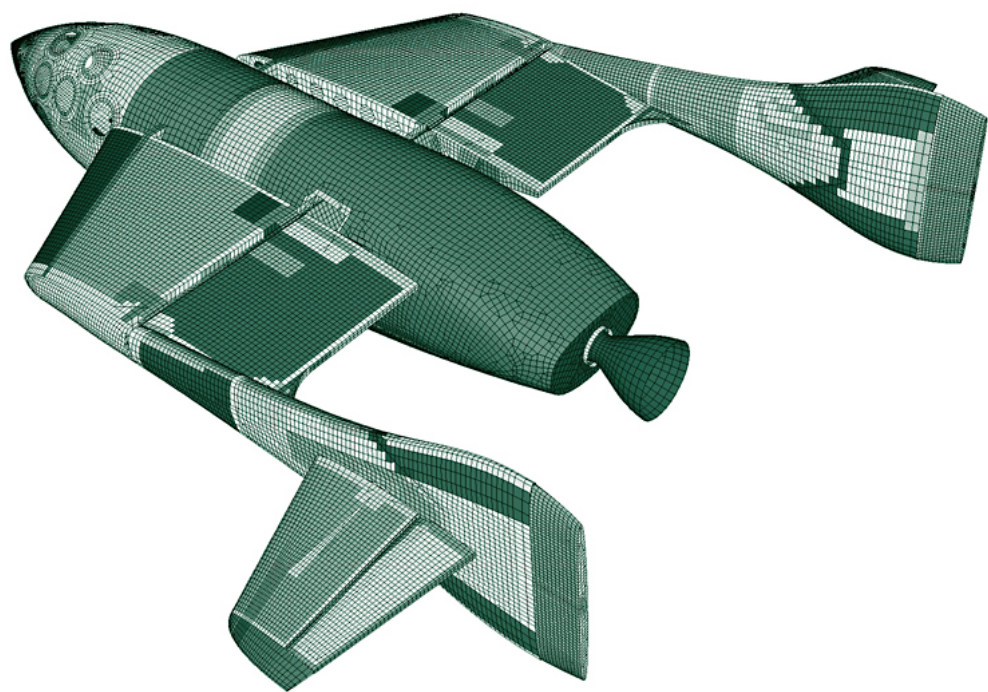


Applied Calculus of Variations for Engineers



Louis Komzsik

Applied Calculus of Variations for Engineers

Louis Komzsik



CRC Press

Taylor & Francis Group

Boca Raton London New York

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

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2009 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4200-8662-1 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Komzsis, Louis.
Applied calculus of variations for engineers / Louis Komzsis.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4200-8662-1 (hardcover : alk. paper)
1. Calculus of variations. 2. Engineering mathematics. I. Title.

TA347.C3K66 2009
620.001'51564--dc22

2008042179

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>
and the CRC Press Web site at
<http://www.crcpress.com>

To my daughter, Stella

Acknowledgments

I am indebted to Professor Emeritus Bajcsay Pál of the Technical University of Budapest, my alma mater. His superior lectures about four decades ago founded my original interest in calculus of variations.

I thank my coworkers, Dr. Leonard Hoffnung, for his meticulous verification of the derivations, and Mr. Chris Mehling, for his very diligent proofreading. Their corrections and comments greatly contributed to the quality of the book.

I appreciate the thorough review of Professor Dr. John Brauer of the Milwaukee School of Engineering. His recommendations were instrumental in improving the clarity of some topics and the overall presentation.

Special thanks are due to Nora Konopka, publisher of Taylor & Francis books, for believing in the importance of the topic and the timeliness of a new approach. I also appreciate the help of Ashley Gasque, editorial assistant, Amy Blalock and Stephanie Morkert, project coordinators, and the contributions of Michele Dimont, project editor.

I am grateful for the courtesy of Mojave Aerospace Ventures, LLC, and Quartus Engineering Incorporated for the model in the cover art. The model depicts a boosting configuration of SpaceShipOne, a Paul G. Allen project.

Louis Komzsik

About the Author

Louis Komzsik is a graduate of the Technical University of Budapest, Hungary, with a doctorate in mechanical engineering, and has worked in the industry for 35 years. He is currently the chief numerical analyst in the Office of Architecture and Technology at Siemens PLM Software.



Dr. Komzsik is the author of the *NASTRAN Numerical Methods Handbook* first published by MSC in 1987. His book, *The Lanczos Method*, published by SIAM, has also been translated into Japanese, Korean and Hungarian. His book, *Computational Techniques of Finite Element Analysis*, published by CRC Press, is in its second print, and his *Approximation Techniques for Engineers* was published by Taylor & Francis in 2006.

Preface

The topic of this book has a long history. Its fundamentals were laid down by icons of mathematics like Euler and Lagrange. It was once heralded as the panacea for all engineering optimization problems by suggesting that all one needs to do was to apply the Euler-Lagrange equation form and solve the resulting differential equation.

This, as most all encompassing solutions, turned out to be not always true and the resulting differential equations are not necessarily easy to solve. On the other hand, many of the differential equations commonly used by engineers today are derived from a variational problem. Hence, it is important and useful for engineers to delve into this topic.

The book is organized into two parts: theoretical foundation and engineering applications. The first part starts with the statement of the fundamental variational problem and its solution via the Euler-Lagrange equation. This is followed by the gradual extension to variational problems subject to constraints, containing functions of multiple variables and functionals with higher order derivatives. It continues with the inverse problem of variational calculus, when the origin is in the differential equation form and the corresponding variational problem is sought. The first part concludes with the direct solution techniques of variational problems, such as the Ritz, Galerkin and Kantorovich methods.

With the emphasis on applications, the second part starts with a detailed discussion of the geodesic concept of differential geometry and its extensions to higher order spaces. The computational geometry chapter covers the variational origin of natural splines and the variational formulation of B-splines under various constraints.

The final two chapters focus on analytic and computational mechanics. Topics of the first include the variational form and subsequent solution of several classical mechanical problems using Hamilton's principle. The last chapter discusses generalized coordinates and Lagrange's equations of motion. Some fundamental applications of elasticity, heat conduction and fluid mechanics as well as their computational technology conclude the book.

Contents

I	Mathematical foundation	1
1	The foundations of calculus of variations	3
1.1	The fundamental problem and lemma of calculus of variations	3
1.2	The Legendre test	7
1.3	The Euler-Lagrange differential equation	9
1.4	Application: Minimal path problems	11
1.4.1	Shortest curve between two points	12
1.4.2	The brachistochrone problem	14
1.4.3	Fermat's principle	18
1.4.4	Particle moving in the gravitational field	20
1.5	Open boundary variational problems	21
2	Constrained variational problems	25
2.1	Algebraic boundary conditions	25
2.2	Lagrange's solution	27
2.3	Application: Iso-perimetric problems	29
2.3.1	Maximal area under curve with given length	29
2.3.2	Optimal shape of curve of given length under gravity .	31
2.4	Closed-loop integrals	35
3	Multivariate functionals	37
3.1	Functionals with several functions	37
3.2	Variational problems in parametric form	38
3.3	Functionals with two independent variables	39
3.4	Application: Minimal surfaces	40
3.4.1	Minimal surfaces of revolution	43
3.5	Functionals with three independent variables	44
4	Higher order derivatives	49
4.1	The Euler-Poisson equation	49
4.2	The Euler-Poisson system of equations	51
4.3	Algebraic constraints on the derivative	52
4.4	Application: Linearization of second order problems	54

5	The inverse problem of the calculus of variations	57
5.1	The variational form of Poisson's equation	58
5.2	The variational form of eigenvalue problems	59
5.2.1	Orthogonal eigensolutions	61
5.3	Application: Sturm-Liouville problems	62
5.3.1	Legendre's equation and polynomials	64
6	Direct methods of calculus of variations	69
6.1	Euler's method	69
6.2	Ritz method	71
6.2.1	Application: Solution of Poisson's equation	75
6.3	Galerkin's method	76
6.4	Kantorovich's method	78
II	Engineering applications	85
7	Differential geometry	87
7.1	The geodesic problem	87
7.1.1	Geodesics of a sphere	89
7.2	A system of differential equations for geodesic curves	90
7.2.1	Geodesics of surfaces of revolution	92
7.3	Geodesic curvature	95
7.3.1	Geodesic curvature of helix	97
7.4	Generalization of the geodesic concept	98
8	Computational geometry	101
8.1	Natural splines	101
8.2	B-spline approximation	104
8.3	B-splines with point constraints	109
8.4	B-splines with tangent constraints	112
8.5	Generalization to higher dimensions	115
9	Analytic mechanics	119
9.1	Hamilton's principle for mechanical systems	119
9.2	Elastic string vibrations	120
9.3	The elastic membrane	125
9.3.1	Nonzero boundary conditions	130
9.4	Bending of a beam under its own weight	132
10	Computational mechanics	139
10.1	Three-dimensional elasticity	139
10.2	Lagrange's equations of motion	142
10.2.1	Hamilton's canonical equations	146
10.3	Heat conduction	148
10.4	Fluid mechanics	150
10.5	Computational techniques	153

10.5.1	Discretization of continua	153
10.5.2	Computation of basis functions	155
	Closing Remarks	159
	Notation	161
	List of Tables	163
	List of Figures	165
	References	167
	Index	169

Part I

Mathematical foundation

1

The foundations of calculus of variations

The problem of the calculus of variations evolves from the analysis of functions. In the analysis of functions the focus is on the relation between two sets of numbers, the independent (x) and the dependent (y) set. The function f creates a one-to-one correspondence between these two sets, denoted as

$$y = f(x).$$

The generalization of this concept is based on allowing the two sets not being restricted to be real numbers and to be functions themselves. The relationship between these sets is now called a functional. The topic of the calculus of variations is to find extrema of functionals, most commonly formulated in the form of an integral.

1.1 The fundamental problem and lemma of calculus of variations

The fundamental problem of the calculus of variations is to find the extremum (maximum or minimum) of the functional

$$I(y) = \int_{x_0}^{x_1} f(x, y, y') dx,$$

where the solution satisfies the boundary conditions

$$y(x_0) = y_0$$

and

$$y(x_1) = y_1.$$

This problem may be generalized to the cases when higher derivatives or multiple functions are given and will be discussed in [Chapters 3](#) and [4](#), respectively. These problems may also be extended with constraints, the topic of [Chapter 2](#).

A solution process may be arrived at with the following logic. Let us assume that there exists such a solution $y(x)$ for the above problem that satisfies

the boundary conditions and produces the extremum of the functional. Furthermore, we assume that it is twice differentiable. In order to prove that this function results in an extremum, we need to prove that any alternative function does not attain the extremum.

We introduce an alternative solution function of the form

$$Y(x) = y(x) + \epsilon\eta(x),$$

where $\eta(x)$ is an arbitrary auxiliary function of x , that is also twice differentiable and vanishes at the boundary:

$$\eta(x_0) = \eta(x_1) = 0.$$

In consequence the following is also true:

$$Y(x_0) = y(x_0) = y_0$$

and

$$Y(x_1) = y(x_1) = y_1.$$

A typical relationship between these functions is shown in [Figure 1.1](#) where the function is represented by the solid line and the alternative function by the dotted line. The dashed line represents the arbitrary auxiliary function.

Since the alternative function $Y(x)$ also satisfies the boundary conditions of the functional, we may substitute into the variational problem.

$$I(\epsilon) = \int_{x_0}^{x_1} f(x, Y, Y') dx.$$

where

$$Y'(x) = y'(x) + \epsilon\eta'(x).$$

The new functional in terms of ϵ is identical with the original in the case when $\epsilon = 0$ and has its extremum when

$$\frac{\partial I(\epsilon)}{\partial \epsilon} \Big|_{\epsilon=0} = 0.$$

Executing the derivation and taking the derivative into the integral, since the limits are fixed, with the chain rule we obtain

$$\frac{\partial I(\epsilon)}{\partial \epsilon} = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Y} \frac{dY}{d\epsilon} + \frac{\partial f}{\partial Y'} \frac{dY'}{d\epsilon} \right) dx.$$

Clearly

$$\frac{dY}{d\epsilon} = \eta(x),$$

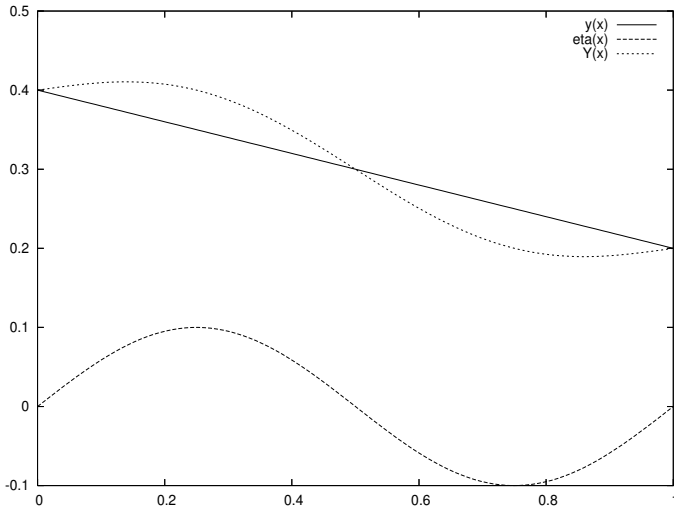


FIGURE 1.1 Alternative solutions example

and

$$\frac{dY'}{d\epsilon} = \eta'(x),$$

resulting in

$$\frac{\partial I(\epsilon)}{\partial \epsilon} = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Y'} \eta(x) + \frac{\partial f}{\partial Y'} \eta'(x) \right) dx.$$

Integrating the second term by parts yields

$$\int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Y'} \eta'(x) \right) dx = \frac{\partial f}{\partial Y'} \eta(x) \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \left(\frac{d}{dx} \frac{\partial f}{\partial Y'} \right) \eta(x) dx.$$

Due to the boundary conditions, the first term vanishes. With substitution and factoring the auxiliary function, the problem becomes

$$\frac{\partial I(\epsilon)}{\partial \epsilon} = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Y} - \frac{d}{dx} \frac{\partial f}{\partial Y'} \right) \eta(x) dx.$$

The extremum is achieved when $\epsilon = 0$ as stated above, hence

$$\frac{\partial I(\epsilon)}{\partial \epsilon} \Big|_{\epsilon=0} = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \eta(x) dx.$$

Let us now consider the following integral:

$$\int_{x_0}^{x_1} \eta(x) F(x) dx,$$

where $x_0 \leq x \leq x_1$ and $F(x)$ is continuous, while $\eta(x)$ is continuously differentiable, satisfying

$$\eta(x_0) = \eta(x_1) = 0.$$

The fundamental lemma of calculus of variations states that if for all such $\eta(x)$

$$\int_{x_0}^{x_1} \eta(x) F(x) dx = 0,$$

then

$$F(x) = 0$$

in the whole interval.

The following proof by contradiction is from [13]. Let us assume that there exists at least one such location $x_0 \leq \zeta \leq x_1$ where $F(x)$ is not zero, for example

$$F(\zeta) > 0.$$

By the condition of continuity of $F(x)$ there must be a neighborhood of

$$\zeta - h \leq \zeta \leq \zeta + h$$

where $F(x) > 0$. In this case, however, the integral becomes

$$\int_{x_0}^{x_1} \eta(x) F(x) dx > 0,$$

for the right choice of $\eta(x)$, which contradicts the original assumption. Hence the statement of the lemma must be true.

Applying the lemma to this case results in the **Euler-Lagrange differential equation** specifying the extremum

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0.$$

1.2 The Legendre test

The Euler-Lagrange differential equation just introduced represents a necessary, but not sufficient, condition for the solution of the fundamental variational problem.

The alternative functional of

$$I(\epsilon) = \int_{x_0}^{x_1} f(x, Y, Y') dx,$$

may be expanded as

$$I(\epsilon) = \int_{x_0}^{x_1} f(x, y + \epsilon\eta(x), y' + \epsilon\eta'(x)) dx.$$

Assuming that the f function has continuous partial derivatives, the mean-value theorem is applicable:

$$\begin{aligned} f(x, y + \epsilon\eta(x), y' + \epsilon\eta'(x)) &= f(x, y, y') + \\ &\epsilon\eta(x) \frac{\partial f(x, y, y')}{\partial y} + \eta'(x) \frac{\partial f(x, y, y')}{\partial y'} + O(\epsilon^2). \end{aligned}$$

By substituting we obtain

$$\begin{aligned} I(\epsilon) &= \int_{x_0}^{x_1} f(x, y, y') dx + \\ &\epsilon \int_{x_0}^{x_1} \left(\eta(x) \frac{\partial f(x, y, y')}{\partial y} + \eta'(x) \frac{\partial f(x, y, y')}{\partial y'} \right) dx + O(\epsilon^2). \end{aligned}$$

With the introduction of

$$\delta I_1 = \epsilon \int_{x_0}^{x_1} \left(\eta(x) \frac{\partial f(x, y, y')}{\partial y} + \eta'(x) \frac{\partial f(x, y, y')}{\partial y'} \right) dx,$$

we can write

$$I(\epsilon) = I(0) + \delta I_1 + O(\epsilon^2),$$

where δI_1 is called the first variation. The vanishing of the first variation is a necessary, but not sufficient, condition to have an extremum. To establish a sufficient condition, assuming that the function is thrice continuously differentiable, we further expand as

$$I(\epsilon) = I(0) + \delta I_1 + \delta I_2 + O(\epsilon^3).$$

Here the newly introduced second variation is

$$\begin{aligned}\delta I_2 = & \frac{\epsilon^2}{2} \int_{x_0}^{x_1} (\eta^2(x) \frac{\partial^2 f(x, y, y')}{\partial y^2} + \\ & 2\eta(x)\eta'(x) \frac{\partial^2 f(x, y, y')}{\partial y \partial y'} + \\ & \eta'^2(x) \frac{\partial^2 f(x, y, y')}{\partial y'^2}) dx.\end{aligned}$$

We now possess all the components to test for the existence of the extremum (maximum or minimum). The **Legendre test** in [5] states that if independently of the choice of the auxiliary $\eta(x)$ function

- the Euler-Lagrange equation is satisfied,
- the first variation vanishes ($\delta I_1 = 0$), and
- the second variation does not vanish ($\delta I_2 \neq 0$)

over the interval of integration, then the functional has an extremum. This test manifests the necessary conditions for the existence of the extremum. Specifically, the extremum will be a maximum if the second variation is negative, and conversely a minimum if it is positive. Certain similarities to the extremum evaluation of regular functions by the teaching of classical calculus are obvious.

We finally introduce the variation of the function as

$$\delta y = Y(x) - y(x) = \epsilon \eta(x),$$

and the variation of the derivative as

$$\delta y' = Y'(x) - y'(x) = \epsilon \eta'(x).$$

Based on these variations, we distinguish between the following cases:

- strong extremum occurs when δy is small, however, $\delta y'$ is large, while
- weak extremum occurs when both δy and $\delta y'$ are small.

On a final note: the above considerations did not ever state the finding or presence of an absolute extremum; only the local extremum in the interval of the integrand is obtained.

1.3 The Euler-Lagrange differential equation

Let us expand the derivative in the second term of the Euler-Lagrange differential equation as follows:

$$\frac{d}{dx} \frac{\partial f}{\partial y'} = \frac{\partial^2 f}{\partial x \partial y'} + \frac{\partial^2 f}{\partial y \partial y'} y' + \frac{\partial^2 f}{\partial y'^2} y''.$$

This demonstrates that the Euler-Lagrange equation is usually of second order.

$$\frac{\partial f}{\partial y} - \frac{\partial^2 f}{\partial x \partial y'} - \frac{\partial^2 f}{\partial y \partial y'} y' - \frac{\partial^2 f}{\partial y'^2} y'' = 0.$$

The above form is also called the extended form. Consider the case when the multiplier of the second derivative term vanishes:

$$\frac{\partial^2 f}{\partial y'^2} = 0.$$

In this case f must be a linear function of y' , in the form of

$$f(x, y, y') = p(x, y) + q(x, y)y'.$$

For this form, the other derivatives of the equation are computed as

$$\frac{\partial f}{\partial y} = \frac{\partial p}{\partial y} + \frac{\partial q}{\partial y} y',$$

$$\frac{\partial f}{\partial y'} = q,$$

$$\frac{\partial^2 f}{\partial x \partial y'} = \frac{\partial q}{\partial x},$$

and

$$\frac{\partial^2 f}{\partial y \partial y'} = \frac{\partial q}{\partial y}.$$

Substituting results in the Euler-Lagrange differential equation of the form

$$\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x} = 0,$$

or

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}.$$

In order to have a solution, this must be an identity, in which case there must be a function of two variables

$$u(x, y)$$

whose total differential is of the form

$$du = p(x, y)dx + q(x, y)dy = f(x, y, y')dx.$$

The functional may be evaluated as

$$I(y) = \int_{x_0}^{x_1} f(x, y, y')dx = \int_{x_0}^{x_1} du = u(x_1, y_1) - u(x_0, y_0).$$

It follows from this that the necessary and sufficient condition for the solution of the Euler-Lagrange differential equation is that the integrand of the functional be the total differential with respect to x of a certain function of both x and y .

Considering furthermore, that the Euler-Lagrange differential equation is linear with respect to f , it also follows that a term added to f will not change the necessity and sufficiency of that condition.

Another special case may be worthy of consideration. Let us assume that the integrand does not explicitly contain the x term. Then by executing the differentiations

$$\begin{aligned} & \frac{d}{dx}(y' \frac{\partial f}{\partial y'} - f) = \\ & y' \frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} y' = \\ & y' (\frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y}) - \frac{\partial f}{\partial x}. \end{aligned}$$

With the last term vanishing in this case, the differential equation simplifies to

$$\frac{d}{dx}(y' \frac{\partial f}{\partial y'} - f) = 0.$$

Its consequence is the expression also known as Beltrami's formula:

$$y' \frac{\partial f}{\partial y'} - f = c_1, \tag{1.1}$$

where the right-hand side term is an integration constant. The classical problem of the brachistochrone, discussed in the next section, belongs to this class.

Finally, it is also often the case that the integrand does not contain the y term explicitly. Then

$$\frac{\partial f}{\partial y} = 0$$

and the differential equation has the simpler

$$\frac{d}{dx} \frac{\partial f}{\partial y'} = 0$$

form. As above, the result is

$$\frac{\partial f}{\partial y'} = c_2$$

where c_2 is another integration constant. The geodesic problems, also the subject of [Chapter 7](#), represent this type of Euler-Lagrange equations.

We can surmise that the Euler-Lagrange differential equation's general solution is of the form

$$y = y(x, c_1, c_2),$$

where the c_1, c_2 are constants of integration, and are solved from the boundary conditions

$$y_0 = y(x_0, c_1, c_2)$$

and

$$y_1 = y(x_1, c_1, c_2).$$

1.4 Application: Minimal path problems

This section deals with several classical problems to illustrate the methodology. The problem of finding the minimal path between two points in space will be addressed in different senses.

The first problem is simple geometry, the shortest geometric distance between the points. The second one is the well-known classical problem of the brachistochrone, originally posed and solved by Bernoulli. This is the path of the shortest time required to move from one point to the other under the force of gravity.

The third problem considers a minimal path in an optical sense and leads to Snell's law of reflection in optics. The fourth example finds the path of minimal kinetic energy of a particle moving under the force of gravity.

All four problems will be presented in two-dimensional space, although they may also be posed and solved in three dimensions with some more algebraic difficulty but without any additional instructional benefit.

1.4.1 Shortest curve between two points

First we consider the rather trivial variational problem of finding the solution of the shortest curve between two points, P_0, P_1 , in the plane. The form of the problem using the arclength expression is

$$\int_{P_0}^{P_1} ds = \int_{x_0}^{x_1} \sqrt{1 + y'^2} dx = \text{extremum}.$$

The obvious boundary conditions are the curve going through its endpoints:

$$y(x_0) = y_0,$$

and

$$y(x_1) = y_1.$$

It is common knowledge that the solution in Euclidean geometry is a straight line from point (x_0, y_0) to point (x_1, y_1) . The solution function is of the form

$$y(x) = y_0 + m(x - x_0),$$

with slope

$$m = \frac{y_1 - y_0}{x_1 - x_0}.$$

To evaluate the integral, we compute the derivative as

$$y' = m$$

and the function becomes

$$f(x, y, y') = \sqrt{1 + m^2}.$$

Since the integrand is constant, the integral is trivial

$$I(y) = \sqrt{1 + m^2} \int_{x_0}^{x_1} dx = \sqrt{1 + m^2} (x_1 - x_0).$$

The square of the functional is

$$I^2(y) = (1 + m^2)(x_1 - x_0)^2 = (x_1 - x_0)^2 + (y_1 - y_0)^2.$$

This is the square of the distance between the two points in plane, hence the extremum is the distance between the two points along the straight line. Despite the simplicity of the example, the connection of a geometric problem to a variational formulation of a functional is clearly visible. This will be the most powerful justification for the use of this technique.

Let us now solve the

$$\int_{x_0}^{x_1} \sqrt{1 + y'^2} dx = \text{extremum}$$

problem via its Euler-Lagrange equation form. Note that the form of the integrand dictates the use of the extended form.

$$\begin{aligned}\frac{\partial f}{\partial y} &= 0, \\ \frac{\partial^2 f}{\partial x \partial y'} &= 0, \\ \frac{\partial^2 f}{\partial y \partial y'} &= 0,\end{aligned}$$

and

$$\frac{\partial^2 f}{\partial y'^2} = \frac{1}{(1 + y'^2)^{3/2}}.$$

Substituting into the extended form gives

$$\frac{1}{(1 + y'^2)^{3/2}} y'' = 0,$$

which simplifies into

$$y'' = 0.$$

Integrating twice, one obtains

$$y(x) = c_0 + c_1 x,$$

clearly the equation of a line. Substituting into the boundary conditions we obtain two equations,

$$y_0 = c_0 + c_1 x_0,$$

and

$$y_1 = c_0 + c_1 x_1.$$

The solution of the resulting linear system of equations is

$$c_0 = y_0 - c_1 x_0,$$

and

$$c_1 = \frac{y_1 - y_0}{x_1 - x_0}.$$

It is easy to reconcile that

$$y(x) = y_0 - \frac{y_1 - y_0}{x_1 - x_0} x_0 + \frac{y_1 - y_0}{x_1 - x_0} x$$

is identical to

$$y(x) = y_0 + m(x - x_0).$$

The noticeable difference between the two solutions of this problem is that using the Euler-Lagrange equation required no a priori assumption on the shape of the curve and the geometric know-how was not used. This is the case in most practical engineering applications and this is the reason for the utmost importance of the Euler-Lagrange equation.

1.4.2 The brachistochrone problem

The problem of the brachistochrone may be the first problem of variational calculus, already solved by Johann Bernoulli in the late 1600s. The name stands for the shortest time in Greek, indicating the origin of the problem.

The problem is elementary in a physical sense. Its goal is to find the shortest path of a particle moving in a vertical plane from a higher point to a lower point under the (only) force of gravity. The sought solution is the function $y(x)$ with boundary conditions $y(x_0) = y_0$ and $y(x_1) = y_1$ where

$$P_0 = (x_0, y_0)$$

and

$$P_1 = (x_1, y_1)$$

are the starting and terminal points, respectively. Based on elementary physics considerations, the problem represents an exchange of potential energy with kinetic energy.

A moving body's kinetic energy is related to its velocity and its mass. The higher the velocity and the mass, the bigger the kinetic energy. A body can gain kinetic energy using its potential energy, and conversely, can use its kinetic energy to build up potential energy. At any point during the movement, the total energy is at equilibrium. This is the principle of Hamilton's that will be discussed in more detail in [Chapter 9](#).

The potential energy of the particle at any x, y point during the motion is

$$E_p = mg(y_0 - y),$$

where m is the mass of the particle and g is the acceleration of gravity. The kinetic energy is

$$E_k = \frac{1}{2}mv^2$$

assuming that the particle at the (x, y) point has velocity v . They are in balance as

$$E_k = E_p,$$

resulting in an expression of the velocity as

$$v = \sqrt{2g(y_0 - y)}.$$

The velocity by definition is

$$v = \frac{ds}{dt},$$

where s is the arclength of the yet unknown curve. The time required to run the length of the curve is

$$t = \int_{P_0}^{P_1} dt = \int_{P_0}^{P_1} \frac{1}{v} ds.$$

Using the arclength expression from calculus, we get

$$t = \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{v} dx.$$

Substituting the velocity expression yields

$$t = \frac{1}{\sqrt{2g}} \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{\sqrt{y_0 - y}} dx.$$

Since we are looking for the minimal time, this is a variational problem of

$$I(y) = \frac{1}{\sqrt{2g}} \int_{x_0}^{x_1} \frac{\sqrt{1 + y'^2}}{\sqrt{y_0 - y}} dx = \text{extremum}.$$

The integrand does not contain the independent variable, hence we can apply Beltrami's formula of equation (1.1). This results in the form of

$$\frac{y'^2}{\sqrt{(y_0 - y)(1 + y'^2)}} - \frac{\sqrt{1 + y'^2}}{\sqrt{y_0 - y}} = c_0.$$

Creating a common denominator on the left-hand side produces

$$\frac{y'^2 \sqrt{y_0 - y} - \sqrt{1 + y'^2} \sqrt{(y_0 - y)(1 + y'^2)}}{\sqrt{(y_0 - y)(1 + y'^2)} \sqrt{y_0 - y}} = c_0.$$

Grouping the numerator simplifies to

$$\frac{-\sqrt{y_0 - y}}{\sqrt{(y_0 - y)(1 + y'^2)} \sqrt{y_0 - y}} = c_0.$$

Canceling and squaring results in the solution for y' as

$$y'^2 = \frac{1 - c_0^2(y_0 - y)}{c_0^2(y_0 - y)}.$$

Since

$$y'^2 = \left(\frac{dy}{dx}\right)^2,$$

the differential equation may be separated as

$$dx = \frac{\sqrt{y_0 - y}}{\sqrt{2c_1 - (y_0 - y)}} dy.$$

Here the new constant is introduced for simplicity as

$$c_1 = \frac{1}{2c_0^2}.$$

Finally x may be expressed directly by integrating

$$x = \int \frac{\sqrt{y_0 - y}}{\sqrt{2c_1 - (y_0 - y)}} dy + c_2.$$

The usual trigonometric substitution of

$$y_0 - y = 2c_1 \sin^2\left(\frac{t}{2}\right)$$

yields the integral of

$$x = 2c_1 \int \sin^2\left(\frac{t}{2}\right) dt = c_1(t - \sin(t)) + c_2,$$

where c_2 is another constant of integration. Reorganizing yields

$$y = y_0 - c_1(1 - \cos(t)).$$

The final solution of the brachistochrone problem therefore is a cycloid. [Figure 1.2](#) depicts the problem of the point moving from $(0, 1)$ until it reaches the x axis.

The resulting curve seems somewhat counter-intuitive, especially in view of the earlier example of the shortest geometric distance between two points in the plane and demonstrated by the straight line chord between the two points. The shortest time, however, when the speed obtained during the traversal of the interval depends on the path taken, is an entirely different matter.

