{i'}, \qquad i = 1, \ldots, r.$$

Hence it follows from the third equation in (4.2) that $d(\mu^i \xi^i)/dt = 0$ along K_2^*. This and the continuity of $F_{\xi'}$ at corners of K_2^* imply that $\mu^i \xi^{i'}$ is constant along K_2^*. From the transversality condition (4.3) we get $F_{\xi'} \xi_{1\tau} = 0$; from (3.7) we obtain $\xi_{1\tau} = I$, where I is the $r \times r$ identity matrix. Therefore, we have $F_{\xi'} = 0$ at the right-hand end point of K_2^*, and consequently $\mu^i \xi^{i'} = 0$ along K_2^*. It now follows from the second equation in (3.6) that along K_2^*,

$$(5.2) \qquad \mu^i R^i = 0 \qquad\qquad i = 1, \ldots, r.$$

A similar argument shows that along K_2^*, we have $F_{y'} = 0$. We now introduce the function

$$(5.3) \qquad H(t,x,y',\lambda_0,\lambda) = \lambda_0 f(t,x,y') + \lambda G(t,x,y').$$

Clearly,

$$(5.4) \qquad F = H - \lambda x' + \mu(R - \xi'^2).$$

The following are immediate consequences of (5.4):

$$(5.5) \qquad \begin{gathered} F_x = H_x + \mu R_x, \quad F_{y'} = H_{y'} + \mu R_{y'}, \\ F_{x'} = -\lambda. \end{gathered}$$

Since $F_{y'} = 0$ along K_2^*, we see that along K_2^* we have

$$(5.6) \qquad H_{y'} + \mu R_{y'} = 0.$$

From (4.2) and (5.5), along K_2^* we also have

$$(5.7) \qquad \lambda' = -(H_x + \mu R_x).$$

It follows from the vanishing of $F_{y'}$ and $F_{\xi'}$ along K_2^*, and from (5.4), (5.5), and the second equation of (3.6), that

$$(5.8) \qquad F - x' F_{x'} - y' F_{y'} - \xi' F_{\xi'} = H$$

along K_2^*. Hence it follows that the transversality condition becomes

$$(5.9) \qquad \lambda_0 g_\sigma + Ht_{1\sigma} - \lambda x_{1\sigma} = 0.$$

The relationships used to establish (5.8) and the fact that (t, x, Y', ζ') satisfies the second equation of (3.6)

enable us to translate the Weierstrass condition (4.4) into the condition

$$(5.10) \qquad H(t,x,Y',\lambda_0,\lambda) \geq H(t,x,y',\lambda_0,\lambda).$$

It is an immediate consequence of (5.4) that (4.6) becomes

$$(5.11) \qquad \rho((H + \mu R)_{y'y'})\rho - 2\sum_{i=1}^{r} \mu^i(\kappa^i)^2 \geq 0.$$

If $R^i > 0$ at a point of K_2^*, then by (5.2), $\mu^i = 0$. If $R^i = 0$ at this point, let $\pi = 0$, let $\rho = 0$, and let κ be a vector with i-th component 1 and other components 0. Then $(\pi,\rho,\kappa) \neq 0$; and since $\xi^i = 0$, (π,ρ,κ) is a solution of (4.5). Hence from (5.11) we get $\mu^i \leq 0$ at this point. Consequently, we always have

$$\mu^i \leq 0 \text{ along } K_2^*, \qquad i = 1, \ldots, r.$$

Let (t,x,y) be a point of K_2^* such that at most r_1, where $r_1 < m$, components of $R(t,x,y')$ vanish; we suppose for definiteness that these are the first r_1 components. It follows from (2.2)-(ii) that the system of linear equations

$$\sum_{j=1}^{m} \frac{\partial R^i}{\partial y^{j\prime}} \rho^j - 2\xi^i\kappa^i = 0, \qquad i = 1, \ldots, r_1,$$

has a solution in ρ and $\hat{\kappa} = (\kappa^1, \ldots, \kappa^{r_1})$ such that $\rho \neq 0$ and $\hat{\kappa} = 0$. It now follows from the second system of equations in (3.6) and the assumption that $R^j > 0$ for $j > r_1$, that

the system (4.5) has a solution (ρ, π, κ) such that $\rho \neq 0$ and $\hat{\kappa} = 0$. Let $j > r_1$. Since indices $j > r_1$ correspond to components $R^j > 0$, it follows from (5.2) that $\mu^j = 0$ for $j > r_1$. Hence each term in the second summation in (5.11) vanishes, and from (4.5) we have

$$(5.12) \qquad \rho((H + \mu R)_{y'y'})\rho \geq 0$$

for any solution vector ρ of the system

$$(5.13) \qquad \sum_{j=1}^{m} \frac{\partial R^i}{\partial y^{j'}} \rho^j = 0, \qquad i = 1, \ldots, r_1.$$

The conclusion just stated holds, of course, even if m components of R vanish. In that case, however, the system (5.13) has only the trivial solution.

6. NECESSARY CONDITIONS FOR THE CONTROL PROBLEM

The following theorem, in which necessary conditions for optimal control are given, is an immediate consequence of the conclusion obtained in Sec. 5, and of the use of (3.1) to justify the replacing of the argument y' by u, wherever y' occurs. The function H is now

$$H(t,x,u,\lambda_0,\lambda) = \lambda_0 f(t,x,u) + \lambda G(t,x,u).$$

<u>Theorem 2. Let $u^* \in \mathcal{A}$ be an optimal control, let K^* be the corresponding curve, and let $x^*(t)$ be the function</u>

defining K^* on $[t_0,t_1]$. Then there exist a constant $\lambda_0 \geq 0$, an n-dimensional vector $\lambda(t)$ defined and continuous on $[t_0,t_1]$, and an r-dimensional vector $\mu(t) \leq 0$ defined and continuous on the interval $[t_0,t_1]$, except perhaps at values of t corresponding to corners of K^*, where it possesses unique right-hand and left-hand limits, such that the vector $(\lambda_0,\lambda(t))$ never vanishes, and such that the following conditions are fulfilled:

Condition I. Along K^* the following equations hold:

$$x'(t) = H_\lambda, \tag{6.1}$$

$$\lambda'(t) = -(H_x + \mu R_x), \tag{6.2}$$

$$H_u + \mu R_u = 0, \tag{6.3}$$

$$\mu^i R^i = 0, \qquad i = 1, \ldots, r. \tag{6.4}$$

At the end point (t_1,x_1^*) of K^* the transversality condition holds:

$$\lambda_0 g_\sigma + Ht_{1\sigma} - \lambda x_{1\sigma} = 0. \tag{6.5}$$

Along K^*, the function H is continuous.

