

# The Calculus of Variations and Optimal Control

An Introduction

George Leitmann

## Preface

When the Tyrian princess Dido landed on the North African shore of the Mediterranean sea she was welcomed by a local chieftain. He offered her all the land that she could enclose between the shoreline and a rope of knotted cowhide. While the legend does not tell us, we may assume that Princess Dido arrived at the correct solution by stretching the rope into the shape of a circular arc and thereby maximized the area of the land upon which she was to found Carthage. This story of the founding of Carthage is apocryphal. Nonetheless it is probably the first account of a problem of the kind that inspired an entire mathematical discipline, the calculus of variations and its extensions such as the theory of optimal control.

This book is intended to present an introductory treatment of the calculus of variations in Part I and of optimal control theory in Part II. The discussion in Part I is restricted to the simplest problem of the calculus of variations. The topic is entirely classical; all of the basic theory had been developed before the turn of the century. Consequently the material comes from many sources; however, those most useful to me have been the books of Oskar Bolza and of George M. Ewing. Part II is devoted to the elementary aspects of the modern extension of the calculus of variations, the theory of optimal control of dynamical systems. Here the approach is not variational but rather geometric; it is based on a theory developed in collaboration with Austin Blaquièrre of the University of Paris.

This volume is the outgrowth of lecture notes for a course on the variational calculus and optimal control which has been taught at the University of California at Berkeley for over twenty years. Based on this experience, I believe that a first-year graduate student in an engineering or

applied science curriculum should possess the requisite mathematical sophistication required for a reading of this text.

Over the years I have benefited greatly from fruitful discussions with many colleagues and students, too numerous to list here; they know who they are. However, two of them merit special mention. I am deeply grateful to Martin Corless and to Wolfram Stadler for their critical reading of the manuscript and for their constructive suggestions. I am also indebted to David G. Luenberger, William E. Schmitendorf, and Thomas L. Vincent for allowing me to quote from their work in Sections 13.12, 15.8, and 15.9 of the book.

George Leitmann

## Symbols and Notation

Standard mathematical symbols and notation are used in this book. The most commonly used symbols are defined first. Thereafter we give the definitions of the basic notation employed in the text.

### Symbols

$\triangleq$	equals by definition; denotes
$=$	equals, is equivalent to
$\neq$	does not equal; is not equivalent to
$\equiv$	equals identically; is the same as
$\not\equiv$	does not equal identically; is not the same as
$\leq (\geq)$	is less (greater) than or equal to
$< (>)$	is less (greater) than
$\forall$	for all, for every
$\in$	is an element (member) of; belongs to
$\notin$	is not an element (member) of; does not belong to
$\emptyset$	empty set
$\subset$	is a subset of; is contained in
$\supset$	contains
$\cup$	union
$\cap$	intersection
$\times$	Cartesian product
$\{e P\}$	set of all $e$ having property $P$
$R^1$	set of real numbers; real line
$[a, b]$	$\{x \in R^1   a \leq x \leq b\}$

$(a, b)$	$\{x \in R^1   a < x < b\}$
$[a, b)$	$\{x \in R^1   a < x \leq b\}$
$[a, b]$	$\{x \in R^1   a \leq x < b\}$
$\setminus$	subtraction of sets; that is, $A \setminus B \triangleq \{e   e \in A, e \notin B\}$
$  $	absolute value
$   $	Euclidean norm
$\inf(\sup)$	infimum (supremum)
$\min(\max)$	minimum (maximum)
$\text{sgn } x$	signum; that is, for $x \in R^1$ , $\text{sgn } x = 1$ if $x > 0$ , $\text{sgn } x = -1$ if $x < 0$
$T$	transpose (superscript)

## Spaces

The set of all ordered  $n$ -tuples of real numbers is denoted by  $R^n$ ; that is,

$$R^n \triangleq R^1 \times R^1 \times \cdots \times R^1 \quad (n \text{ times}).$$

Thus, given an ordered  $n$ -tuple of real numbers,  $x_1, x_2, \dots, x_n$ , we consider it to be a *vector*  $x \in R^n$ . We let all vectors be column vectors; that is,

$$x \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \ x_2 \ \cdots \ x_n]^T.$$

