

Liming Dai · Reza N. Jazar *Editors*

# Nonlinear Approaches in Engineering Applications

Applied Mechanics, Vibration Control,  
and Numerical Analysis

 Springer

# Nonlinear Approaches in Engineering Applications



Liming Dai • Reza N. Jazar  
Editors

# Nonlinear Approaches in Engineering Applications

Applied Mechanics, Vibration Control,  
and Numerical Analysis

 Springer



*Editors*

Liming Dai  
Department of Industrial  
Systems Engineering  
University of Regina  
Regina, SK, Canada

Reza N. Jazar  
School of Aerospace, Mechanical  
and Manufacturing Engineering  
RMIT University Bundoora East Campus  
Bundoora, VIC, Australia

ISBN 978-3-319-09461-8

ISBN 978-3-319-09462-5 (eBook)

DOI 10.1007/978-3-319-09462-5

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014953220

© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

Dedicated to *Xinming* and *Mojgan*.



# Preface

This book follows the same concept of the first two volumes and is based on the ASME 2013 and 2012 Congress in the track of Dynamics Systems and Control, Optimal Approaches in Nonlinear Dynamics, which were organized by the editors. The nonlinear approaches and techniques have been developed since decades ago. The research on nonlinear science and dynamics have brought new insights into our modeling methods of natural and engineering phenomena. Although many of these approaches and techniques have been brought into research and engineering practices, linearization and simplification are still the dominating approaches existing in physics and engineering. Nature is nonlinear in general as the responses of physics and engineering systems are nonlinear. Linearization ends up with simplification, and it damages the original characteristics of the systems. Such simplifications usually lead to inaccuracy, misunderstanding, or even incorrect conclusions. For example, Hooke's law including the generalized Hooke's law is linear and it composes the foundation of linear elasticity and dominates numerous solutions of physical systems and mechanical designs. However, no material is perfectly linear. Any material used in the real world can actually be a nonlinear and complex system, not only due to its material or structure nonlinearity but also due to the inhomogeneous and anisotropy of the materials.

Another challenge facing the scientists and engineers in our time is the generation of the solutions and characterization of the nonlinear systems modeled from the physical systems in reality. It would be greatly beneficial in accurately evaluating the behavior of nonlinear systems and revealing the actual nature of the systems, with utilization of the existing mathematical tools and analytical means, if the analytical solutions of nonlinear systems could be pursued. Due to the nonlinearity and complexity of the nonlinear systems, unfortunately, it is very difficult or impossible to derive the analytical and closed-loop solutions for the systems. In solving or simulating the nonlinear systems, one may have to rely on approximate or numerical methods, which may only provide approximate results for the systems while errors are unavoidable during the processes of generating the approximate results.

Approximation and inaccuracy are the inescapable shadows following the current research and engineering practices involving nonlinearity or nonlinear systems.

In the role of the editors as well as the chapter contributors of this book, we have tried to present a collection of chapters showing the theoretically and practically sound nonlinear approaches and their engineering applications in various areas, in hoping that this book may provide useful tools and comprehensible examples of solving, modeling, and simulating the nonlinear systems existing in the real world. The carefully selected chapters contained in the present book reflect recent advances in nonlinear approaches and their engineering applications. The book intends to feature in particular the fundamental concepts and approaches of nonlinear science and their applications in engineering and physics fields. It is anticipated that this book may help to promote the development of nonlinear science and nonlinear dynamics in engineering, as well as to stimulate research and applications of nonlinear science and nonlinear dynamics in physics and engineering practices. It is also expected that the book will further enhance the comprehension of nonlinear science and stimulate interactions among scientists and engineers who are interested in nonlinear science and who find that nonlinearity and complexity of systems play an important role in their respective fields.

With the theme of the book, nonlinear approaches and engineering applications: Applied Mechanics, Vibration Control and Numerical Analysis, the book covers interdisciplinary studies on theories and methods of nonlinear science and their applications in complex systems such as those in nonlinear dynamics, vehicle dynamics, rigid body, solid mechanics, safety, visco-elastic mechanics, control engineering, ocean engineering, mechatronic engineering, acoustic engineering, and material science. Examples include: steady-state dynamic analysis of vehicles; rigid body dynamics and Razi acceleration discussion; a deep discussion about piecewise linear vibration isolators and challenges in discovering its time and steady-state determination; active vibration control for axially translation cables to limit their vibrating reactions; a new numerical solution method for ordinary differential equations where other methods fail; visco-elastic materials analysis and determination of their loss tangent; investigation on the frontal human-vehicle impact and re-designing the hood of the car to save pedestrians; a new treatment and analysis of the post-bukling of thin tube structures; study in acoustics of exhaust system to reduce their noise; parametric segmentation of nonlinear structures in image data; prediction and analysis of sea level.

## **Level of the Book**

This book aims at engineers, scientists, researchers, engineering and physics students of graduate levels, together with the interested individuals in engineering, physics, and mathematics. This chapter-book focuses on application of the nonlinear approaches representing a wide spectrum of disciplines of engineering and science. Throughout the book, great emphases are placed on engineering applications,

physical meaning of the nonlinear systems, and methodologies of the approaches in analyzing and solving for the systems. Topics that have been selected are of high interest in engineering and physics. An attempt has been made to expose the engineers and researchers to a broad range of practical topics and approaches. The topics contained in the present book are of specific interest to engineers who are seeking expertise in nonlinear analysis, mathematical modeling of complex systems, optimization of nonlinear systems, non-classical engineering problems, and future of engineering.

The primary audience of this book is the researchers, graduate students and engineers in mechanical engineering, engineering mechanics, electrical engineering, civil engineering, aerospace engineering, ocean engineering, mathematics, and science disciplines. In particular, the book can be used as a book for the graduate students as well as senior undergraduate students to enhance their knowledge by taking a graduate or advanced undergraduate course in the areas of nonlinear science, dynamics and vibration of discrete and continuous system, structure dynamics, and engineering applications of nonlinear science. It can also be utilized as a guide to the readers' fulfilment in practices. The covered topics are also of interest to engineers who are seeking to expand their expertise in these areas.

## **Organization of the Book**

The main structure of the book consists of two parts of analytical and practical nonlinearity, including 12 chapters. Each of the chapters covers an independent topic along the line of nonlinear approach and engineering applications of nonlinear science. The main concepts in nonlinear science and engineering applications are explained fully with necessary derivatives in detail. The book and each of the chapters is intended to be organized as essentially self-contained. All necessary concepts, proofs, mathematical background, solutions, methodologies and references are supplied except for some fundamental knowledge well known in the general fields of engineering and physics. The readers may therefore gain the main concepts of each chapter with as less as possible the need to refer to the concepts of the other chapters. Readers may hence start to read one or more chapters of the book for their own interests.

## **Method of Presentation**

The scope of each chapter is clearly outlined and the governing equations are derived with an adequate explanation of the procedures. The covered topics are logically and completely presented without unnecessary overemphasis. The topics are presented in a book form rather than in the style of a handbook. Tables, charts, equations, and references are used in abundance. Proofs and derivations are emphasized in such a

way that they can be straightforwardly followed by the readers with fundamental knowledge of engineering science and university physics. The physical model and final results provided in the chapters are accompanied with necessary illustrations and interpretations. Specific information that is required in carrying out the detailed theoretical concepts and modeling processes has been stressed.

## **Prerequisites**

The present book is primarily intended for researchers, engineers, and graduate students, so the assumption is that the readers are familiar with the fundamentals of dynamics, calculus, and differential equations associated with dynamics in engineering and physics, as well as a basic knowledge of linear algebra and numerical methods. The presented topics are given in a way to establish as conceptual framework that enables the readers to pursue further advances in the field. Although the governing equations and modeling methodologies will be derived with adequate explanations of the procedures, it is assumed that the readers have a working knowledge of dynamics, university mathematics and physics together with theory of linear elasticity.

## **Acknowledgements**

This book is made available under the close and effective collaborations of all the enthusiastic chapter contributors who have the expertise and experience in various disciplines of nonlinear science and engineering applications. They deserve sincere gratitude for the motivation of creating such book, encouragement in completing the book, scientific and professional attitude in constructing each of the chapters of the book, and the continuous efforts toward improving the quality of the book. Without the collaboration and consistent efforts of the chapter contributors, the completion of this book would have been impossible. What we have at the end is a book that we have every reason to be proud of.

It has been gratifying to work with the staff of Springer-Verlag through the development of this book. The assistance provided by the staff members has been valuable and efficient. We thank Springer-Verlag for their production of an elegant book.

Regina, SK, Canada  
Bundoora, VIC, Australia

Liming Dai  
Reza N. Jazar

# Contents

## Part I Analytical Nonlinearity

<b>1</b>	<b>Steady-State Vehicle Dynamics</b> .....	3
	Hormoz Marzbani, Reza N. Jazar, and M. Fard	
<b>2</b>	<b>On the Razi Acceleration</b> .....	31
	A. Salahuddin M. Harithuddin, Pavel M. Trivailo, and Reza N. Jazar	
<b>3</b>	<b>Challenges in Exact Response of Piecewise Linear Vibration Isolator</b> .....	59
	Oleksandr Pogorilyi, Reza N. Jazar, and Pavel M. Trivailo	
<b>4</b>	<b>Active Vibration Control for Nonlinear Axially Translating Cable Systems of Multi-Dimensions</b> .....	81
	L. Dai and L. Sun	
<b>5</b>	<b>Nonlinear Initial Value Ordinary Differential Equations</b> .....	117
	Mohammad M. Aghdam, Ali Fallah, and Poorya Haghi	
<b>6</b>	<b>The Loss Tangent of Visco-Elastic Models</b> .....	137
	Franz Konstantin Fuss	

## Part II Practical Nonlinearity

<b>7</b>	<b>Optimization of Hood Design to Minimise Pedestrian Head Injury in Impact</b> .....	161
	Revathi Krishnamoorthy, Monir Takla, Aleksandar Subic, and Derek Scott	
<b>8</b>	<b>Nonlinear Finite Element and Post-Buckling of Large Diameter Thin Walled Tubes</b> .....	219
	Toh Yen Pang, Dale Waterson, Rees Veltjens, and Tristan Garcia	



**9 Exhaust System Acoustic Modeling** ..... 235  
Milan Simic

**10 Nonlinear Approaches in Three Dimensional Medical  
Image Registration** ..... 251  
Ruwan Tennakoon, Alireza Bab-Hadiashar, and Zhenwei Cao

**11 Pose Estimation of a Dioptric Imaging Sensor with a  
Circle-Projecting Collimated Laser Moving Inside a Pipeline** ..... 281  
Abbasali Dehghan Tezerjani, Mehran Mehrandezh,  
and Raman Paranjape

**12 Analysis of Tide Gauge Sea Level Time Series** ..... 299  
Albert Parker

**Index**..... 317

# List of Figures

Fig. 1.1	Kinematics of a moving vehicle at steady-state conditions .....	4
Fig. 1.2	The steady-state responses as functions of the velocity $v_x$ .....	7
Fig. 1.3	The steady-state values of $R$ , $\beta$ , and $r$ , as functions of forward velocities .....	7
Fig. 1.4	The coordinate $(x_O, y_O)$ of the dynamic rotation center of the vehicle for $0 \leq v_x \leq 40$ m/s and its critical speed .....	9
Fig. 1.5	The location of the steady-state rotation center for $0 \leq v_x \leq 40$ m/s .....	9
Fig. 1.6	A realistic step change of steer angle .....	10
Fig. 1.7	The instantaneous time history of $r$ , $v_y$ , and $R$ for a step steering change .....	11
Fig. 1.8	The steady-state time history of $r$ , $v_y$ , and $R$ for a step steering change .....	12
Fig. 1.9	The difference between the instantaneous and steady-state values of $r$ during a step steering input .....	12
Fig. 1.10	The difference between the instantaneous and steady-state values of $v_y$ during a step steering input .....	12
Fig. 1.11	The difference between the instantaneous and steady-state values of $R$ during a step steering input .....	13
Fig. 1.12	The difference between the instantaneous and steady-state values of $\psi$ during a step steering input .....	13
Fig. 1.13	The instantaneous path of motion of the vehicle and the steady-state rotation center .....	15
Fig. 1.14	The instantaneous and steady-state global velocity components of the vehicle and their differences .....	16
Fig. 1.15	The instantaneous and steady-state paths of motion of the vehicle .....	17
Fig. 1.16	Plot of $X$ , $X_s$ , $Y$ , $Y_s$ , as functions of $t$ .....	17
Fig. 1.17	Plot of $X - X_s$ , $Y - Y_s$ , as functions of $t$ .....	18
Fig. 1.18	Plot of $x - x_s$ , as functions of $t$ .....	18

Fig. 1.19	Plot of $y - y_s$ , as functions of $t$ .....	18
Fig. 1.20	Cycloidal Motion lane change maneuver—moving 3 m from one lane to the adjacent lane .....	19
Fig. 1.21	Radius of curvature for Cycloidal Motion Lane Change Maneuver Road.....	20
Fig. 1.22	Radius of curvature for Cycloidal Motion Lane Change Maneuver Road.....	20
Fig. 1.23	A front wheel steering vehicle in a right turn .....	21
Fig. 1.24	Illustration of a car which is moving on a road at a point that $C_c$ is the center of curvature .....	21
Fig. 1.25	The kinematic steering angles— <i>Black</i> : $\delta_i$ and <i>Red</i> : $\delta_o$ .....	22
Fig. 1.26	The average of the kinematic steering angles .....	24
Fig. 1.27	The side slip velocity $v_y$ of the vehicle for the kinematic steering case.....	24
Fig. 1.28	The rate of change of the heading angle $r(t)$ of the vehicle for the kinematic steering case .....	25
Fig. 1.29	The $x$ -component of the resulted motion expressed in body frame for the kinematic case.....	25
Fig. 1.30	The $y$ -component of the resulted motion expressed in body frame for the kinematic case.....	25
Fig. 1.31	Resulted road as seen in the body frame for the kinematic case .....	26
Fig. 1.32	The $X$ -component of velocity in the global frame for the kinematic case .....	26
Fig. 1.33	The $Y$ -component of velocity in the global frame for the kinematic case .....	26
Fig. 1.34	Kinematic steering road .....	26
Fig. 1.35	The kinematic steering angles vs. steady-state steering angles, <i>red</i> : kinematic— <i>green</i> : steady-state .....	27
Fig. 1.36	The difference between kinematic and steady-state steering angles .....	27
Fig. 1.37	The resulted road in the global frame .....	27
Fig. 1.38	Actual road ( <i>red</i> ) vs. kinematic steering road ( <i>black</i> ) .....	27
Fig. 1.39	Actual road ( <i>red</i> ) vs. steady-state road ( <i>green</i> ) .....	28
Fig. 1.40	The difference between the steady-state and kinematic angles—high speed maneuver .....	28
Fig. 1.41	Actual road ( <i>red</i> ) vs. steady-state road ( <i>black</i> )—high speed maneuver.....	28
Fig. 1.42	Actual road ( <i>red</i> ) vs. steady-state road ( <i>green</i> )—high speed manoeuvre .....	28
Fig. 2.1	Three coordinate frames system .....	32
Fig. 2.2	Binary coordinate frames system—a global-fixed frame $G$ and a rotating body frame $B$ .....	35
Fig. 2.3	Three relatively rotating coordinate frames .....	38

Fig. 2.4 Multiple coordinate frames system with frame 0 as the inertial frame ..... 39

Fig. 2.5 Describing compound rotation motion using three coordinate frames system ..... 50

Fig. 2.6 Razi acceleration vector ..... 51

Fig. 2.7  ${}^B_A\omega_B = \dot{\phi} = 5 \text{ rad/s}; {}^C_B\omega_C = \dot{\theta} = 1 \text{ rad/s}$  ..... 51

Fig. 2.8  ${}^B_A\omega_B = \dot{\phi} = 4 \text{ rad/s}; {}^C_B\omega_C = \dot{\theta} = 2 \text{ rad/s}$  ..... 52

Fig. 2.9  ${}^B_A\omega_B = \dot{\phi} = 3 \text{ rad/s}; {}^C_B\omega_C = \dot{\theta} = 3 \text{ rad/s}$  ..... 52

Fig. 2.10  ${}^B_A\omega_B = \dot{\phi} = 2 \text{ rad/s}; {}^C_B\omega_C = \dot{\theta} = 4 \text{ rad/s}$  ..... 53

Fig. 2.11  ${}^B_A\omega_B = \dot{\phi} = 1 \text{ rad/s}; {}^C_B\omega_C = \dot{\theta} = 5 \text{ rad/s}$  ..... 53

Fig. 2.12  ${}^B_A\omega_B = \dot{\phi} = 5 \text{ rad/s}; {}^C_B\omega_C = \dot{\theta} = 0 \text{ rad/s}$  ..... 54

Fig. 3.1 A dual rate bilinear vibrating system ..... 60

Fig. 3.2 A vibrating system with gap-damper engagement ..... 60

Fig. 3.3 Mechanical model of the piecewise linear system with symmetric constraints ..... 61

Fig. 3.4 A steady state periodic response of the system ..... 63

Fig. 3.5 The frequency response for  $Z = \frac{U}{\Delta}$  ..... 68

Fig. 3.6 The frequency response curve of the primary for  $\omega_1 = 1, \delta = 1, \xi_1 = 0.1$  ..... 69

Fig. 3.7 The implicit plot of the two transcendental equations (3.46) and (3.47) as functions of  $t_1$  and  $\varphi$  for  $\omega_2 = \sqrt{2}, \xi_2 = 0.1, \omega_1 = 1, w = 0.5, \xi_1 = 0.1$  ..... 70

Fig. 3.8 A sample of the steady response of the system for  $\omega_2 = \sqrt{2}, \xi_2 = 0.1, \omega_1 = 1, w = 0.5, \xi_1 = 0.1$  ..... 70

Fig. 3.9 The implicit plot of the Eqs. (3.46) and (3.47) for  $\omega = 1, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{17}, \xi_2 = 0.1$  ..... 71

Fig. 3.10 The steady response of the system for  $\varphi = -1.248353815, t_1 = .9049692487$  ..... 72

Fig. 3.11 The steady response of the system for  $\varphi = -1.314128509, t_1 = 2.229699384$  ..... 72

Fig. 3.12 The implicit plot of the Eqs. (3.46) and (3.47) for  $\omega = 1, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{2}, \xi_2 = 0.1$  ..... 73

Fig. 3.13 The steady response of the system for  $\omega = 1, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{2}, \xi_2 = 0.1$  ..... 73

Fig. 3.14 The implicit plot of the Eqs. (3.46) and (3.47) for  $\omega = 1.4, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{17}, \xi_2 = 0.1$  ..... 74

Fig. 3.15 The steady response of the system for  $\omega = 1, w = 3$  and  $\omega_1 = 1.4, \xi_1 = 0.1, \omega_2 = \sqrt{2}, \xi_2 = 0.1$  ..... 75

Fig. 3.16 One set of solution of Eqs. (3.46) and (3.47) for  $w = 3, \omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{17}, \xi_2 = 0.1$  and varying  $\omega$  ..... 75

Fig. 3.17 The mathematical model of an asymmetric piecewise linear vibration isolator ..... 76

Fig. 3.18	The mathematical model of a one-sided piecewise linear vibration isolator .....	77
Fig. 4.1	The model of the axially translating cable .....	85
Fig. 4.2	The wave diagram of $w_p$ without the application of the control strategy .....	95
Fig. 4.3	(a) The contribution of the first vibration mode to $w_p$ without the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ without the application of the control strategy. (c) The wave diagram of $w_3$ without the application of the control strategy .....	96
Fig. 4.4	The wave diagram of $w_p$ with the application of the control strategy .....	97
Fig. 4.5	(a) The contribution of the first vibration mode to $w_p$ with the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ with the application of the control strategy. (c) The contribution of the third vibration mode to $w_p$ with the application of the control strategy .....	98
Fig. 4.6	The wave diagram of $w_p$ (the <i>blue continuous line</i> ) and the reference signal $w_r$ (the <i>green dash line</i> ) .....	99
Fig. 4.7	The control input .....	99
Fig. 4.8	The wave diagram of $w_p$ with the application of the control strategy .....	99
Fig. 4.9	(a) The contribution of the first vibration mode to $w_p$ with the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ with the application of the control strategy. (c) The contribution of the third vibration mode to $w_p$ with the application of the control strategy .....	100
Fig. 4.10	The wave diagram of $w_p$ (the <i>blue continuous line</i> ) and the reference signal $w_r$ (the <i>green dash line</i> ) .....	101
Fig. 4.11	The control input .....	102
Fig. 4.12	The wave diagram of $w_p$ with the application of the control strategy .....	102
Fig. 4.13	(a) The contribution of the first vibration mode to $w_p$ with the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ with the application of the control strategy. (c) The contribution of the third vibration mode to $w_p$ with the application of the control strategy .....	103
Fig. 4.14	The wave diagram of $w_p$ (the <i>blue continuous line</i> ) and the reference signal $w_r$ (the <i>green dash line</i> ) .....	104
Fig. 4.15	The control input .....	104

Fig. 4.16	The wave diagram of $w_p$ with the application of the control strategy .....	105
Fig. 4.17	(a) The contribution of the first vibration mode to $w_p$ with the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ with the application of the control strategy. (c) The contribution of the third vibration mode to $w_p$ with the application of the control strategy .....	106
Fig. 4.18	The wave diagram of $w_p$ (the <i>blue continuous line</i> ) and the reference signal $w_r$ (the <i>green dash line</i> ) .....	107
Fig. 4.19	The control input .....	107
Fig. 4.20	The wave diagram of $w_p$ with the application of the control strategy .....	108
Fig. 4.21	(a) The contribution of the first vibration mode to $w_p$ with the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ with the application of the control strategy. (c) The contribution of the third vibration mode to $w_p$ with the application of the control strategy .....	109
Fig. 4.22	The wave diagram of $w_p$ (the <i>blue continuous line</i> ) and the reference signal $w_r$ (the <i>green dash line</i> ) .....	110
Fig. 4.23	The control input .....	110
Fig. 4.24	The wave diagram of $w_p$ with the application of the control strategy .....	111
Fig. 4.25	(a) The contribution of the first vibration mode to $w_p$ with the application of the control strategy. (b) The contribution of the second vibration mode to $w_p$ with the application of the control strategy. (c) The contribution of the third vibration mode to $w_p$ with the application of the control strategy .....	112
Fig. 4.26	The wave diagram of $w_p$ (the <i>blue continuous line</i> ) and the reference signal $w_r$ (the <i>green dash line</i> ) .....	113
Fig. 4.27	The control input .....	113
Fig. 5.1	Comparison between Euler method and second-order Runge–Kutta method .....	119
Fig. 5.2	Domain of k-step method .....	119
Fig. 5.3	Bernstein polynomials (a) quadratic $n = 2$ , (b) cubic $n = 3$ .....	122
Fig. 5.4	Comparing Newton interpolation with Bézier curves and Convex hull property .....	123
Fig. 5.5	Nonlinear mass–spring–damper system .....	128
Fig. 5.6	Solution of nonlinear mass–spring system with damping for (a) $h = 0.1$ , (b) $h = 0.3$ , (c) $h = 0.5$ .....	130
Fig. 5.7	Solution of Lotka–Volterra equation for (a) $h = 0.1$ , (b) $h = 0.3$ , (c) $h = 0.5$ .....	131

Fig. 5.8 Solution of Van der Pol equation in case of low damping  $\mu = 0.5$  for (a)  $h = 0.1$ , (b)  $h = 0.2$ , (c)  $h = 0.4$  ..... 133

Fig. 5.9 Solution of Van der Pol equation in case of high damping  $\mu = 5$  for (a)  $h = 0.001$ , (b)  $h = 0.005$ , (c)  $h = 0.05$ ..... 134

Fig. 6.1 Sinusoidal strain curve (*blue*) imposed on a visco-elastic material and resulting stress curve (*pink*);  $\sigma_0$ : stress amplitude, maximal stress;  $\varepsilon_0$ : strain amplitude, maximal strain;  $\delta$ : phase shift ..... 139

Fig. 6.2 Hysteresis loop of stress–strain ellipse;  $\sigma_0$ : maximal stress;  $\varepsilon_0$ : maximal strain;  $\delta$ : phase angle ..... 140

Fig. 6.3 Stress–time curve (*red*) during the first load cycles; *blue curve*: transient component; *green curve*: steady state component; the stress–time curve (*red*) is the sum of transient and steady state components ..... 140

Fig. 6.4 Standard linear solid of Voight form;  $\sigma$ : stress;  $\varepsilon$ : strain;  $\eta$ : viscosity constant;  $E_1$ : modulus of series spring;  $E_2$ : modulus of parallel spring;  $\varepsilon_1$ : strain of series spring;  $\varepsilon_2$ : strain of Kelvin–Voight model ..... 141

Fig. 6.5 Maximal stress  $\sigma_0$  against modulus  $E_2$  of parallel spring;  $f$ : frequency;  $\eta$ : viscosity constant;  $E_2$ :  $E_2$  at which a local minimum of  $\sigma_0$  is found ..... 145

Fig. 6.6 Phase angle  $\delta$  against frequency  $f$ ;  $\eta$ : viscosity constant;  $E_1$  and  $E_2$ : moduli of series and parallel spring;  $f_0$ :  $f$  at which a local maximum of  $\delta$  is found ..... 146

Fig. 6.7 Peak stress  $\sigma_0$  against frequency  $f$ ;  $\eta$ : viscosity constant;  $E$ : velocity independent elasticity parameter ..... 152

Fig. 6.8 Ratio of  $E/\eta$  against frequency  $f$ ;  $\eta$ : viscosity constant;  $E$ : velocity independent elasticity parameter;  $\tan \delta$ : loss tangent;  $f_0$ : frequency at which  $\tan \delta = \pi/2$  ..... 154

Fig. 6.9 Phase angle  $\delta$  against log frequency;  $\eta$ : viscosity constant;  $E$ : velocity independent elasticity parameter ..... 154

Fig. 7.1 Unsafe conditions for pedestrians (World Health Organization 2011) ..... 162

Fig. 7.2 Kinematics of a pedestrian in PPCFC (Synaptic analysis consulting group inc) ..... 163

Fig. 7.3 Direction of impact for pedestrian fatalities due to collision with a passenger car in the US (National highway traffic safety administration 2009)..... 164

Fig. 7.4 Pedestrian protection technical requirements (Euro NCAP 2010) ..... 167

Fig. 7.5 Subsystem test method (Hardy et al. 2007) ..... 168

Fig. 7.6 Comparison between TRL-PLI and Flex-PLI (World Health Organization 2010) ..... 168

Fig. 7.7	TRL-PLI lower leg impactor .....	169
Fig. 7.8	Leg impact and head impact test machines (Centre of automotive safety research 2007) .....	169
Fig. 7.9	POLAR III pedestrian dummy (Honda Worldwide 2008) .....	171
Fig. 7.10	Comparison of peak acceleration magnitude .....	172
Fig. 7.11	Comparison of peak acceleration duration .....	172
Fig. 7.12	Primary impact of head with hood panels with excessive stiffness .....	173
Fig. 7.13	Effect of inertial loading of the head form on HIC value .....	174
Fig. 7.14	Secondary impact of head with hard components within engine bay .....	174
Fig. 7.15	Waveform with secondary impact .....	174
Fig. 7.16	Waveform efficiency ( $\eta$ ) of a half-sine waveform (Lim et al. 1995) .....	176
Fig. 7.17	Realistic optimal waveform for pedestrian head impact (Wu and Beaudet 2007) .....	178
Fig. 7.18	Underside view of a hood assembly .....	179
Fig. 7.19	Inlaid hood (A2Mac1 2010) .....	179
Fig. 7.20	Wraparound hood (A2Mac1 2010) .....	180
Fig. 7.21	FEA model for optimisation .....	182
Fig. 7.22	FEA model with head impactor for optimisation .....	182
Fig. 7.23	FEA model simulation using LS-DYNA for optimisation .....	182
Fig. 7.24	Contour plot of hood displacement using HYPERVIEW .....	183
Fig. 7.25	Plot of hood displacement vs. time using HYPERGRAPH .....	183
Fig. 7.26	Plot of resultant acceleration vs. time using HYPERGRAPH .....	184
Fig. 7.27	Energy plot from the finite element model .....	184
Fig. 7.28	Pedestrian head impact finite element model .....	185
Fig. 7.29	Correlation between ANCAP test and FE simulation .....	186
Fig. 7.30	Correlation of hood deformation between ANCAP test and FE simulation .....	186
Fig. 7.31	Coordinates for intrusion computation .....	187
Fig. 7.32	Comparison of impactor intrusion .....	188
Fig. 7.33	Accelerometer rotation due to impactor rebound .....	188
Fig. 7.34	'Intrusion T' calculation method .....	189
Fig. 7.35	Design space for hood assembly .....	189
Fig. 7.36	Scope of the project .....	190
Fig. 7.37	Hood leading edge (United Nations Economic Commission for Europe 1998) .....	190
Fig. 7.38	Hood rear reference line (United Nations Economic Commission for Europe 1998) .....	191
Fig. 7.39	Hood side reference line (United Nations Economic Commission for Europe 1998) .....	191
Fig. 7.40	Head impact test area on hood .....	192
Fig. 7.41	Impact positions considered for optimisation .....	192
Fig. 7.42	Concepts for improving pedestrian protection performance .....	193



Fig. 7.43	Pugh matrix .....	194
Fig. 7.44	Inner hood geometries chosen for developing an optimal solution .....	194
Fig. 7.45	Control factors in scope for refinement .....	195
Fig. 7.46	Parameter diagram .....	196
Fig. 7.47	Predictability of response surface .....	197
Fig. 7.48	ANOVA chart for mean value of HIC .....	198
Fig. 7.49	Evaluation of efficiency .....	199
Fig. 7.50	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point A .....	200
Fig. 7.51	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point B .....	200
Fig. 7.52	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point C .....	201
Fig. 7.53	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point D .....	201
Fig. 7.54	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point E.....	202
Fig. 7.55	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point F.....	202
Fig. 7.56	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point G .....	203
Fig. 7.57	HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point H .....	203
Fig. 7.58	Efficiency and combined hood panel thickness relationship .....	204
Fig. 7.59	Influence of deformation space on HIC value (steel) .....	205
Fig. 7.60	Head impact performance with 80 mm deformation space.....	206
Fig. 7.61	Influence of hood panel thickness on HIC value (aluminium) ....	206
Fig. 7.62	Influence of combined hood panel thickness on HIC value (aluminium) .....	207
Fig. 7.63	Combined hood panel thickness vs. the sum of the mean value of HIC and twice its standard deviation (aluminium) .....	207
Fig. 7.64	Combined thickness of aluminium outer and inner hood panels in various vehicles .....	208
Fig. 7.65	Influence of deformation space on HIC value (aluminium) .....	208
Fig. 7.66	Waveform comparison-70 mm deformation space.....	209

Fig. 7.67 Waveform comparison aluminium with 85 mm deformation space ..... 210

Fig. 7.68 ANOVA chart for peak acceleration ..... 211

Fig. 7.69 Resultant acceleration against vertical intrusion ..... 211

Fig. 7.70 Comparison of rebound velocities ..... 212

Fig. 8.1 Thin walled cylindrical tube used for the experiment ..... 221

Fig. 8.2 Experimental setup of test specimen in the load press..... 221

Fig. 8.3 Images of experiment demonstrating mode of failure ..... 222

Fig. 8.4 Force vs. displacement graph of experiment ..... 222

Fig. 8.5 Stress–strain behaviour of material damage where softening occurs, as well as degradation of material..... 224

Fig. 8.6 Comparison of force–displacement plots for numerical and experimental analyses ..... 229

Fig. 8.7 Deformation of aluminium tubing using ductility and MSFLD damage criteria for Aluminium 6060 T5 ..... 229

Fig. 8.8 Close-up deformation of aluminium tubing using ductility and MSFLD damage criteria for Aluminium 6060 T5 ..... 230

Fig. 9.1 An exhaust system structure ..... 236

Fig. 9.2 Power spectrum measurement program ..... 240

Fig. 9.3 Exhaust system sound power spectrum ..... 241

Fig. 9.4 Basic acoustic subsystem ..... 242

Fig. 9.5 Gas temperature distribution along the system ..... 244

Fig. 9.6 A segment from the acoustic ladder network ..... 245

Fig. 9.7 Distributed parameters calculations ..... 245

Fig. 9.8 Header and its equivalent physical network: Flow sources  $F(t)$  are represented by arrows, acoustic inductors as  $L_a$  and capacitors as  $C_a$ . Equivalent circuit is given on the right where equivalent source flow is  $4F(t)$  and equivalent components are as given in the text and on the figure ..... 246

Fig. 9.9 Simulink model of the cascaded acoustic filter ..... 246

Fig. 9.10 Bode plots of the exhaust system ..... 247

Fig. 10.1 In this illustration the moving image needs to be both translated and rotated to align with the fixed image ..... 252

Fig. 10.2 (a) CT, (b), MRI, (c) PET images of the same subject. Images are from the RIRE dataset (<http://www.insight-journal.org/rire>) ..... 253

Fig. 10.3 Registration flow of a typical intensity information based registration algorithm ..... 255

Fig. 10.4 Joint histograms between the original image (*top-left*) and rotated versions of it. (*Top-right*) No rotation, (*bottom-left*) rotated by  $2^\circ$ , (*bottom-right*) rotated by  $15^\circ$  ..... 259

Fig. 10.5 Inverse transformation vs. forward transformation. The figure shows that using inverse transform is computationally advantageous. (a) Inverse transform. (b) Forward transform ..... 263

Fig. 10.6 Original image (c) and the image after an affine transform (d). The transformation of both the grid (a, b) and the intensities are shown ..... 264

Fig. 10.7 B-Spline kernels ..... 267

Fig. 10.8 Original image and the image after a B-Spline transform. The transformation of both the grid and the intensities are shown together with the control point locations ... 269

Fig. 11.1 Robotic pipe inspection system. (a) Robotic inspection system. (b) Imaging sensor mounted on the robot. (c) Snapshot of the robot inside a PVC pipe ..... 284

Fig. 11.2 Main components of the imaging unit..... 284

Fig. 11.3 Proposed imaging platform and its configuration inside the pipe ..... 285

Fig. 11.4 Projection geometry of the imaging platform ..... 286

Fig. 11.5 Projection of the laser ring on the image plane of an equidistant fish-eye lens for a sample camera pose and configuration  $f = 0.14$  cm,  $t_x = -6$ ,  $t_y = -4$ ,  $\alpha_x = \pi/6$ ,  $\alpha_y = \pi/8$ ,  $d = 5$  cm, and  $r_p = 20$  cm ..... 288

Fig. 11.6 Snapshot of developed GUI for generating sample data..... 291

Fig. 11.7 Observation and analytical reference signals during the optimization process, where the actual pose is  $Q_{actual} = \langle x_c = -1.2988 \ y_c = 5.6033 \ \theta_p = 0.2496 \ \theta_t = -0.2153 \ \rangle$ : (a) after 15 iterations, (b) after 50 iterations, (c) after 100 iterations, and (d) after 140 iterations ..... 292

Fig. 11.8 Performance of the pose estimation algorithm vs. number of iterations and initial seed value. (a) Pose estimation error vs. the number of iterations. (b) Success rate of the pose estimation algorithm vs. the offset value of the initial seed for the camera pose ..... 293

Fig. 11.9 Comparison of the actual and estimated camera pose for 60 consecutive frames of the camera moving in a spiral path. In this experiment the last estimated pose (last-known configuration) is used as the initial seed for the pose estimation on the next frame. (a) Trajectories in X-Y plane. (b) Trajectories in pan-tilt plane ..... 294

Fig. 11.10 Computation time and optimization error vs. number of iterations for two initial seeds with 40 and 60 % offset. (a) Computation time vs. number of iterations. (b) Value of the objective function  $J$  vs. number of iterations .... 294

Fig. 12.1 Glacial isostatic adjustment (from [http://en.wikipedia.org/wiki/Glacial\\_isostatic\\_adjustment](http://en.wikipedia.org/wiki/Glacial_isostatic_adjustment)) ..... 302

Fig. 12.2 Demography of PSMSL stations (from <http://www.psmsl.org/data/obtaining/>, updated: 22 August 2012): stations with at least 40 years of data (*top*) and stations from Northern and Southern Hemisphere (*bottom*) ..... 303

Fig. 12.3 Sea levels in Sydney (*left*) and Fremantle (*right*): (a) monthly sea levels; (b) linear trend; (c) oscillations of monthly sea levels about the linear trend; (d) land motion in nearby GPS stations ..... 305

Fig. 12.4 Sea levels in Sydney (*left*) and Fremantle (*right*): (a) time history of SLR computed with 20, 30 and 60 years windows; (b) time history of SLR computed by using all the data available at any time—this is the conventional sea level velocity; (c) time history of SLA computed as the time rate of change of the above SLR; (d) apparent present velocity that could have been computed with the linear model in case the measuring record could have started in different times ..... 309

Fig. 12.5 Sea levels in Sydney (*left*) and Fremantle (*right*): (a) time history of SLA computed with 20, 30 and 60 years windows; (b) time history of SLA computed by using all the data available at any time—this is the average sea level acceleration since the start of the record; (c) apparent present acceleration that could have been computed as twice the second order coefficient of a polynomial fitting in case the measuring record could have started in different times ..... 310

Fig. 12.6 Sea levels in Sydney (*left*) and Fremantle (*right*): (a) periodogram of the monthly oscillations about the linear trend; (b), (c) and (d) measured and computed monthly sea levels by the proposed multiple nonlinear fitting over different time windows—better accuracy may be obtained by using other periodic functions ..... 312

**Part I**  
**Analytical Nonlinearity**

# Chapter 1

## Steady-State Vehicle Dynamics

Hormoz Marzbani, Reza N. Jazar, and M. Fard

**Keywords** Vehicle dynamics • Autodriver algorithm • Autonomous vehicle  
Steering dynamics

### 1.1 Dynamic Equations of Motion

The equations of motion of the planar bicycle car model, expressed in the principal body coordinate frame  $B$ , are governed by the following set of nonlinear coupled ordinary differential equations. In the equations, the steering angle  $\delta$  acts as the input and the mass center forward velocity  $v_x$ , lateral velocity  $v_y$ , and yaw rate  $r$  act as the outputs (Jazar 2014).

$$\dot{v}_x = \frac{1}{m} F_x + r v_y \quad (1.1)$$

$$\dot{v}_y = \frac{C_\beta}{m v_x} v_y - \left( v_x - \frac{C_r}{m} \right) r + \frac{C_\delta}{m} \delta(t) \quad (1.2)$$

$$\dot{r} = \frac{D_\beta}{I_z v_x} v_y + \frac{D_r}{I_z} r + \frac{D_\delta}{I_z} \delta(t) \quad (1.3)$$

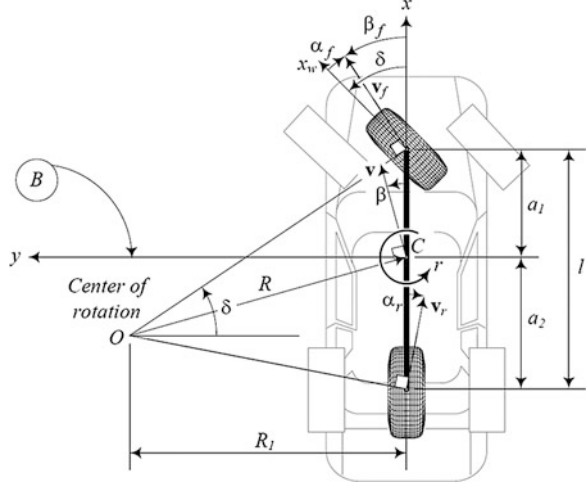
where

$$C_r = \frac{\partial F_y}{\partial r} = -\frac{a_1}{v_x} C_{\alpha f} + \frac{a_2}{v_x} C_{\alpha r} \quad (1.4)$$

---

H. Marzbani (✉) • R.N. Jazar • M. Fard  
School of Aerospace, Mechanical, and Manufacturing Engineering  
RMIT University, Melbourne, VIC, Australia  
e-mail: [hormoz.marzbani@rmit.edu.au](mailto:hormoz.marzbani@rmit.edu.au); [reza.jazar@rmit.edu.au](mailto:reza.jazar@rmit.edu.au);  
[mohamad.fard@rmit.edu.au](mailto:mohamad.fard@rmit.edu.au)

**Fig. 1.1** Kinematics of a moving vehicle at steady-state conditions



$$C_\beta = \frac{\partial F_y}{\partial \beta} = -C_{\alpha_f} - C_{\alpha_r} \quad (1.5)$$

$$C_\delta = \frac{\partial F_y}{\partial \delta} = C_{\alpha_f} \quad (1.6)$$

$$D_r = \frac{\partial M_z}{\partial r} = -\frac{a_1^2}{v_x} C_{\alpha_f} - \frac{a_2^2}{v_x} C_{\alpha_r} \quad (1.7)$$

$$D_\beta = \frac{\partial M_z}{\partial \beta} = a_2 C_{\alpha_r} - a_1 C_{\alpha_f} \quad (1.8)$$

$$D_\delta = \frac{\partial M_z}{\partial \delta} = a_1 C_{\alpha_f} \quad (1.9)$$

$$\beta = \frac{v_y}{v_x} \quad (1.10)$$

The orthogonal body coordinate frame  $B$  is fixed to the vehicle at its mass center  $C$  and is set such that the  $x$ -axis is longitudinal,  $y$ -axis is lateral, and  $z$ -axis is vertical upward. The body coordinate and kinematics of the bicycle car model in a forward motion on a positive turn is illustrated in Fig. 1.1.

The coefficients  $C_r$ ,  $C_\beta$ ,  $C_\delta$ ,  $D_r$ ,  $D_\beta$ ,  $D_\delta$  in the equations of motion are slopes of the curves for lateral force  $F_y$  and yaw moment  $M_z$  as a function of  $r$ ,  $\beta$ , and  $\delta$ , respectively. The coefficients  $C_r$  and  $D_r$  are functions of  $v_x$ , and the coefficients  $C_\beta$ ,  $C_\delta$ ,  $D_\beta$ ,  $D_\delta$  are constant for a given vehicle. The solution of the equations to a given time dependent steer angle  $\delta(t)$  would be

$$v_x = v_x(t) \quad (1.11)$$

$$v_y = v_y(t) \quad (1.12)$$

$$r = r(t) \quad (1.13)$$

The equations of motion are not analytically solvable, however, researchers traditionally assume that the forward velocity  $v_x$  is constant to reduce the equation to the following linear set.

Generally, this happens when the forward velocity is known and there is no need for calculating  $v_x$ .

$$F_x = -mr v_y \quad (1.14)$$

$$\begin{bmatrix} \dot{v}_y \\ \dot{r} \end{bmatrix} = \begin{bmatrix} \frac{C_\beta}{mv_x} & \frac{C_r}{m} - v_x \\ \frac{D_\beta}{I_z v_x} & \frac{D_r}{I_z} \end{bmatrix} \begin{bmatrix} v_y \\ r \end{bmatrix} + \begin{bmatrix} \frac{C_\delta}{m} \\ \frac{D_\delta}{I_z} \end{bmatrix} \delta(t) \quad (1.15)$$

When  $v_x$  remains constant or in more general cases is known, Eq. (1.15) become independent from (1.14). The set of Eq. (1.15) can then be written in the form

$$\dot{\mathbf{q}} = [A] \mathbf{q} + \mathbf{u} \quad (1.16)$$

in which  $[A]$  is a constant coefficient matrix,  $\mathbf{q}$  is the vector of control variables, and  $\mathbf{u}$  is the vector of inputs.

$$\begin{aligned} [A] &= \begin{bmatrix} \frac{C_\beta}{mv_x} & \frac{C_r}{m} - v_x \\ \frac{D_\beta}{I_z v_x} & \frac{D_r}{I_z} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{C_{\alpha f} + C_{\alpha r}}{mv_x} & -a_1 C_{\alpha f} + a_2 C_{\alpha r} - v_x \\ \frac{a_1 C_{\alpha f} - a_2 C_{\alpha r}}{I_z v_x} & -\frac{a_1^2 C_{\alpha f} + a_2^2 C_{\alpha r}}{I_z v_x} \end{bmatrix} \end{aligned} \quad (1.17)$$

$$\mathbf{q} = \begin{bmatrix} v_y \\ r \end{bmatrix} \quad (1.18)$$

$$\mathbf{u} = \begin{bmatrix} \frac{C_\delta}{m} \\ \frac{a_1 C_{\alpha f}}{I_z} \end{bmatrix} \delta(t) = \begin{bmatrix} \frac{C_\delta}{m} \\ \frac{D_\delta}{I_z} \end{bmatrix} \delta(t) \quad (1.19)$$



## 1.2 Steady-State Responses

When the vehicle is turning at a steady-state condition, its behavior is governed by the following equations:

$$F_x = -\frac{m}{R} v_x v_y \quad (1.20)$$

$$C_\beta \beta + (C_r v_x - m v_x^2) \frac{1}{R} = -C_\delta \delta \quad (1.21)$$

$$D_\beta \beta + D_r v_x \frac{1}{R} = -D_\delta \delta \quad (1.22)$$

from which, we can define a set of steady-state responses (Jazar 2013):

1. Curvature response,  $S_\kappa$

$$S_\kappa = \frac{\kappa}{\delta} = \frac{1}{R\delta} = \frac{C_\delta D_\beta - C_\beta D_\delta}{v_x (D_r C_\beta - C_r D_\beta + m v_x D_\beta)} \quad (1.23)$$

2. Sideslip response,  $S_\beta$

$$S_\beta = \frac{\beta}{\delta} = \frac{D_\delta (C_r - m v_x) - D_r C_\delta}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} \quad (1.24)$$

3. Yaw rate response,  $S_r$

$$S_r = \frac{r}{\delta} = \frac{\kappa}{\delta} v_x = S_\kappa v_x = \frac{C_\delta D_\beta - C_\beta D_\delta}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} \quad (1.25)$$

4. Centripetal acceleration response,  $S_a$

$$S_a = \frac{v_x^2/R}{\delta} = \frac{\kappa}{\delta} v_x^2 = S_\kappa v_x^2 = \frac{(C_\delta D_\beta - C_\beta D_\delta) v_x}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} \quad (1.26)$$

5. Lateral velocity response,  $S_y$

$$S_y = \frac{v_y}{\delta} = S_\beta v_x = \frac{D_\delta (C_r - m v_x) - D_r C_\delta}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} v_x \quad (1.27)$$

Figure 1.2 illustrates a sample of steady-state responses with respect to  $v_x$  for a car with the following characteristics.

$$\begin{aligned} C_{\alpha f} &= 50,000 \text{ N/rad} & C_{\alpha r} &= 50,000 \text{ N/rad} \\ m &= 1,000 \text{ kg} & I_z &= 1,650 \text{ kg m}^2 \\ a_1 &= 1.0 \text{ m} & a_2 &= 1.5 \text{ m} \end{aligned} \quad (1.28)$$

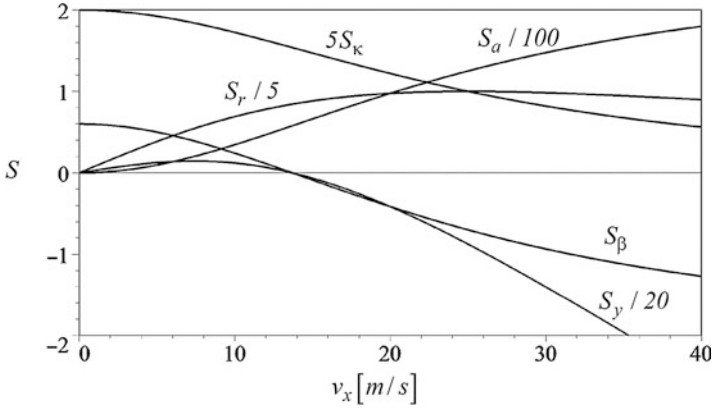


Fig. 1.2 The steady-state responses as functions of the velocity  $v_x$

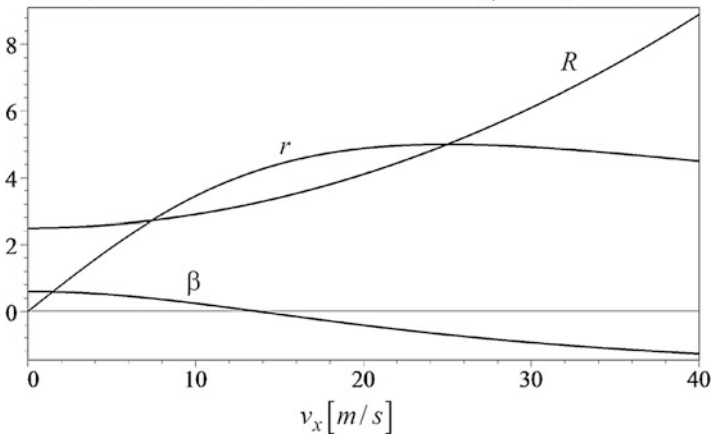


Fig. 1.3 The steady-state values of  $R$ ,  $\beta$ , and  $r$ , as functions of forward velocities

The vehicle is understeer because of a positive stability parameter  $K$ .

$$K = \frac{m}{l^2} \left( \frac{a_2}{C_{\alpha f}} - \frac{a_1}{C_{\alpha r}} \right) = 0.0016 > 0 \tag{1.29}$$

Employing the steady-state responses, we can determine the variation of the kinematics of motion as functions of forward velocity  $v_x$ . Figure 1.3 depicts the steady-state radius of rotation  $R$  of the mass center of the car, vehicle side slip angle  $\beta$ , and yaw rate  $r$ , at different forward velocities.