Condition II. For every element (t,x^*,u^*) of K^* and every u such that $u = u(t)$ for some u in $\mathcal{A}$, we have

$$H(t,x^*,u,\lambda_0,\lambda) \geq H(t,x^*,u^*,\lambda_0,\lambda). \tag{6.6}$$

<u>Condition III</u>. <u>At each point of</u> K^* <u>let</u> $\hat{R}$ <u>denote the vector formed from</u> R <u>by taking those components of</u> R <u>that vanish at that point. Let</u> $e = (e^1, \ldots, e^m)$ <u>be a nonzero solution vector of the linear system</u> $\hat{R}_u e = 0$ <u>at a point of</u> K^*. <u>Then</u> $e((H + \mu R)_{uu})e \geq 0$ <u>at this point</u>.

Equations (6.1) – (6.4) are the Euler equations, Condition II follows from the Weierstrass condition (5.10), and Condition III follows from the Clebsch condition (5.12). The continuity of H along K^* follows from the continuity of the left-hand member of (5.8) along K_2^*, and the continuity of λ follows from (5.5) and the continuity of $F_{x'}$ (Weierstrass–Erdmann corner conditions). The nonvanishing of (λ_0, λ) along K^* is established as follows. If (λ_0, λ) were zero at a point of K^*, then from (6.3) we would have $\mu R_u = 0$ at this point. For the sake of definiteness, suppose that the indexing is such that $R^i = 0$ for $i = 1, \ldots, r_1$, where by (2.2), $r_1 \leq m$. Hence, by (6.4), we have $\mu^i = 0$ for $i > r_1$. Thus the condition $\mu R_u = 0$ reduces to a system of linear equations in $\mu^1, \ldots, \mu^{r_1}$ with coefficient matrix $(\partial R^i/\partial u^k)$, $i = 1, \ldots, r_1$, $k = 1, \ldots, m$. From (2.2) – (ii), this matrix has rank r_1. Hence $\mu^1 = \ldots = \mu^{r_1} = 0$ is the only solution of the linear system. Thus, we have shown that if (λ^0, λ) is zero at a point, then the vector $(\lambda^0, \lambda, \mu)$ must also be zero, contradicting the assertion of Theorem 1.

If the constraints are specialized, then important simplifications can be effected in the Euler equations.

<u>Corollary 1</u>. <u>Let the constraints be of the form</u>

$$B^i(t,x) \leq u^i \leq A^i(t,x), \qquad i = 1, \ldots, m,$$

<u>where</u> $A^i > B^i$ <u>and each</u> A^i <u>and</u> B^i <u>is of class</u> C'' <u>on</u> $\mathcal{D}$. <u>Then at each point of</u> K^* <u>we have</u>

$$H_{u^i} \begin{cases} \geq 0 & \text{if } u^{*i} = B^i, \\ = 0 & \text{if } B^i < u^{*i} < A^i, \\ \leq 0 & \text{if } u^{*i} = A^i, \end{cases} \qquad i = 1, \ldots, m.$$

If we write the constraints as $A^i - u^i \geq 0$ and $u^i - B^i \geq 0$, $i = 1, \ldots, m$, we obtain a 2m-dimensional constraint vector with components $A^i - u^i$ and $u^i - B^i$. It follows from the condition $A^i > B^i$ and the form of the constraints that (2.2) is satisfied. The conclusion of the corollary follows from (6.3) and (6.4) by straightforward calculation and use of the condition $\mu \leq 0$.

<u>Remark</u>. If the i-th component of u is constrained only from one side, say $u^i \leq A^i(t,x)$, then $H_{u^i} = 0$ if $u^i < A^i$ and $H_{u^i} \leq 0$ if $u^i = A^i$. Similar statements hold for $u^i \geq B^i$.

Another important special case is one in which the constraints are independent of the state, that is, $R(t,x,u) \equiv R(t,u)$. Since $R_x = 0$ in this case, we have the following corollary:

<u>Corollary 2</u>. <u>If</u> R <u>is independent of</u> x, <u>then equation</u> (6.2) <u>becomes</u>

$$(6.2)' \qquad \lambda' = -H_x.$$

In the problem considered by Pontryagin [13], the constraints required u to lie in a fixed closed set, independent of time t and position x. Equations (6.1), (6.2)', and (6.6) constitute the maximum principle as stated by Pontryagin. Our function H is the negative of Pontryagin's, so that his maximum appears as a minimum in our paper. Note, however, that the Euler equations and Condition II of Theorem 2, which is the Weierstrass condition, give a minimum principle for a wider class of problems.

<u>Remark</u>. Note that if the A^i and B^i of Corollary 1 are constants, then the results of both corollaries are valid.

7. INTEGRABLE CONTROLS

Instead of considering functions $u = u(t)$ that are piecewise C'', we can consider functions that are merely assumed to be Lebesgue integrable. In this way we can define a class of admissible controls $\mathcal{A}^+$, and we can look for an optimal control u^* in $\mathcal{A}^+$. The curves K corresponding to functions u in $\mathcal{A}^+$ will be defined by absolutely continuous functions $x(t)$, and so will be rectifiable. We can reduce the control problem with constraints to a Bolza problem without constraints as we did before, except that the functions

$(x(t),y(t),\xi(t))$ are now absolutely continuous. To this problem we can apply a theorem of McShane (Theorem 16.1, [12]). We can then translate back to the original control problem and obtain the result that the conclusions of Theorem 2, appropriately modified, hold almost everywhere along a curve K^* corresponding to a control u^* that minimizes (2.3) over all u in $\mathcal{A}^+$.

8. NORMALITY

A piecewise C'' minimizing curve K^*, or equivalently the corresponding curve K_2^* of Problem II, is said to be <u>normal</u> if there are no sets of multipliers with $\lambda_0 = 0$. (See [3], pp. 213–219.) If the minimizing curve is normal, then the multipliers can be chosen so that $\lambda_0 = 1$, and with this choice of λ_0 they are unique. If the curve is not normal, there may be no neighboring curves that satisfy the differential equations, constraints, and end conditions. Necessary and sufficient conditions for normality are given in [3]. These criteria applied to the present problem would involve variations along K_2^* and would generally be difficult to apply in practice. We shall give a condition for normality in the control problem that is sufficient, but not necessary. It is, however, easier to apply in practice, and reduces to a very simple condition in the special case that the terminal manifold $\mathcal{T}$ is n-dimensional.