By endowing  $R^n$  with the *natural basis*  $\{e^1, e^2, \dots, e^n\}$ , where  $e^i \in R^n$  and

$$e^{iT}e^j = \delta_{ij}, \quad \delta_{ij} \triangleq \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j, \end{cases}$$

we assure that  $R^n$  is a Euclidean space. In particular it follows that

$$x^T x = \sum_{i=1}^n x_i^2 = \|x\|^2.$$

The *closure of a set*  $X \subset R^n$ , denoted by  $\bar{X}$ , is the set together with all of its accumulation points; that is,

$$\bar{X} \triangleq X \cup \{x | \text{there is a sequence } x_i \in X, i=1, 2, \dots, \\ \text{such that } x_i \text{ converges to } x\}.$$

## Functions

Given the nonempty sets  $X$  and  $Y$ ,

$$f(\cdot): X \rightarrow Y$$

denotes a *function* (mapping) from the domain  $X$  into the range  $Y$ ; that is, associated with  $x \in X$  there is one and only one  $y \in Y$ . We write  $y = f(x)$ , and we term  $f(x)$  the value of the function at  $x$ .

The scalar-valued function  $f(\cdot): [a, b] \rightarrow R^1$ ,  $[a, b] \subset R^1$ ,  $a < b$ , is of class  $C^k$  if and only if it and its first  $k$  derivatives are continuous on  $[a, b]$ . Such a function of class  $C^1$  is also called *smooth*.

The scalar-valued function  $f(\cdot): [a, b] \rightarrow R^1$ ,  $[a, b] \subset R^1$ ,  $a < b$ , is *piecewise continuous* if and only if it is continuous on  $[a, b]$  with the exception of a finite number of points of  $(a, b)$  where it possesses defined left and right limits; that is, if  $f(\cdot)$  is discontinuous at  $\bar{x} \in (a, b)$ , then

$$f(\bar{x}-0) \triangleq \lim_{\substack{x \rightarrow \bar{x} \\ x < \bar{x}}} f(x)$$

and

$$f(\bar{x}+0) \triangleq \lim_{\substack{x \rightarrow \bar{x} \\ x > \bar{x}}} f(x)$$

are defined. In order to have  $f(x)$  defined for all  $x \in [a, b]$ , we take

$$f(a) = f(a+0),$$

$$f(b) = f(b-0),$$

and if  $f(\cdot)$  is discontinuous at  $\bar{x} \in (a, b)$  we take

$$f(\bar{x}) = f(\bar{x}-0)$$

or

$$f(\bar{x}) = f(\bar{x}+0).$$

The function  $f(\cdot): [a, b] \rightarrow R^1$ ,  $[a, b] \subset R^1$ ,  $a < b$ , is *piecewise smooth* if and only if it is continuous and its first derivative is piecewise continuous on  $[a, b]$ . If the first derivative is discontinuous at  $\bar{x} \in (a, b)$ , then the point  $(x, y) = (\bar{x}, f(\bar{x}))$  is termed a *corner* of  $f(\cdot)$ .

The same notation is used for a vector-valued function  $f(\cdot): [a, b] \rightarrow \mathbb{R}^n$ ,  $[a, b] \subset \mathbb{R}^1$ ,  $a < b$ , provided the appropriate conditions are satisfied by its components which are scalar-valued; for instance,  $f(\cdot): [a, b] \rightarrow \mathbb{R}^n$  is of class  $C^k$  if and only if the functions  $f_i(\cdot): [a, b] \rightarrow \mathbb{R}^1$ ,  $i = 1, 2, \dots, n$ , are of class  $C^k$ , where  $f(x) \triangleq [f_1(x) \ f_2(x) \ \cdots \ f_n(x)]^T$ .

The function  $f(\cdot): X \rightarrow \mathbb{R}^1$ ,  $X \subset \mathbb{R}^m$ , is of class  $C^k$  if and only if it and its partial derivatives up to and including order  $k$  are continuous on  $X$ .

Given a function  $f(\cdot): X \rightarrow Y$  and  $Z \subset X$ , the *restriction* of  $f(\cdot)$  to  $Z$ , denoted by  $f(\cdot)|_Z$ , is the function  $f(\cdot)|_Z: Z \rightarrow Y$  such that  $f(x)|_Z = f(x)$  for all  $x \in Z$ .

Consider the function  $o(\cdot): [a, b] \rightarrow \mathbb{R}^n$ ,  $[a, b] \subset \mathbb{R}^1$ ,  $a < 0$ ,  $b > 0$ , such that

$$\lim_{x \rightarrow 0} \frac{o(x)}{x} = 0.$$

Every function having this property is denoted by  $o(\cdot)$ .