### 1.3 Rotation Center

Having the steady-state responses  $S_k = 1/R/\delta$  and  $S_\beta = \beta/\delta$ , we are able to determine the steady-state position of the rotation center  $(x_O, y_O)$  of a vehicle in the vehicle body coordinate frame as illustrated in Fig. 1.1.

$$x_O = -R \sin \beta = -\frac{1}{S_k \delta} \sin (S_\beta \delta) \quad (1.30)$$

$$y_O = R \cos \beta = \frac{1}{S_k \delta} \cos (S_\beta \delta) \quad (1.31)$$

At steady-state conditions the radius of rotation  $R$  can be found from the curvature response  $S_\kappa$ , and the vehicle sideslip angle  $\beta$  from the sideslip response  $S_\beta$ .

$$R = \frac{1}{\delta S_\kappa} = \frac{v_x (D_r C_\beta - C_r D_\beta + m v_x D_\beta)}{(C_\delta D_\beta - C_\beta D_\delta) \delta} \quad (1.32)$$

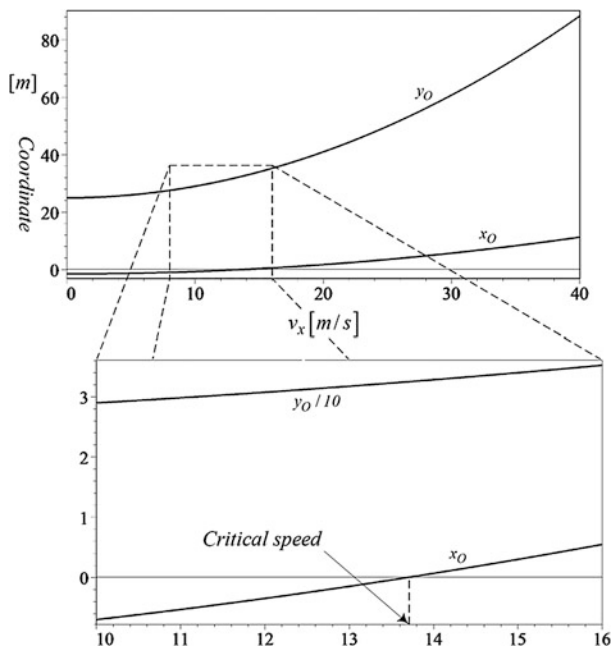
$$\beta = \delta S_\beta = \frac{D_\delta (C_r - m v_x) - D_r C_\delta}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} \delta \quad (1.33)$$

Therefore, the steady-state position of the dynamic center of rotation  $O$  in the body coordinate frame  $B$ , about which the vehicle will actually turn, is at

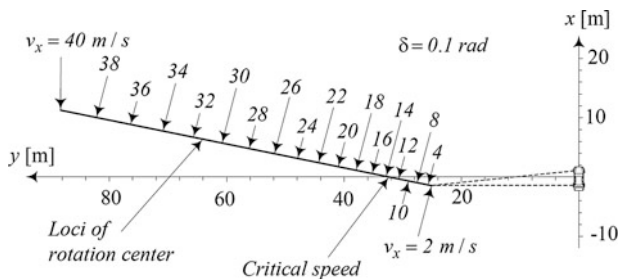
$$x_O = -\frac{v_x (D_r C_\beta - C_r D_\beta + m v_x D_\beta)}{(C_\delta D_\beta - C_\beta D_\delta) \delta} \sin \left( \frac{D_\delta (C_r - m v_x) - D_r C_\delta}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} \delta \right) \quad (1.34)$$

$$y_O = \frac{v_x (D_r C_\beta - C_r D_\beta + m v_x D_\beta)}{(C_\delta D_\beta - C_\beta D_\delta) \delta} \cos \left( \frac{D_\delta (C_r - m v_x) - D_r C_\delta}{D_r C_\beta - C_r D_\beta + m v_x D_\beta} \delta \right) \quad (1.35)$$

Figure 1.4 illustrates the coordinates  $x_O$  and  $y_O$  of the vehicle for  $0 \leq v_x \leq 40$  m/s at a constant steer angle  $\delta = 0.1$  rad. At a constant steer angle, the dynamic rotation center of the understeer vehicle moves away and forward by increasing the forward velocity. The rate of the displacement of the rotation center increases by speed. The figure also depicts the critical speed  $v_x = (C_r D_\delta - D_r C_\delta) / m / D_\delta = 13.693$  m/s at which the velocity vector of the vehicle at its mass center is in the  $x$ -direction  $\mathbf{v} = v_x \hat{i}$ . At the critical speed, we have  $\beta = 0$  and the dynamic rotation center is on the  $y$ -axis. Figure 1.5 illustrates the location of the steady-state rotation center for  $0 \leq v_x \leq 40$  m/s, relative to the vehicle.



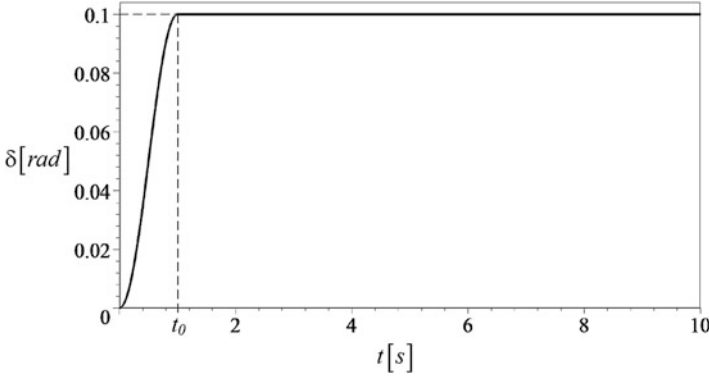
**Fig. 1.4** The coordinate  $(x_O, y_O)$  of the dynamic rotation center of the vehicle for  $0 \leq v_x \leq 40$  m/s and its critical speed



**Fig. 1.5** The location of the steady-state rotation center for  $0 \leq v_x \leq 40$  m/s

### 1.4 Effect of Changing Steering Angle on Transient and Steady-State Responses

In the following, we calculate the transition response of an understeer passenger car that is moving at a constant forward velocity, under variable steering angle. The transition behavior of the car will be determined by solving the equations of motion. The same transition behavior will be determined using steady-state responses to examine the proximity of the two analyses. It will be shown that steady-state response equations are good enough to predict the transition behavior of the car.



**Fig. 1.6** A realistic step change of steer angle

The results of this investigation will be useful in speeding the online calculation of autodrivers algorithm, as well as obtaining data to determine the steady-state behavior of vehicles (Tahami et al. 2003). The maneuvers can be considered as the steady state circular path which is a classical vehicle dynamic parameter determination test. The steady-state circular test is an open-loop test in which either the circular track radius or the steering angle or the vehicle's forward velocity must be kept constant (ISO4138 2012).

A realistic step change of steer angle can be expressed by

$$\delta = \delta_0 \left( H(t - t_0) + \sin^2 \left( \frac{t}{2t_0/\pi} \right) H(t_0 - t) \right) \quad (1.36)$$

as illustrated in Fig. 1.6 for  $t_0 = 1$  s, where  $H(t - t_0)$  is the Heaviside function

$$H(t - t_0) = \begin{cases} 0 & t < t_0 \\ 1 & t \geq t_0 \end{cases} \quad (1.37)$$

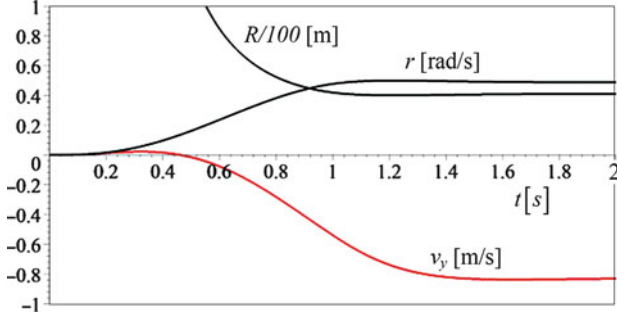
and  $t_0$  is the response time and  $\delta_0$  is the maximum constant steer angle.

We determine the reaction of the vehicle for

$$\delta_0 = 0.1 \text{ rad} \quad t_0 = 1 \text{ s} \quad v_x = 20 \text{ m/s} \quad (1.38)$$

and determine the solution of the equations of motion.

$$\begin{aligned} r = & H(t - t_0) e^{4.962(1-t)} (0.0588 \sin 3.7687(t - 1) - 0.01048 \cos 3.7687(t - 1)) \\ & + H(t - t_0) (0.233 \cos(\pi t) + 0.0869 \sin(\pi t) + 0.2439) \\ & + e^{-4.962t} (0.0588 \sin(\pi t) - 0.01048 \cos(\pi t)) \\ & - 0.233 \cos(\pi t) - 0.0869 \sin(\pi t) + 0.2439 \end{aligned} \quad (1.39)$$



**Fig. 1.7** The instantaneous time history of  $r$ ,  $v_y$ , and  $R$  for a step steering change

$$\begin{aligned}
 v_y = & H(t - t_0) e^{4.962(1-t)} (0.0492 \sin 3.7687(t - 1) + 0.2924 \cos 3.7687(t - 1)) \\
 & - H(t - t_0) (0.1222 \cos(\pi t) + 0.4028 \sin(\pi t) + 0.4146341465) \\
 & + e^{-4.962t} (0.2924 \cos(3.7687t) + 0.0492 \sin(3.7687t)) \\
 & + 0.1222 \cos(\pi t) + 0.4028 \sin(\pi t) - 0.414
 \end{aligned} \tag{1.40}$$

$$R = \frac{v_x}{r} \tag{1.41}$$

Let us call the dynamic variables that are calculated by solving the equations of motion (1.1)–(1.3) instantaneous variable and instantaneous vehicle, and the variables that are calculated based on responses (1.14)–(1.15) the steady-state variable and steady-state vehicle. Figure 1.7 depicts the instantaneous time history of  $r$ ,  $v_y$ , and  $R$ , during and after steering change. The steer angle is variable for  $0 < t \leq 1$  s and is fixed at  $\delta_0 = 0.1$  rad for  $t > 0$ . It takes less than one second for the vehicle to achieve the final steady-state conditions, while their states are not too far from the steady-states.

The steady-state expressions of  $r$ ,  $v_y$ , and  $R$  as functions of time for the step change of steer angle are:

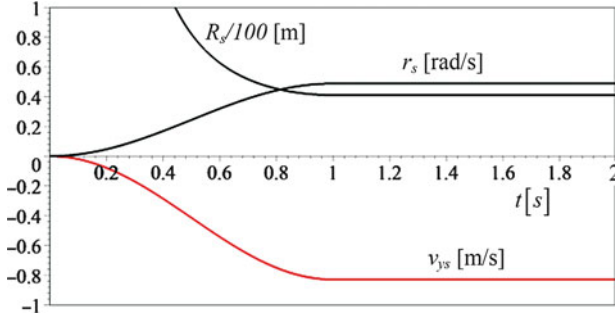
$$r_s = 0.4878 \sin^2 \left( \frac{t}{2/\pi} \right) H(1 - t) + 0.4878 H(t - 1) \tag{1.42}$$

$$v_{y_s} = -0.829 \sin^2 \left( \frac{t}{2/\pi} \right) H(1 - t) - 0.8292 H(t - 1) \tag{1.43}$$

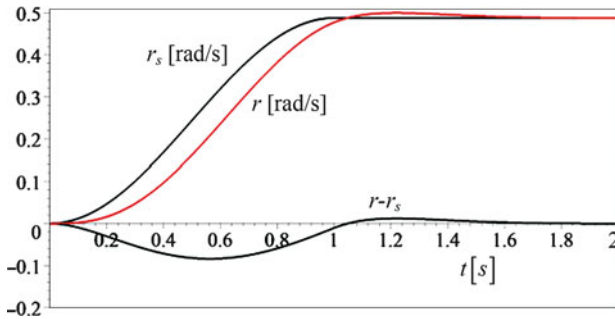
$$R_s = \frac{4.1}{0.1 \sin^2 \left( \frac{t}{2/\pi} \right) H(1 - t) + 0.1 H(t - 1)} \tag{1.44}$$

which are shown in Fig. 1.8.

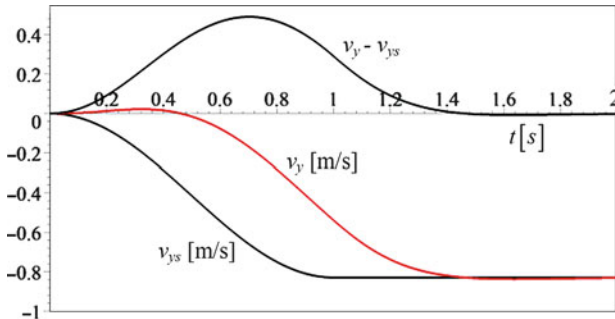
The difference between the instantaneous and steady-state values of  $r$ ,  $v_y$ , and  $R$  is shown in Figs. 1.9, 1.10, 1.11. The maximum difference of  $r - r_s$  happens around



**Fig. 1.8** The steady-state time history of  $r$ ,  $v_y$ , and  $R$  for a step steering change

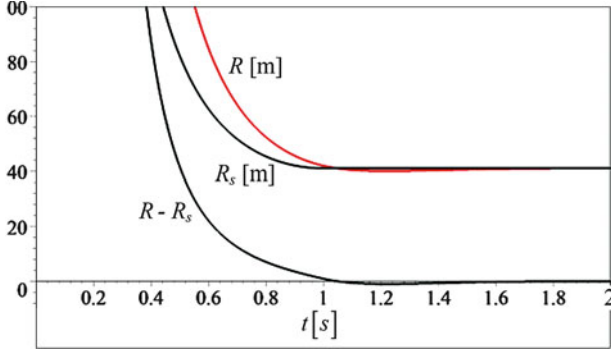


**Fig. 1.9** The difference between the instantaneous and steady-state values of  $r$  during a step steering input

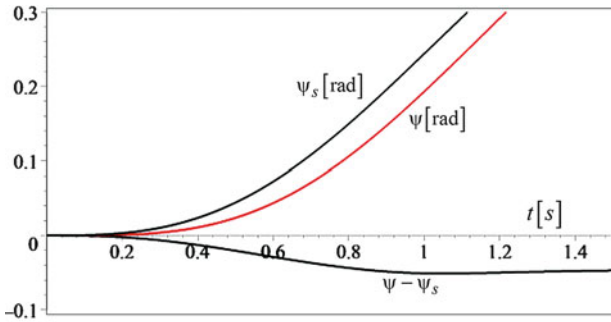


**Fig. 1.10** The difference between the instantaneous and steady-state values of  $v_y$  during a step steering input

middle of the steering change at  $t \approx 0.56$  s and is  $r - r_s \approx -0.083913$  rad/s which is 17.2% of the steady-state value of 0.4878 rad/s. As  $r < r_s$  during  $t < t_0$  the instantaneous yaw rate of the vehicle is higher than the steady-state, and therefore, the instantaneous vehicle is turning faster than the steady-state vehicle. This yields that the angular position of the steady-state vehicle has a constant lag.



**Fig. 1.11** The difference between the instantaneous and steady-state values of  $R$  during a step steering input



**Fig. 1.12** The difference between the instantaneous and steady-state values of  $\psi$  during a step steering input

The angular orientation of the body coordinate frame  $B(xyz)$  with respect to the global frame  $G(XYZ)$  is indicated by the angle  $\psi$  between axes  $X$  and  $x$  measured about the  $Z$ -axis. Figure 1.12 depicts  $\psi$ ,  $\psi_s$ , and  $\psi - \psi_s$  indicating that steady-state vehicle is a few degrees behind the instantaneous vehicle.

$$\psi = \int_0^t r(t) dt \tag{1.45}$$

$$\begin{aligned} &= H(t - t_0) (-e^{4.962(1-t)} (0.00435 \cos 3.769(t - 1) - 0.0085 \sin 3.769(t - 1))) \\ &+ H(t - t_0) (-0.0277 \cos(\pi t) + 0.074 \sin(\pi t) - 0.267) + 0.244t \\ &- e^{4.962(1-t)} (0.004356 \cos 3.769(t - 1) - 0.0085 \sin 3.769(t - 1)) \\ &+ 0.0277 \cos(\pi t) - 0.0743 \sin(\pi t) + 0.244t - 0.02332 \end{aligned} \tag{1.46}$$



The global frame  $G$  is fixed on the ground, and  $B$  coincides with  $G$  at the start of maneuvers. The  $B$ -frame moves with the vehicle while the axes  $z$  and  $Z$  are always parallel. Therefore, the velocity vector of the vehicle in the global frame is

$$\begin{aligned} {}^G \mathbf{v} &= {}^G R_B {}^B \mathbf{v} = \begin{bmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} \\ &= \begin{bmatrix} v_x \cos \psi - v_y \sin \psi \\ v_y \cos \psi + v_x \sin \psi \end{bmatrix} = \begin{bmatrix} v_X \\ v_Y \end{bmatrix} \end{aligned} \quad (1.47)$$

where  ${}^G R_B$  is the transformation between  $B$  and  $G$  and the velocity vector of the vehicle in the body frame is

$${}^B \mathbf{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad (1.48)$$

Therefore, the global coordinates of the mass center of the vehicle would be

$$X = \int_0^t v_X dt = \int_0^t (v_x \cos \psi - v_y \sin \psi) dt \quad (1.49)$$

$$Y = \int_0^t v_Y dt = \int_0^t (v_y \cos \psi + v_x \sin \psi) dt \quad (1.50)$$

When the steer angle is kept constant, the vehicle will eventually be turning on a constant circular path. The position of the steady-state rotation center of the vehicle in the  $B$ -frame is at

$$x_O = -R \sin \beta = 1.6994 \text{ m} \quad (1.51)$$

$$y_O = R \cos \beta = 40.965 \text{ m} \quad (1.52)$$

because

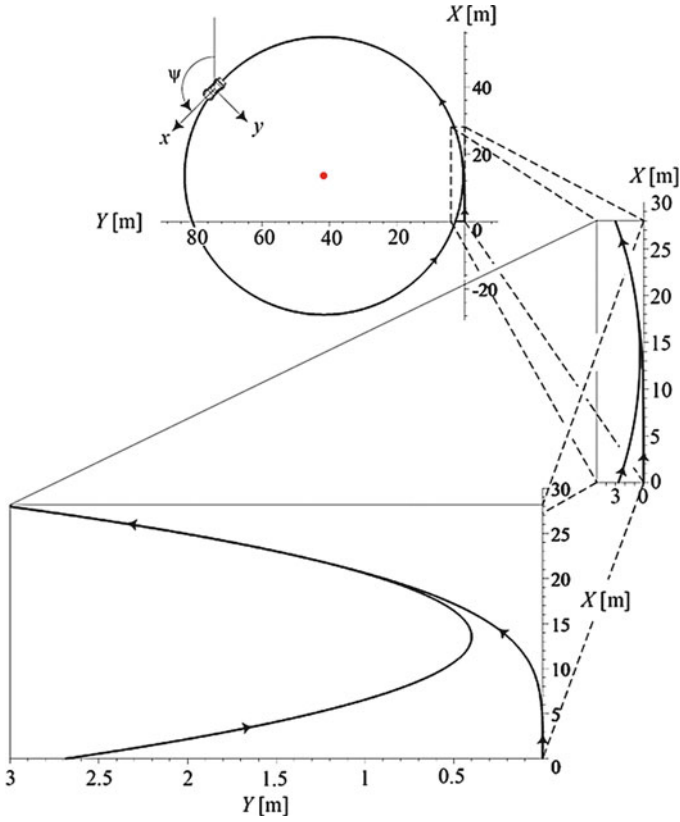
$$\lim_{t \rightarrow \infty} R = 41 \text{ m} \quad (1.53)$$

$$\lim_{t \rightarrow \infty} \beta = -0.04146 \text{ rad} \quad (1.54)$$

The global coordinates of the steady-state rotation center is

$$\begin{bmatrix} X_O \\ Y_O \end{bmatrix} = {}^G \mathbf{r}_B + {}^G R_B \begin{bmatrix} x_O \\ y_O \end{bmatrix} = \begin{bmatrix} 13.617 \text{ m} \\ 41.388 \text{ m} \end{bmatrix} \quad (1.55)$$

where  ${}^G \mathbf{r}_B$  is the  $G$ -expression of the position vector of the origin of the  $B$ -frame with respect to the origin of the  $G$ -frame at any point on the steady-state conditions (Jazar 2011). In this example, we used the calculated data at  $t = 2$  s when we have



**Fig. 1.13** The instantaneous path of motion of the vehicle and the steady-state rotation center

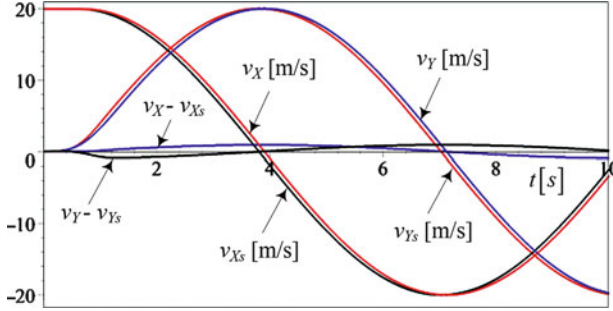
$$\begin{bmatrix} x_O \\ y_O \end{bmatrix} = \begin{bmatrix} 1.6994 \text{ m} \\ 40.965 \text{ m} \end{bmatrix} \tag{1.56}$$

$$\psi = 0.6851 \text{ rad} \tag{1.57}$$

$${}^G_G \mathbf{r}_B = \begin{bmatrix} 38.2226 \text{ m} \\ 8.5916 \text{ m} \end{bmatrix} \tag{1.58}$$

Figure 1.13 illustrates the instantaneous path of motion of the vehicle in global frame and its final steady-state rotation center. The magnification of the transient stage depicts how the vehicle approaches its steady-state circular path.

To compare the instantaneous and steady-state, or in other words, the exact and approximate maneuvers, we substitute  $v_{y_s}$  from (1.19) in (1.47) and calculate the steady-state velocity components of the vehicle in  $G$ -frame.



**Fig. 1.14** The instantaneous and steady-state global velocity components of the vehicle and their differences

$${}^G \mathbf{v}_s = {}^G R_B {}^B \mathbf{v}_s = \begin{bmatrix} v_x \cos \psi_s - v_{y_s} \sin \psi_s \\ v_{y_s} \cos \psi_s + v_x \sin \psi_s \end{bmatrix} = \begin{bmatrix} v_{X_s} \\ v_{Y_s} \end{bmatrix} \quad (1.59)$$

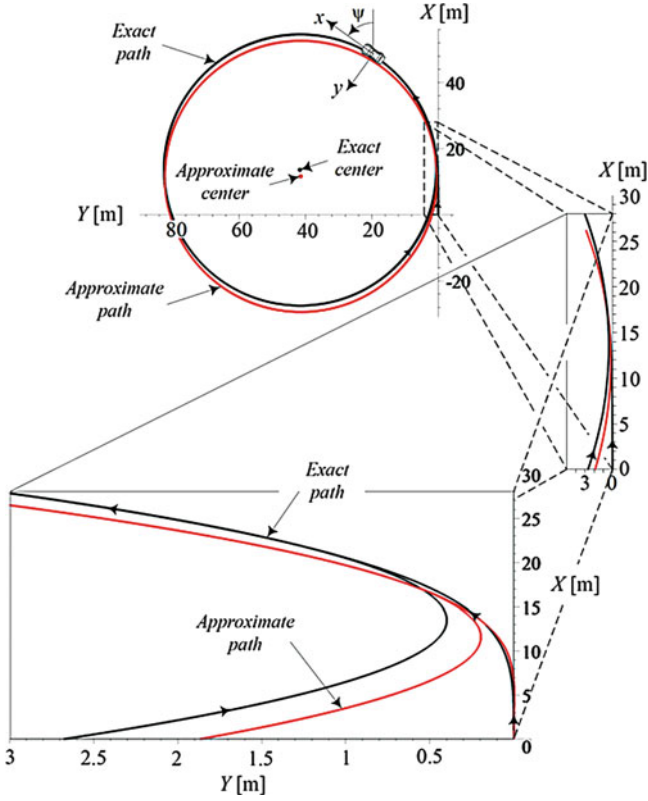
Figure 1.14 depicts the instantaneous and steady-state global velocity components of the vehicle and their differences.

Integration of  $v_{X_s}$  and  $v_{Y_s}$  provides us with the path of motion based on steady-state calculation. Figure 1.15 illustrates the instantaneous and steady-state paths of motion of the vehicle in the global frame. The magnification of the transient stage depicts how different the paths are. Figure 1.16 depicts  $X$ ,  $X_s$ ,  $Y$ ,  $Y_s$ , as functions of  $t$ , and Fig. 1.17 depicts  $X - X_s$ ,  $Y - Y_s$ . The maximum error of  $X - X_s$  is 3.808 m which is less than 6.89% of the maximum  $X$  of 54.522 m. The maximum error of  $Y - Y_s$  is also 2.116 m which is less than 2.59% of the maximum  $Y$  of 82.468 m. The errors in the local coordinate are much less, as are shown in Figs. 1.18 and 1.19.

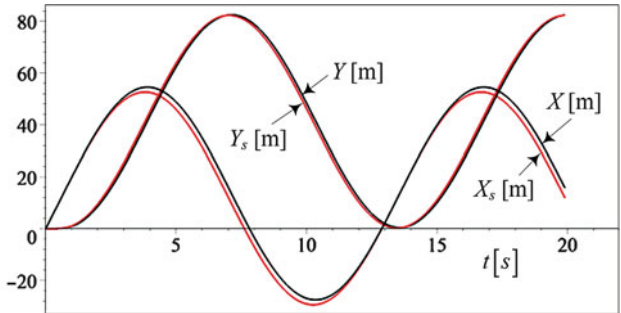
The difference between instantaneous and steady-state calculation decreases exponentially by increasing  $t_0$ . When  $t_0 = 2$  s the maximum error of  $X - X_s$  is 3.796 m which is less than 5.89% of the maximum  $X$  of 64.45 m and the maximum error of  $Y - Y_s$  is also 2.114 m which is less than 2.55% of the maximum  $Y$  of 83.154 m. When  $t_0 = 10$  s the maximum error of  $X - X_s$  is 3.188 m which is less than 2.33% of the maximum  $X$  of 136.886 m and the maximum error of  $Y - Y_s$  is also 0.0002 m which is less than 0.00025% of the maximum  $Y$  of 90.161 m.

## 1.5 Steady-State Dynamic Steering

According to the findings of the previous section the steady-state responses of a vehicle are in a good approximation range compared with the results of solving the dynamic equations of motion of the vehicle. This concept, initiated the idea of finding the steering angles required to keep a vehicle on a specific road by using the steady-state responses instead of the dynamic equations. The mentioned idea is tested in this section and is called the Steady-State Dynamic Steering (*SSDS*).

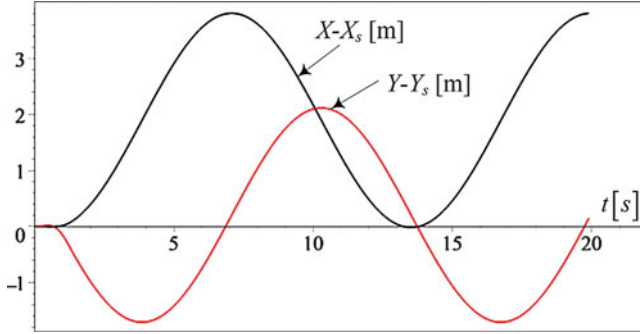


**Fig. 1.15** The instantaneous and steady-state paths of motion of the vehicle

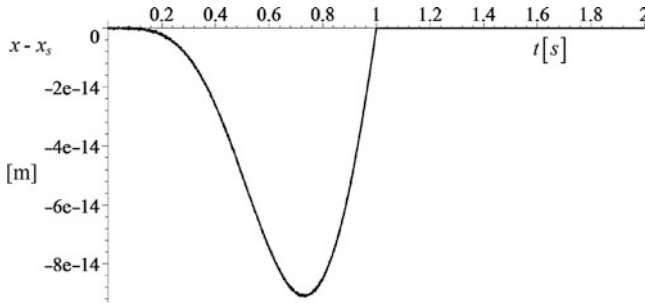


**Fig. 1.16** Plot of  $X$ ,  $X_s$ ,  $Y$ ,  $Y_s$ , as functions of  $t$

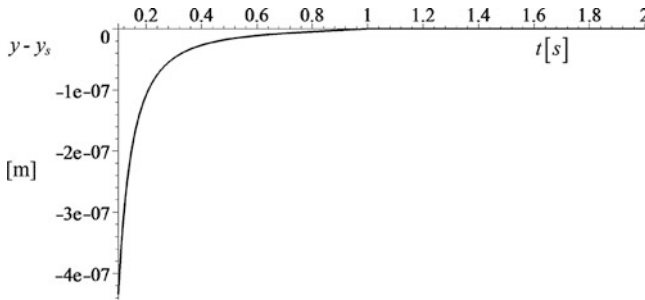
After examining the output–input relationships of the steady-state responses it was found that the curvature response of the vehicle could be used to find the steering angles of the vehicle.



**Fig. 1.17** Plot of  $X - X_s$ ,  $Y - Y_s$ , as functions of  $t$



**Fig. 1.18** Plot of  $x - x_s$ , as functions of  $t$

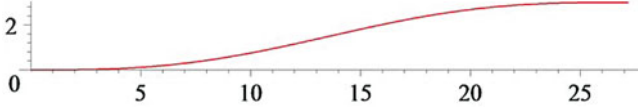


**Fig. 1.19** Plot of  $y - y_s$ , as functions of  $t$

$$s_k = \frac{k}{\delta} = \frac{1}{R\delta} \quad (1.60)$$

where  $R$  is the radius of curvature of the road,  $k = \frac{1}{R}$  is the curvature, and  $\delta$  is the steady-state steering angle of the vehicle.

To determine the road curvature center, it is necessary to find the radius of curvature of the road at any time during the travel. Now, consider having the



**Fig. 1.20** Cycloidal Motion lane change maneuver-moving 3 m from one lane to the adjacent lane

equation of a road which could be used to find the radius of curvature of the road. By replacing the radius of curvature found from the equation of the road in Eq. (1.60) we can find the steady-state steering angle of the vehicle and validate the findings by solving the equations of motion of the vehicle using the resulted steering angles. The equation for lane change maneuvering which will be examined is the following:

$$\begin{aligned} y(t) &= d\left(\frac{\alpha}{\pi}x(t) - \frac{1}{2\pi} \sin(2\alpha x(t))\right) \\ x(t) &= v_x t \end{aligned} \quad (1.61)$$

where  $d$  is the scaling factor and  $\alpha$  is

$$\alpha = \sqrt{\frac{2a_m}{d}} \quad (1.62)$$

where  $a_m$  is the maximum acceleration.

Equation (1.61) indicates a Cycloid motion. This produces a sine acceleration curve. The only problem with such a curve is the existence of jerk in the start of the motion, but the acceleration and velocity start and end at zero. Figure 1.20 shows the mentioned curve drawn using the following values:

$$a_m = 0.02\text{m/s}^2 \quad d = 3\text{m} \quad v_x = 1\text{m/s} \quad (1.63)$$

As mentioned above the equation for the steering angles can be easily calculated by the use of the curvature response, but in order to check the differences we'll solve the equations for both kinematic and steady-state situations.

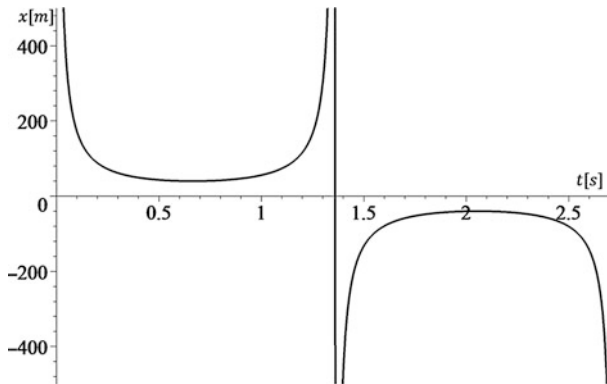
The forward velocity has been adjusted on 1m/s to start with. This will be changed to higher velocities to check the effect of the speed change as well.

The first step is to find the center of rotation of the road, and radius of curvature.

$$R = \frac{\sqrt{(\dot{x}^2 + \dot{y}^2)^3}}{\ddot{y}\dot{x} - \ddot{x}\dot{y}} = 0.78 \frac{\sqrt{(v_x^2 + 0.45d v_x^2 a_m (1 - \cos(2.82v_x t \sqrt{\frac{a_m}{d}}))^2)^3}}{\sin(2.82\sqrt{\frac{a_m}{d}} t v_x)} \quad (1.64)$$

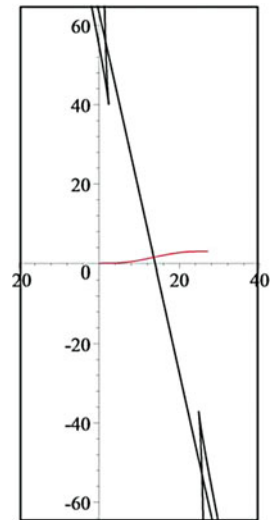
Radius of curvature from Eq. (1.64) is shown in Fig. 1.21.

Center of curvature of the road can be found using the radius of curvature and the unit normal vectors. The center of curvature of the road is shown in Fig. 1.22.



**Fig. 1.21** Radius of curvature for Cycloidal Motion Lane Change Maneuver Road

**Fig. 1.22** Radius of curvature for Cycloidal Motion Lane Change Maneuver Road

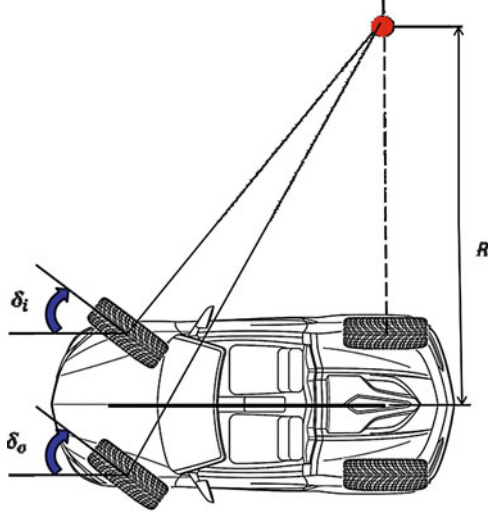


From Fig. 1.22 it can be seen that for the start of the motion the car is on a straight line which has the center of rotation in the positive side towards infinity. Then the radius of curvature gets closer to the road and starts getting far from the road again as the car gets closer to the center of the road. Right at the center the radius of curvature moves toward infinity once again and as soon as the side of the road curve changes it repeats the same motion backwards but this time on the negative side.

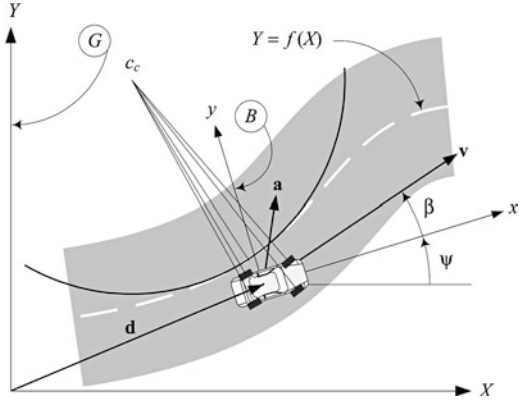
The steering angles of a front-wheel-steering vehicle can be found using the following two equations:

$$\delta_i = \arctan\left(\frac{l}{R_1 - \frac{w}{2}}\right) \tag{1.65}$$

**Fig. 1.23** A front wheel steering vehicle in a right turn



**Fig. 1.24** Illustration of a car which is moving on a road at a point that  $C_c$  is the center of curvature



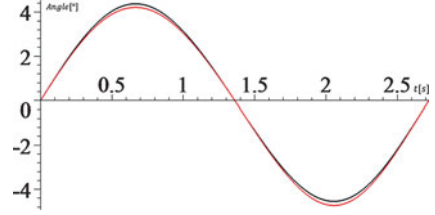
$$\delta_o = \arctan\left(\frac{l}{R_1 + \frac{w}{2}}\right) \tag{1.66}$$

where  $\delta_i$  is the angle for the wheel in the inside of the turn,  $\delta_o$  is the angle for the wheel in the outside of the turn,  $w$  is the width of the front of the car, and  $R_1$  is the vertical distance of the center of mass of the vehicle from the center of rotation of the car. These are illustrated in Fig. 1.23.

There is a global coordinate frame  $G$  attached to the ground, and a vehicle coordinate frame  $B$  attached to the car at its mass center  $C$ , as shown in Fig. 1.24. The  $z$  and  $Z$  axes are assumed to be parallel and the angle  $\psi$  indicates the orientation of  $B$  in  $G$ .  $\psi$  is the angle between  $x$  and  $X$  axes, and is called the heading angle of the car. If  $(X_c, Y_c)$  indicates the coordinates of the road curvature center  $c_c$  in the global coordinate frame  $G$ , then the coordinate of  $c_c$  in  $B$  would be



**Fig. 1.25** The kinematic steering angles—Black:  $\delta_i$  and Red:  $\delta_o$



$${}^B r_C = R_{z,\psi} ({}^G r_C - {}^G d) \quad (1.67)$$

$$\begin{aligned} \begin{bmatrix} x_c \\ y_c \\ 0 \end{bmatrix} &= \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \left( \begin{bmatrix} X_c \\ Y_c \\ 0 \end{bmatrix} - \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} (X_c - X) \cos \psi + (Y_c - Y) \sin \psi \\ (Y_c - Y) \cos \psi - (X_c - X) \sin \psi \\ 0 \end{bmatrix} \end{aligned} \quad (1.68)$$

When the equation of the road is given, Eq. (1.68) determines the point that the car should turn about. We use Eq. (1.68) to calculate the position of the curvature center and set the steer angles to have the same point as the turning center of the car. Having the coordinates of the road curvature center  $c_c$  in the vehicle coordinate frame is enough to determine the kinematic characteristic  $R_1$ .

$$R_1 = y_c = (Y_c - Y) \cos \psi - (X_c - X) \sin \psi \quad (1.69)$$

Figure 1.25 illustrates the outer and inner wheel's kinematic steering angle in order for the car to be kept on the road shown in Fig. 1.20. These are calculated using nominal values 1.5 m and 3 m for the width and total length of the vehicle, respectively.

Recalling from introduction, we can use the following equations of motion of the vehicle, and the kinematic steering angles found above to find out the actual road resulted from applying these steering angles, which we will call the kinematic road from now on.

$$\dot{v}_x = \frac{1}{m} F_x + r v_y \quad (1.70)$$

$$\dot{v}_y = \frac{C_\beta}{m v_x} v_y - \left( v_x - \frac{C_r}{m} \right) r + \frac{C_\delta}{m} \delta(t) \quad (1.71)$$

$$\dot{r} = \frac{D_\beta}{I_z v_x} v_y + \frac{D_r}{I_z} r + \frac{D_\delta}{I_z} \delta(t) \quad (1.72)$$

where

$$C_r = \frac{\partial F_y}{\partial r} = -\frac{a_1}{v_x} C_{\alpha f} + \frac{a_2}{v_x} C_{\alpha r} \quad (1.73)$$

$$C_\beta = \frac{\partial F_y}{\partial \beta} = -C_{\alpha f} - C_{\alpha r} \quad (1.74)$$

$$C_\delta = \frac{\partial F_y}{\partial \delta} = C_{\alpha f} \quad (1.75)$$

$$D_r = \frac{\partial M_z}{\partial r} = -\frac{a_1^2}{v_x} C_{\alpha f} - \frac{a_2^2}{v_x} C_{\alpha r} \quad (1.76)$$

$$D_\beta = \frac{\partial M_z}{\partial \beta} = a_2 C_{\alpha r} - a_1 C_{\alpha f} \quad (1.77)$$

$$D_\delta = \frac{\partial M_z}{\partial \delta} = a_1 C_{\alpha f} \quad (1.78)$$

$$\beta = \frac{v_y}{v_x} \quad (1.79)$$

The solution of the equations to a given time dependent steer angle  $\delta(t)$  would be

$$v_x = v_x(t) \quad (1.80)$$

$$v_y = v_y(t) \quad (1.81)$$

$$r = r(t) \quad (1.82)$$

Equations of motion of the vehicle are calculated using the following nominal values:

$$\begin{aligned} C_{\alpha f} &= 100,000 & C_{\alpha r} &= 100,000 & m &= 1,000 \text{ kg} \\ a_1 &= 1.2 \text{ m} & a_2 &= 1.8 \text{ m} & I_z &= 1,650 \text{ kg m}^2 \end{aligned} \quad (1.83)$$

and  $\delta(t)$  can be calculated by getting the average of  $\delta_i$  and  $\delta_o$  according to the following equation:

$$\delta(t) = \arctan \left( \frac{2 \tan \delta_i \tan \delta_o}{\tan \delta_i + \tan \delta_o} \right) \quad (1.84)$$

$\delta(t)$  is illustrated in Fig. 1.26.

The equations of motion will be solved for  $\delta(t)$  from Eq. (1.84) to find out what will be the final path taken by the vehicle having these steering angles. By solving the equations we will get to see the kinematic answers for  $v_y$  which indicates the lateral velocity, and  $r$  which is the rate of change of heading angle of the vehicle. These can be seen in Figs. 1.27, 1.28, 1.29, 1.30, 1.31, 1.32, 1.33.

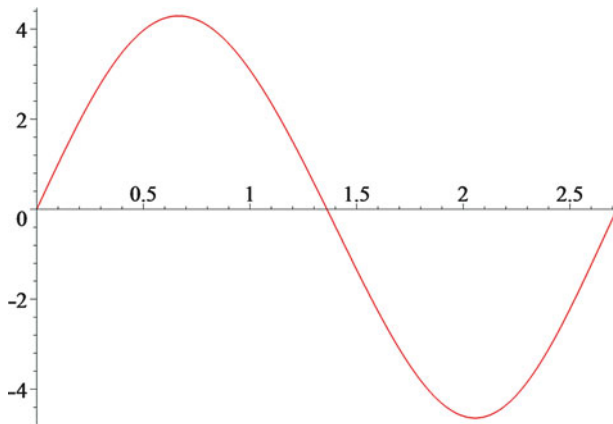


Fig. 1.26 The average of the kinematic steering angles

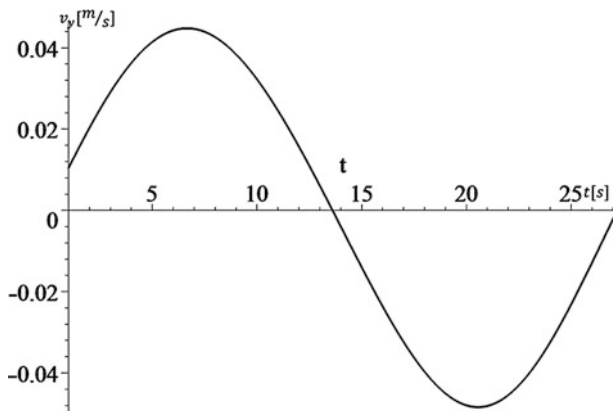


Fig. 1.27 The side slip velocity  $v_y$  of the vehicle for the kinematic steering case

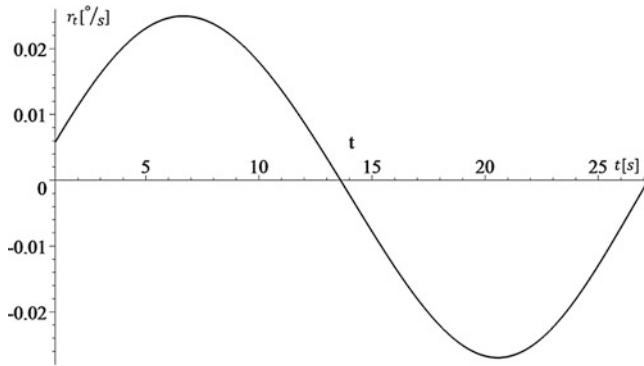
The resulted road from using the kinematic steering angles is shown in Fig. 1.34.

Now the problem will be solved using the new method based on the steady-state responses of the vehicle as mentioned above. More specifically the response used here is the curvature response of the vehicle.

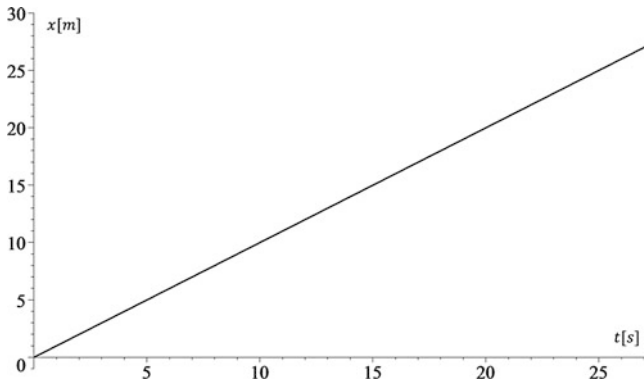
The curvature response  $s_k$  of the vehicle may be calculated using Eq. (1.85) and nominal values (1.83)

$$s_k = \frac{C_d D_b - C_b D_d}{v_x (D_r C_b - C_r D_b + m V_x D_b)} = \frac{1}{R_{steady} \delta(t)} = 0.3331112592 \quad (1.85)$$

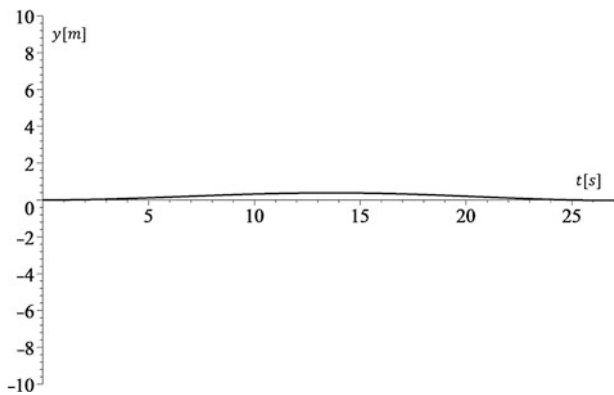
where  $R_{steady}$  is the radius of curvature using steady-state conditions, and  $\delta(t)$  is the new steering angle which will be found by putting the steady-state curvature radius equal to the radius of curvature of the road. This will result in the following equations:



**Fig. 1.28** The rate of change of the heading angle  $r(t)$  of the vehicle for the kinematic steering case



**Fig. 1.29** The  $x$ -component of the resulted motion expressed in body frame for the kinematic case



**Fig. 1.30** The  $y$ -component of the resulted motion expressed in body frame for the kinematic case

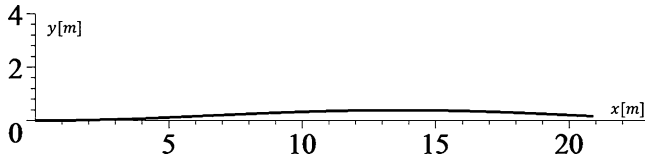


Fig. 1.31 Resulted road as seen in the body frame for the kinematic case

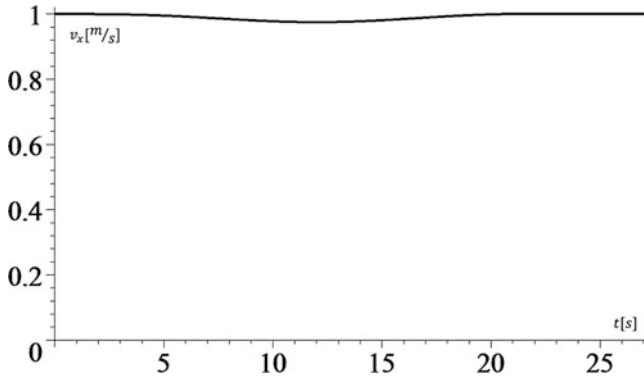


Fig. 1.32 The X-component of velocity in the global frame for the kinematic case

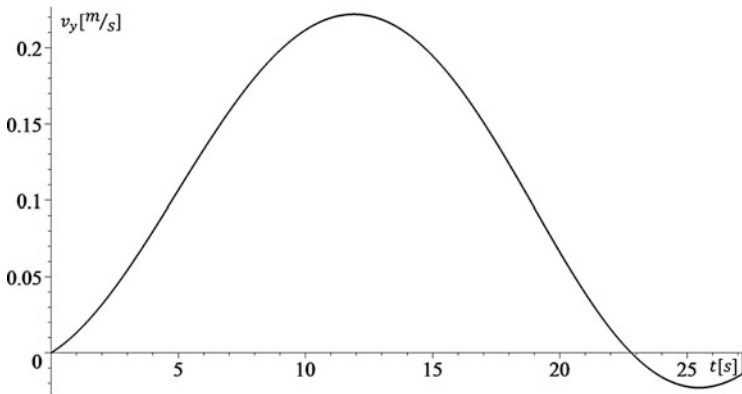


Fig. 1.33 The Y-component of velocity in the global frame for the kinematic case

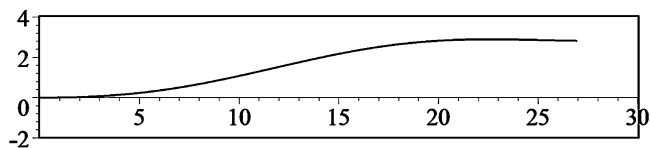
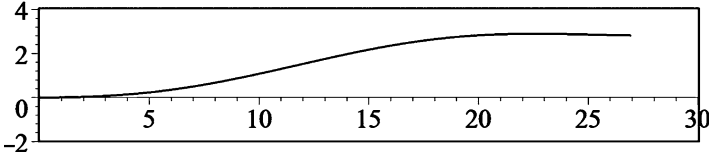
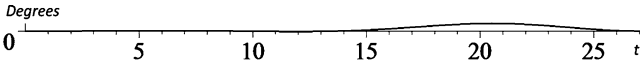


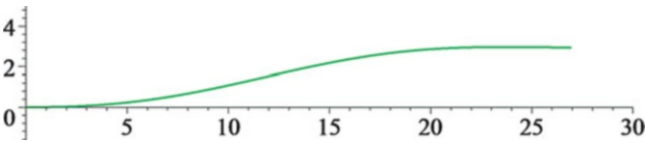
Fig. 1.34 Kinematic steering road



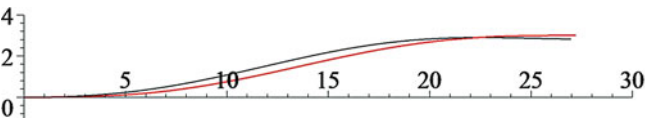
**Fig. 1.35** The kinematic steering angles vs. steady-state steering angles, *red*: kinematic—*green*: steady-state



**Fig. 1.36** The difference between kinematic and steady-state steering angles



**Fig. 1.37** The resulted road in the global frame



**Fig. 1.38** Actual road (*red*) vs. kinematic steering road (*black*)

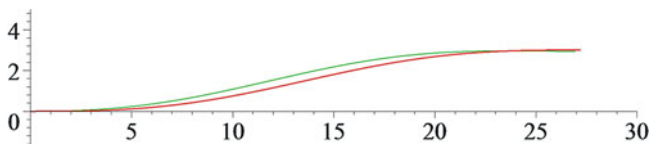
$$R_{steady} = \frac{3.002}{\delta(t)} \tag{1.86}$$

$$\delta(t) = \frac{0.7644530225 \sin(0.2309401076t)}{\sqrt{(1 + (0.1102657791 - 0.110265779 \cos(0.2309401076t))^2)^3}} \tag{1.87}$$

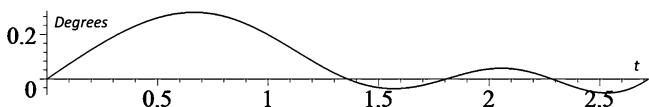
The equation for the steering angles has been found and illustrated in comparison with the kinematic steering angles in Fig. 1.35.