At (t_1,x_1^*), the end point of K^*, let r_1 components of $R(t_1,x_1^*,u^*(t_1))$ vanish. From (2.2) – (i) we get $r_1 \leq m$. Let $\hat{R}$ denote the r_1-dimensional vector formed from R by taking those components of R that vanish at (t_1,x_1^*), and

let $\hat{\mu}$ be the vector formed from μ by taking the corresponding components. Then from (6.4) we have $\mu^j(t_1) = 0$ for those components of μ that are not in $\hat{\mu}$. Let M denote the n by p matrix with typical element

$$(8.1) \qquad \left(G^i \frac{\partial t_1}{\partial \sigma^j} - \frac{\partial x_1^i}{\partial \sigma^j}\right), \qquad i = 1, \ldots, n, \quad j = 1, \ldots, p,$$

where the elements are evaluated at the end point of K^*. Let C denote the $(n + r_1)$ by $(m + p)$ matrix

$$\begin{pmatrix} G_u & M \\ \hat{R}_u & 0 \end{pmatrix},$$

where G_u and $\hat{R}_u$ are evaluated at (t_1, x_1^*).

If K^* is not normal, then there exists a set of multipliers $(\lambda_0, \lambda, \mu)$ with $\lambda_0 = 0$. From (6.3) and (6.5) we see that, at the end point (t_1, x_1^*) of K^*, the vector $(\lambda, \hat{\mu})$ is a solution of the linear system $(\lambda, \hat{\mu})C = 0$. The following theorem is now a consequence of a standard theorem concerning the solutions of homogeneous linear systems and the fact that (λ, μ) cannot be zero if $\lambda_0 = 0$:

<u>Theorem 3</u>. <u>If the rank of</u> C <u>is</u> $(n + r_1)$, <u>then</u> K^* <u>is normal</u>.

Note that C can have rank $(n + r_1)$ only when $(n + r_1) \leq (m + p)$, and that Theorem 3 is not a necessary condition.

<u>Corollary</u>. <u>If</u> $\mathcal{T}$ <u>is</u> n-<u>dimensional and</u> K^* <u>is not tangent to</u> $\mathcal{T}$, <u>then</u> K^* <u>is normal</u>.

If $\mathcal{T}$ is n-dimensional, the matrix M is an n by n matrix. By (2.2) – (ii), the r_1 by m matrix $\hat{R}_u$ has rank r_1. Hence, C has rank $(n + r_1)$ whenever M has rank n. Since K^* is not tangent to $\mathcal{T}$, the matrix

$$\begin{pmatrix} 1 & t_{1\sigma} \\ G & x_{1\sigma} \end{pmatrix}$$

has rank $n + 1$. If for each $j = 1, \ldots, n$ we multiply the first column of this matrix by $-\partial t_1/\partial\sigma^j$ and add the result to the j-th column, we get the matrix

$$\begin{pmatrix} 1 & 0 \\ G & -M \end{pmatrix}.$$

Hence M has rank n and the corollary follows.

9. DISCONTINUOUS f, G, AND R

Let $\mathcal{M}$ be a manifold of dimension n, lying in $\mathcal{D}$, and dividing $\mathcal{D}$ into two regions, such that some or all of the functions f, G, and R are discontinuous across $\mathcal{M}$. Let the discontinuity of a function be such that the function and its derivatives have unique one-sided limits. Further, let us assume that K^* intersects $\mathcal{M}$ at $(t_2, x_2) = (t_2, x_2(t))$

and is not tangent to $\mathcal{M}$ at this point. It can be shown by appropriate modifications of the arguments in [3] (pp. 196–202) that the multipliers λ and μ of Problem II need not be continuous at $t = t_1$, but will have unique right-hand and left-hand limits at (t_2, x_2) as will F and its various partial derivatives when evaluated along K_2^*. Although $F_{x'}$, $F_{y'}$, and $F_{\xi'}$ need not be continuous across $\mathcal{M}$, the expression

$$(F - x' F_{x'} - y' F_{y'} - \xi' F_{\xi'})dt + F_{x'}dx_2 + F_{y'}dy_2 + F_{\xi'}d\xi_2$$

has equal right-hand and left-hand limits along K_2^* at (t_2, x_2) for all differentials dt_2, dx_2 on $\mathcal{M}$ and all dy_2 and $d\xi_2$. For the original control problem this translates into the condition that

$$(9.1) \qquad (H^+ - H^-)dt_2 - (\lambda^+ - \lambda^-)dx_2 = 0$$

at (t_2, x_2), where the one-sided limits are evaluated along K^*.

10. DEFINITION OF SYNTHESIS

Consider a point (t_1, x_1) of the p-dimensional terminal manifold $\mathcal{T}$, where $0 \leq p \leq n$. Let $\mathcal{K}'$ denote a region in $(n - p)$-dimensional space over which a vector φ ranges. If $p = n$, then φ is the zero vector. Let $u^*(t; t_1, x_1, \varphi)$ be a function defined in some interval $[t_0, t_1]$, where $t_0 = t_0(t_1, x_1, \varphi)$, such that the following condition holds:

Assumption 1. (i) The function u^* is piecewise C'' on $[t_0,t_1]$, and its range lies in $\mathcal{U}$. (ii) If u^* is substituted into (6.1) (or equivalently into (1.1)), the resulting differential equation

$$(10.1) \qquad x' = G(t,x,u^*(t;t_1,x_1,\varphi)), \quad x(t_1) = x_1,$$

has a continuous solution $x^*(t;t_1,x_1,\varphi)$ on $[t_0,t_1]$ such that (t,x^*) lies in $\mathcal{D}$ and $R(t,x^*,u^*) \geq 0$.

We denote the curve corresponding to $x^*(t;t_1,x_1,\varphi)$ by $K(t_1,x_1,\varphi)$.

We now suppose that the assumptions just made for a particular point (t_1,x_1) hold for all points (t_1,x_1) of $\mathcal{T}$. From (2.1), we have $(t_1,x_1) = (t_1(\sigma),x_1(\sigma))$, where σ ranges over an open cube $\mathcal{K}$ in a p-dimensional space. Let θ be an n-dimensional vector defined as follows:

$$(10.2) \qquad \theta = (\sigma,\varphi), \quad \sigma \in \mathcal{K}, \quad \varphi \in \mathcal{K}'.$$

We define functions

$$(10.3) \qquad \begin{aligned} t_0(\theta) &\equiv t_0(t_1(\sigma), x_1(\sigma),\varphi), \\ t_1(\theta) &\equiv t_1(\sigma), \end{aligned}$$

and functions

$$(10.4)\qquad \begin{aligned} u^*(t,\theta) &\equiv u^*(t;t_1(\sigma),x_1(\sigma),\varphi), \\ x^*(t,\theta) &\equiv x^*(t;t_1(\sigma),x_1(\sigma),\varphi), \end{aligned}$$

for σ in $\mathcal{K}$, φ in $\mathcal{K}'$, and $t_0(\theta) \leq t \leq t_1(\theta)$. The differential equation (10.1) can now be written as

$$(10.5)\qquad x' = G(t,x,u^*(t,\theta)), \quad x(t_1(\theta)) = x_1(\sigma).$$

Clearly, $x^*(t,\theta)$ is a solution of (10.5). We shall denote the curve $K(t_1,x_1,\varphi)$ by $K(\theta)$.

