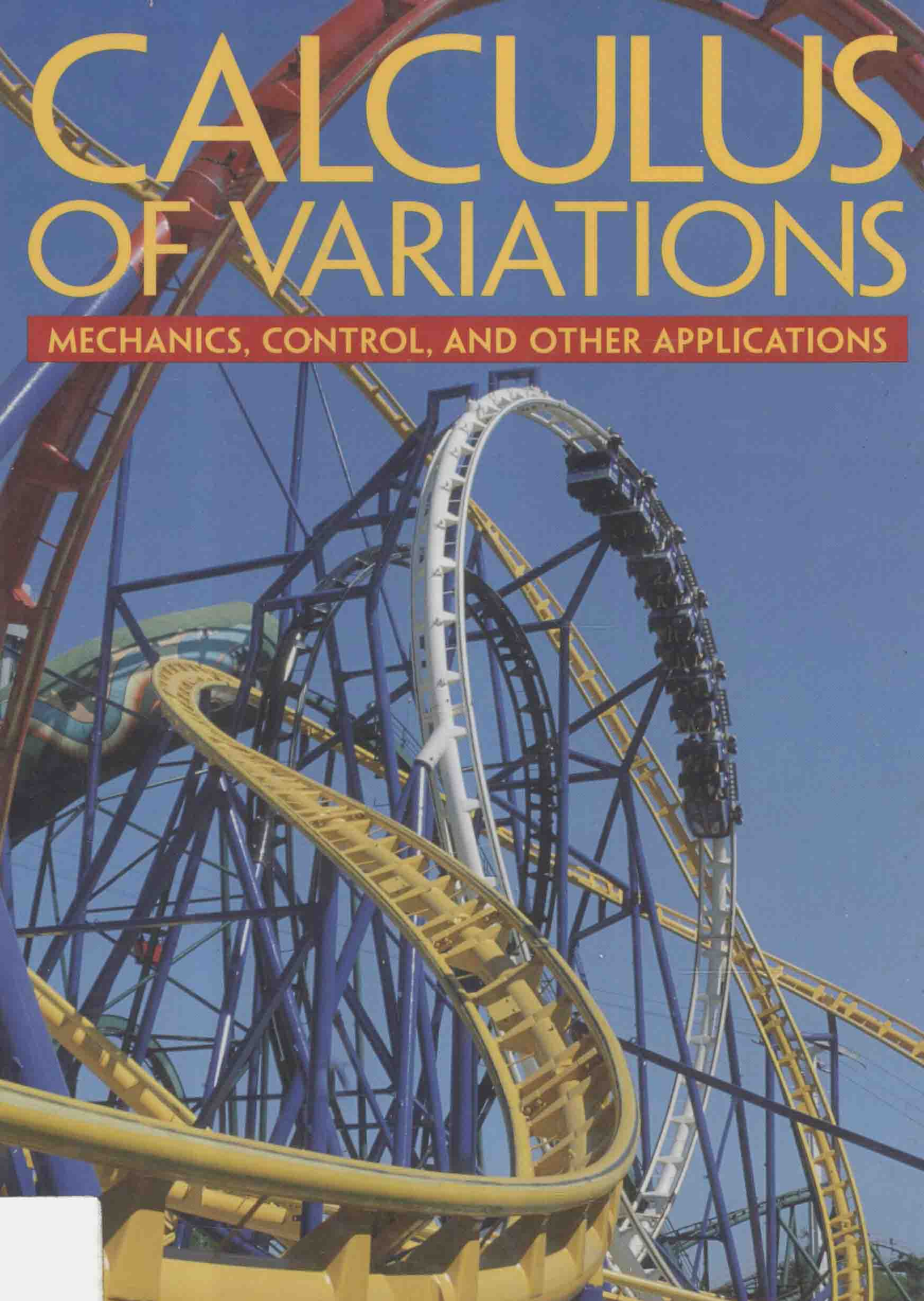
CALCULUS OF VARIATIONS

MECHANICS, CONTROL, AND OTHER APPLICATIONS



HARLES R. MACCLUER

Calculus of Variations

Mechanics, Control, and Other Applications

Charles R. MacCluer

Michigan State University

PEARSON
Prentice Hall

Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publications Data

MacCluer, Charles R.
Calculus of variations: mechanics, control, and other applications/
Charles R. MacCluer.
p. cm
Includes index.
ISBN 0-13-142383-5
1. Calculus of variations. I. Title.