The constants of the integration may be solved by substituting the boundary points. From the x equation above at $t = 0$ we easily find

$$x = c_2 = x_0.$$

Substituting the endpoint location into the y equation, we obtain

$$y = y_0 - c_1(1 - \cos(t)) = y_1,$$

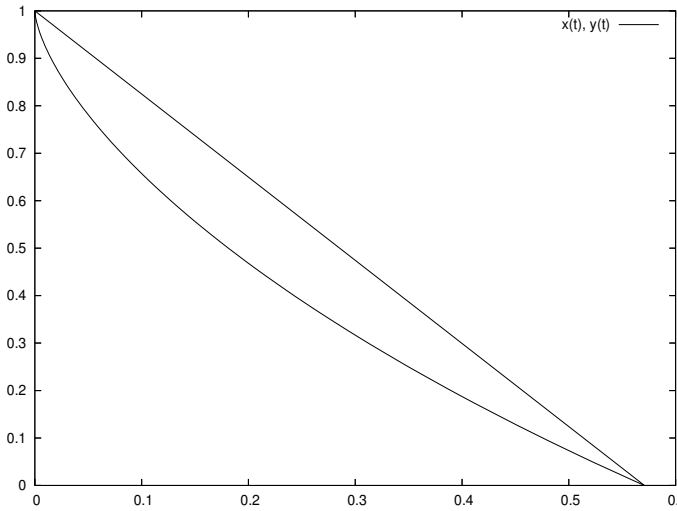


FIGURE 1.2 Solution of the brachistochrone problem

which is inconclusive, since the time of reaching the endpoint is not known. For a simple conclusion of this discussion, let us assume that the particle reaches the endpoint at time $t = \Pi/2$. Then

$$c_1 = y_0 - y_1,$$

and the final solution is

$$x = x_0 + (y_0 - y_1)(t - \sin(t))$$

and

$$y = y_1(1 - \cos(t)).$$

For the case shown in Figure 1.2, the point moving from $(0, 1)$ until it reaches the x axis, the solution curve is

$$x = (t - \sin(t))$$

and

$$y = \cos(t).$$

Another intriguing characteristic of the brachistochrone particle is that when two particles are let go from two different points of the curve they will

reach the terminal point of the curve at the same time. This is also counter-intuitive, since clearly they have different geometric distances to cover; however, since they are acting under the gravity and the slope of the curve is different at the two locations, the particle starting from a higher location gathers much bigger speed than the particle starting at a lower location.

This, so-called tautochrone, behavior may be proven by calculation of the time of the particles using the formula developed earlier. Evaluation of this integral between points (x_0, y_0) and (x_1, y_1) as well as between (x_2, y_2) and (x_1, y_1) (where (x_2, y_2) lies on the solution curve anywhere between the starting and terminal point) will result in the same time.

Hence the brachistochrone problem may also be posed with a specified terminal point and a variable starting point, leading to the class of variational problems with open boundary, subject of [Section 1.5](#).

1.4.3 Fermat's principle

Fermat's principle states that light traveling through an inhomogeneous medium chooses the path of minimal optical length. The path's optimal length depends on the speed of light in the medium, which is defined as a continuous function of

$$c = c(y),$$

where y is the vertical component of the path. Here

$$c(y) = \frac{ds}{dt},$$

as the derivative of the length of the path and its inverse will be in the variational form. The time required to cover the distance between two points is

$$t = \int \frac{1}{c(y)} ds.$$

The problem is now posed as a variational problem of

$$I(y) = \int_{(x_1, y_1)}^{(x_2, y_2)} \frac{ds}{c(y)}.$$

Substituting the arclength

$$\int_{x_1}^{x_2} \frac{\sqrt{1 + y'^2}}{c(y)} dx = \text{extremum},$$

with boundary conditions given at the two points P_1, P_2 .

$$y(x_1) = y_1; y(x_2) = y_2.$$

The functional does not contain the x term explicitly, allowing the use of Beltrami's formula of equation (1.1) and resulting in the simplified form of

$$y' \frac{\partial f}{\partial y'} - f = k_1$$

where k_1 is a constant of integration and its notation is chosen to distinguish from the speed of light value c . Substituting f , differentiating and simplifying yields

$$\frac{1}{c(y)\sqrt{1+y'^2}} = -k_1.$$

Reordering and separating results

$$\int dx = \pm k_1 \int \frac{c(y)}{\sqrt{1 - k_1^2 c^2(y)}} dy.$$

Depending on the particular model of the speed of light in the medium, the result varies. In the case of the inhomogeneous optical medium consisting of two homogeneous media in which the speed of light is piecewise constant, the result is the well-known Snell's law describing the scenario of the breaking path of light at the water's surface.

Assume the speed of light is c_1 between points P_1 and P_0 and c_2 between points P_0 and P_2 , both constant in their respective medium. The boundary between the two media is represented by

$$P_0(x, y_0),$$

where the notation signifies the fact that the x location of the light ray is not known yet. The known y_0 location specifies the distance of the points in the two separate media from the boundary.

Then the time to run the full path between P_1 and P_2 is simply

$$t = \frac{\sqrt{(x - x_1)^2 + (y_0 - y_1)^2}}{c_1} + \frac{\sqrt{(x_2 - x)^2 + (y_2 - y_0)^2}}{c_2}.$$

The minimum of this is simply obtained by classical calculus as

$$\frac{dt}{dx} = 0,$$

or

$$\frac{x - x_1}{c_1 \sqrt{(x - x_1)^2 + (y_0 - y_1)^2}} - \frac{x_2 - x}{c_2 \sqrt{(x_2 - x)^2 + (y_2 - y_0)^2}} = 0.$$

The solution of this equation yields the x location of the ray crossing the boundary, and produces the well-known Snell's law of

$$\frac{\sin\phi_1}{c_1} = \frac{\sin\phi_2}{c_2},$$

where the angles are measured with respect to the normal of the boundary between the two media. The preceding work generalizes to multiples of homogeneous media, which is a practical application in lens systems of optical machinery.

1.4.4 Particle moving in the gravitational field

The motion of a particle moving in the gravitational field of the Earth is computed based on the principle of least action. The principle, a subcase of Hamilton's principle, has been known for several hundred years, and was first proven by Euler. The principle states that a particle under the influence of a gravitational field moves on a path along which the kinetic energy is minimal. As such, it is a variational problem of

$$I = 2 \int E_k dt = \text{extremum}.$$

Here the multiplier is introduced for computational convenience. The kinetic energy is expressed as

$$E_k = \frac{1}{2}mv^2,$$

where v is the velocity of the particle. Substituting

$$v dt = ds$$

and

$$ds = \sqrt{1 + (y')^2} dx$$

the functional may be written as

$$I = m \int v \sqrt{1 + (y')^2} dx.$$

Since the gravitational field induces the particle's motion, its speed is related to its height known from elementary physics as

$$v^2 = u^2 - 2gy,$$

where u is an initial speed with yet undefined direction. Substituting into the functional yields

$$I = m \int \sqrt{u^2 - 2gy} \sqrt{1 + (y')^2} dx = \text{extremum}.$$

Since the functional does not contain x explicitly, we can use Beltrami's formula of equation (1.1), resulting in the Euler-Lagrange equation of

$$\frac{\sqrt{u^2 - 2gy}}{\sqrt{1 + y'^2}} = c_1,$$

where c_1 is an arbitrary constant. Expressing y' , separating and integrating yields

$$c_1^2(u^2 - 2gy - c_1^2) = g^2(x - c_2)^2,$$

with c_2 being another constant of integration. Reordering yields the well-known parabolic trajectory of

$$y = \frac{u^2 - c_1^2}{2g} - \frac{g}{2c_1^2}(x - c_2)^2.$$

The resolution of the constants may be by giving boundary conditions of the initial location and velocity of the particle. The constant c_1 is related to the latter and the constant c_2 is related to the location. Assuming the origin as initial location and an angle α of the initial velocity u with respect to the horizontal axis, the formula may be simplified into

$$y = x \tan(\alpha) - \frac{gx^2}{2u^2 \cos^2(\alpha)}.$$

Figure 1.3 demonstrates the path of the particle. The upper three curves show the path with a 60-degree angle of the initial velocity and with different magnitudes. The lower three curves demonstrate the paths obtained by the same magnitude (10 units), but different angles of the initial velocity. For visualization purposes the gravity constant was chosen to be 10 units as well.

A similarity between the four problems of this section is apparent. This recognition is a very powerful aspect of variational calculus. There are many instances in engineering applications when one physical problem may be solved in an analogous form using another principle. The common variational formulation of both problems is the key to such recognition in most cases.

1.5 Open boundary variational problems

Let us consider the variational problem of Section 1.1:

$$I(y) = \int_{x_0}^{x_1} f(x, y, y') dx,$$

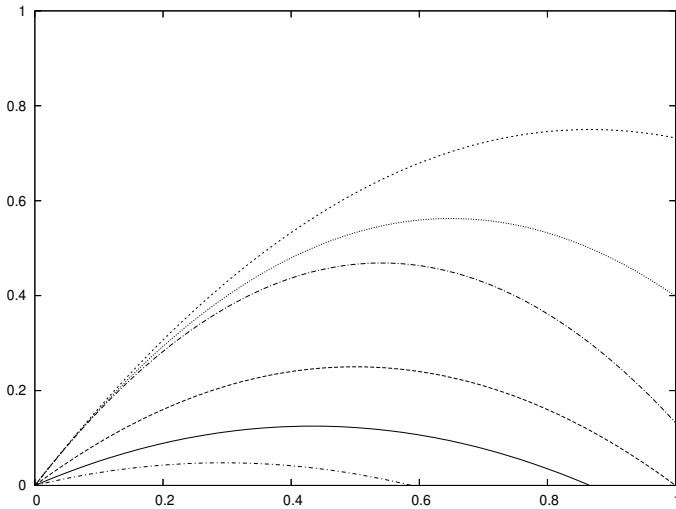


FIGURE 1.3 Trajectory of particle

with boundary condition

$$y(x_0) = y_0.$$

Let the boundary condition at the upper end be undefined. We introduce an auxiliary function $\eta(x)$ that in this case only satisfies

$$\eta(x_0) = 0.$$

The extremum in this case is obtained from the same concept as earlier:

$$\frac{\partial I(\epsilon)}{\partial \epsilon} \Big|_{\epsilon=0} = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial y} \eta(x) + \frac{d}{dx} \frac{\partial f}{\partial y'} \eta'(x) \right) dx = 0,$$

while recognizing the fact that x_1 is undefined. Integrating by parts and considering the one-sided boundary condition posed on the auxiliary function yields

$$\frac{\partial I(\epsilon)}{\partial \epsilon} \Big|_{\epsilon=0} = \frac{\partial f}{\partial y'} \Big|_{x=x_1} \eta(x_1) + \int_{x_0}^{x_1} \left(\left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \eta'(x) \right) dx = 0.$$

The extremum is obtained when the Euler-Lagrange equation of

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0$$

along with the given boundary condition of

$$y(x_0) = y_0$$

is satisfied, in addition to obeying the constraint of

$$\frac{\partial f}{\partial y'}|_{x=x_1} = 0.$$

Similar arguments may be applied when the starting point is open. This problem is the predecessor of the more generic constrained variational problems, the topic of the next chapter.

2

Constrained variational problems

The boundary values applied in the prior discussion may also be considered as constraints. The subject of this chapter is to generalize the constraint concept in two senses. The first is to allow more difficult, algebraic boundary conditions, and the second is to allow constraints imposed on the interior of the domain as well.

2.1 Algebraic boundary conditions

There is the possibility of defining the boundary condition at one end of the integral of the variational problem with an algebraic constraint. Let the

$$\int_{x_0}^{x_1} f(x, y, y') dx = \text{extremum}$$

variational problem subject to the customary boundary condition

$$y(x_0) = y_0,$$

on the lower end, and on the upper end an algebraic condition of the following form is given:

$$g(x, y) = 0.$$

We again consider an alternative solution of the form

$$Y(x) = y(x) + \epsilon \eta(x).$$

The given boundary condition in this case is

$$\eta(x_0) = 0.$$

Then, following [7], the intersection of the alternative solution and the algebraic curve is

$$X_1 = X_1(\epsilon)$$

and

$$Y_1 = Y_1(\epsilon).$$

The notation is to distinguish from the fixed boundary condition values given via x_1, y_1 . Therefore the algebraic condition is

$$g(X_1, Y_1) = 0.$$

This must be true for any ϵ , hence applying the chain rule yields

$$\frac{dg}{d\epsilon} = \frac{\partial g}{\partial X_1} \frac{dX_1}{d\epsilon} + \frac{\partial g}{\partial Y_1} \frac{dY_1}{d\epsilon} = 0. \quad (2.1)$$

Since

$$Y_1 = y(X_1) + \epsilon \eta(X_1),$$

we expand the last derivative of the second term of equation (2.1) as

$$\frac{dY_1}{d\epsilon} = \left. \frac{dy}{dx} \right|_{x=X_1} \frac{dX_1}{d\epsilon} + \eta(X_1) + \epsilon \left. \frac{d\eta}{dx} \right|_{x=X_1} \frac{dX_1}{d\epsilon}.$$

Substituting into equation (2.1) results in

$$\frac{dg}{d\epsilon} = \frac{\partial g}{\partial X_1} \frac{dX_1}{d\epsilon} + \frac{\partial g}{\partial Y_1} \left(\left. \frac{dy}{dx} \right|_{x=X_1} \frac{dX_1}{d\epsilon} + \eta(X_1) + \epsilon \left. \frac{d\eta}{dx} \right|_{x=X_1} \frac{dX_1}{d\epsilon} \right) = 0.$$

Since (X_1, Y_1) becomes (x_1, y_1) when $\epsilon = 0$,

$$\left. \frac{dX_1}{d\epsilon} \right|_{\epsilon=0} = - \frac{\eta(x_1) \left. \frac{\partial g}{\partial y} \right|_{y=y_1}}{\left. \frac{\partial g}{\partial x} \right|_{x=x_1} + \left. \frac{\partial g}{\partial y} \right|_{y=y_1} \left. \frac{dy}{dx} \right|_{x=x_1}}. \quad (2.2)$$

We now consider the variational problem of

$$I(\epsilon) = \int_{x_0}^{X_1} f(x, Y, Y') dx.$$

The derivative of this is

$$\frac{\partial I(\epsilon)}{\partial \epsilon} = \left. \frac{dX_1}{d\epsilon} f \right|_{x=X_1} + \int_{x_0}^{X_1} \left(\frac{\partial f}{\partial Y} \eta + \frac{\partial f}{\partial Y'} \eta' \right) dx.$$

Integrating by parts and taking $\epsilon = 0$ yields

$$\left. \frac{\partial I(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} = \left. \frac{dX_1}{d\epsilon} f \right|_{\epsilon=0} \Big|_{x=x_1} + \left. \frac{\partial f}{\partial y'} \right|_{x=x_1} \eta(x_1) + \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \eta dx.$$

Substituting the first expression with equation (2.2) results in

$$\left(\left. \frac{\partial f}{\partial y'} \right|_{x=x_1} - \frac{\left. \frac{\partial g}{\partial y} \right|_{y=y_1} f \Big|_{x=x_1}}{\left. \frac{\partial g}{\partial x} \right|_{x=x_1} + \left. \frac{\partial g}{\partial y} \right|_{y=y_1} \left. \frac{dy}{dx} \right|_{x=x_1}} \right) \eta(x_1) + \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \eta dx = 0.$$

Due to the fundamental lemma of calculus of variations, to find the constrained variational problem's extremum the Euler-Lagrange differential equation of

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0,$$

with the given boundary condition

$$y(x_0) = y_0,$$

and the algebraic constraint condition of the form

$$\frac{\partial f}{\partial y'}|_{x=x_1} = \frac{\frac{\partial g}{\partial y}|_{y=y_1} f|_{x=x_1}}{\frac{\partial g}{\partial x}|_{x=x_1} + \frac{\partial g}{\partial y}|_{y=y_1} \frac{dy}{dx}|_{x=x_1}}$$

all need to be satisfied.

2.2 Lagrange's solution

We now further generalize the variational problem and impose both boundary conditions as well as an algebraic condition on the whole domain as follows:

$$I(y) = \int_{x_0}^{x_1} f(x, y, y') dx = \text{extremum},$$

with

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

while

$$J(y) = \int_{x_0}^{x_1} g(x, y, y') dx = \text{constant}.$$

Following the earlier established pattern, we introduce an alternative solution function, at this time, however, with two auxiliary functions as

$$Y(x) = y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x).$$

Here the two auxiliary functions are arbitrary and both satisfy the boundary conditions:

$$\eta_1(x_0) = \eta_1(x_1) = \eta_2(x_0) = \eta_2(x_1) = 0.$$

Substituting these into the integrals gives

$$I(Y) = \int_{x_0}^{x_1} f(x, Y, Y') dx,$$

and

$$J(Y) = \int_{x_0}^{x_1} g(x, Y, Y') dx.$$

Lagrange's ingenious solution is to tie the two integrals together with a, yet unknown, multiplier (now called the Lagrange multiplier) as follows:

$$I(\epsilon_1, \epsilon_2) = I(Y) + \lambda J(Y) = \int_{x_0}^{x_1} h(x, Y, Y') dx,$$

where

$$h(x, y, y') = f(x, y, y') + \lambda g(x, y, y').$$

The condition to solve this variational problem is

$$\frac{\partial I}{\partial \epsilon_i} = 0$$

when

$$\epsilon_i = 0; i = 1, 2.$$

The derivatives are of the form

$$\frac{\partial I}{\partial \epsilon_i} = \int_{x_0}^{x_1} \left(\frac{\partial h}{\partial Y} \eta_i + \frac{\partial h}{\partial Y'} \eta'_i \right) dx.$$

The extremum is obtained when

$$\frac{\partial I}{\partial \epsilon_i} |_{\epsilon_i=0, i=1,2} = \int_{x_0}^{x_1} \left(\frac{\partial h}{\partial Y} \eta_i + \frac{\partial h}{\partial Y'} \eta'_i \right) dx = 0.$$

Considering the boundary conditions and integrating by parts yields

$$\int_{x_0}^{x_1} \left(\frac{\partial h}{\partial y} - \frac{d}{dx} \frac{\partial h}{\partial y'} \right) \eta_i dx = 0,$$

which, due to the fundamental lemma of calculus of variations, results in the relevant Euler-Lagrange differential equation

$$\frac{\partial h}{\partial y} - \frac{d}{dx} \frac{\partial h}{\partial y'} = 0.$$

This equation contains three undefined coefficients: the two coefficients of integration satisfying the boundary conditions and the Lagrange multiplier, enforcing the constraint.

2.3 Application: Iso-perimetric problems

Iso-perimetric problems use a given perimeter of a certain object as the constraint of some variational problem. The perimeter may be a curve in the two-dimensional case, as in the example of the next section. It may also be the surface of a certain body, in the three-dimensional case.

2.3.1 Maximal area under curve with given length

This problem is conceptually very simple, but useful to illuminate the process just established. It is also a very practical problem with more difficult geometries involved. Here we focus on the simple case of finding the curve of given length between two points in the plane. Without restricting the generality of the discussion, we'll position the two points on the x axis in order to simplify the arithmetic.

The given points are $(x_0, 0)$ and $(x_1, 0)$ with $x_0 < x_1$. The area under any curve going from the start point to the endpoint in the upper half-plane is

$$I(y) = \int_{x_0}^{x_1} y dx.$$

The constraint of the given length L is presented by the equation

$$J(y) = \int_{x_0}^{x_1} \sqrt{1 + y'^2} dx = L.$$

The Lagrange multiplier method brings the function

$$h(x, y, y') = y(x) + \lambda \sqrt{1 + y'^2}.$$

The constrained variational problem is

$$I(y) = \int_{x_0}^{x_1} h(x, y, y') dx$$

whose Euler-Lagrange equation becomes

$$1 - \lambda \frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2}} = 0.$$

Integration yields

$$\frac{\lambda y'}{\sqrt{1 + y'^2}} = x - c_1.$$

First we separate the variables

$$dy = \pm \frac{x - c_1}{\sqrt{\lambda^2 - (x - c_1)^2}} dx,$$

and integrate again to produce

$$y(x) = \pm \sqrt{\lambda^2 - (x - c_1)^2} + c_2.$$

It is easy to reorder this into

$$(x - c_1)^2 + (y - c_2)^2 = \lambda^2,$$

which is the equation of a circle. Since the two given points are on the x axis, the center of the circle must lie on the perpendicular bisector of the chord, which implies that

$$c_1 = \frac{x_0 + x_1}{2}.$$

To solve for the value of the Lagrange multiplier and the other constant, we consider that the circular arc between the two points is the given length:

$$L = \lambda\theta,$$

where θ is the angle of the arc. The angle is related to the remaining constant as

$$2\Pi - \theta = \text{atan}\left(\frac{x_1 - x_0}{2c_2}\right).$$

The two equations may be simultaneously satisfied with

$$\theta = \Pi,$$

resulting in the shape being a semi-circle. This yields the solutions of

$$c_2 = 0$$

and

$$\lambda = \frac{L}{\pi}.$$

The final solution function in implicit form is

$$\left(x - \frac{x_0 + x_1}{2}\right)^2 + y^2 = \left(\frac{L}{\pi}\right)^2,$$

or explicitly

$$y(x) = \sqrt{\left(\frac{L}{\pi}\right)^2 - \left(x - \frac{x_0 + x_1}{2}\right)^2}.$$

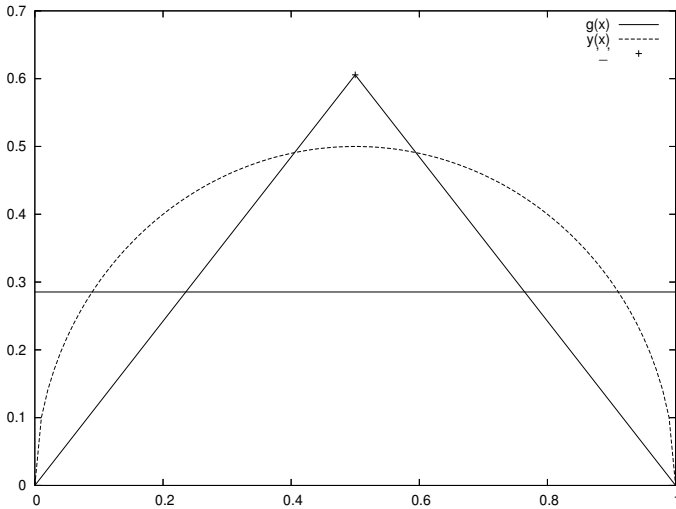


FIGURE 2.1 Maximum area under curves

It is simple to verify that the solution produces the extremum of the original variational problem.

Figure 2.1 visibly demonstrates the phenomenon with three curves of equal length ($\pi/2$) over the same interval. None of the solid curves denoted by $g(x)$, the triangle, or the rectangle cover as much area as the semi-circle $y(x)$ marked by the dashed lines.

2.3.2 Optimal shape of curve of given length under gravity

Another constrained variational problem, whose final result is often used in engineering, is the rope hanging under its weight. The practical importance of the problem regarding power lines and suspended cables is well-known. Here we derive the solution of this problem from a variational origin.

A body in a force field is in static equilibrium when its potential energy has a stationary value. Furthermore, if the stationary value is a minimum, then the body is in stable equilibrium. This is also known as principle of minimum potential energy.

Assume a body of a homogeneous cable with a given weight per unit length of $\rho = \text{constant}$, and suspension point locations of

$$P_0 = (x_0, y_0),$$

and

$$P_1 = (x_1, y_1).$$

These constitute the boundary conditions. A constraint is also given on the length of the curve: L . The potential energy of the cable is

$$E_p = \int_{P_0}^{P_1} \rho y ds$$

where y is the height of the infinitesimal arc segment above the horizontal base line and ρds is its weight. Using the arc length formula we obtain

$$E_p = \rho \int_{x_0}^{x_1} y \sqrt{1 + y'^2} dx.$$

The principle of minimal potential energy dictates that the equilibrium position of the cable is the solution of the variational problem of

$$I(y) = \rho \int_{x_0}^{x_1} y \sqrt{1 + y'^2} dx = \text{extremum},$$

under boundary conditions

$$y(x_0) = y_0; y(x_1) = y_1$$

and constraint of

$$\int_{x_0}^{x_1} \sqrt{1 + y'^2} dx = L.$$

Introducing the Lagrange multiplier and the constrained function

$$h(y) = \rho y \sqrt{1 + y'^2} + \lambda \sqrt{1 + y'^2},$$

the Euler-Lagrange differential equation of the problem after the appropriate differentiations becomes

$$\rho \sqrt{1 + y'^2} - \frac{d}{dx} \frac{(\rho y + \lambda) y'}{\sqrt{1 + y'^2}} = 0.$$

Some algebraic activity, which does not add anything to the discussion, and hence is undetailed, yields

$$(\rho y + \lambda) \left(\frac{y'^2}{\sqrt{1 + y'^2}} - \sqrt{1 + y'^2} \right) = c_1,$$

where the right-hand side is a constant of the integration. Another integration results in the solution of the so-called catenary curve

$$y = -\frac{\lambda}{\rho} - \frac{c_1}{\rho} \cosh\left(\frac{\rho(x - c_2)}{c_1}\right),$$

with c_2 being another constant of integration. The constants of integration may be determined by the boundary conditions albeit the calculation, due to the presence of the hyperbolic function, is rather tedious. Let us consider the specific case of the suspension points being at the same height and symmetric with respect to the origin. This is a typical engineering scenario for the span of suspension cables. This results in the following boundary conditions:

$$P_0 = (x_0, y_0) = (-s, h)$$

and

$$P_1 = (x_1, y_1) = (s, h).$$

Without the loss of the generality, we can consider unit weight ($\rho = 1$) and by substituting above boundary conditions we obtain

$$h + \lambda = c_1 \cosh\left(\frac{-s + c_2}{c_1}\right) = c_1 \cosh\left(\frac{s + c_2}{c_1}\right).$$

This implies that

$$c_2 = 0.$$

The value of the second coefficient is solved by adhering to the length constraint. Integrating the constraint equation yields

$$L = 2c_1 \sinh\left(\frac{s}{c_1}\right)$$

whose only unknown is the integration constant c_1 . This problem is not solvable by analytic means; however, it can be solved by an iterative procedure numerically by considering the unknown coefficient as a variable:

$$c_1 = x,$$

and intersecting the curve

$$y = x \sinh\left(\frac{s}{x}\right)$$

and the horizontal line

$$y = \frac{L}{2}.$$

The minimal cable length must exceed the width of the span, hence we expect the cable to have some slack. Then for example using

$$L = 3s$$

will result in an approximate solution of

$$c_1 = 0.6175.$$

Clearly, depending on the length of the cable between similarly posted suspension locations, different catenary curves may be obtained.

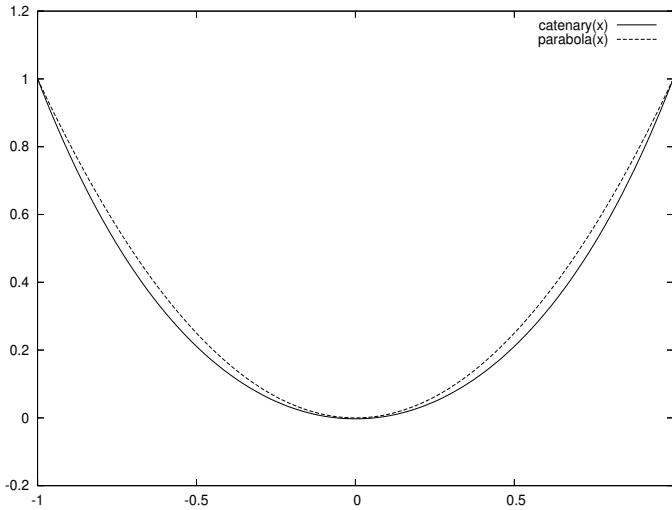


FIGURE 2.2 The catenary curve

The Lagrange multiplier may finally be resolved by the expression

$$\lambda = c_1 \cosh\left(\frac{s}{c_1}\right) - h.$$

Assuming a cable suspended with a unit half-span ($s = 1$) and from unit height ($h = 1$) and length of three times the half-span ($L = 3$), the value of the Lagrange multiplier becomes

$$\lambda = 0.6175 \cosh\left(\frac{1}{0.6175}\right) - 1 = 0.6204.$$

The final catenary solution curve, shown with a solid line in Figure 2.2, is represented by

$$y = 0.6175 \cosh\left(\frac{1}{0.6175}\right) - 0.6204.$$

For comparison purposes, the figure also shows a parabola with dashed lines, representing an approximation of the catenary and obeying the same boundary conditions.

2.4 Closed-loop integrals

As a final topic in this chapter, we briefly view variational problems posed in terms of closed-loop integrals, such as

$$I = \oint f(x, y, y') dx = \text{extremum},$$

subject to the constraint of

$$J = \oint g(x, y, y') dx.$$

Note that there are no boundary points of the path given since it is a closed loop. The substitution of

$$x = a \cos(t), \quad y = a \sin(t),$$

changes the problem to the conventional form of

$$I = \int_{t_0}^{t_1} F(x, y, \dot{x}, \dot{y}) dt,$$

subject to

$$J = \int_{t_0}^{t_1} G(x, y, \dot{x}, \dot{y}) dt.$$

The arbitrary t_0 and the specific $t_1 = t_0 + 2\pi$ boundary points clearly cover a complete loop.

3

Multivariate functionals

3.1 Functionals with several functions

The variational problem of multiple dependent variables is posed as

$$I(y_1, y_2, \dots, y_n) = \int_{x_0}^{x_1} f(x, y_1, y_2, \dots, y_n, y'_1, y'_2, \dots, y'_n) dx$$

with a pair of boundary conditions given for all functions:

$$y_i(x_0) = y_{i,0}$$

and

$$y_i(x_1) = y_{i,1}$$

for each $i = 1, 2, \dots, n$. The alternative solutions are:

$$Y_i(x) = y_i(x) + \epsilon_i \eta_i(x); i = 1, \dots, n$$

with all the arbitrary auxiliary functions obeying the conditions:

$$\eta_i(x_0) = \eta_i(x_1) = 0.$$

The variational problem becomes

$$I(\epsilon_1, \dots, \epsilon_n) = \int_{x_0}^{x_1} f(x, \dots, y_i + \epsilon_i \eta_i, \dots, y'_i + \epsilon_i \eta'_i, \dots) dx,$$

whose derivative with respect to the auxiliary variables is

$$\frac{\partial I}{\partial \epsilon_i} = \int_{x_0}^{x_1} \frac{\partial f}{\partial \epsilon_i} dx = 0.$$

Applying the chain rule we get

$$\frac{\partial f}{\partial \epsilon_i} = \frac{\partial f}{\partial Y_i} \frac{\partial Y_i}{\partial \epsilon_i} + \frac{\partial f}{\partial Y'_i} \frac{\partial Y'_i}{\partial \epsilon_i} = \frac{\partial f}{\partial Y_i} \eta_i + \frac{\partial f}{\partial Y'_i} \eta'_i.$$

Substituting into the variational equation yields, for $i = 1, 2, \dots, n$:

$$I(\epsilon_i) = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Y_i} \eta_i + \frac{\partial f}{\partial Y'_i} \eta'_i \right) dx.$$

Integrating by parts and exploiting the alternative function form results in

$$I(\epsilon_i) = \int_{x_0}^{x_1} \eta_i \left(\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y'_i} \right) dx.$$

To reach the extremum, based on the fundamental lemma, we need the solution of a set of n Euler-Lagrange equations of the form

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y'_i} = 0; i = 1, \dots, n.$$

3.2 Variational problems in parametric form

Most of the discussion insofar was focused on functions in explicit form. The concepts also apply to problems posed in parametric form. The explicit form variational problem of

$$I(y) = \int_{x_0}^{x_1} f(x, y, y') dx$$

may be reformulated with the substitutions

$$x = u(t), \quad y = v(t).$$

The parametric variational problem becomes of the form

$$I(x, y) = \int_{t_0}^{t_1} f(x, y, \frac{\dot{y}}{\dot{x}}) \dot{x} dt,$$

or

$$I(x, y) = \int_{t_0}^{t_1} F(t, x, y, \dot{x}, \dot{y}) dt.$$

The Euler-Lagrange differential equation system for this case becomes

$$\frac{\partial F}{\partial x} - \frac{d}{dt} \frac{\partial F}{\partial \dot{x}} = 0,$$

and

$$\frac{\partial F}{\partial y} - \frac{d}{dt} \frac{\partial F}{\partial \dot{y}} = 0.$$

It is proven in [7] that an explicit variational problem is invariant under parameterization. In other words, independently of the algebraic form of the parameterization, the same explicit solution will be obtained.