Let Ω denote the domain of definition of $u^*(t,\theta)$ and $x^*(t,\theta)$; that is, the set of points (t,θ) in $(n+1)$-dimensional space with θ as in (10.2) and $t_0(\theta) \leq t \leq t_1(\theta)$. Clearly, Ω has nonvoid interior, which we denote by Ω^0. It follows from (2.1) and (10.3) that $t_1(\theta)$ defines a C'' manifold $\mathcal{N}_1$ of dimension n in (t,θ) space and that $\mathcal{N}_1$ is part of the boundary of Ω. We also suppose that $t_0(\theta)$ defines a C'' manifold of dimension n.

A set of functions

$$t = t_i(\theta), \qquad i = 1, 2, \ldots, \alpha,$$

defined and C'' on the region defined in (10.2), with $t_1(\theta)$ as in (10.3) and such that

$$t_0(\theta) < t_{i+1}(\theta) < t_1(\theta), \quad i = 1, \ldots, \alpha - 1,$$

will be said to induce a <u>regular decomposition</u> of Ω. Clearly, each $t_i(\theta)$, $i \geq 2$, defines a C'' manifold $\mathcal{N}_i$ of dimension n lying in Ω^0. We let $\mathcal{N}_{\alpha+1}$ denote the manifold defined by $t_0(\theta)$. We define subregions Ω_i of Ω^0 as follows:

$$\Omega_i = E\{(t,\theta) \in \Omega^0 | t_{i+1}(\theta) < t < t_i(\theta)\}, \qquad i = 1, \ldots, \alpha - 1.$$

We shall say that a function $h(t,\theta)$ is piecewise $C^{(k)}$ on Ω if on each subregion Ω_i it agrees with a function $h_{(i)}(t,\theta)$ that is $C^{(k)}$ on $\overline{\Omega_i}$, the closure of Ω_i.

Two more assumptions can now be stated.

<u>Assumption 2</u>. The function $x^*(t,\theta)$ maps Ω^0 in a one-to-one fashion onto a subregion $\mathcal{R}$ of the region $\mathcal{D}$ in (t,x)-space, and maps $\mathcal{N}_{\alpha+1}$ in a one-to-one fashion onto an n-dimensional manifold that forms part of the boundary of $\mathcal{R}$.

<u>Assumption 3</u>. There exist functions $t_i(\theta)$ that induce a regular decomposition of Ω such that: (i) The function $u^*(t,\theta)$ is piecewise C'' on Ω. (ii) If f, G, or R possess manifolds of discontinuity that lie in $\mathcal{R}$ (as discussed in Sec. 9), then each of these manifolds is coincident with the image of some set $\mathcal{N}_i$, $i = 2, \ldots, \alpha$. (iii) For each component R^j of the constraint vector R, we have either $R^j(t,x^*(t,\theta),u^*(t,\theta)) \equiv 0$ on Ω_i, or, with the possible exception of a finite number of points, $R^j(t,x^*(t,\theta),u^*(t,\theta)) > 0$ on Ω_i.

We shall denote the image of Ω_i by $\mathcal{R}_i$, $i = 1, \ldots, \alpha$, and the image of $\mathcal{N}_i$ by $\mathcal{M}_i$, $i = 2, \ldots, \alpha + 1$. The function x^* also maps $\mathcal{N}_1$ onto $\mathcal{T}$, whence we may set $\mathcal{M}_1 = \mathcal{T}$. Note, however, that the mapping of $\mathcal{N}_1$ onto $\mathcal{T}$ in general is not one-to-one.

<u>Lemma 1</u>. <u>The function</u> $x^*(t,\theta)$ <u>is continuous and is piecewise</u> C'' <u>on</u> Ω. <u>The sets</u> $\mathcal{M}_i$, $i = 1, \ldots, \alpha + 1$, <u>are manifolds of class</u> C''.

Let $u^*_{(1)}$ denote the function that is C'' on $\overline{\Omega_1}$ and that agrees with u^* on Ω_1. Let $G_{(1)}$ denote the function that is C'' on $\overline{\mathcal{R}_1} \times \mathcal{U}$ and that agrees with G on $\mathcal{R}_1 \times \mathcal{U}$. We may extend the function $u^*_{(1)}$ to a function $\tilde{u}^*_{(1)}$ that has range in $\mathcal{U}$ and that is C'' on a region containing $\overline{\Omega_1}$ (and hence $\mathcal{N}_1$ and $\mathcal{N}_2$) in its interior. We may also extend $G_{(1)}$ to a function $\tilde{G}_{(1)}$ that is C'' on a region containing $\overline{\mathcal{R}_1} \times \mathcal{U}$ in its interior. It now follows from (10.5), the properties of $t_1(\sigma)$ and $x_1(\sigma)$, Assumption 3-(i), and standard theorems about the behavior of solutions of differential equations with respect to parameters and initial conditions, that $x^*(t,\theta)$ is C'' on $\overline{\Omega_1}$. Since $\mathcal{M}_2$ is given by $t = t_2(\theta)$ and $x = x_2(\theta) \equiv x^*(t_2(\theta),\theta)$, it follows that $t_2(\theta)$ and $x_2(\theta)$ are C''. The argument just given can be repeated with the appropriate modifications on Ω_2 and $\mathcal{R}_2$, with $t = t_2(\theta)$ and $x = x_2(\theta)$ as the boundary conditions for (10.5). We then see that x^* has the desired properties

on $\overline{\Omega_2}$ and is continuous on $\overline{\Omega_1} \vee \overline{\Omega_2}$, and that $\mathcal{M}_3$ is given by $t = t_3(\theta)$, $x = x_3(\theta) \equiv x^*(t_3(\theta),\theta)$. Proceeding inductively in this fashion, we can establish the desired properties for x^*. We note that the sets $\mathcal{M}_i$, $i = 2, \ldots, \alpha + 1$, are given by functions

$$(10.6) \qquad t = t_1(\theta), \quad x = x_1(\theta) \equiv x^*(t_1(\theta),\theta),$$

and hence are manifolds of class C^r.