As it can be seen in both Figs. 1.35 and 1.36, the difference between the two is very small in case of a slow forward velocity ( $v_x = 1$ ), and it happens mostly while returning to the straight section of the road. In order to validate the new method and the steering angles found using this method, the equation of motion of the vehicle was solved once again and this time by substituting the new steering angles found using the steady-state responses of the vehicle. The road resulted using the steady-state dynamic angles is shown in Fig. 1.37. The figure proves to be very similar to the actual road but new plots will show how accurate the two methods are.

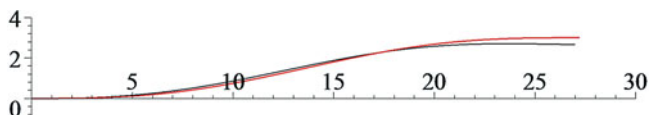
It can be seen from Figs. 1.38 and 1.39 that the road resulted from using the steady-state steering angles is generating less error and matches the actual road



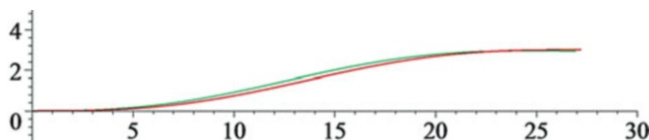
**Fig. 1.39** Actual road (*red*) vs. steady-state road (*green*)



**Fig. 1.40** The difference between the steady-state and kinematic angles—high speed maneuver



**Fig. 1.41** Actual road (*red*) vs. steady-state road (*black*)—high speed maneuver



**Fig. 1.42** Actual road (*red*) vs. steady-state road (*green*)—high speed manoeuvre

much better than the kinematic steering case. In order to take the testing of the new method further, same maneuvers will be repeated with a higher forward velocity equal to 10 m/s this time. The results are shown in Figs. 1.40 and 1.41.

Comparing the difference between the angles in this case with the previous case proves that the deviation is much more than before in the case of high speed maneuvers.

Figures 1.40 and 1.41 show that the major source of the bigger difference in the steering angles as a result of increasing the speed is the kinematic steering angles calculations. As seen in Fig. 1.42 the steering angles calculated using the steady-state dynamics are consistent after the change of speed but the deviation from the road gets much bigger for the kinematic steering case.

To have a closer look, the final position of the actual road will be compared with the final position of the vehicles traveling on the road using the two different methods which are being investigated. Although this might not be the place on which the maximum deviation happens at, it gives a good understanding of the results. On the other hand, from the figures it is obvious that all of these maneuvers will take the car in the desired direction but the main difference happens at the end of the road. The values of the end point of the actual road can be easily found by substituting the time at which the traveling ends which is equal to  $T = 2.72$  s in the case with the higher forward velocity case.

The position of the end point of the road is at  $(x, y) = (27.2, 3)$  m. This is shown in figures of the road. The car travels 27.2 m horizontally and this results in a 27.36 m total travel.

The final point of travel for the case of the kinematic steering angles is  $(x, y) = (26.98, 2.66)$  m which is a total travel of about 27.1 m. This point for the case of steady state steering angles is at  $(x, y) = (26.95, 2.94)$  m which is again 27.1 m of travel.

Looking back at the main purpose of the maneuver, which was changing the lane that the vehicle is moving in to the adjacent lane with 3 m of vertical displacement, one can come to the conclusion that the steady-state steering angles is providing a much more accurate and faster response. The displacement in the vertical direction in the steady-state case is only deviating about 1.6 % from the actual road which is very accurate. The amount of deviation in the other case, however, is about 11.3 % from the actual road.

## 1.6 Conclusion

The proximity of the transient and steady-state responses has been checked and validated by applying a step change input to the steering angle of a vehicle with constant velocity. This proved that if using the steady-state responses could help to solve any problem, they are close enough to be used instead of the transient responses of the vehicle. A unique method for finding steering angles for the use of the autodrivers algorithm was introduced using the curvature response of the vehicle. This method which is called Steady-State Dynamic Steering (*SSDS*) helps to find the angles which are accurate enough compared to the dynamic steering angles, with short calculations and for every kind of road equations. The angles found using the *SSDS* method were applied on different scenarios, which proved its high accuracy, less complication, and consistency in the resulted road. The use of this method will result in much shorter calculation times, less complicated and smaller steering angles equations, and a much smaller work load on the control feedback system as a result. Besides, the *SSDS* method proves to be consistent even in the case of completed road edition which made it impossible to solve the problem dynamically (Marzbani 2012; Jazar 2010; Pacejka 1973).

## 1.7 Key Symbols

$d$	Translation scalar
$\mathbf{d}$	Translation vector
$d_i$	Element $i$ of $\mathbf{d}$
$D$	Homogeneous translation matrix
${}^A D_B$	Homogeneous translation matrix from $B$ -frame to $A$ -frame
$E$	Total illumination



$E_r$	Distance-dependent illuminance
$E_\varphi$	Incident angle-dependent illuminance
$I, I_i$	Light intensity
$\hat{\mathbf{n}}$	Normal unit vector to the collector
$O$	Origin point of a coordinate frame
$r$	Distance between target and source of light, position scalar
$\mathbf{r}$	Position vector
$r_i$	Element $i$ of $\mathbf{r}$
$r_{ij}$	Element of row $i$ and column $j$ of matrix $R$
${}^A\mathbf{r}$	Vector $\mathbf{r}$ expressed in $A$ -frame
$\hat{\mathbf{r}}$	Direction of light ray with respect to the collector
$r, \theta, \phi$	Local spherical coordinates
$R$	Rotation transformation matrix
${}^A R_B$	Rotation transformation matrix from $B$ -frame to $A$ -frame
$R_{x,\alpha}$	Rotation transformation about $x$ -axis with $\alpha$ angle
$\mathbf{S}$	Sun vector
$T$	Homogeneous transformation matrix
${}^A T_B$	Homogeneous transformation matrix from $B$ -frame to $A$ -frame
$x, y, z$	Local coordinate axes
$X, Y, Z$	Global coordinate axes
$\alpha$	Yaw angle
$\beta$	Pitch angle
$\gamma$	Orientation of the vehicle body-frame in the global frame
$\theta$	Azimuth angle
$\phi$	Elevation angle
$\varphi$	Angle of incidence

## References

- Jazar RN (2013) *Vehicle dynamics: theory and applications*, 2nd edn. Springer, New York
- Jazar RN (2014) *Vehicle dynamics: theory and applications*, 2nd edn. Springer, New York
- Jazar RN (2011) *Advanced dynamics: rigid body, multibody, and aerospace application*. Wiley, New York
- Jazar RN (2010) Mathematical theory of autodrivers for autonomous vehicles. *J Vib Control* 16(2):253–279.
- Marzbani H (2012) Flat ride; problems and solutions in vehicle. *Nonlinear Eng* 1(3–4):101–108
- Pacejka, HB (1973) Simplified analysis of steady-state turning behaviour of motor vehicles. Part 1. Handling diagrams of simple systems. *Vehicle Syst Dyn* 2:161–172
- Tahami F, Kazemi R, Farhanghi, S (2003) A novel driver assist stability system for all-wheel drive electric vehicles. *IEEE Trans. Vehicular Technol* 52(3):683–692
- ISO4138 (2012) Passenger Cars Steady-state circular driving behaviour, *Open Loop test procedure, ISO 4138*

# Chapter 2

## On the Razi Acceleration

A. Salahuddin M. Harithuddin, Pavel M. Trivailo, and Reza N. Jazar

**Keywords** Razi acceleration • Rigid body kinematics • Multiple coordinate system kinematics • Coordinate transformation

### 2.1 Introduction

Jazar presented a revised method of calculating vector derivatives in multiple coordinate frames using the extended variation of the Euler derivative transformation formula (Jazar 2012). The paper derives the general expression for vector derivatives for a system of three relatively rotating coordinate frames called the mixed derivative transformation formula. The formula is demonstrated as follows: Let  $A$ ,  $B$ , and  $C$  be three arbitrary, relatively rotating coordinate frames sharing an origin as shown in Fig. 2.1. Let  ${}^B \square$  be a generic vector expressed in the frame  $B$ ,  ${}_A \omega_B$  be the angular velocity of the frame  $B$  with respect to the frame  $A$ , and  ${}_C \omega_B$  be the angular velocity of the frame  $B$  with respect to the frame  $C$ . The time-derivative of the  $B$ -vector as seen from the  $A$  frame can be expressed as

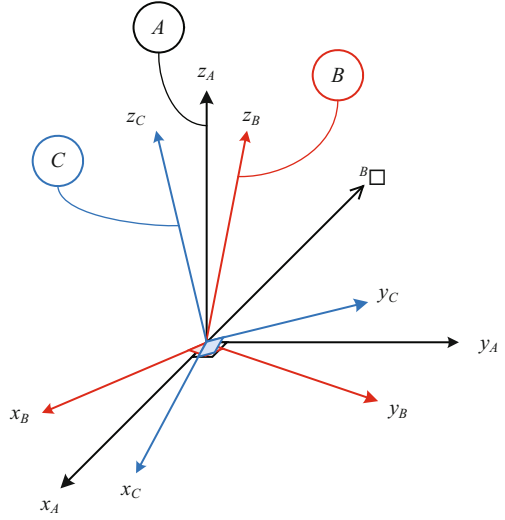
$$\frac{{}^A d}{{}^A dt} {}^B \square = \frac{{}^C d}{{}^C dt} {}^B \square + ({}^B \omega_B - {}^C \omega_B) \times {}^B \square \quad (2.1)$$

where the left superscript indicates the frame in which the vector is expressed. The formula relates the  $A$ -derivative of a  $B$ -vector to the  $C$ -derivative of the same  $B$ -vector by means of the angular velocity differences between the three

---

A.S.M. Harithuddin • P.M. Trivailo (✉) • R.N. Jazar  
School of Aerospace, Mechanical, and Manufacturing Engineering,  
RMIT University, Melbourne, VIC, Australia  
e-mail: [salah.harithuddin@student.rmit.edu.au](mailto:salah.harithuddin@student.rmit.edu.au); [pavel.trivailo@rmit.edu.au](mailto:pavel.trivailo@rmit.edu.au);  
[reza.jazar@rmit.edu.au](mailto:reza.jazar@rmit.edu.au)

**Fig. 2.1** Three coordinate frames system



coordinate frames. The formula presents a more general expression of the classical derivative transformation formula between two relatively rotating frames

$$\frac{A d}{dt} {}^B \square = \frac{B d}{dt} {}^B \square + {}^B \omega_B \times {}^B \square \tag{2.2}$$

where the frame  $C$  is assumed to coincide with  $B$  ( $C \equiv B$ ), thus  ${}^B \omega_B = 0$ . One advantage of the general mixed derivative transformation formula is that it leaves the need to compute the derivative of  ${}^B d {}^B \square / dt$  in its local coordinate frame.

An interesting result appears from the application of a reference frame system with three coordinate frames. A new acceleration term, given the name the Razi acceleration (Jazar 2011), is shown to appear when the derivatives of an acceleration vector are taken from two different coordinate frames. Such differentiation, which resulting expression is called *mixed* acceleration, contains the Razi term,  $({}^A \omega_C \times {}^B \omega_C) \times \mathbf{r}$ . To demonstrate the Razi acceleration term, let there be three relatively rotating coordinate frames  $A$ ,  $B$ , and  $C$ , shown in Fig. 2.1. Assume a position vector  $\mathbf{r}$  is expressed in its local frame which is  $C$ . Taking the first derivative of  $\mathbf{r}$  in  $C$  from  $B$  using the vector derivative transformation formula Eq. (2.2) yields

$$\frac{B d}{dt} {}^C \mathbf{r} = \frac{C d}{dt} {}^C \mathbf{r} + {}^C \omega_C \times {}^C \mathbf{r} \tag{2.3}$$

Taking the derivative of  $\frac{B d}{dt} {}^C \mathbf{r}$  again, but from the frame  $A$  yields:

$$\frac{A d}{dt} \frac{B d}{dt} {}^C \mathbf{r} = {}^C \mathbf{a}$$

$$\begin{aligned}
&= {}^C C \mathbf{a} + {}^C_A \boldsymbol{\omega}_C \times {}^C_C \mathbf{v} + {}^C_{CB} \boldsymbol{\alpha}_C \times {}^C_C \mathbf{r} + ({}^C_A \boldsymbol{\omega}_C \times {}^C_B \boldsymbol{\omega}_C) \times {}^C_C \mathbf{r} \\
&\quad + {}^C_B \boldsymbol{\omega}_C \times {}^C_C \mathbf{v} + {}^C_B \boldsymbol{\omega}_C \times ({}^C_A \boldsymbol{\omega}_C \times {}^C_C \mathbf{r})
\end{aligned} \tag{2.4}$$

where

$$\begin{aligned}
{}^C C \mathbf{a} &\equiv \frac{d^2}{dt^2} {}^C_C \mathbf{r} \\
{}^C_C \mathbf{v} &\equiv \frac{d}{dt} {}^C_C \mathbf{r} \\
{}^C_{CB} \boldsymbol{\alpha}_C &\equiv \frac{d}{dt} {}^C_B \boldsymbol{\omega}_C
\end{aligned} \tag{2.5}$$

and the left superscripts in (2.4) show that every vector is expressed in the frame  $C$ . The term  $({}^C_A \boldsymbol{\omega}_C \times {}^C_B \boldsymbol{\omega}_C) \times {}^C_C \mathbf{r}$  is a new term in the acceleration expression called the Razi acceleration.

For comparison, let us assume that the frame  $B$  is not rotating and coincides with the frame  $A$ , making  $B \equiv A$ . Equation (2.4) is now reduced to the classical expression of acceleration between two relatively rotating frames  $B$  and  $C$

$$\begin{aligned}
\frac{d^2}{dt^2} {}^C_C \mathbf{r} &= {}^C_{BB} \mathbf{a} = {}^C C \mathbf{a} + 2 {}^C_B \boldsymbol{\omega}_C \times {}^C_C \mathbf{v} + {}^C_{CB} \boldsymbol{\alpha}_C \times {}^C_C \mathbf{r} \\
&\quad + {}^C_B \boldsymbol{\omega}_C \times ({}^C_B \boldsymbol{\omega}_C \times {}^C_C \mathbf{r})
\end{aligned} \tag{2.6}$$

Equation (2.6) shows that there are three supplementary terms to the absolute acceleration of a point in local frame  $C$ : The Coriolis acceleration  $2 {}^C_B \boldsymbol{\omega}_C \times {}^C_C \mathbf{v}$ , the tangential acceleration  ${}^C_{CB} \boldsymbol{\alpha}_C \times {}^C_C \mathbf{r}$ , and the centripetal acceleration  ${}^C_B \boldsymbol{\omega}_C \times ({}^C_B \boldsymbol{\omega}_C \times {}^C_C \mathbf{r})$ . The difference between the expressions in Eqs. (2.4) and (2.6) is the Razi term which has the form of a vector product between a product of two angular velocities and a position. This shows that adding another rotating coordinate frame and differentiating a displacement vector from two different frames reveals a new acceleration term.

It is shown in the published work of Jazar that the Razi acceleration is a novel finding in classical mechanics. In its current status, however, the Razi acceleration has only found a limited application in relative kinematics. This paper seeks to reveal the aspect of the Razi acceleration in mechanics of relative rigid body motion specifically in the multiple coordinate system kinematics. We show the inclusion of the new Razi term in the equation of motion of a rigid body described in non-inertial rotating coordinate frames. To do this, we redefine and extend the Euler derivative transformation formula to be applied in a system with arbitrary number of coordinate frames. We provide an illustrative example of a rigid body in compound rotation motion about a fixed point to show the appearance of the Razi acceleration. We compare the magnitudes of the Razi acceleration with the classical centripetal acceleration to show that in the case of complex rotating motion, the Razi acceleration may appear as a result of rotational inertia.

## 2.2 Time-Derivative and Coordinate Frame

The time-derivative of a vector depends on the coordinate frame in which it is differentiated. In a multiple coordinate frames environment, it is important to know from which coordinate the derivative vectors, e.g. velocity and acceleration, are calculated and to which coordinate frame it is related. In this section we introduce the derivative transformation formula and expand it to include more than two reference frames.

### 2.2.1 Euler Derivative Transformation Formula

Any vector  $\mathbf{r}$  in a coordinate frame can be represented by its components using the basis vectors of the frame. Let  $B(O\hat{i}\hat{j}\hat{k})$  be an arbitrary coordinate frame in which  $\mathbf{r}$  is expressed. The vector is written using the basis vectors of  $B$  as

$$\mathbf{r} = x\hat{i} + y\hat{j} + z\hat{k} \quad (2.7)$$

For  $\mathbf{r}$ , a vector function of time, its components  $x$ ,  $y$ , and  $z$  are scalar functions of time. The time-derivative of the vector  $\mathbf{r}$  is

$$\frac{d}{dt}\mathbf{r} = \dot{x}\hat{i} + \dot{y}\hat{j} + \dot{z}\hat{k} + x\dot{\hat{i}} + y\dot{\hat{j}} + z\dot{\hat{k}} \quad (2.8)$$

Since the basis vector is constant with respect to its own frame, the time-derivative of a vector expressed in  $B$  differentiated from its own frame  $B$  is

$${}^B_B\dot{\mathbf{r}} = \frac{{}^B d}{dt} {}^B_B\mathbf{r} = \dot{x}\hat{i} + \dot{y}\hat{j} + \dot{z}\hat{k} \quad (2.9)$$

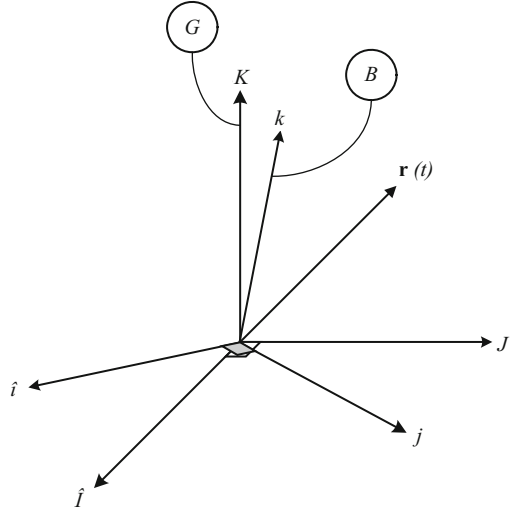
This is called a *local* or *simple* derivative since the vector derivative is taken from the coordinate frame in which it is also expressed. The result is a simple derivative of its scalar components.

Now consider global-fixed coordinate frame  $G(O\hat{I}\hat{J}\hat{K})$  and a rotating body-fixed coordinate frame  $B(O\hat{i}\hat{j}\hat{k})$  as shown in Fig. 2.2. Due to the relative motion of the coordinate frames, the basis vectors of one frame are not constant in relation to the other frame. The time-derivative of a vector expressed in  $B$  differentiated from frame  $G$  is expressed as

$${}^B_B\dot{\mathbf{r}} = \frac{{}^G d}{dt} {}^B_B\mathbf{r} = \dot{x}\hat{i} + \dot{y}\hat{j} + \dot{z}\hat{k} + x\frac{{}^G d}{dt}\hat{i} + y\frac{{}^G d}{dt}\hat{j} + z\frac{{}^G d}{dt}\hat{k} \quad (2.10)$$

Similarly, the time-derivative of a vector expressed in  $G$  differentiated from frame  $B$  is expressed as

**Fig. 2.2** Binary coordinate frames system—a global-fixed frame  $G$  and a rotating body frame  $B$



$${}^G_B \dot{\mathbf{r}} = \frac{{}^B d}{{}^G dt} {}^G \mathbf{r} = \dot{X} \hat{I} + \dot{Y} \hat{J} + \dot{Z} \hat{K} + X \frac{{}^B d}{{}^G dt} \hat{I} + Y \frac{{}^B d}{{}^G dt} \hat{J} + Z \frac{{}^B d}{{}^G dt} \hat{K} \quad (2.11)$$

An angular velocity vector can be defined in terms of a time-derivative of basis vectors coordinate frame in relation to another coordinate frame. The angular velocity of a rotating frame  $B$  in coordinate frame  $G$ ,  ${}^G \boldsymbol{\omega}_B$ , expressed using the  $B$ -coordinates, is written as

$${}^B \boldsymbol{\omega}_B = \left( \hat{k} \cdot \frac{{}^G d \hat{j}}{dt} \right) \hat{i} + \left( \hat{i} \cdot \frac{{}^G d \hat{k}}{dt} \right) \hat{j} + \left( \hat{j} \cdot \frac{{}^G d \hat{i}}{dt} \right) \hat{k} \quad (2.12)$$

And the angular velocity of a rotating frame  $G$  in coordinate frame  $B$ ,  ${}^B \boldsymbol{\omega}_G$ , expressed using the  $G$ -coordinates, is written as

$${}^G \boldsymbol{\omega}_G = \left( \hat{K} \cdot \frac{{}^B d \hat{J}}{dt} \right) \hat{I} + \left( \hat{I} \cdot \frac{{}^B d \hat{K}}{dt} \right) \hat{J} + \left( \hat{J} \cdot \frac{{}^B d \hat{I}}{dt} \right) \hat{K} \quad (2.13)$$

Equations (2.10) and (2.11) can be represented as

$${}^B \dot{\mathbf{r}} = \frac{{}^G d}{{}^B dt} {}^B \mathbf{r} = \frac{{}^B d}{{}^G dt} {}^B \mathbf{r} + {}^B \boldsymbol{\omega}_B \times {}^B \mathbf{r} \quad (2.14)$$

$${}^G \dot{\mathbf{r}} = \frac{{}^B d}{{}^G dt} {}^G \mathbf{r} = \frac{{}^G d}{{}^B dt} {}^G \mathbf{r} + {}^G \boldsymbol{\omega}_G \times {}^G \mathbf{r} \quad (2.15)$$

An alternate proof of Eqs. (2.14) and (2.15) is given in Appendix 1.

The resulting equations show the method of calculating and transforming the derivative of a vector expressed in one frame to another. This results in the most fundamental result of kinematics: Let  $A$  and  $B$  be two arbitrary frames which motion is related by the angular velocity  $\boldsymbol{\omega}$ , for every vector  $\square$  expressed in  $B$ , its time derivative with respect to  $A$  is given by

$$\frac{{}^A d}{{}^A dt} {}^B \square = \frac{{}^B d}{{}^B dt} {}^B \square + {}^B \boldsymbol{\omega}_B \times {}^B \square \quad (2.16)$$

This formula is called the *Euler derivative transformation formula*. It relates the first time-derivative of a vector with respect to another coordinate frame by means of angular velocity.

Despite its importance and ubiquitous application in classical mechanics, there is no universally accepted name for the Euler derivative transformation formula in Eq. (2.16). Several terminologies are used: kinematic theorem (Kane et al. 1983), transport theorem (Rao 2006), and transport equation (Kasdin and Paley 2011). These terms, although terminologically correct, are more prevalent in the subject of fluid mechanics to refer to entirely different concepts. We use Zipfel's (Zipfel 2000) and Jazar's (Jazar 2011) terminology, which we opine as more descriptive and unique to the subject of derivative kinematics. The name also is a homage to Leonard Euler who first used the technique of coordinate transformation in his works on rigid body dynamics (Koetsier 2007; Tenenbaum 2004; Baruh 1999; Van Der Ha and Shuster 2009).

### 2.2.2 Second Derivative

One of the most important applications resulting from the vector derivative transformation method is the development of inertial accelerations. Inertial accelerations, or sometimes called d'Alembert accelerations or, misnomerly, "fictitious" accelerations, are the acceleration terms that appear when the double-derivatives of a position vector are transformed from a rotating frame to an inertial frame or vice versa.

To demonstrate this, let  $G$  be an inertial coordinate frame and  $B$  a non-inertial coordinate frame rotating with respect to  $G$  with an angular velocity  ${}_G \boldsymbol{\omega}_B$ . Using the Euler derivative transformation formula, the first derivative of a vector  ${}^B \mathbf{r}$  as seen from  $G$  is

$$\frac{{}^G d}{{}^G dt} {}^B \mathbf{r} = \frac{{}^B d}{{}^B dt} {}^B \mathbf{r} + {}^B \boldsymbol{\omega}_B \times {}^B \mathbf{r} \quad (2.17)$$

Using the formula again to get the second derivative of the vector  ${}^B \mathbf{r}$  from  $G$  yields

$$\begin{aligned}
\frac{{}^G d}{{}^G dt} \frac{{}^G d}{{}^G dt} {}^B \mathbf{r} &= \frac{{}^G d}{{}^G dt} \left( \frac{{}^B d}{{}^B dt} {}^B \mathbf{r} + {}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r} \right) \\
&= \frac{{}^B d}{{}^B dt} \left( \frac{{}^B d}{{}^B dt} {}^B \mathbf{r} + {}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r} \right) + {}^B_G \boldsymbol{\omega}_B \times \left( \frac{{}^B d}{{}^B dt} {}^B \mathbf{r} + {}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r} \right) \\
&= \frac{{}^B d^2}{{}^B dt^2} {}^B \mathbf{r} + {}^B_G \boldsymbol{\alpha}_B \times {}^B \mathbf{r} + 2 {}^B_G \boldsymbol{\omega}_B \times \frac{{}^B d}{{}^B dt} {}^B \mathbf{r} + {}^B_G \boldsymbol{\omega}_B \times ({}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r}) \\
&= {}^B_{BB} \mathbf{a} + {}^B_G \boldsymbol{\alpha}_B \times {}^B \mathbf{r} + 2 {}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{v} \\
&\quad + {}^B_G \boldsymbol{\omega}_B \times ({}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r})
\end{aligned} \tag{2.18}$$

Equation (2.18) is the classical expression for the acceleration of a point in a rotating frame as seen from an inertial frame. The term  ${}^B_{BB} \mathbf{a}$  is the local acceleration of a point in  $B$  with respect to the frame itself regardless of the rotation of  $B$  in  $G$ . The second term  ${}^B_G \boldsymbol{\alpha}_B \times {}^B \mathbf{r}$  is called the tangential acceleration as its direction is always tangential to the rotation path. Tangential acceleration term is represented by a vector product of an angular acceleration and a position vector. The third term  $2 {}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{v}$  is an acceleration term traditionally named after Coriolis. The Coriolis acceleration term appears from the two different sources: half of the term comes from a local differentiation of  ${}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r}$ , and the other half is from the product of angular velocity and the local linear velocity vector. The Coriolis acceleration in the classical second derivative between two frames is written as twice the vector product of an angular velocity and a linear velocity. The final term in the expression  ${}^B_G \boldsymbol{\omega}_B \times ({}^B_G \boldsymbol{\omega}_B \times {}^B \mathbf{r})$  is called the centripetal acceleration as its direction is towards the center of rotation. The centripetal acceleration has the form of a vector product of an angular velocity with another vector product of an angular velocity and a position vector.

### 2.2.3 Derivative Transformation Formula in Three Coordinate Frames

The Euler derivative transformation formula is now extended to include a third coordinate frame. Let  $A$ ,  $B$ , and  $C$  be three relatively-rotating coordinate frames sharing a same origin as shown in Fig. 2.3.

The first time-derivative of a  $B$ -vector in relative to the  $A$ -frame is written as

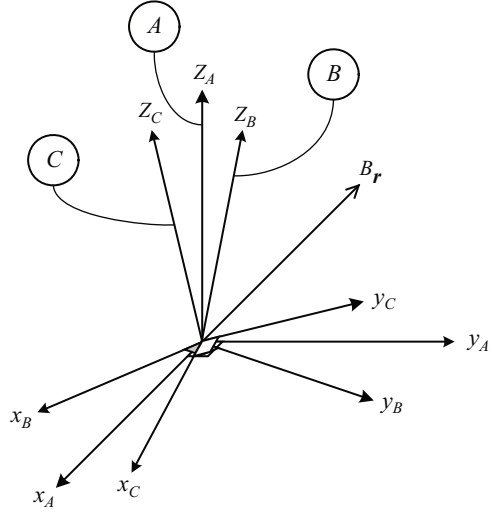
$${}^B_A \mathbf{v} = {}^B_B \mathbf{v} + {}^B_A \boldsymbol{\omega}_B \times {}^B \mathbf{r} \tag{2.19}$$

And the first time derivative of the same  $B$ -vector in relative to the  $C$ -frame is written as

$${}^B_C \mathbf{v} = {}^B_B \mathbf{v} + {}^B_C \boldsymbol{\omega}_B \times {}^B \mathbf{r} \tag{2.20}$$



**Fig. 2.3** Three relatively rotating coordinate frames



Combining Eqs. (2.19) and (2.20) yields

$${}^B_A \dot{\mathbf{v}} = {}^B_C \dot{\mathbf{v}} + ({}^B_A \boldsymbol{\omega}_B - {}^B_C \boldsymbol{\omega}_B) \times {}^B \mathbf{r} \quad (2.21)$$

The resulting equation relates the  $A$ -derivative of a  $B$ -vector (2.19) and the  $C$ -derivative of the same  $B$ -vector (2.20) by means of the angular velocity differences between the three coordinate frames. This result also shows that if the relative velocities and angular velocities with respect to one of the three coordinate frames are known, the simple derivative in local frame can be neglected. In another perspective, by introducing an auxiliary coordinate frame, one can forgo altogether the calculation of the vector derivatives in the local coordinate frame.

The extended Euler derivative transformation in three coordinate frames can be generalized as

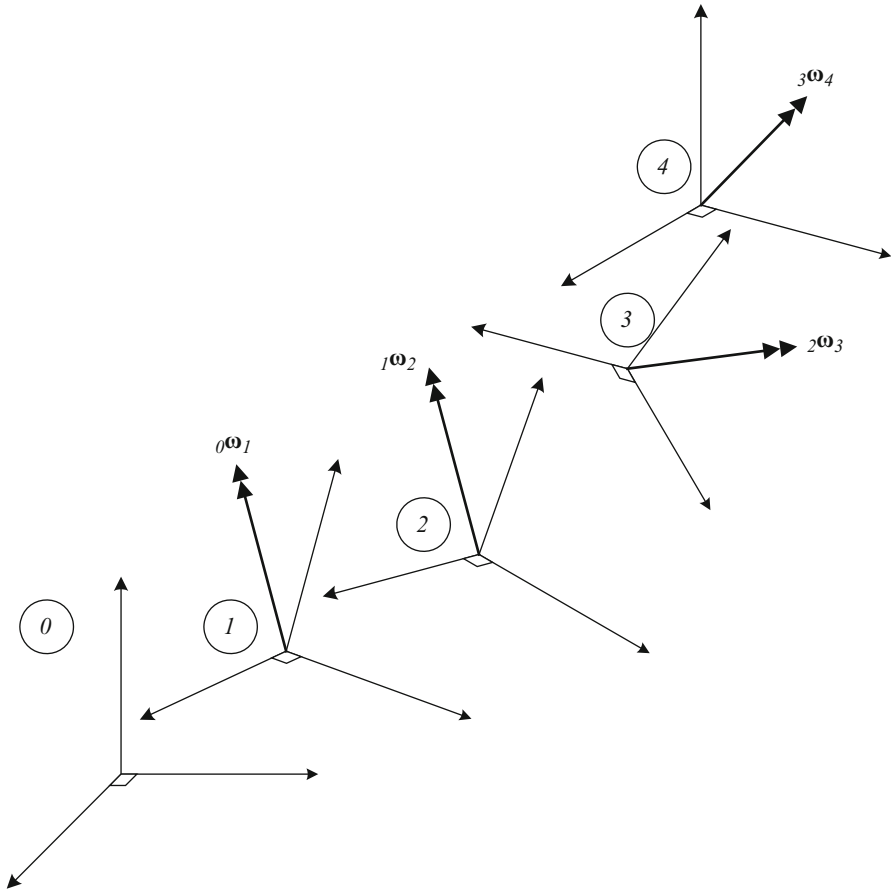
$${}^B_A \dot{\square} = {}^B_C \dot{\square} + ({}^B_A \boldsymbol{\omega}_B - {}^B_C \boldsymbol{\omega}_B) \times {}^B \square \quad (2.22)$$

where  $\square$  can be only generic vector.

This formula is called the mixed derivative transformation formula (Jazar 2012). It can be used to relate vector derivatives in three frames by means of angular velocities difference between the three coordinate frames.

## 2.2.4 Kinematic Chain Rule

To extend the Euler derivative transformation to account for arbitrary number of moving coordinate frames, we derive the general formula for the composition of



**Fig. 2.4** Multiple coordinate frames system with frame 0 as the inertial frame

multiple angular velocity vectors. This is useful for, but not limited to, a system with a chain of coordinate frames where the motion of one frame is subject to another. One example of such system is depicted in Fig. 2.4.

Consider now  $n$  angular velocities representing a set of  $n$  ( $n = 1, 2, \dots, n$ ) non-inertial rotating coordinate frames in an inertial frame indicated by  $n = 0$ . Their relative angular velocities satisfy the relation

$${}^0\omega_n = {}^0\omega_1 + {}^1\omega_2 + {}^2\omega_3 + \dots + {}^{n-1}\omega_n = \sum_{i=1}^n {}^{i-1}\omega_i \quad (2.23)$$

if and only if they all are expressed in the same coordinate frame, which in this case frame 0. Using the rule, the composition of the angular velocity  ${}^0\omega_n$  can be represented in any coordinate frame  $f$  in the set.

$${}^f_0 \boldsymbol{\omega}_n = {}^f R_{00} {}^0 \boldsymbol{\omega}_n = \sum_{i=1}^n {}^{f}_{i-1} \boldsymbol{\omega}_i \quad (2.24)$$

Equation (2.24) represents the additional rule of the composition of multiple angular velocity vectors. Any equation of angular velocities that has no indication of the coordinate frame in which they are expressed is incomplete and generally incorrect.

$${}_0 \boldsymbol{\omega}_n \neq {}_0 \boldsymbol{\omega}_1 + {}_1 \boldsymbol{\omega}_2 + {}_2 \boldsymbol{\omega}_3 + \cdots + {}_{n-1} \boldsymbol{\omega}_n \quad (2.25)$$

Angular acceleration is the first time-derivative of an angular velocity vector. For an angular velocity vector of  $B$  with respect to  $G$ , for example,

$${}^B_G \boldsymbol{\omega}_B = \omega_x \hat{i} + \omega_y \hat{j} + \omega_z \hat{k} \quad (2.26)$$

the angular acceleration is defined as

$${}^B_G \boldsymbol{\alpha}_B = \frac{{}^B d}{{}^G dt} {}^B_G \boldsymbol{\omega}_B = \dot{\omega}_x \hat{i} + \dot{\omega}_y \hat{j} + \dot{\omega}_z \hat{k} \quad (2.27)$$

Suppose there are  $n$  angular velocity vectors representing a set of  $n$  coordinate frames expressed in an arbitrary frame  $f$ .

$${}^f_0 \boldsymbol{\omega}_n = {}^f_0 \boldsymbol{\omega}_1 + {}^f_1 \boldsymbol{\omega}_2 + {}^f_2 \boldsymbol{\omega}_3 + \cdots + {}^f_{n-1} \boldsymbol{\omega}_n \quad (2.28)$$

To differentiate the composition of angular velocities, which are expressed in  $f$ , from an arbitrary frame  $g$  in the set, one has to apply the Euler derivative transformation formula in Eq. (2.16).

$$\begin{aligned} \frac{{}^g d}{{}^g dt} \left( \sum_{i=1}^n {}^{f}_{i-1} \boldsymbol{\omega}_i \right) &= \frac{{}^g d}{{}^g dt} {}^f_0 \boldsymbol{\omega}_n \\ &= \frac{{}^g d}{{}^g dt} \left( {}^f_0 \boldsymbol{\omega}_1 + {}^f_1 \boldsymbol{\omega}_2 + \cdots + {}^f_{n-1} \boldsymbol{\omega}_n \right) \\ &= \left( {}^f_0 \boldsymbol{\alpha}_1 + {}^f_g \boldsymbol{\omega}_1 \times {}^f_0 \boldsymbol{\omega}_1 \right) + \left( {}^f_1 \boldsymbol{\alpha}_2 + {}^f_g \boldsymbol{\omega}_2 \times {}^f_1 \boldsymbol{\omega}_2 \right) + \cdots \\ &\quad \cdots + \left( {}^f_{n-1} \boldsymbol{\alpha}_n + {}^f_g \boldsymbol{\omega}_n \times {}^f_{n-1} \boldsymbol{\omega}_n \right) \\ &= \sum_{i=1}^n {}^{f}_{i-1} \boldsymbol{\alpha}_i + {}^f_g \boldsymbol{\omega}_i \times {}^{f}_{i-1} \boldsymbol{\omega}_i \end{aligned} \quad (2.29)$$

where  $\alpha$  is defined as the differentiation of  $\omega$  in its own frame  ${}^f d {}^f \boldsymbol{\omega} / dt$ . The formula provides a method to express the first derivative of a composition of angular velocity vectors.

As an example, let the number of rotating coordinate frames be  $n = 3$ . The angular velocity of frame 3 with respect to the inertial frame 0 is written as a composition of three angular velocities, expressed in frame 3.

$${}^3_0\boldsymbol{\omega}_3 = {}^3_0\boldsymbol{\omega}_1 + {}^3_1\boldsymbol{\omega}_2 + {}^3_2\boldsymbol{\omega}_3 \quad (2.30)$$

Applying differentiation from frame 0, using the formula in Eq. (2.29), one obtains

$$\begin{aligned} \frac{{}^0d}{{}^0dt} {}^3_0\boldsymbol{\omega}_3 &= ({}^3_0\boldsymbol{\alpha}_1 + {}^3_0\boldsymbol{\omega}_1 \times {}^3_0\boldsymbol{\omega}_1) + ({}^3_1\boldsymbol{\alpha}_2 + {}^3_0\boldsymbol{\omega}_2 \times {}^3_1\boldsymbol{\omega}_2) \\ &\quad + ({}^3_2\boldsymbol{\alpha}_3 + {}^3_0\boldsymbol{\omega}_3 \times {}^3_2\boldsymbol{\omega}_3) \\ &= {}^3_0\boldsymbol{\alpha}_1 + ({}^3_1\boldsymbol{\alpha}_2 + {}^3_0\boldsymbol{\omega}_2 \times {}^3_1\boldsymbol{\omega}_2) + ({}^3_2\boldsymbol{\alpha}_3 + {}^3_0\boldsymbol{\omega}_3 \times {}^3_2\boldsymbol{\omega}_3) \end{aligned} \quad (2.31)$$

Looking at the third term in the right-hand side,  ${}^3_2\boldsymbol{\alpha}_3$  is the rate of *magnitudinal* change of  ${}^3_2\boldsymbol{\omega}_3$  in frame 3, and  ${}^3_0\boldsymbol{\omega}_3 \times {}^3_2\boldsymbol{\omega}_3$  is the *convective* rate of change of  ${}^3_2\boldsymbol{\omega}_3$  due to its angular velocity of frame 3 with respect to the frame it is differentiated,  ${}^3_0\boldsymbol{\omega}_3$ . The second right-hand side term can be interpreted similarly. The total angular acceleration of frame 1 with respect to frame 0 is only  ${}^3_0\boldsymbol{\alpha}_1$  as the direction of the angular velocity of frame 1 with respect to frame 0  ${}^3_0\boldsymbol{\omega}_1$  does not change in frame from which it is differentiated as  ${}^3_0\boldsymbol{\omega}_1 \times {}^3_0\boldsymbol{\omega}_1 = 0$ .

Therefore, any equation of angular acceleration that has no indication of in which coordinate frame they are expressed, or without the convective rate of change term  $\sum_{i=1}^n {}^f_g \boldsymbol{\omega}_i \times {}_{i-1}^f \boldsymbol{\omega}_i$  is incomplete and generally incorrect.

$${}^0\boldsymbol{\alpha}_n \neq {}^0\boldsymbol{\alpha}_1 + {}^1\boldsymbol{\alpha}_2 + {}^2\boldsymbol{\alpha}_3 + \cdots + {}_{n-1}\boldsymbol{\alpha}_n \quad (2.32)$$

$$\frac{{}^g d}{{}^g dt} \left( \sum_{i=1}^n {}_{i-1}^f \boldsymbol{\omega}_i \right) \neq {}^f_0 \boldsymbol{\alpha}_1 + {}^f_1 \boldsymbol{\alpha}_2 + {}^f_2 \boldsymbol{\alpha}_3 + \cdots + {}^f_{n-1} \boldsymbol{\alpha}_n \quad (2.33)$$

### 2.3 The Razi Acceleration

Jazar (2012) showed that the Razi term appears as a relative acceleration due to transforming vector derivative in multiple reference frames. Here, we show that the Razi term can be exhibited by the same method that the inertial accelerations in a fixed-axis rotation motion are calculated, such as the Coriolis and the centripetal acceleration. We examine the system of three nested coordinate frames—a body-fixed coordinate frame  $C$  is rotating in another coordinate frame  $B$ , which in turn is rotating in a globally fixed, inertial frame  $A$ , with all coordinate frames sharing a common origin. The position vector of a body point  ${}^C\mathbf{r}$  is expressed in the body frame  $C$  and its double derivative, hence the acceleration of the body

point, is taken from the inertial frame  $A$ . Any resulting acceleration term, apart from the local absolute acceleration  ${}^C C \mathbf{a}$ , that appears from the transformation of the vector derivatives is defined as an inertial acceleration. The derivative transformation between two coordinate frames produces inertial accelerations such as centripetal, Coriolis, and tangential acceleration as shown in Eq. (2.18). We show that nested rotations involving more than two coordinate frames produce more inertial accelerations terms, in which one of them is the Razi acceleration.

### 2.3.1 Razi Acceleration in Mixed Acceleration Expression

We reproduce the derivation of the Razi acceleration as shown in Jazar (2011, 2012) as a review. Consider three relatively rotating coordinate frames  $A$ ,  $B$ , and  $C$ , where the frame  $A$  is considered as a global-fixed frame, while the frame  $B$  is rotating in  $A$ , and a body frame  $C$  is rotating in  $B$ . This motion is called compound or nested rotation since the rotation of  $C$  is inside the frame  $B$ , which in turn is also rotating inside another frame,  $A$ . Assume a position vector  ${}^C \mathbf{r}$  is expressed in its local frame  $C$ . Taking the first derivative of  ${}^C \mathbf{r}$  from the  $B$ -frame using the vector derivative transformation formula yields

$$\frac{{}^B d}{{}^B dt} {}^C \mathbf{r} = \frac{{}^C d}{{}^C dt} {}^C \mathbf{r} + {}^C \boldsymbol{\omega}_C \times {}^C \mathbf{r} \quad (2.34)$$

Taking the derivative of  $\frac{{}^B d}{{}^B dt} {}^C \mathbf{r}$  from the  $A$ -frame now yields

$$\begin{aligned} {}^{AB} \mathbf{a} &= \frac{{}^A d}{{}^A dt} \frac{{}^B d}{{}^B dt} {}^C \mathbf{r} \\ &= \frac{{}^A d}{{}^A dt} \left( \frac{{}^C d}{{}^C dt} {}^C \mathbf{r} + {}^C \boldsymbol{\omega}_C \times {}^C \mathbf{r} \right) \\ &= \frac{{}^C d}{{}^C dt} \left( \frac{{}^C d}{{}^C dt} {}^C \mathbf{r} + {}^C \boldsymbol{\omega}_C \times {}^C \mathbf{r} \right) + {}^C \boldsymbol{\omega}_C \times \left( \frac{{}^C d}{{}^C dt} {}^C \mathbf{r} + {}^C \boldsymbol{\omega}_C \times {}^C \mathbf{r} \right) \\ &= {}^{CC} \mathbf{a} + {}^C \boldsymbol{\alpha}_C \times {}^C \mathbf{r} + {}^C \boldsymbol{\omega}_C \times {}^C \mathbf{v} + {}^C \boldsymbol{\omega}_C \times {}^C \mathbf{v} \\ &\quad + ({}^C \boldsymbol{\omega}_C \times {}^C \boldsymbol{\omega}_C) \times {}^C \mathbf{r} + {}^C \boldsymbol{\omega}_C \times ({}^C \boldsymbol{\omega}_C \times {}^C \mathbf{r}) \end{aligned} \quad (2.35)$$

From the expression, the mixed acceleration expression contains the local acceleration term  ${}^{CC} \mathbf{a}$ ; two acceleration terms that have the Coriolis form of a product of an angular velocity and a linear velocity, which are  ${}^C \boldsymbol{\omega}_C \times {}^C \mathbf{v}$  and  ${}^C \boldsymbol{\omega}_C \times {}^C \mathbf{v}$ ; the tangential acceleration term  ${}^C \boldsymbol{\alpha}_C \times {}^C \mathbf{r}$ ; and a mixed centripetal acceleration term  ${}^C \boldsymbol{\omega}_C \times ({}^C \boldsymbol{\omega}_C \times {}^C \mathbf{r})$ .

The term  $({}^C \boldsymbol{\omega}_C \times {}^C \boldsymbol{\omega}_C) \times {}^C \mathbf{r}$  is the Razi acceleration. Note the difference between the Razi acceleration and the mixed centripetal acceleration is the order

of cross product. This indicates that when a coordinate body frame is rotating relative to other two rotating coordinate frame, the mixed double derivative of  ${}^C \mathbf{r}$  contains these inertial acceleration terms: mixed Coriolis acceleration, tangential acceleration, mixed centripetal acceleration, and the Razi acceleration. An application example of mixed derivative transformation formula with this Razi acceleration is illustrated on Jazar (2012, pp. 858–864).

However, the dynamical interpretation of the mixed acceleration is not clear. In practice, the double derivative of a position vector is usually directly transformed from the local frame to the inertial frame without undergoing a mid-frame differentiation. The reason is that a double-differentiation of a point position in a rotating frame from the inertial frame gives the additional acceleration terms, which contribute to inertial effect such as centrifugal and Coriolis forces. In mixed acceleration, in contrast, the differentiation is taken from two different frames. The mixed acceleration  ${}_{AB} {}^C \mathbf{a} = \frac{A d}{dt} {}^C \mathbf{v}$  can be used to complement the acceleration of a point in  $C$  from  $B$  using the mixed derivative transformation formula in Eq. (2.22).

$$\frac{B d}{dt} {}^C \mathbf{v} = \frac{A d}{dt} {}^C \mathbf{v} + ({}^C \boldsymbol{\omega}_C - {}^C \boldsymbol{\omega}_A) \times {}^C \mathbf{v} \quad (2.36)$$

Next, we show that if there are more than one nested rotation, the Razi acceleration would appear even when the vector derivative is transformed directly from the local frame to the inertial frame without using the mixed acceleration technique.