The function  $f(\cdot): [a, b] \rightarrow \mathbb{R}^n$ ,  $[a, b] \subset \mathbb{R}^1$ , is *convex* if and only if, for every  $x$  and  $y$  in  $[a, b]$  and for every  $\alpha \in [0, 1]$ , we have

$$f(z) \leq \alpha f(y) + (1 - \alpha)f(x),$$

where

$$z = \alpha y + (1 - \alpha)x.$$

Finally, given a function  $f(\cdot): X \rightarrow \mathbb{R}^1$ ,  $X \subset \mathbb{R}^n$ , that is differentiable at  $x \in X$ , we let

$$\text{grad } f(x) \triangleq \left[ \frac{\partial f(x)}{\partial x_1} \quad \frac{\partial f(x)}{\partial x_2} \quad \cdots \quad \frac{\partial f(x)}{\partial x_n} \right]^T.$$

# Contents

Symbols and Notation .....	xi
----------------------------	----

## PART I. CALCULUS OF VARIATIONS

<b>1. Introduction .....</b>	<b>3</b>
<b>2. Problem Statement and Necessary Conditions for an Extremum.....</b>	<b>7</b>
2.1. Introduction .....	7
2.2. The Simplest Problem with Fixed End Points .....	7
2.3. Minima of Ordinary Functions .....	8
2.4. Local Minima of Integrals .....	11
2.5. The DuBois-Reymond Lemma .....	14
2.6. The Necessary Condition of Euler .....	15
2.7. Remarks .....	19
2.8. Stationarity .....	22
<i>Exercises</i> .....	23
<b>3. Integration of the Euler-Lagrange Equation .....</b>	<b>25</b>
3.1. Introduction .....	25
3.2. The Function $f(\cdot)$ Independent of $x$ : $f_x(t, x, r) \equiv 0$ .....	25
3.3. The Function $f(\cdot)$ Independent of $t$ : $f_t(t, x, r) \equiv 0$ .....	26
3.4. The Function $f(\cdot)$ Independent of $t$ and $x$ : $f_t(t, x, r) \equiv 0$ and $f_x(t, x, r) \equiv 0$ .....	27
3.5. Examples .....	27
3.6. Remarks .....	31
3.7. The Function $f(\cdot)$ Linear in $r$ : $f_{rr}(t, x, r) \equiv 0$ .....	33
<i>Exercises</i> .....	37
<b>4. An Inverse Problem .....</b>	<b>39</b>
4.1. Problem Statement .....	39
4.2. Solution .....	39



4.3. Examples .....	41
<i>Exercises</i> .....	44
<b>5. The Weierstrass Necessary Condition</b> .....	<b>47</b>
5.1. Introduction .....	47
5.2. The Excess Function and a Necessary Condition .....	47
5.3. Example .....	50
5.4. The Legendre Necessary Condition .....	51
5.5. Example .....	52
<i>Exercises</i> .....	53
<b>6. Jacobi's Necessary Condition</b> .....	<b>55</b>
6.1. Introduction .....	55
6.2. The Accessory Minimum Problem .....	55
6.3. The Integration of Jacobi's Equation .....	60
6.4. Example .....	64
<i>Exercises</i> .....	65
<b>7. Corner Conditions</b> .....	<b>67</b>
7.1. Necessary Conditions .....	67
7.2. Example .....	68
<i>Exercises</i> .....	70
<b>8. Concluding Remarks</b> .....	<b>71</b>

## PART II. OPTIMAL CONTROL

<b>9. Introduction</b> .....	<b>77</b>
<b>10. Problem Statement and Optimality</b> .....	<b>79</b>
10.1. Introduction .....	79
10.2. Problem Statement .....	79
10.3. Joining Controls and Additivity of Costs .....	83
10.4. Optimal Cost and an Optimality Principle .....	85
10.5. Augmented State Space and Trajectories .....	88
10.6. Limiting and Optimal Isocost Surfaces .....	90
10.7. Fundamental Properties of Trajectories and of Limiting Surfaces .....	92
10.8. An Illustrative Example .....	95
<i>Exercises</i> .....	98
<b>11. Regular Optimal Trajectories</b> .....	<b>99</b>
11.1. Regular Interior Points of a Limiting Surface .....	99
11.2. Necessary Conditions at a Regular Interior Point .....	105
11.3. A Linear Transformation .....	107