<u>Assumption 4</u>. For every subregion $\Omega^1 \subset \Omega^0$ at positive distance from $\mathcal{N}_1$, there exists a positive constant $d(\Omega^1)$ such that $||x^*_\theta(t,\theta)|| \geq d(\Omega^1)$ on $\overline{\Omega}^1$. (At boundary points of Ω^1 and at points of $\mathcal{N}_i$, $i \geq 2$, the bounding away from zero of the determinant is to be interpreted for the various limits.)

It can be shown that if $\mathcal{T}$ is n-dimensional, then the assumption that each curve $K(\theta)$ is not tangent to $\mathcal{T}$ implies the existence of a constant $d > 0$ such that $||x^*_\theta(t,\theta)|| \geq d$ on all of $\overline{\Omega}$.

It is an immediate consequence of Assumption 4 and (10.6) that the manifolds $\mathcal{M}_i$, $i = 2, \ldots, \alpha + 1$, have dimension n. It also follows from Assumption 4 that the curves $K(\theta)$ are not tangent (from either side) to a manifold $\mathcal{M}_i$, $i \geq 2$.

From Assumption 2, it follows that on Ω^0 the relation $x = x^*(t,\theta)$ can be inverted to give a relation

(10.7) $\theta = \theta(t,x)$,

where θ is a single-valued function on $\mathcal{R}$. It further follows from Lemma 1, Assumption 4, and the implicit-function theorem that θ is C'' on each $\mathcal{R}_i$, $i \geq 2$, and on the set $\bar{\mathcal{R}}_1 - \mathcal{T}$. Since $x^*(t,\theta)$ is one-to-one, it follows that θ is continuous on $\mathcal{R}$. From the identity $\theta = \theta(t,x^*(t,\theta))$ it follows that as (t,x) tends to $\mathcal{T}$ along $K(\theta)$, the function θ will tend to the value θ. In general, θ will not tend to a unique limit at points of $\mathcal{T}$. It can be shown, however, that if $\mathcal{T}$ is n-dimensional and the curves $K(\theta)$ are not tangent to $\mathcal{T}$, then θ is C'' on $\bar{\mathcal{R}}_1$ as well as on $\bar{\mathcal{R}}_i$, $i \geq 2$.

<u>Assumption 5</u>. (i) For every point $(\bar{t},\bar{x}) = (\bar{t},x^*(\bar{t},\theta))$ in $\mathcal{R}$, the control problem (2.3) with initial point $(\bar{t},\bar{x})$ has a unique solution in which the optimal control is $u^*(t,\theta)$, $\bar{t} \leq t \leq t_1(\theta)$, and the corresponding curve is $K(\theta)$. (ii) There exists a multiplier vector $(\lambda_0(\theta),\lambda(t,\theta),\mu(t,\theta))$ along each $K(\theta)$ such that $\lambda_0 = 1$ and the functions $\lambda_1(\theta) \equiv \lambda(t_1(\theta),\theta)$ and $\mu_1(\theta) = \mu(t_1(\theta),\theta)$ are C' on $\mathcal{K} \times \mathcal{K}'$.

The existence of multipliers along each $K(\theta)$ follows from Theorem 2; the assumption concerns the properties of λ_0, λ_1, and μ_1.

A function $u^*(t,\theta)$ such that Assumptions 1-5 hold will be called a <u>normal parametric synthesis</u> of the control.

Remark. If $\mathcal{T}$ is n-dimensional and each $K(\theta)$ is not tangent to $\mathcal{T}$, then (ii) follows from the Corollary of Theorem 3 and the transversality condition (6.5).

Define

$$U^*(t,x) = u^*(t,\theta(t,x)). \tag{10.8}$$

It follows from the preceding discussion that U^* is C'' on each $\bar{\mathcal{R}}_i$, for $i \geq 2$, and is C'' on $\bar{\mathcal{R}}_1 - \mathcal{T}$. Along each $K(\theta)$, however, $U^*(t,x)$ does tend to a limit as $\mathcal{T}$ is approached. If $\mathcal{T}$ is n-dimensional and the curves $K(\theta)$ are not tangent to $\mathcal{T}$, then U^* is C'' on $\bar{\mathcal{R}}_1$ as well. We call U^* a normal synthesis of the control.

11. THE FUNCTIONS λ, μ, AND L

Lemma 2. The functions $\lambda(t,\theta)$ and $\mu(t,\theta)$ are piecewise C' on Ω. Across every manifold $\mathcal{M}_i$, $i = 2, \ldots, \alpha$, equation (9.1) holds. If f, G, and R are continuous across $\mathcal{M}_i$, then so is $\lambda(t,\theta)$.

Let $\hat{R}$ denote the vector formed by taking those components R^j of R such that $R^j(t,x^*(t,\theta),u^*(t,\theta)) \equiv 0$ on Ω_1. Let $\hat{\mu}$ be the vector obtained from μ by taking the corresponding components. From (2.2)-(ii), it follows that $\hat{R}_u$ has maximum rank, say r_1, on Ω_1. Let $\hat{R}_{\hat{u}}$ be an r_1 by r_1 nonsingular submatrix of $\hat{R}_u$. Let $H_{\hat{u}}$ denote the vector obtained from H_u by selecting the components corresponding to the columns of

$\hat{R}_u$ used to obtain $\hat{R}_{\hat{u}}$. In order to simplify the exposition we shall assume that the same submatrix is nonsingular at all points of $\bar{\Omega}_1$. It will be seen from the ensuing discussion that this restriction can be easily overcome.

From (6.3), we get

$$(11.1) \qquad \hat{\mu} = -(H_{\hat{u}})(\hat{R}_{\hat{u}})^{-1}.$$

Since by (6.4) those components of μ that are not included in $\hat{\mu}$ vanish on $\bar{\Omega}_1$, we may write (6.2) along each $K(\theta)$ as follows:

$$(11.2) \qquad \lambda_t(t,\theta) = -H_x + (H_{\hat{u}})(\hat{R}_{\hat{u}})^{-1}\hat{R}_x, \quad \lambda(t_1,\theta) = \lambda_1(\theta),$$

where the arguments of the functions on the right-hand side are $(t,x^*(t,\theta),u^*(t,\theta))$. A proof similar to that used in Lemma 1 can now be used to show that $\lambda(t,\theta)$ is of class C' on $\bar{\Omega}_1$. It then follows from (11.1) that $\hat{\mu}$ is also C' on $\bar{\Omega}_1$. Since the other components of μ vanish on $\bar{\Omega}_1$, the vector μ is C' on $\bar{\Omega}_1$.