Parametrically given problems may be considered as functionals with several functions. As an example, we consider the following twice differentiable

functions

$$x = x(t), y = y(t), z = z(t).$$

The variational problem in this case is presented as

$$I(x, y, z) = \int_{t_0}^{t_1} f(t, x, y, z, \dot{x}, \dot{y}, \dot{z}) dt.$$

Here the independent variable t is the parameter, and there are three dependent variables : x, y, z . Applying the steps just explained for this specific case results in the system of Euler-Lagrange equations

$$\begin{aligned} \frac{\partial f}{\partial x} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}} &= 0, \\ \frac{\partial f}{\partial y} - \frac{d}{dt} \frac{\partial f}{\partial \dot{y}} &= 0, \end{aligned}$$

and

$$\frac{\partial f}{\partial z} - \frac{d}{dt} \frac{\partial f}{\partial \dot{z}} = 0.$$

The most practical applications of this case are variational problems in three-dimensional space, presented in parametric form. This is usual in many geometry problems and will be exploited in [Chapters 7](#) and [8](#).

3.3 Functionals with two independent variables

All our discussions so far were confined to a single integral of the functional. The next step of generalization is to allow a functional with multiple independent variables. The simplest case is that of two independent variables, and this will be the vehicle to introduce the process. The problem is of the form

$$I(z) = \int_{y_0}^{y_1} \int_{x_0}^{x_1} f(x, y, z, z_x, z_y) dx dy = \text{extremum}.$$

Here the derivatives are

$$z_x = \frac{\partial z}{\partial x}$$

and

$$z_y = \frac{\partial z}{\partial y}.$$

The alternative solution is also a function of two variables

$$Z(x, y) = z(x, y) + \epsilon \eta(x, y).$$

The now familiar process emerges as

$$I(\epsilon) = \int_{y_0}^{y_1} \int_{x_0}^{x_1} f(x, y, Z, Z_x, Z_y) dx dy = \text{extremum}.$$

The extremum is obtained via the derivative

$$\frac{\partial I}{\partial \epsilon} = \int_{y_0}^{y_1} \int_{x_0}^{x_1} \frac{\partial f}{\partial \epsilon} dx dy.$$

Differentiating and substituting yields

$$\frac{\partial I}{\partial \epsilon} = \int_{y_0}^{y_1} \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Z} \eta + \frac{\partial f}{\partial Z_x} \eta_x + \frac{\partial f}{\partial Z_y} \eta_y \right) dx dy.$$

The extremum is reached when $\epsilon = 0$:

$$\frac{\partial I}{\partial \epsilon} \Big|_{\epsilon=0} = \int_{y_0}^{y_1} \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial z} \eta + \frac{\partial f}{\partial z_x} \eta_x + \frac{\partial f}{\partial z_y} \eta_y \right) dx dy = 0.$$

Applying Green's identity for the second and third terms produces

$$\int_{y_0}^{y_1} \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial z} - \frac{\partial}{\partial x} \frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y} \frac{\partial f}{\partial z_y} \right) \eta dx dy + \int_{\partial D} \left(\frac{\partial f}{\partial z_x} \frac{dy}{ds} - \frac{\partial f}{\partial z_y} \frac{dx}{ds} \right) \eta ds = 0.$$

Here ∂D is the boundary of the domain of the problem and the second integral vanishes by the definition of the auxiliary function. Due to the fundamental lemma of calculus of variations, the Euler-Lagrange differential equation becomes

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x} \frac{\partial f}{\partial z_x} - \frac{\partial}{\partial y} \frac{\partial f}{\partial z_y} = 0.$$

3.4 Application: Minimal surfaces

Minimal surfaces occur in intriguing applications. For example, soap films spanned over various type of wire loops intrinsically attain such shapes, no matter how difficult the boundary curve is. Various biological cell interactions also manifest similar phenomena.

From a differential geometry point-of-view a minimal surface is a surface for which the mean curvature of the form

$$\kappa_m = \frac{\kappa_1 + \kappa_2}{2}$$

vanishes, where κ_1 and κ_2 are the principal curvatures. A subset of minimal surfaces are the surfaces of minimal area, and surfaces of minimal area passing

through a closed space curve are minimal surfaces. Finding minimal surfaces is called the problem of Plateau.

We seek the surface of minimal area with equation

$$z = f(x, y), (x, y) \in D,$$

with a closed-loop boundary curve

$$g(x, y, z) = 0; (x, y) \in \partial D.$$

The boundary condition represents a three-dimensional curve defined over the perimeter of the domain. The curve may be piecewise differentiable, but continuous and forms a closed loop, a Jordan curve.

The corresponding variational problem is

$$I(z) = \int \int_D \sqrt{1 + \frac{\partial z^2}{\partial x} + \frac{\partial z^2}{\partial y}} dxdy = \text{extremum}.$$

subject to the constraint of the boundary condition above. The Euler-Lagrange equation for this case is of the form

$$-\frac{\partial}{\partial x} \frac{\frac{\partial z}{\partial x}}{\sqrt{1 + (\frac{\partial z}{\partial x})^2 + (\frac{\partial z}{\partial y})^2}} - \frac{\partial}{\partial y} \frac{\frac{\partial z}{\partial y}}{\sqrt{1 + (\frac{\partial z}{\partial x})^2 + (\frac{\partial z}{\partial y})^2}} = 0.$$

After considerable algebraic work, this equation becomes

$$(1 + (\frac{\partial z}{\partial y})^2) \frac{\partial^2 z}{\partial x^2} - 2 \frac{\partial z}{\partial x} \frac{\partial z}{\partial y} \frac{\partial^2 z}{\partial x \partial y} + (1 + (\frac{\partial z}{\partial x})^2) \frac{\partial^2 z}{\partial y^2} = 0.$$

This is the differential equation of minimal surfaces, originally obtained by Lagrange himself. The equation is mainly of verification value as this is one of the most relevant examples for the need of a numerical solution. Most of the problems of finding minimal surfaces are solved by Ritz type methods, the subject of [Chapter 6](#).

The simplest solutions for such problems are the so-called saddle surfaces, such as, for example shown in [Figure 3.1](#), whose equation is

$$z = x^3 - 2xy^2.$$

It is easy to verify that this satisfies the equation. The figure also shows the level curves of the surface projected to the $x - y$ plane. The straight lines on the plane correspond to geodesic paths, a subject of detailed discussion in [Chapter 7](#). It is apparent that the $x = 0$ planar cross-section of the surface is the $z = 0$ line in the $x - y$ plane, as indicated by the algebra. The intersection with the $y = 0$ plane produces the $z = x^3$ curve, again in full adherence to

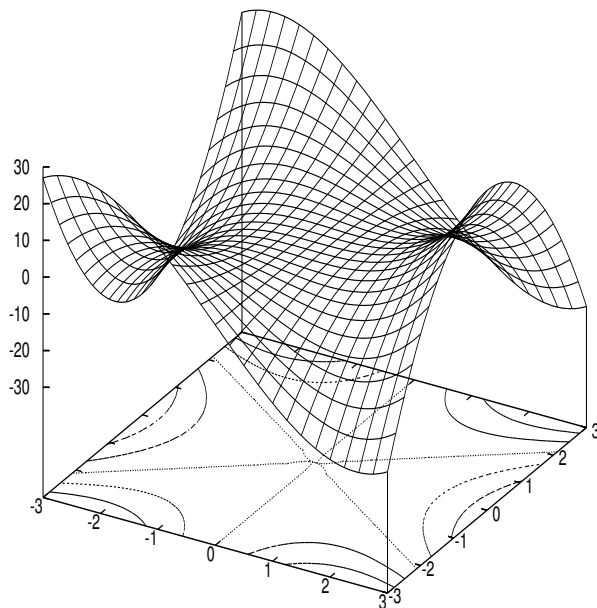


FIGURE 3.1 Saddle surface

the equation.

When a minimal surface is sought in a parametric form

$$\underline{r} = x(u, v)\underline{i} + y(u, v)\underline{j} + z(u, v)\underline{k}.$$

the variational problem becomes

$$I(\underline{r}) = \int \int_D \sqrt{EF - G^2} dA,$$

where the so-called first fundamental quantities are defined as

$$E(u, v) = (\underline{r}'_u)^2,$$

$$F(u, v) = \underline{r}'_u \underline{r}'_v,$$

and

$$G(u, v) = (\underline{r}'_v)^2.$$

The solution may be obtained from the differential equation

$$\frac{\partial}{\partial u} \frac{F \underline{r}'_u - G \underline{r}'_v}{\sqrt{EF - G^2}} + \frac{\partial}{\partial v} \frac{E \underline{r}'_v - G \underline{r}'_u}{\sqrt{EF - G^2}} = 0.$$

Finding minimal surfaces for special boundary arrangements arising from revolving curves is discussed in the next section.

3.4.1 Minimal surfaces of revolution

The problem has obvious relevance in mechanical engineering and computer-aided manufacturing (CAM). Let us now consider two points

$$P_0 = (x_0, y_0), P_1 = (x_1, y_1),$$

and find the function $y(x)$ going through the points that generates an object of revolution $z = f(x, y)$ when rotated around the x axis with minimal surface area. The surface of that object of revolution is

$$S = 2\pi \int_{x_0}^{x_1} y \sqrt{1 + y'^2} dx.$$

The corresponding variational problem is

$$I(y) = 2\pi \int_{x_0}^{x_1} y \sqrt{1 + y'^2} dx = \text{extremum},$$

with the boundary conditions of

$$y(x_0) = y_0, y(x_1) = y_1.$$

The Beltrami formula of equation (1.1) produces

$$y \sqrt{1 + y'^2} - \frac{yy'^2}{\sqrt{1 + y'^2}} = c_1.$$

Reordering and another integration yields

$$x = c_1 \int \frac{1}{\sqrt{y^2 - c_1^2}} dy.$$

Hyperbolic substitution enables the integration as

$$x = c_1 \cosh^{-1}\left(\frac{y}{c_1}\right) + c_2.$$

Finally the solution curve generating the minimal surface of revolution between the two points is

$$y = c_1 \cosh\left(\frac{x - c_2}{c_1}\right),$$

where the integration constants are resolved with the boundary conditions as

$$y_0 = c_1 \cosh\left(\frac{x_0 - c_2}{c_1}\right),$$

and

$$y_1 = c_1 \cosh\left(\frac{x_1 - c_2}{c_1}\right).$$

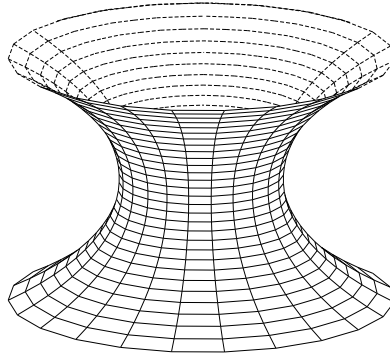


FIGURE 3.2 Catenoid surface

An example of such a surface of revolution, the catenoid, is shown in Figure 3.2 where the meridian curves are catenary curves.

3.5 Functionals with three independent variables

The generalization to functionals with multiple independent variables is rather straightforward from the last section. The case of three independent variables, however, has such enormous engineering importance that it is worthy of a special section. The problem is of the form

$$I(u(x, y, z)) = \int \int \int_D f(x, y, z, u, u_x, u_y, u_z) dx dy dz = \text{extremum}.$$

The solution function $u(x, y, z)$ may be some engineering quantity describing a physical phenomenon acting on a three-dimensional body. Here the domain is generalized as well to

$$x_0 \leq x \leq x_1, \quad y_0 \leq y \leq y_1, \quad z_0 \leq z \leq z_1.$$

The alternative solution is also a function of three variables

$$U(x, y, z) = u(x, y, z) + \epsilon \eta(x, y, z).$$

As usual

$$I(\epsilon) = \int \int \int_D f(x, y, z, U, U_x, U_y, U_z) dx dy dz.$$

The extremum is reached when:

$$\frac{\partial I}{\partial \epsilon} \Big|_{\epsilon=0} = \int \int \int_D \left(\frac{\partial f}{\partial u} \eta + \frac{\partial f}{\partial u_x} \eta_x + \frac{\partial f}{\partial u_y} \eta_y + \frac{\partial f}{\partial u_z} \eta_z \right) dx dy dz = 0.$$

Applying Green's identity for the last three terms and a considerable amount of algebra produces the Euler-Lagrange differential equation for this case

$$\frac{\partial f}{\partial u} - \frac{\partial}{\partial x} \frac{\partial f}{\partial u_x} - \frac{\partial}{\partial y} \frac{\partial f}{\partial u_y} - \frac{\partial}{\partial z} \frac{\partial f}{\partial u_z} = 0.$$

An even more practical three-variable case, important in engineering dynamics, is when the Euclidean spatial coordinates are extended with time. Let us consider the variational problem of one temporal and two spatial dimensions as

$$I(u) = \int_{t_0}^{t_1} \int \int_D f(x, y, t, u, u_x, u_y, u_t) dx dy dt = \text{extremum}.$$

Here again

$$u_x = \frac{\partial u}{\partial x}; u_y = \frac{\partial u}{\partial y},$$

and

$$u_t = \frac{\partial u}{\partial t}.$$

We introduce the alternative solution as

$$U(x, y, t) = u(x, y, t) + \epsilon \eta(x, y, t),$$

with the temporal boundary conditions of

$$\eta(x, y, t_0) = \eta(x, y, t_1) = 0.$$

As above

$$I(\epsilon) = \int_{t_0}^{t_1} \int \int_D f(x, y, t, U, U_x, U_y, U_t) dx dy dt,$$

and the extremum is reached when:

$$\frac{\partial I}{\partial \epsilon}|_{\epsilon=0} = \int_{t_0}^{t_1} \int \int_D \left(\frac{\partial f}{\partial u} \eta + \frac{\partial f}{\partial u_x} \eta_x + \frac{\partial f}{\partial u_y} \eta_y + \frac{\partial f}{\partial u_t} \eta_t \right) dx dy dt = 0. \quad (3.1)$$

The last member of the integral may be written as

$$\int_{t_0}^{t_1} \int \int_D \frac{\partial f}{\partial u_t} \eta_t dx dy dt = \int \int_D \int_{t_0}^{t_1} \frac{\partial f}{\partial u_t} \eta_t dt dx dy.$$

Integrating by parts yields

$$\int \int_D \left(\frac{\partial f}{\partial u_t} \eta \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \eta \frac{\partial}{\partial t} \left(\frac{\partial f}{\partial u_t} \right) dt \right) dx dy.$$

Due to the temporal boundary condition the first term vanishes and

$$- \int_{t_0}^{t_1} \int \int_D \eta \frac{\partial}{\partial t} \left(\frac{\partial f}{\partial u_t} \right) dx dy dt$$

remains. The second and third terms of equation (3.1) may be rewritten by Green's identity as follows:

$$\begin{aligned} & \int_{t_0}^{t_1} \int \int_D \left(\frac{\partial f}{\partial u_x} \eta_x + \frac{\partial f}{\partial u_y} \eta_y \right) dx dy dt = \\ & - \int_{t_0}^{t_1} \int \int_D \eta \left(\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial u_x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial u_y} \right) \right) dx dy dt + \\ & \int_{t_0}^{t_1} \int_{\partial D} \eta \left(\frac{\partial f}{\partial u_x} \frac{dy}{ds} + \frac{\partial f}{\partial u_y} \frac{dx}{ds} \right) ds dt. \end{aligned}$$

With these changes, equation (3.1) becomes

$$\begin{aligned} \frac{\partial I}{\partial \epsilon}|_{\epsilon=0} = & \int_{t_0}^{t_1} \left(\int \int_D \eta \left(\frac{\partial f}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial u_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial u_y} \right) - \frac{\partial}{\partial t} \left(\frac{\partial f}{\partial u_t} \right) \right) dx dy + \right. \\ & \left. \int_{\partial D} \eta \left(\frac{\partial f}{\partial u_x} \frac{dy}{ds} - \frac{\partial}{\partial u_y} \frac{dx}{ds} \right) ds \right) dt = 0. \end{aligned}$$

Since the auxiliary function η is arbitrary, by the fundamental lemma of calculus of variations the first integral is only zero when

$$\frac{\partial f}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial u_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial u_y} \right) - \frac{\partial}{\partial t} \left(\frac{\partial f}{\partial u_t} \right) = 0,$$

in the interior of the domain D . This is the Euler-Lagrange differential equation of the problem. Since the boundary conditions of the auxiliary function

were only temporal, the second integral is only zero when

$$\frac{\partial f}{\partial u_x} \frac{dy}{ds} - \frac{\partial}{\partial u_y} \frac{dx}{ds} = 0,$$

on the boundary ∂D . This is the constraint of the variational problem. This result will be utilized in [Chapter 9](#) to solve the elastic membrane problem.

The case of functionals with four independent variables

$$u(x, y, z, t)$$

will also be discussed in [Chapter 10](#) in connection with elasticity problems in solids.

The generalization of the process to even more independent variables is algebraically straightforward. Generalization to more spatial coordinates is not very frequent, although in some manufacturing applications five-dimensional hyperspaces do occur.

4

Higher order derivatives

The fundamental problem of the calculus of variations involved the first derivative of the unknown function. In this chapter we will allow the presence of higher order derivatives.

4.1 The Euler-Poisson equation

First let us consider the variational problem of a functional with a single function, but containing its higher derivatives:

$$I(y) = \int_{x_0}^{x_1} f(x, y, y', \dots, y^{(m)}) dx.$$

Accordingly, boundary conditions for all derivatives will also be given as

$$\begin{aligned} y(x_0) &= y_0, y(x_1) = y_1, \\ y'(x_0) &= y'_0, y'(x_1) = y'_1, \\ y''(x_0) &= y''_0, y''(x_1) = y''_1, \end{aligned}$$

and so on until

$$y^{(m-1)}(x_0) = y_0^{(m-1)}, y^{(m-1)}(x_1) = y_1^{(m-1)}.$$

As in the past chapters, we introduce an alternative solution of

$$Y(x) = y(x) + \epsilon \eta(x),$$

where the arbitrary auxiliary function $\eta(x)$ is continuously differentiable on the interval $x_0 \leq x \leq x_1$ and satisfies

$$\eta(x_0) = 0, \eta(x_1) = 0.$$

The variational problem in terms of the alternative solution is

$$I(\epsilon) = \int_{x_0}^{x_1} f(x, Y, Y', \dots, Y^{(m)}) dx.$$

The differentiation with respect to ϵ follows

$$\frac{dI}{d\epsilon} = \int_{x_0}^{x_1} \frac{d}{d\epsilon} f(x, Y, Y', \dots, Y^{(m)}) dx,$$

and by using the chain rule the integrand is reshaped as

$$\frac{\partial f}{\partial Y} \frac{dY}{d\epsilon} + \frac{\partial f}{\partial Y'} \frac{dY'}{d\epsilon} + \frac{\partial f}{\partial Y''} \frac{dY''}{d\epsilon} + \dots + \frac{\partial f}{\partial Y^{(m)}} \frac{dY^{(m)}}{d\epsilon}.$$

Substituting the alternative solution and its derivatives with respect to ϵ the integrand yields

$$\frac{\partial f}{\partial Y} \eta + \frac{\partial f}{\partial Y'} \eta' + \frac{\partial f}{\partial Y''} \eta'' + \dots + \frac{\partial f}{\partial Y^{(m)}} \eta^{(m)}.$$

Hence the functional becomes

$$\frac{dI}{d\epsilon} = \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial Y} \eta + \frac{\partial f}{\partial Y'} \eta' + \frac{\partial f}{\partial Y''} \eta'' + \dots + \frac{\partial f}{\partial Y^{(m)}} \eta^{(m)} \right) dx.$$

Integrating by terms results in

$$\frac{dI}{d\epsilon} = \int_{x_0}^{x_1} \frac{\partial f}{\partial Y} \eta dx + \int_{x_0}^{x_1} \frac{\partial f}{\partial Y'} \eta' dx + \int_{x_0}^{x_1} \frac{\partial f}{\partial Y''} \eta'' dx + \dots + \int_{x_0}^{x_1} \frac{\partial f}{\partial Y^{(m)}} \eta^{(m)} dx,$$

and integrating by parts produces

$$\begin{aligned} \frac{dI}{d\epsilon} = & \int_{x_0}^{x_1} \eta \frac{\partial f}{\partial Y} dx - \int_{x_0}^{x_1} \eta \frac{d}{dx} \frac{\partial f}{\partial Y'} dx + \int_{x_0}^{x_1} \eta \frac{d^2}{dx^2} \frac{\partial f}{\partial Y''} dx - \\ & \dots (-1)^m \int_{x_0}^{x_1} \eta \frac{d^{(m)}}{dx^{(m)}} \frac{\partial f}{\partial Y^{(m)}} dx. \end{aligned}$$

Factoring the auxiliary function and combining the terms again simplifies to

$$\frac{dI}{d\epsilon} = \int_{x_0}^{x_1} \eta \left(\frac{\partial f}{\partial Y} - \frac{d}{dx} \frac{\partial f}{\partial Y'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial Y''} - \dots (-1)^m \frac{d^{(m)}}{dx^{(m)}} \frac{\partial f}{\partial Y^{(m)}} \right) dx.$$

Finally the extremum at $\epsilon = 0$ and the fundamental lemma produces the **Euler-Poisson equation**

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y''} - \dots (-1)^m \frac{d^{(m)}}{dx^{(m)}} \frac{\partial f}{\partial y^{(m)}} = 0.$$

The Euler-Poisson equation is an ordinary differential equation of order $2m$ and requires the aforementioned $2m$ boundary conditions, where m is the highest order derivative contained in the functional.

For example, the simple $m = 2$ functional

$$I(y) = \int_{x_0}^{x_1} (y^2 - (y'')^2) dx$$

results in the derivatives

$$\frac{\partial f}{\partial y''} = -2y'',$$

and

$$\frac{\partial f}{\partial y} = 2y.$$

The corresponding Euler-Poisson equation derivative term is

$$\frac{d^2}{dx^2} \frac{\partial f}{\partial y''} = \frac{d^2}{dx^2} (-2y'') = -2 \frac{d^4}{dx^4} y,$$

and the equation, after cancellation by -2 , becomes

$$\frac{d^4}{dx^4} y - y = 0.$$

Clearly the solution of this may be achieved by classical calculus tools with four boundary conditions. Application problems exploiting this will be addressed in [Chapters 8](#) (the natural spline) and [9](#) (the bending beam).

4.2 The Euler-Poisson system of equations

In the case of a functional with multiple functions along with their higher order derivatives, the problem gets more difficult. Assuming p functions in the functional, the problem is posed in the form of

$$I(y_1, \dots, y_p) = \int_{x_0}^{x_1} f(x, y_1, y_1', \dots, y_1^{(m_1)}, \dots, y_p, y_p', \dots, y_p^{(m_p)}) dx.$$

Note that the highest order of the derivative of the various functions is not necessarily the same. This is a rather straightforward generalization of the case of the last section, leading to a system of Euler-Poisson equations as follows:

$$\begin{aligned} \frac{\partial f}{\partial y_1} - \frac{d}{dx} \frac{\partial f}{\partial y_1'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y_1''} - \dots (-1)^{m_1} \frac{d^{(m_1)}}{dx^{(m_1)}} \frac{\partial f}{\partial y_1^{(m_1)}} &= 0, \\ \dots, \\ \frac{\partial f}{\partial y_p} - \frac{d}{dx} \frac{\partial f}{\partial y_p'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y_p''} - \dots (-1)^{m_p} \frac{d^{(m_p)}}{dx^{(m_p)}} \frac{\partial f}{\partial y_p^{(m_p)}} &= 0. \end{aligned}$$

This is a set of p ordinary differential equations that may or may not be coupled, hence resulting in a varying level of ease of the solution.

4.3 Algebraic constraints on the derivative

It is also common in engineering applications to impose boundary conditions on some of the derivatives (von Neumann boundary conditions). These result in algebraic constraints posed on the derivative, such as

$$I(y) = \int_{x_0}^{x_1} f(x, y, y') dx$$

subject to

$$g(x, y, y') = 0.$$

In order to be able to solve such problems, we need to introduce a Lagrange multiplier as a function of the independent variable as

$$h(x, y, y', \lambda) = f(x, y, y') + \lambda(x)g(x, y, y').$$

The use of this approach means that the functional now contains two unknown functions and the variational problem becomes

$$I(y, \lambda) = \int_{x_0}^{x_1} h(x, y, y', \lambda) dx,$$

with the original boundary conditions, but without a constraint. The solution is obtained for the unconstrained, but two function case by a system of two Euler-Lagrange equations.

Derivative constraints may also be applied to the case of higher order derivatives. The second order problem of

$$I(y) = \int_{x_0}^{x_1} f(x, y, y', y'') dx$$

may be subject to a constraint

$$g(x, y, y', y'') = 0.$$

In order to be able to solve such problems, we also introduce a Lagrange multiplier function as

$$h(x, y, y', y'', \lambda) = f(x, y, y', y'') + \lambda(x)g(x, y, y', y'').$$

The result is a variational problem of two functions with higher order derivatives as

$$I(y, \lambda) = \int_{x_0}^{x_1} h(x, y, y', y'', \lambda) dx.$$

Hence the solution may be obtained by the application of a system of two Euler-Poisson equations.

Finally, derivative constraints may also be applied to a variational problem originally exhibiting multiple functions, such as

$$I(y, z) = \int_{x_0}^{x_1} f(x, y, y', z, z') dx$$

subject to

$$g(x, y, y', z, z') = 0.$$

Here the new functional is

$$h(x, y, y', z, z', \lambda) = f(x, y, y', z, z') + \lambda(x)g(x, y, y', z, z').$$

Following above, this problem translates into the unconstrained form of

$$I(y, z, \lambda) = \int_{x_0}^{x_1} f(x, y, y', z, z', \lambda) dx$$

that may be solved by a system of three Euler-Lagrange differential equations

$$\frac{\partial h}{\partial y} - \frac{d}{dx} \frac{\partial h}{\partial y'} = 0,$$

$$\frac{\partial h}{\partial z} - \frac{d}{dx} \frac{\partial h}{\partial z'} = 0,$$

and

$$\frac{\partial h}{\partial \lambda} - \frac{d}{dx} \frac{\partial h}{\partial \lambda'} = 0.$$

For example, the variational problem of

$$I(y, z) = \int_{x_0}^{x_1} (y^2 - z^2) dx = \text{extremum},$$

under the derivative constraint of

$$y' - y + z = 0,$$

results in

$$h(x, y, y', z, z', \lambda) = y^2 - z^2 + \lambda(x)(y' - y + z).$$

The solution is obtained from the following three equations

$$2y - \lambda + \lambda' = 0,$$

$$-2z + \lambda = 0,$$

and

$$y' - y + z = 0.$$

The elimination of the Lagrange multiplier results in the system of

$$y - z + z' = 0,$$

and

$$y' - y + z = 0,$$

whose solution follows from classical calculus.

4.4 Application: Linearization of second order problems

It is very common in engineering practice that the highest derivative of interest is of second order. Accelerations in engineering analysis of motion, curvature in description of space curves, and other important application concepts are tied to the second derivative.

This specific case of quadratic problems may be reverted to a linear problem involving two functions. Consider

$$I(y) = \int_{x_0}^{x_1} f(x, y, y', y'') dx = \text{extremum}.$$

with the following boundary conditions given

$$y(x_0), y(x_1), y'(x_0), y'(x_1).$$

By introducing a new function

$$z(x) = y'(x),$$

we can reformulate the unconstrained second order variational problem as a variational problem of the first order with multiple functions in the integrand

$$I(y, z) = \int_{x_0}^{x_1} f(x, y, z, z') dx = \text{extremum},$$

but subject to a constraint involving the derivative

$$g(x, y, z) = z - y' = 0.$$

Using a Lagrange multiplier function in the form of

$$h(x, y, z, z', \lambda) = f(x, y, z, z') + \lambda(x)(z - y'),$$

and following the process laid out in the last section we can produce a system of three Euler-Lagrange differential equations.

$$\begin{aligned}\frac{\partial h}{\partial y} - \frac{d}{dx} \frac{\partial h}{\partial y'} &= \frac{\partial f}{\partial y} - \frac{d\lambda}{dx} = 0, \\ \frac{\partial h}{\partial z} - \frac{d}{dx} \frac{\partial h}{\partial z'} &= \frac{\partial f}{\partial z} + \lambda - \frac{d}{dx} \frac{\partial f}{\partial z'} = 0,\end{aligned}$$

and

$$\frac{\partial h}{\partial \lambda} - \frac{d}{dx} \frac{\partial h}{\partial \lambda'} = z - y'.$$

This may, of course, be turned into the Euler-Poisson equation by expressing

$$\lambda = \frac{d}{dx} \frac{\partial f}{\partial z'} - \frac{\partial f}{\partial z}$$

from the middle equation and differentiating as

$$\frac{d\lambda}{dx} = \frac{d^2}{dx^2} \frac{\partial f}{\partial z'} - \frac{d}{dx} \frac{\partial f}{\partial z}.$$

Substituting this and the third equation into the first yields the Euler-Poisson equation we could have achieved, had we approached the original quadratic problem directly:

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y''} = 0.$$

Depending on the particular application circumstance, the linear system of Euler-Lagrange equations may be more conveniently solved than the quadratic Euler-Poisson equation.

5

The inverse problem of the calculus of variations

It is often the case that the engineer starts from a differential equation with certain boundary conditions which is difficult to solve. Executing the inverse of the Euler-Lagrange process and obtaining the variational formulation of the boundary value problem may also be advantageous.

It is not necessarily easy, or may not even be possible to reconstruct the variational problem from a differential equation. For differential equations, partial or ordinary, containing a linear, self-adjoint, positive operator, the task may be accomplished. Such an operator exhibits

$$(Au, v) = (u, Av),$$

where the parenthesis expression denotes a scalar product in the function space of the solution of the differential equation. Positive definiteness of the operator means

$$(Au, u) \geq 0,$$

with zero attained only for the trivial ($u = 0$) solution. Let us consider the differential equation of

$$Au = f,$$

where the operator obeys the above conditions and f is a known function. If the differential equation has a solution, it corresponds to the minimum value of the functional

$$I(u) = (Au, u) + 2(u, f).$$

This may be proven by simply applying the appropriate Euler-Lagrange equation to this functional.