11.4. Transformation of the Tangent Plane.....	109
11.5. Regular Optimal Trajectories .....	111
11.6. An Illustrative Example.....	114
11.7. The Terminal Transversality Condition .....	115
11.8. A Maximum Principle .....	118
11.9. Remarks.....	119
11.10. Extremal Control .....	121
11.11. An Illustrative Example .....	122
<i>Exercises</i> .....	123
<b>12. Examples of Extremal Control.....</b>	<b>125</b>
12.1. Time-Optimality for a Constant-Power Rocket .....	125
12.2. A Problem of Time-Optimal Navigation .....	130
12.3. The Minimum Distance to a Given Curve.....	135
<i>Exercises</i> .....	138
<b>13. Some Generalizations .....</b>	<b>139</b>
13.1. Introduction .....	139
13.2. The Initial Transversality Condition .....	140
13.3. Minimum Distance between Given Curves .....	144
13.4. A Terminal Cost .....	145
13.5. An Extremal Thrust Control for a Rocket.....	150
13.6. Nonautonomous Systems .....	155
13.7. Problems with a Fixed Interval .....	160
13.8. A Minimum Fuel Rendezvous of a Constant-Power Rocket .....	161
13.9. The Simplest Problem of the Calculus of Variations .....	164
13.10. State-Dependent Control Constraints .....	167
13.11. A Time-Optimal Regulator with Velocity-Dependent Control Bounds .....	174
13.12. Isoperimetric Constraints.....	181
13.13. Dido's Problem .....	188
13.14. Parameter Optimization .....	190
13.15. An Illustrative Example .....	195
13.16. End Point Inequality Constraints .....	196
13.17. An Illustrative Example .....	198
<i>Exercises</i> .....	205
<b>14. Special Systems.....</b>	<b>211</b>
14.1. Introduction .....	211
14.2. Linear Time-Invariant State Equations .....	211
14.3. The Switching Function.....	212
14.4. Time-Optimality and Bang-Bang Control .....	213
14.5. The Number of Switches .....	215
14.6. A One-Dimensional Time-Optimal Regulator.....	217
14.7. A Three-Dimensional Time-Optimal Regulator .....	222
14.8. Singular Control .....	225
14.9. The Maximum Range of a Thrust-Limited Rocket .....	226

14.10. The Maximum Range of a Rocket in Horizontal Flight.....	231
<i>Exercises</i> .....	237
<b>15. Sufficient Conditions</b> .....	<b>241</b>
15.1. Introduction .....	241
15.2. A Field Theorem .....	241
15.3. An Illustrative Example.....	244
15.4. Another Field Theorem .....	245
15.5. Time-Optimality for a Constant-Power Rocket .....	247
15.6. A Direct Sufficiency Theorem .....	249
15.7. An Illustrative Example.....	251
15.8. Life History Strategies of Plants .....	252
15.9. An Economic Control Problem .....	256
<i>Exercises</i> .....	262
<b>16. Feedback Control</b> .....	<b>265</b>
16.1. Introduction .....	265
16.2. The Synthesis of Optimal Feedback Control .....	268
16.3. The Linear-Quadratic Problem .....	269
16.4. The Existence of Feedback Solutions .....	275
16.5. An Illustrative Example.....	278
<i>Exercises</i> .....	283
<b>17. Optimization with Vector-Valued Cost</b> .....	<b>285</b>
17.1. Introduction .....	285
17.2. How to Choose a Cheese .....	286
17.3. Pareto-Optimal Control.....	292
17.4. Necessary Conditions for Pareto-Optimality .....	293
17.5. Sufficient Conditions for Pareto-Optimality .....	295
17.6. An Illustrative Example.....	296
<i>Exercises</i> .....	298
<b>REFERENCES</b> .....	<b>301</b>
<b>BIBLIOGRAPHY</b> .....	<b>305</b>
<b>INDEX</b> .....	<b>309</b>

# The Calculus of Variations and Optimal Control

## An Introduction

George Leitmann

*University of California  
Berkeley, California*

PLENUM PRESS · NEW YORK AND LONDON

---

Library of Congress Cataloging in Publication Data

Leitmann, George.

The calculus of variations and optimal control.

(Mathematical concepts and methods in science and engineering; v. 24)

Bibliography: p.

Includes index.