The same arguments applied to Ω_2, with $\hat{R}_u$, $\hat{R}_{\hat{u}}$, $H_{\hat{u}}$, and $\hat{\mu}$ appropriately redefined and with the proper initial data $\lambda(t_2(\theta),\theta)$, show that $\lambda(t,\theta)$ and $\mu(t,\theta)$ are C' on $\bar{\Omega}_2$. The initial data $\lambda(t_2(\theta),\theta)$ are defined by continuity or by (9.1) if $\mathcal{N}_2$ corresponds to a manifold of discontinuity of f, G, or R. Proceeding backwards in this fashion, we see that λ and μ are piecewise C' on Ω and have the requisite continuity properties.

Define

$$(11.3) \qquad L(t,x) = \lambda(t,\theta(t,x)), \quad (t,x) \in \mathcal{R}_i, \quad i = 1, \ldots, \alpha.$$

We list the properties of $L(t,x)$ in the following Lemma:

Lemma 3. The function L is C' on each $\bar{\mathcal{R}}_i$, $i = 2, \ldots, \alpha$, and on the set $\bar{\mathcal{R}}_1 - \mathcal{T}$. Moreover, if f, G, and R are continuous across a manifold $\mathcal{M}_i$, $i = 2, \ldots, \alpha$, then so is L. Across a manifold $\mathcal{M}_i$, (9.1) holds with λ replaced by L, where $+$ now indicates a limit from the interior of $\mathcal{R}_{i-1}$ and $-$ indicates a limit from the interior of $\mathcal{R}_i$. If $\mathcal{T}$ is n-dimensional and the curves $K(\theta)$ are not tangent to $\mathcal{T}$, then L is C' on $\bar{\mathcal{R}}_1$ also.

The proof of this lemma, except for the next to the last sentence, is an immediate consequence of Lemma 2 and the properties of $\theta(t,x)$. It is clear from the properties of f, G, $\lambda(t,\theta)$, $x^*(t,\theta)$, and $u^*(t,\theta)$ that $H(t,x^*(t,\theta),\lambda(t,\theta),u^*(t,\theta))$ is continuous on each of the sets $\Omega_i \cup \mathcal{N}_i$ and $\Omega_{i-1} \cup \mathcal{M}_i$, $i = 2, \ldots, \alpha$. Hence, $H(t,x,L,U^*)$ is continuous on each $\mathcal{R}_i \cup \mathcal{M}_i$ and $R_{i-1} \cup \mathcal{M}_i$, $i = 2, \ldots, \alpha$. If $\mathcal{M}_i$ is not a manifold of discontinuity of f, G, or R, then by Theorem 2, H is continuous across $\mathcal{M}_i$ along each $K(\theta)$. Hence from the continuity of H on $R_i \cup \mathcal{M}_i$ and $R_{i-1} \cup \mathcal{M}_i$ it follows that H is continuous across $\mathcal{M}_i$, unrestrictedly in this case. Since L is continuous across $\mathcal{M}_i$, (9.1) holds across $\mathcal{M}_i$ unrestrictedly. A similar argument shows that if $\mathcal{M}_i$

is a manifold of discontinuity of f, G or R, then (9.1) holds across $\mathcal{M}_i$, also without the restriction that the limits be taken along $K(\theta)$.

12. THE VALUE AND THE HAMILTON-JACOBI EQUATION

Let Assumptions 1 to 5 of Sec. 10 hold. Then we can define a function $W(t,x)$ on $\mathcal{R}$ by assigning to each point (t,x) in $\mathcal{R}$ the value that the functional (2.3) with $(t_0,x_0) = (t,x)$ takes along the optimal curve $K(\theta)$ through (t,x). Thus we have

$$(12.1) \qquad \begin{aligned} W(t,x) &= W(t,x^*(t,\theta)) \\ &= g(\sigma) + \int_t^{t_1(\sigma)} f(t,x^*(t,\theta),u^*(t,\theta))dt, \end{aligned}$$

where θ and σ are related by (10.2). We shall call W the <u>value function</u>, or simply the <u>value</u> of the control problem. We summarize the properties of W in the following theorem:

<u>Theorem 4</u>. <u>The value</u> W <u>is continuous on</u> $\mathcal{R}$, <u>is</u> C'' <u>on each</u> $\bar{\mathcal{R}}_i$, $i \geq 2$, <u>and is</u> C'' <u>on</u> $\mathcal{R}_1 - \mathcal{T}$. <u>On each</u> $\mathcal{R}_i$, $i = 1, \ldots, \alpha$, <u>we have</u>

$$(12.2) \qquad \begin{aligned} W_t(t,x) &= -f(t,x,U^*(t,x)) - L(t,x)G(t,x,U^*(t,x)), \\ W_x(t,x) &= L(t,x). \end{aligned}$$

<u>At points of a manifold</u> $\mathcal{M}_i$, $i \geq 2$, (12.2) <u>holds for the one-sided limits. If</u> $\mathcal{M}_i$ <u>is not a manifold of discontinuity</u>

of f, G, *or* R, *then* W_t *and* W_x *are continuous across* $\mathcal{M}_1$. *Across every manifold* $\mathcal{M}_j$, $j = 2, \ldots, \alpha$, *the relation*

$$W_t^+ \, dt_j - W_x^+ \, dx_j = W_t^- \, dt_j - W_x^- \, dx_j$$

holds for all differentials dt_j, dx_j *along* $\mathcal{M}_j$.

Remark 1. If we substitute the second equation of (12.2) into the first, we see that the value satisfies the Hamilton-Jacobi equation on each $\mathcal{R}_i$.

Remark 2. It follows from the properties of L, U^*, and θ that both W_t and W_x possess limits as (t,x) tends to $\mathcal{T}$ along a curve $K(\theta)$, even though in general W_t and W_x do not possess limits as (t,x) tends to $\mathcal{T}$. If, however, $\mathcal{T}$ is n-dimensional and the curves $K(\theta)$ are not tangent to $\mathcal{T}$, then W is C'' on $\bar{\mathcal{R}}_1$.

Remark 3. In Assumption 5-(ii) we supposed that along each $K(\theta)$ there was one set of multipliers with $\lambda_0 = 1$. The second equation in (12.2) now shows that if there is one such set satisfying the other requirements of Assumption 5, then it must be unique.

The proof that we now give for Theorem 4 is an extension of an argument used in the calculus of variations to prove the invariance of Hilbert's integral in certain fields.