5.1 The variational form of Poisson's equation

We demonstrate the inverse process through the example of Poisson's equation, a topic of much interest for engineers:

$$\Delta u(x) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y).$$

Here the left-hand side is the well-known Laplace operator. We impose Dirichlet type boundary conditions on the boundary of the domain of interest.

$$u(x, y) = 0; (x, y) \in \partial D,$$

where D is the domain of solution and ∂D is its boundary. According to the above proposition, we need to compute

$$(Au, u) = \int \int_D u \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) dx dy.$$

Applying some vector calculus results in

$$(Au, u) = \int_{\partial D} \left(u \frac{\partial u}{\partial y} dx - u \frac{\partial u}{\partial x} dy \right) + \int \int_D \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right) dx dy.$$

Due to the boundary conditions, the first term vanishes and we obtain

$$(Au, u) = \int \int_D \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right) dx dy.$$

The right-hand side term of the differential equation is processed as

$$(u, f) = \int \int_D u f(x, y) dx dy.$$

The variational formulation of Poisson's equation finally is

$$\int \int_D \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + 2uf \right) dx dy = \int \int_D F(x, y, u, u_x, u_y) dx dy = \text{extremum}.$$

To prove this, we will apply the Euler-Lagrange equation developed in [Section 3.3](#). The terms for this particular case are:

$$\frac{\partial F}{\partial u} = 2f,$$

$$\frac{\partial}{\partial x} \frac{\partial F}{\partial u_x} = \frac{\partial}{\partial x} 2u_x = 2 \frac{\partial^2 u}{\partial x^2},$$

and

$$\frac{\partial}{\partial y} \frac{\partial F}{\partial u_y} = \frac{\partial}{\partial y} 2u_y = 2 \frac{\partial^2 u}{\partial y^2}.$$

The resulting equation of

$$2f - 2\frac{\partial^2 u}{\partial x^2} - 2\frac{\partial^2 u}{\partial y^2} = 0$$

is clearly equivalent with Poisson's equation.

5.2 The variational form of eigenvalue problems

Eigenvalue problems of various kinds may also be formulated as variational problems. We consider the equation of the form

$$\Delta u(x) - \lambda u(x) = 0, \quad (5.1)$$

where the unknown function $u(x)$ defined on domain D is the eigensolution and λ is the eigenvalue. The boundary condition is imposed as

$$u(x, y) = 0,$$

on the perimeter ∂D of the domain D . The corresponding variational problem is of the form

$$I = \int \int_D \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right) dx dy = \text{extremum}, \quad (5.2)$$

under the condition of

$$g(x, y) = \int \int_D u^2(x, y) dx dy = 1.$$

This relation is proven as follows. Following the Lagrange solution of constrained variational problems introduced in [Section 2.2](#), we can write

$$h(x, y) = u(x, y) + \lambda g(x, y),$$

and

$$I = \int \int_D \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \lambda u^2(x, y) \right) dx dy.$$

Note that the λ is only in the role of the Lagrange multiplier yet, although its name hints at its final meaning as well. Introducing

$$U(x, y) = u(x, y) + \epsilon \eta(x, y)$$

the variational form becomes

$$I(\epsilon) = \int \int_D \left(\left(\frac{\partial u}{\partial x} + \epsilon \frac{\partial \eta}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} + \epsilon \frac{\partial \eta}{\partial y} \right)^2 + \lambda (u + \epsilon \eta)^2 \right) dx dy.$$

The extremum is reached when

$$\frac{dI(\epsilon)}{d\epsilon}\bigg|_{\epsilon=0} = 0,$$

which gives rise to the equation

$$2 \int \int_D \left(\frac{\partial u}{\partial x} \frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \eta}{\partial y} + \lambda u \eta \right) dx dy = 0. \quad (5.3)$$

Green's identity in its original three-dimensional form was exploited on several occasions earlier; here we apply it for the special vector field

$$\eta \nabla u$$

in a two-dimensional domain. The result is

$$\int \int_D (\nabla \eta \cdot \nabla u) dA = \int_{\partial D} \eta (\nabla u \cdot \underline{n}) ds - \int \int_D \eta \nabla^2 u dA.$$

Since the tangent of the circumference is in the direction of

$$dx \underline{i} + dy \underline{j},$$

the unit normal may be computed as

$$\underline{n} = \frac{dy \underline{i} - dx \underline{j}}{\sqrt{dx^2 + dy^2}}.$$

Finally utilizing the arc length formula of

$$ds = \sqrt{dx^2 + dy^2},$$

the line integral over the circumference of the domain becomes

$$\int_{\partial D} \eta \left(\frac{\partial u}{\partial x} dy - \frac{\partial u}{\partial y} dx \right).$$

Applying above for the first two terms of equation (5.3) results in

$$\begin{aligned} & \int \int_D \left(\frac{\partial u}{\partial x} \frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \eta}{\partial y} \right) dx dy = \\ & \int \int_{\partial D} \left(\eta \left(\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} \right) \cdot \underline{n} \right) ds - \int \int_D \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \eta dx dy = \\ & \int_{\partial D} \frac{\partial u}{\partial x} \eta dy - \frac{\partial u}{\partial y} \eta dx - \int \int_D \Delta u \eta dx dy. \end{aligned}$$

The integral over the boundary vanishes due to the assumption on η and substituting the remainder part into equation (5.3) we obtain

$$-2 \int \int_D (\Delta u - \lambda u) \eta dx dy = 0.$$

Since $\eta(x, y)$ is arbitrarily chosen, in order to satisfy this equation

$$\Delta u - \lambda u = 0$$

must be satisfied. Thus we have established that equation (5.2) is indeed the variational form of equation (5.1) and the Lagrange multiplier is the eigenvalue.

5.2.1 Orthogonal eigensolutions

The eigenvalue problem has an infinite sequence of eigenvalues and for each eigenvalue there exists a corresponding eigensolution that is unique apart from a constant factor. Hence the variational form should also provide means for the solution of multiple pairs.

Let us denote the series of eigenpairs as

$$(\lambda_1, u_1), (\lambda_2, u_2), \dots (\lambda_n, u_n).$$

Assuming that we have already found the first pair satisfying

$$\Delta u_1 - \lambda_1 u_1 = 0,$$

we seek the second solution $u_2, \lambda_2 \neq \lambda_1$ following the process laid out in the last section. Then for any arbitrary auxiliary function η it follows that

$$\int \int_D \left(\frac{\partial u_2}{\partial x} \frac{\partial \eta}{\partial x} + \frac{\partial u_2}{\partial y} \frac{\partial \eta}{\partial y} + \lambda_2 u_2 \eta \right) dx dy = 0.$$

Applying an auxiliary function of the special form of

$$\eta = u_1,$$

we obtain

$$\int \int_D \left(\frac{\partial u_2}{\partial x} \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} \frac{\partial u_1}{\partial y} + \lambda_2 u_2 u_1 \right) dx dy = 0.$$

The same argument for the first solution results in

$$\int \int_D \left(\frac{\partial u_1}{\partial x} \frac{\partial \eta}{\partial x} + \frac{\partial u_1}{\partial y} \frac{\partial \eta}{\partial y} + \lambda_1 u_1 \eta \right) dx dy = 0.$$

Applying an auxiliary function of the special form of

$$\eta = u_2,$$

we obtain

$$\int \int_D \left(\frac{\partial u_1}{\partial x} \frac{\partial u_2}{\partial x} + \frac{\partial u_1}{\partial y} \frac{\partial u_2}{\partial y} + \lambda_1 u_1 u_2 \right) dx dy = 0.$$

Subtracting the equations and canceling the identical terms results in

$$(\lambda_2 - \lambda_1) \int \int_D u_1 u_2 dx dy = 0.$$

Since

$$\lambda_1 \neq \lambda_2,$$

it follows that

$$\int \int_D u_2 u_1 dx dy = 0$$

must be true. The two eigensolutions are orthogonal. With similar arguments and specially selected auxiliary functions, it is also easy to show that the second solutions also satisfy

$$\Delta u_2 - \lambda_2 u_2 = 0.$$

The subsequent eigensolutions may be found by the same procedure and the sequence of the eigenpairs attain the extrema of the variational problem under the successive conditions of the orthogonality against the preceding solutions.

5.3 Application: Sturm-Liouville problems

The process demonstrated in the last section in connection with Laplace's operator may be applied to arrive at eigenvalues and eigensolutions of other differential equations as well. The differential equations of the form

$$-\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y(x) = \lambda y(x),$$

are called the Sturm-Liouville differential equations. Here the unknown solution function $y(x)$ is the eigensolution and λ is the eigenvalue. The known functions $p(x)$ and $q(x)$ are continuous and continuously differentiable, respectively. The boundary conditions imposed are

$$y(x_0) = 0$$

and

$$y(x_1) = 0.$$

The corresponding variational problem is posed as

$$I(y) = \int_{x_0}^{x_1} (p(x)y'^2(x) + q(x)y^2(x))dx = \text{extremum},$$

subject to above boundary conditions and the additional constraint of

$$\int_{x_0}^{x_1} y^2(x) dx = 1.$$

The engineering importance of these problems lies in the fact that, depending on the choice of the coefficient functions $p(x), q(x)$, various influential functions may be generated as the eigensolutions. For example, the Bessel functions, Hermite, Chebishev and Laguerre polynomials may be derived from Sturm-Liouville equations with various selections of the coefficient functions.

The simplest form of the Sturm-Liouville equations is with

$$p(x) = 1$$

and

$$q(x) = 0,$$

leading to

$$-\frac{d}{dx}\left(\frac{dy}{dx}\right) = \lambda y(x).$$

Straightforward integration indicates the possibility of a solution in the form of trigonometric functions. Surely, a solution of the form

$$y(x) = c_i \sin(ix), i = 1, 2, \dots$$

would satisfy the equation with a judicious choice of the constant c_i and boundary conditions. We will assume boundary conditions of

$$y(0) = 0$$

and

$$y(\pi) = 0.$$

The selection of this boundary is for the convenience of dealing with this function family and does not restrict the generality of the discussion. Furthermore, it is well established that

$$\int_0^\pi \sin(ix) \sin(jx) dx = 0; i \neq j,$$

hence these solutions are orthogonal. The constraint equation is easy to evaluate, since for $i = 1$ for example

$$\int_0^\pi (c_1 \sin(1x))^2 dx = c_1^2 \int_0^\pi \sin^2(x) dx = c_1^2 \left(\frac{1}{2} x \Big|_0^\pi - \frac{1}{4} \sin(2x) \Big|_0^\pi \right).$$

Substituting the boundary values results in

$$c_1^2 \frac{\pi}{2} = 1,$$

producing the constant necessary to satisfy the constraint:

$$c_1 = \sqrt{\frac{2}{\pi}}.$$

The generic solution may be obtained from the form

$$\int_0^\pi (c_i \sin(ix))^2 dx = c_i^2 \int_0^\pi \sin^2(ix) dx = c_i^2 \left(\frac{1}{2} x \Big|_0^\pi - \frac{1}{4i} \sin(2ix) \Big|_0^\pi \right).$$

Considering the boundary conditions, the second term vanishes and

$$c_i^2 \frac{\pi}{2} = 1,$$

which yields the generic coefficients as

$$c_i = \sqrt{\frac{2}{\pi}}.$$

In order to establish the eigenvalues, we execute the operations posted by the differential equation for the generic solution.

$$-\frac{d}{dx}(c_i \sin'(ix)) = -\frac{d}{dx}(ic_i \cos(ix)) = i^2 c_i \sin(ix),$$

which must be equal to the right-hand side of

$$\lambda c_i \sin(ix).$$

This results in the eigenvalue of

$$\lambda = i^2.$$

A more generic discussion of this case is presented in [8]. It is noteworthy that even this simplest form of Sturm-Liouville problems leads to an engineering application, the vibrating string problem, the subject of [Section 9.1](#).

5.3.1 Legendre's equation and polynomials

A very important sub-case of the Sturm-Liouville problems is when

$$p(x) = 1 - x^2,$$

along with

$$q(x) = 0.$$

The form of the equation becomes

$$\frac{d}{dx}((1 - x^2) \frac{dy}{dx}) + \lambda y(x) = 0,$$

subject to the boundary conditions

$$y(-1) = 0,$$

and

$$y(1) = 0.$$

The associated variational problem becomes

$$I(y) = \int_{-1}^{+1} (1 - x^2) \left(\frac{dy}{dx} \right)^2 dx = \text{extremum},$$

subject to the constraint

$$\int_{-1}^{+1} y^2 dx = 1.$$

It is easy to obtain the Euler-Lagrange equation of this constrained variational problem as

$$\frac{\partial}{\partial y}((1 - x^2)y'^2 - \lambda y^2) - \frac{d}{dx} \left(\frac{\partial}{\partial y'}((1 - x^2)y'^2 - \lambda y^2) \right) = 0.$$

Here λ again is only the Lagrange multiplier connecting the constraint, but its final disposition is pre-ordained. The equation with some algebra simplifies to

$$\frac{d}{dx}((1 - x^2)y') + \lambda y = 0,$$

which is of course the equation we started from. The eigenvalues of this problem are of the form

$$\lambda = i(i + 1),$$

and the eigensolutions are

$$y(x) = c_i Le_i(x).$$

Here c_i are constants and Le_i are the Legendre polynomials. The first few values of the eigenpairs are shown in [Table 5.1](#).

The eigensolution functions are shown graphically in [Figure 5.1](#).

The first eigensolution with

$$\lambda_0 = 0$$

and

$$y_0(x) = 1$$

TABLE 5.1
Eigenpairs of Legendre
equation

i	λ	$Le_i(x)$	c_i
0	0	1	$\sqrt{1/2}$
1	2	x	$\sqrt{3/2}$
2	6	$(3x^2 - 1)/2$	$\sqrt{5/2}$
3	12	$(5x^3 - 3x)/2$	$\sqrt{7/2}$

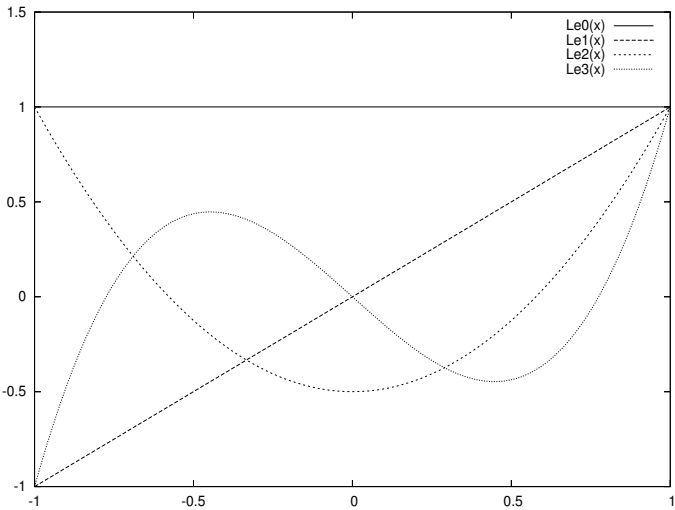


FIGURE 5.1 Eigensolutions of Legendre’s equation

is clearly a trivial solution, since both terms of the left-hand side of the equation vanish. The first non-trivial solution may be verified as

$$\frac{d}{dx}((1 - x^2)(x)') + 2x = \frac{d}{dx}(1 - x^2) + 2x = -2x + 2x = 0.$$

Furthermore, the satisfaction of the constraint equations is seen as

$$\int_{-1}^{+1} y^2 dx = c_1^2 \int_{-1}^{+1} x^2 dx = c_1^2 \frac{x^3}{3} \Big|_{-1}^{+1} = c_1^2 \frac{2}{3}.$$

The constant of

$$c_1 = \sqrt{\frac{3}{2}}$$

will enforce the satisfaction of

$$\int_{-1}^{+1} y^2 dx = 1.$$

In fact, the generic form of the constants may be obtained by

$$c_i = \sqrt{\frac{2i+1}{2}}.$$

Finally, Sturm-Liouville problems may also be presented with functions of multiple variables. For example in three dimensions, the equation becomes

$$-\nabla(p(x, y, z)\nabla u(x, y, z)) + q(x, y, z)u(x, y, z) = \lambda u(x, y, z)$$

leading to various elliptic partial differential equations that all have engineering implications.

6

Direct methods of calculus of variations

Most of the applications insofar were solved in analytical forms. Application problems in engineering practice, however, may not be easily solved by such techniques, if solvable at all. Hence, before we embark on applications, it seems prudent to discuss solution techniques that are amenable for practical problems.

It was mentioned in the introduction that the solution of the Euler-Lagrange differential equation resulting from a certain variational problem may not be easy. This gave rise to the idea of directly solving the variational problem.

The direct methods produce approximate solutions and as such, sometimes are also called numerical solutions. The two fundamental techniques are Euler's method and the Ritz method. The methods of Galerkin and Kantorovich, both described in [9], could be considered extensions of Ritz's, are the most well-known by engineers and used in the industry.

6.1 Euler's method

Euler proposed a direct solution for the variational problem of

$$I(y) = \int_{x_0}^{x_n} f(x, y, y') dx = \text{extremum}$$

with the boundary conditions

$$y(x_0) = y_0; y(x_n) = y_n,$$

by subdividing the interval of the independent variable as

$$x_i = x_0 + i \frac{x_n - x_0}{n}; i = 1, 2, \dots, n.$$

Introducing

$$h = \frac{x_n - x_0}{n},$$

the functional may be approximated as

$$I(y_i) = \int_{x_0}^{x_1} f(x_i, y_i, y'_i) = h \sum_{i=1}^{n-1} f(x_0 + ih, y_i, \frac{y_{i+1} - y_i}{h}) dx = \text{extremum}.$$

Here the approximated solution values y_i are the unknowns and the extremum may be found by differentiation:

$$\frac{\partial I}{\partial y_i} = 0.$$

The process is rather simple and follows from Euler's other work in the numerical solution of differential equations. For illustration, we consider the following problem:

$$I(y) = \int_0^1 (2xy + y^2 + y'^2) dx = \text{extremum},$$

with the boundary conditions

$$y(0) = y(1) = 0.$$

Let us subdivide the interval into $n = 5$ equidistant segments with

$$h = 0.2,$$

and

$$x_i = 0.2i.$$

The approximate functional with the appropriate substitutions becomes

$$I(y_i) = 0.2 \sum_{i=1}^4 (0.4iy_i + y_i^2 + (5(y_{i+1} - y_i))^2).$$

The computed partial derivatives are

$$\begin{aligned} \frac{\partial I}{\partial y_1} &= 0.2(0.4 + 2y_1 - \frac{2(y_2 - y_1)}{0.04}) = 0, \\ \frac{\partial I}{\partial y_2} &= 0.2(0.8 + 2y_2 - \frac{2(y_3 - y_2)}{0.04} + \frac{2(y_2 - y_1)}{0.04}) = 0, \\ \frac{\partial I}{\partial y_3} &= 0.2(1.2 + 2y_3 - \frac{2(y_4 - y_3)}{0.04} + \frac{2(y_3 - y_2)}{0.04}) = 0, \end{aligned}$$

and

$$\frac{\partial I}{\partial y_4} = 0.2(1.6 + 2y_4 + \frac{2y_4}{0.04} + \frac{2(y_4 - y_3)}{0.04}) = 0.$$

This system of four equations yields the values of the approximate solution. The analytic solution of this problem is

$$y(x) = -x + e \frac{e^x - e^{-x}}{e^2 - 1}.$$

The comparison of the Euler solution (y_i) and the analytic solution ($y(x_i)$) at the four discrete points is shown in Table 6.1.

TABLE 6.1
Accuracy of Euler's
method

i	x_i	y_i	$y(x_i)$
1	0.2	-0.0286	-0.0287
2	0.4	-0.0503	-0.0505
3	0.6	-0.0580	-0.0583
4	0.8	-0.0442	-0.0444

The boundary solutions of $y(0)$ and $y(1) = 0$ are not shown since they are in full agreement by definition.

6.2 Ritz method

Let us consider the variational problem of

$$I(y) = \int_{x_0}^{x_1} f(x, y, y') dx = \text{extremum},$$

under the boundary conditions

$$y(x_0) = y_0; y(x_1) = y_1.$$

The Ritz method is based on an approximation of the unknown solution function with a linear combination of certain basis functions. Finite element or spline-based approximations are the most commonly used and will be subject of detailed discussion in [Chapters 8](#) and [10](#). Let the unknown function be approximated with

$$\overline{y}(x) = \alpha_0 b_0(x) + \alpha_1 b_1(x) + \dots + \alpha_n b_n(x),$$

where the basis functions are also required to satisfy the boundary conditions and the coefficients are as yet unknown. Substituting the approximate solution into the variational problem results in

$$I(\overline{y}) = \int_{x_0}^{x_1} f(x, \overline{y}, \overline{y}') dx = \text{extremum}.$$

In order to reach an extremum of the functional, it is necessary that the derivatives with respect to the unknown coefficients vanish:

$$\frac{\partial I(\bar{y})}{\partial \alpha_i} = 0; i = 0, 1, \dots, n.$$

It is not intuitively clear that the approximated function approaches the extremum of the original variational problem, but it has been proven, for example in [9]. Let us just demonstrate the process with a small analytic example. Consider the variational problem of

$$I(y) = \int_0^1 y'^2(x) dx = \text{extremum},$$

with the boundary conditions

$$y(0) = y(1) = 0,$$

and constraint of

$$\int_0^1 y^2(x) dx = 1.$$

Since this is a constrained problem, we apply the Lagrange multiplier technique and rewrite the variational problem as

$$I(y) = \int_0^1 (y'^2(x) - \lambda y^2) dx = \text{extremum}.$$

Let us use, for example, the basis functions of

$$b_0(x) = x(x-1)$$

and

$$b_1(x) = x^2(x-1).$$

It is trivial to verify that these also obey the boundary conditions. The approximated solution function is

$$\bar{y} = \alpha_0 x(x-1) + \alpha_1 x^2(x-1).$$

The functional of the constrained, approximated variational problem is

$$I(\bar{y}) = \int_0^1 (\bar{y}'^2 - \lambda \bar{y}^2) dx.$$

Evaluating the integral yields

$$I(\bar{y}) = \frac{1}{3}(\alpha_0^2 + \alpha_0\alpha_1 + \frac{2}{5}\alpha_1^2) - \lambda(\frac{1}{30}\alpha_0^2 + \frac{1}{30}\alpha_0\alpha_1 + \frac{1}{105}\alpha_1^2).$$

The extremum requires the satisfaction of

$$\frac{\partial I}{\partial \alpha_0} = \alpha_0 \left(\frac{2}{3} - \frac{\lambda}{15} \right) + \alpha_1 \left(\frac{1}{3} - \frac{\lambda}{30} \right) = 0$$

and

$$\frac{\partial I}{\partial \alpha_1} = \alpha_0 \left(\frac{1}{3} - \frac{\lambda}{30} \right) + \alpha_1 \left(\frac{4}{15} - \frac{2\lambda}{105} \right) = 0.$$

A nontrivial solution of this system of equations is obtained by setting its determinant to zero, resulting in the following quadratic equation

$$\frac{\lambda^2 - 52\lambda + 420}{6300} = 0.$$

Its solutions are

$$\lambda_1 = 10; \lambda_2 = 42.$$

Using the first value and substituting into the second condition yields

$$\alpha_1 = 0$$

with arbitrary α_0 . Hence

$$\bar{y}(x) = \alpha_0 x(x-1).$$

The condition

$$\int_0^1 \bar{y}^2 dx = \int_0^1 \alpha_0^2 x^2 (x-1)^2 dx = 1$$

results in

$$\alpha_0 = \pm \sqrt{30}.$$

The approximate solution of the variational problem is

$$\bar{y}(x) = \pm \sqrt{30} x(x-1).$$

It is very important to point out that the solution obtained as a function of the chosen basis functions is not the analytic solution of the variational problem. For this particular example the corresponding Euler-Lagrange differential equation is

$$y'' + \lambda y = 0$$

whose analytic solution, based on [Section 5.3](#), is

$$y = \pm \sqrt{2} \sin(\pi x).$$

[Figure 6.1](#) compares the analytic and the approximate solutions and plots the error of the latter.

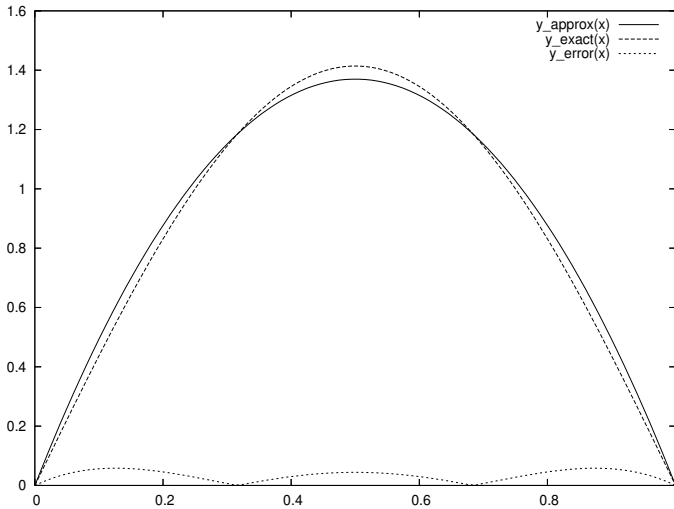


FIGURE 6.1 Accuracy of the Ritz solution

The figure demonstrates that the Ritz solution satisfies the boundary conditions and shows acceptable differences in the interior of the interval. Finally, the variational problem's extremum is computed for both cases. The analytical solution is based on the derivative

$$y' = \sqrt{2\pi} \cos(\pi x),$$

and obtained as

$$\int_0^1 y'^2(x) dx = 2\pi^2 \int_0^1 \cos^2(\pi x) dx = \pi^2 = 9.87.$$

The Ritz solution's derivative is

$$\bar{y}' = -\sqrt{30}(2x - 1).$$

and the approximate extremum is

$$\int_0^1 \bar{y}'^2(x) dx = 30 \int_0^1 (2x - 1)^2 dx = \frac{30}{3} = 10.$$

The approximate extremum is slightly higher than the analytic extremum, but by only a very acceptable error.

6.2.1 Application: Solution of Poisson's equation

The second order boundary value problem of Poisson's, introduced earlier, is presented in the variational form of

$$I(y) = \int \int_D \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + 2f(x, y)u(x, y) \right) dx dy$$

whose Euler-Lagrange equation leads to the form

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y),$$

as was shown in the last chapter. For the simplicity of the discussion and without the loss of generality, we impose the boundary condition of

$$u = 0$$

on the perimeter of the domain D . Ritz's method indicates the use of the basis functions

$$b_i(x, y)$$

and demands that they also vanish on the boundary. The approximate solution in this two-dimensional case is

$$\bar{u}(x, y) = \sum_{i=1}^n \alpha_i b_i(x, y).$$

The partial derivatives are

$$\frac{\partial \bar{u}}{\partial x} = \sum_{i=1}^n \alpha_i \frac{\partial b_i(x, y)}{\partial x},$$

and

$$\frac{\partial \bar{u}}{\partial y} = \sum_{i=1}^n \alpha_i \frac{\partial b_i(x, y)}{\partial y}.$$

Substituting the approximate solution into the functional yields

$$I(\bar{u}) = \int \int_D \left(\left(\frac{\partial \bar{u}}{\partial x} \right)^2 + \left(\frac{\partial \bar{u}}{\partial y} \right)^2 + 2f(x, y)\bar{u}(x, y) \right) dx dy.$$

Evaluating the derivatives, this becomes

$$I(\bar{u}) = \int \int_D \left(\left(\sum_{i=1}^n \alpha_i \frac{\partial b_i}{\partial x} \right)^2 + \left(\sum_{i=1}^n \alpha_i \frac{\partial b_i}{\partial y} \right)^2 + 2f(x, y) \right) \sum_{i=1}^n \alpha_i b_i dx dy,$$

which may be reordered into the form

$$I(\bar{u}) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} \alpha_i \alpha_j + \sum_{i=1}^n d_i \alpha_i.$$

The coefficients are

$$c_{ij} = \int \int_D \left(\frac{\partial b_i}{\partial x} \frac{\partial b_j}{\partial x} + \frac{\partial b_i}{\partial y} \frac{\partial b_j}{\partial y} \right) dx dy$$

and

$$d_i = \int \int_D f(x, y) b_i dx dy.$$

As above, the unknown coefficients are solved from the conditions

$$\frac{\partial I(\bar{u})}{\partial \alpha_i} = 0, i = 1, 2, \dots, n,$$

resulting in the linear system of equations

$$\sum_{j=1}^n c_{ij} \alpha_j + d_j = 0, i = 1, 2, \dots, n.$$

It may be shown that the system is nonsingular and always yields a nontrivial solution assuming that the basis functions form a linearly independent set. The computation of the terms of the equations, however, is rather tedious and resulted in the emergence of the next method.

6.3 Galerkin's method

The difference between the Ritz method and that of Galerkin's is in the fact that the latter addresses the differential equation form of a variational problem. Galerkin's method minimizes the residual of the differential equation integrated over the domain with a weight function; hence, it is also called the method of weighted residuals.

This difference lends more generality and computational convenience to Galerkin's method. Let us consider a linear differential equation in two variables:

$$L(u(x, y)) = 0$$

and apply Dirichlet boundary conditions. Galerkin's method is also based on the Ritz approximation of the solution as

$$\bar{u} = \sum_{i=1}^n \alpha_i b_i(x, y),$$

in which case, of course there is a residual of the differential equation

$$L(\bar{u}) \neq 0.$$

Galerkin proposed using the basis functions of the approximate solution also as the weights, and requires that the integral vanishes with a proper selection of the coefficients:

$$\int \int_D L(\bar{u}) b_j(x, y) dx dy = 0; j = 1, 2, \dots, n.$$

This yields a system for the solution of the coefficients as

$$\int \int_D L\left(\sum_{i=1}^n \alpha_i b_i(x, y)\right) b_j(x, y) dx dy = 0; j = 1, 2, \dots, n.$$

This is also a linear system and produces the unknown coefficients α_i . We illustrate the computational technique of Galerkin's method also in connection with Poisson's equation:

$$L(u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - f(x, y) = 0.$$

For this Galerkin's method is presented as

$$\int \int_D \left(\frac{\partial^2 \bar{u}}{\partial x^2} + \frac{\partial^2 \bar{u}}{\partial y^2} - f(x, y) \right) b_j dx dy = 0, j = 1, \dots, n.$$

Therefore

$$\int \int_D \left(\sum_{i=1}^n \alpha_i \frac{\partial^2 b_i}{\partial x^2} + \sum_{i=1}^n \alpha_i \frac{\partial^2 b_i}{\partial y^2} - f(x, y) \right) b_j dx dy = 0, j = 1, \dots, n.$$

Reordering yields

$$\sum_{i=1}^n \alpha_i \int \int_D \left(\frac{\partial^2 b_i}{\partial x^2} + \frac{\partial^2 b_i}{\partial y^2} \right) b_j dx dy - \int \int_D f(x, y) b_j dx dy = 0, j = 1, \dots, n.$$

The system of equations becomes

$$B\mathbf{a} = \mathbf{b}$$

with solution vector of

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix},$$

The system matrix is of the form

$$B = \begin{bmatrix} B_{1,1} & B_{1,2} & \dots & B_{1,n} \\ B_{2,1} & B_{2,2} & \dots & B_{2,n} \\ \dots & \dots & \dots & \dots \\ B_{n,1} & B_{n,2} & \dots & B_{n,n} \end{bmatrix}$$

whose terms are defined as

$$B_{j,i} = \int \int_D \left(\frac{\partial^2 b_i}{\partial x^2} + \frac{\partial^2 b_i}{\partial y^2} \right) b_j dx dy.$$

Finally, the right-hand side vector is

$$\underline{b} = \begin{bmatrix} \int \int_D f(x,y) b_1 dx dy \\ \int \int_D f(x,y) b_2 dx dy \\ \dots \\ \int \int_D f(x,y) b_n dx dy \end{bmatrix}.$$

Note that if Galerkin's and Ritz's methods are applied to the same problem, the approximate solutions found are identical.

6.4 Kantorovich's method

Both the Ritz and Galerkin methods are restricted in their choices of basis functions, because their basis functions are required to satisfy the boundary conditions. The clever method of Kantorovich, described in [9], relaxes this restriction and enables the use of simpler basis functions.