### 2.3.2 Razi Acceleration as an Inertial Acceleration

The derivation of the Razi acceleration here uses the Euler derivative transformation formula in conjunction with the kinematic chain rule. It differs from the mixed acceleration method in such a way that the both derivatives of a body point in a local coordinate frame are taken from the inertial frame, the same way the Coriolis, centripetal, and tangential accelerations are derived. To demonstrate, consider a two coordinate frames system with  $G$  as the global-fixed inertial frame and  $B$  as the rotating body-fixed frame. Since the relative rotation between the two frames can be represented by one angular velocity  ${}_G \boldsymbol{\omega}_B$ , we can use Eq. (2.16) twice to find the double derivative of a body-fixed position vector  ${}^B \mathbf{r}$  from the inertial frame  $G$ . Looking at the resulting expression of the acceleration in two frames, shown in Eq. (2.18), one can separate the acceleration terms as follows

$$\begin{aligned}
{}_{BB}^B \mathbf{a} &\equiv \text{local acceleration in } B \\
{}_{G}^B \boldsymbol{\alpha}_B \times {}^B \mathbf{r} &\equiv \text{tangential acceleration} \\
2{}_{G}^B \boldsymbol{\omega}_B \times {}_{B}^B \mathbf{v} &\equiv \text{Coriolis acceleration} \\
{}_{G}^B \boldsymbol{\omega}_B \times ({}_{G}^B \boldsymbol{\omega}_B \times {}^B \mathbf{r}) &\equiv \text{centripetal acceleration}
\end{aligned} \tag{2.37}$$

Now, consider a slightly complex case of nested rotations. Consider three relatively rotating frames  $A$ ,  $B$ , and  $C$  where the  $A$ -frame is treated as the inertial frame, the  $B$ -frame is rotating in  $A$ , and the  $C$ -frame, which is the body-fixed frame, is rotating in  $B$ . The motion of a rigid body which is spinning in one or more rotating coordinate frames is called compound rotation or nested rotation. Compound rotation is usually seen in the gyroscopic motion, or the precession and nutation of a rigid body. To express the velocity of a point in  $C$  as seen from the inertial frame, the Euler derivative transformation formula in (2.16) is invoked to transform the velocity from the local frame  $C$  to the inertial frame  $A$

$$\frac{{}^A d}{{}^A dt} {}^C \mathbf{r} = {}^C \dot{\mathbf{r}} = {}^C \dot{\mathbf{r}} + {}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{r} \tag{2.38}$$

Since, the frame  $C$  is in  $B$ , and the frame  $B$  is in  $A$ , the angular velocity  ${}^C_A \boldsymbol{\omega}_C$  can be expanded using the angular velocity addition rule from Eq. (2.24).

$${}^C_A \boldsymbol{\omega}_C = {}^C_A \boldsymbol{\omega}_B + {}^C_B \boldsymbol{\omega}_C \tag{2.39}$$

The expression of the acceleration of a point in  $C$  from the inertial frame can be expressed as

$$\frac{{}^A d}{{}^A dt} \frac{{}^A d}{{}^A dt} {}^C \mathbf{r} = {}^C \mathbf{a} + \frac{{}^A d}{{}^A dt} {}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{r} + 2{}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{v} + {}^C_A \boldsymbol{\omega}_C \times ({}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{r}) \tag{2.40}$$

The angular velocity derivative  $\frac{{}^A d}{{}^A dt} {}^C_A \boldsymbol{\omega}_C$  can be expanded using the rule of addition for angular acceleration in (2.29) with  $n = 2$

$$\begin{aligned}
\frac{{}^A d}{{}^A dt} {}^C_A \boldsymbol{\omega}_C &= \frac{{}^A d}{{}^A dt} ({}^C_A \boldsymbol{\omega}_B + {}^C_B \boldsymbol{\omega}_C) \\
&= {}^C_A \boldsymbol{\alpha}_B + {}^C_A \boldsymbol{\omega}_B \times {}^C_A \boldsymbol{\omega}_B + {}^C_B \boldsymbol{\alpha}_C + {}^C_A \boldsymbol{\omega}_C \times {}^C_B \boldsymbol{\omega}_C \\
&= {}^C_A \boldsymbol{\alpha}_B + {}^C_B \boldsymbol{\alpha}_C + {}^C_A \boldsymbol{\omega}_C \times {}^C_B \boldsymbol{\omega}_C
\end{aligned} \tag{2.41}$$

The expanded expression of the acceleration of a point in  $C$  from the inertial frame now becomes

$$\begin{aligned}
\frac{{}^A d}{{}^A dt} \frac{{}^A d}{{}^A dt} {}^C \mathbf{r} &= {}^C_C \mathbf{a} + {}^C_A \boldsymbol{\alpha}_B \times {}^C \mathbf{r} + ({}^C_A \boldsymbol{\omega}_C \times {}^C_B \boldsymbol{\omega}_C) \times {}^C \mathbf{r} + {}^C_B \boldsymbol{\alpha}_C \times {}^C \mathbf{r} \\
&\quad + 2{}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{v} + {}^C_A \boldsymbol{\omega}_C \times ({}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{r})
\end{aligned} \tag{2.42}$$

One can now list the acceleration terms expressed in local frame  $C$  and include the Razi term

$$\begin{aligned}
 {}^C_C \mathbf{a} &\equiv \text{local acceleration in } C \\
 {}^C_A \boldsymbol{\alpha}_C \times {}^C \mathbf{r} &\equiv \text{tangential acceleration} \\
 2 {}^C_A \boldsymbol{\omega}_C \times {}^C_C \mathbf{v} &\equiv \text{Coriolis acceleration} \\
 {}^C_A \boldsymbol{\omega}_C \times ({}^C_A \boldsymbol{\omega}_C \times {}^C \mathbf{r}) &\equiv \text{centripetal acceleration} \\
 ({}^C_A \boldsymbol{\omega}_C \times {}^C_B \boldsymbol{\omega}_C) \times {}^C \mathbf{r} &\equiv \text{Razi acceleration}
 \end{aligned} \tag{2.43}$$

The Razi acceleration  $({}^C_A \boldsymbol{\omega}_C \times {}^C_A \boldsymbol{\omega}_B) \times {}^C \mathbf{r}$  comes from differentiating the composition of angular velocity vectors from a different frame as shown in Eq. (2.41). Next, we show that for the case of  $n > 2$  nested rotating frames, Razi acceleration terms appear in the equation of motion.

It is important to note that the Razi acceleration can only appear in compound rotation motion. A rigid body in a rotation about a *fixed* axis does not experience an inertial force caused by the Razi acceleration. This is because a rotation about a fixed axis can always be simplified to a two coordinate frames system. In such case the Razi acceleration vanishes as there is only one angular velocity, making  ${}^A \boldsymbol{\omega}_B \times {}^A \boldsymbol{\omega}_B = 0$ .

### 2.3.3 Acceleration Transformation in Multiple Coordinate Frames with a Common Origin

Given a set of  $n$  non-inertial rotating coordinate frames in nested rotations with  ${}_{i-1} \boldsymbol{\omega}_i$  representing the angular velocity of  $i$ -th coordinate frame with respect to the preceding coordinate frame, we present the general expression for the acceleration (the double derivative of a vector) of the  $n$ -th coordinate frame as seen from the inertial coordinate frame 0.

$$\begin{aligned}
 \frac{{}^0 d^2}{{}^0 dt^2} {}^n \mathbf{r} &= {}^n_0 \mathbf{a} = {}^n_n \mathbf{a} + \sum_{i=1}^n {}_{i-1}^n \boldsymbol{\alpha}_i \times {}^n \mathbf{r} + \sum_{i=1}^n ({}^n_0 \boldsymbol{\omega}_i \times {}_{i-1}^n \boldsymbol{\omega}_i) \times {}^n \mathbf{r} \\
 &\quad + 2 \sum_{i=1}^n {}^n_0 \boldsymbol{\omega}_i \times {}^n \mathbf{v} + \sum_{i=1}^n {}^n_0 \boldsymbol{\omega}_i \times \left( \sum_{i=1}^n {}^n_0 \boldsymbol{\omega}_i \times {}^n \mathbf{r} \right)
 \end{aligned} \tag{2.44}$$



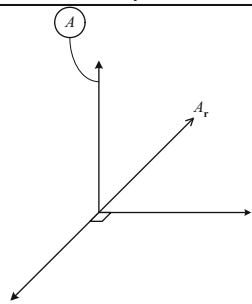
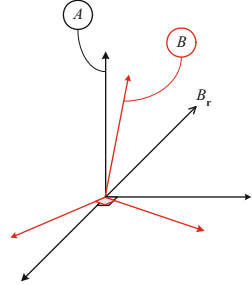
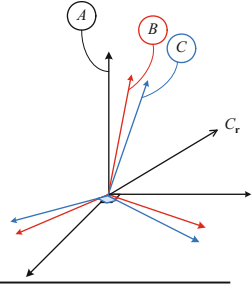
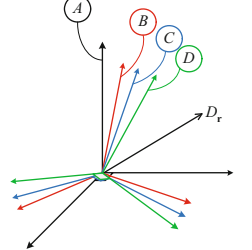
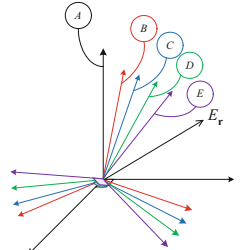
where

$$\begin{aligned}
 {}_n^n \mathbf{a} &\equiv \text{local acceleration} \\
 \sum_{i=1}^n {}_{i-1}^n \boldsymbol{\alpha}_i \times {}^n \mathbf{r} &\equiv \text{tangential acceleration} \\
 2 \sum_{i=1}^n {}_0^n \boldsymbol{\omega}_i \times {}_n^n \mathbf{v} &\equiv \text{Coriolis acceleration} \\
 \sum_{i=1}^n {}_0^n \boldsymbol{\omega}_i \times \left( \sum_{i=1}^n {}_0^n \boldsymbol{\omega}_i \times {}^n \mathbf{r} \right) &\equiv \text{centripetal acceleration} \\
 \sum_{i=1}^n ({}_0^n \boldsymbol{\omega}_i \times {}_{i-1}^n \boldsymbol{\omega}_i) \times {}^n \mathbf{r} &\equiv \text{Razi acceleration}
 \end{aligned} \tag{2.45}$$

The general expressions for acceleration when a position vector in a local frame is differentiated twice from the inertial frame are shown in Table 2.1. Examples are given for up to a five coordinate frames system. The frame  $A$  is dedicated to the inertial frame and the successive relatively rotating frames are named  $B$ ,  $C$ , and so on, with the last letter indicates the local coordinate frame. For example, in a system of three coordinate frames,  $A$  is the inertial frame,  $B$  is the intermediate frame, and  $C$  is the local frame. All vectors are expressed in the local coordinate frame as indicated by the left superscript on each vector. Addition of angular acceleration vectors is simplified as  ${}_0^n \boldsymbol{\alpha}_n \equiv \sum_{i=1}^n {}_{i0}^n \boldsymbol{\alpha}_i$ .

The resulting acceleration expression in Table 2.1 shows that if a rigid body moves linearly with respect to an inertial space, its motion can be sufficiently described by a linear acceleration vector  ${}^A \mathbf{a}$ . When a rotating reference frame is introduced and the motion is described in this frame, the additional acceleration terms—centripetal, Coriolis, and tangential accelerations—act as “supplementary” accelerations to describe the motion in a rotating reference frame. When another rotating frame is introduced in the existing rotating frame making a three coordinate frames system, the expression adds another acceleration term: the Razi accelerations. As more coordinate frames are introduced in the nested rotating frames system, the acceleration expressions becomes more complex to accommodate the relative motion between the coordinate frames. In general, Eq. (2.44) includes all of the relative accelerations acting on a rigid body in compound motion which are otherwise unobservable with the classical derivative kinematics method and not intuitive to the analyst.

**Table 2.1** Inertial acceleration terms for describing rigid body in multiple frames system

No. of frames	Acceleration terms	Enclosed system
Inertial frame ( $n = 0$ )	$A_a$	
Inertial frame + 1 rotating frame ( $n = 1$ )	$B_a$ $+ {}^B_A \omega_B \times ({}^B_A \omega_B \times B_r)$ $+ 2 {}^B_A \omega_B \times B_v$ $+ {}^B_A \alpha_B \times B_r$	
Inertial frame + 1 rotating frame ( $n = 1$ )	$B_a$ $+ {}^B_A \omega_B \times ({}^B_A \omega_B \times B_r)$ $+ 2 {}^B_A \omega_B \times B_v$ $+ {}^B_A \alpha_B \times B_r$	
Inertial frame + 3 rotating frames ( $n = 3$ )	$D_a$ $+ {}^D_A \omega_D \times ({}^D_A \omega_D \times D_r)$ $+ 2 {}^D_A \omega_D \times D_v$ $+ {}^D_A \alpha_D \times D_r$ $+ ({}^D_A \omega_D \times {}^D_C \omega_C) \times C_r$ $+ ({}^D_A \omega_D \times {}^D_B \omega_B) \times B_r$	
Inertial frame + 4 rotating frames ( $n = 4$ )	$E_a$ $+ {}^E_A \omega_E \times ({}^E_A \omega_E \times E_r)$ $+ 2 {}^E_A \omega_E \times E_v$ $+ {}^E_A \alpha_E \times E_r$ $+ ({}^E_A \omega_E \times {}^E_C \omega_C) \times C_r$ $+ ({}^E_A \omega_E \times {}^E_B \omega_B) \times B_r$ $+ ({}^E_A \omega_E \times {}^E_D \omega_D) \times D_r$	

## 2.4 Mechanical Interpretation of the Razi Acceleration

Classical dynamics asserts that there exists a rigid frame of reference in which Newton's second law of motion is valid. Assuming the mass is constant at all times, the law of motion is written as

$$\mathbf{F} = m\mathbf{a} \quad (2.46)$$

where  $m$  indicates a mass scalar and  $\mathbf{a}$  represents the acceleration vector. The reference frame in which this relationship is valid is called an *inertial* frame. An inertial frame is fixed in space, or at most, moves in relation to fixed points in space with a constant velocity. On the other hand, a *non-inertial* reference frame can be defined as a frame which accelerates with respect to the inertially fixed points. The calculation of motion observed from a non-inertial reference frame does not follow Newton's second law due to the relative acceleration between the observer's frame and the inertial frame.

Starting from an inertial frame, the method of derivative kinematics is used to calculate and transform the vector derivatives to the non-inertial observer. Such transformation produces additional acceleration terms to the Newton's law in Eq. (2.46). For the classical case of one stationary inertial frame and a body frame which rotates about a fixed axis, these additional acceleration terms are generally divided into tangential acceleration, Coriolis acceleration, and centripetal acceleration, as shown in the expression in Eq. (2.18).

$$\mathbf{F} = m(\mathbf{a} + \mathbf{a}_{\text{tangential}} + \mathbf{a}_{\text{Coriolis}} + \mathbf{a}_{\text{centripetal}}) \quad (2.47)$$

Multiplying the acceleration terms with mass gives the expression of the forces that is acting on a point mass in the rotating coordinate frame.

In the case of multiple nested coordinate frames, the classical method of transforming derivative between two frames is limiting. This method is sufficient for any motion consideration in two-dimensional space, due to the assumption that the direction of a rigid body's angular velocity is fixed with respect to the inertial frame at all time. However, a general formula for a three-dimensional space is needed to account for the subtle accelerations due to a non-constant angular velocity.

The derivative transformation method can be extended to include the effects of moving angular velocities. To do this, the complex rotation caused by the non-constant angular velocity vector is decomposed into several simpler coordinate frame representation such that each coordinate frame then will have a fixed-axis angular velocity vector. Such technique results on multiple coordinate frames in nested rotations. To analyze such system, we incorporate the extended derivative transformation formula and the kinematic chain rule, which method reveals new results and application.

For the case of three coordinate frames system (two nested rotating frames in an inertial frame), we have seen that the acceleration expression includes

a new rotational inertia called which is the Razi acceleration. We also have shown the general expression for all four inertial acceleration—tangential, Coriolis, centripetal, and Razi—for an arbitrary number of nested rotating frames.

$$\mathbf{F} = m (\mathbf{a} + \mathbf{a}_{\text{tangential}} + \mathbf{a}_{\text{Coriolis}} + \mathbf{a}_{\text{centripetal}} + \mathbf{a}_{\text{Razi}}) \quad (2.48)$$

Extending the derivative transformation formula and incorporating the kinematic chain rule reveal several new results and applications. The Razi acceleration is given a focus due to its appearance as one of the additional acceleration terms in the system of more than two coordinate frames.

We take the simplest case of the Razi acceleration where there are three nested coordinate frames— $A$ ,  $B$ , and  $C$ —sharing an origin with the body coordinate frame  $C$  being the “innermost” frame. Such system can be better illustrated by a flat disc in compound rotation motion, as shown in Fig. 2.5. We fix the  $C$  coordinate frame in the disc, which is spinning on its own axis with an angular speed  $\dot{\theta}$  and tilted on a turning shaft by  $\psi$  degrees. We fix the  $B$  coordinate frame on the shaft, which is turning with respect to the global-fixed frame with an angular speed  $\dot{\phi}$ . The two angular velocities can be written as  ${}_B\boldsymbol{\omega}_C = \dot{\theta}\hat{i}$  and  ${}_A\boldsymbol{\omega}_B = \dot{\phi}\hat{j}$ .

The total acceleration acting on a point  $\mathbf{r}$  on the disc, as seen from the inertial frame, can be expressed by Eq. (2.42). Decomposing the expression into multiple terms, the Razi acceleration, if expressed in the local body coordinates of  $C$ , is written as

$$\mathbf{a}_{\text{Razi}} = ({}_A\boldsymbol{\omega}_C \times {}_B\boldsymbol{\omega}_C) \times {}^C\mathbf{r} \quad (2.49)$$

The cross product of two angular velocities  ${}_A\boldsymbol{\omega}_C \times {}_B\boldsymbol{\omega}_C$  is the convective rate of change of  ${}_B\boldsymbol{\omega}_C$  due to the angular velocity of  $C$  in  $A$ ,  ${}_A\boldsymbol{\omega}_C$ . Therefore, the Razi acceleration can be described as an inertial effect caused by the change of the local angular velocity vector direction. This is separable from the tangential direction which is due to the change of an angular velocity magnitude, hence producing acceleration in the direction which is tangent to the rotation curve.

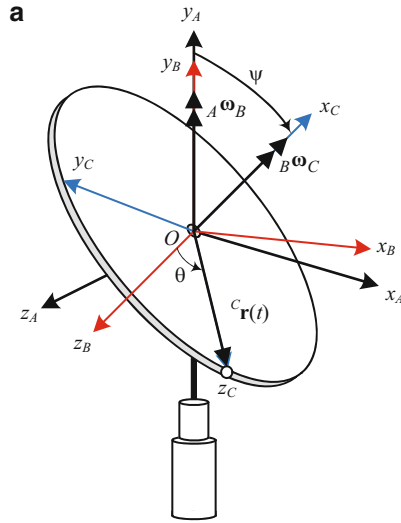
To examine the direction of action of the Razi acceleration, we visualize the resulting vectors in Fig. 2.6. Assuming the position vector of the point of interest  $\mathbf{r}$  is constant in  $C$ . Since the angular velocity of  $C$  in  $A$  can be decomposed using the addition rule, i.e.  ${}_A\boldsymbol{\omega}_C = {}_A\boldsymbol{\omega}_B + {}_B\boldsymbol{\omega}_C$ , the Razi acceleration is written as

$$\mathbf{a}_{\text{Razi}} = ({}_A\boldsymbol{\omega}_B \times {}_B\boldsymbol{\omega}_C) \times {}^C\mathbf{r} \quad (2.50)$$

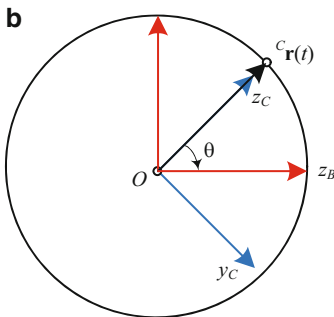
The total acceleration, thus, can be expressed as

$$\frac{{}^A d}{{}^A dt} \frac{{}^A d}{{}^A dt} {}^C\mathbf{r} = ({}_A\boldsymbol{\omega}_B \times {}_B\boldsymbol{\omega}_C) \times {}^C\mathbf{r} + {}_A\boldsymbol{\omega}_C \times ({}_A\boldsymbol{\omega}_C \times {}^C\mathbf{r}) \quad (2.51)$$

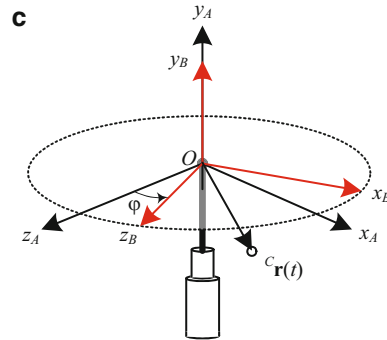
In Fig. 2.6, the total acceleration vector of a body point in  $C$  is decomposed into the centripetal acceleration vector  ${}_A\boldsymbol{\omega}_C \times ({}_A\boldsymbol{\omega}_C \times {}^C\mathbf{r})$  and the Razi acceleration vector



Body coordinate frame  $C$  spinning in shaft coordinate frame  $B$ . The  $B$ -frame, in turn, is rotating in inertial frame  $A$ .



$C$ -frame rotating in  $B$ -frame with  ${}^C_B \omega_C = \dot{\theta} \hat{i}$



$B$ -frame rotating in  $A$ -frame with  ${}^B_A \omega_B = \dot{\phi} \hat{i}$

**Fig. 2.5** Describing compound rotation motion using three coordinate frames system

$({}^C_A \omega_B \times {}^C_B \omega_C) \times {}^C \mathbf{r}$ . Using the disc to represent the plane of rotation of the  $C$ -frame, we can see the resultant vector of the Razi acceleration is always in the out-of-plane direction. This can be tested by finding the direction of the cross products in the Razi term using the right-hand rule.

Apart from the magnitudes of  ${}^C_A \omega_B$ ,  ${}^C_B \omega_C$ , and  ${}^C \mathbf{r}$ , the magnitude of the Razi acceleration vector depends on the angle between the two angular velocity vectors,  $\psi$ . Figures 2.7, 2.8, 2.9, 2.10, 2.11, 2.12 show the magnitudes of Razi

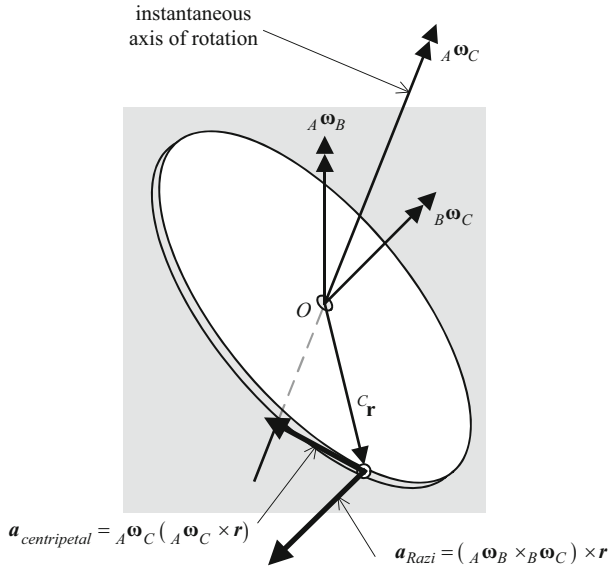


Fig. 2.6 Razi acceleration vector

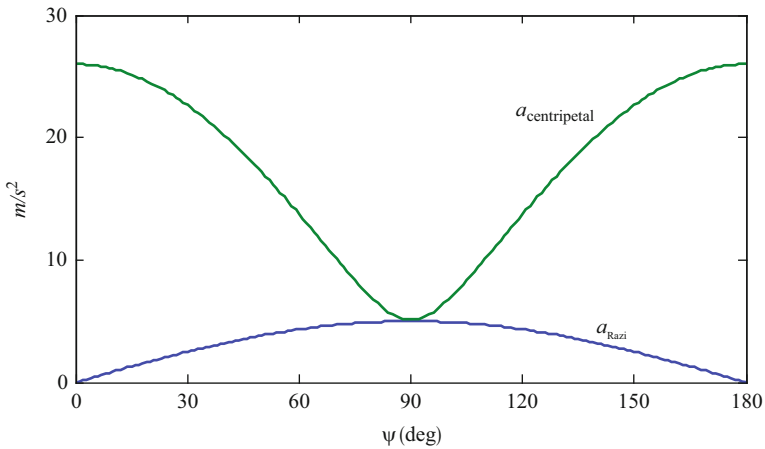
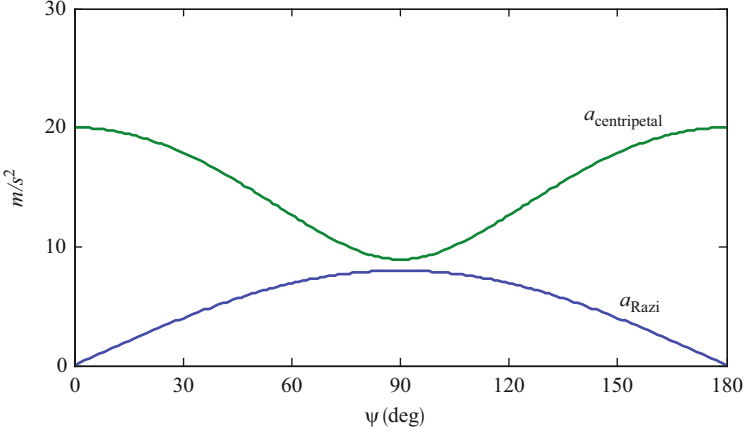
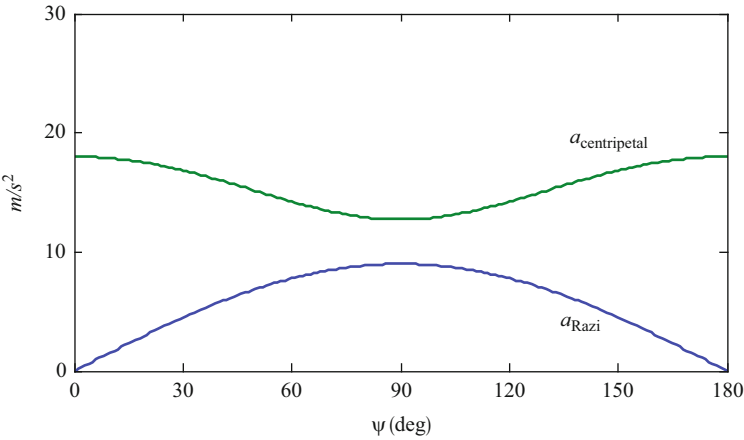


Fig. 2.7  ${}^B_A\omega_B = \dot{\phi} = 5 \text{ rad/s}$ ;  ${}^C_B\omega_C = \dot{\theta} = 1 \text{ rad/s}$

and centripetal acceleration experienced by the body point  ${}^C\mathbf{r}$  in the example. The plots show the effect of  $\psi$  to the magnitude of Razi acceleration for different combinations of angular velocities. Since most readers are more familiar with the centripetal acceleration effect and its influence on bodies under rotation, we compare its magnitude with the magnitude of Razi acceleration.



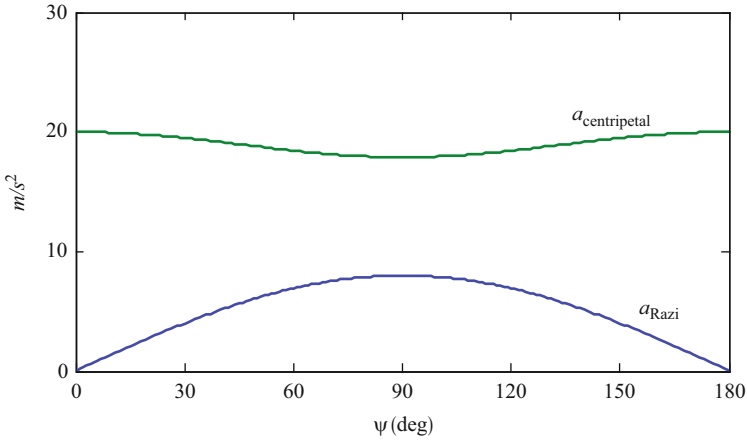
**Fig. 2.8**  ${}^B_A\omega_B = \dot{\phi} = 4 \text{ rad/s}$ ;  ${}^C_B\omega_C = \dot{\theta} = 2 \text{ rad/s}$



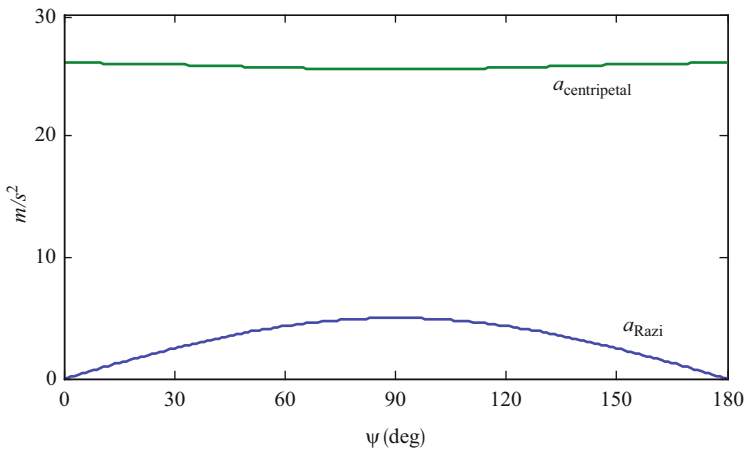
**Fig. 2.9**  ${}^B_A\omega_B = \dot{\phi} = 3 \text{ rad/s}$ ;  ${}^C_B\omega_C = \dot{\theta} = 3 \text{ rad/s}$

### 2.5 Conclusion

By revising and extending the classical Euler vector derivative transformation formula, we update the interpretation of the Razi acceleration and reveal its application in rigid body motion. The Razi acceleration is revealed as one of the inertial effects, apart from the classical Coriolis, centripetal, and tangential accelerations, affecting rigid bodies in compound rotation motion. It is shown to appear as a result of the convective rate of change of a body-frame’s angular velocity in nested rotation. Razi’s mathematical expression,  $\sum_{i=1}^n ({}^n\omega_i \times {}_{i-1}^n\omega_i) \times {}^n\mathbf{r}$ , can only



**Fig. 2.10**  ${}^B_A\omega_B = \dot{\phi} = 2 \text{ rad/s}$ ;  ${}^C_B\omega_C = \dot{\theta} = 4 \text{ rad/s}$



**Fig. 2.11**  ${}^B_A\omega_B = \dot{\phi} = 1 \text{ rad/s}$ ;  ${}^C_B\omega_C = \dot{\theta} = 5 \text{ rad/s}$

be discovered by using multiple coordinate frames in analyzing the kinematics of complexly rotating bodies. A practical example is illustrated to show the typical direction of action of Razi acceleration. Its magnitude is compared to the more widely acknowledged centripetal acceleration to show that the Razi acceleration can have a significant impact to the structures of a body in compound rotation.



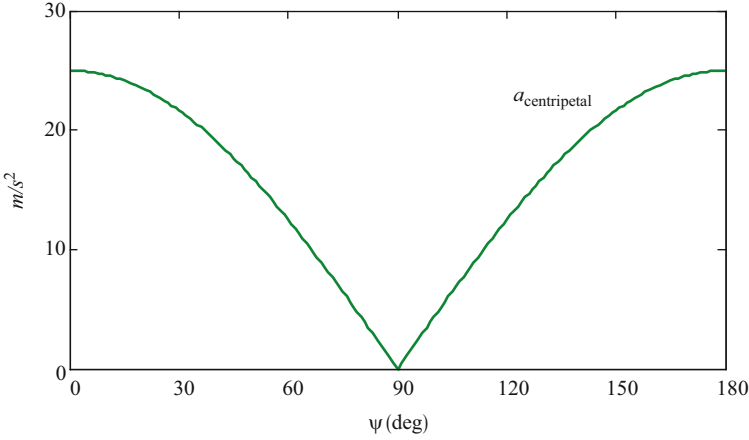


Fig. 2.12  ${}^B\omega_B = \dot{\phi} = 5 \text{ rad/s}$ ;  ${}^C\omega_C = \dot{\theta} = 0 \text{ rad/s}$

## Appendix 1: Alternative Proof of Derivative Transformation Formula

Angular velocity can be defined in terms of time-derivative of rotation matrix. Consider the two coordinate frames system as shown in Fig. 2.2. The angular velocity of  $B$  in relative to  $G$ ,  ${}^G\omega_B$  is written as follows

$$\begin{aligned} {}^G\omega_B &= {}^G\dot{R}_B {}^G R_B^T && \text{if expressed in } G \\ {}^B\omega_B &= {}^G R_B^T {}^G\dot{R}_B && \text{if expressed in } B \end{aligned} \quad (2.52)$$

To prove Eq. (2.52), we start with the relation between a vector expressed in two arbitrary coordinate frames  $G$  and  $B$ .

$${}^G\mathbf{r} = {}^G R_B {}^B\mathbf{r} \quad (2.53)$$

Finding the derivative of the  $G$ -vector from  $G$ , we get

$${}^G\dot{\mathbf{r}} = {}^G\dot{R}_B {}^B\mathbf{r} = {}^G\dot{R}_B {}^G R_B^T {}^G\mathbf{r} = {}^G\omega_B \times {}^G\mathbf{r} \quad (2.54)$$

Here, it is assumed that the vector  ${}^B\mathbf{r}$  is constant in  $B$ . If it is not constant, we have to include the simple derivative of the vector in  $B$  and rotate it to  $G$

$${}^G\dot{\mathbf{r}} = {}^G\dot{R}_B {}^B\mathbf{r} + {}^G R_B {}^B\dot{\mathbf{r}} = {}^G\omega_B \times {}^G\mathbf{r} + {}^G\dot{\mathbf{r}} \quad (2.55)$$

The angular velocities between two arbitrary frames are equal and oppositely directed given that both are expressed in the same frame

$${}^G\omega_B = -{}^B\omega_G \quad (2.56)$$

Using this angular velocity property, rearranging Eq. (2.54) gives

$${}^G_B \dot{\mathbf{r}} = {}^G_G \dot{\mathbf{r}} + {}^G_B \boldsymbol{\omega}_G \times {}^G_G \mathbf{r} \quad (2.57)$$

## Appendix 2: Proof of the Kinematic Chain Rule for Angular Acceleration Vectors

To prove Eq. (2.29), we start with a composition of  $n$  angular velocity vectors which are all expressed in an arbitrary frame  $g$ .

$$\begin{aligned} {}^g_0 \boldsymbol{\omega}_n &= \sum_{i=1}^n {}^{g}_{i-1} \boldsymbol{\omega}_i = {}^g_0 \boldsymbol{\omega}_1 + {}^g_1 \boldsymbol{\omega}_2 + \cdots + {}^{g}_{n-1} \boldsymbol{\omega}_n \\ &= {}^g R_1 {}^1_0 \boldsymbol{\omega}_1 + {}^g R_2 {}^2_1 \boldsymbol{\omega}_2 + \cdots + {}^g R_n {}^{n}_{(n-1)} \boldsymbol{\omega}_n \end{aligned} \quad (2.58)$$

We use the technique of differentiating from another frame in Appendix 1. Differentiating the angular velocities from an arbitrary frame  $g$  gives

$$\begin{aligned} \frac{{}^g d}{dt} \sum_{i=1}^n {}^{g}_{i-1} \boldsymbol{\omega}_i &= \frac{{}^g d}{dt} \left( {}^g R_1 {}^1_0 \boldsymbol{\omega}_1 + {}^g R_2 {}^2_1 \boldsymbol{\omega}_2 + \cdots + {}^g R_n {}^{n}_{(n-1)} \boldsymbol{\omega}_n \right) \\ &= ({}^g R_1 {}^1_0 \boldsymbol{\alpha}_1 + {}^g \dot{R}_1 {}^1_0 \boldsymbol{\omega}_1) + ({}^g R_2 {}^2_1 \boldsymbol{\alpha}_2 + {}^g \dot{R}_2 {}^2_1 \boldsymbol{\omega}_2) + \cdots \\ &\quad \cdots + ({}^g R_n {}^{n}_{(n-1)} \boldsymbol{\alpha}_n + {}^g \dot{R}_n {}^{n}_{(n-1)} \boldsymbol{\omega}_n) \\ &= ({}^g R_1 {}^1_0 \boldsymbol{\alpha}_1 + {}^g \dot{R}_1 {}^g R_1^T {}^g_0 \boldsymbol{\omega}_1) + ({}^g R_2 {}^2_1 \boldsymbol{\alpha}_2 + {}^g \dot{R}_2 {}^g R_2^T {}^g_1 \boldsymbol{\omega}_2) + \cdots \\ &\quad \cdots + ({}^g R_n {}^{n}_{(n-1)} \boldsymbol{\alpha}_n + {}^g \dot{R}_n {}^g R_n^T {}^{g}_{(n-1)} \boldsymbol{\omega}_n) \\ &= ({}^g_0 \boldsymbol{\alpha}_1 + {}^g_0 \boldsymbol{\omega}_1 \times {}^g_0 \boldsymbol{\omega}_1) + ({}^g_1 \boldsymbol{\alpha}_2) + {}^g_0 \boldsymbol{\omega}_2 \times {}^g_1 \boldsymbol{\omega}_2 + \cdots \\ &\quad \cdots + ({}^{g}_{(n-1)} \boldsymbol{\alpha}_n + {}^g_0 \boldsymbol{\omega}_n \times {}^{g}_{(n-1)} \boldsymbol{\omega}_n) \end{aligned} \quad (2.59)$$

The whole expression can be transformed to any arbitrary frame  $f$  by applying the rotation matrix from  $g$  to  $f$ ,  ${}^f R_g$ .

$$\begin{aligned} \frac{{}^g d}{dt} \left( \sum_{i=1}^n {}^{g}_{i-1} \boldsymbol{\omega}_i \right) &= \left( {}^f R_g \frac{{}^g d}{dt} \sum_{i=1}^n {}^{g}_{i-1} \boldsymbol{\omega}_i \right) \\ &= \frac{{}^g d}{dt} \left( {}^f R_g \sum_{i=1}^n {}^{g}_{i-1} \boldsymbol{\omega}_i \right) \\ &= \sum_{i=1}^n {}^{f}_{i-1} \boldsymbol{\alpha}_i + {}^f_g \boldsymbol{\omega}_i \times {}^{f}_{i-1} \boldsymbol{\omega}_i \end{aligned} \quad (2.60)$$

## Notations

In this article, the following symbols and notations are being used. Lowercase bold letters are used to indicate a vector. Dot accents on a vector indicate that is a time-derivative of the vector.

- $\mathbf{r}$ ,  $\mathbf{v}$  (or  $\dot{\mathbf{r}}$ ), and  $\mathbf{a}$  (or  $\ddot{\mathbf{r}}$ ) are the position, velocity, and acceleration vectors
- $\boldsymbol{\omega}$  and  $\boldsymbol{\alpha}$  (or  $\dot{\boldsymbol{\omega}}$ ) are angular velocity and angular acceleration vectors.

Capital letter  $A$ ,  $B$ ,  $C$ , and  $G$  are used to denote a reference frame. In examples where only  $B$  and  $G$  are used, the former indicates a rotating, body coordinate frame and the latter indicates the global-fixed, inertial body frame. When a reference frame is introduced, its origin point and three basis vectors are indicated. For example:

$$A(O\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3) \quad G(OXYZ) \quad B(o\hat{i}\hat{j}\hat{k})$$

Capital letters  $R$  is reserved for rotation matrix and transformation matrix. The right subscript on a rotation matrix indicates its departure frame, and the left superscript indicates its destination frame. For example:

$${}^B R_A = \text{rotation matrix from frame } A \text{ to frame } B$$

Left superscript is used to denote the coordinate frame in which the vector is expressed. For example, if a vector  $\mathbf{r}$  is expressed in an arbitrary coordinate frame  $A$  which has the basis vector  $(\hat{i}, \hat{j}, \hat{k})$ , it is written as

$$\begin{aligned} {}^A \mathbf{r} &= r_1 \hat{i} + r_2 \hat{j} + r_3 \hat{k} \\ &= \text{vector } \mathbf{r} \text{ expressed in frame } A \end{aligned}$$

Left subscript is used to denote the coordinate frame from which the vector is differentiated. For example,

$$\begin{aligned} {}^A \dot{\mathbf{r}} &= \frac{d}{dt} {}^A \mathbf{r} \\ &= \text{vector } \mathbf{r} \text{ expressed in frame } A \text{ and differentiated in frame } B \\ {}^{BB} \ddot{\mathbf{r}} &= \frac{d}{dt} \frac{d}{dt} {}^A \mathbf{r} \\ &= \text{vector } \mathbf{r} \text{ expressed in frame } A \text{ and differentiated twice in frame } B \end{aligned}$$

$${}_{CB}{}^A\ddot{\mathbf{r}} = \frac{{}^C d}{{}^C dt} \frac{{}^B d}{{}^B dt} {}^A \mathbf{r}$$

= vector  $\mathbf{r}$  expressed in frame  $A$  and differentiated in frame  $B$   
and differentiated again in frame  $C$

Right subscript is used to denote the coordinate frame to which the vector is referred, and left subscript is the coordinate frame to which it is related. For example, the angular velocity of frame  $A$  with respect to frame  $B$  is written as

$${}^B\boldsymbol{\omega}_A = \text{angular velocity vector of } A \text{ with respect to } B$$

If the coordinate frame in which the angular velocity is expressed is specified

$${}^C_B\boldsymbol{\omega}_A = \text{angular velocity vector of } A \text{ with respect to } B \\ \text{expressed in } C$$

## 2.6 Key Symbols

$\mathbf{a} = \ddot{\mathbf{r}}$	Acceleration vector
$\mathbf{v} = \dot{\mathbf{r}}$	Velocity vector
$\mathbf{r}$	Displacement vector
$\boldsymbol{\alpha}$	Angular acceleration vector
$\boldsymbol{\omega}$	Angular velocity vector
$\mathbf{F}$	Force vector
$m$	Mass scalar
$n$	Number of elements in the set of noninertial coordinate frames
$i$	$i$ th element in a set
$f, g$	Elements in the set of noninertial coordinate frames
$\hat{I}, \hat{J}, \hat{K}$	Orthonormal unit vectors of an inertial frame
$\hat{i}, \hat{j}, \hat{k}$	Orthonormal unit vectors of a noninertial frame
$O$	Origin point of a coordinate frame
$\dot{\theta}$	Angular speed of body (spin)
$\dot{\phi}$	Angular speed of body (precession)
$\psi$	Nutation angle, angle between spin vector and precession vector
$R$	Rotation transformation matrix
$\square$	Arbitrary vector
${}^A \mathbf{r}$	Vector $\mathbf{r}$ expressed in $A$ -frame
${}^A_B \dot{\mathbf{r}}$	Vector $\dot{\mathbf{r}}$ expressed in $A$ -frame and differentiated in $B$ -frame

- ${}_{CB}^A \dot{\mathbf{r}}$  Vector  $\ddot{\mathbf{r}}$  expressed in  $A$ -frame, first-differentiated in  $B$ -frame, and second-differentiated from  $C$ -frame
- ${}^A R_B$  Rotation transformation matrix from  $B$ -frame to  $A$ -frame

## References

- Baruh H (1999) Analytical dynamics. WCB/McGraw-Hill, Boston
- Jazar RN (2012) Derivative and coordinate frames *Nonlinear Eng* 1(1):25–34
- Jazar RN (2011) Advanced dynamics: rigid body, multibody, and aerospace applications. Wiley, Hoboken, pp 858–864
- Kane TR, Likins PW, Levinson DA (1983) Spacecraft dynamics. McGraw-Hill, New York, pp 1–90
- Kasdin NJ, Paley DA (2011) Engineering dynamics: a comprehensive introduction. Princeton University Press, Princeton
- Koetsier T (2007) Euler and kinematics. *Stud Hist Philos Math* 5:167–194
- Rao A (2006) Dynamics of particles and rigid bodies: a systematic approach. Cambridge University Press, Cambridge
- Tenenbaum RA (2004) Fundamentals of applied dynamics. Springer, New York
- Van Der Ha JC, Shuster MD (2009) A tutorial on vectors and attitude (Focus on Education). *Control Syst IEEE* 29(2):94–107
- Zipfel PH (2000) Modeling and simulation of aerospace vehicle dynamics. AIAA, Virginia

# Chapter 3

## Challenges in Exact Response of Piecewise Linear Vibration Isolator

Oleksandr Pogorilyi, Reza N. Jazar, and Pavel M. Trivailo

**Keywords** Piecewise linear systems • Nonlinear vibration isolator • Frequency response • Exact methods • Numerical methods

### 3.1 Introduction

Piecewise linear vibration isolator is a vibration isolator whose system has nonlinear geometric characteristics of stiffness and damping, where nonlinearity is a result of moving among finite number of linear segments.

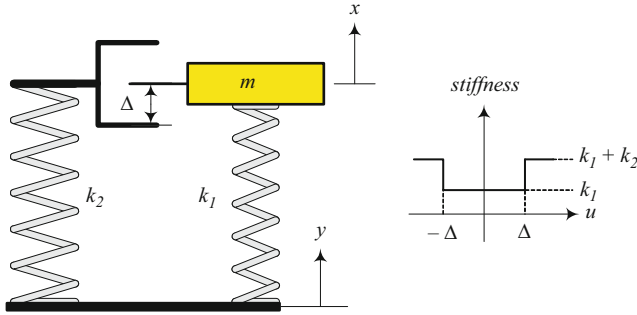
The scientific story of piecewise linear vibration isolator starts with the work of Den Hartog who analyzed an undamped bilinear stiffness system as is shown in Fig. 3.1 (Den Hartog and Mikina 1932). He showed that the exact solution for the frequency response of the system can be obtained by solving two transcendental equations. Solving those equations, Den Hartog presented a set of frequency response plots for a few different values of stiffness. His calculations were confirmed by experimental results of Jacobsen and Jaspersen (1935) who explored the bilinear stiffness system and received results close to the Den Hartog mathematical results.

The next step in study of piecewise linear vibration isolator was analysis of a single degree of freedom system with clearance in the viscous damping as shown in Fig. 3.2. Gurtin presented an approximate method for solving of the system (Gurtin 1961). The method is based on finding an equivalent viscous damper.

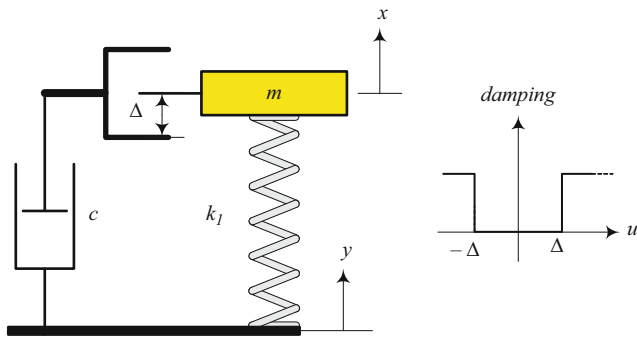
The mathematical modeling and analysis of piecewise linear systems have been grown in parallel to their application. Aizerman tried to ground the mathematical

---

O. Pogorilyi • R.N. Jazar (✉) • P.M. Trivailo  
School of Aerospace, Mechanical, and Manufacturing Engineering,  
RMIT University, Melbourne, VIC, Australia  
e-mail: [s3373806@student.rmit.edu.au](mailto:s3373806@student.rmit.edu.au); [reza.jazar@rmit.edu.au](mailto:reza.jazar@rmit.edu.au); [pavel.trivailo@rmit.edu.au](mailto:pavel.trivailo@rmit.edu.au)



**Fig. 3.1** A dual rate bilinear vibrating system



**Fig. 3.2** A vibrating system with gap-damper engagement

foundation for the piecewise linear systems (Aizerman and Gantmakher 1957, 1958; Aizerman and Lur'e 1963) and developed a number of methods to deal with these systems. Also, Fleishman (1965) analyzed a system of coupled piecewise linear equations from a control viewpoint and described the required conditions for a periodic response.

The simplicity of the piecewise linear vibrating systems helped scientists to perform many experiments. They not only discovered validation of the mathematical predictions, but they also found several unexpected behaviors that were not predictable by the simple mathematical models. Hayashi (1964) did a notable work in application direction. His results show that great complication is to be expected even in simple modeling cases. Masri (1965, 1978), Masri and Stott (1978) used a general one DOF vibrating system to explain the difficulties that introduce in formulation of the steady state motion. Marsi used an experimental device to analyze the sensitivity of the physical system for some of the parameters involved. Watabene (1978) used the method of Den Hartog to plot the frequency response of an equivalent bilinear mass-spring vibrator to a base excited cantilever beam with clearance. He examined the effect of the amplitude of harmonic excitation in his plots.

Iwan (1973) tried to use approximation methods using averaged periodic solutions. Nguyen et al. (1986a,b) used numerical integration methods and described the time domain behavior of the system.

Rosenberg (1966) developed the mathematical foundations for the existence of steady state solution for equation  $\ddot{x} + f(x) = p \cdot g(t) = p \cdot g(t + T)$ , where  $g(t)$  is a cosine-like periodic function. He showed that the necessary condition for the existence of steady state periodic solution of the equation is that  $f(x)$  be everywhere analytic, odd, monotonically increasing function  $xf(x) > 0$  for  $x = 0$ . The piecewise linear system without damping is of this type.

Natsiavas (1989), Natsiavas and Gonzalez (1992) developed a stability analysis for asymmetric piecewise linear force excited system using methodology that was firstly found by Den Hartog and received system of transcendental equations that can be used to calculate amplitudes of relative displacement analytically. Narimani et al. (2004a) predicted the frequency response of piecewise linear system using perturbation methods. They experimentally verified the perturbation results and discovered a new phenomenon called the frequency island (Narimani et al. 2004b). Deshpande et al. (2006) optimized the secondary isolator employing the RMS method (Jazar et al. 2003; Alkhatib et al. 2004; Christopherson and Jazar 2005; Jazar et al. 2006a). Jump avoidance conditions of the system were investigated by Jazar et al. (2006b, 2007). The importance and application of the system were reviewed by several investigators (Jazar and Golnaraghi 2002; Marzbani et al. 2012, 2013).

Schulman (1983) is the first scientist who discovered chaos in piecewise linear systems. Later this phenomenon was analyzed in vast systems of this type.

The simplest practical model of a piecewise linear vibration isolator is a system with bilinear stiffness and damping characteristics as is shown in Fig. 3.3. The first spring and damper that are directly attached to  $m$  is called “Primary suspension” and the second stage, which is effective beyond the clearance amplitude  $\Delta$ , is called “Secondary suspension.” The clearance  $\Delta$  generates a switch to engage the secondary suspension. This represents a sudden change in the system properties

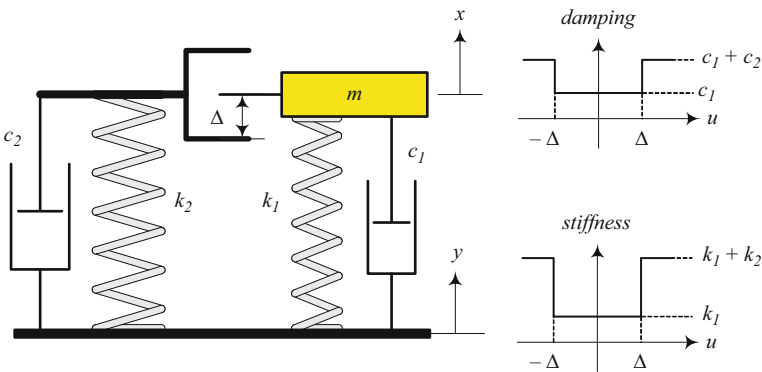


Fig. 3.3 Mechanical model of the piecewise linear system with symmetric constraints



which provides a hard nonlinearity of the piecewise linear system. In this article the examined system is subjected to a periodic base excitation  $y(t)$  with period  $T > 0$ .