It is clear from (12.1) that W is continuous. Let

$$(12.3) \qquad t = T_0(s), \; x = X_0(s), \qquad 0 \leq s \leq 1,$$

define a curve Γ that does not intersect itself and that, with the possible exception of end points, lies entirely within some $\mathcal{R}_1$. For definiteness we take $i = \alpha$. It follows from Assumptions 2 and 4 that the system of equations

$$(12.4) \qquad T_0(s) = t, \quad X_0(s) = x^*(t,\theta), \qquad 0 \leq s \leq 1,$$

defines a function $\theta = \theta(s)$ that is C'' on $[0,1]$. Hence if we traverse Γ as s goes from 0 to 1, we obtain a family of curves $K(s) = K(\theta(s))$ by means of the function $x^*(t,\theta(s))$, where $T_0(s) \leq t \leq t_1(s)$. Since the manifolds $\mathcal{M}_j$, $j = 1, \ldots, \alpha$, are given by (10.6), it follows that the intersections of the curves $K(s)$ with the manifolds $\mathcal{M}_j$ are given by

$$(12.5) \qquad \begin{aligned} x &= T_j(s) \equiv t_j(\theta(s)), \\ x &= X_j(s) \equiv x^*(t_j(\theta(s)),\theta(s)), \qquad j = 1, \ldots, \alpha. \end{aligned}$$

The functions T_j and X_j, $j = 2, \ldots, \alpha$, are clearly C'' on $[0,1]$. For $j = 2, \ldots, \alpha$, we can compute dX_j/ds from (12.5) in two ways,

$$(12.6) \qquad \frac{dX_j}{ds} = x_t^{*+} \frac{dT_j}{ds} + x_\theta^{*+} \frac{d\theta}{ds} = x_t^{*-} \frac{dT_j}{ds} + x_\theta^{*-} \frac{d\theta}{ds},$$

where the superscript $+$ indicates that we are taking limits from the interior of $\mathcal{R}_{j-1}$, and the superscript $-$ indicates limits from the interior of $\mathcal{R}_j$. Equation (12.6) also holds

for $j = 0$, without the superscripts $+$ and $-$. From (12.5), (10.2) – (10.5), and standard theorems on the differentiation of solutions of differential equations with respect to initial data, we get

$$(12.7) \qquad \frac{dT_1}{ds} = t_{1\sigma}\frac{d\sigma}{ds}, \quad x^*_\theta(t_1,\theta(s)) = (-M,\ 0),$$

where M is the matrix (8.1) and 0 is the n by $(n-p)$ zero matrix.

We now consider W along Γ. From (12.1) we obtain

$$W(T_0(s),X_0(s)) =$$

$$g(\sigma(s)) + \int_{T_0(s)}^{T_1(s)} f(t,x^*(t,\theta(s)),u^*(t,\theta(s)))dt.$$

Hence dW/ds exists and is given by

$$\frac{dW}{ds} = \left[g_\sigma \frac{d\sigma}{ds} + f\frac{dT_1}{ds}\right]_{T_1(s)} - \left[f\frac{dT_0}{ds}\right]_{T_0(s)}$$

(12.8)

$$+ \sum_{j=2}^{\alpha}\left[(f^- - f^+)\frac{dT_j}{ds}\right]_{T_j(s)} + \int_{T_0(s)}^{T_1(s)} \frac{\partial f}{\partial s}\,dt,$$

where $\partial f/\partial s = (f_x x^*_\theta + f_u u^*_\theta)(d\theta/ds)$, the superscripts $+$ and $-$ have the same meaning as in (12.6), and the arguments of the functions are $(t,\theta(s))$. From (6.2) we get

$f_x = -(\lambda_t + \lambda G_x + \mu R_x)$, and from (6.3) we have $f_u = -(\lambda G_u + \mu R_u)$. Hence we obtain

$$\frac{\partial f}{\partial s} = -\Big[\lambda_t\, x^*_\theta + \lambda(G_x\, x^*_\theta + G_u\, u^*_\theta) + \mu(R_x\, x^*_\theta + R_u\, u^*_\theta)\Big] \frac{d\theta}{ds}\,. \tag{12.9}$$

The components of the vector $\mu(R_x\, x^*_\theta + R_u\, u^*_\theta)$ can be written as follows:

$$\sum_{k=1}^{r} \mu^k(R^k_x\, x^*_{\theta^i} + R^k_u\, u^*_{\theta^i}) \equiv \sum_{k=1}^{r} \mu^k \frac{\partial R^k}{\partial \theta^i}, \quad i = 1, \ldots, n. \tag{12.10}$$

If at a point (t,θ) in Ω, we have $R^k(t,x^*(t,\theta),u^*(t,\theta)) > 0$, then from (6.4) we obtain $\mu^k(t,\theta) = 0$. On the other hand, since $R^k(t,x^*(t,\theta),u^*(t,\theta)) \geq 0$ on Ω, if $R^k = 0$ at (t,θ) then R^k, as a function of (t,θ), has a minimum at this point. Since (t,θ) is interior to Ω, $\partial R^k/\partial\theta^i$ vanishes at this point for all $i = 1, \ldots, n$. Hence (12.10) is zero for all (t,θ).

If we set $\partial G/\partial\theta = (G_x\, x^*_\theta + G_u\, u^*_\theta)$, then from (10.5) we have $x^*_{t\theta} = \partial G/\partial\theta$. Furthermore, we have $x^*_{t\theta} = x^*_{\theta t}$. Hence we may write (12.9) as

$$\frac{\partial f}{\partial s} = -(\lambda\, x^*_\theta)_t \frac{d\theta}{ds}\,.$$

Substituting this expression into the integral in (12.8) and performing the integration gives

$$- \left[\lambda\, x_\theta^* \frac{d\theta}{ds}\right]_{T_0(s)}^{T_1(s)} - \sum_{j=2}^{\alpha} \left[(\lambda\, x_\theta^*)^- - (\lambda\, x_\theta^*)^+\right]_{T_j(s)} \frac{d\theta}{ds} .$$

If we now use (12.6), (12.7), and the relation $x_t = G$, and substitute the resulting expression into (12.8), using the definition of H, we get

$$\frac{dW}{ds} = \left[(g_\sigma + Ht_{1\sigma} - \lambda x_{1\sigma}) \frac{d\sigma}{ds}\right]_{T_1(s)} - \left[H \frac{dT_0}{ds} - \lambda \frac{dX_0}{ds}\right]_{T_0(s)}$$

$$+ \sum_{j=2}^{\alpha} \left[(H^- - H^+) \frac{dT_j}{ds} - (\lambda^- - \lambda^+) \frac{dX_j}{ds}\right]_{T_j(s)} .$$

From (6.5) we see that the first square bracket vanishes. From Theorem 2 and (9.1) it follows that every square bracket in $\sum_{j=2}^{\alpha}$ vanishes. Hence, since Γ is arbitrary, we have

$$(12.11) \qquad dW = - H\, dT + L\, dX$$

for arbitrary differentials (dT, dX). The theorem is an immediate consequence of (12.11), the properties of f, G, L, and U^*, Theorem 2, and (9.1).