Consider the variational problem of two variables

$$I(u) = \text{extremum}, (x, y) \in D,$$

with boundary conditions

$$u(x, y) = 0, (x, y) \in \partial D.$$

Here ∂D again denotes the boundary of the domain.

The method proposes the construction of a function, such that

$$\omega(x, y) \geq 0, (x, y) \in D,$$

and

$$\omega(x, y) = 0, (x, y) \in \partial D.$$

This function assumes the role of enforcing the boundary condition and the following set of simpler, power functions based, basis functions are adequate to present the solution:

$$\begin{aligned}b_1(x, y) &= \omega(x, y), \\b_2(x, y) &= \omega(x, y)x, \\b_3(x, y) &= \omega(x, y)y, \\b_4(x, y) &= \omega(x, y)x^2, \\b_5(x, y) &= \omega(x, y)xy, \\b_6(x, y) &= \omega(x, y)y^2,\end{aligned}$$

and so on, following the same pattern. It is clear that all these basis functions vanish on the boundary by the virtue of $\omega(x, y)$, even though the power function components do not.

The question is how to generate $\omega(x, y)$ for various shapes of domains. For a centrally located circle with radius r , the equation

$$x^2 + y^2 = r^2$$

implies very intuitively the form of

$$\omega(x, y) = r^2 - x^2 - y^2.$$

Obviously the function is zero everywhere on the circle and nonzero in the interior of the domain. It is also nonzero on the outside of the domain, but that is irrelevant in connection with our problem.

One can also consider a domain consisting of overlapping circular regions, some of which represent voids in the domain. [Figure 6.2](#) shows a domain of two circles with equations

$$x^2 + y^2 = r^2$$

and

$$(x - r/2)^2 + y^2 = (r/2)^2.$$

Reordering the latter yields

$$x^2 - xr + y^2 = 0,$$

and in turn results in

$$\omega(x, y) = (r^2 - x^2 - y^2)(x^2 - rx + y^2).$$

Clearly on the boundary of the larger circle the left term is zero and on the boundary of the smaller circle the right term is zero. Hence the product

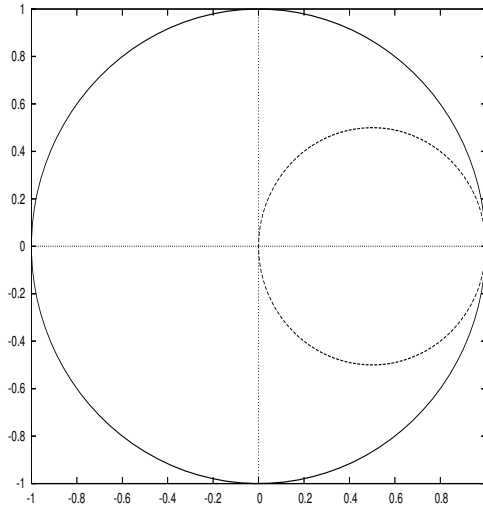


FIGURE 6.2 Domain with overlapping circular regions

function vanishes on the perimeter of both circles, which constitutes the now nontrivial boundary.

Let us now consider the boundary of a rectangle of width $2w$ and height $2h$, also centrally located around the origin. The equations of the sides

$$x = \pm w,$$

and

$$y = \pm h,$$

imply the very simple form of

$$\omega(x, y) = (w^2 - x^2)(h^2 - y^2).$$

The verification is very simple,

$$\omega(x, y) = 0, (x, y) = (\pm w, \pm h).$$

The construction technique clearly shows signs of difficulties to come with very generic, and especially three-dimensional domains. In fact such difficulties limited the practical usefulness of this otherwise innovative method until more recent work enabled the automatic creation of the ω functions for generic

two- or three-dimensional domains with the help of spline functions, a topic which will be discussed in [Chapter 8](#) at length.

We shall now demonstrate the correctness of such a solution. For this we consider the solution of a specific Poisson's equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2,$$

with

$$u(x, y) = 0, (x, y) \in \partial D,$$

where we designate the domain to be the rectangle whose ω function was specified above. We will search for Kantorovich's solution in the form of

$$u(x, y) = (w^2 - x^2)(h^2 - y^2)(\alpha_1 + \alpha_2 x + \alpha_3 y + \dots).$$

The method, as all direct methods, is approximate, so we may truncate the sequence of power function terms at a certain order. It is sufficient for the demonstration to use only the first term.

We will apply the method in connection with Galerkin's method of the last section. Therefore the extremum is sought from

$$\int_{-w}^{+w} \int_{-h}^{+h} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 2 \right) \omega(x, y) dy dx = 0.$$

Executing the posted differentiations and substituting results in

$$\begin{aligned} \int_{-w}^{+w} \int_{-h}^{+h} & -2\alpha_1(w^2 - x^2)(h^2 - y^2)^2 - 2\alpha_1(w^2 - x^2)^2(h^2 - y^2) + \\ & 2(w^2 - x^2)(h^2 - y^2) dy dx = 0. \end{aligned}$$

Since we only have a single coefficient, the system of equations developed earlier boils down to a single scalar equation of

$$b\alpha_1 = f,$$

with

$$b = \int_{-w}^{+w} \int_{-h}^{+h} ((w^2 - x^2)(h^2 - y^2)^2 + (w^2 - x^2)^2(h^2 - y^2)) dy dx,$$

and

$$f = \int_{-w}^{+w} \int_{-h}^{+h} (w^2 - x^2)(h^2 - y^2) dy dx.$$

After the (tedious) evaluation of the integrals, the value of

$$\alpha_1 = \frac{5}{4(w^2 + h^2)}$$

emerges. In turn, the approximate Kantorovich-Galerkin solution is

$$u(x, y) = \frac{5}{4} \frac{(w^2 - x^2)(h^2 - y^2)}{w^2 + h^2}.$$

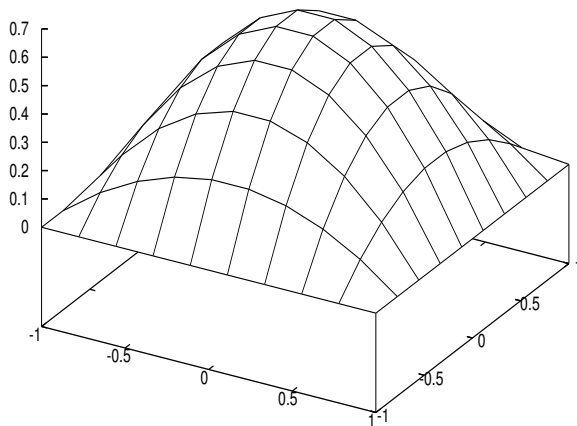


FIGURE 6.3 Solution of Poisson's equation

The solution is depicted graphically in Figure 6.3 using

$$w = h = 1.$$

The figure demonstrates that the solution function satisfies the zero boundary condition on the circumference of the square. To increase accuracy, additional terms of the power series may be used. The method also enables the exploitation of the symmetry of the domain. For example, if the above domain would exhibit the same height as width,

$$s = w = h,$$

the solution may be sought in the form of

$$u(x, y) = (s^2 - x^2)(s^2 - y^2)(\alpha_1 + \alpha_{23}(x + y)),$$

where α_{23} denotes the single constant accompanying both the second and third terms.

A generalization of this approach is necessary to eliminate the difficulties of producing an analytic ω function for practical domains with convoluted boundaries. The idea is to use an approximate solution to generate the ω function as well.

Let us consider the two-dimensional domain case and generate a surface approximation over the domain in the form of

$$\omega(x, y) = \sum_{i=0}^n \sum_{j=0}^m C_{i,j} B_i(x) B_j(y),$$

where the two sets of B basis functions are of common form, but dependent on either of the independent variables. The coefficients $C_{i,j}$ are either sampling points of the domain, or control points used to generate the surface. The latter case applies mainly to the interior points, and the earlier to the boundary.

This requires a simple Cartesian discretization of the domain along topological (possibly even equidistant) lines. The B-spline fitting technique introduced in [Chapter 8](#) will provide the means for accomplishing this.

Part II

Engineering applications

Differential geometry

Differential geometry is a classical mathematical area that has become very important for engineering applications in the recent decades. This importance is based on the rise of computer-aided visualization and geometry generation technologies.

The fundamental problem of differential geometry, the finding of geodesic curves, has practical implications in manufacturing. Development of non-mathematical surfaces used in ships and airplanes has serious financial impact in reducing material waste and improving the quality of the surfaces.

While the discussion in this chapter will focus on analytically solvable problems, the methods and concepts we introduce will provide a foundation applicable in various engineering areas.

7.1 The geodesic problem

Finding a geodesic curve on a surface is a classical problem of differential geometry. Variational calculus seems uniquely applicable to this problem. Let us consider a parametrically given surface

$$\underline{r} = x(u, v)\underline{i} + y(u, v)\underline{j} + z(u, v)\underline{k}.$$

Let two points on the surface be

$$\underline{r}_0 = x(u_0, v_0)\underline{i} + y(u_0, v_0)\underline{j} + z(u_0, v_0)\underline{k},$$

and

$$\underline{r}_1 = x(u_1, v_1)\underline{i} + y(u_1, v_1)\underline{j} + z(u_1, v_1)\underline{k}.$$

The shortest path on the surface between these two points is the **geodesic curve**. Consider the square of the arclength

$$ds^2 = (dx)^2 + (dy)^2 + (dz)^2,$$

and compute the differentials related to the parameters.

$$ds^2 = E(u, v)(du)^2 + 2F(u, v)dudv + G(u, v)(dv)^2.$$

Here the so-called first fundamental quantities are defined as

$$E(u, v) = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2 = (r'_u)^2,$$

$$F(u, v) = \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} + \frac{\partial y}{\partial u} \frac{\partial y}{\partial v} + \frac{\partial z}{\partial u} \frac{\partial z}{\partial v} = r'_u r'_v,$$

and

$$G(u, v) = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2 = (r'_v)^2.$$

Assume that the equation of the geodesic curve in the parametric space is described by

$$v = v(u).$$

Then the geodesic curve is the solution of the variational problem

$$I(v) = \int_{u_0}^{u_1} \sqrt{E(u, v) + 2F(u, v) \frac{dv}{du} + G(u, v) \left(\frac{dv}{du}\right)^2} du = \text{extremum}$$

with boundary conditions

$$v(u_0) = v_0,$$

and

$$v(u_1) = v_1.$$

The corresponding Euler-Lagrange differential equation is

$$\frac{E_v + 2v'F_v + v'^2G_v}{2\sqrt{E(u, v) + 2F(u, v)v' + G(u, v)v'^2}} - \frac{d}{du} \frac{F + Gv'}{\sqrt{E(u, v) + 2F(u, v)v' + G(u, v)v'^2}} = 0,$$

with the notation of

$$E_v = \frac{\partial E}{\partial v}, F_v = \frac{\partial F}{\partial v}, G_v = \frac{\partial G}{\partial v},$$

and

$$v' = \frac{dv}{du}.$$

The equation is rather difficult in general, and exploitation of special cases arising from the particular surface definitions is advised.

When the first fundamental quantities are only functions of the u parameter, the equation simplifies to

$$\frac{F + Gv'}{\sqrt{E(u, v) + 2F(u, v)v' + G(u, v)v'^2}} = c_1.$$

A further simplification is based on the practical case when the u and v parametric lines defining the surface are orthogonal. In this case

$$F = 0,$$

and the equation may easily be integrated as

$$v = c_1 \int \frac{\sqrt{E}}{\sqrt{G^2 - c_1^2 G}} du + c_2.$$

The integration constants may be resolved from the boundary conditions.

When the function in the integral only contains the v function explicitly, and the $F = 0$ assumption still holds, then the equation becomes

$$\frac{Gv'^2}{\sqrt{E + Gv'^2}} - \sqrt{E + Gv'^2} = c_1.$$

Reordering and another integration brings

$$u = c_1 \int \frac{\sqrt{G}}{\sqrt{E^2 - c_1^2 E}} dv + c_2.$$

7.1.1 Geodesics of a sphere

For an enlightening example, we consider a sphere, given by

$$x(u, v) = R \sin(v) \cos(u),$$

$$y(u, v) = R \sin(v) \sin(u),$$

and

$$z(u, v) = R \cos(v).$$

The first fundamental quantities encapsulating the surface information are

$$E = R^2 \sin^2(v),$$

$$F = 0,$$

and

$$G = R^2.$$

Since this is the special case consisting of only v , the equation of the geodesic curve becomes

$$u = c_1 \int \frac{R}{\sqrt{R^4 \sin^4(v) - c_1^2 R^2 \sin^2(v)}} dv + c_2.$$

After the integration by substitution and some algebraic manipulations, we get

$$u = -a \sin \frac{\cot(v)}{\sqrt{\left(\frac{R}{c_1}\right)^2 - 1}} + c_2.$$

It follows that

$$\sin(c_2)(R \sin(v) \cos(u)) - \cos(c_2)(R \sin(v) \sin(u)) - \frac{R \cos(v)}{\sqrt{\left(\frac{R}{c_1}\right)^2 - 1}} = 0.$$

Substituting the surface definition of the sphere yields

$$x \sin(c_2) + y \cos(c_2) - \frac{z}{\sqrt{\left(\frac{R}{c_1}\right)^2 - 1}} = 0$$

and that represents an intersection of the sphere with a plane. By substituting boundary conditions, it would be easy to see that the actual plane contains the origin and defines the great circle going through the two given points. This fact is manifested in everyday practice by the transoceanic airplane routes' well-known northern swing in the Northern Hemisphere.

7.2 A system of differential equations for geodesic curves

Let us now seek the geodesic curve in the parametric form of

$$u = u(t),$$

and

$$v = v(t).$$

The curve goes through two points

$$P_0 = (u(t_0), v(t_0)),$$

and

$$P_1 = (u(t_1), v(t_1)).$$

Then the following variational problem provides the solution:

$$I(u, v) = \int_{t_0}^{t_1} \sqrt{Eu'^2 + 2Fu'v' + Gv'^2} dt = \text{extremum}.$$

Here

$$u' = \frac{du}{dt}, v' = \frac{dv}{dt}.$$

The corresponding Euler-Lagrange system of differential equations is

$$\frac{E_u u'^2 + 2F_u u'v' + G_u v'^2}{\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}} - \frac{d}{dt} \frac{2(Eu' + Fv')}{\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}} = 0,$$

and

$$\frac{E_v u'^2 + 2F_v u'v' + G_v v'^2}{\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}} - \frac{d}{dt} \frac{2(Fu' + Gv')}{\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}} = 0.$$

In the equations the notation

$$E_u = \frac{\partial E}{\partial u}, F_u = \frac{\partial F}{\partial u}, G_u = \frac{\partial G}{\partial u}$$

was used.

A more convenient and practically useful formulation, without the explicit derivatives, based on [6] is

$$u'' + \Gamma_{11}^1 u'^2 + 2\Gamma_{12}^1 u'v' + \Gamma_{22}^1 v'^2 = 0,$$

and

$$v'' + \Gamma_{11}^2 u'^2 + 2\Gamma_{12}^2 u'v' + \Gamma_{22}^2 v'^2 = 0.$$

The Γ are the Christoffel symbols that are defined from the following system of equations

$$\Gamma_{11}^1 E + \Gamma_{11}^2 F = (\mathbf{r}_{uu}, \mathbf{r}_u),$$

$$\Gamma_{11}^1 F + \Gamma_{11}^2 G = (\mathbf{r}_{uu}, \mathbf{r}_v),$$

$$\Gamma_{12}^1 E + \Gamma_{12}^2 F = (\mathbf{r}_{uv}, \mathbf{r}_u),$$

$$\Gamma_{12}^1 F + \Gamma_{12}^2 G = (\mathbf{r}_{uv}, \mathbf{r}_v),$$

$$\Gamma_{22}^1 E + \Gamma_{22}^2 F = (\mathbf{r}_{vv}, \mathbf{r}_u),$$

and

$$\Gamma_{22}^1 F + \Gamma_{22}^2 G = (\mathbf{r}_{vv}, \mathbf{r}_v).$$

The right-hand sides are inner products of the various derivatives

$$r_u = \frac{\partial}{\partial u} \mathbf{r},$$

$$r_{uu} = \frac{\partial^2}{\partial u^2} \mathcal{L},$$

$$r_{uv} = \frac{\partial^2}{\partial u \partial v} \mathcal{L},$$

and so forth. This system constitutes three pairs of equations with two unknowns each and they all may be solved when

$$EG - F^2 \neq 0$$

which is true when a parameterization is regular.

7.2.1 Geodesics of surfaces of revolution

Another practically important special case is represented by surfaces of revolution. Their generic description may be of the form

$$x = u \cos(v),$$

$$y = u \sin(v),$$

and

$$z = f(u).$$

Here the last equation describes the meridian curve generating the surface. The first order fundamental terms are

$$E = 1 + f'^2(u),$$

$$F = 0,$$

and

$$G = u^2.$$

The solution following the discussion in [Section 7.1](#) becomes

$$v = c_1 \int \frac{\sqrt{1 + f'^2(u)}}{u \sqrt{u^2 - c_1^2}} du + c_2.$$

A simple subcase of this class is a unit cylinder, described as

$$x = \cos(v),$$

$$y = \sin(v),$$

and

$$z = u.$$

The geometric meaning of the v parameter is the rotation angle generating the cylinder's circumference and the u parameter is the axial direction of the

surface. The fundamental terms are

$$E = 1,$$

$$F = 0,$$

and

$$G = 1.$$

The equation of the geodesic curve on the cylinder following above is

$$v = c_1 \int \frac{1}{1\sqrt{1-c_1^2}} du + c_2,$$

or

$$v = c_1 \frac{1}{\sqrt{1-c_1^2}} \int du + c_2.$$

With

$$c_3 = c_1 \frac{1}{\sqrt{1-c_1^2}}$$

and integration we obtain

$$v = c_3 u + c_2.$$

In the general case, this is a helix on the surface of the cylinder going through the two points. This is also a line in the parametric space. This fact is geometrically easy to explain because the cylinder is a developable surface. Such a surface may be rectified onto a plane. In such a case the geodesic curve is a straight line on the rectifying plane. The only curvature of the helix will be that of the cylinder.

The constants of integration may be determined from the boundary conditions. For example, assume the case shown in [Figure 7.1](#), where the starting point is at

$$P_0 = (x_0, y_0, z_0) = (1, 0, 0)$$

corresponding to parametric coordinates

$$u(t_0) = 0, v(t_0) = 0.$$

The endpoint is located at

$$P_1 = (x_1, y_1, z_1) = (0, 1, 1)$$

corresponding to parametric coordinates

$$u(t_1) = 1, v(t_1) = \pi/2.$$

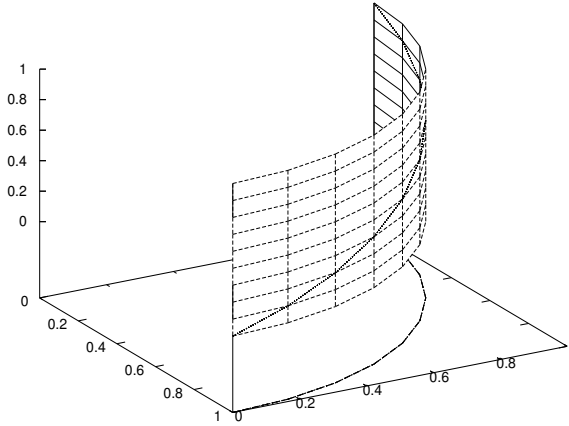


FIGURE 7.1 Geodesic curve of a cylinder

Substituting the starting point yields

$$0 = c_3 \cdot 0 + c_2,$$

which results in

$$c_2 = 0.$$

The endpoint substitution produces

$$\pi/2 = c_3 \cdot 1 + c_2,$$

and in turn

$$c_3 = \frac{\pi}{2}.$$

The specific solution for this case in the parametric space is

$$v = \frac{\pi}{2}u.$$

The Cartesian solution is obtained in the form of

$$x = \cos(v),$$

$$y = \sin(v),$$

and

$$z = \frac{v}{\pi/2}.$$

It is easy to see that the latter equation makes the elevation change from zero to 1, in accordance with the turning of the helix.

Since the parametric space of the cylinder is simply the rectangle of the developed surface, it is easy to see some special subcases. If the two points are located at the same rotational position ($v=\text{constant}$), but at different heights, the geodesic curve is a straight line. If the two points are on the same height ($u=\text{constant}$), but at different rotational angles, the geodesic curve is a circular arc.

The last two sections demonstrated the difficulties of finding the geodesic curves even on regular surfaces like the sphere or the cylinder. On a general three-dimensional surface these difficulties increase significantly and may render using the differential equation of the geodesic curve unfeasible.

7.3 Geodesic curvature

Let us consider the parametric curve

$$\underline{r}(t) = x(t)\underline{i} + y(t)\underline{j} + z(t)\underline{k}$$

on the surface

$$S(u, v) = x(u, v)\underline{i} + y(u, v)\underline{j} + z(u, v)\underline{k}.$$

Let \underline{n} denote the unit normal of the surface. The curvature vector of a three-dimensional curve is defined as

$$\underline{k} = \frac{d\underline{t}}{dt} = \underline{t}',$$

where \underline{t} is the tangent vector computed as

$$\underline{t} = \frac{d\underline{r}}{dt},$$

and also assumed to be a unit vector (a unit speed curve) for the simplicity of the derivation. Then the unit bi-normal vector is

$$\underline{b} = \underline{n} \times \underline{t}.$$

We represent the curvature vector with components along the bi-normal vector and the normal vector \underline{n} at any point as

$$\underline{k} = \kappa_n \underline{n} + \kappa_g \underline{b}.$$

The coefficients are the normal curvature and the geodesic curvature, respectively. Taking the inner product of the last equation with the \underline{b} vector and exploiting the perpendicularity conditions present, we obtain

$$\underline{b} \cdot \underline{k} = \kappa_g.$$

Substituting the definition of the bi-normal and the curvature vector results in

$$\kappa_g = (\underline{n} \times \underline{t}) \cdot \underline{t}'.$$

For the more generic case when the tangent and normal vectors are not of unit length, the geodesic curvature of a curve is defined as

$$\kappa_g = \frac{\underline{r}''(t) \cdot (\underline{n} \times \underline{r}'(t))}{\|\underline{r}'(t)\|^3}.$$

A curve on a surface is called geodesic if at each point of the curve its principal normal and the surface normal are collinear. Therefore:

A curve $\underline{r}(t)$ on the surface $S(u, v)$ is geodesic if the geodesic curvature of the curve is zero.

In order to prove that, the terms are computed from the surface information, such as

$$\underline{r}' = \underline{t} = \frac{\partial f}{\partial u} u' + \frac{\partial f}{\partial v} v' = f_u u' + f_v v'.$$

The application of the chain rule results in

$$\underline{r}'' = \underline{t}' = f_{uu}(u')^2 + 2f_{uv}u'v' + f_{vv}(v')^2 + f_u u'' + f_v v''.$$

After substitution into the equation of the geodesic curvature and some algebraic work, while employing the Christoffel symbols, [6] produces the form

$$\begin{aligned} \kappa_g = & \sqrt{EG - F^2}(\Gamma_{11}^2(u')^3 + (2\Gamma_{12}^1 - \Gamma_{11}^1)(u')^2v' + \\ & (\Gamma_{22}^2 - 2\Gamma_{12}^1)u'(v')^2 - \Gamma_{22}^1(v')^3 + u'v'' - u''v'). \end{aligned}$$

Since

$$EG - F^2 \neq 0,$$

the term in the parenthesis must be zero for zero geodesic curvature. That happens when the following terms vanish

$$u'' + \Gamma_{11}^1 u'^2 + 2\Gamma_{12}^1 u'v' + \Gamma_{22}^1 v'^2 = 0,$$

and

$$v'' + \Gamma_{11}^2 u'^2 + 2\Gamma_{12}^2 u'v' + \Gamma_{22}^2 v'^2 = 0.$$

This result is the decoupled system of equations of the geodesic, introduced in [Section 7.1](#), hence the vanishing of the geodesic curvature is indeed a characteristic of a geodesic curve.

Finally, since the recent discussions were mainly on parametric forms, the equation of the geodesic for an explicitly given surface

$$z = z(x, y(x))$$

is quoted from [6] for completeness' sake without derivation:

$$\begin{aligned} (1 + (\frac{\partial z}{\partial x})^2 + (\frac{\partial z}{\partial y})^2) \frac{d^2 y}{dx^2} &= \frac{\partial z}{\partial x} \frac{\partial^2 z}{\partial y^2} (\frac{dy}{dx})^3 + \\ & (2 \frac{\partial z}{\partial x} \frac{\partial^2 z}{\partial x \partial y} - \frac{\partial z}{\partial y} \frac{\partial^2 z}{\partial y^2}) (\frac{dy}{dx})^2 + \\ & (\frac{\partial z}{\partial x} \frac{\partial^2 z}{\partial x^2} - 2 \frac{\partial z}{\partial y} \frac{\partial^2 z}{\partial x \partial y}) \frac{dy}{dx} - \frac{\partial z}{\partial y} \frac{\partial^2 z}{\partial x^2}. \end{aligned}$$

The formula is rather overwhelming and useful only in connection with the simplest surfaces.

7.3.1 Geodesic curvature of helix

Let us enlighten this further by reconsidering the case of the geodesic curve of the cylinder discussed in the last section. The geodesic curve we obtained was the helix:

$$r(t) = \cos(t)\underline{i} + \sin(t)\underline{j} + \frac{t}{\pi/2}\underline{k}.$$

The appropriate derivatives are

$$r'(t) = -\sin(t)\underline{i} + \cos(t)\underline{j} + \frac{1}{\pi/2}\underline{k}$$

and

$$r''(t) = -\cos(t)\underline{i} - \sin(t)\underline{j} + 0\underline{k}.$$

The surface normal is computed as

$$\underline{n} = \frac{\partial S}{\partial u} \times \frac{\partial S}{\partial v}.$$

In the specific case of the cylinder

$$S(u, v) = \cos(v)\underline{i} + \sin(v)\underline{j} + u\underline{k},$$

it becomes

$$\underline{n} = \cos(t)\underline{i} + \sin(t)\underline{j} + 0\underline{k}.$$

The cross-product and substitution of $v = t$ yields

$$\underline{n}(t) \times r'(t) = \frac{2}{\pi} \sin(t)\underline{i} - \frac{2}{\pi} \cos(t)\underline{j} + (\sin^2(t) + \cos^2(t))\underline{k}.$$

The numerator of the curvature becomes zero, as

$$(-\cos(t)\underline{i} - \sin(t)\underline{j} + 0\underline{k}) \cdot \left(\frac{2}{\pi} \sin(t)\underline{i} - \frac{2}{\pi} \cos(t)\underline{j} + (\sin^2(t) + \cos^2(t))\underline{k} \right) = 0.$$

Since the denominator

$$\|r'(t)\|^3 = \left(\sqrt{1 + \left(\frac{1}{\pi/2} \right)^2} \right)^3$$

is not zero, the geodesic curvature becomes zero. Hence the helix is truly the geodesic curve of the cylinder.

The concept of geodesic curves may be generalized to higher-dimensional spaces. The geodesic curve notation, however, while justified on a three-dimensional surface, needs to be generalized as well. In such cases one talks about finding a geodesic object, or just a geodesic, as opposed to a curve on a surface.

7.4 Generalization of the geodesic concept

Insofar we confined the geodesic problem to finding a curve on a three-dimensional surface, but the concept may be generalized to higher dimensions. Physicists use the space-time continuum as a four-dimensional (Minkowski) space and find geodesic paths in that space. The arc length definition of this space is

$$ds^2 = dx^2 + dy^2 + dz^2 - cdt^2,$$

where t is the time dimension and c is the speed of light. The variational problem of minimal arclength may be posed similarly as in [Section 7.1](#) and may be solved with similar techniques. Einstein used this generalization to explain planetary motion as a geodesic phenomenon in the four-dimensional space.

Another important generalization is in the area of dynamical analysis of mechanical systems, a topic of [Chapter 10](#). The motion of the system will be described in terms of n generalized coordinates q_i and their time derivatives \dot{q}_i , the generalized velocities. These generalized coordinates, along with their constraint relationships, define an n -dimensional space, sometimes called the mechanical space.

The system will strive to reach an energy minimal equilibrium described by a variational problem in the form of

$$I = \int_{t_0}^{t_1} f(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) dt = \textit{extremum}.$$

The solution of this is mathematically equivalent to finding the geodesic path in the n -dimensional mechanical space.

To demonstrate the meaning of a mechanical space, let us consider a particle that is constrained to move on a two-dimensional surface, say a cylinder, enforced by some constraint. Then the particle's move in this two-dimensional mechanical space (the surface) is along the geodesic curve of the surface. This abstraction to higher dimensions is sometimes called the geometrization of mechanics.

Hamilton spearheaded this generalization and this will lead to Lagrange's equations of motion. The approach is of utmost importance in computational mechanics, the subjects of extensive discussion in [Chapter 10](#).

Computational geometry

The geodesic concept, introduced in the last chapter purely on variational principles, has interesting engineering aspects. On the other hand, the analytic solution of a geodesic curve by finding the extremum of a variational problem may not be easy in practical cases.

It is reasonable to assume, however, that the quality of a curve in a geodesic sense is related to its curvature. This observation proposes a strategy for creating good quality (albeit not necessarily geodesic) curves by minimizing the curvature.

Since the curvature is difficult to compute, one can use the second derivative of the curve in lieu of the curvature. This results in the following variational problem statement for a smooth curve: Find the curve $s(t)$ that results in

$$I(s) = \int_{t_0}^{t_n} k(s'')^2 dt = \text{extremum}.$$

The constant k represents the fact that this is an approximation of the curvature, but will be left out from our work below. This variational problem leads to the well-known spline functions.

8.1 Natural splines

Let us consider the following variational problem. Find the curve between two points P_0, P_3 such that

$$I(s) = \int_{t_0}^{t_3} \left(\frac{d^2 s}{dt^2}\right)^2 dt = \text{extremum},$$

with boundary conditions of

$$P_0 = s(t_0), P_3 = s(t_3),$$

and additional discrete internal constraints of

$$P_1 = s(t_1), P_2 = s(t_2).$$

In essence, we are constraining two interior points of the curve, along with the fixed beginning and endpoints. We will, for the sake of simplicity, assume unit equidistant parameter values as

$$t_i = i, i = 0, \dots, 3.$$

While the functional does not contain the independent variable t and the dependent variable $s(t)$, it is of higher order. Hence the Euler-Poisson equation of second order applies:

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y''} = 0,$$

and in this case it simplifies to

$$\frac{d^4}{dt^4} s(t) = 0.$$

Straightforward integration yields the solution of the form

$$s(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3,$$

where c_i are constants of integration to be satisfied by the boundary conditions. Imposing the boundary conditions and constraints yields the system of equations

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix}.$$

The inversion of the system matrix results in the generating matrix

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -11/6 & 3 & -3/2 & 1/3 \\ 1 & -5/2 & 2 & -1/2 \\ -1/6 & 1/2 & -1/2 & 1/6 \end{bmatrix}$$

for the natural spline. For any given set of four points

$$P = \begin{bmatrix} x_0 & y_0 & z_0 \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}$$

the coefficients of the solution may be obtained by

$$C = MP,$$

with distinct coefficients for all coordinate directions as

$$C = \begin{bmatrix} c_0^x & c_0^y & c_0^z \\ c_1^x & c_1^y & c_1^z \\ c_2^x & c_2^y & c_2^z \\ c_3^x & c_3^y & c_3^z \end{bmatrix}.$$

For example, the points

$$P = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 2 \\ 4 & 3 \end{bmatrix}$$

result in coefficients

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 13/6 \\ 0 & -3/2 \\ 0 & 1/3 \end{bmatrix}.$$

The parametric solution curve is of the form

$$\begin{aligned} x(t) &= 1 + t, \\ y(t) &= 1 + 13/6t - 3/2t^2 + 1/3t^3. \end{aligned}$$

The example solution curve is shown in [Figure 8.1](#), where the input points are connected by the straight lines that represent the chords of the spline. The spline demonstrates a good smoothness while satisfying the constraints.