QA315.M23 2005
515'.64—dc22 2004040069

Executive Acquisitions Editor: *George Lobell*
Editor-in-Chief: *Sally Yagan*
Vice President/Director of Production and Manufacturing: *David W. Riccardi*
Production Editor: *Debbie Ryan*
Senior Managing Editor: *Linda Mihatov Behrens*
Assistant Managing Editor: *Bayani Mendoza de Leon*
Executive Managing Editor: *Kathleen Schiaparelli*
Assistant Manufacturing Manager/Buyer: *Michael Bell*
Manufacturing Manager: *Trudy Pisciotti*
Marketing Manager: *Halee Dinsey*
Marketing Assistant: *Rachael Beckman*
Cover Designer: Bruce Kenselaar
Creative Director: *Jayne Conte*
Director of Creative Services: *Paul Belfanti*
Editorial Assistant: *Jennifer Brady*
Front Cover Image: *Susan Van Etten/Stock Boston*
Back Cover Image: *Jim Olive/Peter Arnold Inc.*

PEARSON Prentice Hall

Pearson Prentice Hall
Pearson Education, Inc.
Upper Saddle River, NJ 07458

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN: 0-13-142383-5

Pearson Education, Ltd., *London*
Pearson Education Australia PTY. Limited, Sydney
Pearson Education Singapore, Pte., Ltd
Pearson Education North Asia Ltd, *Hong Kong*
Pearson Education Canada, Ltd., *Toronto*
Pearson Education de Mexico, S.A. de C.V.
Pearson Education - Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd

RELATED TITLES OF INTEREST

Partial Differentials with Fourier Series and Boundary Value Problems (0-13-148096-0) *Nakhle Asmar*

An Introduction to Dynamical Systems: Continuous and Discrete (0-13-143140-4) *R. Clark Robinson*

Modeling Differential Equations in Biology (0-13-017325-8) *Clifford Taubes*

The Theory of Differential Equations: Classical and Qualitative (0-13-102026-9) *Walter Kelley and Allan Peterson*

A First Course in the Qualitative Theory of Differential Equations (0-13-008380-1) *James Liu*

Applied Partial Differential Equations with Fourier Series and Boundary Value Problems, 4e (0-13-065243-1) *Richard Haberman*

A Very Applied First Course in Partial Differential Equations (0-13-030417-4) *Michael Keane*

Partial Differential Equations (0-13-674359-5) *Arthur David Snider*

Partial Differential Equations, 2e (0-13-009335-1) *Robert McOwen*

A First Course in Fourier Analysis (0-13-578782-3) *David Kammler*

A First Course in Wavelets with Fourier Analysis (0-13-022809-5) *Albert Boggess and Francis Narcowich*

Classical and Modern Fourier Analysis (0-13-035399-X) *Loukas Grafakos*

Introduction to the Mathematics of Medical Imaging (0-13-067548-2) *Charles Epstein*

Advanced Engineering Mathematics (0-13-321431-1) *Michael Greenberg*

Fundamentals of Complex Analysis with Applications to Engineering, Science, and Mathematics, 3e (0-13-907874-6) *E.B. Saff and A. David Snider*

Applied Complex Analysis with Partial Differential Equations (0-13-089239-4) *Nakhle Asmar*

Real Analysis with Real Applications (0-13-041647-9) *Ken Davidson and Alan Donsig*

Differential Equations and Linear Algebra (0-13-086250-9) *Farlow, Hall, McDill, and West*