The equations governing the motion of the system could be written as

$$m\ddot{x} + g_1(x, \dot{x}) = f_1(y, \dot{y}) \quad (3.1)$$

where  $g_1(x, \dot{x})$  and  $f_1(y, \dot{y})$  are piecewise linear functions presenting sudden changing characteristics of the system and sudden changing excitation, respectively:

$$g_1(x, \dot{x}) = \begin{cases} (c_1 + c_2)\dot{x} + (k_1 + k_2)x - k_2\Delta & x - y > \Delta \\ c_1\dot{x} + k_1x & |x - y| < \Delta \\ (c_1 + c_2)\dot{x} + (k_1 + k_2)x + k_2\Delta & x - y < -\Delta \end{cases} \quad (3.2)$$

$$f_1(y, \dot{y}) = \begin{cases} (c_1 + c_2)\dot{y} + (k_1 + k_2)y & x - y > \Delta \\ c_1\dot{y} + k_1y & |x - y| < \Delta \\ (c_1 + c_2)\dot{y} + (k_1 + k_2)y & x - y < -\Delta \end{cases} \quad (3.3)$$

The equation of motion for the system shown in Fig. 3.3 may also be written in a nondimensional form:

$$\begin{cases} \ddot{z} + 2\xi_2\omega_2\dot{z} + \omega_2^2z = w\omega^2 \sin(\omega t - \varphi) + \omega_3^2 & z > 1 \\ \ddot{z} + 2\xi_1\omega_1\dot{z} + \omega_1^2z = w\omega^2 \sin(\omega t - \varphi) & |z| < 1 \\ \ddot{z} + 2\xi_2\omega_2\dot{z} + \omega_2^2z = w\omega^2 \sin(\omega t - \varphi) - \omega_3^2 & z < -1 \end{cases} \quad (3.4)$$

where

$$\begin{aligned} z &= \frac{u}{\Delta} & w &= \frac{Y}{\Delta} & u &= x - y \\ \omega_1^2 &= \frac{k_1}{m} & \omega_2^2 &= \frac{k_1 + k_2}{m} \\ \xi_1 &= \frac{c_1}{2\sqrt{k_1m}} & \xi_2 &= \frac{c_1 + c_2}{2\sqrt{(k_1 + k_2)m}} \end{aligned} \quad (3.5)$$

We seek the frequency response of the system (3.4) by developing and detecting its exact steady state time response.

### 3.2 Exact Solution

As long as the relative displacement of  $m$  is less than  $\Delta$ , the system is a one degree-of-freedom base exciting system with well-known time and frequency responses (Jazar 2013, 2014). However, there is no general closed form solution for the system when the relative displacement exceeds  $\Delta$ . In order to determine the frequency response of the system, we search for possible steady state time response of the system and detect its maximum amplitude. Repeating this method ends up with a series of amplitudes for different values of excitation frequencies.

A steady state periodic response of the system should look like Fig. 3.4. To ensure that the secondary suspension is engaged, we seek a periodic solution of the system (3.4) with the following initial conditions:

$$z(0) = 1 \quad \dot{z}(0) > 0 \tag{3.6}$$

To set the time axis to begin at this condition, we introduce a phase lag  $\varphi$  in the excitation function  $y$ .

$$y = Y \sin(\omega t - \varphi) \tag{3.7}$$

Furthermore, we assume that there is a steady state periodic response with exactly the same frequency  $\omega$  as the excitation frequency. We also assume that in the first half of the period, the response passes through  $z = 1$  with  $\dot{z}(0) < 0$  at a time  $t = t_1$ ,  $0 \leq t_1 \leq 2\pi/\omega$ .

Due to these assumptions, the solution will be  $|z(t)| = 1$  at  $t = 0$ ,  $t = t_1$ ,  $t = \pi/\omega$ ,  $t = \pi/\omega + t_1$ , and  $t = 2\pi/\omega$  in the period starting at  $t = 0$ . Let us call the solution in domain  $z > 1$  as  $z_2(t)$  and the solution in domain  $|z| < 1$  as  $z_1(t)$  (see Fig. 3.4). The continuity and compatibility of  $z_1(t)$  and  $z_2(t)$  would be satisfied by imposing the following conditions:

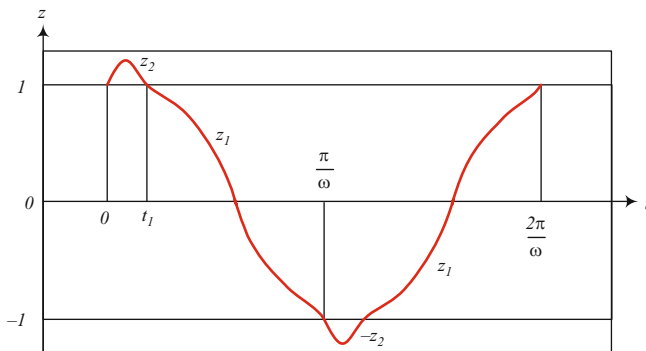


Fig. 3.4 A steady state periodic response of the system

$$\begin{aligned}
z_2(0) = 1 & & \dot{z}_2(0) = v_2 & & z_2(t_1) = 1 & & \dot{z}_2(t_1) = v_1 \\
z_1(t_1) = 1 & & \dot{z}_1(t_1) = v_1 & & z_1\left(\frac{\pi}{\omega}\right) = -1 & & \dot{z}_1\left(\frac{\pi}{\omega}\right) = -v_2
\end{aligned} \quad (3.8)$$

There is no however a one-to-one relationship between the input amplitude  $Y$  and the output response  $z$ . For some values of  $Y$  there may be no such solutions, while for the other values of  $Y$  there may be several.

The underdamped solutions  $z_2(t)$  and  $z_1(t)$  are:

$$\begin{aligned}
z_1(t) = e^{-\xi_1\omega_1 t} (A_1 \sin(\omega_{d_1} t) + B_1 \cos(\omega_{d_1} t)) \\
+ C_1 \sin(\omega t) + D_1 \cos(\omega t)
\end{aligned} \quad (3.9)$$

$$\begin{aligned}
z_2(t) = e^{-\xi_2\omega_2 t} (A_2 \sin(\omega_{d_2} t) + B_2 \cos(\omega_{d_2} t)) \\
+ C_2 \sin(\omega t) + D_2 \cos(\omega t) + \frac{\omega_3^2}{\omega_2^2}
\end{aligned} \quad (3.10)$$

and therefore,

$$\begin{aligned}
\dot{z}_1(t) = -\xi_1\omega_1 e^{-\xi_1\omega_1 t} (A_1 \sin(\omega_{d_1} t) + B_1 \cos(\omega_{d_1} t)) \\
+ \omega_{d_1} e^{-\xi_1\omega_1 t} (A_1 \cos(\omega_{d_1} t) - B_1 \sin(\omega_{d_1} t)) \\
+ \omega (C_1 \cos(\omega t) - D_1 \sin(\omega t))
\end{aligned} \quad (3.11)$$

$$\begin{aligned}
\dot{z}_2(t) = -\xi_2\omega_2 e^{-\xi_2\omega_2 t} (A_2 \sin(\omega_{d_2} t) + B_2 \cos(\omega_{d_2} t)) \\
+ \omega_{d_2} e^{-\xi_2\omega_2 t} (A_2 \cos(\omega_{d_2} t) - B_2 \sin(\omega_{d_2} t)) \\
+ \omega (C_2 \cos(\omega t) - D_2 \sin(\omega t))
\end{aligned} \quad (3.12)$$

where

$$\omega_{d_2} = \omega_2 \sqrt{1 - \xi_2^2} \quad \omega_{d_1} = \omega_1 \sqrt{1 - \xi_1^2} \quad (3.13)$$

The coefficients  $A_1$ ,  $B_1$ ,  $A_2$ ,  $B_2$  will be found by imposing the initial and compatibility conditions (3.8). The coefficients  $C_1$ ,  $D_1$ ,  $C_2$ ,  $D_2$  depend on the forcing function and will be found by collecting the coefficients of  $\sin(\omega t)$  and  $\cos(\omega t)$  in  $z_1(t)$  and  $z_2(t)$

$$C_1 = Q_1 \cos \varphi - Q_2 \sin \varphi \quad (3.14)$$

$$D_1 = -Q_2 \cos \varphi - Q_1 \sin \varphi \quad (3.15)$$

$$C_2 = Q_3 \cos \varphi - Q_4 \sin \varphi \quad (3.16)$$

$$D_2 = -Q_4 \cos \varphi - Q_3 \sin \varphi \quad (3.17)$$

where

$$Q_1 = \frac{(\omega_1^2 - \omega^2)\omega^2}{(\omega_1^2 - \omega^2)^2 + (2\xi_1\omega\omega_1)^2} \omega \quad (3.18)$$

$$Q_2 = \frac{2\xi_1\omega^3\omega_1}{(\omega_1^2 - \omega^2)^2 + (2\xi_1\omega\omega_1)^2} \omega \quad (3.19)$$

$$Q_3 = \frac{(\omega_2^2 - \omega^2)\omega^2}{(\omega_2^2 - \omega^2)^2 + (2\xi_2\omega\omega_2)^2} \omega \quad (3.20)$$

$$Q_4 = \frac{2\xi_2\omega^3\omega_2}{(\omega_2^2 - \omega^2)^2 + (2\xi_2\omega\omega_2)^2} \omega \quad (3.21)$$

Imposing the eight boundary conditions (3.8) on Eqs. (3.9)–(3.12), produces eight transcendental equations with eight unknowns  $A_2$ ,  $B_2$ ,  $A_1$ ,  $B_1$ ,  $t_1$ ,  $v_1$ ,  $v_2$ , and  $\varphi$ :

$$B_2 + D_2 + \frac{\omega_3^2}{\omega_2^2} = 1 \quad (3.22)$$

$$-\xi_2\omega_2 B_2 + \omega_{d_2} A_2 + C_2\omega = v_2 \quad (3.23)$$

$$\begin{aligned} & e^{-\xi_2\omega_2 t_1} (A_2 \sin(\omega_{d_2} t_1) + B_2 \cos(\omega_{d_2} t_1)) \\ & + C_2 \sin(\omega t_1) + D_2 \cos(\omega t_1) + \frac{\omega_3^2}{\omega_2^2} = 1 \end{aligned} \quad (3.24)$$

$$\begin{aligned} & -\xi_2\omega_2 e^{-\xi_2\omega_2 t_1} (A_2 \sin(\omega_{d_2} t_1) + B_2 \sin(\omega_{d_2} t_1)) \\ & + \omega_{d_2} e^{-\xi_2\omega_2 t_1} (A_2 \cos(\omega_{d_2} t_1) - B_2 \sin(\omega_{d_2} t_1)) \\ & + \omega (C_2 \cos(\omega t_1) - D_2 \sin(\omega t_1)) = v_1 \end{aligned} \quad (3.25)$$

$$\begin{aligned} & e^{-\xi_1\omega_1 t_1} (A_1 \sin(\omega_{d_1} t_1) + B_1 \cos(\omega_{d_1} t_1)) \\ & + C_1 \sin(\omega t_1) + D_1 \cos(\omega t_1) = 1 \end{aligned} \quad (3.26)$$

$$\begin{aligned} & -\xi_1\omega_1 e^{-\xi_1\omega_1 t_1} (A_1 \sin(\omega_{d_1} t_1) + B_1 \cos(\omega_{d_1} t_1)) \\ & + \omega_{d_1} e^{-\xi_1\omega_1 t_1} (A_1 \cos(\omega_{d_1} t_1) - B_1 \sin(\omega_{d_1} t_1)) \\ & + \omega (C_1 \cos(\omega t_1) - D_1 \sin(\omega t_1)) = v_1 \end{aligned} \quad (3.27)$$

$$e^{-\xi_1\omega_1 \frac{\pi}{\omega}} (A_1 \sin(\omega_{d_1} \frac{\pi}{\omega}) + B_1 \cos(\omega_{d_1} \frac{\pi}{\omega})) - D_1 = -1 \quad (3.28)$$

$$\begin{aligned} & -\xi_1\omega_1 e^{-\xi_1\omega_1 \frac{\pi}{\omega}} (A_1 \sin(\omega_{d_1} \frac{\pi}{\omega}) + B_1 \cos(\omega_{d_1} \frac{\pi}{\omega})) \\ & + \omega_{d_1} e^{-\xi_1\omega_1 \frac{\pi}{\omega}} (A_1 \cos(\omega_{d_1} \frac{\pi}{\omega}) - B_1 \sin(\omega_{d_1} \frac{\pi}{\omega})) \\ & - C_1\omega = -v_2 \end{aligned} \quad (3.29)$$

The parameters  $B_2$  and  $A_2$  may be found from the Eqs. (3.22) and (3.23),

$$B_2 = 1 - D_2 - \frac{\omega_3^2}{\omega_2^2} \quad (3.30)$$

$$A_2 = \frac{v_2 + \xi_2 \omega_2 B_2 - C_2 \omega}{\omega_{d_2}} = -Q_5 - \frac{\xi_2}{\omega_2} \frac{\omega_3^2}{\omega_{d_2}} \quad (3.31)$$

and the parameters  $B_1$  and  $A_1$  from the Eqs. (3.28) and (3.29)

$$B_1 = e^{\frac{\xi_1 \omega_1 \pi}{\omega}} \left( -Q_6 \sin\left(\frac{\omega_{d_1} \pi}{\omega}\right) + (D_1 - 1) \cos\left(\frac{\omega_{d_1} \pi}{\omega}\right) \right) \quad (3.32)$$

$$A_1 = e^{\frac{\xi_1 \omega_1 \pi}{\omega}} \left( Q_6 \cos\left(\frac{\omega_{d_1} \pi}{\omega}\right) + (D_1 - 1) \sin\left(\frac{\omega_{d_1} \pi}{\omega}\right) \right) \quad (3.33)$$

where

$$Q_5 = \frac{\xi_2 \omega_2 (D_2 - 1) + C_2 \omega - v_2}{\omega_{d_2}} \quad (3.34)$$

$$Q_6 = \frac{\xi_1 \omega_1 (D_1 - 1) + C_1 \omega - v_2}{\omega_{d_1}} \quad (3.35)$$

Now, substitution of  $A_1$ ,  $B_1$ ,  $A_2$ , and  $B_2$  in (3.24)–(3.27) produces four equations for  $v_1$ ,  $v_2$ ,  $t_1$ , and  $\varphi$

$$\begin{aligned} & Q_7 \sin(\omega_{d_2} t_1) + Q_8 \cos(\omega_{d_2} t_1) \\ & + C_2 \sin(\omega t_1) + D_2 \cos(\omega t_1) + \frac{\omega_3^2}{\omega_2^2} = 1 \end{aligned} \quad (3.36)$$

$$\begin{aligned} & -Q_9 \sin(\omega_{d_2} t_1) + Q_{10} \cos(\omega_{d_2} t_1) \\ & + \omega (C_2 \cos(\omega t_1) - D_2 \sin(\omega t_1)) = v_1 \end{aligned} \quad (3.37)$$

$$\begin{aligned} & Q_{11} \sin(\omega_{d_1} t_1) + Q_{12} \cos(\omega_{d_1} t_1) \\ & + C_1 \sin(\omega t_1) + D_1 \cos(\omega t_1) = 1 \end{aligned} \quad (3.38)$$

$$\begin{aligned} & -Q_{13} \sin(\omega_{d_1} t_1) + Q_{14} \cos(\omega_{d_1} t_1) \\ & + \omega (C_1 \cos(\omega t_1) - D_1 \sin(\omega t_1)) = v_1 \end{aligned} \quad (3.39)$$

where

$$Q_7 = A_2 e^{-\xi_2 \omega_2 t_1} \quad Q_8 = B_2 e^{-\xi_2 \omega_2 t_1} \quad (3.40)$$

$$Q_9 = e^{-\xi_2 \omega_2 t_1} (\xi_2 \omega_2 A_2 + \omega_{d_2} B_2) \quad (3.41)$$

$$Q_{10} = e^{-\xi_2 \omega_2 t_1} (\omega_{d_2} A_2 - \xi_2 \omega_2 B_2) \quad (3.42)$$

$$Q_{11} = A_1 e^{-\xi_1 \omega_1 t_1} \quad Q_{12} = B_1 e^{-\xi_1 \omega_1 t_1} \quad (3.43)$$

$$Q_{13} = e^{-\xi_1 \omega_1 t_1} (\xi_1 \omega_1 A_1 + \omega_{d_1} B_1) \quad (3.44)$$

$$Q_{14} = e^{-\xi_1 \omega_1 t_1} (\omega_{d_1} A_1 + \xi_1 \omega_1 B_1) \quad (3.45)$$

Eliminating  $v_1$  between Eqs. (3.37) and (3.39) reduces the number of equations to three to find  $v_2$ ,  $t_1$ , and  $\varphi$ .

$$\begin{aligned} & Q_7 \sin(\omega_{d_2} t_1) + Q_8 \cos(\omega_{d_2} t_1) \\ & + C_2 \sin(\omega t_1) + D_2 \cos(\omega t_1) + \frac{\omega_3^2}{\omega_2^2} = 1 \end{aligned} \quad (3.46)$$

$$\begin{aligned} & - Q_9 \sin(\omega_{d_2} t_1) + Q_{10} \cos(\omega_{d_2} t_1) + Q_{13} \sin(\omega_{d_1} t_1) - Q_{14} \cos(\omega_{d_1} t_1) \\ & + \omega ((C_2 - C_1) \cos(\omega t_1) - (D_2 - D_1) \sin(\omega t_1)) = 0 \end{aligned} \quad (3.47)$$

$$\begin{aligned} & Q_{11} \sin(\omega_{d_1} t_1) + Q_{12} \cos(\omega_{d_1} t_1) \\ & + C_1 \sin(\omega t_1) + D_1 \cos(\omega t_1) = 1 \end{aligned} \quad (3.48)$$

The parameter  $v_2$  is embedded in  $A_2$ ,  $B_1$ , and  $A_1$ . Therefore,  $Q_5$  to  $Q_7$  and  $Q_9$  to  $Q_{14}$  are functions of  $v_2$ . Substituting  $Q_{11}$  and  $Q_{12}$  in Eq. (3.48) produces a linear equation of  $Q_6$

$$Q_{15} Q_6 + Q_{16} = Q_{17} \quad (3.49)$$

where

$$Q_{15} = -e^{\xi_1 \omega_1 \left(\frac{\pi}{\omega} - t_1\right)} \sin\left(\frac{\pi - \omega t_1}{\omega} \omega_{d_1}\right) \quad (3.50)$$

$$\begin{aligned} Q_{16} = & e^{-\xi_1 \omega_1 t_1} (C_1 \sin(\omega t_1) + D_1 \cos(\omega t_1)) \\ & + e^{\xi_1 \omega_1 \left(\frac{\pi}{\omega} - t_1\right)} (D_1 - 1) \cos\left(\frac{\pi - \omega t_1}{\omega} \omega_{d_1}\right) \end{aligned} \quad (3.51)$$

$$Q_{17} = D_1 e^{\frac{\pi}{\omega} \xi_1 \omega_1} \quad (3.52)$$

$Q_6$  is solvable for  $v_2$

$$v_2 = \omega C_1 + (D_1 - 1) \xi_1 \omega_1 - Q_6 \omega_{d1} \tag{3.53}$$

and therefore, eliminating  $v_2$  in Eqs. (3.46) and (3.47) reduces the number of transcendental equations to two to determine  $t_1$  and  $\varphi$  for a set of given system and excitation.

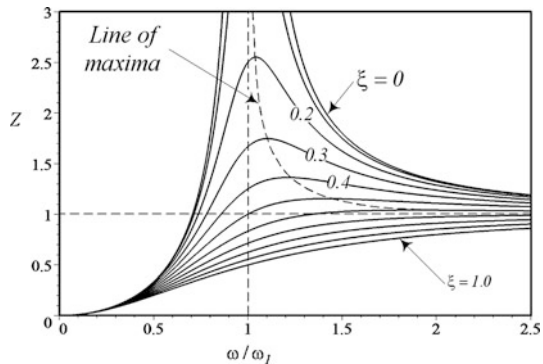
### 3.3 Time Response

If we were able to solve Eqs. (3.46)–(3.48) analytically, then we could complete the solution  $z_2(t)$  and  $z_1(t)$ . In that case the maximum value of  $z_2(t)$  would be the steady state amplitude. To determine the steady state response of the system, we need to numerically determine the parameters.

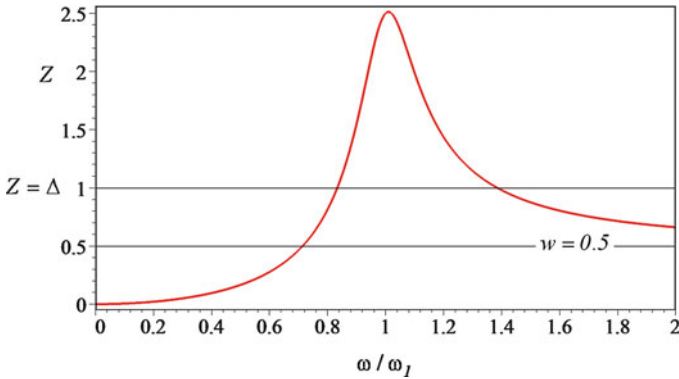
The primary suspension’s frequency response is

$$Z_1 = \sqrt{C_1^2 + D_1^2} = \frac{\omega^2}{\sqrt{(\omega_1^2 - \omega^2)^2 + (2\xi_1 \omega \omega_1)^2}} w \tag{3.54}$$

and shows in Fig. 3.5 for  $w = 1$ . The area above the line  $Z = 1$  indicates the engagement of the secondary suspension. The curves will approach the line  $Z = w$  asymptotically. If  $w < 1$ , then there would be a frequency span in which the secondary will engage, and there would be no engagement outside the span. If  $w > 1$ , then the engagement will begin at a frequency and never let go. As a result, the curves above  $Z = 1$  will change because of the sudden change of stiffness and damping.



**Fig. 3.5** The frequency response for  $Z = \frac{U}{\Delta}$



**Fig. 3.6** The frequency response curve of the primary for  $\omega_1 = 1, \delta = 1, \xi_1 = 0.1$

Without losing generality we may assume

$$\omega_1 = 1 \tag{3.55}$$

and set

$$w = 0.5 < 1 \quad \xi_1 = 0.1 \tag{3.56}$$

and plot the associated frequency response curve of the primary in Fig. 3.6. The amplitude of the primary will be greater than  $Z_1 = 1$  when the excitation frequency is in the following range.

$$0.8337115427 < \omega < 1.385012057 \tag{3.57}$$

To express the method and determine a sample of steady state time response, let us set the secondary suspension parameters as

$$\omega_2 = \sqrt{2} \quad \xi_2 = 0.1 \tag{3.58}$$

Substituting the system characteristics we should solve the two equations of (3.46) and (3.47)  $t_1$  and  $\varphi$ . Figure 3.7 illustrates the implicit plot of the equations, with a unique solution at

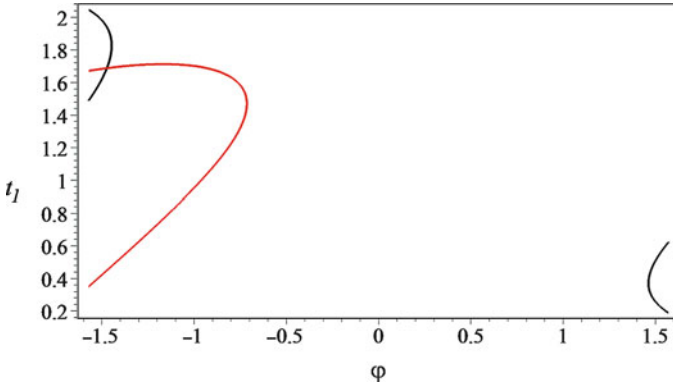
$$\varphi = -1.472642330 \quad t_1 = 1.688156972 \tag{3.59}$$

Employing  $t_1$  and  $\varphi$  we calculate

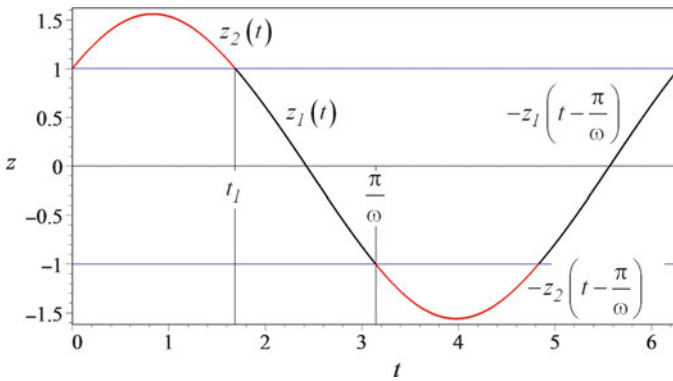
$$v_1 = -1.154853554 \quad v_2 = 1.218400685$$

and determine  $z_2(t)$  and  $z_1(t)$  to plot the steady state solution as shown in Fig. 3.8.





**Fig. 3.7** The implicit plot of the two transcendental equations (3.46) and (3.47) as functions of  $t_1$  and  $\varphi$  for  $\omega_2 = \sqrt{2}$ ,  $\xi_2 = 0.1$ ,  $\omega_1 = 1$ ,  $w = 0.5$ ,  $\xi_1 = 0.1$



**Fig. 3.8** A sample of the steady response of the system for  $\omega_2 = \sqrt{2}$ ,  $\xi_2 = 0.1$ ,  $\omega_1 = 1$ ,  $w = 0.5$ ,  $\xi_1 = 0.1$

### 3.4 Challenges and Unsolved Problems

Regardless of the simple configuration and arrangement of the piecewise linear vibration isolator, determination of the frequency and steady state time responses of the system is not completely covered to date. There are several challenges and unsolved problems that we review in this section.

#### 3.4.1 Multiple Solutions

The eight conditions (3.8) are not enough to guarantee a single solution. Depending on the parameters, we may get multiple possible steady state responses. This

multiplicity is different than the existence of multiple solutions for some values of frequency in the frequency response of nonlinear vibrating systems. As an example, let us consider a system with the following characteristics

$$\omega_1 = 1 \quad \xi_1 = 0.1 \quad \omega_2 = \sqrt{17} \quad \xi_2 = 0.1 \quad (3.60)$$

and the excitation with

$$\omega = 1 \quad w = 3 \quad (3.61)$$

Figure 3.9 illustrates the intersection of Eqs. (3.46) and (3.47) for the possible solutions of  $t_1$  and  $\varphi$ .

The intersections provide us with two sets of solutions:

$$\varphi = -1.248353815 \quad t_1 = 0.9049692487 \quad (3.62)$$

$$\varphi = -1.314128509 \quad t_1 = 2.229699384 \quad (3.63)$$

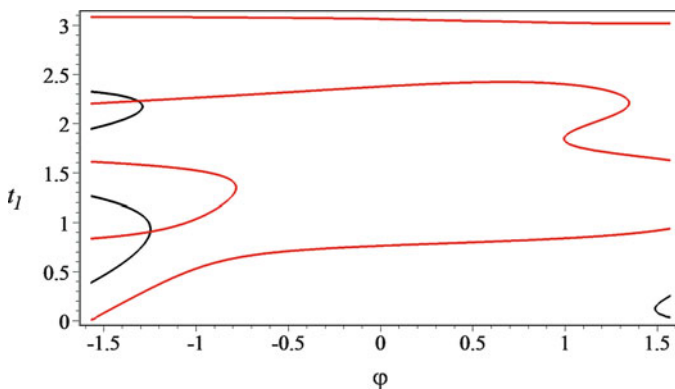
The steady response of the system for solution (3.62) is illustrated in Fig. 3.10, and for solution (3.63) illustrated in Fig. 3.11.

Considering

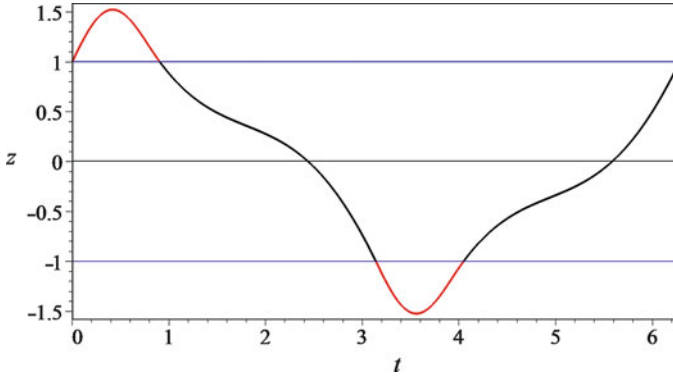
$$\omega_1 = 1 \quad \xi_1 = 0.1 \quad \omega_2 = \sqrt{2} \quad \xi_2 = 0.1 \quad (3.64)$$

we find the illustration of Eqs. (3.46) and (3.47) as shown in Fig. 3.12, with a unique solution for the  $t_1$  and  $\varphi$ .

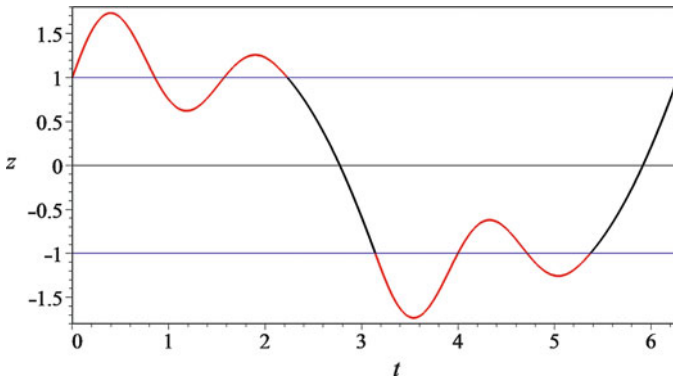
$$\varphi = -0.6226982402 \quad t_1 = 2.620975511$$



**Fig. 3.9** The implicit plot of the Eqs. (3.46) and (3.47) for  $\omega = 1, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{17}, \xi_2 = 0.1$



**Fig. 3.10** The steady response of the system for  $\varphi = -1.248353815, t_1 = .9049692487$

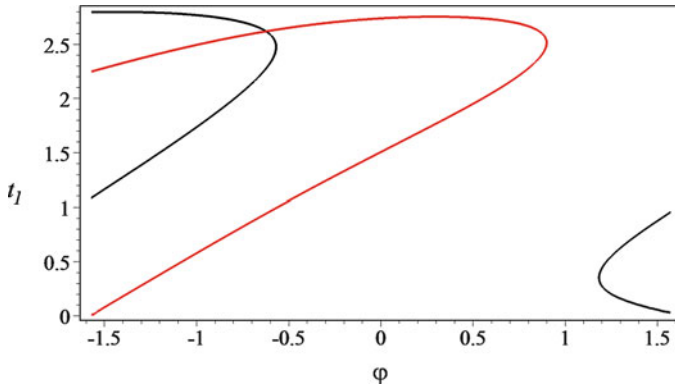


**Fig. 3.11** The steady response of the system for  $\varphi = -1.314128509, t_1 = 2.229699384$

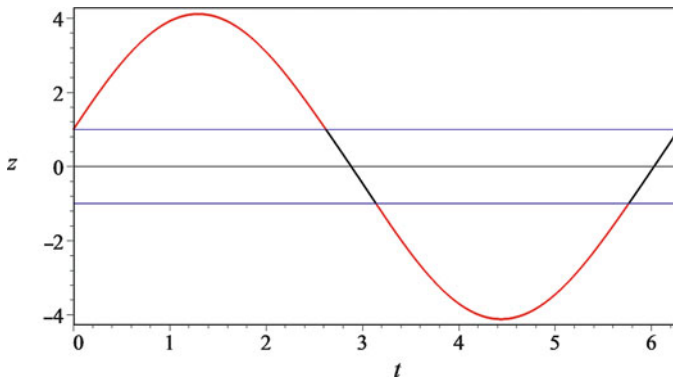
The associated steady state response of the system is shown in Fig. 3.13.

Comparison of Figs. 3.13 and 3.10, 3.11 suggests some open problems to be solved in the future. It indicates that there would be a periodic bifurcation point in the interval  $\sqrt{2} < \omega_2 < \sqrt{17}$  for  $\omega_1 = 1, \xi_1 = 0.1, \xi_2 = 0.1, w = 3, \omega = 1$ . Any other combination may have the same feature as well. Double hitting or multiple hitting may also happen without violating the imposed boundary conditions. Searching only for the responses with single hitting needs more conditions to be introduced. In case of multiple hitting, such as the one in Fig. 3.11, it must be discussed to clarify if the first or the highest amplitude should be recorded.

Furthermore, when plot of Eqs. (3.46) and (3.47) show two intersections, one of them must be associated with a stable and the other one associated with an unstable solution. Based on the nonlinear vibration understanding, the solution with higher amplitude must be the stable solution, and the lower amplitude be associated with unstable solution. In case of three possible amplitudes, the middle one is unstable.



**Fig. 3.12** The implicit plot of the Eqs. (3.46) and (3.47) for  $\omega = 1, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{2}, \xi_2 = 0.1$



**Fig. 3.13** The steady response of the system for  $\omega = 1, w = 3$  and  $\omega_1 = 1, \xi_1 = 0.1, \omega_2 = \sqrt{2}, \xi_2 = 0.1$

However, the stability of the solutions cannot be determined by analyzing one period of the steady state solutions. It must be done by a stability analysis.

### 3.4.2 Range of Nonlinearity

Although the system is piecewise linear, it shows the behavior of nonlinear systems because of the sudden change in stiffness and damping. One of the common features of nonlinear vibrating systems is the jump phenomenon. The jump is in general a harmful phenomenon in real systems. It is the reason why jump avoidance design must be taken (Jazar et al. 2006a). In the piecewise linear vibrating system, the jump may appear when the secondary suspension is still engaged while the excitation

frequency is beyond the upper limit of the frequency interval of the line  $Z_1 = 1$  intersecting the frequency response of the primary. As an example, consider a system with

$$\omega_1 = 1 \quad \xi_1 = 0.1 \quad \omega_2 = \sqrt{17} \quad \xi_2 = 0.1 \quad (3.65)$$

and the excitation with

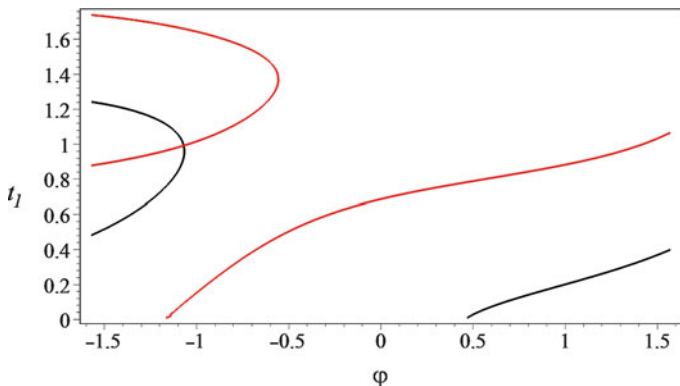
$$\omega = 1.4 \quad w = 3 \quad (3.66)$$

The excitation frequency is higher than upper limit of (3.57). Therefore, the system should not ever engage with the secondary suspension. We expect that if such an engagement happens, the system let the secondary go and settle down on a steady state vibration supported only by the primary suspension. However, Fig. 3.14 indicates that there is a possible solution for system to vibrate on a steady state condition while the secondary suspension engages. The response is depicted in Fig. 3.15.

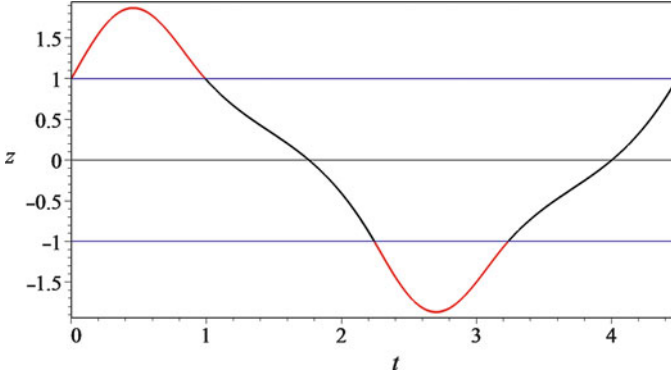
Theoretically, variation of  $\omega$  with some small step and solution of Eqs. (3.46)–(3.48) allow us to calculate the steady state amplitude of the system. In the case of nonlinear systems the highest value of the frequency for which a steady state response happens cannot be predicted exactly. The current practical way is to increase the excitation frequency  $\omega$  little by little until no solution can be achieved from Eqs. (3.46) and (3.47). Additionally, we should monitor the tangent slope of

$$\frac{dZ_2}{d\omega} \simeq \frac{Z_{2n} - Z_{2n-1}}{\omega_n - \omega_{n-1}} \quad (3.67)$$

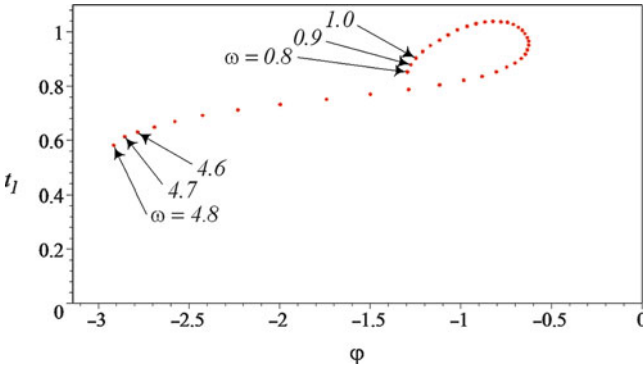
The point where the slope of the tangent to the response curve is almost vertical the jump will occur. However, when the slope gets close to vertical, the sensitivity of the solution to the variation of  $\omega$  becomes higher. Therefore, the step size of  $\omega$  must become finer. One set of solution of Eqs. (3.46) and (3.47) for different values of



**Fig. 3.14** The implicit plot of the Eqs. (3.46) and (3.47) for  $\omega = 1.4$ ,  $w = 3$  and  $\omega_1 = 1$ ,  $\xi_1 = 0.1$ ,  $\omega_2 = \sqrt{17}$ ,  $\xi_2 = 0.1$



**Fig. 3.15** The steady response of the system for  $\omega = 1$ ,  $w = 3$  and  $\omega_1 = 1.4$ ,  $\xi_1 = 0.1$ ,  $\omega_2 = \sqrt{2}$ ,  $\xi_2 = 0.1$



**Fig. 3.16** One set of solution of Eqs. (3.46) and (3.47) for  $w = 3$ ,  $\omega_1 = 1$ ,  $\xi_1 = 0.1$ ,  $\omega_2 = \sqrt{17}$ ,  $\xi_2 = 0.1$  and varying  $\omega$

$\omega$  is shown in Fig. 3.16. No solution for  $\omega < 0.8$  or  $\omega > 4.8$  could be found. At  $\omega = 4.8$  the slope of the frequency response curve is close to vertical.

$$\frac{dZ_2}{d\omega} \simeq \frac{6.21014567 - 7.928497661}{0.1} = -17.184 \quad (3.68)$$

$$\arctan(-17.184) \simeq -1.5127 \text{ rad} \simeq -86.673 \text{ deg} \quad (3.69)$$

### 3.4.3 Asymmetric Systems

The symmetric piecewise linear vibration isolator is a limited portion of the piecewise linear isolators. The system may easily become asymmetric in practice. Any change of mass will make the positive and negative gap sizes unequal. The

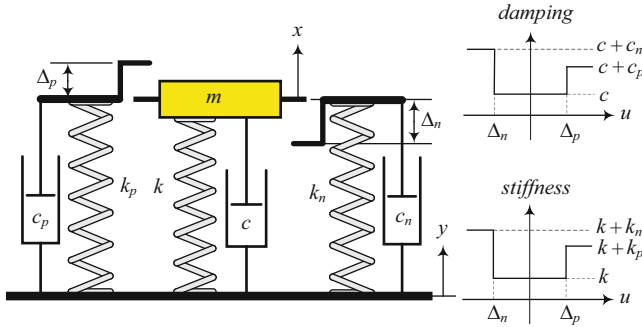


Fig. 3.17 The mathematical model of an asymmetric piecewise linear vibration isolator

spring may behave different in compression and extension. Similarly, real dampers resist differently in bound and rebound.

Figure 3.17 illustrates the mathematical model of an asymmetric piecewise linear vibration isolator. Determination of the frequency and steady state time responses of the system is more complicated, as the number of boundary conditions and transcendental equations increases.

The equations governing the motion of the system could be written as

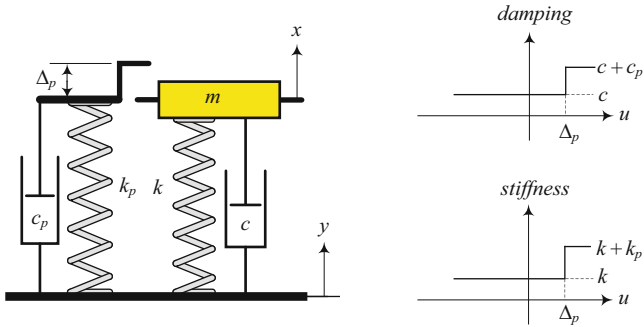
$$m\ddot{x} + g_1(x, \dot{x}) = f_1(y, \dot{y}) \tag{3.70}$$

where  $g_1(x, \dot{x})$  and  $f_1(y, \dot{y})$  are piecewise linear functions presenting nonlinear characteristics of the system and nonlinear excitation, respectively:

$$g_1(x, \dot{x}) = \begin{cases} (c + c_p)\dot{x} + (k + k_p)x - k_p\Delta_p & x - y > \Delta_p \\ c\dot{x} + kx & \Delta_n < |x - y| < \Delta_p \\ (c + c_n)\dot{x} + (k + k_n)x + k_n\Delta_n & x - y < -\Delta_n \end{cases} \tag{3.71}$$

$$f_1(y, \dot{y}) = \begin{cases} (c + c_p)\dot{y} + (k + k_p)y & x - y > \Delta_p \\ c\dot{y} + ky & \Delta_n < |x - y| < \Delta_p \\ (c + c_n)\dot{y} + (k + k_n)y & x - y < -\Delta_n \end{cases} \tag{3.72}$$

In an extreme case of the asymmetric system, the secondary suspension might be engaged only in one direction. Figure 3.18 illustrates the mathematical model of such system.



**Fig. 3.18** The mathematical model of a one-sided piecewise linear vibration isolator

### 3.5 Conclusion

In this chapter we have shown that obtaining exact frequency response of piecewise linear vibration isolator associates with a few numerical problems. It has been shown in the past that the exact solution of the system depends on a set of transcendental equations. For the first time, in this study the number of independent transcendental equations has been reduced to two. This reduction simplifies many of the numerical problems. However, not all of the problems can be resolved at the moment. The main remained challenges were discussed in this chapter. Among them are: difficulties in determining range of nonlinearity, multiple solutions and transition from one amplitude/frequency curve to the other (from stable solution to unstable and vice versa). It is still unclear how to designate the nonlinearity range, to calculate the maximum value of frequency for double engagement. The other challenges are solved and discussed in detail in this paper.

### 3.6 Key Symbols

- $A_i$  sin coefficient in  $z_i(t)$
- $B_i$  cos coefficient in  $z_i(t)$
- $c$  Damping coefficient [N s / m]
- $c_1$  Damping coefficient of the primary suspension [N s / m]
- $c_2$  Damping of the system when secondary engages [N s / m]
- $C_i$  Damping sine coefficient in  $z_i(t)$
- $D_i$  Damping cosine coefficient in  $z_i(t)$
- $k$  Stiffness [N / m]
- $k_1$  Stiffness of the primary suspension [N / m]
- $k_2$  Stiffness of the system when secondary engages [N / m]



$m$	Mass [kg]
$Q_i$	Short notation of mathematical expression
$t$	Time [s]
$t_1$	Time when $z_2 = z_1 = +1$ [s]
$u$	Relative displacement [m]
$w$	Dimensionless base excitation amplitude
$x$	Absolute displacement [m]
$y$	Base excitation function [m]
$Y$	Base excitation amplitude [m]
$z$	Dimensionless relative displacement
$Z$	Amplitude of relative displacement
$Z_1$	Amplitude of relative displacement of primary suspension
$Z_2$	Amplitude of relative displacement of secondary suspension
$\delta$	Gap distance
$v_1$	Velocity in $t_1$
$v_2$	Velocity in $z_2(t) = 0$
$\xi$	Damping ratio
$\xi_1$	Damping ratio of the primary suspension
$\xi_2$	Damping ratio of the system when secondary engages
$\phi$	Phase lag
$\omega$	Excitation frequency [1/s]
$\omega_1$	Natural frequency of the primary suspension
$\omega_2$	Natural frequency of the system when secondary engages
$\omega_{d_i}$	Damped natural frequency

## References

- Aizerman M, Gantmakher F (1957) Determination of periodic modes in systems with piecewise-linear characteristics, composed of segments parallel to two specified straight lines. *Autom Remote Control* 18:105–120
- Aizerman M, Gantmakher F (1958) Determination of stability by linear approximation of a periodic solution of a system of differential equations with discontinuous right-hand sides. *Q J Mech Appl Math* 11(4):385–398
- Aizerman M, Lur'e A (1963) Methods for construction of periodic motions in piecewise-linear systems. In *Symposium on Non-Linear Vibrations*, vol. 1, pp. 27–50. International Union for Theoretical and Applied Mechanics, Kiev
- Alkhatib R, Jazar, RN, Golnaraghi MF (2004) Optimal design of passive linear mounts with genetic algorithm method. *J Sound Vib* 275(3–5):665–691
- Christopherson J, Jazar, RN (2005) Optimization of classical hydraulic engine mounts based on RMS method. *J Shock Vib* 12(2):119–147
- Den Hartog J, Mikina S (1932) Forced vibrations with nonlinear spring constraints. *Trans ASME Appl Mech* 54-APM-15:157
- Deshpande S, Mehta S, Jazar RN (2006) Optimization of secondary suspension of piecewise linear vibration isolation systems. *Int J Mech Sci* 48(4):341–377
- Fleishman B (1965) Forced oscillations and convex superposition in piecewise-linear systems. *SIAM Rev* 7(2):205

- Gurtin M (1961) On the use of clearance in viscous dampers to limit high frequency force transmission. *J Eng Ind* 83(1):50–52
- Hayashi C (1964) *Nonlinear oscillations in physical systems*. McGraw-Hill, New York, McGraw-Hill
- Iwan WD (1973) A generalization of the concept of equivalent linearization. *Int J Nonlinear Mech* 8(3):279–287
- Jacobsen L, Jaspersen H (1935) Steady forced vibration of a single mass system with a nonlinear restoring force. *J Franklin Inst* 220:467–496
- Jazar RN (2013) *Advanced vibrations: a modern approach*. Springer, New York
- Jazar RN (2014) *Vehicle dynamics: theory and application*, 2nd edn. Springer, New York
- Jazar RN, Alkhatib R, Golnaraghi MF (2006a) Root mean square optimization criterion for vibration behavior of linear quarter car using analytical methods. *J Vehicle Syst Dyn* 44(6):477–512
- Jazar RN, Golnaraghi MF (2002) Engine mounts for automotive applications: a survey. *Shock Vib Digest* 34(5):363–379
- Jazar RN, Houim R, Narimani A, Golnaraghi F (2006b) Nonlinear passive engine mount, frequency response and jump avoidance. *J Vib Control* 12(11):1205–1237
- Jazar RN, Mahinfalah M, Deshpande S (2007) Design of a piecewise linear vibration isolator for jump avoidance. *IMEchE Part K J Multi Body Dyn* 221(K3):441–450
- Jazar RN, Narimani A, Golnaraghi MF, Swanson DA (2003) Practical frequency and time optimal design of passive linear vibration isolation mounts. *J Vehicle Syst Dyn* 39(6):437–466
- Marzbani H, Jazar RN, Khazaei A (2012) Smart passive vibration isolation: requirements and unsolved problems. *J Appl Nonlinear Dyn* 1(4):341–386
- Marzbani H, Jazar RN, Fard M (2013) Hydraulic engine mounts: a survey. *J Vib Control* 19(16):253–279
- Masri S (1965) *Analytical and experimental studies of impact dampers*. California Institute of Technology
- Masri S (1978) Analytical and experimental studies of a dynamic system with a gap. *J Mech Des* 100(3):480–486
- Masri S, Stott S (1978) Random excitation of a nonlinear vibration neutralizer. *J Mech Des* 100(4):681–689
- Narimani A, Golnaraghi MF, Jazar RN (2004a) Frequency response of a piecewise linear system. *J Vib Control* 10(12):1775–1894
- Narimani A, Jazar RN, Golnaraghi MF (2004b) Sensitivity analysis of frequency response of a piecewise linear system in frequency island. *J Vib Control* 10(2):175–198
- Natsiavas S (1989) Periodic response and stability of oscillators with symmetric trilinear restoring force. *J Sound Vib* 134(2):315–331
- Natsiavas S, Gonzalez H (1992) Vibration of harmonically excited oscillators with asymmetric constraints. *J Appl Mech Trans ASME* 59(2):S284–S290
- Nguyen D, Noah S, Kettleborough C (1986a) Impact behaviour of an oscillator with limiting stops, part I: a parametric study. *J Sound Vib* 109(2):293–307
- Nguyen D, Noah S, Kettleborough C (1986b) Impact behaviour of an oscillator with limiting stops, part II: Dimensionless design parameters. *J Sound Vib* 109(2):309–325
- Rosenberg R (1996) Steady-state forced vibrations. *Int J Nonlinear Mech* 1(2):95–108
- Schulman J (1983) Chaos in piecewise-linear systems. *Phys Rev A* 28(1):477
- Watabene T (1978) Forced vibration of a continuous system with nonlinear boundary conditions. *Trans ASME J Mech Des* 100:487–491

# Chapter 4

## Active Vibration Control for Nonlinear Axially Translating Cable Systems of Multi-Dimensions

L. Dai and L. Sun

**Keywords** Active vibration control • Nonlinear behaviors • Chaos • Axially translating cable • Dynamic systems of multi-dimensions

### 4.1 Introduction

Axially translating systems, such as cables, strings, belts, blades, and beams, have been widely studied for their important applications in the fields of civil, mechanical, architecture, and industrial engineering. The research on the dynamical responses and vibration controls of the systems, usually highly flexible and nonlinear, has attracted the attentions of the researchers and engineers in the fields and significant contributions have been made toward theoretical development and engineering applications. Regardless the research contributions in the fields, systematical and thorough investigations on the behavior of the axially translating systems and the control of the nonlinear behavior of the systems are still in demanding. It is the effective and reliable control of such nonlinear systems that has great significances in engineering applications.