1. Calculus of variations. 2. Control theory. I. Title. II. Series.

QA315.L36

515'.64

81-4582

ISBN 0-306-40707-8

AACR2

---

© 1981 Plenum Press, New York  
A Division of Plenum Publishing Corporation  
233 Spring Street, New York, N.Y. 10013

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher

Printed in the United States of America

# Part I

## Calculus of Variations

1

### Introduction

All of us “know” the answer to the question: What is the shape of the shortest plane curve connecting two given points? Of course it is a straight line. In mathematical terms one may pose this question as follows. Consider the family of all piecewise smooth functions

$$y(\cdot): [x_0, x_1] \rightarrow R^1, \quad x_0 < x_1,$$

satisfying

$$y(x_0) = y_0, \quad y(x_1) = y_1,$$

where  $x_0, x_1, y_0,$  and  $y_1$  are prescribed. Find a function  $y^*(\cdot)$  in the family defined above that yields the curve of minimum length joining points  $(x_0, y_0)$  and  $(x_1, y_1)$ . For given  $y(\cdot)$ , the length of the curve is

$$\int_{x_0}^{x_1} [1 + (dy/dx)^2]^{1/2} dx \quad (1.1)$$

so that we seek  $y^*(\cdot)$  such that

$$\int_{x_0}^{x_1} [1 + (dy^*/dx)^2]^{1/2} dx \leq \int_{x_0}^{x_1} [1 + (dy/dx)^2]^{1/2} dx \quad (1.2)$$

for all  $y(\cdot)$  in the class specified above.

A less trivial and considerably more difficult problem is that of determining the thrust program which results in maximizing the flight distance or range of a rocket plane in horizontal flight; see also Exercise 3.6

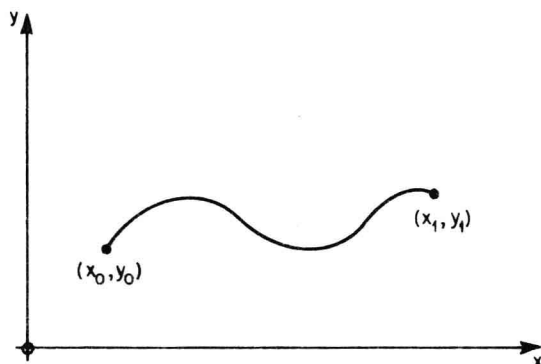


Figure 1.1. A curve connecting given points.

and Section 14.10. Let

$t$  = time,

$x$  = horizontal distance (range),

$v$  = speed,

$m$  = mass,

$T = -c \, dm/dt$  = thrust,  $c = \text{constant} > 0$ ,

$L$  = lift,

$D$  = drag.

We assume that the lift,  $L$ , is adjusted to balance the weight,  $mg$ ,  $g = \text{constant} > 0$ , so that the rocket moves horizontally. We assume further that the aerodynamic drag depends on the speed and lift as follows (Ref. 1.1):

$$D = Av^2 + BL^2, \quad A \text{ and } B = \text{constants} > 0.$$

Then the equations of motion of the rocket are

$$\begin{aligned} \frac{dx}{dt} &= v, \\ m \frac{dv}{dt} &= c\beta - D, \quad D = Av^2 + Bg^2 m^2, \\ \frac{dm}{dt} &= -\beta. \end{aligned} \tag{1.3}$$

Now, given the initial and terminal values of the speed  $v$  and mass  $m$  (and

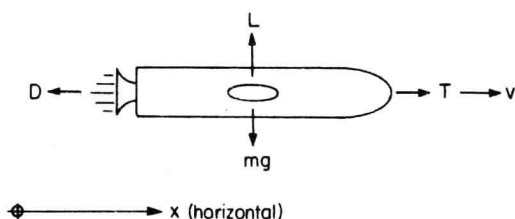


Figure 1.2. Forces acting on a rocket.

hence of the fuel consumed), we wish to determine the thrust program—that is, how  $T$  must be varied—in order to maximize the range  $x(t_1) - x(0)$ .

From (1.3) one obtains

$$dx = -\frac{cv}{D} \left[ dm + \frac{m}{c} dv \right]$$

so that the range is

$$x(t_1) - x(0) = \int_{m_1}^{m_0} \frac{cv}{D} \left[ 1 + \frac{m}{c} \frac{dv}{dm} \right] dm \quad (1.4)$$

where  $m(0) = m_0$ ,  $m(t_1) = m_1 < m_0$ ,  $v(0) = v_0$  and  $v(t_1) = v_1$  are prescribed. We may now restate the problem more simply; namely, determine the speed  $v$  as a function of the mass  $m$ , satisfying the given end conditions and maximizing the value of the integral (1.4). From (1.3) one then has

$$T = c\beta = \frac{cD}{c + m \, dv/dm} \quad (1.5)$$

yielding the thrust program as a function of  $m$  and  $v$ , and thence of  $m$ .

These examples typify the simplest problems of the calculus of variations. In the next chapter we state such problems in a general way.