Differential Equations and Linear Algebra, 2e (0-13-263757-X) *Stephen Goode*

Differential Equations and Linear Algebra (0-13-011118-X) *Michael Greenberg*

Multivariable Mathematics, 4e (0-13-067276-9) *Williamson/Trotter*

Differential Equations: An Applied Approach (0-13-044930-X) *J.M. Cushing*

Differential Equations (0-13-598137-9) *Polking, Boggess, Arnold*

Differential Equations with Boundary Value Problems (0-13-091106-2) *Polking, Boggess, Arnold*

Differential Equations: Computing & Modeling, 3e (0-13-067667-4) *Edwards/Penney*

Differential Equations and Boundary Value Problems: Computing & Modeling, 3e (0-13-065245-8) *Edwards/Penney*

Elementary Differential Equations, 5e (0-13-145773-X) *Edwards/Penney*

Elementary Differential Equations with Boundary Value Problems, 5e (0-13-145774-8) *Edwards/Penney*

Differential Equations: A Systems Approach (0-13-046026-5) *Bruce Conrad*

Differential Equations and Boundary Value Problems: A Systems Approach (0-13-093419-4) *Bruce Conrad*

Differential Equations with Boundary Value Problems (0-13-015927-1) *Selwyn Hollis*

Differential Equations with Graphical and Numerical Methods (0-13-084376-8) *Bernard Banks*

Differential Equations: Modeling with MATLAB (0-13-736539-X) *Paul Davis*

Elementary Differential Equations, 8e (0-13-508011-8) *Rainville/Bedient/Bedient*

This book is dedicated to my wife Ann.

Preface

About This Book

This book is aimed at the junior- or senior-level student of mathematics, science, and engineering. It can also be used as an amusing summer course for graduate students by a judicious use of the starred exercises and proofs. Chapters 1–7 form a leisurely undergraduate semester course.

The difficulty of the book ramps up gradually—Chapter 8 is at a strong senior level, while Chapters 9 and 10 (Weak and Strong Sufficiency) and Chapter 11 (Corner Points) are more abstract and at very strong senior or graduate level.

The charm of this subject is found in its classical applications accessible to any student with calculus. We have attempted to downplay (at first) the technical details, to instead develop technique. As a result, even a modestly equipped student can carry away a strong understanding of the subject based on practice with the calculations. The starred proofs employ advanced machinery but are sketched in an expository style that may be comprehensible to undergraduates.

Why This Book?

There is no modern text at this level that is accessible to students armed only with calculus. There are of course the fine classic Dover editions of Fox, Sagan, Weinstock, Ewing, and Gelfand/Fomin. But these books are all showing their age, and, unlike our book, none of these incorporate a simple introduction to optimal control, the bang-bang theorem, Pontryagin's maximum principle, or linear-quadratic control design. Some of the most entertaining applications of the calculus of variations are found in optimal control.

To the Instructor

At times much of the detail is thrown into the Exercises. This is to facilitate flow and better display the attractive big picture. You may include some of these solutions in your lectures or assign them in some proportion consonant with your degree of commitment to the Moore system. A disk of solutions is available upon request. Additions and corrections to the text will be updated at `http://www.math.msu.edu/~maccluer/PrenHall/additions.pdf`.

Acknowledgments

One great joy of University life is having living resources available for the mere asking. I thank my colleagues David E. Blair, William C. Brown, B-Y Chen, Leonid Freidovich, Milan Miklavcic, Boris Mordukhovich, Fedor Nazarov, Sheldon Newhouse, George Pappus, Jacob Plotkin, Clark Radcliffe, Elias Strangas, Ralph Svetic, Lal Tummala, Clifford Weil, Peter R. Wolenski, Lijian Yang, Vera Zeidan, and Zhengfang Zhou.

Many students helped shape this book. They suffered through the early write-ups and typos and often suggested valuable improvements. I thank the undergraduate students Daniel Brian Bouk, Lynne M. Evasic, Leonard Joseph Ford, Tanya Christine Hopkalo, Harold Leatherman Hunt, Rachel C. McCormick, Megan Jayne Mercer, John E. Mills, Christopher Thomas Morling, Jessica A. Munger, Jacquelyn M. Ormiston, Lindsay J. D. Radke, Stephanie L. Semann, Michael J. Stinson, Pieter C. vanRooyen, Julie K. Waibel, and Mengmeng Yu.