Several extensions of this problem are noteworthy. It is possible to pose the variational problem in two-parameter form as

$$I(s) = \int \int_D \left(\left(\frac{\partial}{\partial u} s(u, v) \right)^2 + \left(\frac{\partial}{\partial v} s(u, v) \right)^2 \right) dudv = \text{extremum}.$$

The Euler-Lagrange equation corresponding to this problem arrives at Laplace's equation:

$$\frac{\partial^2}{\partial u^2} s(u, v) + \frac{\partial^2}{\partial v^2} s(u, v) = 0.$$

This is sometimes called the harmonic equation, hence the splines so obtained are also called harmonic splines.

It is also often the case in practice that many more than four points are given:

$$P_i = (x_i, y_i, z_i), i = 0, \dots, n.$$

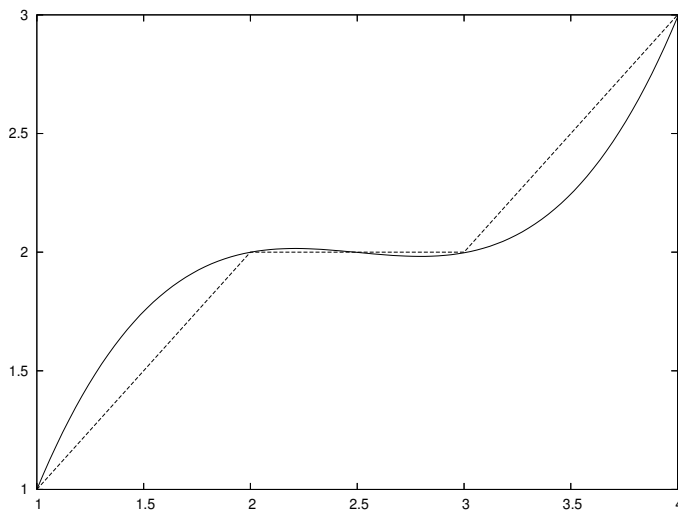


FIGURE 8.1 Natural spline approximation

This enables the generation of a multitude of natural spline segments, and a curvature continuity condition between the segments may also be enforced. Finally, the direct (for example Ritz) solution of the above variational problem leads to the widely used B-splines, a topic of the next chapter.

8.2 B-spline approximation

As shown in [Chapter 6](#), when using direct methods an approximate solution is sought in terms of suitable basis functions:

$$\bar{s}(t) = \sum_{i=0}^n B_{i,k}(t)Q_i,$$

where Q_i are the yet unknown control points ($i=0, \dots, n$) and $B_{i,k}$ are the basis functions of degree k in the parameter t . For industrial computational work, the class of basis functions resulting in the so-called B-splines are defined in [1] as

$$B_{i,0}(t) = \begin{cases} 1, & t_i \leq t < t_{i+1} \\ 0, & t < t_i, t \geq t_{i+1} \end{cases}$$

The higher order terms are recursively computed:

$$B_{i,k}(t) = \frac{t - t_i}{t_{i+k} - t_i} B_{i,k-1}(t) + \frac{t_{i+k+1} - t}{t_{i+k+1} - t_{i+1}} B_{i+1,k-1}(t).$$

The basis functions are computed from specific parameter values, called knot values. If their distribution is not equidistant, the splines are called non-uniform B-splines. If they are uniformly distributed, they are generating uniform B-splines.

The knot values are a subset of the parameter space and their selection enables a unique control of the behavior of the spline. For example, the use of duplicate knot values inside the parameters span of the spline results in a local change. The use of multiple knot values at the boundaries enforces various end conditions, such as the frequently used clamped end condition. This control mechanism contributes to the popularity of the method in computer-aided design (CAD), but will not be further explored here.

Smoothing a B-spline is defined by the variational problem

$$I_s(Q) = \int_{t_0}^{t_n} \left(\sum_{i=0}^n B''_{i,k}(t) Q_i \right)^2 dt = \text{extremum}.$$

The derivative of the basis functions may be recursively computed. For $k = 1$, since $B_{i,0}$ are constant

$$\frac{d}{dt} B_{i,1}(t) = B'_{i,1}(t) = \frac{1}{t_{i+1} - t_i} B_{i,0}(t) - \frac{1}{t_{i+2} - t_{i+1}} B_{i+1,0}(t).$$

For $k = 2$

$$\begin{aligned} \frac{d}{dt} B_{i,2}(t) = B'_{i,2}(t) &= \frac{1}{t_{i+2} - t_i} B_{i,1}(t) + \frac{t - t_i}{t_{i+2} - t_i} B'_{i,1}(t) - \\ &\quad \frac{1}{t_{i+3} - t_i} B_{i+1,1}(t) + \frac{t_{i+3} - t}{t_{i+3} - t_{i+1}} B'_{i+1,1}(t). \end{aligned}$$

Substituting the $k = 1$ derivative into the second term results in

$$\begin{aligned} \frac{t - t_i}{t_{i+2} - t_i} B'_{i,1}(t) &= \frac{t - t_i}{t_{i+2} - t_i} \left(\frac{1}{t_{i+1} - t_i} B_{i,0}(t) - \frac{1}{t_{i+2} - t_{i+1}} B_{i+1,0}(t) \right) = \\ &\quad \frac{1}{t_{i+2} - t_i} \left(\frac{t - t_i}{t_{i+1} - t_i} B_{i,0}(t) + \frac{t_i - t}{t_{i+2} - t_{i+1}} B_{i+1,0}(t) \right). \end{aligned}$$

The content of the parenthesis is easily recognizable as $B_{i,1}(t)$, hence this term is identical to the first. Similar arithmetic on the second two terms results in

the derivative for $k = 2$ as

$$\frac{d}{dt}B_{i,2}(t) = B'_{i,2}(t) = \frac{2}{t_{i+2} - t_i}B_{i,1} - \frac{2}{t_{i+3} - t_{i+1}}B_{i+1,1}.$$

By induction, for any k value the first derivative is as follows:

$$\frac{d}{dt}B_{i,k}(t) = B'_{i,k}(t) = \frac{k}{t_{i+k} - t_i}B_{i,k-1}(t) - \frac{k}{t_{i+k+1} - t_{i+1}}B_{i+1,k-1}(t).$$

A repeated application of the same step will produce the needed second derivative B'' as

$$\frac{d}{dt}B'_{i,k}(t) = B''_{i,k}(t) = \frac{k}{t_{i+k} - t_i}B'_{i,k-1}(t) - \frac{k}{t_{i+k+1} - t_{i+1}}B'_{i+1,k-1}(t).$$

The spline, besides being smooth (minimal in curvature), is expected to approximate a given set of points $P_j; j = 0, \dots, m$, with associated prescribed parameter values (not necessarily identical to the knot values) of $t_j; j = 0, \dots, m$. If such parameter values are not given, the parameterization may be via the simple method of uniform spacing defined as $t_j = j; 0 \leq j \leq m$. Assuming that the point set defined is geometrically semi-equidistant this is proven in industry to be a good method for the problem at hand. If that is not the case, a parameterization based on the chord length may also be used.

Approximating the given points with the spline is another variational problem that requires finding a minimum of the squares of the distances between the spline and the points.

$$I_a(\bar{s}) = \sum_{j=0}^m (\bar{s}(t_j) - P_j)^2 = \text{extremum}.$$

Substituting the B-spline formulation and the basis functions results in

$$I_a(Q) = \sum_{j=0}^m \left(\sum_{i=0}^n B_{i,k}(t_j) Q_i - P_j \right)^2.$$

Similarly, in the smoothing variational problem we also replace the integral with a sum over the given points in the parameter span, resulting in

$$I_s(Q) = \sum_{j=0}^m \left(\sum_{i=0}^n B''_{i,k}(t_j) Q_i \right)^2.$$

Finally, the functional to produce a smooth spline approximation is the sum of the two functionals

$$I(Q) = I_a(Q) + I_s(Q).$$

The notation is to demonstrate the dependence upon the yet unknown control points of the spline.

According to the direct method developed in [Chapter 6](#), the problem has an extremum when

$$\frac{\partial I}{\partial Q_i} = 0,$$

for each $i = 0, \dots, n$. The control points will be, of course, described by Cartesian coordinates; hence, each of the above equations represents three scalar equations.

The derivative of the approximating component with respect to an unknown control point Q_p yields

$$2 \sum_{j=0}^m B_{p,k}(t_j) \left(\sum_{i=0}^n B_{i,k}(t_j) Q_i - P_j \right),$$

where $p = 0, 1, \dots, n$. This is expressed in matrix form as

$$B^T B Q - B^T P$$

with the matrices

$$B = \begin{bmatrix} B_{0,k}(t_0) & B_{1,k}(t_0) & B_{2,k}(t_0) & \dots & B_{n,k}(t_0) \\ B_{0,k}(t_1) & B_{1,k}(t_1) & B_{2,k}(t_1) & \dots & B_{n,k}(t_1) \\ \dots & \dots & \dots & \dots & \dots \\ B_{0,k}(t_m) & B_{1,k}(t_m) & B_{2,k}(t_m) & \dots & B_{n,k}(t_m) \end{bmatrix},$$

$$P = \begin{bmatrix} P_0 \\ P_1 \\ \dots \\ P_m \end{bmatrix},$$

and

$$Q = \begin{bmatrix} Q_0 \\ Q_1 \\ \dots \\ Q_n \end{bmatrix}.$$

For degree $k = 3$ the basis functions may be analytically computed as:

$$B_{0,3} = \frac{1}{6}(1-t)^3,$$

$$B_{1,3} = \frac{1}{6}(3t^3 - 6t^2 + 4),$$

$$B_{2,3} = \frac{1}{6}(-3t^3 + 3t^2 + 3t + 1),$$

and

$$B_{3,3} = \frac{1}{6}t^3.$$

Furthermore, for uniform parameterization the B matrix is easily computed by hand as

$$B = \frac{1}{6} \begin{bmatrix} 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ -1 & 4 & -5 & 8 \\ -8 & 31 & -44 & 27 \\ -27 & 100 & -131 & 64 \end{bmatrix}.$$

The derivative of the smoothing component of the functional, with respect to an unknown control point Q_p yields

$$2 \sum_{j=1}^m B''_{p,k}(t_j) \sum_{i=0}^n B''_{i,k}(t_j) Q_i,$$

where $p \in [0, \dots, n]$. This results in a smoothing matrix

$$D = \begin{bmatrix} B''_{0,k}(t_0) & B''_{1,k}(t_0) & \dots & B''_{n,k}(t_0) \\ B''_{0,k}(t_1) & B''_{1,k}(t_1) & \dots & B''_{n,k}(t_1) \\ \dots & \dots & \dots & \dots \\ B''_{0,k}(t_m) & B''_{1,k}(t_m) & \dots & B''_{n,k}(t_m) \end{bmatrix}.$$

These second derivatives for the cubic case are

$$B''_{0,3} = 1 - t,$$

$$B''_{1,3} = 3t - 2,$$

$$B''_{2,3} = -3t + 1,$$

and

$$B''_{3,3} = t.$$

For uniform parameterization, the smoothing matrix is computed as

$$D = \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ -1 & 4 & -5 & 2 \\ -2 & 7 & -8 & 3 \\ -3 & 10 & -11 & 4 \end{bmatrix}.$$

The simultaneous solution of both the smoothing and approximating problem is now represented by the matrix equation

$$AQ = B^T P$$

where the system matrix is

$$A = B^T B + D^T D.$$

The solution of this system produces the control points for a smooth approximation.

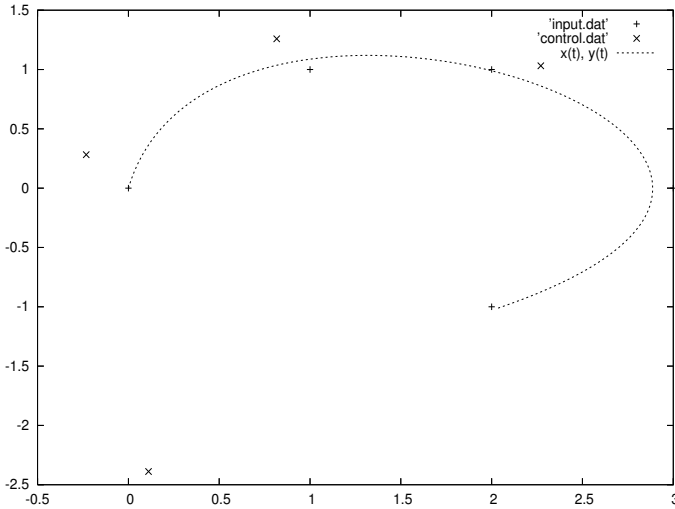


FIGURE 8.2 Smooth B-spline approximation

Figure 8.2 shows the smooth spline approximation for a set of given points. The input points as well as the computed control points are also shown. Note that, as opposed to the natural spline, the curve does not go through the points exactly, but it is very smooth.

8.3 B-splines with point constraints

It is possible to require some of the points to be exactly satisfied. For the generic case of multiple enforced points, a constrained variational problem is

formed.

$$I(Q_i) = \text{extremum},$$

subject to

$$\bar{s}(t_l) = R_l; l = 0, 1, \dots, o.$$

Here the enforced points R_l represent a subset of the given points (P_j), while the remainder are to be approximated. The subset is specified as

$$R_l = MP_j; l = 0, 1, \dots, o; j = 0, 1, \dots, m, o < m,$$

where the mapping matrix M has o rows and m columns and contains a single nonzero term in each row, in the column corresponding to a selected interpolation point. For example, the matrix

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

would specify the two internal points of four P_j points, i.e.,

$$R_0 = P_1$$

and

$$R_1 = P_2.$$

This approach could be used to specify any pattern, such as every second or third term, or some specific points at intermittent locations.

Introducing the specifics of the splines and Lagrange multipliers, the constrained variational problem is presented as

$$I(Q_i, \lambda_l) = I(Q_i) + \sum_{l=0}^o \lambda_l \left(\sum_{i=0}^n B_{i,k}(t_l) Q_i - R_l \right).$$

The derivatives of $I(Q_i)$ with respect to the Q_p control point were computed earlier, but need to be extended with the derivative of the term containing the Lagrange multiplier:

$$\sum_{l=0}^o B_{p,k}(t_l) \lambda_l \sum_{i=0}^n B_{i,k}(t_l).$$

Utilizing the earlier introduced matrices, this term is

$$B^T M^T \Lambda,$$

where Λ is a column vector of $o + 1$ Lagrange multipliers,

$$\begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \dots \\ \lambda_o \end{bmatrix}.$$

The derivatives with respect to the Lagrange multipliers λ_l produce

$$\left(\sum_{i=0}^n B_{i,k}(t_l) Q_i - R_l \right) = 0; l = 0, 1, \dots, o.$$

This results in $o + 1$ new equations of the form

$$\sum_{i=0}^n B_{i,k}(t_l) Q_i = R_l,$$

or in matrix form, using the earlier matrices:

$$MBQ = R,$$

where R is a vector of the interpolated points. The two sets of equations may be assembled into a single matrix equation with $n + 1 + o + 1$ rows and columns of the form

$$\begin{bmatrix} A & B^T M^T \\ MB & 0 \end{bmatrix} \begin{bmatrix} Q \\ \Lambda \end{bmatrix} = \begin{bmatrix} B^T P \\ MP \end{bmatrix}.$$

The first matrix row represents the constrained functional and the second row represents the constraints. The simultaneous solution of this (symmetric, indefinite, but still linear) system produces the optimized (approximated and smoothed) and selectively interpolated solution.

The solution of this problem is accomplished in the following steps. First the unknown control points are expressed from the first row of

$$AQ + B^T M^T \Lambda = B^T P$$

as functions of the unknown Lagrange multipliers. Substituting into the second equation is the way to compute the multipliers:

$$\Lambda = (MBA^{-1}B^T M^T)^{-1}(MBA^{-1}B^T P - MP).$$

Finally, the set of control points, which are solutions of the constrained variational problem, are obtained explicitly from the first equation as

$$Q = A^{-1}(B^T P - B^T M^T \Lambda).$$

The effect of point constraints is shown in [Figure 8.3](#) in connection with the earlier example, constraining the spline to go through the second and the

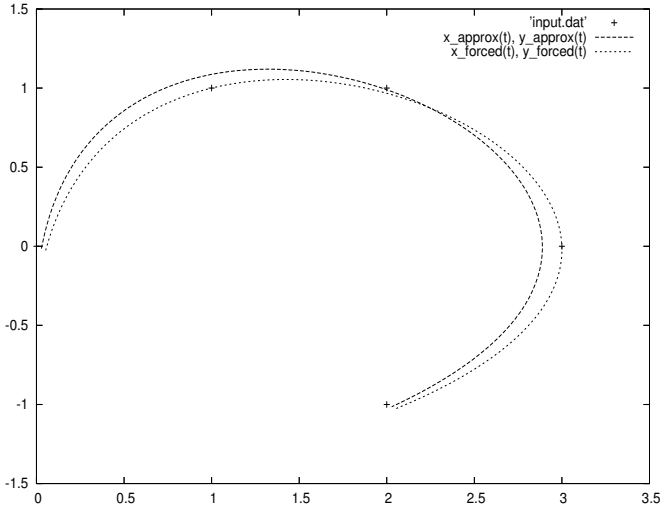


FIGURE 8.3 Point constrained B-spline

fourth points. The dashed curve is the original curve while the dotted curve is the new curve and it demonstrates the adherence to the constraint, at the same time maintaining the quality of the approximation and the smoothness.

8.4 B-splines with tangent constraints

It may be desirable for the engineer to be able to enforce constraints posed by specifying tangents at certain points. These are of the form

$$\vec{s}'(t_l) = T_l; l = 0, 1, \dots, o,$$

assuming for now that the tangents are given at the same points where interpolation constraints were also given. The constrained problem shown in the prior section will be extended with the additional constraints and Lagrange multipliers:

$$\sum_{l=0}^o \lambda_l \left(\sum_{i=0}^n B'_{i,k}(t_l) Q_i - T_l \right); l = 0, 1, \dots, o.$$

The derivative with respect to the new Lagrange multipliers is

$$\sum_{i=0}^n B'_{i,k}(t_l) Q_i - T_l = 0; l = 0, 1, \dots, o.$$

This results in $o + 1$ new equations of the form

$$\sum_{i=0}^n B'_{i,k}(t_l) Q_i = T_l,$$

or in matrix form, using some of the earlier matrices:

$$MCQ = T,$$

where T is a vector of the given tangents and the matrix of first derivatives is

$$C = \begin{bmatrix} B'_{0,k}(t_0) & B'_{1,k}(t_0) & \dots & B'_{n,k}(t_0) \\ B'_{0,k}(t_1) & B'_{1,k}(t_1) & \dots & B'_{n,k}(t_1) \\ \dots & \dots & \dots & \dots \\ B'_{0,k}(t_m) & B'_{1,k}(t_m) & \dots & B'_{n,k}(t_m) \end{bmatrix}.$$

The first derivatives of the basis functions for the cubic case are

$$B'_{0,3} = -\frac{1}{2}(1-t)^2,$$

$$B'_{1,3} = \frac{3}{2}t^2 - 2t,$$

$$B'_{2,3} = -\frac{3}{2}t^2 + t + \frac{1}{2},$$

and

$$B'_{3,3} = \frac{1}{2}t^2.$$

For the uniform case the C matrix containing the first derivatives is

$$C = \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 4 & -7 & 4 \\ -4 & 15 & -20 & 9 \\ -9 & 32 & -39 & 16 \end{bmatrix}.$$

The three sets of equations may be assembled into one matrix equation with $n + 1 + 2(o + 1)$ rows and columns of the form

$$\begin{bmatrix} A & B^T M^T & C^T M^T \\ MB & 0 & 0 \\ MC & 0 & 0 \end{bmatrix} \begin{bmatrix} Q \\ \Lambda_p \\ \Lambda_t \end{bmatrix} = \begin{bmatrix} B^T P \\ MP \\ MT \end{bmatrix}.$$

The index of the Lagrange multipliers refers to points (p) or tangents (t).

The restriction of giving tangents at all the same points where interpolation constraints are also given may be relaxed and the points with tangential constraints may only be a subset of the points where interpolation constraints are placed. In this case the final equation is

$$\begin{bmatrix} A & B^T M^T & C^T N^T \\ MB & 0 & 0 \\ NC & 0 & 0 \end{bmatrix} \begin{bmatrix} Q \\ \Lambda_p \\ \Lambda_t \end{bmatrix} = \begin{bmatrix} B^T P \\ MP \\ NT \end{bmatrix}.$$

Here the N mapping matrix is a subset of the M mapping matrix. The solution of this problem is similar to the solution of the simply constrained case, albeit a bit more tedious, due to the fact that the constraints are now of two different kinds. First the solution in terms of the multipliers is expressed

$$Q = A^{-1} B^T P - A^{-1} [B^T M^T \ C^T N^T] \begin{bmatrix} \Lambda_p \\ \Lambda_t \end{bmatrix}.$$

Then there is a matrix equation to compute the multipliers from

$$\begin{bmatrix} MB \\ NC \end{bmatrix} Q = \begin{bmatrix} MP \\ NT \end{bmatrix}.$$

The sets of multipliers are obtained as

$$\begin{bmatrix} \Lambda_p \\ \Lambda_t \end{bmatrix} = \left(\begin{bmatrix} MB \\ NC \end{bmatrix} A^{-1} [B^T M^T \ C^T N^T] \right)^{-1} \left(\begin{bmatrix} MB \\ NC \end{bmatrix} A^{-1} (B^T P) - \begin{bmatrix} MP \\ NT \end{bmatrix} \right).$$

Finally, the desired set of control points satisfying the constrained variational problem are

$$Q = A^{-1} (B^T P - B^T M^T \Lambda_p - C^T N^T \Lambda_t).$$

Let us use again the same set of points, but enforce the curve going through the second point with a tangent of 45 degrees.

The dotted curve in [Figure 8.4](#) demonstrates the satisfaction of both constraints, going through the second point and having a 45-degree tangent. It is also very clear that such a strong constraint imposed upon one point has a significant, and in this case detrimental, effect on the shape of the curve at least as far as the approximation is concerned. The smoothness of the curve is still impeccable.

In practical applications, some heuristics, like setting the tangent at a certain point parallel to the chord between the two neighboring points, can be used successfully. Then

$$T_i = \frac{\overline{P_{i+1} - P_i}}{\| \overline{P_{i+1} - P_i} \|}.$$

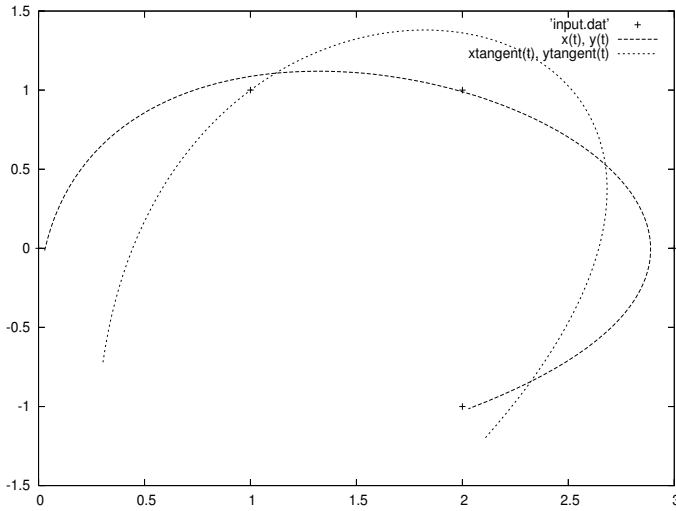


FIGURE 8.4 Tangent constrained B-spline

This would result in different control points and a much smaller deformation of the overall curve may be obtained. Systematic and possibly interactive application of this concept should result in good shape preservation and general smoothness.

8.5 Generalization to higher dimensions

The spline technology discussed above is easily generalized to higher-dimensional spaces. Let us consider surfaces given in the form of

$$z(x, y) = f(x, y)$$

first. A B-spline surface is defined by a set of control points as

$$\bar{s}(u, v) = \sum_{i=0}^n \sum_{j=0}^m B_{i,k}(u) B_{j,k}(v) Q_{ij},$$

where now we have two distinct knot value sequences of

$$u_i; i = 0, 1, \dots, n,$$

and

$$v_j; j = 0, 1, \dots, m.$$

The topological rectangularity of the control points is not necessary and may be overcome by coalescing certain points or adjustments of the knot points. The control points to provide a smooth approximation of the given geometric surface are selected from the variational problem of

$$I(\bar{s}) \int \int (f(x, y) - \bar{s}(x, y))^2 dx dy = \text{extremum}.$$

Substituting the surface spline definition and sampling of the given surface results in another, albeit more difficult, system of linear equations from which the control point locations may be resolved in a similar fashion as in the case of spline curves before.

Finally, it is also possible to describe some geometrical volumes with the B-spline technology. Consider the form

$$\bar{s}(u, v, t) = \sum_{i=0}^n \sum_{j=0}^m \sum_{l=0}^p B_{i,k}(u) B_{j,k}(v) B_{l,k}(t) Q_{ijl},$$

where now the third set of knot values

$$t_l; l = 0, 1, \dots, p$$

defines the direction through the volume starting from a surface. Topological rectangularity considerations again apply, but may be overcome with some inconvenience.

Both of these generalizations are important in computer-aided design (CAD) tools in the industry. They represent an efficient way to describe general (non-mathematical) surfaces and volumes.

Another, very important higher than three-dimensional extension occurs in computer-aided manufacturing (CAM). The calculation of smooth tool-paths of five-axis machines includes the three Cartesian coordinates and two additional quantities related to the tool position. This is important in machining of surface parts comprised of valleys and walls.

The positioning of the cutting tool is customarily described by two angles. The tool's "leaning" in the normal plane is one which may be construed as a rotation with respect to the bi-normal of the path curve. The tool's "swaying" from the normal plane, which constitutes a rotation around the path tangent

as an axis, may be another angle.

Abrupt changes in the tool axes are detrimental to the machined surface quality as well as to the operational efficiency. Hence it is a natural desire to smooth these quantities as well. The variational formulation for the geometric smoothing of the spline, shown above, accommodates any number of such additional considerations.

Analytic mechanics

Analytic mechanics is a mathematical science, but it is of high importance for engineers as it enables the analytic solution of some fundamental problems of engineering mechanics. At the same time it establishes generally applicable procedures. Mathematical physics texts, such as [4], laid the foundation for these analytic approaches addressing physical problems.

9.1 Hamilton's principle for mechanical systems

Hamilton's principle was briefly mentioned earlier in [Section 1.4.2](#) in connection with the problem of a particle moving under the influence of a gravity field. The principle, however, is much more general and it is applicable to complex mechanical systems. For conservative systems, Hamilton's principle states that the motion between two points is defined by the variational problem of

$$\int_{t_0}^{t_1} (E_k - E_p) dt = \text{extremum},$$

where E_k and E_p are the kinetic and potential energy, respectively. Introducing the so-called Lagrange function

$$L = E_k - E_p,$$

the principle may also be stated as

$$\int_{t_0}^{t_1} L dt = \text{extremum},$$

where the extremum is not always zero. The advantage of using Hamilton's principle is that it is stated in terms of energies, which are independent of the selection of coordinate systems. This will be further exploited and extended in the next chapter.

In the following sections we find analytic solutions for classical mechanical problems of elasticity utilizing Hamilton's principle. The most fitting applica-

tion is the excitation of an elastic system by displacing it from its equilibrium position. In this case the system will vibrate with a frequency characteristic to its geometry and material, while constantly exchanging kinetic and potential energy.

The case of non-conservative systems, where energy loss may occur due to dissipation of the energy, will not be discussed. Hamilton's principle may be extended to non-conservative systems, but the added difficulties do not enhance the discussion of the variational aspects, which is our main focus.

9.2 Elastic string vibrations

We now consider the vibrations of an elastic string. Let us assume that the equilibrium position of the string is along the x axis, and the endpoints are located at $x = 0$ and $x = L$. We will stretch the string (since it is elastic) by displacing it from its equilibrium with some

$$\Delta L$$

value, resulting in a certain force F exerted on both endpoints to hold it in place. We assume there is no damping and the string will vibrate indefinitely if displaced, i.e., the system is conservative.

The particle of the string located at the coordinate value x at the time t has a yet unknown displacement value of $y(x, t)$. The boundary conditions are:

$$y(0, t) = y(L, t) = 0,$$

in other words the string is clamped at the ends. In order to use Hamilton's principle, we need to compute the kinetic and potential energies.

With unit length mass of ρ , the kinetic energy is of the form

$$E_k = \frac{1}{2} \int_0^L \rho \left(\frac{\partial y}{\partial t} \right)^2 dx.$$

The potential energy is related to the elongation (stretching) of the string. The arc length of the elastic string is

$$\int_0^L \sqrt{1 + \left(\frac{\partial y}{\partial x} \right)^2} dx,$$

and the elongation due to the transversal motion is

$$\Delta L = \int_0^L \sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2} dx - L.$$

Assuming that the elongation is small, i.e.,

$$\left|\frac{\partial y}{\partial x}\right| < 1,$$

it is reasonable to approximate

$$\sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2} \approx 1 + \frac{1}{2}\left(\frac{\partial y}{\partial x}\right)^2.$$

The elongation by substitution becomes

$$\Delta L \approx \frac{1}{2} \int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx.$$

Hence the potential energy contained in the elongated string is

$$E_p = \frac{1}{2} F \Delta L = \frac{F}{2} \int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx.$$

We are now in the position to apply Hamilton's principle. The variational problem becomes

$$I(y) = \int_{t_0}^{t_1} (E_k - E_p) dt = \frac{1}{2} \int_{t_0}^{t_1} \int_0^L \left(\rho \left(\frac{\partial y}{\partial t}\right)^2 - F \left(\frac{\partial y}{\partial x}\right)^2 \right) dx dt = \text{extremum}.$$

The Euler-Lagrange differential equation for a function of two variables, derived in [Section 3.2](#), is applicable and results in

$$F \frac{\partial^2 y}{\partial x^2} = \rho \frac{\partial^2 y}{\partial t^2}. \quad (9.1)$$

This is the well-known differential equation of the elastic string, also known as the **wave equation**.