13. AN EQUATION OF DYNAMIC PROGRAMMING

For each (t,x) in $\mathcal{R}$, let $\mathcal{A}(t,x)$ denote the set of admissible controls u at (t,x). Since U^* is a normal synthesis, it follows from (6.6) that for any (t,x) in $\mathcal{R}_i$, $i = 1, \ldots, \alpha$, we have

$$(13.1) \qquad H(t,x,L(t,x)U^*(t,x)) = \min_{u \in \mathcal{A}(t,x)} H(t,x,L(t,x)u).$$

If we apply (12.2) to (13.1), we see that on $\mathcal{R}_i$,

$$(13.2) \qquad W_t = - \min_{u \in \mathcal{A}(t,x)} [f(t,x,u) + W_x G(t,x,u)].$$

If (t,x) lies on a manifold $\mathcal{R}_i$, $i = 1, \ldots, \alpha$, then the relations (13.1) and (13.2) hold for the one-sided limits.

Equation (13.2) is the functional equation obtained formally by Bellman [1] for control problems in which $\mathcal{T}$ is the n-dimensional manifold t_1 = constant and f, G, and R are C''. We note that (13.2) holds for more general problems than these. Since (13.1) is a restatement of the Weierstrass condition, since (12.2) says that on each $\mathcal{R}_i$, W satisfies the Hamilton-Jacobi equation, and since Pontryagin's principle derives from the Weierstrass condition, the relationship between these items and (13.2) is clear.

We remark that computational schemes based directly on (13.1) in the case that $\mathcal{T}$ is of dimension p, with $p < n$, will encounter difficulties because, in general, W_t and W_x do not exist at $\mathcal{T}$. (See Remark 2, Theorem 4.)

14. THE PROBLEM OF SYNTHESIS

Let $u^*(t,\theta)$ and $x^*(t,\theta)$ be as in Assumptions 1 to 4, and let us replace Assumption 5 by the following:

Assumption 6. Along each $K(\theta)$, let equations (6.1) – (6.5) hold with $\lambda_0 = 1$, and let $\lambda(t,\theta)$ and $\mu(t,\theta)$ have

the properties described in Theorem 2. Let the functions $\lambda_1(\theta)$ and $\mu_1(\theta)$ be as in Assumption 5. Let the function H be such that (9.1) holds for all manifolds $\mathcal{M}_j$, $j = 2, \ldots, \alpha$.

Assumption 6 consists of those consequences of Assumption 5 that enter into the discussion of Sec. 10. Hence, if we now look upon W as being defined by (12.1), then Theorem 4 still holds. In particular, (12.11) holds. Moreover, if we take Γ to lie entirely on a manifold $\mathcal{M}_i$, $i = 2, \ldots, \alpha + 1$, then the arguments used to establish (12.11) for Γ in some $\mathcal{R}_i$ show that (12.11) holds for Γ on a manifold $\mathcal{M}_i$, $i \geq 2$. For curves Γ on $\mathcal{J} = \mathcal{M}_1$, the validity of (12.11) follows from (6.5). Hence the integral

$$(14.1) \qquad \int_\Gamma H(t,x,L\ U^*)dT - L(t,x)dX$$

is independent of path in $\mathcal{R}$ for all curves Γ consisting of a finite number of arcs, each arc lying entirely in some $\mathcal{R}_i$ or on a manifold $\mathcal{M}_i$, $i = 1, \ldots, \alpha$.

From the preceding discussion we see that Assumptions 1 to 4 and Assumption 6 determine for the control problem the analogue of a field in the calculus of variations, with (14.1) as the Hilbert invariant integral. The following theorem can now be established by using the same argument as is used for the analogous fundamental sufficiency theorem in the calculus of variations:

Theorem 5. Let Assumptions 1 to 4 and Assumption 6 hold. Furthermore, let (6.6) hold on $\mathcal{R}$ for $u^* = U(t,x)$. Then $u^*(t,\theta)$ is a normal parametric synthesis of the control and $U^*(t,x)$ is a normal synthesis of the control.

REFERENCES

1. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

2. Bellman, R., I. Glicksberg, and O. Gross, "On the 'bang-bang' control problem," Q. Appl. Math., Vol. 14, 1956, pp. 11-18.

3. Bliss, G. A., Lectures on the Calculus of Variations, University of Chicago Press, Chicago, Illinois, 1946.

4. Breakwell, J. V., "The optimization of trajectories," J. Soc. Indust. Appl. Math., Vol. 7, 1959, pp. 215-247.

5. Bushaw, D. W., "Optimal discontinuous forcing terms," Contributions to the Theory of Nonlinear Oscillations, Vol. IV, pp. 29-52, Princeton University Press, Princeton, New Jersey, 1958.

6. Gamkrelidze, R. V., "Theory of time-optimal processes for linear systems," Izvestia Akad. Nauk, SSSR, Vol. 22, 1958, pp. 449-474 (Russian).

7. Hestenes, M. R., A General Problem in the Calculus of Variations with Applications to Paths of Least Time, The RAND Corporation, Research Memorandum RM-100, February 1949.

8. Krasovskii, N. N., "On the theory of optimum regulation," Avtomat. i Telemek., Vol. 18, 1957, pp. 960-970 (Russian); translated as Automation and Remote Control, Vol. 18, pp. 1005-1016.

9. Krasovskii, N. N., "On a problem of optimum control of non-linear systems," Prikl. Mat. Meh., Vol. 23, 1959, pp. 209-229 (Russian); translated as J. Appl. Math. Mech., Vol. 23, pp. 303-332.

10. LaSalle, J. P., "The time optimal control problem," Contributions to the Theory of Nonlinear Oscillations, Vol. V, pp. 1-24, Princeton University Press, Princeton, New Jersey, 1960.

11. McShane, E. J., "On multipliers for Lagrange problems," Amer. J. Math., Vol. 61, 1939, pp. 809-819.

12. McShane, E. J., "Necessary conditions in generalized-curve problems of the calculus of variations," Duke Math. J., Vol. 7, 1940, pp. 1-27.

13. Pontryagin, L. S., "Optimal control processes," Uspehi Mat. Nauk, Vol. 14, 1959, pp. 3–20.

14. Rozonoèr, L. I., "L. S. Pontryagin's maximum principle in the theory of optimum systems, I, II, III," Avtomat. i Telemeh., Vol. 20, 1959, pp. 1320–1334, 1441–1458, 1561–1578 (Russian); translated as Automation and Remote Control, Vol. 20, pp. 1288–1302, 1405–1421, 1517–1532.

15. Valentine, F. A., "The problem of Lagrange with differential inequalities as added side conditions," Contributions to the Calculus of Variations, 1933–1937, pp. 407–448, The University of Chicago Press, Chicago, Illinois, 1937.