My special thanks go to the graduate students Michelle L. Boorom, Matthew T. Brenneman, Alberto A. Condori, Chinthaka V. Hettitantri, Ki-Moon Lee, Laura M. Stadelman, Steven W. Sy, Brian J. Vessell, and Jared Wasburn-Moses.

The reviewers of this book offered many helpful suggestions. For their insights I thank Mark Coffey of the Colorado School of Mines, Gregor Kovacic of Rensselaer Polytechnic Institute, Boris Mordukovich of Wayne State University, and Eduardo Sontag of Rutgers University.

Finally, I thank George Lobell, Executive Editor, Prentice Hall, for his continued help and support.

Charles R. MacCluer
maccluer@msu.edu

Contents

Preface ix
Acknowledgments x

1 Preliminaries
1.1 Directional Derivatives and Gradients 1
1.2 Calculus Rules 3
1.3 Contour Surfaces and Sublevel Sets 5
1.4 Lagrange Multipliers 7
1.5 Convexity 9
Exercises 12

2 Optimization
2.1 Mathematical Programming 17
2.2 Linear Programming 22
2.3 Statistical Problems 23
2.4 Variational Problems 24
Exercises 24

3 Formulating Variational Problems
3.1 Shortest Distance between Two Points (CVP 1) 31
3.2 Graph with Least Surface of Revolution (CVP 2) 32
3.3 The Catenary (CVP 3) 32
3.4 The Brachistochrone (CVP 4) 33
3.5 Cruise-Climb (CVP 5) 34
3.6 Shapes of Minimum Resistance (CVP 6) 35
3.7 Hamilton's Principle 36
3.8 Isoperimetric Problems 39
Exercises 40

4 The Euler-Lagrange Equation
4.1 One Degree of Freedom 45
4.2 Two Special Cases: No y, No x 48
4.3 Multiple Degrees of Freedom 52
4.4 The Hamiltonian 54
4.5 A Closer Look 58
Exercises 60

5 Constrained Problems
5.1 Dido's Problem 67
5.2 Statement of the Problem 68
5.3 The Inverse Function Theorem 68
5.4 The Euler-Lagrange Equation for Constrained Problems 69
5.5 Example Applications 71
5.6 Multiple Degrees of Freedom 76
5.7 Nonintegral Constraints 78
5.8 Hamilton's Principle with Constraints 81
Exercises 86

6 Extremal Surfaces
6.1 A Soap Film (CVP 15) 91
6.2 Stable Flows (CVP 17) 96
6.3 Schrödinger's Equation (CVP 18) 97
6.4 Eigenvalue Problems 98
6.5 Rayleigh-Ritz Numerics 99
Exercises 104

7 Optimal Control
7.1 A Rolling Cart (OCP 1) 109
7.2 General Formulation 112
7.3 Reinvestments (OCP 2) 113
7.4 Average Voltage (OCP 3) 115
7.5 A Time-Optimal Problem (OCP 4) 116
7.6 The Bang-Bang Principle 118
7.7 The Maximum Principle 120
7.8 Example Applications 124
Exercises 129

8 The LQ Problem
8.1 Problem Statement 139
8.2 State Feedback 142
8.3 Stability 143
8.4 The LQR Problem 146
8.5 A Tracking Servo 152
Exercises 156

9 Weak Sufficiency
9.1 Weak versus Strong Extrema 161
9.2 First and Second Variations 162
9.3 In Application 166
9.4 The Integrand $p\eta'^2 + q\eta^2$ 170
9.5 Weak Local Sufficiency 173
Exercises 177

10 Strong Sufficiency
10.1 The Goal 185
10.2 Flows 186
10.3 Flows of the Euler-Lagrange Equation 190
10.4 The E-Function and Strong Sufficiency 192
10.5 The Existence of Flows 196
Exercises 205