Among the research works on the axially translating structures, a significant majority of the research is on axially translating beams, plates, and blades to which the modeling and analysis approaches of axially translating cables are similar. An axially traveling plate, with a constant length along the traveling direction, was studied (Luo and Hamidzadeh 2004) in designing the aerospace and aeronautical structures. In the study, the analytical solutions of high-speed traveling plates, as well as the buckling stability, were derived. In analyzing the axially translating

---

L. Dai (✉) • L. Sun  
Industrial Systems Engineering, University of Regina, Regina, SK, S4S 0A2, Canada  
e-mail: [liming.dai@uregina.ca](mailto:liming.dai@uregina.ca)

beams, the shear deformation of the beam plays an important role (Carrera and Giunta 2011). An axially moving Timoshenko beam was introduced in (Ghayesh and Amabili 2013) with the considerations of the shear deformations. In this study, a 20-dimension nonlinear dynamic system of Timoshenko beams with invariant length was modeled to study the bifurcations as well as chaos of the system.

The class of the axially moving structures with varying length has also been investigated widely for its various applications in the areas of engineering. Coming from the dynamics of spacecraft antenna, of which the length is varying with time, the equation of motion of a cantilevered beam was established (Tabarrok et al. 1974). In the study, a 16-dimension system of the beam has been derived, and a decreasing frequency has been discovered with respect to the increasing length of the beam. In (Fung et al. 1998), the four nonlinear axially moving cantilevered beam models, including Timoshenko, Euler, simple-flexible, and rigid-body beam models were established in considering a tip mass. In analyzing the hydraulic and motor driven systems, the authors pointed out that the rigid-body motion and the flexible vibration of the beam could be nonlinearly coupled and there existed Coriolis forces in the systems.

When cables and highly flexible structures are considered, the dynamic system modeling is an important part for studying the systems' dynamic behavior and stabilities which are critical to the control of the systems. With the interests in self-spinning tethered satellites, an Euler–Bernoulli beam was adopted to represent a tether with varying length (Tang et al. 2011). The deployment process of two-self-spinning tethered satellite systems was successfully simulated through a newly proposed hybrid Eulerian and Lagrangian frame work. A beam model together with a cable model was established by Zhu and Ni (2000) in order to investigate the dynamics of the flexible robot arms and elevator hoist cables considered in their study. The energetics and stability of translating media of the systems considered were investigated by the authors. The rates of energy changes of the established models were studied in both the extension and retraction processes of the systems considered. It is interesting to notice that a linear five-dimensional dynamic system was derived to ensure the reliability of the conducted study, as shown in their numerical simulation of a specified elevator hoist cable.

In the dynamic analysis of some mechanical axially traveling systems, such as saw blades and magnetics, the nonlinear vibration of a traveling tensioned beam was modeled and studied by Wickert (1992). The reported results showed that the inclusion of the nonlinearity in the proposed system was most important at near-critical speeds, where the modal stiffness was small and dominated by a nonlinear extensional stiffness. Based on the results regarding the supercritical stability characteristics, it was pointed out by the authors that the investigation needed to be improved in the case of higher-order equilibria of the translating beam.

With the models established for the axially traveling systems, the control of the dynamic motion of the systems can be analyzed, simulated, and applied. In the study by Zhu (2002), the equations of motion of translating media were derived via the Eulerian frame of reference, and then the rate of energy change of translating media was distinguished with control volume. Therefore, translating media could be

divided into two classes: the translating media with invariant control volume, and those with variable control volume. The rate of energy change was emphasized in characterizing the dynamic stability of translating media and stationary media with moving boundaries.

The control volume of an axially translating beam, which is located in between two spatially fixed supports, was discussed in the study of Humer (2013). As demonstrated in the study, this type of problems, in which kinematic constraints and loads were inherently non-material, could be found in many industrial processes such as belt drives for power transmission. In this study, the free vibration of an extended beam was investigated, and the length of the protruding part showed significant influence on the behavior of the structure.

An energy-based control scheme was proposed by Yang et al. (2005). In this study, the applications of an axially translating beam were restricted in high-speed and precision systems, due to the unwanted vibrations of moving continua. In the application of the control scheme, as shown in the study, the vibration of the translating beam could be suppressed.

For theoretically and numerically analyzing the vibrations of the axially translating systems and simulating the control of the vibrations, almost all the studies available in the archived literature are based on discretizing the governing equations of the systems modeled via the Galerkin method. On the basis of the discretized system, some of the control strategies were developed. A typical such approach can be seen in Yang et al. (2005), in which large-amplitude of the vibrations of axially translating beams was considered and an energy-based control strategy was developed. It should be noticed that, as indicated in the works (Zhu and Ni 2000; Wickert 1992), a multi-dimensional dynamic system will be favored in order to derive reliable results.

In engineering applications, corresponding to the nonlinear vibrations discovered in cables, strings, beams, and plates, several control schemes have been developed in the last several decades. Utkin (1992) proposed a control scheme in 1992, named sliding mode control (SMC) strategy, the SMC strategy has been applied in controlling the vibrations of nonlinear engineering structures, and this control strategy has gradually shown its applicability and efficiency in the works following Utkin's work.

Corresponding to the nonlinear dynamic systems with uncertain external disturbance, the fuzzy sliding mode control (FSMC) strategy was developed, in which the fuzzy logic rule was introduced into the original SMC strategy. The FSMC strategy showed its effectiveness in controlling nonlinear and chaotic motions discovered in the nonlinear dynamics systems (Yau and Kuo 2006; Kuo and Shieh 2007; Yau et al. 2010). In the study by Haghghi and Markazi (2009), for example, both the geometric nonlinearity and the nonlinear electrostatic force are considered, and a chaotic motion was found and controlled by applying the FSMC control strategy. The authors showed the elimination of the chaotic motion and stabilization of the nonlinear system with the control strategy.

In the studies (Yau and Kuo 2006; Kuo and Shieh 2007; Yau et al. 2010; Haghghi and Markazi 2009), although the effectiveness of the FSMC strategy

has been demonstrated in controlling the nonlinear vibrations, it can be found that the application of FSMC strategy is restricted on one-dimensional nonlinear dynamics systems. Corresponding to the vibrations of the multi-dimensional dynamic system, there is currently no control strategy available in the literature. Therefore, for controlling the nonlinear vibrations of multi-dimensional systems, the FSMC control strategy currently used needs to be modified. The multi-dimensional systems, developed through the Galerkin method, can be used to provide more accurate and reliable results through numerical simulations in comparing with that of single-dimensional system. The higher the order of the multi-dimensional system, theoretically, the more accurate will be the simulations of the nonlinear axially translating systems. This has been recognized by the researchers in this field (Zhu and Ni 2000; Wickert 1992).

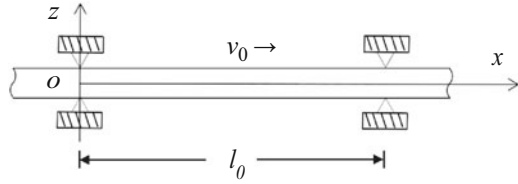
With the considerations described above, this chapter intends to introduce a newly developed approach for quantitatively controlling the vibrations of nonlinear axially translating cable systems. This includes a development of an active control strategy and the application of the control strategy in controlling the nonlinear cable systems. For developing such a control strategy, a cable system model consisting of the equations of motion is to be established based on von Karman-type equations. In developing for the solutions of the cable system and for the sake of applying the strategy in the nonlinear vibration control, the governing equations in the forms of partial differential equations will be non-dimensionalized and then transformed into three ordinary differential equations via a third-order Galerkin method. Corresponding to the derived multi-dimensional dynamic system, an active control strategy is developed based on the FSMC strategy. The applicability of the control strategy developed will be demonstrated in some numerical simulations based on the model established. A few cases of chaotic vibrations of the cable system are considered for applying the control strategy. The suppression and stabilization of the chaotic vibrations are to be demonstrated graphically to show the application and efficiency of the active control strategy in controlling the nonlinear axially translating cable systems.

## 4.2 Equations of Motion

The axially translating cable considered in this research is sketched in Fig. 4.1. The equations of motion of the cable are to be derived based on Hamilton's principle and von Karman-type equations. As can be seen from Fig. 4.1, the translating cable with pinned–pinned boundaries is allowed to move axially at a constant rate  $v_0$ , and the length of the cable is given as  $l$ . The displacement of any point of the cable along the  $x$ - and  $z$ - axes is designated as  $u$  and  $w$ .

Starting from the origin at the left support of the cable, a position vector,  $\mathbf{r}$ , of any point  $x(t)$  of the translating beam without deformation is given as

**Fig. 4.1** The model of the axially translating cable



$$\mathbf{r} = x(t)\mathbf{i} + 0\mathbf{k}, \tag{4.1}$$

where  $\mathbf{i}$  and  $\mathbf{k}$  are the unit vectors of the fixed Cartesian coordinate shown in the figure.

Thus, the displacement field of the axially translating cable can be derived as

$$\mathbf{R} = (x(t) + u_0(x(t), t))\mathbf{i} + w_0(x(t), t)\mathbf{k}, \tag{4.2}$$

where  $u_0(x(t), t)$  and  $w_0(x(t), t)$  are the displacement components along the  $x$ - and  $z$ -directions respectively, of a point of the cable.

Taking full differentiation of  $\mathbf{R}$  with respect to time  $t$ , one may obtain

$$\frac{d\mathbf{R}}{dt} = \left( \frac{dx(t)}{dt} + \frac{du_0(x(t), t)}{dt} \right)\mathbf{i} + \frac{dw_0(x(t), t)}{dt}\mathbf{k}, \tag{4.3}$$

where the derivative of  $x(t)$  with respect to time is equal to the translating rate of the cable, and the full derivative of  $w_0$  is

$$\frac{dw_0(x(t), t)}{dt} = \frac{\partial w_0(x(t), t)}{\partial t} + v_0 \frac{\partial w_0(x(t), t)}{\partial x(t)}, \tag{4.4-a}$$

$$\frac{d^2w_0(x(t), t)}{dt^2} = \frac{\partial^2 w_0(x(t), t)}{\partial t^2} + 2v_0 \frac{\partial w_0(x(t), t)}{\partial x(t)} + v_0^2 \frac{\partial^2 w_0(x(t), t)}{\partial x^2(t)}. \tag{4.4-b}$$

Hence, the kinetic energy of the translating cable over a volume  $V$  of the cable is expressible as

$$T = \int_0^l \frac{1}{2} \left( \rho \frac{d\mathbf{R}}{dt} \frac{d\mathbf{R}}{dt} \right) dx, \tag{4.5}$$

where  $\rho$  denotes the mass per unit length.

The von Karman-type equations of strains of large deflection associated with the displacement field, normal to the cross section of the cable along the  $x$  direction, in Eq. (4.2) can thus be given by

$$\varepsilon_{11} = \frac{\partial u_0(x(t), t)}{\partial x(t)} + \frac{1}{2} \left( \frac{\partial w_0(x(t), t)}{\partial x(t)} \right)^2. \tag{4.6}$$

Therefore, the total strain energy of the cable can be given by

$$U = Q_{11} \int_0^l \frac{1}{2} (\varepsilon_{11} \varepsilon_{11}) dx, \quad (4.7)$$

where  $Q_{11}$  represents the elastic coefficient in the same direction with  $\varepsilon_{11}$ .

The virtual work done by the force is 0,

$$W = 0, \quad (4.8)$$

In the following analysis, the Hamilton's principle will be employed to obtain the nonlinear equations of motion for the axially translating cable. The mathematical statement of the Hamilton's principle is given by

$$\int_{t_1}^{t_2} (\delta L + \delta W) dt = 0, \quad (4.9)$$

where the total Lagrangian function  $L$  is given by

$$L = T - E. \quad (4.10)$$

For the sake of clarity, hereafter, use  $x$ ,  $u_0$ ,  $w_0$  to replace  $x(t)$ ,  $u_0(x(t), t)$ , and  $w_0(x(t), t)$  respectively. Substitute Eqs. (4.5)–(4.8) into Eq. (4.10), and the first term in Eq. (4.9) can be developed as below,

$$\begin{aligned} \int_{t_1}^{t_2} \delta L dt &= \int_{t_1}^{t_2} \delta (T - U) dt \\ &= \int_{t_1}^{t_2} \delta \left\{ \int_0^l \left[ \rho \left( \frac{d\mathbf{R}}{dt} \frac{d\mathbf{R}}{dt} \right) - Q_{11} (\varepsilon_{11} \varepsilon_{11}) \right] \right\} dt \\ &= \int_{t_1}^{t_2} \int_0^l \left[ \rho \left( \frac{d\mathbf{R}}{dt} \frac{d\delta\mathbf{R}}{dt} \right) \right] dt - \int_{t_1}^{t_2} \int_0^l [Q_{11} (\varepsilon_{11} \delta\varepsilon_{11})] dt \\ &= 0 - \int_V \int_{t_1}^{t_2} \rho \left( \frac{d^2\mathbf{R}}{dt^2} \right) \delta\mathbf{R} dt dV - \int_{t_1}^{t_2} \int_V Q_{11} \varepsilon_{11} \delta\varepsilon_{11} dV dt \\ &= - \int_{t_1}^{t_2} \int_0^l \rho \delta (x + u_0) \left( \frac{d^2x}{dt^2} + \frac{d^2u_0}{dt^2} \right) dx dt \\ &\quad - \int_{t_1}^{t_2} \int_0^l \rho \delta w_0 \frac{d^2w_0}{dt^2} dx dt \end{aligned}$$



$$\begin{aligned}
& - \int_{t_1}^{t_2} \int_0^l Q_{11} \delta \left( \frac{\partial u_0}{\partial x} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \varepsilon_{11} dx dt \\
& = L_1 + L_2 + L_3
\end{aligned}$$

where

$$L_1 = - \int_{t_1}^{t_2} \int_0^l \rho \delta (x + u_0) \left( \frac{d^2 x}{dt^2} + \frac{d^2 u_0}{dt^2} \right) dx dt, \quad (4.11-a)$$

$$L_2 = - \int_{t_1}^{t_2} \int_0^l \rho \delta w_0 \frac{d^2 w_0}{dt^2} dx dt, \quad (4.11-b)$$

$$L_3 = - \int_{t_1}^{t_2} \int_0^l Q_{11} \delta \left( \frac{\partial u_0}{\partial x} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \varepsilon_{11} dx dt. \quad (4.11-c)$$

Since the cable is moving axially at a constant velocity ( $\frac{d^2 v_0}{dt^2} = 0$ ), from Eq. (4.11-a), it can be derived

$$\begin{aligned}
L_1 &= - \int_{t_1}^{t_2} \int_0^l \rho \delta (x + u_0) \left( \frac{d^2 x}{dt^2} + \frac{d^2 u_0}{dt^2} \right) dx dt \\
&= - \int_{t_1}^{t_2} \int_0^l \rho (0 + \delta u_0) \left( 0 + \frac{d^2 u_0}{dt^2} \right) dx dt \\
&= - \int_{t_1}^{t_2} \int_0^l \rho \frac{d^2 u_0}{dt^2} \delta u_0 dx dt
\end{aligned} \quad (4.12-a)$$

From Eq. (4.11-c), it can be derived in as follows

$$\begin{aligned}
L_3 &= - \int_{t_1}^{t_2} \int_0^l Q_{11} \delta \left( \frac{\partial u_0}{\partial x} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \varepsilon_{11} dx dt \\
&= - \int_{t_1}^{t_2} \int_0^l Q_{11} \left( \frac{\partial \delta u_0}{\partial x} + \frac{1}{2} \delta \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \varepsilon_{11} dx dt \\
&= - \int_{t_1}^{t_2} \int_0^l Q_{11} \varepsilon_{11} \frac{\partial \delta u_0}{\partial x} dx dt - \int_{t_1}^{t_2} \int_0^l Q_{11} \varepsilon_{11} \frac{\partial w_0}{\partial x} \frac{\partial \delta w_0}{\partial x} dx dt \\
&= - \int_{t_1}^{t_2} \left[ (Q_{11} \varepsilon_{11} \delta u_0) \Big|_0^l - \int_0^l Q_{11} \delta u_0 d \varepsilon_{11} \right] dt
\end{aligned}$$

$$\begin{aligned}
& - \int_{t_1}^{t_2} \left[ \left( Q_{11} \varepsilon_{11} \frac{\partial w_0}{\partial x} \delta w_0 \right) \Big|_0^l - \int_0^l Q_{11} \varepsilon_{11} \delta w_0 d \frac{\partial w_0}{\partial x} \right] dt \\
& = - \int_{t_1}^{t_2} \left( 0 - \int_0^l Q_{11} \delta u_0 d \varepsilon_{11} \right) dt \\
& \quad - \int_{t_1}^{t_2} \left( 0 - \int_0^l Q_{11} \delta w_0 d \left( \varepsilon_{11} \frac{\partial w_0}{\partial x} \right) \right) dt \\
& = - \int_{t_1}^{t_2} \left( - \int_0^l Q_{11} \frac{\partial \varepsilon_{11}}{\partial x} \delta u_0 dx \right) dt - \int_{t_1}^{t_2} \left( - \int_0^l Q_{11} \varepsilon_{11} \frac{\partial^2 w_0}{\partial x^2} \delta w_0 dx \right) dt \\
& \quad - \int_{t_1}^{t_2} \left( - \int_0^l Q_{11} \frac{\partial \varepsilon_{11}}{\partial x} \frac{\partial w_0}{\partial x} \delta w_0 dx \right) dt \\
& = - \int_{t_1}^{t_2} \left( - \int_0^l Q_{11} \frac{\partial}{\partial x} \left( \frac{\partial u_0}{\partial x} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \delta u_0 dx \right) dt \\
& \quad - \int_{t_1}^{t_2} \left( - \int_0^l Q_{11} \left( \frac{\partial u_0}{\partial x} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \frac{\partial^2 w_0}{\partial x^2} \delta w_0 dx \right) dt \\
& \quad - \int_{t_1}^{t_2} \left( - \int_0^l Q_{11} \frac{\partial}{\partial x} \left( \frac{\partial u_0}{\partial x} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \frac{\partial w_0}{\partial x} \delta w_0 dx \right) dt \\
& = \int_{t_1}^{t_2} \left( \int_0^l Q_{11} \left( \frac{\partial^2 u_0}{\partial x^2} + \frac{\partial w_0}{\partial x} \frac{\partial^2 w_0}{\partial x^2} \right) \delta u_0 dx \right) dt \\
& \quad + \int_{t_1}^{t_2} \left( \int_0^l Q_{11} \left( \frac{\partial u_0}{\partial x} \frac{\partial^2 w_0}{\partial x^2} + \frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 \frac{\partial^2 w_0}{\partial x^2} \right) \delta w_0 dx \right) dt \\
& \quad + \int_{t_1}^{t_2} \left( \int_0^l Q_{11} \left( \frac{\partial^2 u_0}{\partial x^2} \frac{\partial w_0}{\partial x} + \frac{\partial^2 w_0}{\partial x^2} \left( \frac{\partial w_0}{\partial x} \right)^2 \right) \delta w_0 dx \right) dt \tag{4.12-b}
\end{aligned}$$

From Eqs. (4.12-a), (4.12-b), and (4.11-b), the nonlinear governing equation of an axially translating cable can be derived in the following,

$$-\rho \frac{d^2 u_0}{dt^2} + Q_{11} \left( \frac{\partial^2 u_0}{\partial x^2} + \frac{\partial w_0}{\partial x} \frac{\partial^2 w_0}{\partial x^2} \right) = 0, \tag{4.13-a}$$

$$-\rho \frac{d^2 w_0}{dt^2} + Q_{11} \frac{\partial^2 u_0}{\partial x^2} \frac{\partial w_0}{\partial x} + Q_{11} \frac{\partial u_0}{\partial x} \frac{\partial^2 w_0}{\partial x^2} + \frac{3}{2} Q_{11} \left( \frac{\partial w_0}{\partial x} \right)^2 \frac{\partial^2 w_0}{\partial x^2} = 0, \tag{4.13-b}$$

Associated with the nonlinear dynamic equations and the boundary conditions (Nayfeh and Mook 1979; Abou-Rayan et al. 1993), the strain can be obtained as

$$\frac{\partial u_0}{\partial x} = -\frac{1}{2} \left( \frac{\partial w_0}{\partial x} \right)^2 + \frac{1}{2l} \int_0^l \left( \frac{\partial w_0}{\partial x} \right)^2 dx. \quad (4.14)$$

Then, substitute Eq. (4.14) into Eq. (4.13-b), and the nonlinear differential governing equation of the axially translating cable in  $z$  direction is derived as

$$-\rho \frac{d^2 w_0}{dt^2} + \int_0^l \frac{1}{2l} Q_{11} \left( \frac{\partial w_0}{\partial x} \right)^2 dx \frac{\partial^2 w_0}{\partial x^2} = 0. \quad (4.15)$$

### 4.3 Non-dimensionalization

To validate the governing equation Eq. (4.15) and facilitate the numerical simulations in the consequent sections, the following non-dimensional variables are introduced,

$$\bar{t} = \sqrt{\frac{Q_{11}}{\rho}} t = \tau t, \quad \bar{x} = \frac{x}{l}, \quad (4.16)$$

and

$$\bar{w}_0 = \frac{w_0}{l}, \quad \frac{d\bar{w}_0}{d\bar{t}} = \frac{1}{\tau l} \frac{dw_0}{dt}, \quad \frac{d^2\bar{w}_0}{d\bar{t}^2} = \frac{1}{\tau^2 l} \frac{d^2w_0}{dt^2}. \quad (4.17)$$

With the non-dimensional variables shown in Eqs. (4.16) and (4.17) introduced into Eq. (4.14), the non-dimensional governing equation of the axially translating cable can be expressed as

$$-\frac{d^2\bar{w}_0}{d\bar{t}^2} + \eta \int_0^1 \left( \frac{\partial\bar{w}_0}{\partial\bar{x}} \right)^2 d\bar{x} \frac{\partial^2\bar{w}_0}{\partial\bar{x}^2} = 0, \quad (4.18)$$

where

$$\eta = \frac{1}{2} \frac{1}{l} Q_{11} \frac{1}{\rho l \tau^2}.$$

#### 4.4 Series Solutions

Based on the Galerkin method of discretization, the transverse displacement  $\bar{w}_0$  is expanded in a series form, in terms of a set of comparison functions as,

$$\bar{w}_0 = \sum_{n=1}^{\infty} \varphi_n(\bar{x}) \bar{w}_n(\bar{t}). \quad (4.18)$$

Corresponding to the pinned–pinned boundaries of the axially translating string,  $\varphi_n(\bar{x})$  can be given as follows,

$$\varphi_n(\bar{x}) = \sin(n\pi\bar{x}). \quad (4.19)$$

Substitute the series solution of Eq. (4.19) into Eq. (4.17), and to assist presentation, replace  $\varphi_n$ ,  $w_n$ ,  $\dot{w}_n$ ,  $\ddot{w}_n$ ,  $v$ , and  $t$  for  $\varphi_n(\bar{x})$ ,  $\bar{w}_n(\bar{t})$ ,  $\frac{dw_n}{dt}$ ,  $\frac{d^2w_n}{dt^2}$ ,  $\bar{v}_0$ , and  $\bar{t}$  respectively. With the application of the Galerkin method at  $n = 3$ , and

$$\begin{aligned} w_1 &= w_{1,1}, \quad w_2 = w_{2,1}, \quad w_3 = w_{3,1}, \\ \dot{w}_{1,1} &= w_{1,2}\dot{w}_{2,1} = w_{2,2}, \quad \dot{w}_{3,1} = w_{3,2}. \end{aligned}$$

The discretized governing equations of the axially translating cable with the pinned–pinned boundaries can be obtained as the following.

$$\begin{cases} \dot{w}_{1,1} = w_{1,2} \\ \dot{w}_{1,2} = 2(\chi_1 + \eta\zeta_1) \\ \dot{w}_{2,1} = w_{2,2} \\ \dot{w}_{2,2} = 2(\chi_2 + \eta\zeta_2) \\ \dot{w}_{3,1} = w_{3,2} \\ \dot{w}_{3,2} = 2(\chi_3 + \eta\zeta_3) \end{cases} \quad (4.20)$$

where,

$$\begin{aligned} \chi_1 &= \frac{8}{3}vw_{2,1} + \frac{1}{2}v^2\pi^2w_{1,1}, \\ \chi_2 &= \frac{24}{5}vw_{3,1} - \frac{8}{3}vw_{1,1} + 2v^2\pi^2w_{2,1}, \\ \chi_3 &= -\frac{24}{5}vw_{3,1} + \frac{9}{2}v^2\pi^2w_{3,1}, \\ \zeta_1 &= \left( -\frac{1}{4}\eta\pi^4w_{1,1}^3 - \frac{9}{4}\eta\pi^4w_{1,1}w_{3,1}^2 - \eta\pi^4w_{1,1}w_{2,1}^2 \right), \end{aligned}$$

$$\begin{aligned}\zeta_2 &= (-\eta\pi^4 w_{2,1} w_{1,1}^2 - 9\eta\pi^4 w_{2,1} w_{3,1}^2 - 4\eta\pi^4 w_{2,1}^3), \\ \zeta_3 &= \left(-\frac{9}{4}\eta\pi^4 w_{3,1} w_{1,1}^2 - \frac{81}{4}\eta\pi^4 w_{3,1}^3 - 9\eta\pi^4 w_{3,1} w_{2,1}^2\right),\end{aligned}$$

## 4.5 Control Strategy Development

With the developed governing equations, boundary conditions, and the solutions of the governing equations, an active vibration control strategy can be developed. Based on the previous works (Utkin 1992; Haghghi and Markazi 2009; Kuo and Shieh 2007), the proposed active control strategy is developed for the vibration control of a multi-dimensional nonlinear dynamic system shown in Eq. (4.20).

For a nonlinear governing equation in the following general form

$$\ddot{w} = \Phi(w, \dot{w}, t), \quad (4.21)$$

denote  $U$  as the control input and  $\Delta P(w, \dot{w})$  the unknown external disturbance applying on the beam, the governing equation Eq. (4.21) for the beam with the control input and the external disturbance can be given by

$$\ddot{w} = \Phi(w, \dot{w}, t) + U + \Delta P(w, \dot{w}). \quad (4.22)$$

With application of the control, it is expected that the vibration of the beam can be controlled

If the  $n$ th Galerkin method is applied in the discretization of the governing equation given in Eq. (4.22), a series of second order ordinary differential equations considering the control input  $U$  and the unknown external disturbance will be derived as follows,

$$\left\{ \begin{array}{l} \dot{w}_{1,1} = w_{1,2} \\ \dot{w}_{1,2} = \phi_1(\mathbf{W}, t) + u_1 + \Delta f_1(\mathbf{W}, t) \\ \dot{w}_{2,1} = w_{2,2} \\ \dot{w}_{2,2} = \phi_2(\mathbf{W}, t) + u_2 + \Delta f_2(\mathbf{W}, t) \\ \vdots \\ \dot{w}_{i,1} = w_{i,2} \\ \dot{w}_{i,2} = \phi_i(\mathbf{W}, t) + u_i + \Delta f_i(\mathbf{W}, t) \\ \vdots \\ \dot{w}_{n,1} = w_{n,2} \\ \dot{w}_{n,2} = \phi_n(\mathbf{W}, t) + u_n + \Delta f_n(\mathbf{W}, t) \end{array} \right., \quad (4.23)$$

where  $\phi_i(\mathbf{W}, t)$ ,  $u_i$ , and  $\Delta f_i(\mathbf{W}, t)$  represent the expressions of  $\Phi(w, \dot{w}, t)$ ,  $U$ , and  $\Delta P(w, \dot{w})$  after the application of the Galerkin discretization. With the Galerkin discretization, the column vector  $\mathbf{W}$  in Eq. (4.23) is given below,

$$\mathbf{W} = [w_{1,1} \ w_{1,2} \ w_{2,1} \ w_{2,2} \ \cdots \ w_{i,1} \ w_{i,2} \ \cdots \ w_{n,1} \ w_{n,2}]^T.$$

Considering the response of a point of the axially translating cable along the  $x$  direction, based on Eq. (4.23) and the expression in Eq. (4.18), the non-dimensional response of the selected point  $w_p$  can be given as

$$w_p = \sum_{n=1}^{\infty} \varphi_n(x_p) w_n(t), \tag{4.24}$$

where  $x_p$  denotes the location of the selected point.

For a desired vibration or reference vibration expressed as

$$w_r = A_r \sin(\omega_r t), \tag{4.25}$$

where,  $A_r$  and  $\omega_r$  are the amplitude and frequency of the reference signal.

The control input  $U$  can be given as,

$$U = U_{eq} - U_r, \tag{4.26}$$

where  $U_{eq}$  and  $U_r$  are expressed as below,

$$U_{eq} = -((\dot{w}_p - \dot{\Psi}) + \kappa(w_p - \Psi)), \quad U_r = k_{fs} U_{fs}. \tag{4.27}$$

In Eq. (4.27),  $\kappa$  designates the control parameters governing the sliding surface,  $k_{fs}$  is given as  $|\Delta F(w, \dot{w})| < k_{fs} \in R^+$ , and the value of  $U_{fs}$  depends on the fuzzy rule shown in the table below.

The fuzzy rule of  $U_{fs}$

$U_{fs}$		$U_{eq}$						
		1	2/3	1/3	0	-1/3	-2/3	-1
$\frac{dU_{eq}}{dt}$	1	-1	-1	-1	-1	-2/3	-1	0
	2/3	-1	-1	-1	-2/3	-1	0	1/3
	1/3	-1	-1	-2/3	-1	0	1/3	2/3
	0	-1	-2/3	-1	0	1/3	2/3	1
	-1/3	-2/3	-1	0	1/3	2/3	1	1
	-2/3	-1	0	1/3	2/3	1	1	1
	-1	0	1/3	2/3	1	1	1	1

With the control strategy demonstrated in Eqs. (4.22)–(4.27), the control of an axially translating cable governed with the general governing equation Eq. (4.21) can be realized.

Take the axially translating beam governed by Eq. (4.17) as an example. Use the control strategy developed and apply the control input as shown in Eq. (4.27), the governing equation with the control input for the cable can be given by the following expression.

$$\ddot{w}_0 = \eta \int_0^1 \left( \frac{\partial \bar{w}_0}{\partial \bar{x}} \right)^2 d\bar{x} \frac{\partial^2 \bar{w}_0}{\partial \bar{x}^2} + U + \Delta F(w, \dot{w}), \quad (4.28)$$

With the application of the third-order Galerkin discretization, Eq. (4.29) may have the following form:

$$\begin{cases} \dot{w}_{1,1} = w_{1,2} \\ \dot{w}_{1,2} = 2(\chi_1 + \eta \zeta_1) + u_1 + \Delta f_1(\mathbf{W}, t) \\ \dot{w}_{2,1} = w_{2,2} \\ \dot{w}_{2,2} = 2(\chi_2 + \eta \zeta_2) + u_2 + \Delta f_2(\mathbf{W}, t) \\ \dot{w}_{3,1} = w_{3,2} \\ \dot{w}_{3,2} = 2(\chi_3 + \eta \zeta_3) + u_3 + \Delta f_3(\mathbf{W}, t) \end{cases}, \quad (4.29)$$

where  $u_1, u_2, u_3, u_4, u_5,$  and  $u_6$  are derived as follows through the third-order Galerkin method,

$$u_1 = \frac{2}{\pi}U, u_2 = 0 \cdot U, u_3 = \frac{2}{3\pi}U.$$

It should be noticed that due to the pinned–pinned boundaries of the axially translating cable, the application of the Galerkin method based on Eq. (4.19) makes the coefficients of  $u_2$  equal to zero. In the next section, it will be demonstrated in the numerical simulation that the actual response of the axially translating cable at a selected point can be well synchronized with a desired reference signal in the case that the coefficients of  $u_2$  equals to zero.

## 4.6 Vibration Control

To demonstrate the applicability and effectiveness of the control strategy developed in the previous section, numerical simulations are conducted for controlling an axially translating cable governed with Eq. (4.17). The nonlinear responses of the cable are emphasized in this research. With the numerical simulations performed, a chaotic motion is found when the cable is translating at certain rates. The proposed active control strategy is found not only effectively reduces the amplitude of the chaotic motion, but also stabilizes the motion so that the response of

the translating cable is controlled to a desired periodic motion. To facilitate the numerical simulation, the fourth-order P–T method by Dai (2008) is implemented.

The parameters used for the simulations for the responses of the axially translating cable are given as follows,

$$Q_{11} = 2.9 \times 10^4 N, \quad l = 0.5m, \quad \rho = 1.00kg/m,$$

and the constant axially translating rate is given as below,

$$v_0 = 2m/s.$$

The non-dimensionalized initial conditions, corresponding to the displacements described by Eqs. (4.29) after the implementation of the third-order Galerkin method, are taken as

$$w_{1,1}(0) = 0.001, \quad w_{1,2}(0) = 0.005,$$

$$w_{2,1}(0) = 0.0001, \quad w_{2,2}(0) = 0.0025,$$

$$w_{3,1}(0) = 0.00005, \quad w_{3,2}(0) = 0.00125.$$

If the response of a point at  $0.3125m$  along the  $x$ -axis of the cable is selected, based on Eqs. (4.18) and (4.29) the non-dimensional response of the selected point  $w_p$  can be derived as,

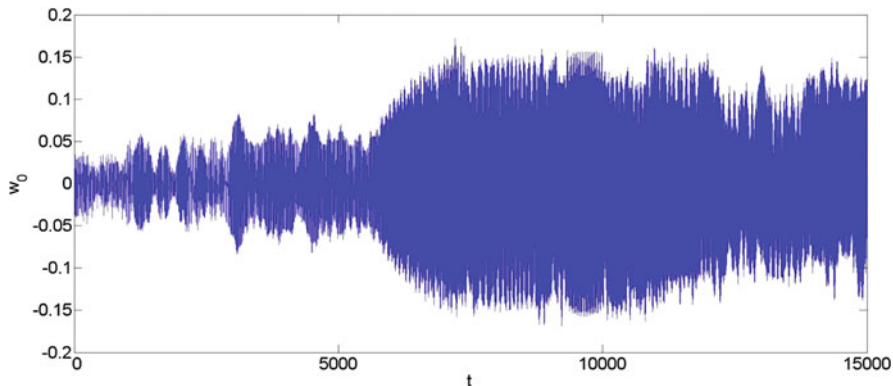
$$w_p = \sum_{n=1}^6 \varphi_n w_{n,1} = 0.92388w_{1,1} - 0.70711w_{2,1} + 0.38268w_{3,1}. \quad (4.30)$$

### 4.6.1 Chaotic Motion

The response of the cable translating at the speed  $v_0 = 2m/s$  is shown in Fig. 4.2, corresponding to the non-dimensional time from  $t = 0$  to  $t = 15,000$ . During this period of time, one may notice: in Fig. 4.2, it is a chaotic motion discovered; the maximum amplitude of the vibration of the cable can exceed 0.15. In considering that the displacement shown in the figure is non-dimensional, the amplitude is very large. Thus, reduction and stabilization of the chaotic motion may improve the operation of the beam.

Besides, from Fig. 4.3a–c, it can be learned that although the contribution of the first vibration mode in Fig. 4.3a is larger than those of the other two vibration modes as shown in Fig. 4.3b, c, the contributions of the other two vibration modes are obviously not negligible. Actually it can be learned that the other two vibration modes also significantly contributes to the actual response of the selected point.





**Fig. 4.2** The wave diagram of  $w_p$  without the application of the control strategy

Thus, the development of a multi-dimensional dynamic system is necessary for the accurate prediction of the dynamics of the extending nonlinear elastic cable.

#### 4.6.2 Amplitude Synchronization

The control strategy will be applied to show its effectiveness and efficiency in the amplitude synchronization. Three different desired amplitudes will be specified as the amplitude of the reference signal.

$$A_r = 0.045$$

In the application of the control strategy, the desired amplitude of the reference is set as

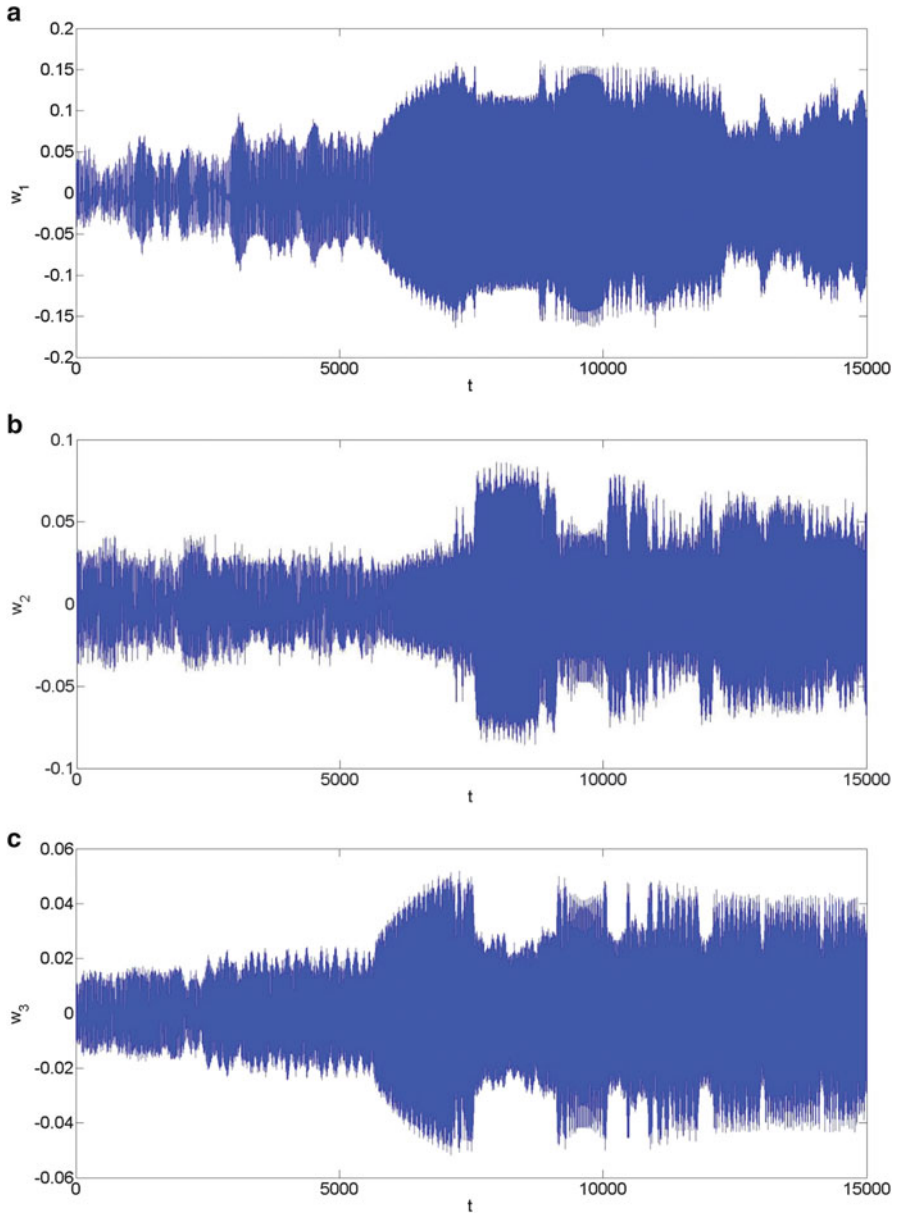
$$w_r = 0.045 \sin(0.0369t).$$

and other the control parameters and the unknown external disturbance take the following values:

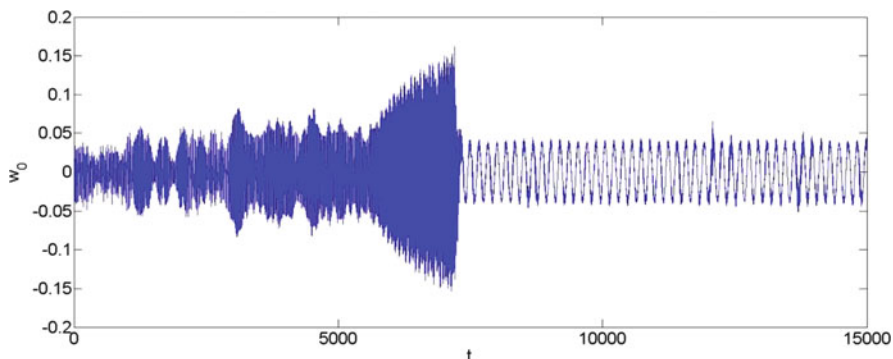
$$\kappa = 3.5, k_{fs} = 0.01, \Delta F(w, \dot{w}) = 0.001 \sin(w_p).$$

As shown in Fig. 4.4, the proposed control strategy is applied at  $t = 7,200$ . After the application of the control strategy, the responses of the cable at the translating speed  $v_0 = 2m/s$  are shown in Figs. 4.4 and 4.5.

In Fig. 4.4 the responses of the selected point,  $w_p$ , is shown, for the period of time from  $t = 7,200$  to  $t = 15,000$ . It can be seen from Fig. 4.4 that a short period of



**Fig. 4.3** (a) The contribution of the first vibration mode to  $w_p$  without the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  without the application of the control strategy. (c) The wave diagram of  $w_3$  without the application of the control strategy



**Fig. 4.4** The wave diagram of  $w_p$  with the application of the control strategy

time is needed for stabilizing the beam after the application of the control strategy. After the short period, the chaotic motion will then become a periodic one, of which the amplitude is 0.045.

From Fig. 4.5a–c, the displacements are shown in terms of  $w_1$ ,  $w_2$ , and  $w_3$ . Based on these figures, through the application of the control strategy, each of the displacements of the axially translating cable is gradually stabilized from a chaotic motion into a periodic one.

Figure 4.6 shows the comparison between the actual response of the beam  $w_p$  and the reference signal  $w_r$  applied. One may notice that the reference signal  $w_r$  is well periodic with respect to time  $t$ . One may also see from the figure, the maximum amplitude of the cable's response slightly varies after the stabilization of the cable with the application of the control strategy. As shown in the figure, the maximum amplitude of the cable is very close to that of the reference signal.

Figure 4.7 shows the control input  $U$ . Initially, the control input reaches a peak very quickly after the application of the control strategy. Once the system is stabilized, the control input displays a periodic wave diagram as shown in the figure and the maximum value of the control input is significantly decreased to a very small value.

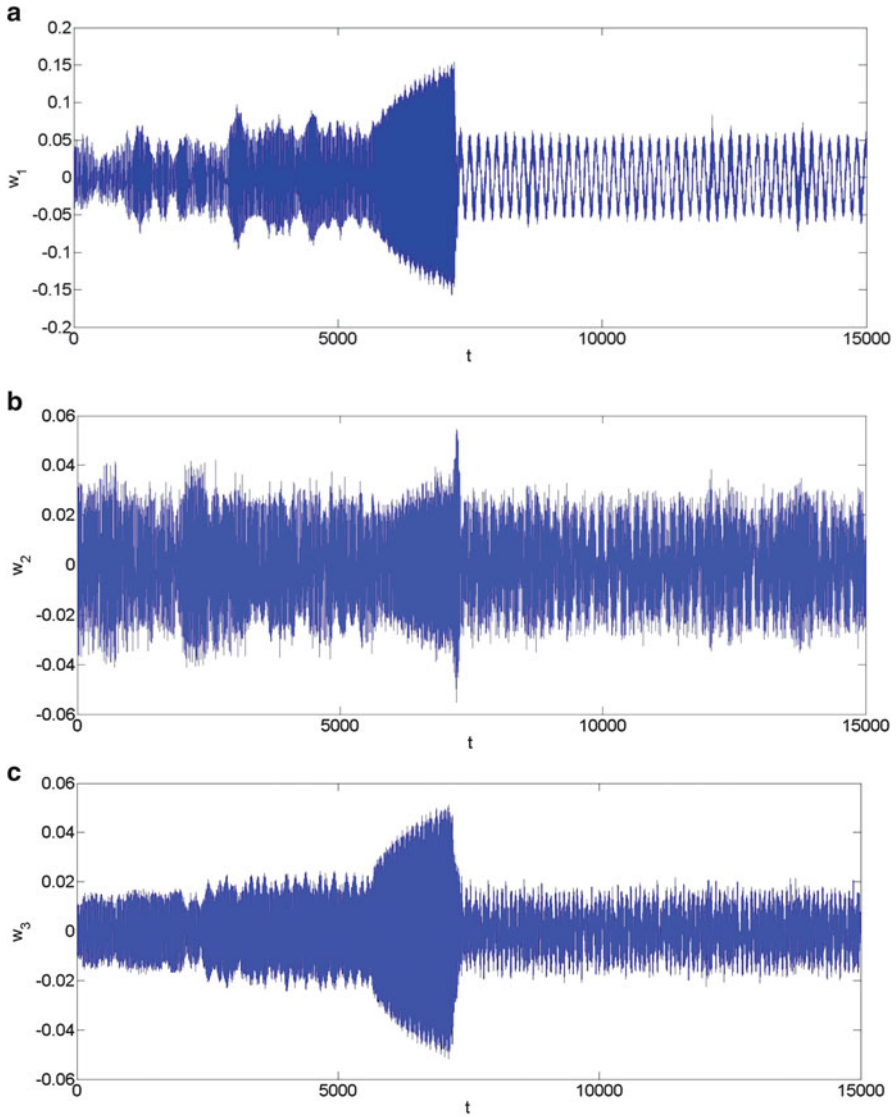
### $A_r = 0.015$

In the application of the control strategy, the desired amplitude of the reference is set as

$$w_r = 0.015 \sin(0.0369t).$$

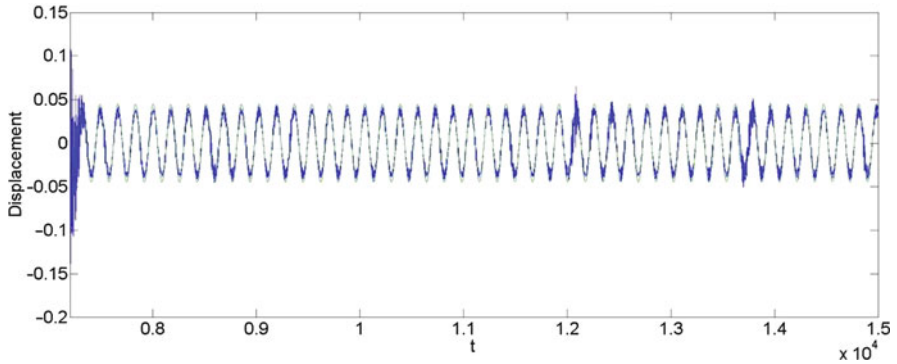
and other the control parameters and the unknown external disturbance take the following values:

$$\kappa = 2, k_{fs} = 0.01, \Delta F(w, \dot{w}) = 0.001 \sin(w_p).$$

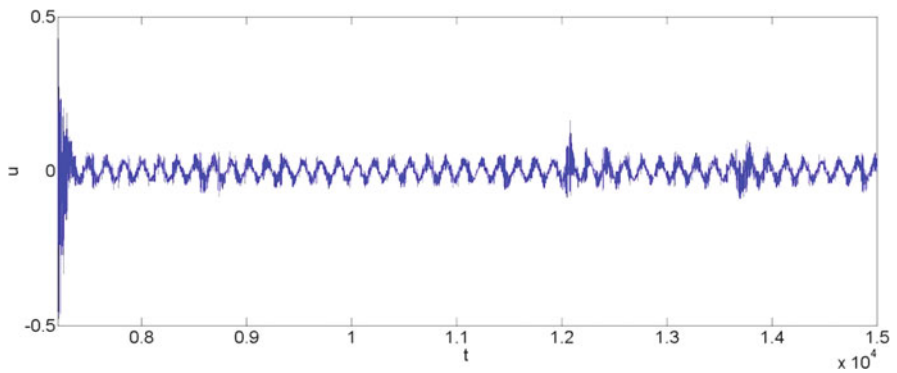


**Fig. 4.5** (a) The contribution of the first vibration mode to  $w_p$  with the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  with the application of the control strategy. (c) The contribution of the third vibration mode to  $w_p$  with the application of the control strategy

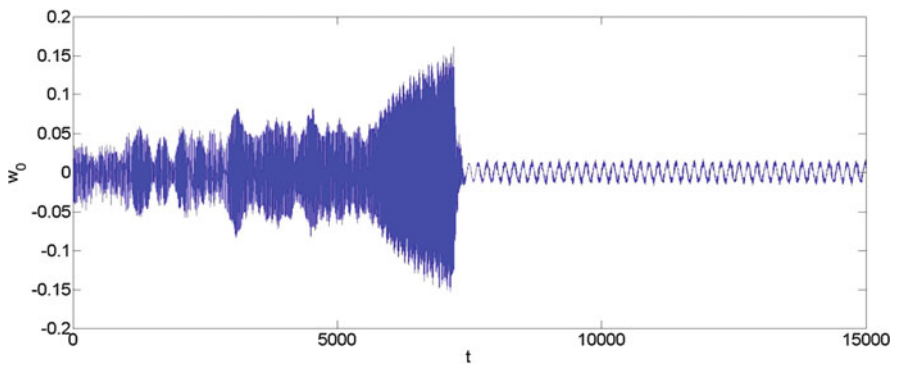
As shown in Fig. 4.8, the proposed control strategy is applied at  $t = 7,200$ . After the application of the control strategy, the responses of the cable at the translating speed  $v_0 = 2m/s$  are shown in Figs. 4.8 and 4.9.