The solution of the problem may be solved by separation. We seek a solution in the form of

$$y(x, t) = a(t)b(x),$$

separating it into time and geometry dependent components. Then

$$\frac{\partial^2 y}{\partial x^2} = b''(x)a(t)$$

and

$$\frac{\partial^2 y}{\partial t^2} = a''(t)b(x),$$

where

$$b''(x) = \frac{\partial^2 b}{\partial x^2},$$

$$a''(t) = \frac{\partial^2 a}{\partial t^2}.$$

Substituting into equation (9.1) yields

$$\frac{b''(x)}{b(x)} = \frac{1}{f^2} \frac{a''(t)}{a(t)},$$

where for future convenience we introduced

$$f^2 = \frac{F}{\rho}.$$

The two sides of this differential equation are dependent on x and t , respectively. Their equality is required at any x and t values which implies that the two sides are constant. Let's denote the constant by $-\lambda$ and separate the (partial) differential equation into two ordinary differential equations:

$$\frac{\partial^2 b}{\partial x^2} + \lambda b(x) = 0,$$

and

$$\frac{\partial^2 a}{\partial t^2} + f^2 \lambda a(t) = 0.$$

The solution of these equations may be obtained by the techniques learned in [Chapter 5](#) for the eigenvalue problems. The first equation has the eigensolutions of the form

$$b_k(x) = \sin\left(\frac{k\pi}{L}x\right); k = 1, 2, \dots,$$

corresponding to the eigenvalues

$$\lambda_k = \frac{k^2\pi^2}{L^2}.$$

Applying these values we obtain the time dependent solution from the second equation by means of classical calculus in the form of

$$a_k(t) = c_k \cos\left(\frac{k\pi f}{L}t\right) + d_k \sin\left(\frac{k\pi f}{L}t\right),$$

with c_k, d_k arbitrary coefficients. Considering that at $t = 0$ the string is in a static equilibrium position

$$a'(t = 0) = 0$$

we obtain $d_k = 0$ and the solution of

$$a_k(t) = c_k \cos\left(\frac{k\pi f}{L}t\right).$$

The fundamental solutions of the problem become

$$y_k(x, t) = c_k \cos\left(\frac{k\pi f}{L}t\right) \sin\left(\frac{k\pi}{L}x\right); k = 1, 2, \dots$$

For any specific value of k the solution is a periodic function with period

$$\frac{2L}{kf}$$

and frequency

$$\frac{k\pi f}{L}.$$

The quantities

$$\lambda_k = \frac{k\pi}{L}$$

for a specific k value are the natural frequencies of the string. The corresponding fundamental solutions are natural vibration modes shapes, or the normal modes. The first three normal modes are shown in [Figure 9.1](#) for an elastic spring of unit tension force, mass density, and span. The figure demonstrates that the half-period decreases along with increasing mode number.

The motion is initiated by displacing the string and releasing it. Let us define this initial enforced amplitude as

$$y(x_m, 0) = y_m,$$

where the x_m describes the location of the initial stationary displacement of the string as an internal value of the span

$$x_m \in (0, L).$$

Then the initial shape of the string is a triangle over the span, described by the function

$$f(x) = \begin{cases} \frac{y_m}{x_m}x, & 0 \leq x \leq x_m, \\ y_m + \frac{y_m}{x_m-L}(x - x_m), & x_m \leq x \leq L. \end{cases}$$

The unknown coefficient may be solved from the initial condition as

$$y(x_m, 0) = f(x_m) = c_k \cos\left(\frac{k\pi f}{L}0\right) \sin\left(\frac{k\pi}{L}x_m\right) = y_m,$$

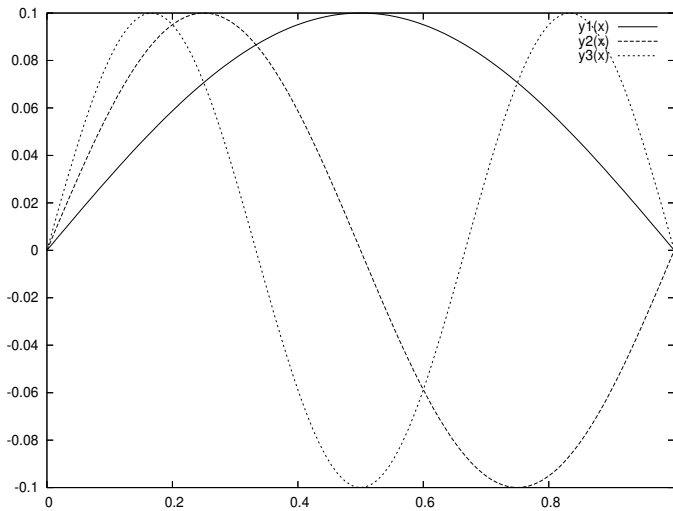


FIGURE 9.1 Normal modes of elastic string

from which

$$c_k = \frac{y_m}{\sin(\frac{k\pi}{L}x_m)}.$$

Note that if the interior point is the middle point of the span,

$$x_m = \frac{L}{2},$$

then the first coefficient will be simply the y_m amplitude:

$$c_1 = y_m,$$

since

$$\sin(\frac{\pi}{L} \frac{L}{2}) = \sin(\frac{\pi}{2}) = 1.$$

Similar, but not identical, considerations may be applied for the coefficients of the higher normal modes.

The natural frequencies depend on the physical conditions, such as the pre-applied tension force distribution and the material characteristics embodied in the unit weight ρ . Specifically, the higher the tension force F in the string,

the higher the frequency becomes. A very tight string vibrates very quickly (with high frequency), while a very loose string vibrates slowly.

9.3 The elastic membrane

We now turn our attention to the case of an elastic membrane. We assume that the membrane is fixed on its perimeter L which surrounds the domain D of the membrane. We further assume that the initial, equilibrium position of the membrane is coplanar with the $x - y$ plane.

$$z(x, y, t) = 0.$$

The membrane is displaced by a certain amount and released. The ensuing vibrations are the subject of our interest. The vibrations are a function of the location of the membrane and the time as

$$z = z(x, y, t).$$

We will again use Hamilton's principle after the kinetic and potential energy of the membrane are found. Let us assume that the unit area mass of the membrane does not change with time, and is not a function of the location:

$$\rho(x, y) = \rho = \text{constant}.$$

The velocity of the membrane point at (x, y) is

$$v = \frac{\partial z}{\partial t},$$

resulting in kinetic energy of

$$E_k = \frac{1}{2} \int \int_D \rho v^2 dx dy$$

or

$$E_k = \frac{1}{2} \int \int_D \rho \left(\frac{\partial z}{\partial t} \right)^2 dx dy.$$

We consider the source of the potential energy to be the stretching of the surface of the membrane. The initial surface is

$$\int \int_D dx dy,$$

and the extended surface is

$$\int \int_D \sqrt{1 + \left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2} dx dy.$$

Assuming small vibrations, we approximate as earlier in the case of the string

$$\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \approx 1 + \frac{1}{2} \left(\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 \right).$$

Hence the surface change is

$$\frac{1}{2} \iint_D \left(\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 \right) dx dy.$$

The stretching of the surface results in a surface tension σ per unit surface area. The potential energy is equal to σ multiplied by the surface change.

$$E_p = \sigma \frac{1}{2} \iint_D \left(\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 \right) dx dy.$$

We are now in the position to apply Hamilton's principle. Since

$$I(z) = \int_{t_0}^{t_1} (E_k - E_p) dt = \text{extremum},$$

substitution yields the variational problem of the elastic membrane:

$$I(z) = \frac{1}{2} \int_{t_0}^{t_1} \iint_D \rho \left(\frac{\partial z}{\partial t} \right)^2 - \sigma \left(\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 \right) dx dy dt = \text{extremum}.$$

The Euler-Lagrange differential equation for this class of problems following [Section 3.4](#) becomes

$$\sigma \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right) = \rho \frac{\partial^2 z}{\partial t^2},$$

or using Laplace's symbol

$$\sigma \Delta z = \rho \frac{\partial^2 z}{\partial t^2}.$$

The solution will follow the insight gained at the discussion of the elastic string and we seek a solution in the form of

$$z(x, y, t) = a(t)b(x, y).$$

The derivatives of this solution are

$$\Delta z(x, y, t) = a(t)\Delta b(x, y),$$

and

$$\frac{\partial^2 z(x, y, t)}{\partial t^2} = b(x, y) \frac{\partial^2 a(t)}{\partial t^2}.$$

Substitution yields

$$\frac{\sigma \Delta b}{\rho b} = \frac{\partial^2 a(t)}{\partial t^2} / a(t).$$

Again, since the left-hand side is only a function of the spatial coordinates, the right-hand side is of time, they must be equal and constant, assumed to be $-\lambda$. This separates the partial differential equation into two ordinary ones:

$$\frac{\partial^2 a(t)}{\partial t^2} + \lambda a(t) = 0,$$

and

$$\sigma \Delta b(x, y) + \lambda \rho b(x, y) = 0.$$

The solution of the first differential equation is

$$a(t) = c_1 \cos(\sqrt{\lambda}t) + c_2 \sin(\sqrt{\lambda}t).$$

Since initially the membrane is in equilibrium,

$$\left. \frac{da}{dt} \right|_{t=0} = 0$$

this indicates that

$$c_2 = 0.$$

Hence

$$a(t) = c_1 \cos(\sqrt{\lambda}t).$$

In order to demonstrate the solution for the second equation, let us omit the tension and material density for ease of discussion. The differential equation of the form

$$\Delta b(x, y) + \lambda b(x, y) = 0,$$

is the same we solved analytically in the case of the elastic string; however, it is now with a solution function of two variables. The solution strategy will consider the variational form of this eigenvalue problem introduced in [Section 5.2](#):

$$I(b) = \int \int_D \left(\left(\frac{\partial b}{\partial x} \right)^2 + \left(\frac{\partial b}{\partial y} \right)^2 - \lambda b^2(x, y) \right) dx dy = \text{extremum}.$$

We further restrict ourselves to the domain of the unit circle for simplicity. The domain D is defined by

$$D : (1 - x^2 - y^2 \geq 0).$$

We use Kantorovich's method and seek an approximate solution in the form of

$$b(x, y) = \alpha \omega(x, y) = \alpha (x^2 + y^2 - 1),$$

where α is a yet unknown constant. It follows that on the boundary ∂D

$$\omega(x, y) = x^2 + y^2 - 1 = 0,$$

hence the approximate solution satisfies the zero boundary condition. With this choice

$$I(\alpha) = \alpha^2 \int \int_D (4x^2 + 4y^2 - \lambda(x^2 + y^2 - 1)^2) dx dy = \text{extremum}.$$

Introducing polar coordinates for ease of integration yields

$$I(\alpha) = \alpha^2 \int_0^{2\pi} \int_0^1 4r^3 - \lambda r(r^2 - 1)^2 dr d\phi = \text{extremum}.$$

The evaluation of the integral results in the form

$$I(\alpha) = (2\pi - \lambda \frac{\pi}{3}) \alpha^2 = \text{extremum}.$$

The necessary condition of the extremum is

$$\frac{\partial I(\alpha)}{\partial \alpha} = 0,$$

which yields an equation for λ

$$2\alpha(2\pi - \lambda \frac{\pi}{3}) = 0.$$

The eigenvalue as the solution of this equation is

$$\lambda = 6.$$

The unknown solution function coefficient may be solved by normalizing the eigensolution as

$$\int \int_D b^2(x, y) dx = 1.$$

Substituting yields

$$\alpha^2 \int_0^{2\pi} \int_0^1 r(r^2 - 1)^2 dr d\phi = 1.$$

Integrating results in

$$\alpha^2 \frac{\pi}{3} = 1.$$

Hence

$$\alpha = \sqrt{\frac{3}{\pi}}.$$

The solution of the second equation is

$$b(x, y) = \sqrt{\frac{3}{\pi}}(x^2 + y^2 - 1).$$

The complete solution of the differential equation of the elastic membrane of the unit circle is finally

$$z(x, y, t) = c_1 \cos(\sqrt{6}t) \sqrt{\frac{3}{\pi}} (x^2 + y^2 - 1).$$

The remaining coefficient may be established by the initial condition. Assuming the center of the membrane is displaced by an amplitude A ,

$$z(0, 0, 0) = A = c_1 \sqrt{\frac{3}{\pi}} (-1).$$

from which follows

$$c_1 = -A \sqrt{\frac{\pi}{3}}.$$

The final solution is

$$z(x, y, t) = -A \cos(\sqrt{6}t) (x^2 + y^2 - 1).$$

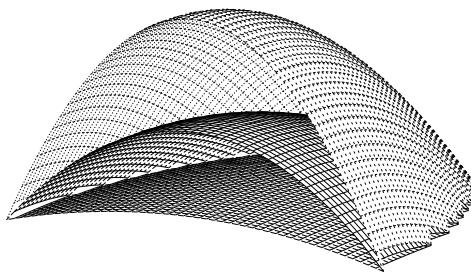


FIGURE 9.2 Vibration of elastic membrane

The shape of the solution is shown in [Figure 9.2](#). The figure shows the solution of the half-membrane at three distinct time steps. The jagged edges are artifacts of the discretization; the shape of membrane was the unit circle.

9.3.1 Nonzero boundary conditions

So far we restricted ourselves to trivial boundary conditions for the sake of clarity. In engineering practice, however, nonzero boundary conditions are very often imposed. These, also called enforced motion, boundary conditions are the subject of our focus here.

Let us consider the membrane with flexible boundary allowing some or all of the boundary points to attain nonzero displacement from the plane. Let the arclength differential of a section of the boundary in equilibrium denoted by ds . The reactive force on the section due to displacement z is

$$-p(s)z(x, y, t)ds,$$

where the negative sign indicates the force's effort to pull the boundary back towards the equilibrium position and opposite from the displacement. The potential energy of the boundary section may be computed by

$$p(s)ds \int z dz = \frac{1}{2}p(s)z^2 ds.$$

The total potential energy due to the reactive force on the boundary L is

$$E_p^L = \frac{1}{2} \int_L p(s)z^2 ds.$$

Applying Hamilton's principle for this scenario now yields

$$I(z) = \frac{1}{2} \int_{t_0}^{t_1} \left(\int \int_D \rho \left(\frac{\partial z}{\partial t} \right)^2 - \sigma \left(\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 \right) dx dy - \int_L p(s)z^2 ds \right) dt.$$

The newly introduced boundary integral's inconvenience may be avoided as follows. First, it may also be written as

$$\int_L p(s)z^2 ds = \frac{1}{2} \int_L \left(p(s)z^2 \frac{ds}{dy} dy + p(s)z^2 \frac{ds}{dx} dx \right).$$

Introducing the twice differentiable

$$P = \frac{1}{2} p z^2 \frac{ds}{dy}$$

and

$$Q = -\frac{1}{2} p z^2 \frac{ds}{dx}$$

functions that are defined on the boundary curve L the integral further changes to

$$\int_L pz^2 ds = \int_L (Pdy - Qdx).$$

Finally, with the help of Green's theorem we obtain

$$\int_L pz^2 ds = \int \int_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) dx dy.$$

Hence the variational form of this problem becomes

$$I(z) = \frac{1}{2} \int_{t_0}^{t_1} \int \int_D \left(\rho \left(\frac{\partial z}{\partial t} \right)^2 - \sigma \left(\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 \right) - \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) \right) dx dy dt.$$

This problem is identical to the one in [Section 3.5](#), the case of functionals with three independent variables. The two spatial independent variables are augmented in this case with time as the third independent variable. The corresponding Euler-Lagrange differential equation becomes the same as in the case of the fixed boundary

$$\sigma \Delta z = \rho \left(\frac{\partial^2 z}{\partial t^2} \right),$$

with the addition of the constraint on the boundary as

$$\sigma \frac{\partial z}{\partial n} + pz = 0,$$

where n is the normal of the boundary. The solution may again be sought in the form of

$$z(x, y, t) = a(t)b(x, y),$$

and as before, based on the same reasoning

$$a(t) = c_1 \cos(\sqrt{\lambda}t) + c_2 \sin(\sqrt{\lambda}t).$$

The $b(x, y)$ now must satisfy the following two equations.

$$\sigma \Delta b + \lambda \rho b = 0; (x, y) \in D,$$

and

$$\sigma \frac{\partial b}{\partial n} + pb = 0; (x, y) \in L.$$

The solution of these two equations follows the procedure established in the last section.

9.4 Bending of a beam under its own weight

The two analytic elasticity examples presented so far were one- and two-dimensional, respectively. The additional dimensions (the string's cross-section or the thickness of the membrane) were negligible and ignored in the presentation. In this section we address the phenomenon of the bending of a beam with a non-negligible cross-section and consider all three dimensions.

In order to deal with the problem of the beam, we introduce some basic concepts of elasticity for this specific case only. A fuller exposition of the topic will be in the next chapter. Let us consider an elastic beam with length L and cross-section area A . We consider the beam fully constrained at one end and free on the other, known as a cantilever beam, with a rectangular cross-section of width a along the z axis and height b along the y axis as shown in Figure 9.3. The axis of the beam is aligned along the x axis.

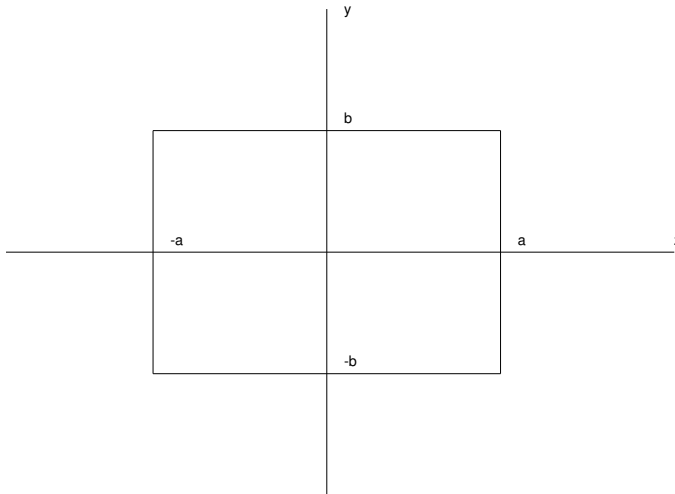


FIGURE 9.3 Beam cross-section

The relationship between the stress resulting from an axial force exerted on

the free end of the beam and its subsequent deformation is expressed by the well-known Hooke's law

$$\sigma = E\epsilon,$$

where the constant E , called Young's modulus, expresses the inherent elasticity of the material with regards to elongation. The relationship between the stress (σ) and the force (F) is

$$\sigma = \frac{F}{A},$$

The strain (ϵ) describes the relative deformation of the beam and in the axial case this is

$$\epsilon = \frac{dl}{l},$$

where dl is the elongation along the beam's longitudinal direction. In different deformation scenarios, like the ensuing bending, the formulation for the strain may vary and will be discussed in more detail later.

The energy equilibrium of this problem is again based on Hamilton's principle; however, since in this particular example we consider a static equilibrium, there is no kinetic energy. In this case Hamilton's principle becomes the principle of minimal potential energy. The two components of the potential energy are the internal strain energy and the work of forces acting on the body.

The internal energy related to the strain is

$$E_s = \frac{1}{2} \int_V \sigma \epsilon dV.$$

Substitution of Hooke's law yields

$$E_s = \frac{1}{2} E \int_V \epsilon \epsilon dA dx.$$

The strain energy of a particular cross-section is obtained by integrating as

$$E_s(x) = \frac{1}{2} E \int_{-b}^b \int_{-a}^a \epsilon^2 dz dy.$$

The bending will result in a curved shape with a radius of curvature r and a strain in the beam. Note that the radius of curvature is a function of the lengthwise location since the shape of the beam (the subject of our interest) is not a circle.

The relationship between the radius of curvature and the strain is established as follows. Above the neutral plane of the bending, that is the $x - z$ plane in our case, the beam is elongated and it is compressed below the plane.

Based on that at a certain distance y above or below the plane the strain is

$$\epsilon = \frac{y}{r}.$$

Note that since y is a signed quantity, above yields zero strain in the neutral plane, positive (tension) above the plane and negative (compression) below. Using this in the strain energy of a particular cross-section yields

$$E_s(x) = \frac{E}{2} \int_{-b}^b \int_{-a}^a \frac{y^2}{r^2} dz dy = \frac{E}{2} \frac{4ab^3}{3} \frac{1}{r^2} = \frac{EI}{2} \frac{1}{r^2},$$

where

$$I = \frac{4ab^3}{3}$$

is the moment of inertia of the rectangular cross-section. The total strain energy in the volume is

$$E_s = \frac{1}{2} EI \int_0^L \frac{1}{r^2} dx.$$

We assume that the only load on the beam is its weight. We denote the weight of the unit length with w . The bending moment generated by the weight of a cross-section with respect to the neutral plane is

$$dM = y dG,$$

where y is the distance from the neutral plane and dG is the weight of the cross-section. Using the unit length weight of the beam, we obtain the moment as

$$dM = y w dx = M(x) dx.$$

The total work of bending will be obtained by integrating along the length of the beam:

$$W = \int dM = \int_0^L M(x) dx = w \int_0^L y dx,$$

since the unit weight is constant. We are now ready to state the equilibrium of the beam

$$E_s = W$$

as a variational problem of the form

$$I(y) = \int_0^L (E_s(x) - M(x)) dx = \text{extremum},$$

or

$$I(y) = \int_0^L \left(\frac{1}{2} EI \frac{1}{r^2} - wy \right) dx = \text{extremum}.$$

Since the radius of curvature is reciprocal of the second derivative of the bent curve of the beam,

$$r = \frac{1}{y''(x)}$$

it follows that

$$I(y) = \int_0^L \left(\frac{1}{2} EI y''^2(x) - wy \right) dx = \text{extremum}.$$

The problem above is a special case of the form

$$I(y) = \int_L f(y, y'') dx = \text{extremum},$$

where neither the y nor the y' term exists explicitly. It is also a case of higher order derivatives discussed in [Section 4.2](#) and results in the Euler-Poisson equation of order 2:

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y''} = 0.$$

Since in this case

$$f(y, y'') = \frac{1}{2} EI y''^2 - wy,$$

the first term is simply

$$\frac{\partial f}{\partial y} = -w.$$

The second term vanishes as the first derivative of the unknown function is not explicitly present. With

$$\frac{\partial f}{\partial y''} = 2 \frac{1}{2} EI y'',$$

the third term becomes

$$\frac{d^2}{dx^2} \frac{\partial f}{\partial y''} = EI \frac{d^4}{dx^4} y.$$

Hence the solution obtained from the Euler-Poisson equation tailored for this case is

$$\frac{d^4 y}{dx^4} = \frac{w}{EI}.$$

Direct integration yields the final solution of

$$y(x) = \frac{w}{24EI} (x^4 + 4c_1 x^3 + 12c_2 x^2 + 24c_3 x + c_4),$$

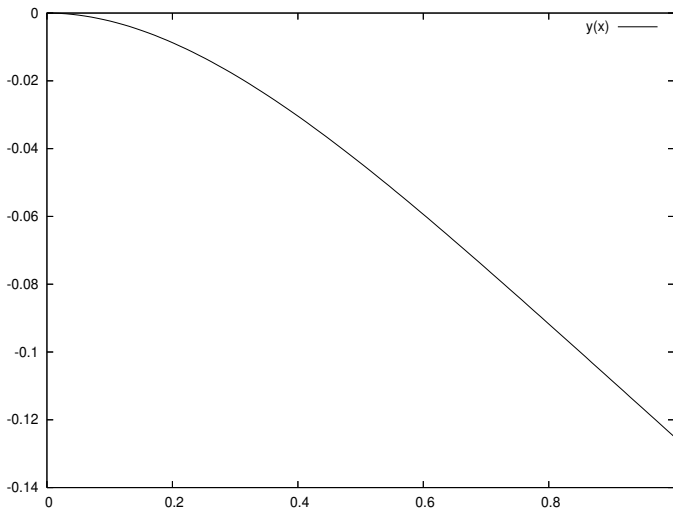


FIGURE 9.4 Beam profile under its weight

where the c_i are constants of integrations. The solution curve yields the shape of the bent beam shown in Figure 9.4.

In the figure unit physical coefficients were used for the sake of simplicity and the coefficients of integration are resolved from the boundary conditions as follows. At the fixed end, the beam is not deflected, hence

$$y(x = 0) = 0,$$

which implies

$$c_4 = 0.$$

Furthermore, at the fixed end, the tangent of the curve is horizontal as

$$y'(x = 0) = 0,$$

implying

$$c_3 = 0.$$

Finally at the free end the beam has no curvature, hence both second and third derivatives vanish. Therefore

$$y'''(L) = 0$$

yields

$$c_1 = -L,$$

and

$$y''(L) = 0$$

results in

$$c_2 = \frac{L^2}{2}.$$

With these, the final solution becomes

$$y(x) = \frac{w}{24EI}(x^4 - 4Lx^3 + 6L^2x^2),$$

which is the source of maximum deflection often quoted in engineering handbooks:

$$y(L) = \frac{wL^4}{8EI}.$$

Finally, it is worthwhile to point out the intriguing similarities between this problem and the natural spline solution of [Chapter 8](#).

The algebraic difficulties of analytic mechanical applications are rather overwhelming and become insurmountable with many real-world geometries. Computational mechanics is based on a discretization of the geometric continuum and describing its physical behavior in terms of generalized coordinates. Its focus is on computing numerical solutions to practical problems of engineering mechanics.

10.1 Three-dimensional elasticity

One of the fundamental concepts necessary to understanding continuum mechanical systems is a generic treatment of elasticity described in detail in the classical reference of the subject [12]. When an elastic continuum undergoes a one-dimensional deformation, like in the case of the elongated bar discussed in [Section 9.3](#), Young's modulus was adequate to describe the changes.

For a general three-dimensional elastic continuum we need another coefficient, introduced by Poisson, to capture the three-dimensional elastic behavior. Poisson's ratio measures the contraction of the cross-section while an object such as a beam is stretched. The ratio ν is defined as the ratio of the relative contraction and the relative elongation:

$$\nu = -\frac{dr}{r} / \frac{dl}{l}.$$

Here a beam with circular cross-section and radius r is assumed. Poisson's ratio is in the range of zero to $1/2$ and expresses the compressibility of the material. The two constants are also often related as

$$(mu = \frac{E}{2(1 + \nu)},$$

and

$$\lambda = \frac{E\nu}{(1 + \nu)(1 - 2\nu)}.$$

Here μ and λ are the so-called Lamé constants.

In a three-dimensional elastic body, the elasticity relations could vary significantly. Let us consider isotropic materials, whose elastic behavior is independent of the material orientation. In this case Young's modulus is replaced by an elasticity matrix whose terms are only dependent on the Lamé constants as follows

$$D = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}.$$

Viewing an infinitesimal cube of the three-dimensional body, there are six stress components on the element,

$$\underline{\sigma} = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \\ \tau_{yz} \\ \tau_{xz} \\ \tau_{xy} \end{bmatrix}.$$

The first three are normal and the second three are shear stresses. There are also six strain components

$$\underline{\epsilon} = \begin{bmatrix} \epsilon_x \\ \epsilon_y \\ \epsilon_z \\ \gamma_{yz} \\ \gamma_{xz} \\ \gamma_{xy} \end{bmatrix}.$$

The first three are extensional strains and the last three are rotational strains. The stress-strain relationship is described by the generalized Hooke's law as

$$\underline{\sigma} = D\underline{\epsilon}.$$

This will be the fundamental component of the computational techniques for elastic bodies. Let us further designate the location of an interior point of the elastic body with

$$\underline{r}(x, y, z) = x\underline{i} + y\underline{j} + z\underline{k} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

and the displacements of the point with

$$\underline{u}(x, y, z) = u\underline{i} + v\underline{j} + w\underline{k} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}.$$

Then the following strain relations hold:

$$\epsilon_x = \frac{\partial u}{\partial x},$$

$$\epsilon_y = \frac{\partial v}{\partial y},$$

and

$$\epsilon_z = \frac{\partial w}{\partial z}.$$

These extensional strains manifest the change of rate of the displacement of an interior point of the elastic continuum with respect to the coordinate directions.

The rotational strains are computed as

$$\gamma_{yz} = \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y},$$

$$\gamma_{xz} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x},$$

and

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}.$$

These terms define the rate of change of the angle between two lines crossing at the interior point that were perpendicular in the undeformed body and get distorted during the elastic deformation.

The strain energy contained in the three-dimensional elastic continuum is

$$E_s = \frac{1}{2} \int_V \underline{\sigma}^T \underline{\epsilon} dV = \frac{1}{2} \int_V \begin{bmatrix} \sigma_x & \sigma_y & \sigma_z & \tau_{yz} & \tau_{xz} & \tau_{xy} \end{bmatrix} \begin{bmatrix} \epsilon_x \\ \epsilon_y \\ \epsilon_z \\ \gamma_{yz} \\ \gamma_{xz} \\ \gamma_{xy} \end{bmatrix} dV.$$

We will also consider distributed forces acting at every point of the volume (like the weight of the beam in [Section 9.4](#)), described by

$$\underline{f} = f_x \underline{i} + f_y \underline{j} + f_z \underline{k} = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix}.$$

The work of these forces is based on the displacements they caused at the certain points and computed as

$$W = \int_V \underline{u}^T \underline{f} dV = \int_V \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} dV.$$

The above two energy components constitute the total potential energy of the volume as

$$E_p = E_s - W.$$

In order to evaluate the dynamic behavior of the three-dimensional body, the kinetic energy also needs to be computed. Let the velocities at every point of the volume be described by

$$\underline{\dot{u}}(x, y, z) = \dot{u} \underline{i} + \dot{v} \underline{j} + \dot{w} \underline{k} = \begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix}.$$

With a mass density of ρ , assumed to be constant throughout the volume, the kinetic energy of the body is

$$E_k = \frac{1}{2} \rho \int_V \underline{\dot{u}}^T \underline{\dot{u}} dV.$$

We are now in the position to write the variational statement describing the equilibrium of the three-dimensional elastic body:

$$I(\underline{u}(x, y, z)) = \int_V (E_k - E_p) dV = \text{extremum},$$

with the expressed unknown displacement solution of the body at every (x, y, z) point which is the subject of the computational solution discussed in the next sections.

10.2 Lagrange's equations of motion

The equations will be obtained by finding an approximate solution of the variational problem based on the total energy of the system as follows. For

$$I(\underline{u}) = \int_V f(\underline{u}(x, y, z)) dV = \text{extremum},$$

we seek the (approximate) solution in the form

$$\underline{u}(x, y, z) = \sum_{i=1}^n \underline{q}_i N_i(x, y, z).$$

The yet unknown coefficients, the \underline{q}_i values are displacements at $i = 1, 2, \dots$ discrete locations inside the volume. These are also known as generalized displacements and discussed eloquently in [11].

The variational problem, based on the total energy of the system, in this case is

$$I(\underline{u}) = \int_V \left(\frac{1}{2} \rho \dot{\underline{u}}^T \dot{\underline{u}} - \left(\frac{1}{2} \underline{\sigma}^T \underline{\epsilon} - \underline{u}^T \underline{f} \right) \right) dV = \text{extremum}.$$

Let us organize the generalized displacements as

$$\underline{q} = \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix},$$

where, in adherence to our three-dimensional focus

$$\underline{q}_i = \begin{bmatrix} q_{i,x} \\ q_{i,y} \\ q_{i,z} \end{bmatrix}.$$

Using this, the approximate solution becomes

$$\underline{u}(x, y, z) = N \underline{q}$$

with the matrix of basis functions

$$N(x, y, z) = N_x \underline{i} + N_y \underline{j} + N_z \underline{k} = \begin{bmatrix} N_1 & 0 & 0 & \dots & N_n & 0 & 0 \\ 0 & N_1 & 0 & \dots & 0 & N_n & 0 \\ 0 & 0 & N_1 & \dots & 0 & 0 & N_n \end{bmatrix}.$$

The basis functions are usually low order polynomials of x, y, z , and will be discussed in detail in [Section 10.5.2](#).