11 Corner Points
11.1 Corners and Extremals 209
11.2 First Erdmann Corner Condition 211
11.3 The Figurative 213
11.4 Second Erdmann Corner Condition 217
Exercises 221

Appendix A. The Inverse Function Theorem 223
Appendix B. Picard's Theorem 229
Appendix C. The Divergence Theorem 233
Appendix D. A MATLAB Cookbook 235
References 249
Index 251

Chapter 1

Preliminaries

This chapter reviews basic tools from calculus that are used in the calculus of variations—directional derivatives, gradients, the chain rules, contour surfaces, sublevel sets, Lagrange multipliers, and the basic notion of convexity. All of these concepts form the basic toolset for attacking optimization problems.

1.1 Directional Derivatives and Gradients

A point in $\mathbb{R}^n$ is denoted by $x = (x_1, x_2, \ldots, x_n)^\top$, which is an $n \times 1$ column vector (the superscript $^\top$ denotes transpose). Suppose a function f of x represents the profit of a commercial enterprise, where x is a vector of the parameters of the operation such as labor costs, production output levels, price of the commodity, and so on. The manager naturally desires to know in which direction from the present operating point x^0 should the company move in order to obtain the maximum increase in profit. The desired direction is found by using a multi-variable notion of a derivative: For each unit vector $u \in \mathbb{R}^n$, the *directional derivative* of f at x^0 in direction u is given by

$$D_u f(x^0) = \lim_{h \to 0+} \frac{f(x^0 + hu) - f(x^0)}{h}, \tag{1.1}$$

provided the limit exists. Geometrically, this limit is the slope of the line tangent to the curve above x^0 obtained by cutting the hypersurface $z = f(x)$ with the hyperplane determined by u and the z-axis. See Figure 1.1. The directional derivatives in the directions parallel to the coordinate axes are, of course, the familiar *partial derivatives*

$$D_u f(x^0) = \frac{\partial f(x^0)}{\partial x_k},$$

when $u = (0, 0, \ldots, 0, 1\ (k\text{th position}), 0, \ldots, 0)^\top$.

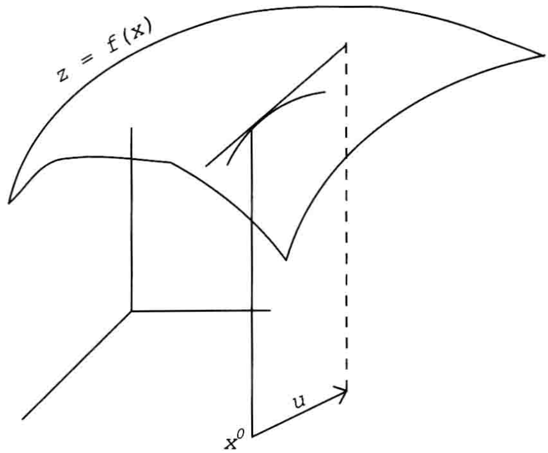


Figure 1.1 The graph of $y = f(x)$ is cut by a plane perpendicular to the coordinate space in the direction u. The slope of the line tangent to the curve thus cut out is given by the directional derivative $D_u f(x^0)$.

The function f is *differentiable* at x^0 provided it is linearly approximated by its tangent plane near x^0. This means there exists a constant (row) vector $a = (a_1, a_2, \ldots, a_n)$ such that for all x in some open ball $B = \{x;\ |x - x^0| < r\}$ of radius r about x^0,

$$f(x) = f(x^0) + a(x - x^0) + \epsilon(x) \tag{1.2a}$$

$$= f(x_1^0, x_2^0, \ldots, x_n^0) + a_1(x_1 - x_1^0) + a_2(x_2 - x_2^0) + \cdots + a_n(x_n - x_n^0) + \epsilon(x),$$

where $\epsilon(x)$ is such that

$$\lim_{x \to x^0} \frac{\epsilon(x)}{|x - x^0|} = 0. \tag{1.2b}$$

If f is differentiable at x^0, the row vector a is called the *gradient* of f at x^0, is necessarily unique (Exercise 1.1), and is denoted by $\nabla f(x^0)$.