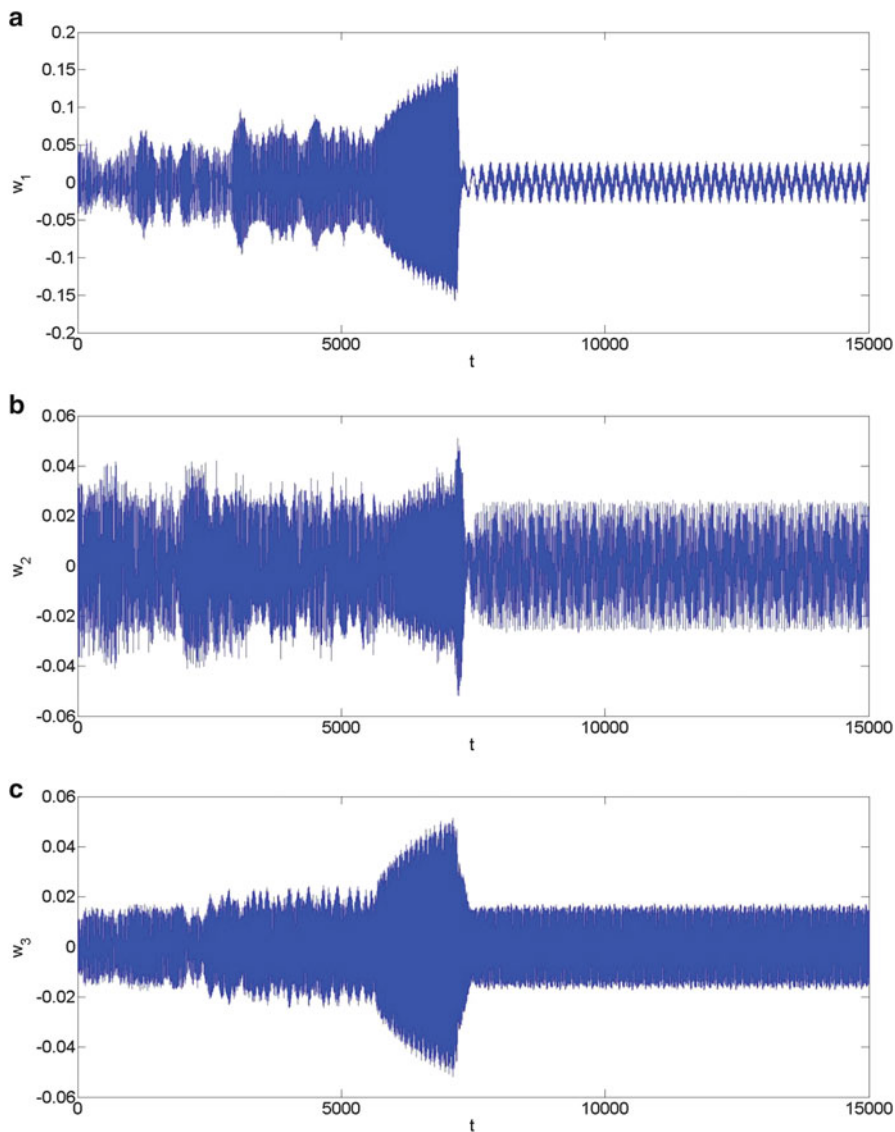
**Fig. 4.6** The wave diagram of  $w_p$  (the blue continuous line) and the reference signal  $w_r$  (the green dash line)



**Fig. 4.7** The control input

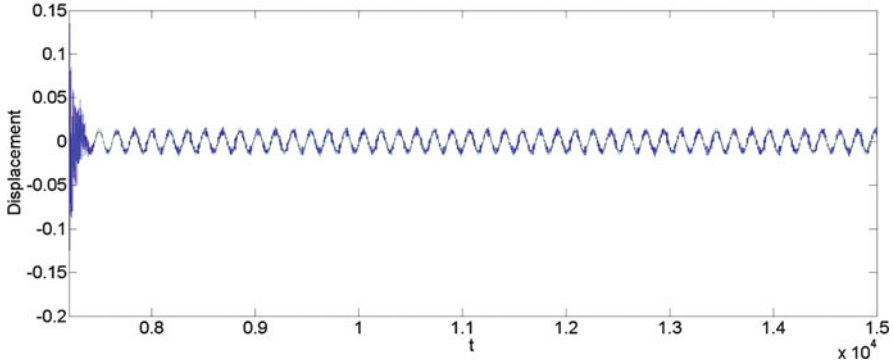


**Fig. 4.8** The wave diagram of  $w_p$  with the application of the control strategy



**Fig. 4.9** (a) The contribution of the first vibration mode to  $w_p$  with the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  with the application of the control strategy. (c) The contribution of the third vibration mode to  $w_p$  with the application of the control strategy

In Fig. 4.8 the responses of the selected point,  $w_p$ , is shown, for the period of time from  $t = 7,200$  to  $t = 15,000$ . It can be seen from Fig. 4.8, a short period of time is needed for stabilizing the beam after the application of the control strategy.



**Fig. 4.10** The wave diagram of  $w_p$  (the blue continuous line) and the reference signal  $w_r$  (the green dash line)

After the short period, the chaotic motion will then become a periodic one, of which the amplitude is 0.015.

From Fig. 4.9a–c, the displacements are shown in terms of  $w_1$ ,  $w_2$ , and  $w_3$ . Based on these figures, through the application of the control strategy, each of the displacements of the axially translating cable is gradually stabilized from a chaotic motion into a periodic one.

Figure 4.10 shows the comparison between the actual response of the beam  $w_p$  and the reference signal  $w_r$  applied. One may notice that the reference signal  $w_r$  is well periodic with respect to time  $t$ . One may also see from the figure, the maximum amplitude of the cable's response slightly varies after the stabilization of the cable with the application of the control strategy. As shown in the figure, the maximum amplitude of the cable is very close to that of the reference signal.

Figure 4.11 shows the control input  $U$ . Initially, the control input reaches a peak very quickly after the application of the control strategy. Once the system is stabilized, the control input displays a periodic wave diagram as shown in the figure and the maximum value of the control input is significantly decreased to a very small value.

$$A_r = 0.005$$

In the application of the control strategy, the desired amplitude of the reference is set as

$$w_r = 0.005 \sin(0.0369t).$$

and other the control parameters and the unknown external disturbance take the following values:

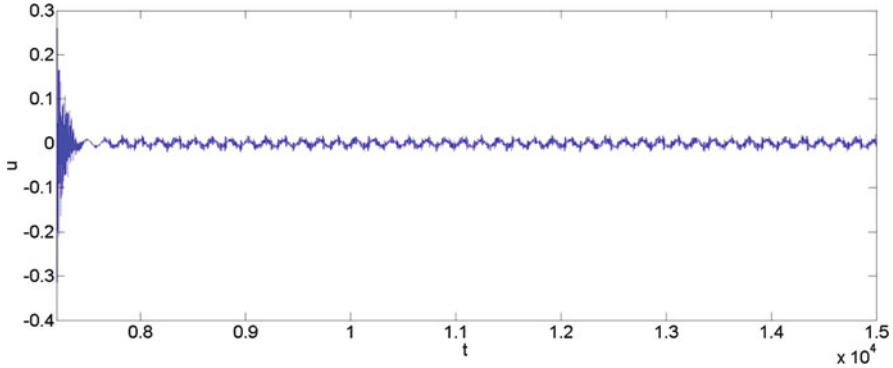


Fig. 4.11 The control input

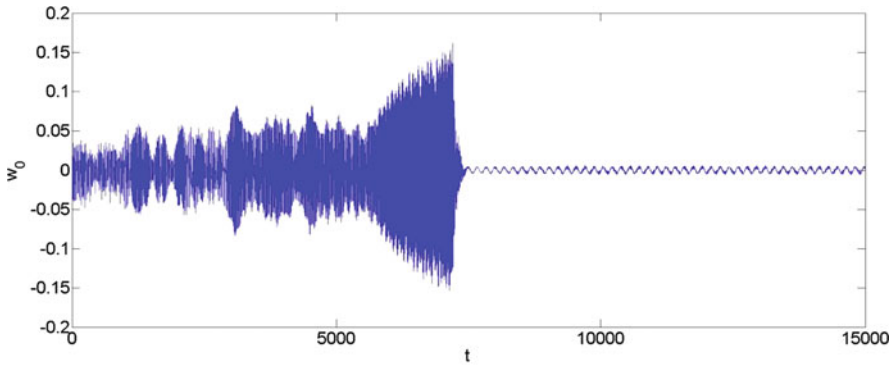


Fig. 4.12 The wave diagram of  $w_p$  with the application of the control strategy

$$\kappa = 2, k_{fs} = 0.01, \Delta F(w, \dot{w}) = 0.001 \sin(w_p).$$

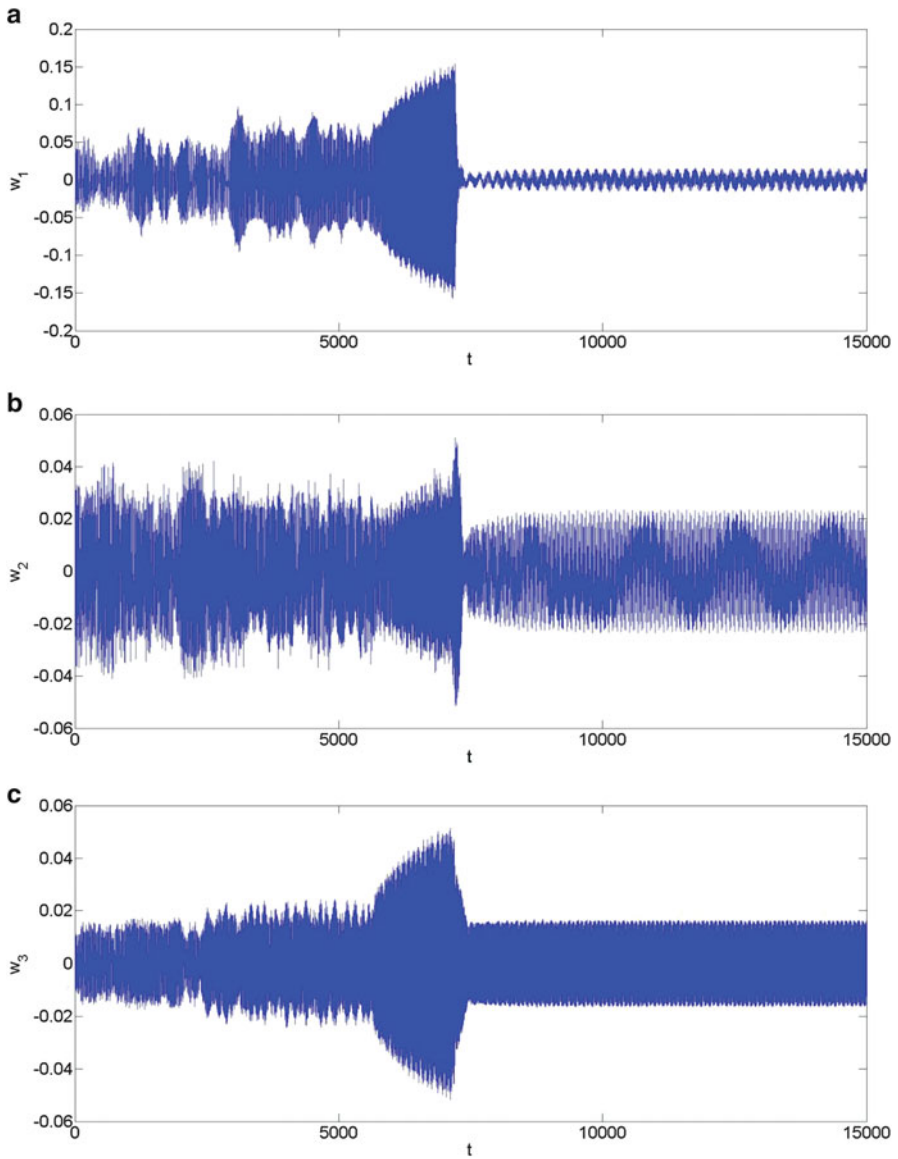
As shown in Fig. 4.12, the proposed control strategy is applied at  $t = 7,200$ . After the application of the control strategy, the responses of the cable at the translating speed  $v_0 = 2m/s$  are shown in Figs. 4.12 and 4.13.

In Fig. 4.12 the responses of the selected point,  $w_p$ , is shown, for the period of time from  $t = 7,200$  to  $t = 15,000$ . It can be seen from Fig. 4.12, a short period of time is needed for stabilizing the beam after the application of the control strategy. After the short period, the chaotic motion will then become a periodic one, of which the amplitude is 0.005.

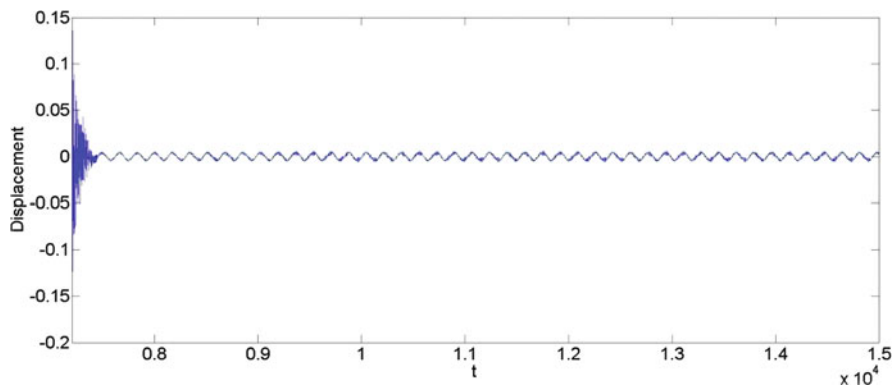
From Fig. 4.13a–c, the displacements are shown in terms of  $w_1$ ,  $w_2$ , and  $w_3$ . Based on these figures, through the application of the control strategy, each of the displacements of the axially translating cable is gradually stabilized from a chaotic motion into a periodic one.

Figure 4.14 shows the comparison between the actual response of the beam  $w_p$  and the reference signal  $w_r$  applied. One may notice that the reference signal  $w_r$  is

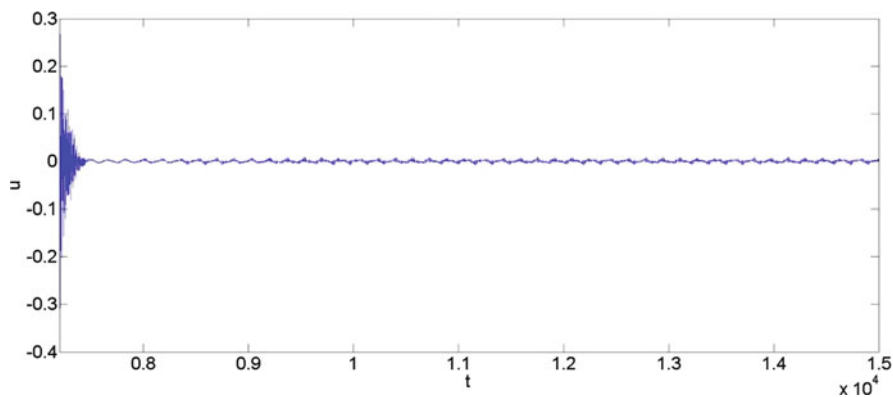




**Fig. 4.13** (a) The contribution of the first vibration mode to  $w_p$  with the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  with the application of the control strategy. (c) The contribution of the third vibration mode to  $w_p$  with the application of the control strategy



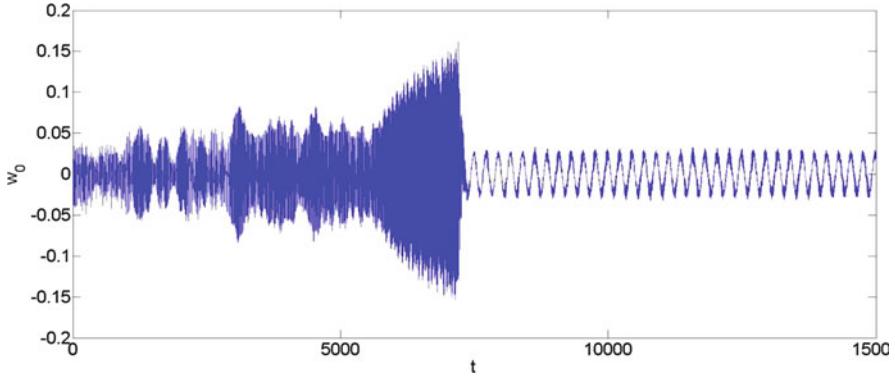
**Fig. 4.14** The wave diagram of  $w_p$  (the blue continuous line) and the reference signal  $w_r$  (the green dash line)



**Fig. 4.15** The control input

well periodic with respect to time  $t$ . One may also see from the figure, the maximum amplitude of the cable's response slightly varies after the stabilization of the cable with the application of the control strategy. As shown in the figure, the maximum amplitude of the cable is very close to that of the reference signal.

Figure 4.15 shows the control input  $U$ . Initially, the control input reaches a peak very quickly after the application of the control strategy. Once the system is stabilized, the control input displays a periodic wave diagram as shown in the figure and the maximum value of the control input is significantly decreased to a very small value.



**Fig. 4.16** The wave diagram of  $w_p$  with the application of the control strategy

### 4.6.3 Frequency Synchronization

The control strategy will be applied to show its effectiveness and efficiency in the frequency synchronization. Three different desired frequencies will be specified as the amplitude of the reference signal.

$$\omega_r = 0.0277$$

In the application of the control strategy, the desired frequency of the reference is set as

$$w_r = 0.03 \sin(0.0277t).$$

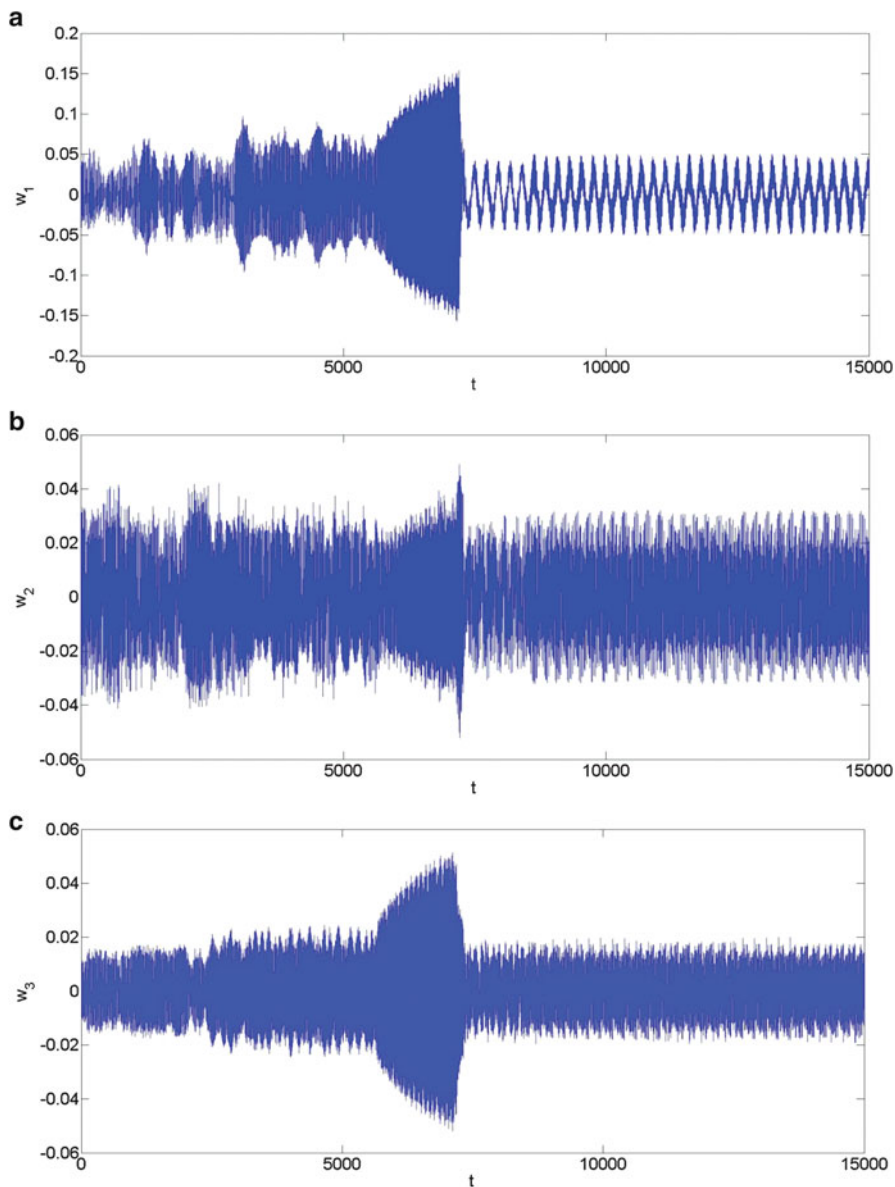
and other the control parameters and the unknown external disturbance take the following values:

$$\kappa = 3, k_{fs} = 0.01, \Delta F(w, \dot{w}) = 0.001 \sin(w_p).$$

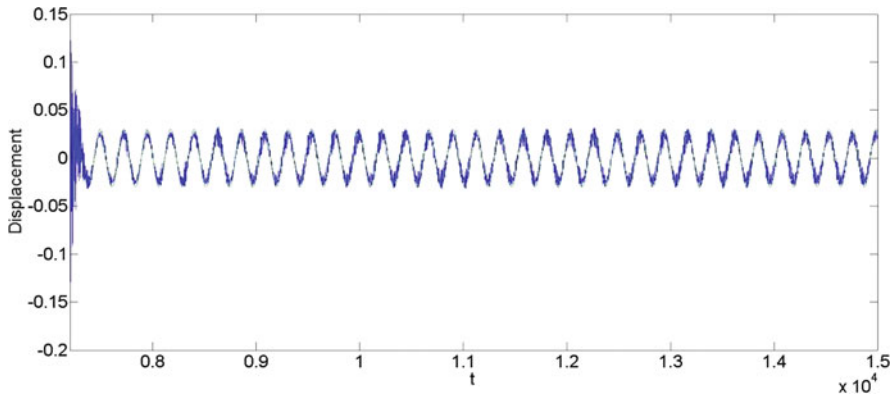
As shown in Fig. 4.16, the proposed control strategy is applied at  $t = 7,200$ . After the application of the control strategy, the responses of the cable at the translating speed  $v_0 = 2m/s$  are shown in Figs. 4.16 and 4.17.

In Fig. 4.16 the responses of the selected point,  $w_p$ , is shown, for the period of time from  $t = 7,200$  to  $t = 15,000$ . It can be seen from Fig. 4.16, a short period of time is needed for stabilizing the beam after the application of the control strategy. After the short period, the chaotic motion will then become a periodic one, of which the frequency is 0.0277.

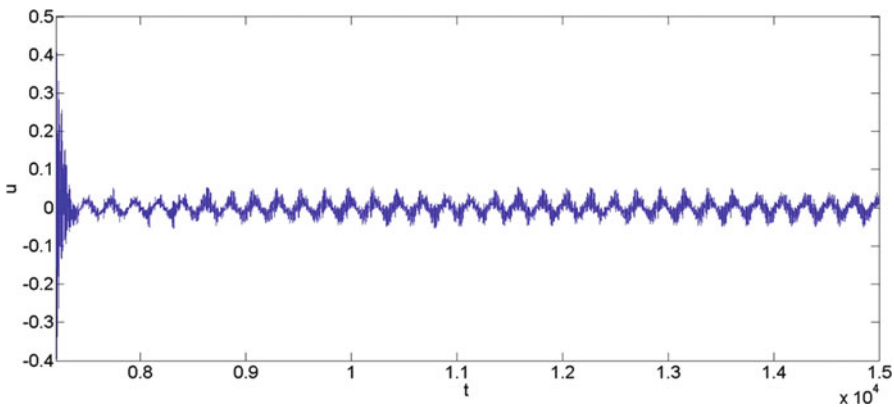
From Fig. 4.17a–c, the displacements are shown in terms of  $w_1$ ,  $w_2$ , and  $w_3$ . Based on these figures, through the application of the control strategy, each of the displacements of the axially translating cable is gradually stabilized from a chaotic motion into a periodic one.



**Fig. 4.17** (a) The contribution of the first vibration mode to  $w_p$  with the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  with the application of the control strategy. (c) The contribution of the third vibration mode to  $w_p$  with the application of the control strategy



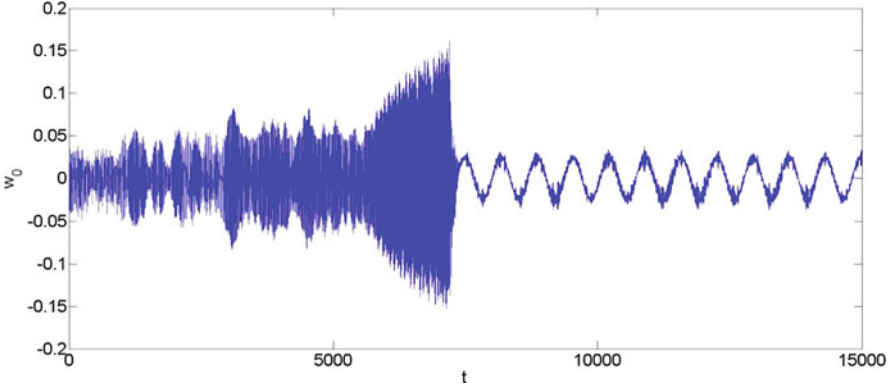
**Fig. 4.18** The wave diagram of  $w_p$  (the blue continuous line) and the reference signal  $w_r$  (the green dash line)



**Fig. 4.19** The control input

Figure 4.18 shows the comparison between the actual response of the beam  $w_p$  and the reference signal  $w_r$  applied. One may notice that the reference signal  $w_r$  is well periodic with respect to time  $t$ . One may also see from the figure, the maximum amplitude of the cable's response slightly varies after the stabilization of the cable with the application of the control strategy. As shown in the figure, the frequency of the cable is very close to that of the reference signal.

Figure 4.19 shows the control input  $U$ . Initially, the control input reaches a peak very quickly after the application of the control strategy. Once the system is stabilized, the control input displays a periodic wave diagram as shown in the figure and the maximum value of the control input is significantly decreased to a very small value.



**Fig. 4.20** The wave diagram of  $w_p$  with the application of the control strategy

$\omega_r = 0.0092$

In the application of the control strategy, the desired frequency of the reference is set as

$$w_r = 0.03 \sin(0.0092t).$$

and other the control parameters and the unknown external disturbance take the following values:

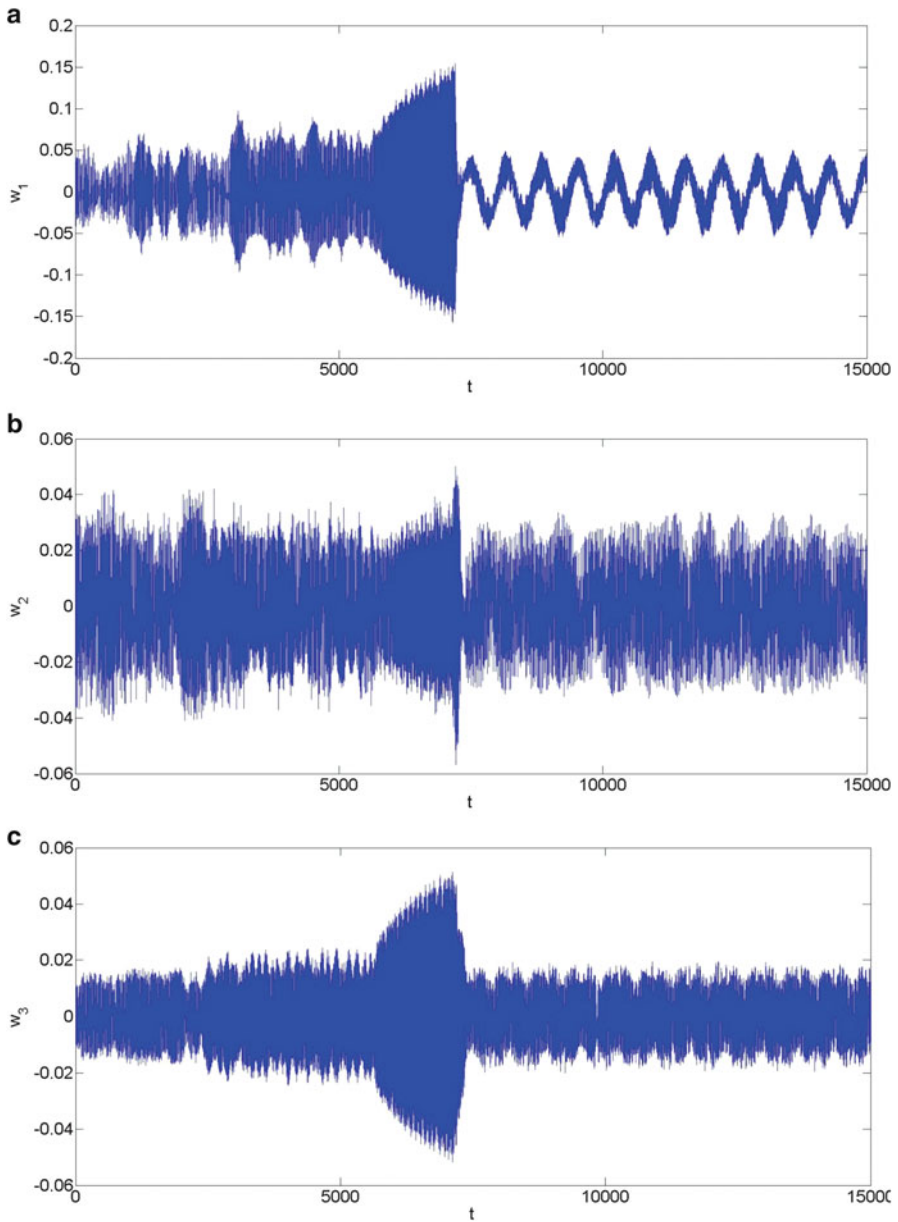
$$\kappa = 3, k_{fs} = 0.01, \Delta F(w, \dot{w}) = 0.001 \sin(w_p).$$

As shown in Fig. 4.20, the proposed control strategy is applied at  $t = 7,200$ . After the application of the control strategy, the responses of the cable at the translating speed  $v_0 = 2m/s$  are shown in Figs. 4.20 and 4.21.

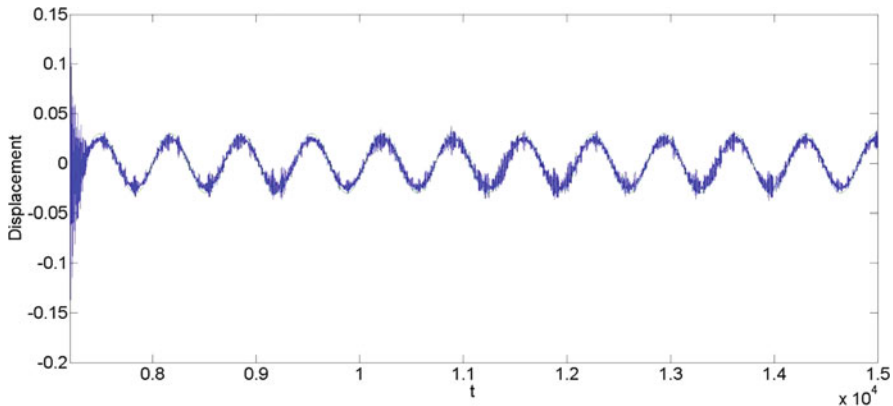
In Fig. 4.20 the responses of the selected point,  $w_p$ , is shown, for the period of time from  $t = 7,200$  to  $t = 1,500$ . It can be seen from Fig. 4.20, a short period of time is needed for stabilizing the beam after the application of the control strategy. After the short period, the chaotic motion will then become a periodic one, of which the frequency is 0.0092.

From Fig. 4.21a–c, the displacements are shown in terms of  $w_1$ ,  $w_2$ , and  $w_3$ . Based on these figures, through the application of the control strategy, each of the displacements of the axially translating cable is gradually stabilized from a chaotic motion into a periodic one.

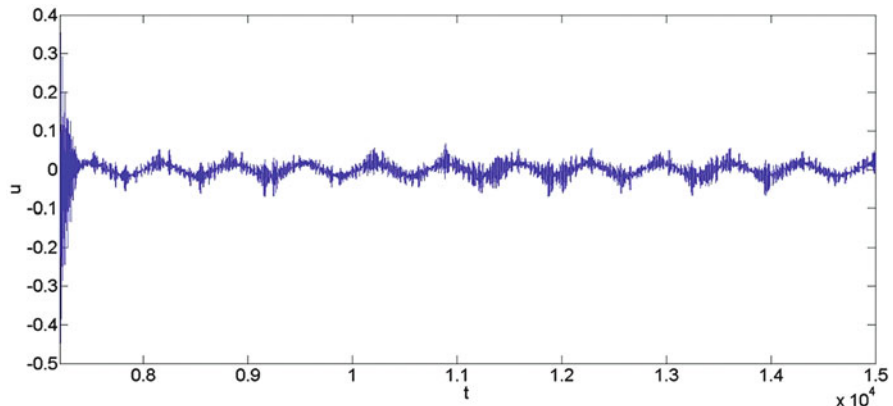
Figure 4.22 shows the comparison between the actual response of the beam  $w_p$  and the reference signal  $w_r$  applied. One may notice that the reference signal  $w_r$  is well periodic with respect to time  $t$ . One may also see from the figure, the maximum amplitude of the cable's response slightly varies after the stabilization of the cable



**Fig. 4.21** (a) The contribution of the first vibration mode to  $w_p$  with the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  with the application of the control strategy. (c) The contribution of the third vibration mode to  $w_p$  with the application of the control strategy



**Fig. 4.22** The wave diagram of  $w_p$  (the blue continuous line) and the reference signal  $w_r$  (the green dash line)

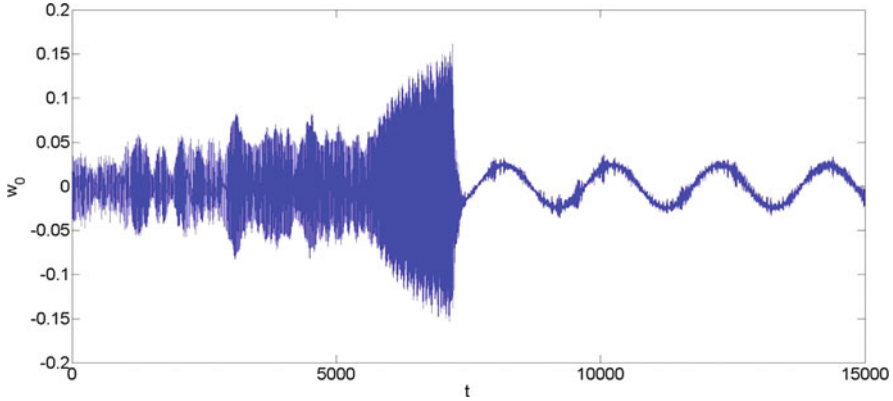


**Fig. 4.23** The control input

with the application of the control strategy. As shown in the figure, the frequency of the cable is very close to that of the reference signal.

Figure 4.23 shows the control input  $U$ . Initially, the control input reaches a peak very quickly after the application of the control strategy. Once the system is stabilized, the control input displays a periodic wave diagram as shown in the figure and the maximum value of the control input is significantly decreased to a very small value.





**Fig. 4.24** The wave diagram of  $w_p$  with the application of the control strategy

$$\omega_r = 0.0031$$

In the application of the control strategy, the desired frequency of the reference is set as

$$w_r = 0.03 \sin(0.0031t).$$

and other the control parameters and the unknown external disturbance take the following values:

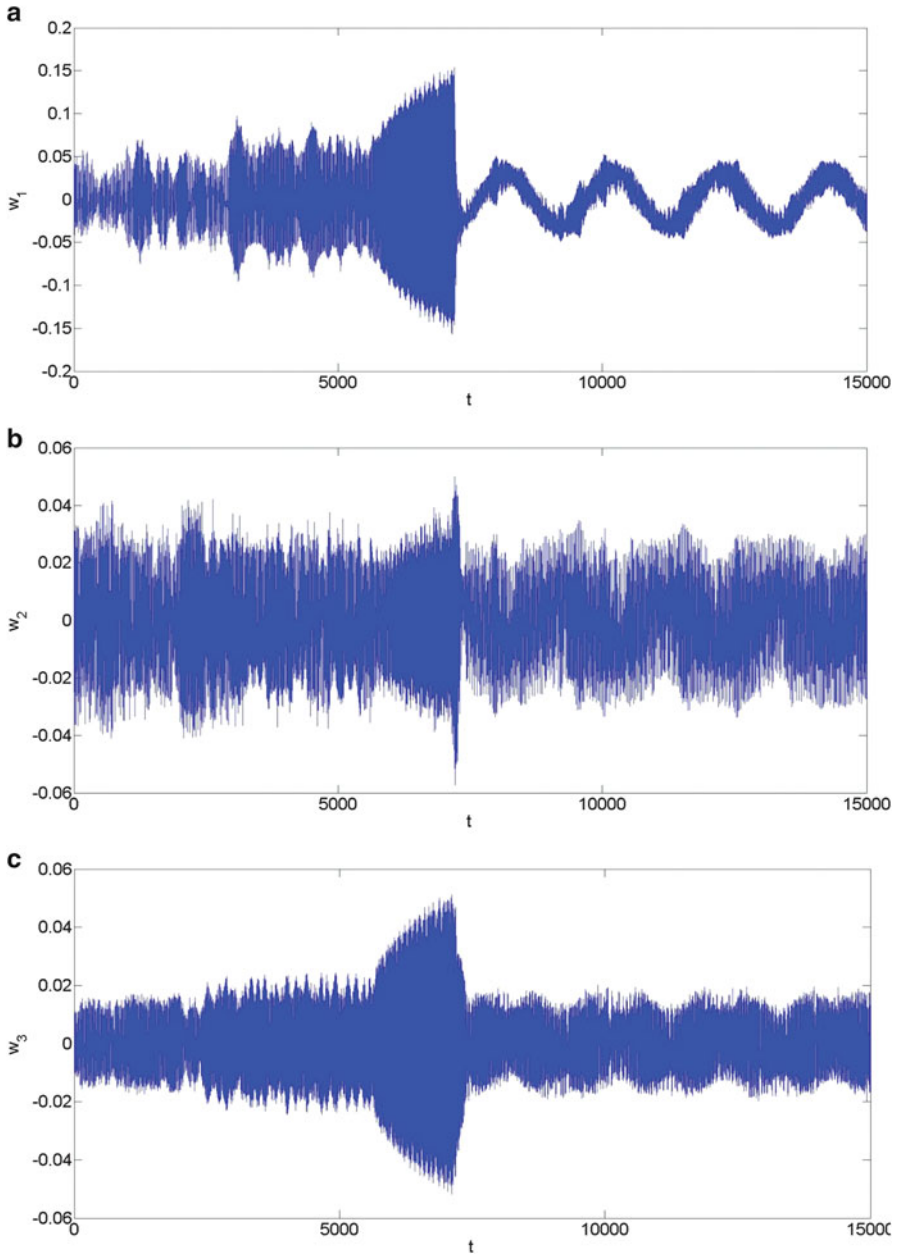
$$\kappa = 3, k_{fs} = 0.01, \Delta F(w, \dot{w}) = 0.001 \sin(w_p).$$

As shown in Fig. 4.24, the proposed control strategy is applied at  $t = 7,200$ . After the application of the control strategy, the responses of the cable at the translating speed  $v_0 = 2m/s$  are shown in Figs. 4.24 and 4.25.

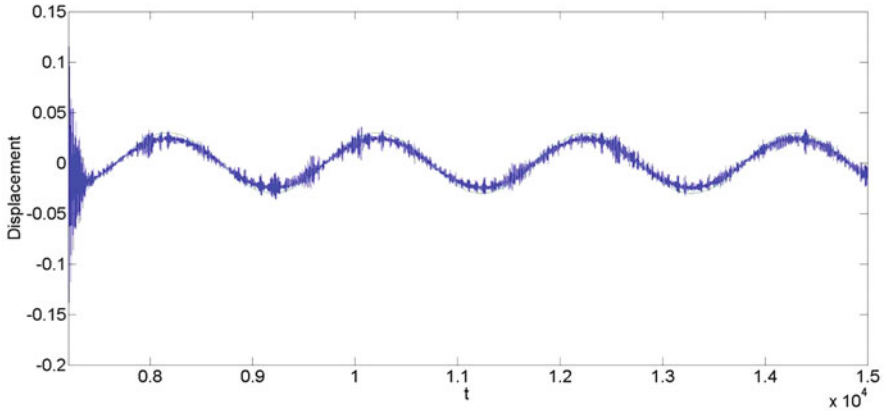
In Fig. 4.24 the responses of the selected point,  $w_p$ , is shown, for the period of time from  $t = 7,200$  to  $t = 15,000$ . It can be seen from Fig. 4.24, a short period of time is needed for stabilizing the beam after the application of the control strategy. After the short period, the chaotic motion will then become a periodic one, of which the frequency is 0.0031.

From Fig. 4.25a–c, the displacements are shown in terms of  $w_1$ ,  $w_2$ , and  $w_3$ . Based on these figures, through the application of the control strategy, each of the displacements of the axially translating cable is gradually stabilized from a chaotic motion into a periodic one.

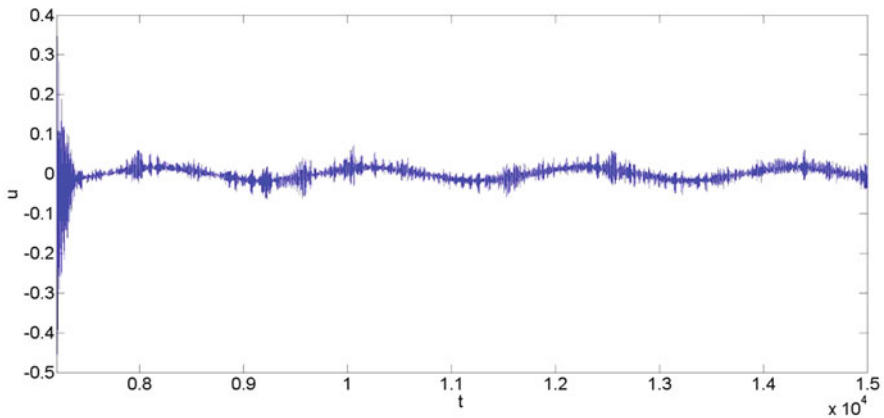
Figure 4.26 shows the comparison between the actual response of the beam  $w_p$  and the reference signal  $w_r$  applied. One may notice that the reference signal  $w_r$  is well periodic with respect to time  $t$ . One may also see from the figure, the maximum amplitude of the cable's response slightly varies after the stabilization of the cable



**Fig. 4.25** (a) The contribution of the first vibration mode to  $w_p$  with the application of the control strategy. (b) The contribution of the second vibration mode to  $w_p$  with the application of the control strategy. (c) The contribution of the third vibration mode to  $w_p$  with the application of the control strategy



**Fig. 4.26** The wave diagram of  $w_p$  (the blue continuous line) and the reference signal  $w_r$  (the green dash line)



**Fig. 4.27** The control input

with the application of the control strategy. As shown in the figure, the frequency of the cable is very close to that of the reference signal.

Figure 4.27 shows the control input  $U$ . Initially, the control input reaches a peak very quickly after the application of the control strategy. Once the system is stabilized, the control input displays a periodic wave diagram as shown in the figure and the maximum value of the control input is significantly decreased to a very small value.

## Key Symbols

$v_0$	The constant rate that the cable is axially moving at
$l$	The length of the cable
$u_0(x(t), t)$	The displacement along the horizontal axis
$w_0(x(t), t)$	The displacement along the vertical axis
$t$	The dimensional time
$x(t)$	The position of any point of the translating cable
$\mathbf{i}$	The unit vector of the fixed Cartesian coordinate along the horizontal axis
$\mathbf{k}$	The unit vector of the fixed Cartesian coordinate along the vertical axis
$\mathbf{r}$	The position vector of any point of the translating cable without deformation
$\mathbf{R}$	The displacement field of the axially translating cable
$V$	The volume of the cable
$T$	The kinetic energy of the cable
$\rho$	The mass of the cable per unit length
$\varepsilon_{11}$	The strain along the horizontal axis
$U$	The total strain energy of the cable
$Q_{11}$	The elastic coefficient along the direction of $\varepsilon_{11}$
$W$	The virtual work done by force
$L$	The Lagrangian function
$\bar{t}$	The non-dimensional time
$\tau$	The non-dimensional coefficients corresponding to time
$\bar{x}$	The non-dimensional position of any point of the translating cable
$\bar{v}_0$	The non-dimensional rate the cable is moving at
$\bar{w}_0$	The non-dimensional displacement along the vertical axis
$\eta$	The non-dimensional coefficients corresponding to the geometric nonlinearity
$\varphi_n(\bar{x})$	The comparison function
$\bar{w}_n(\bar{t})$	The $n$ th vibration mode
$\Phi$	The general form of a nonlinear governing equation
$U$	The total control input
$\Delta P(w, \dot{w})$	The unknown external disturbance
$\phi_i(\mathbf{W}, t)$	The corresponding expression of $\Phi(w, \dot{w}, t)$ via the Galerkin method
$u_i$	The corresponding expression of $U$ via the Galerkin method
$\Delta f_i(\mathbf{W}, t)$	The corresponding expression of $\Delta P(w, \dot{w})$ via the Galerkin method
$w_p$	The non-dimensional response of a selected point of the cable
$x_p$	The non-dimensional location of a selected point of the cable
$w_r$	The reference signal
$w_r$	The amplitude of the reference signal
$\omega_r$	The frequency of the reference signal
$U_{eq}$	A part of the control input

$U_r$	A part of the control input
$U_{fs}$	The control variable via fuzzy rule
$\kappa$	The control parameters governing the sliding surface
$k_{fs}$	The control parameter corresponding to $\Delta P(w, \dot{w})$
$R^+$	The positive real number

## References

- Abou-Rayan AM, Nayfeh AH, Mook DT, Nayfeh AM (1993) Nonlinear response of a parametrically excited bucked beam. *Nonlinear Dyn* 4:499–525
- Carrera E, Giunta G, Petrolo M (2011) *Beam structures: classical and advanced theories*. Wiley, New York
- Dai L (2008) *Nonlinear dynamics of piece constant systems and implementation of piecewise constant arguments*. World Scientific, New Jersey
- Fung RF, Lu PY, Tseng CC (1998) Non-linearly dynamic modelling of an axially moving beam with a tip mass. *J Sound Vib* 291:559–571
- Ghayesh MH, Amabili M (2013) Nonlinear dynamics of an axially moving Timoshenko beam with an internal resonance. *Nonlinear Dyn* 73:39–52
- Haghighi HH, Markazi AHD (2009) Chaos prediction and control in MEMS resonators. *Commun Nonlinear Sci* 15:3091–3099
- Humer A (2013) Dynamic modeling of beams with non-material, deformation-dependent boundary conditions. *J Sound Vib* 332:622–641
- Kuo CL, Shieh CS, Lin CH, Shih SP (2007) Design of an adaptive fuzzy sliding-mode controller for chaos synchronization. *Int J Nonlinear Sci* 8:631–636
- Luo AC, Hamidzadeh HR (2004) Equilibrium and buckling stability for axially traveling plates. *Commun Nonlinear Sci Numer Simulat* 9:343–3360
- Nayfeh AH, Mook DT (1979) *Nonlinear oscillation*. Wiley-Interscience, New York
- Tabarrok B, Leech CM, Kim YI (1974) On the dynamics of an axially moving beam. *J Franklin Inst* 297:201–220
- Tang JL, Ren GX, Zhu WD, Ren H (2011) Dynamics of variable-length tethers with application to tethered satellite deployment. *Commun Nonlinear Sci Numer Simulat* 16:3411–3424
- Utkin VI (1992) *Sliding modes in control and optimization*. Springer, Berlin
- Wickert JA (1992) Non-linear vibration of a travelling tensioned beam. *Int J Mech* 27:503–517
- Yang KJ, Hong KS, Matsuno F (2005) Energy-based control of axially translating beams: varying tension, varying speed, and disturbance adaptation. *IEEE Trans Control Syst Technol* 13:1045–1054
- Yau HT, Kuo CL (2006) Fuzzy sliding mode control for a class of chaos synchronization with uncertainties. *Int J Nonlinear Sci Numer Simulat* 7:333–338
- Yau HT, Wang CC, Hsieh CT, Cho CC (2010) Nonlinear analysis and control of the uncertain micro-electro-mechanical system by using a fuzzy sliding mode control design. *Comput Math Appl* 61:1912–1916
- Zhu WD (2002) Control volume and system formulations for translating media and stationary media with moving boundaries. *J Sound Vib* 254:189–201
- Zhu WD, Ni J (2000) Energetic and stability of translating media with an arbitrarily varying length. *J Vibrot Acoust* 122:295–304

# Chapter 5

## Nonlinear Initial Value Ordinary Differential Equations

Mohammad M. Aghdam, Ali Fallah, and Poorya Haghi

**Keywords** Multi-step method • Nonlinear differential equation • Initial value problem • Error analysis

### 5.1 Introduction

Ordinary differential equations (ODEs) present the mathematical models of many physical problems in different fields of science, engineering, and economy. ODEs are categorized into initial value problems (IVP) and boundary value problem (BVP). BVPs usually explain the equilibrium state of the systems such as steady state heat or mass transfer while IVPs present the dynamic behavior of the system such as motion of system particles subjected to external load or transient heat/mass transfer. In this chapter, IVPs are considered. An initial value ODE in the general case is in the form of:

$$\frac{dy}{dx} = f(x, y), \quad x \in [a, b], \quad y(a) = y_0 \quad (5.1)$$

where  $f(x, y)$  can be linear or nonlinear function of  $x$  and  $y$ , and  $y_0$  is initial condition. It is worth mentioning that, higher order ODEs and systems of ODEs can also be shown in the form of Eq. (5.1). Thus, the presented method is applicable to systems of ODEs of any order.