Let us apply this to the terms of our variational problem, starting with the kinetic energy. Assuming that the velocity is also a function of the generalized velocities,

$$\dot{\underline{u}}(x, y, z) = N \dot{\underline{q}},$$

where

$$\underline{\dot{q}} = \begin{bmatrix} \dot{q}_1 \\ \vdots \\ \dot{q}_n \end{bmatrix},$$

we obtain

$$E_k = \int_V \frac{1}{2} \rho \underline{\dot{u}}^T \underline{\dot{u}} dV = \frac{1}{2} \underline{\dot{q}}^T \int_V N^T \rho N dV \underline{\dot{q}}.$$

Introducing the mass matrix

$$M = \int_V N^T \rho N dV,$$

the final form of the kinetic energy becomes

$$E_k = \frac{1}{2} \underline{\dot{q}}^T M \underline{\dot{q}}.$$

Now let's focus on the strain energy. Note that the strain is now also expressed in terms of the basis functions. Hence

$$\epsilon(N) = \begin{bmatrix} \sum_{i=1}^n \underline{q}_i^T \frac{\partial N}{\partial x} \\ \sum_{i=1}^n \underline{q}_i^T \frac{\partial N}{\partial y} \\ \sum_{i=1}^n \underline{q}_i^T \frac{\partial N}{\partial z} \\ \sum_{i=1}^n \underline{q}_i^T \left(\frac{\partial N}{\partial z} + \frac{\partial N}{\partial y} \right) \\ \sum_{i=1}^n \underline{q}_i^T \left(\frac{\partial N}{\partial z} + \frac{\partial N}{\partial x} \right) \\ \sum_{i=1}^n \underline{q}_i^T \left(\frac{\partial N}{\partial y} + \frac{\partial N}{\partial x} \right) \end{bmatrix},$$

or in matrix form

$$\underline{\epsilon}(N) = B \underline{q},$$

where the columns of B are

$$B_i = \begin{bmatrix} \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} \\ \frac{\partial N_i}{\partial z} \\ \frac{\partial N_i}{\partial z} + \frac{\partial N_i}{\partial y} \\ \frac{\partial N_i}{\partial z} + \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} + \frac{\partial N_i}{\partial x} \end{bmatrix}.$$

With this, the integral becomes

$$\int_V \underline{\epsilon}^T(N) D \underline{\epsilon}(N) dV = \int_V \underline{q}^T B^T D B \underline{q} dV.$$

The total strain energy in the system is

$$E_s = \frac{1}{2} \underline{q}^T \int_V B^T D B dV \underline{q}.$$

Introducing the stiffness matrix of the system as

$$K = \int_V B^T D B dV,$$

the strain energy term is of final form

$$E_s = \frac{1}{2} \underline{q}^T K \underline{q}.$$

A similar approach on the second term of the potential energy yields

$$\int_V \underline{q}^T N^T \underline{f} dV.$$

Introducing the total active force vector on the system as

$$F = \int_V N^T \underline{f} dV,$$

this term becomes

$$\underline{q}^T F.$$

The total potential energy is their difference

$$E_p = \frac{1}{2} \underline{q}^T K \underline{q} - \underline{q}^T F.$$

We are ready to find the value of the unknown coefficients. Since our variational problem contains multiple functions

$$I(\underline{q}, \dot{\underline{q}}, \ddot{\underline{q}}) = \text{extremum},$$

we produce a system of Euler-Lagrange equation, which is called Lagrange's equations of motion. Note the emphasis on the plural. With the Lagrangian expression

$$L = E_k - E_p$$

Lagrange's equations of motion are

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0; i = 1, 2, \dots, n.$$

These equations provide the computational solution for practical mechanical systems. Since in this case

$$\frac{\partial E_k}{\partial q_i} = 0,$$

and

$$\frac{\partial E_p}{\partial \dot{q}_i} = 0,$$

we obtain a simpler version of

$$\frac{d}{dt} \frac{\partial E_k}{\partial \dot{q}_i} + \frac{\partial E_p}{\partial q_i} = 0; i = 1, 2, \dots, n.$$

The first term is evaluated as

$$\frac{\partial E_k}{\partial \dot{\underline{q}}} = M \dot{\underline{q}}.$$

Then

$$\frac{d}{dt}(M \dot{\underline{q}}) = M \ddot{\underline{q}}.$$

Here the generalized accelerations are

$$\ddot{\underline{q}} = \begin{bmatrix} \ddot{q}_1 \\ \vdots \\ \ddot{q}_n \end{bmatrix}.$$

The second term results in the two terms

$$\frac{\partial E_p}{\partial \underline{q}} = K \underline{q} - F = 0.$$

The final result is

$$M \ddot{\underline{q}} + K \underline{q} = F.$$

This is the well-known equation of the forced undamped vibration of a three-dimensional elastic body.

10.2.1 Hamilton's canonical equations

As a final note to the Lagrange equations, we play homage to Hamilton's great contributions to the area and review his ingenious simplification of Lagrange's equations of motion.

Since the Lagrangian is a function of generalized coordinates and time as

$$L = L(\underline{q}, \dot{\underline{q}}, t) = L(q_i, \dot{q}_i, t),$$

for $i = 1, 2, \dots, n$, one can compute the differential of the Lagrangian as

$$dL = \frac{\partial L}{\partial t} dt + \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i} dq_i + \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right).$$

Introducing

$$p_i = \frac{\partial L}{\partial \dot{q}_i},$$

and

$$\dot{p}_i = \frac{\partial L}{\partial q_i},$$

the differential becomes

$$dL = \frac{\partial L}{\partial t} dt + \sum_{i=1}^n (\dot{p}_i dq_i + p_i d\dot{q}_i).$$

Reordering yields

$$d\left(\sum_{i=1}^n p_i \dot{q}_i - L\right) = -\frac{\partial L}{\partial t} dt - \sum_{i=1}^n (\dot{p}_i dq_i - \dot{q}_i dp_i).$$

The total differential of the term

$$H = \sum_{i=1}^n (p_i \dot{q}_i - L) = f(p_i, q_i, t),$$

now called the Hamiltonian is

$$dH = \sum_{i=1}^n \left(\frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i \right) + \frac{\partial H}{\partial t} dt.$$

The relationship with the differential of the Lagrangian is

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t},$$

resulting in Hamilton's canonical equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i},$$

and

$$\dot{p}_i = -\frac{\partial H}{\partial q_i},$$

for $i = 1, 2, \dots, n$. The p_i, q_i are called canonical variables. These are $2n$ first order equations as opposed to the n second order equations of Lagrange; otherwise they are equivalent. In many cases these equations are simpler to solve.

10.3 Heat conduction

While staying on the mechanics territory, we now explore the area of heat conduction. This phenomenon occurs when the temperature between two areas of a body differs. In this application every point in space is associated with a scalar quantity, the temperature, hence these type of problems are called scalar field problems.

For our discussion, we will assume that the body does not deform under the temperature load. This assumption, of course, may be violated in real life. Serious warping of objects left in the sun is a strong example of that scenario.

Two more restrictions we impose. We'll consider two-dimensional problems for simplification of the discussion. We will also only consider the steady state solution case, when the temperature at a certain point is independent of the time.

The differential equation of the heat conduction per [3] for this case is

$$\frac{\partial}{\partial x}(k \frac{\partial T}{\partial x}) + \frac{\partial}{\partial y}(k \frac{\partial T}{\partial y}) + Q = 0,$$

where the temperature field

$$T = T(x, y)$$

is a function of the location and k is the thermal conductivity of the material of the object. In general the conductivity may be a function of the location as well. The Q is a heat source in the model, called such when it generates heat, but called a sink when it absorbs heat.

Since the differential equation is given, we'll use the techniques developed in [Chapter 5](#) when dealing with the inverse variational problem. Accordingly, the solution of a differential equation of the form

$$Au + f = 0,$$

corresponds to the extremum value of the functional

$$I(u) = (Au, u) - 2(u, f).$$

For the heat conduction differential equation, this means

$$I(T) = \frac{1}{2} \int \int_D (k((\frac{\partial T}{\partial x})^2 + (\frac{\partial T}{\partial y})^2) - 2QT) dx dy = \text{extremum}.$$

Following the avenue charted in the last section for the elasticity problem, we'll approximate the temperature field in terms of basis functions by

$$T(x, y) = \sum_{i=1}^n T_i N_i.$$

Then

$$\frac{\partial T(x, y)}{\partial x} = \sum_{i=1}^n T_i \frac{\partial N_i}{\partial x},$$

and

$$\frac{\partial T(x, y)}{\partial y} = \sum_{i=1}^n T_i \frac{\partial N_i}{\partial y}.$$

We introduce a vector of the generalized temperatures

$$\underline{T} = \begin{bmatrix} T_1 \\ \vdots \\ T_n \end{bmatrix}$$

and $B_i, i = 1, \dots, n$ columns of a B matrix as

$$B_i = \begin{bmatrix} \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} \end{bmatrix}.$$

We also concatenate the N_i into the matrix N

$$N = [N_1 \dots N_n].$$

This architecture of the N matrix is simpler than in the case of the elasticity, reflecting the fact that this is a scalar field problem. The elasticity was a vector field problem as the solution quantity at each point was the displacement vector of three dimensions.

Then in matrix form we obtain the approximate temperature field solution as

$$\overline{T}(x, y) = N \underline{T},$$

and

$$\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 = (B \underline{T})^T.$$

We apply these to the variational problem. The first part becomes

$$\frac{1}{2} \int \int_D k \left(\left(\frac{\partial T}{\partial x} \right)^2 + \left(\frac{\partial T}{\partial y} \right)^2 \right) dx dy = \frac{1}{2} \int \int_D k \underline{T}^T B^T B \underline{T} dx dy.$$

The second part changes as

$$-\int \int_D Q T dx dy = -\int \int_D Q N \underline{T} dx dy.$$

With the introduction of a “temperature stiffness matrix” of

$$K = \int \int_D k B^T B dx dy,$$

and the load vector of

$$\underline{Q} = \int \int_D Q N dx dy,$$

finally, we obtain the variational problem of

$$I(\bar{T}) = \frac{1}{2} \underline{T}^T K \underline{T} - \underline{Q} \underline{T} = \text{extremum},$$

from which the now familiar result of

$$K \underline{T} = \underline{Q}$$

emerges.

This simpler two-dimensional case will be used to illuminate the computational details of the generalized coordinate selection and processing in [Section 10.5](#), after reviewing another mechanical application involving fluids.

10.4 Fluid mechanics

As a final application we discuss a problem of fluid mechanics discussed in detail by [14], where the fluid is partially or fully surrounded by an external structure and the dissipation of energy into the surrounding space is negligible.

Assuming small motions, the equilibrium of a compressible fluid inside a cavity is governed by the Euler equation of the form

$$\rho \ddot{u} = -\nabla p,$$

where \ddot{u} is the acceleration of the particles and p is the pressure in the fluid. Furthermore, ρ is the density and ∇ is the differential operator.

We also assume locally linear pressure-velocity behavior of the fluid as

$$p = -b \nabla u,$$

where b is the so-called bulk modulus related to the density of the fluid and the speed of sound. Differentiating twice with respect to time and substituting the Euler equation we get Helmholtz's equation describing the behavior of the fluid:

$$\frac{1}{b}\ddot{p} = \nabla \cdot \left(\frac{1}{\rho} \nabla p \right).$$

The following boundary conditions are also applied. At a structure-fluid interface

$$\frac{\partial p}{\partial \underline{n}} = -\rho \ddot{u}_n, \quad (10.1)$$

where \underline{n} is the direction of the outward normal. At free surfaces:

$$u = p = 0.$$

Since the equilibrium differential equation of the physical phenomenon is given at this time, the inverse problem approach introduced in [Chapter 5](#) will be used again. Accordingly, for the differential equation of

$$Au = 0,$$

the variational problem of

$$I(u) = (Au, u)$$

applies. Here the inner product is defined over the continuum. For our case this results in

$$\int \int \int_V \left[\frac{1}{b} \ddot{p} - \frac{1}{\rho} \nabla \cdot \nabla p \right] p dV = 0. \quad (10.2)$$

Following the earlier sections, we will also assume that the pressure field is approximated by basis functions as:

$$p(x, y, z) = \sum_{i=1}^n N_i p_i = N \underline{p}.$$

The same holds for the derivatives:

$$\ddot{p}(x, y, z) = N \ddot{\underline{p}}.$$

Separating the two parts of the equation (10.2), the first yields

$$\int_V \frac{1}{b} \ddot{p} p dV = \int_V \frac{1}{b} p \ddot{p} dV = \underline{p}^T \int_V \frac{1}{b} N^T N dV \ddot{\underline{p}}.$$

Introducing the mass matrix

$$M = \int_V \frac{1}{b} N^T N dV,$$

this term simplifies to

$$\int_V \frac{1}{b} \ddot{p} p dV = \underline{p}^T M \ddot{\underline{p}}.$$

Let us now turn our attention to the second part of equation (10.2). Integrating by parts yields

$$- \int_V \left(\frac{1}{\rho} \nabla \cdot \nabla p p \right) dV = \int_V \frac{1}{\rho} \nabla p \cdot \nabla p dV - \int_S \frac{1}{\rho} \nabla p p dS.$$

From the above assumptions, it follows that

$$\nabla p = \nabla N \underline{p},$$

and we obtain

$$\underline{p}^T \int_V \left(\frac{1}{\rho} \nabla N^T \right) \nabla N dV \underline{p} + \underline{p}^T \int_S N^T \ddot{u}_n dS.$$

Here the boundary condition stated in equation (10.1) was used. Introducing

$$K = \int_V \frac{1}{\rho} \nabla N^T \nabla N dV,$$

the first part simplifies to

$$\underline{p}^T K \underline{p}.$$

The force exerted on the boundary by the surrounding structure is

$$F = \int_S N^T \ddot{u}_n dS.$$

Substituting and reordering yields

$$\underline{p}^T M \ddot{\underline{p}} + \underline{p}^T K \underline{p} + \underline{p}^T F = 0.$$

Finally, the equilibrium equation is

$$M \ddot{\underline{p}} + K \underline{p} + F = 0.$$

This, as all the similar problems of this chapter, may be solved by efficient numerical linear algebra computations and will not be discussed further here.

10.5 Computational techniques

In order to conclude the mechanical applications presented in the past three sections, we need to elucidate their common, and thus far omitted, computational details. They are, the discretization of the continuum by finite elements covering the domain, and the computation of the basis functions used in the approximate solutions. They are following the various applications to further emphasize the point of their commonality, or even application independence.

These topics are components of the computational techniques of finite element technology that has achieved an unparalleled application success. The topic details cover an extensive territory [10], but we'll discuss only the necessary concepts following here.

10.5.1 Discretization of continua

The discretization of generic three- or two-dimensional continua is usually by finite elements of simple shapes, such as tetrahedra or triangles. The foundation of many general methods of discretization (commonly called meshing) is the classical Delaunay triangulation method. The Delaunay triangulation technique in turn is based on Voronoi polygons. The Voronoi polygon, assigned to a certain point of a set of points in the plane, contains all the points that are closer to the selected point than to any other point of the set.

In [Figure 10.1](#) the dots represent such set of points. The irregular (dotted line) hexagon containing one point in the middle is the Voronoi polygon of the point in the center. It is easy to see that the points inside the polygon are closer to the center point than to any other points of the set. It is also quite intuitive that the edges of the Voronoi polygon are the perpendicular bisectors of the line segments connecting the points of the set.

The union of the Voronoi polygons of all the points in the set completely covers the plane. It follows that the Voronoi polygon of two points of the set do not have common interior points; at most they share points on their common boundary.

The definition and process generalizes to three dimensions very easily. If the set of points are in space, the points closest to a certain point define a Voronoi polyhedron.

The Delaunay triangulation process is based on the Voronoi polygons by constructing Delaunay edges connecting those points whose Voronoi polygons have a common edge. Constructing all such possible edges will result in the

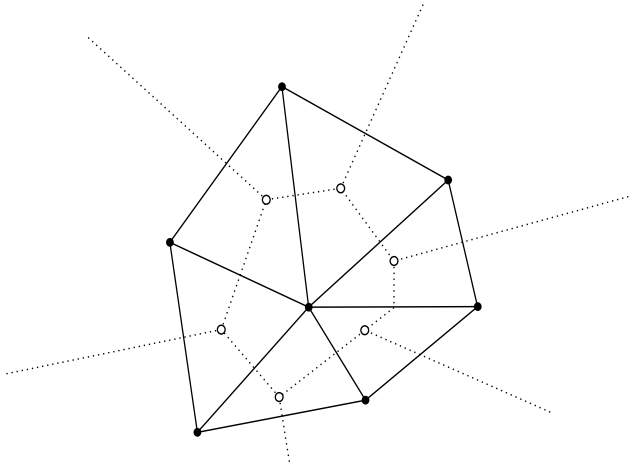


FIGURE 10.1 Delaunay triangularization

covering of the planar region of our interest with triangular regions, the Delaunay triangles.

The process starts with placing vertices on the boundary of the domain in an equally spaced fashion. The Voronoi polygons of all boundary points are created and interior points are generated gradually proceeding inwards by creating Delaunay triangles. This is called the advancing front technique.

The process extends quite naturally and covers the plane as shown in Figure 10.1 with six Delaunay triangles where the dotted lines are the edges of the Voronoi polygons and the solid lines depict the Delaunay edges. It is known that under the given definitions no two Delaunay edges cross each other.

Finally, in three dimensions the Delaunay edges are defined as lines connecting points that share a common Voronoi facet (a face of a Voronoi polyhedron). Furthermore, the Delaunay facets are defined by points that share a common Voronoi edge (an edge of a Voronoi polyhedron). In general each edge is shared by exactly three Voronoi polyhedra; hence, the Delaunay regions' facets are going to be triangles. The Delaunay regions connect points of Voronoi polyhedra that share a common vertex. Since in general the number of such polyhedra is four, the generated Delaunay regions will be tetrahedra.

The triangulation method generalized into three dimensions is called tessellation.

10.5.2 Computation of basis functions

As mentioned before, in order to approximate the solution inside the domain, we use low order polynomial basis functions called shape functions in the finite element nomenclature. For a triangular element discretization of a two-dimensional domain, bilinear interpolation functions of the form

$$u(x, y) = a + bx + cy$$

are commonly used.

In order to find the shape functions, let us consider a triangular element of the $x - y$ plane with corner nodes (x_1, y_1) , (x_2, y_2) and (x_3, y_3) . For this particular triangle we seek three shape functions satisfying

$$N_1 + N_2 + N_3 = 1.$$

We also require that the nonzero shape function at a certain node point reduce to zero at the other two nodes, respectively. The interpolations are continuous across the neighboring elements. On an edge between two triangles, the approximation is linear and the same when approached from either element. Specifically, along the edge between nodes 1 and 2, the shape function N_3 is zero. The shape functions N_1 and N_2 along this edge are the same when calculated from an element on either side of that edge.

The solution for all the nodes of a particular triangular element e can be expressed as

$$u_e = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

This system of equations is solved for the unknown coefficients as

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} N_{1,1} & N_{1,2} & N_{1,3} \\ N_{2,1} & N_{2,2} & N_{2,3} \\ N_{3,1} & N_{3,2} & N_{3,3} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

By substituting into the matrix form of the bilinear interpolation function

$$u(x, y) = \begin{bmatrix} 1 & x & y \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & x & y \end{bmatrix} \begin{bmatrix} N_{1,1} & N_{1,2} & N_{1,3} \\ N_{2,1} & N_{2,2} & N_{2,3} \\ N_{3,1} & N_{3,2} & N_{3,3} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},$$

we get

$$u(x, y) = \begin{bmatrix} N_1 & N_2 & N_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

Here the N_1, N_2, N_3 shape functions are

$$N_1 = N_{1,1} + N_{2,1}x + N_{3,1}y,$$

$$N_2 = N_{1,2} + N_{2,2}x + N_{3,2}y,$$

and

$$N_3 = N_{1,3} + N_{2,3}x + N_{3,3}y.$$

The shape functions, as their name indicates, solely depend on the coordinates of the corner nodes, hence the shape of the particular triangular element of the domain. With these we are now able to approximate the relationship between the solution value inside an element in terms of the solutions at the corner node points, the generalized coordinates, as

$$u(x, y) = N_1u_1 + N_2u_2 + N_3u_3.$$

In order to compute the terms of a particular matrix, for example the $M_{i,j}$ terms of the mass matrix M introduced in [Section 10.2](#), we proceed element by element. We consider all the nodes bounding a particular element and compute all the partial $\overline{M}_{i,j}$ terms produced by that particular element. Thus,

$$M_e = \sum_{(i,j) \in e} \overline{M}_{i,j} = \rho \int \int_e N^T N dx dy,$$

where the N matrix is the local shape function matrix of the particular triangular element. Finally, the M matrix is assembled as

$$M = \sum_{e=1}^m M_e,$$

where the summation is based on the topological relation of elements. If, for example, the element is described by nodes 1, 2 and 3, then the terms in M_e contribute to the terms of the 1st, 2nd and 3rd columns and rows of the global M matrix.

Let us assume that another element is adjacent to the edge between nodes 2 and 3, its other node being 4. Then by similar arguments, the second element's matrix terms (depending on that particular element's shape) will contribute to the 2nd, 3rd and 4th columns and rows of the global matrix. This process is continued for all the elements contained in the finite element discretization.

The derivatives of the local shape functions are as follows:

$$\begin{bmatrix} \frac{\partial u(x,y)}{\partial x} \\ \frac{\partial u(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial x} & \frac{\partial N_3}{\partial x} \\ \frac{\partial N_1}{\partial y} & \frac{\partial N_2}{\partial y} & \frac{\partial N_3}{\partial y} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} B_{1,1} & B_{1,2} & B_{1,3} \\ B_{2,1} & B_{2,2} & B_{3,3} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = B u_e.$$

With this the elemental stiffness matrix may be computed as well as

$$K_e = \int \int B^T D B dx dy.$$

The resulting element matrices are of order 3 by 3 for the above two-dimensional scalar field problem. For a three-dimensional elasticity (vector field) problem the element matrices may be of order 12, a result of the each of the 4 nodes of the tetrahedron having three degrees of freedom.

In practical implementations of the finite element method, the element matrices are generated as a transformation from a standard, parametrically defined element. For these cases the shape functions and their derivatives may be precomputed and appropriately transformed as shown in [15]. The integrals posed above are usually numerically evaluated.

In conclusion, let us emphasize the fact that in all three mechanical engineering disciplines we used the same mathematical techniques to capture the behavior of the physical phenomenon over the geometric domain. Furthermore, it is important to notice the techniques' transcendence of the mechanical disciplines. For example, as demonstrated in [2], the governing equation in electrostatics is also Poisson's equation, albeit the participating terms have different physical meaning.

The applicability of a certain variational problem to unrelated disciplines is straightforward, one only needs to adhere to the differences in the physics. This fact makes the techniques demonstrated in this book extremely useful in a wide variety of engineering applications.

Closing Remarks

Hopefully the engineers reading this book find its theoretical foundation clear and concise, and the analytic and computational examples enlightening. The focus of the applications on mechanical engineering reflects only the author's personal expertise and is not meant to imply any restriction of applicability to other engineering disciplines.

The book was designed to be a self-contained coverage of the topic specifically addressed to the practicing engineer or engineering student reader. Difficult discussions about spaces of functions and rigorous proofs were avoided to make the topic accessible with a standard engineering mathematics foundation.

With more numerical examples and exercises added, the book may be used as a textbook in the engineering curriculum. It could also be a practical alternative to the abstract approaches frequently used in teaching calculus of variations in mathematics courses.

The [reference](#) list reflects the lack of recent attention to the topic of the calculus of variations. The original publications, however, are not only listed here for historical homage. The most readable, despite the archaic style, may be the oldest ones, albeit most of them were not specifically written for an engineering audience. This shortcoming is intended to be corrected by this book.

The reference list is also rather short, containing only those publications that directly influenced the writing of this book. They are all available, some of them in inexpensive reprints, and accessible in English.

Notation

Notation	Meaning
$f(x)$	Function of one variable
$f(x, y), F(x, y)$	Function of two variables
r	Radius of curvature
g	Acceleration of gravity
\underline{r}	Vector in Cartesian coordinates
\underline{r}_x	First partial derivative with respect to x
\underline{r}_y	First partial derivative with respect to y
$\dot{\underline{r}}$	First parametric derivative
$\ddot{\underline{r}}$	Second parametric derivative
s	Arclength
$s(t)$	One-dimensional spline function
$s(u, v)$	Two-dimensional spline function
$y'(x), f'(x)$	First derivative
$y''(x), f''(x)$	Second derivative
\underline{n}	Normal vector
\underline{t}	Tangent vector
\underline{b}	Bi-normal vector
∇	Gradient operator
Δ	Laplace operator
κ	Curvature
κ_g	Geodesic curvature
κ_n	Normal curvature
κ_m	Mean curvature
δI	Variation of integral functional
Γ_{ij}^k	Christoffel symbols
σ	Stress
ϵ	Strain
ν	Poisson's ratio
λ	Eigenvalue
ρ	Material density

$B_{i,k}(t)$	B-spline basis function
E, F, G	First fundamental quantities
E	Young's modulus
F	Active force
E_s	Strain energy
E_k	Kinetic energy
E_p	Potential energy
$I()$	Integral functional
I	Moment of inertia
M	Moment
N	Shape function matrix
S	Surface area
T	Surface (traction) force
V	Volume
W_e	External work

List of Tables

5.1	Eigenpairs of Legendre equation	66
6.1	Accuracy of Euler's method	71

List of Figures

1.1	Alternative solutions example	5
1.2	Solution of the brachistochrone problem	17
1.3	Trajectory of particle	22
2.1	Maximum area under curves	31
2.2	The catenary curve	34
3.1	Saddle surface	42
3.2	Catenoid surface	44
5.1	Eigensolutions of Legendre's equation	66
6.1	Accuracy of the Ritz solution	74
6.2	Domain with overlapping circular regions	80
6.3	Solution of Poisson's equation	82
7.1	Geodesic curve of a cylinder	94
8.1	Natural spline approximation	104
8.2	Smooth B-spline approximation	109
8.3	Point constrained B-spline	112
8.4	Tangent constrained B-spline	115
9.1	Normal modes of elastic string	124
9.2	Vibration of elastic membrane	129
9.3	Beam cross-section	132
9.4	Beam profile under its weight	136
10.1	Delaunay triangularization	154

References

- [1] Barnhill, R. E. and Riesenfeld, R. F.: Computer-aided geometric design, Academic Press, New York, 1974
- [2] Brauer, J. R.: Magnetic actuators and sensors, Wiley, New Jersey, 2006
- [3] Carslaw, H. S. and Jaeger, J. C.: Conduction of heat in solids, Oxford University Press, Cambridge, 1959
- [4] Courant, R. and Hilbert, D.: Methods of mathematical physics, Wiley, New York, 1953
- [5] Ewing, G. M.: Calculus of variations with applications, Norton, New York, 1969
- [6] Forsyth, A. R.: Lectures on differential geometry of curves and surfaces, Cambridge University Press, 1950
- [7] Fox, C.: An introduction to the calculus of variations, Oxford University Press, London, 1950
- [8] Gelfand, I. M. and Fomin, S. V.: Calculus of variations, Prentice-Hall, Englewood Cliffs, New Jersey, 1963
- [9] Kantorovich, L. V. and Krylov, V. I.: Approximate methods of higher analysis, Inter-science, New York, 1958
- [10] Komzsik, L.: Computational techniques of finite element analysis, CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2005
- [11] Lanczos, C.: The variational principles of mechanics, University of Toronto Press, Toronto, 1949
- [12] Timoshenko, S. and Goodier, J. N.: Theory of elasticity, McGraw-Hill, New York, 1951
- [13] Weinstock, R.: Calculus of variations, McGraw-Hill, New York, 1953
- [14] Warsi, Z. U. A.: Fluid dynamics, CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2006
- [15] Zienkiewicz, O. C.: The finite element method, McGraw-Hill, New York, 1968

Index

- auxiliary function, 4
- B-spline, 104
- basis functions, 71
- Beltrami's formula, 10
- Bernoulli, 14
- Bessel functions, 63
- bi-normal vector, 96
- bilinear interpolation, 155
- boundary conditions, 3
- brachistochrone, 10
- bulk modulus, 151
- CAD, 105, 116
- CAM, 43, 116
- catenary curve, 34
- catenoid, 44
- Chebyshev polynomials, 63
- chord, 114
- Christoffel symbols, 91
- compressible fluid, 150
- constrained splines, 109
- constrained variational problem, 29
- control points, 107
- curvature, 96
- curvature vector, 96
- Delaunay triangles, 153
- direct method, 69
- Dirichlet boundary condition, 58
- eigensolution, 61
- eigenvalue, 61
- Euler, 20
- Euler equation, 150
- Euler's method, 69
- Euler-Lagrange equation, 6
- Euler-Poisson equation, 50
- extremum of functional, 3
- Fermat's principle, 18
- finite element matrix assembly, 156
- first fundamental quantities, 42, 88
- first variation, 7
- frequency, 123
- fundamental lemma, 6
- Galerkin's method, 76
- generalized displacements, 143
- geodesic curvature, 96
- geodesic curve, 87
- great circle, 90
- Green's identity, 40, 45, 46, 60
- Green's theorem, 131
- half-period, 123
- Hamilton's principle, 14, 119
- Hamiltonian, 147
- heat conduction, 148
- helix, 93
- Helmholtz equation, 151
- hemisphere, 90
- Hermite polynomials, 63
- Hooke's law, 133
- inverse problem, 57
- Jordan curve, 41
- Kantorovich's method, 78
- kinetic energy, 14
- knot points, 105
- knot values, 105
- Lagrange multiplier, 28, 54, 59, 61, 65
- Lagrange's equations, 145
- Laguerre polynomials, 63
- Lamé constants, 140

- Laplace operator, 57, 58
- Legendre equation, 65
- Legendre polynomials, 65
- Legendre test, 8
- mean curvature, 40
- mean-value theorem, 7
- minimal surfaces, 41
- Minkowski space, 98
- natural frequency, 123
- natural spline, 102
- necessary conditions, 8
- Neumann boundary conditions, 52
- non-uniform splines, 105
- normal curvature, 96
- normal mode, 123
- normal vector, 96
- period, 123
- Plateau's problem, 41
- Poisson's ratio, 139
- potential energy, 14
- pressure, 150
- principal curvature, 40
- principle of least action, 20
- principle of minimal potential energy, 31, 133
- Ritz method, 71
- second variation, 8
- shape functions, 155
- Snell's law, 19
- speed of sound, 151
- steady state solution, 148
- strain, 133
- strain energy, 144
- stress, 133
- strong extremum, 8
- Sturm-Liouville equations, 62
- surface of revolution, 43
- tangent constraints, 112
- tangent vector, 96
- tautochrone, 18
- temperature field, 149
- tessellation, 155
- uniform parameterization, 108
- uniform splines, 105
- velocity, 150
- weak extremum, 8
- weighted residuals, 76
- Young's modulus, 133

Applied Calculus of Variations for Engineers

"Dr. Komzsik's latest book offers the readers, especially engineering graduate students, a very valuable, application-oriented, unique, and clear explanation about the topic. As an engineering university professor (and the author of multiple textbooks), I firmly believe that this is a must-read book for engineers and applied mathematicians who wish to learn the important topic of calculus of variations. I have been a loyal reader of all of Dr. Komzsik's previously published books, due to his consistent and clear presentation style, which is demonstrated in this book as well."

—Professor Duc T. Nguyen

Multidisciplinary Parallel-Vector Computation Center
Old Dominion University, Norfolk, Virginia, USA

About the book

The subject of calculus of variations is to find optimal solutions to engineering problems where the optimum may be a certain quantity, a shape, or a function. **Applied Calculus of Variations for Engineers** addresses this very important mathematical area applicable to many engineering disciplines. Its unique, application-oriented approach sets it apart from the theoretical treatises of most texts. It is aimed at enhancing the engineer's understanding of the topic as well as aiding in the application of the concepts in a variety of engineering disciplines.

86626



CRC Press
Taylor & Francis Group
an **Informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

ISBN: 978-1-4200-8662-1



www.crcpress.com