Directional derivatives may exist in all directions without a function being differentiable (Exercise 1.2). However, there is a simple formula for the directional derivative in terms of the gradient when f is differentiable.

Theorem A. Suppose f is differentiable at x^0. The directional derivative in the direction u is obtained as

$$D_u f(x^0) = \nabla f(x^0) u, \tag{1.3}$$

where the gradient is calculated by

$$\nabla f(x^0) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) \Big|_{x^0}. \tag{1.4}$$

Proof. Exercise 1.3.

The kth partial derivative $\partial f / \partial x_k$ represents the sensitivity of f to changes in the kth component variable x_k.

Corollary A. For $u \in \mathbb{R}^n$ with $|u| = 1$,

$$D_u f(x^0) = |\nabla f(x^0)| \cos \theta, \tag{1.5}$$

where θ is the angle between (the unit vector) u and $\nabla f(x^0)$.

Corollary B. The gradient $\nabla f(x^0)$ points in the direction of maximum increase of f at x^0. This maximal rate of increase is $|\nabla f(x^0)|$.

Example 1. Let $f(x, y) = x^2 + xy + y^3$. The directional derivative of f at $(1, 2)$ in the direction $u = (a, b)$ is then

$$D_{(a,b)} f(1,2) = (2x + y, x + 3y^2)|_{(1,2)} \begin{vmatrix} a \\ b \end{vmatrix} = 4a + 13b,$$

with maximal directional derivative (in the gradient direction) of value $\sqrt{16 + 169}$.

1.2 Calculus Rules

The product formula (1.3) for the directional derivative is actually an instance of a much more general result, the chain rule.

Theorem B. (The First Chain Rule) Suppose that each component of the vector curve $x = x(t)$ is differentiable at $t = t^0$, and that $f = f(x)$ is differentiable at $x^0 = x(t^0)$. Then

$$\frac{d\, f(x(t))}{dt} \Big|_{t=t^0} = \nabla f(x^0) \dot{x}(t^0) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \Big|_{x=x^0} \frac{dx_k}{dt} \Big|_{t=t^0}. \tag{1.6}$$

In more transparent notation,

$$\frac{d}{dt} f(x(t)) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{dx_k}{dt}.$$

Proof. Combine $x = x^0 + (t - t^0)\dot{x}(t^0) + \epsilon$ with equation (1.2) (Exercise 1.5).

Example 2. Let $f(x,y) = x^2 - 2xy + y^3$ and $x(t) = \cos t$, $y(t) = \sin t$. Then

$$\frac{d}{dt} f(x(t), y(t)) = -(2\cos t - 2\sin t)\sin t + (-2\cos t + 3\sin^2 t)\cos t.$$

The First Chain Rule itself is a special case of the following more general rule.

Theorem C. (The Second Chain Rule) Suppose $x = x(u) \in \mathbb{R}^n$ is differentiable at $u = u^0$ where $u = (u_1, \ldots, u_r)^{\top}$ (that is, each component of x is differentiable), and suppose also that $f = f(x)$ is differentiable at $x^0 = x(u^0)$. Then at $u = u^0$,

$$\frac{\partial f}{\partial u_k} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial u_k}. \tag{1.7}$$

Proof. Exercise 1.5.

Example 3. Let $f(x,y) = x^2 - 2xy + y^3$, $x = u^2 - v^2$, $y = u^2 + v^2$. Then

$$\frac{\partial f}{\partial v} = (2x - 2y)(-2v) + (-2x + 3y^2)(2v)$$
$$= -4v^2(-2v) + (-2u^2 + 2v^2 + 3(u^2 + v^2)^2)(2v).$$

Corollary. Suppose $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$, where $x = x(u)$. Then, where differentiable, we have the matrix relation

$$\left[\frac{\partial f_i}{\partial u_j}\right] = \left[\frac{\partial f_i}{\partial x_j}\right]\left[\frac{\partial x_i}{\partial u_j}\right].$$

That is,

$$\frac{\partial f_i}{\partial u_j} = \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k} \frac{\partial x_k}{\partial u_j}. \tag{1.8}$$

Theorem D. (The Mean Value Theorem) Suppose x^0 and x^1 belong to $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at each point on the line segment $[x^0, x^1] = \{tx^1 + (1-t)x^0;\ 0 \le t \le 1\}$. Then there exists a point $x^t \in [x^0, x^1]$ so that

$$f(x^1) - f(x^0) = \nabla f(x^t)(x^1 - x^0).$$

Proof. Apply the usual one-dimensional mean value theorem and the chain rule to $t \mapsto f(tx^1 + (1-t)x^0)$ defined on $[0, 1]$.

1.3 Contour Surfaces and Sublevel Sets

The locus of all points x satisfying $f(x) = f(x^0)$ is called the *contour surface* (or a *contour curve* if $n = 2$) of f through the point $x = x^0$.

Under mild assumptions on the differentiability of f near $x = x^0$, we may generically, in theory, solve for one of the components x_j of x, say for x_n, in terms of the remaining x_k [i.e., $x_n = x_n(x_1, \ldots, x_{n-1})$ so that the portion of the contour surface $f(x) = f(x^0)$ near $x = x^0$ is the graph of $z = f(x_1, x_2, \ldots, x_{n-1}, x_n(x_1, \ldots, x_{n-1}))$ near $(x_1^0, \ldots, x_{n-1}^0)$].

Example 4. Consider the contour curve $x^2 - y^2 = 1$ of $f(x, y) = x^2 - y^2$ through the point $(1, 0)$.

The locus has two disjoint branches—one in the first and fourth quadrant, the other in the second and and third quadrant. But only one branch passes through $(1, 0)$, where we may solve for x in terms of y :

$$x = \sqrt{1 + y^2},$$

valid for all y. See Figure 1.2.

The technical result that validates the preceding intuition in the general case is a workhorse of mathematics, the so-called *Implicit Function Theorem.* This theorem is easily deducible from another workhorse, the so-called *inverse function theorem.* See Chapter 4 and Appendix A. See also Exercises 1.15 and 5.13.

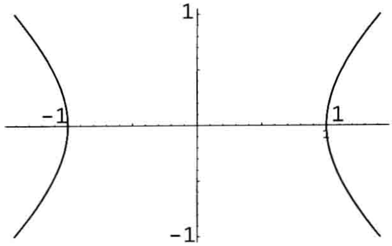


Figure 1.2 The contour curve of $x^2 - y^2 = 1$.

The value $f(x(t))$ along any curve $C : x = x(t)$ passing through $x^0 = x(0)$ that lies within the contour surface $f(x) = f(x^0)$ must, of course, be constantly $f(x^0)$. But then by the first chain rule,

$$\frac{d}{dt}f(x(t)) = \frac{d}{dt}f(x^0) = 0 = \nabla f(x(t))\dot{x}(t), \tag{1.9}$$

where of course $v = \dot{x}$ is tangent to the curve C given by $x = x(t)$ (i.e., the gradient is normal to the curve C). But since C was an arbitrary curve in the contour surface $f(x) = f(x^0)$,

The gradient is normal to the contour surface.

(See Exercise 1.38.) This means precisely that the hyperplane $\mathcal{H}$ that is tangent to the contour surface

$$S = \{x : f(x) = f(x^0)\}$$

at $x = x^0$ has normal vector $\nabla f(x^0)$, and hence the points of $\mathcal{H}$ are those x that satisfy the equation

$$\nabla f(x^0)(x - x^0) = 0. \tag{1.10}$$

Example 6. The plane tangent to the unit sphere $f(x, y, z) = x^2 + y^2 + z^2 = 1$ at $(2\sqrt{3}/5,\ 2/5,\ 3/5)$ has equation

$$(x - 2\sqrt{3}/5,\ y - 2/5,\ z - 3/5) \cdot (4\sqrt{3}/5,\ 4/5,\ 6/5)\ =\ 0,$$

that is,

$$\sqrt{3}x + 10y + 15z = 25.$$