---

M.M. Aghdam (✉) • P. Haghi  
Department of Mechanical Engineering, Amirkabir University of Technology, Tehran, Iran  
e-mail: [aghdam@aut.ac.ir](mailto:aghdam@aut.ac.ir); [phaghi@purdue.edu](mailto:phaghi@purdue.edu)

A. Fallah  
Department of Mechanical Engineering, Sharif University of Technology, Tehran, Iran  
e-mail: [a\\_fallah@mech.sharif.edu](mailto:a_fallah@mech.sharif.edu)

Except in some special cases such as constant coefficient linear ODEs, analytical solutions for ODEs are generally not available. Therefore, various numerical methods are proposed for the solution of initial value ODEs such as methods based on Taylor series, one-step methods such as Euler method, multi-stage single-step methods such as Runge–Kutta methods and linear multi-step methods. Single-step methods only require information about the solution at one previous point such as  $x = x_{n-1}$  to compute the solution at an advanced point such as  $x = x_n$ . For instance, Euler method is a single-step formula for the solution of IVPs that predicts the solution of the problem at  $x = x_n$ , only from slope of the  $y(x)$  at  $x = x_{n-1}$  (Süli and Mayers 2003):

$$y'_n = \frac{y_n - y_{n-1}}{h} = f(x_{n-1}, y_{n-1}) = f_{n-1} \rightarrow y_n = y_{n-1} + hf_{n-1} \quad (5.2)$$

where  $h$  is the step size, i.e.  $h = x_n - x_{n-1}$  and  $y_{n-1} = y(x_{n-1})$ ,  $f_{n-1} = f(x_{n-1}, y_{n-1})$ . Equation (5.2) is called explicit formula because the right hand side of (5.2) can be explicitly computed from known values of  $y_{n-1}$ . In spite of the simplicity of the Euler method, (5.2), it is not generally employed in applied cases. There are two main reasons why Euler's method is *not* generally used in scientific computing. Firstly, the truncation error per step associated with this method is far larger than those associated with other, more advanced methods for a given value of  $h$ . Secondly, Euler's method is too prone to numerical instabilities. The main reason that Euler's method has such large truncation errors per step is that the method only evaluates derivatives at the beginning of the interval: i.e., at  $x_{n-1}$  to compute the solution from  $x_{n-1}$  to  $x_n$ . Therefore, the method is very *asymmetric* with respect to the beginning and the end of the interval.

Multi-stage methods such as Runge–Kutta were developed to overcome this shortcoming. In multi-stage methods, more symmetric integration is employed by evaluating the derivatives in some intermediate points in the considered interval. For instance, in the second-order Runge–Kutta method, derivatives are evaluated by using an Euler-like trial stage to the midpoint of the interval and then using the values of both  $x$  and  $y$  at the midpoint to make the actual step across the interval, see Fig. 5.1. The second-order Runge–Kutta formula is (Süli and Mayers 2003):

$$\begin{aligned} k_1 &= hf_{n-1} \\ k_2 &= hf\left(x_{n-1} + \frac{h}{2}, y_{n-1} + \frac{k_1}{2}\right) \\ y_n &= y_{n-1} + k_2 \end{aligned} \quad (5.3)$$

Therefore, the Euler's method is equivalent to a first-order Runge–Kutta method. Obviously, there is no need to stop at a second-order method and it is possible to achieve higher order of accuracy using more trial stages per interval. The most common Runge–Kutta method is the fourth-order Runge–Kutta (RK4). It is worth mentioning that multi-stage methods such as Runge–Kutta methods are still single-step methods because only the information of the previous point is used to obtain the current point value. In spite of higher order of accuracy of the multi-stage methods with respect to simple single-step methods, they need many derivative calculations in each step.

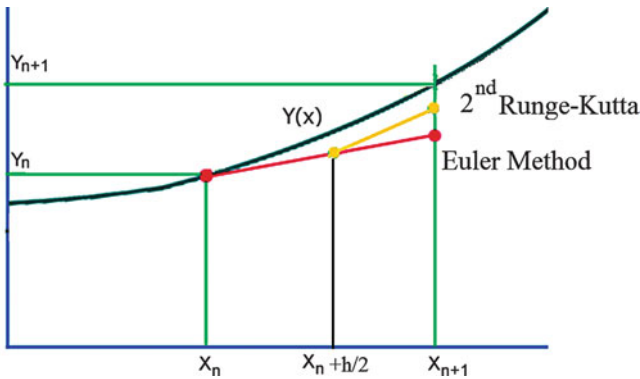


Fig. 5.1 Comparison between Euler method and second-order Runge–Kutta method

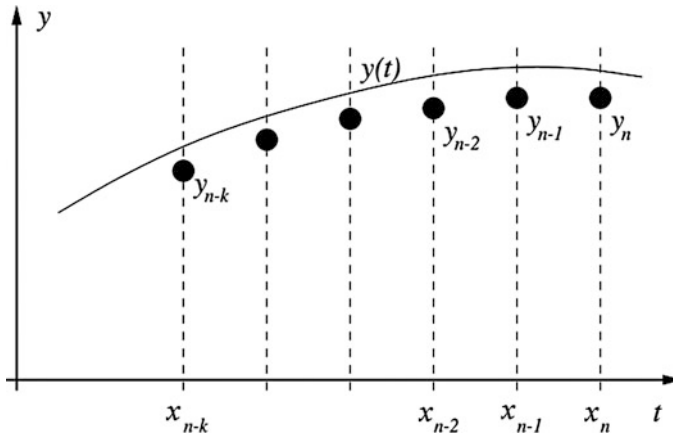


Fig. 5.2 Domain of k-step method

Multi-step methods use previous point information to increase the accuracy of the results instead of using trial stages in each interval, see Fig. 5.2.

Consider the general linear or nonlinear first-order ordinary differential equation (5.1). The integral form of (5.1) can be expressed as:

$$y(x) = y_0 + \int_a^x f(x, y) dx \tag{5.4}$$

In multi-step techniques, numerical solution is obtained by using a discretized version of (5.4) as:

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} f(x, y) dx \tag{5.5}$$



Various multi-step methods use different ways to approximate the integral term in (5.5). For instance, Adams–Moulton method, one of the most common multi-step methods, employs an approximation of  $f(x, y)$  in (5.5) by using the Newton backward difference method to obtain a multi-step formulation.

Although the existing multi-step methods such as Adams–Moulton method are accurate and useful, they also have their own limitations in terms of accuracy and stability at large step sizes and weak performance in the case of stiff ODEs. Therefore, providing more accurate and/or stable predictions while increasing step size to reduce computational cost would be interesting for relevant researchers. In order to achieve this goal, in the present work, the well-known Bernstein polynomials as the basis for Bézier curves are employed. Bernstein polynomials (special case of B-spline) are piecewise polynomials that represent a great variety of functions. They can be differentiated and integrated without difficulty and construct spline curves to approximate functions with desired accuracy. There has been widespread interest in the use of these polynomials to obtain solutions for various types of differential equations (Caglar and Caglar 2006; Jator and Sinkala 2007; Khalifa et al. 2008; Caglar et al. 2008; Zerarka and Nine 2008; Bhattia and Brackenb 2007). In these studies, the solution is expanded in terms of a set of continuous Bernstein or B-spline polynomials over a closed interval and then by employing the weighted residual methods such as Galerkin or collocation to determine the expansion coefficients to construct a solution. However, in the current work, the Bernstein polynomials and Bézier curve are used to evaluate the integral in (5.5) by providing a general  $k$ -step formula.

Results of the presented formula show improved stability in comparison with Adams–Moulton method for the same step size. This leads to significant reduction of computational time. In addition, stability and convergence of the presented formula are discussed through common stability and error analysis of multi-step methods.

## 5.2 Bézier Curves

Before computer graphics ever existed, engineers designed aircraft wings and automobile bodies using splines. A spline is a long flexible piece of wood or plastic with a rectangular cross section held in place at various positions by heavy lead weights with a protrusion called *duck*, where the duck holds the spline in a fixed position against the drawing board (Beach 1991). The spline then conforms to a natural shape between the ducks. By moving the ducks around, the designer can change the shape of the spline. The drawbacks are obvious. Recording duck positions and maintaining the drafting equipment necessary for many complex parts will take up square footage in a storage facility, costs that would be absorbed by a consumer. A not so obvious drawback is that when analyzed mathematically, there is no closed form solution (Buss 2003).

In order to overcome these shortages, in the 1960s a mathematician and engineer named Pierre Bézier introduced a family of smooth curves based on the Bernstein basis polynomials. Bézier used the curves mainly to design various parts of automobile bodies such as exterior car panels for his work at Renault in the 1960s.

The advent of Bézier curves revolutionized the design process and helped the designers draw smooth looking curves on a computer screen, and use less physical storage space for design materials. Bézier's contribution to computer graphics has opened the road for computer-aided design (CAD) software like Maya, Blender, 3D Max, and Corel (Gerald et al. 1999; Lorentz et al. 1986; Farin 2002).

The  $n$ th-degree Bézier curve for  $(n + 1)$  given points  $P_0, P_1, \dots, P_n$  can be written as:

$$B(u) = \sum_{i=0}^n B_{i,n}(u)P_i, \quad u \in [0, 1] \quad (5.6)$$

where the points  $P_i$  are called *control points* for the Bézier curves and  $B_{i,n}$  are known as *Bernstein basis polynomials* of degree  $n$  defined as:

$$B_{i,n}(u) = \binom{n}{i} (1-u)^{n-i} u^i, \quad i = 0, 1, \dots, n \quad (5.7)$$

It is worth mentioning here that quadratic and cubic Bézier curves are the most common types of Bézier curves since higher degree curves are more computationally expensive to evaluate. Figure 5.3 and Table 5.1 show quadratic and cubic Bernstein basis polynomials.

The location of control points in the 2D Cartesian coordinate system is:

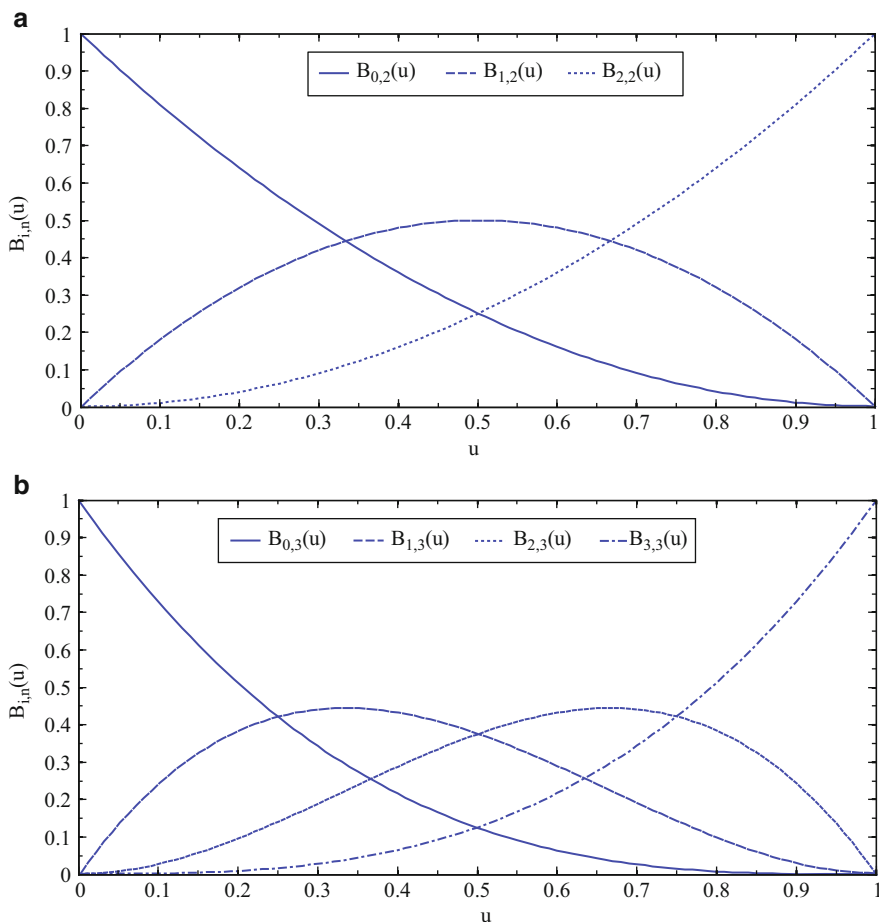
$$P_i = \begin{Bmatrix} x_i \\ y_i \end{Bmatrix} \quad (5.8)$$

Using (5.8), one can rewrite (5.6) in the parametric form of:

$$x(u) = \sum_{i=0}^n B_{i,n}(u)x_i \quad (5.9.1)$$

$$y(u) = \sum_{i=0}^n B_{i,n}(u)y_i \quad (5.9.2)$$

Bézier curves are widely used in computer graphics and computer-aided design, mainly to provide smooth curves. It should be noted that similar to the least-square curves, Bézier curves are not really interpolating curves, as they do not normally pass through all of the control points. However, they have the important property of staying within the polygon determined by the given points as shown in Fig. 5.4.



**Fig. 5.3** Bernstein polynomials (a) quadratic  $n = 2$ , (b) cubic  $n = 3$

**Table 5.1** Linear, quadratic, and cubic Bernstein polynomials

$n$	$B_{i,n}(u)$
1	$B_{0,1}(u) = 1 - u, B_{1,1}(u) = u$
2	$B_{0,2}(u) = (1 - u)^2, B_{1,2}(u) = 2u(1 - u), B_{2,2}(u) = u^2$
3	$B_{0,3}(u) = (1 - u)^3, B_{1,3}(u) = 3u(1 - u)^2, B_{2,3}(u) = 3u^2(1 - u), B_{3,3}(u) = u^3$

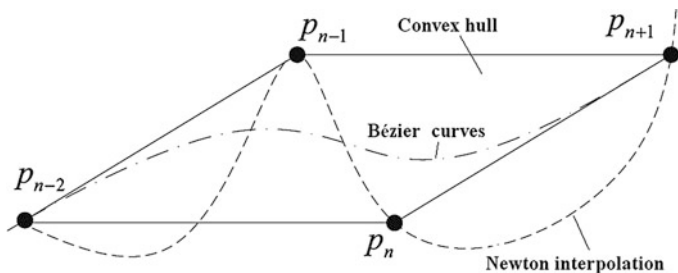


Fig. 5.4 Comparing Newton interpolation with Bézier curves and Convex hull property

More details of Bézier curves and their properties can be found elsewhere (Gerald et al. 1999; Lorentz et al. 1986; Farin 2002).

### 5.3 Technique

In the present study, Bézier curves of any order are used to develop a general  $k$ -step formula that becomes unstable at larger step sizes compared to the existing multi-step methods. To this end, in the integral of (5.5) is approximated by  $m$ th-degree Bézier curves using control points which are defined as:

$$P_{n-j} = \left\{ \begin{matrix} x_{n-j} \\ f_{n-j} \end{matrix} \right\} = \left\{ \begin{matrix} x_n - jh \\ f(x_{n-j}, y_{n-j}) \end{matrix} \right\}, \quad j = n - m \dots n \tag{5.10}$$

Using Eqs. (9) and (5.10), one can write the new parametric form of  $m$ th-degree Bézier curves as:

$$x(u) = \sum_{i=0}^m B_{i,m}(u)x_{n-m+i} = \varnothing(u) \tag{5.11.1}$$

$$f(x, y) = \sum_{i=0}^m B_{i,m}(u)f_{n-m+i} = \varphi(u) \tag{5.11.2}$$

Finally, after approximation of the integral term in (5.5), one may rewrite (5.5) as:

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} f(x, y) dx \cong y_n + \int_{\varnothing^{-1}(x_n)}^{\varnothing^{-1}(x_{n+1})} \varphi(u)\varnothing'(u)du \tag{5.12}$$

Using general formula (5.12) together with various orders of Bézier curves in (5.6) leads to different multi-step formulas. For instance, assuming quadratic Bézier

curve leads to three-step formula. Different degrees of Bézier curves correspond to different step formulas that are:

Two-step formula:

$$y_{n+1} = y_n + \frac{h}{2} (3f_n - f_{n-1}) \quad (5.13)$$

Three-step formula:

$$y_{n+1} = y_n + \frac{h}{12} (19f_n - 8f_{n-1} + f_{n-2}) \quad (5.14)$$

Four-step formula:

$$y_{n+1} = y_n + \frac{h}{108} (175f_n - 81f_{n-1} + 15f_{n-2} - f_{n-3}) \quad (5.15)$$

and general  $m + 1$ -step formula:

$$y_{n+1} = y_n + h \sum_{i=0}^m \left( m \int_1^{\frac{m+1}{m}} \binom{m}{i} (1-u)^{m-i} u^i du \right) f_{n-m+i} \quad (5.16)$$

## 5.4 Consistency, Stability, and Convergence Analyses

The main concepts in the analysis of multi-step methods, and indeed any numerical method for differential equations, are consistency, stability, and convergence. The basic theory for analysis of multi-step methods has been developed by Dahlquist (1963) and is presented in the classical book of Henrici (1962).

It is useful to employ a general framework for the presentation of general  $m + 1$ -step methods as:

$$a_{m+1}y_{n+1} + a_m y_n + \cdots + a_0 y_{n-(m+1)} = h [b_{m+1} f_{n+1} + \cdots + b_0 f_{n-(m+1)}] \quad (5.17)$$

Associated with (5.17) are also two *characteristics* polynomials of the  $m + 1$ -step formula (5.17) as:

$$p(z) = a_{m+1}z^{m+1} + a_m z^m + \cdots + a_1 z + a_0 \quad (5.18)$$

$$q(z) = b_{m+1}z^{m+1} + b_m z^m + \cdots + b_1 z + b_0 \quad (5.19)$$

A multi-step formula is in the implicit form if  $b_{m+1} \neq 0$  and is in explicit configuration if  $b_{m+1} = 0$ . Thus, all presented multi-step formulas in (5.13) through (5.16) are in explicit form.

**Definition 1** For a given  $m + 1$ -step formula Eq. (5.17) the polynomials

$$p(z) = \sum_{i=0}^{m+1} a_i z^i \quad (5.20)$$

$$q(z) = \sum_{i=0}^{m+1} b_i z^i$$

are called first and second characteristic polynomials, respectively.

A multi-step method is called consistent, if it reproduces the exact solution for the differential equation  $y' = 1$ , when exact starting approximations are used.

**Definition 2 (Süli and Mayers 2003)** The multi-step formula (5.17) is consistent if  $p(1) = 0$  and  $p'(1) = q(1)$ .

**Theorem 1** General  $m + 1$ -step formula (5.16) is consistent.

**Proof** From (5.16) and (5.18), it is clear that for the  $m + 1$ -step formula

$$p(z) = z^{m+1} - z^m \quad (5.21)$$

Using (5.21) to determine  $p(1)$  and  $p'(1)$ , one can easily conclude for the  $m + 1$ -step formula (5.16) that

$$\begin{cases} p(1) = 0 \\ p'(1) = 1 \end{cases} \quad (5.22)$$

Therefore, in order for  $m + 1$ -step formula (5.16) to be consistent, the value of  $q(1)$  must be equal to one. Using general  $m + 1$ -step formula (5.16) in conjunction with (5.20), one can determine coefficients of  $q(z)$  in (5.20) as:

$$b_i = m \int_1^{\frac{m+1}{m}} \binom{m}{i} (1-u)^{m-i} u^i du = \binom{m}{i} \left( \sum_{j=0}^{m-i} \binom{m-i}{j} (-1)^j \left( \frac{m}{i+j+1} \right) \left[ \left( \frac{m+1}{m} \right)^{i+j+1} - 1 \right] \right), \quad i = 0 \dots m \quad (5.23)$$

Substitution of (5.23) in (5.20), one can derive  $q(1)$  as:

$$q(1) = \sum_{i=0}^m b_i = \sum_{i=0}^m \left( \binom{m}{i} \left( \sum_{j=0}^{m-i} \binom{m-i}{j} (-1)^j \left( \frac{m}{i+j+1} \right) \left[ \left( \frac{m+1}{m} \right)^{i+j+1} - 1 \right] \right) \right) \quad (5.24)$$

After some mathematical simplification, (5.24) can be rewritten as:

$$\sum_{i=0}^m b_i = \sum_{i=0}^m c_i \quad (5.25)$$

where

$$c_i = \binom{m}{m-i} (-1)^{m-i} \left( \frac{m}{m+1-i} \right) \left[ \left( \frac{m+1}{m} \right)^{m+1-i} - 1 \right] (1 + (-1))^{m-i} \quad (5.26)$$

It can easily be shown that the last term in (5.26) is zero for  $i=0 \dots m-1$ . However, for  $i=m$  the value of  $c_i$  in (5.26) becomes one. Therefore, considering (5.24) and (5.26) results in:

$$q(1) = \sum_{i=0}^m b_i = \sum_{i=0}^m c_i = c_m = 1 \quad (5.27)$$

Consequently, comparing (5.22) and (5.27) with Definition 2, it can be concluded that the general  $m+1$ -step formula (5.16) is consistent.

The numerical solution of a single-step method depends on the initial condition,  $y_0$  while numerical solution of a  $k$ -step method depends on all the  $k$  starting values  $y_1, y_2, \dots, y_k$ . Thus, it is important to know whether the numerical solution is stable with respect to perturbations in the starting values or not. A multi-step formula is *stable* (zero-stable) for a certain differential equation on a given step size,  $h$ , if a perturbation in the starting values of size  $\varepsilon$  causes numerical solution over that interval to change by no more than  $\alpha\varepsilon$  for some value of  $\alpha$  which does not depend on the step size  $h$ . This is called “zero-stability” because it is enough to check the condition for the differential equation  $y' = 0$  (Süli and Mayers 2003).

**Definition 3 (Süli and Mayers 2003)** The multi-step formula in general form of (5.17) is stable (zero-stable) if all roots of  $p(z)$  lie in the disk  $|z| \leq 1$  and if each root of modulus one is simple.

**Theorem 2** General  $m+1$ -step formula (5.16) is zero-stable.

**Proof** Considering (5.21), roots of the first characteristic polynomial of general  $m+1$ -step formula (5.16) are:

$$z_i = \begin{cases} 0 & i = 1 \dots m \\ 1 & i = m + 1 \end{cases} \quad (5.28)$$

Using Definition 3, one may conclude that the  $m + 1$ -step formula (5.16) is zero-stable.

Accuracy of results of the multi-step methods depends on the starting point values and the value of step size. It is important to be assured that as the step size value decreases, difference between the results of the multi-step formula and the exact solution also decreases, and prediction tends to the exact solution. Consider multi-step formula (5.17) is employed for the solution of IVP (5.1) in domain  $a < x < b$  and the sequence of meshes as  $\{a = x_0 < x_1 < x_2 \dots < x_n = b\}$ . As the number of intervals increased, i.e.,  $n \rightarrow \infty$  the step size tends to zeros,  $h \rightarrow 0$ . The multi-step method is said to be convergent if as  $h \rightarrow 0$  the  $y_n \rightarrow y(t_n)$ . Swedish mathematician Dahlquist (1963) proved the relation between the concepts of zero-stability, consistency, and convergence of a multi-step method by the theorem named after him as Dahlquist's equivalence theorem.

**Theorem 3 (Dahlquist's Equivalence Theorem) (Dahlquist 1963)** For the multi-step formula (5.17) to be convergent, it is necessary and sufficient that the formula (5.17) to be zero-stable and consistent.

**Proof** Proof of this theorem is beyond of scope of the presented chapter as details can be found in Dahlquist (1963), Henrici (1962), and Gautschi (1997).

**Theorem 4** General  $m + 1$ -step formula (5.16) is convergent.

**Proof** Since the general  $m + 1$ -step formula (5.16) is both zero-stable and consistent, based on Dahlquist's equivalence theorem the general  $m + 1$ -step formula (5.16) is also convergent.

**Definition 4 (Süli and Mayers 2003)** A convergent multi-step formula (5.17) is said to be strongly stable if  $z = 1$  is the only root of modulus 1 of its first characteristic polynomial. If there is more than one such root, it is said to be relatively stable.

**Theorem 5** General  $m + 1$ -step formula (5.16) is strongly stable.

**Proof** According to (5.28),  $p(z)$  have  $m$  zero roots while the only one nonzero root is  $z = 1$ , thus the general  $m + 1$ -step formula (5.16) is strongly stable.

## 5.5 Numerical Results and Discussions

In this section, efficiency of the presented multi-step method in terms of accuracy, convergence, and computational cost is investigated using various numerical examples. It should be noted that all considered examples are nonlinear ODEs



and therefore, analytical solutions are not available. Predictions of the present work with four-step formula (5.15) are compared with numerical results of ODE45 MATLAB. ODE45 is a single-step solver based on an explicit fourth- or fifth-order Runge–Kutta formula for the solution of ordinary differential equations in MATLAB software (Guide et al. 2004). Furthermore, results of the well-known four-step Adams–Moulton method are also included mainly to compare stability of the presented method.

In examples, three different levels of the step size  $h$  are chosen. These step sizes are carefully selected to show three different stages. At the first stage, both presented and Adams–Moulton numeric methods are stable and accurate while in the second stage, instability initiates in the Adams–Moulton technique and finally at the third stage unstable behavior can be observed in the presented method. Furthermore, in order to examine accuracy of the method, error for both numeric techniques is reported in some examples.

In both methods, four-step formula is employed which indeed information of the first four points is needed to start solution procedure. Since  $y_0$  is the only known initial condition, the ODE45 results are used to obtain the first three points, i.e.,  $y_1, y_2, y_3$  for both presented and the Adams–Moulton method.

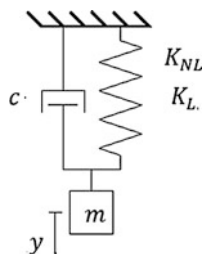
### 5.5.1 Example 1

As first example, a nonlinear mass–spring–damper system as shown in Fig. 5.5 is considered. Governing equation of such system is the well-known Duffing equation (Nayfeh et al. 2007):

$$m \frac{d^2 y}{dt^2} + c \frac{dy}{dt} + K_L y + K_{NL} y^3 = 0 \quad (5.29)$$

where  $m$ ,  $c$ ,  $K_L$ , and  $K_{NL}$  are mass, damping coefficient, linear, and nonlinear spring stiffnesses, respectively.

Introducing new variables  $y_1 = y$  and  $y_2 = \frac{dy}{dt}$ , the second-order governing nonlinear ODE (5.29) can be converted to a first order system of nonlinear ODE as:



**Fig. 5.5** Nonlinear mass–spring–damper system

$$\begin{cases} \frac{dy_1}{dt} = y_2 \\ \frac{dy_2}{dt} = -\frac{1}{m}(cy_2 + K_L y_1 + K_{NL} y_1^3) \end{cases} \quad (5.30)$$

Figure 5.6 shows predictions of presented and Adams–Moulton methods together with numerical solution obtained by ODE45 for different values of step sizes in the case of nonlinear mass–spring system with damping. Numerical results are obtained for  $m = 1$ ,  $c = 0.5$ ,  $K_L = K_{NL} = 1$  and initial condition  $y(0) = 1$ ,  $y'(0) = 0$ . In each case, errors of the Adams–Moulton ( $y_{Adams} - y_{ODE45}$ ) and presented method ( $y_{Bezier} - y_{ODE45}$ ) predictions are provided in the figures. Results show that for small values of step size ( $h = 0.1$ ), both methods are accurate and stable while for higher step size ( $h = 0.3$ ), error of the Adams–Moulton grows rapidly. For step sizes greater than  $h = 0.3$ , the Adams–Moulton becomes unstable whereas the presented method is still stable with reasonable level of accuracy for  $h = 0.5$ .

### 5.5.2 Example 2

In the second example, the Lotka–Volterra equation, which is, also known as the predator–prey equation is considered. This equation describes the dynamics of biological systems in which two species interact, one as a predator and the other as prey. The populations change through time according to the pair of equations (Berryman 1992):

$$\begin{cases} \frac{dx}{dt} = ax - bxy \\ \frac{dy}{dt} = -cy + dxy \end{cases} \quad (5.31)$$

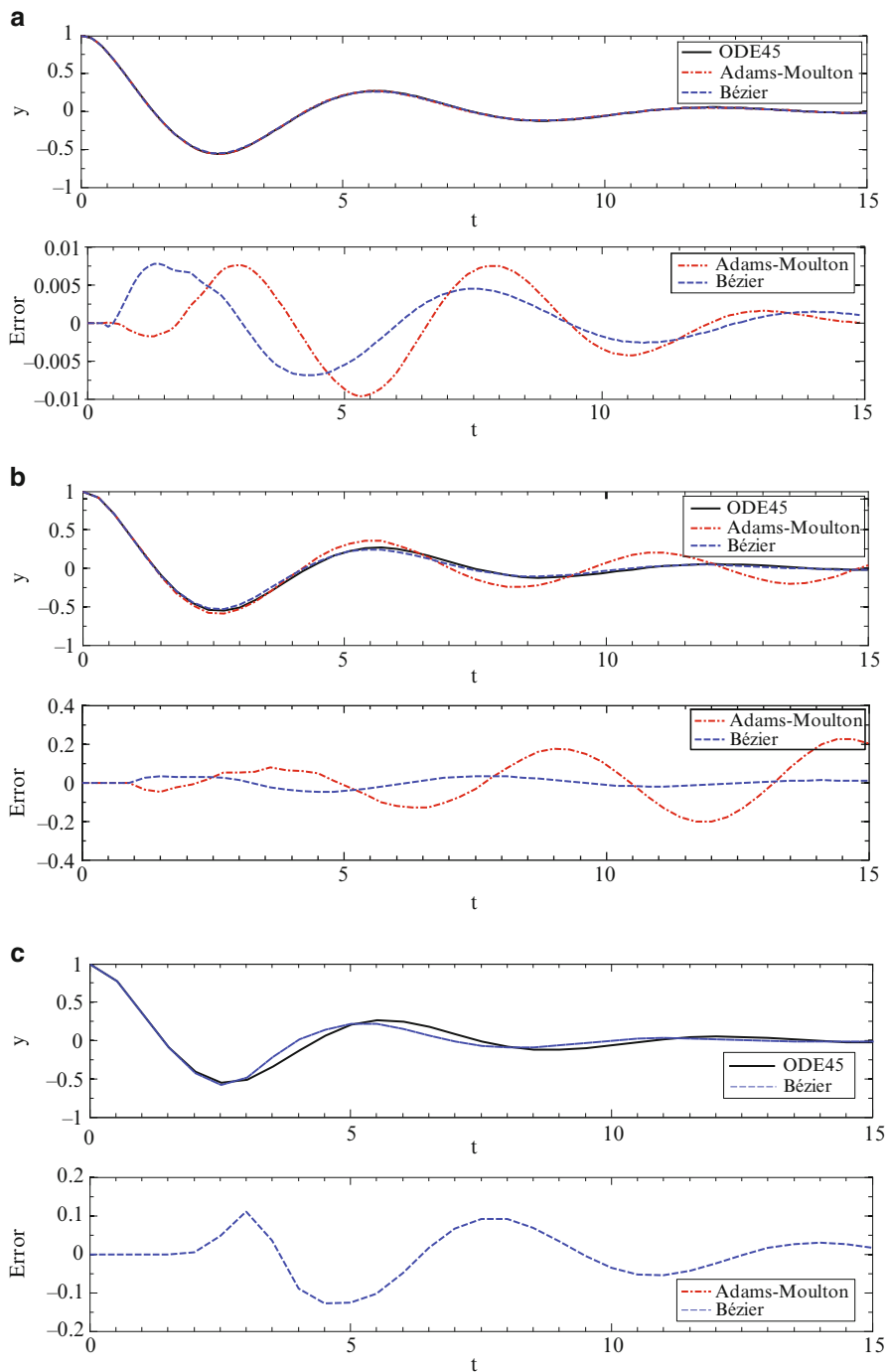
where  $x(t)$  and  $y(t)$  denote population of the prey and predator species, respectively and  $a, b, c$ , and  $d$  are positive constants.

Solution of predator–prey equation with  $a = 1$ ,  $b = 0.03$ ,  $c = 0.4$ ,  $d = 0.01$  using presented and Adams–Moulton methods is shown in Fig. 5.7. For  $h = 0.1$ , results for both methods are accurate and stable. As the step size growth to  $h = 0.3$ , Adams–Moulton shows instability while the presented method become unstable in larger step size of  $h = 0.5$ .

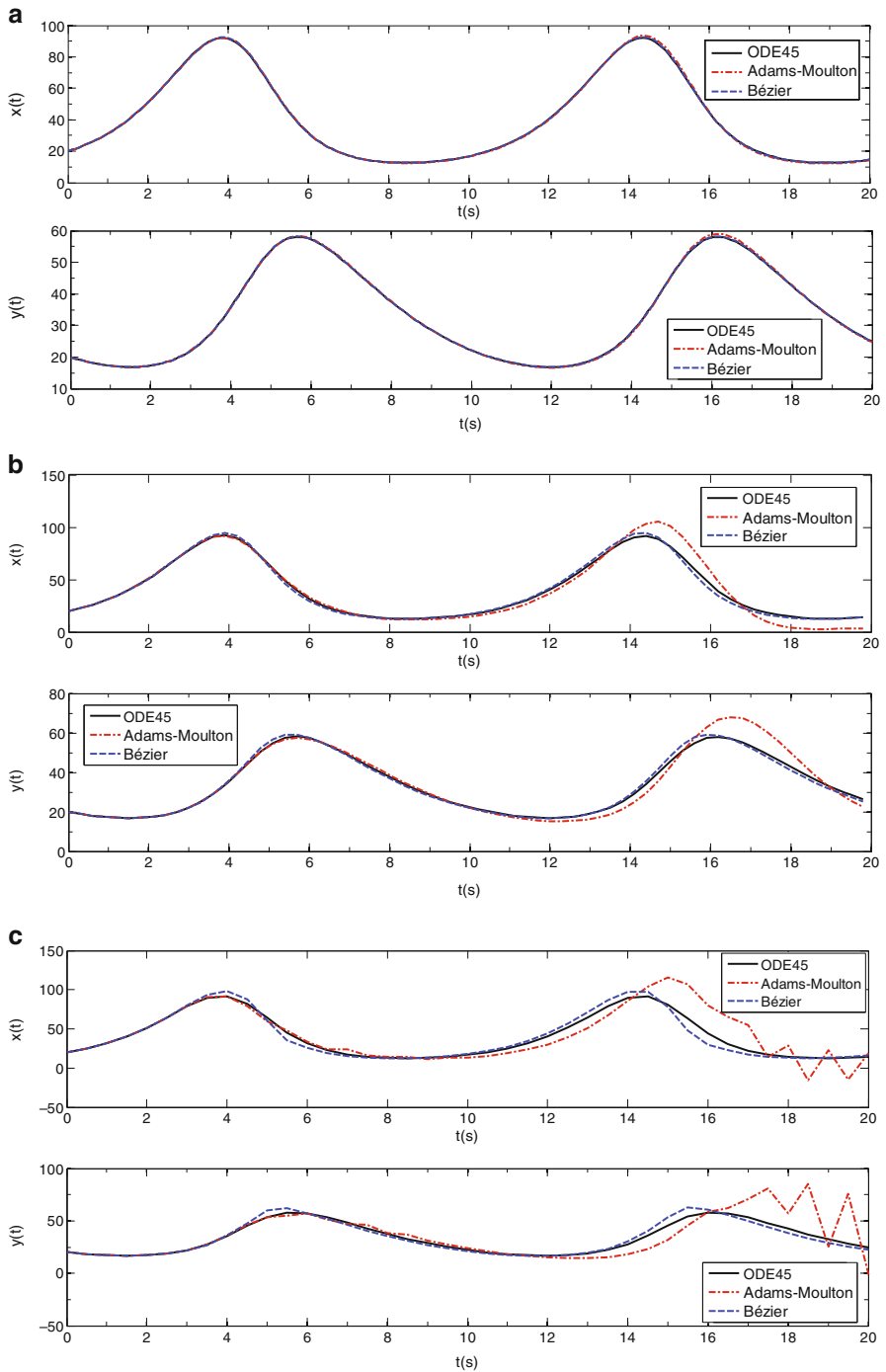
### 5.5.3 Example 3

Consider a Van der Pol oscillator, which is a non-conservative oscillator with nonlinear damping and governing equation of motion as (Nayfeh et al. 2007):

$$\frac{d^2y}{dt^2} - \mu(1 - y^2)\frac{dy}{dt} + y = 0 \quad (5.32)$$



**Fig. 5.6** Solution of nonlinear mass-spring system with damping for (a)  $h = 0.1$ , (b)  $h = 0.3$ , (c)  $h = 0.5$



**Fig. 5.7** Solution of Lotka–Volterra equation for (a)  $h = 0.1$ , (b)  $h = 0.3$ , (c)  $h = 0.5$

where  $\mu$  is a scalar parameter indicating the nonlinearity and the strength of the damping. In the case of  $\mu = 0$ , Eq. (5.32) reduces to  $\frac{d^2y}{dt^2} + y = 0$  which has the solution  $y = c_1 \sin(t) + c_2 \cos(t)$ . For small values of  $\mu$ , one obtain weekly nonlinear oscillation, i.e., oscillation which slightly differ from simple harmonic motion. On the other hand, however, for high values of  $\mu$  relaxation oscillation can be observed which shows strongly nonlinear oscillations exhibiting sharp periodic jumps. In the case of high damping, the Van der Pol equation becomes stiff. Most techniques show weak performance for solution of stiff ODEs. Here the performance of the presented and Adams–Moulton methods in two different conditions of low and high damping is investigated.

Introducing new variables  $y_1 = y$  and  $y_2 = \frac{dy}{dt}$ , the second order governing nonlinear ODE is converted to the following first order system of nonlinear ODE:

$$\begin{cases} \frac{dy_1}{dt} = y_2 \\ \frac{dy_2}{dt} = \mu(1 - y_1^2)y_2 + y_1 \end{cases} \quad (5.33)$$

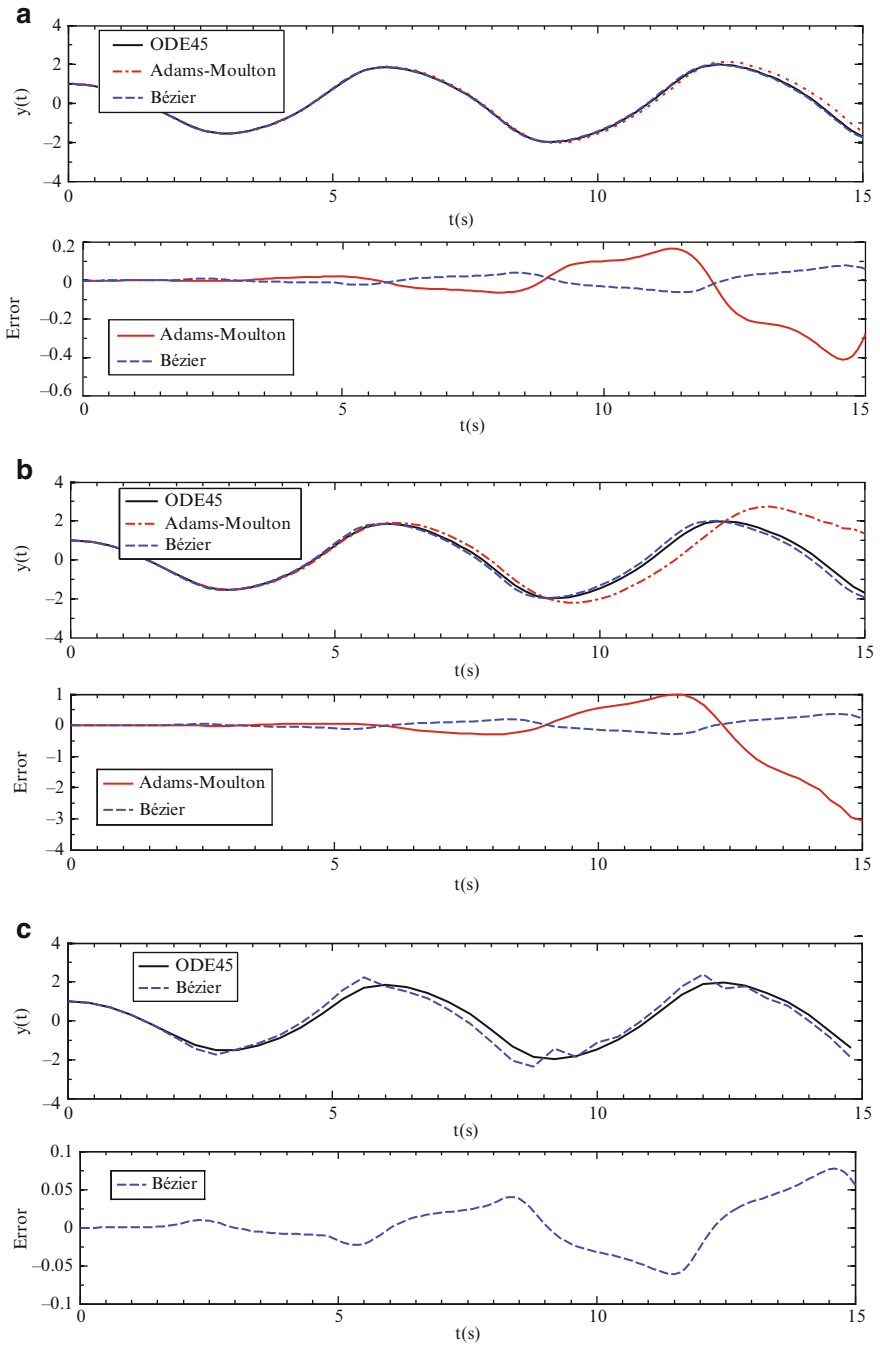
Solution of the Van der Pol equation in the case of  $\mu = 0.5$  and  $\mu = 5$  using presented multi-step method and Adams–Moulton is presented in Figs. 5.8 and 5.9, respectively. Included in the figures are also results of the ODE45 for comparison. In the case of lower damping, i.e.,  $\mu = 0.5$ , predictions of both presented and Adams–Moulton for  $h = 0.1$  are stable and accurate. As the step size increases to  $h = 0.2$ , Adams–Moulton results become unstable while the presented method shows unstable behavior at  $h = 0.4$ .

Furthermore, for higher damping of  $\mu = 5$ , smaller step size is used. For  $h = 0.001$ , both methods are accurate and stable. As the step size growth to  $h = 0.005$ , instability initiates in the results of Adams–Moulton while the presented method becomes unstable at step sizes ten times greater, i.e.,  $h = 0.05$ . Results show that even in the case of stiff ODEs, the presented multi-step method is more accurate and stable in comparison with the well-known Adams–Moulton method.

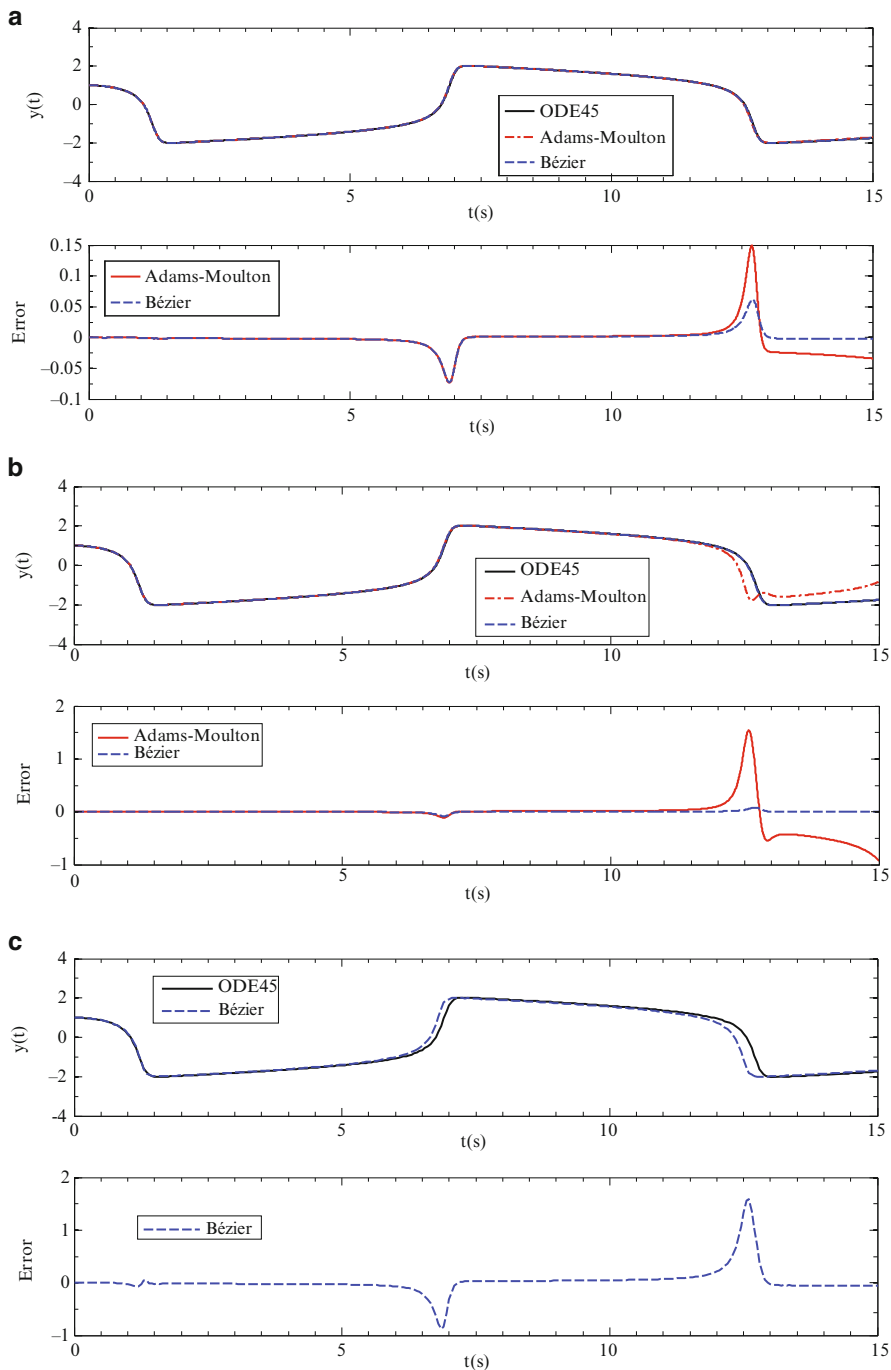
## 5.6 Conclusion

Multi-step methods such as Adams–Moulton are common techniques for solution of initial value ODEs. Besides their accuracy and stability, they also have their own limitations such instability in large step sizes or in the case of stiff differential equation. Thus, multi-step methods with better performance in terms of accuracy and convergence particularly in the case of stiff equations are appealing.

In the present study, a new general multi-step formula developed based on the Bézier curves. The consistency, stability, and convergence of the presented formula are proved by using the common stability and error analysis theorems of multi-step methods. The other advantage of the method is the presentation of a closed form general  $m + 1$ -step formula.



**Fig. 5.8** Solution of Van der Pol equation in case of low damping  $\mu = 0.5$  for (a)  $h = 0.1$ , (b)  $h = 0.2$ , (c)  $h = 0.4$



**Fig. 5.9** Solution of Van der Pol equation in case of high damping  $\mu = 5$  for (a)  $h = 0.001$ , (b)  $h = 0.005$ , (c)  $h = 0.05$

Effects of the step size on the accuracy and stability of the presented method are examined through numerical results of various nonlinear and stiff systems. Results revealed that the presented method shows improved stability behavior in comparison with the well-known Adams–Moulton method, which yields to significant reduction of computational cost. Due to simplicity, minimum computational efforts, and accuracy, the presented multi-step method is expected to be used in nonlinear problems in various science and engineering applications.

## Key Symbols

$B_{i,n}(u)$	The $i$ th Bernstein basis polynomials of degree $n$
$P_i$	Bézier curve control point
$p(z)$	First characteristic polynomials of multi-step formula
$q(z)$	Second characteristic polynomials of multi-step formula
$f_n$	Value of $f(x, y)$ at point $(x_n, y_n)$
$h$	Step size
$x_n$	The $n$ th discretized point $x$ coordinate
$y_n$	The $n$ th discretized point $y$ coordinate and value of $y(x)$ at point $(x_n)$
$x_0$	$x$ coordinate of the first point
$y_0$	Initial condition

## References

- Beach RC (1991) An introduction to curves and surfaces of computer-aided design. Van Nostrand Reinhold, New York
- Beryman AA (1992) The origins and evolution of predator–prey theory. *Ecology* 73(5):1530–1535
- Bhattia MI, Brackenb P (2007) Solutions of differential equations in a Bernstein polynomial basis. *Comput Appl Math* 205:272–280
- Buss SR (2003) 3D computer graphics-a mathematical introduction with open GL. Cambridge University Press, Cambridge, UK
- Caglar N, Caglar H (2006) B-spline solution of singular boundary value problems. *Appl Math Comput* 182:1509–1513
- Caglar H, Ozer M, Caglar N (2008) The numerical solution of the one-dimensional heat equation by using third degree B-spline functions. *Chaos Solit Fract* 38:1197–1201
- Dahlquist GG (1963) A special stability problem for linear multi-step methods. *BIT Numer Math* 3(1):27–43
- Farin G (2002) Curves and surfaces for CAGD. A practical guide, 5th edn. Elsevier Ltd., San Francisco, CA, USA
- Gautschi W (1997) Numerical analysis: an introduction. Birkhäuser, Boston
- Gerald CF, Wheatley PO (1999) Applied numerical analysis, 6th edn. Addison-Wesley, Reading
- Henrici P (1962) Discrete variable methods in ordinary differential equations. Wiley, New York
- Jator S, Sinkala Z (2007) A high order B-spline collocation method for linear boundary value problems. *Appl Math Comput* 191:100–116



- Khalifa AK, Raslan KR, Alzubaidi HM (2008) A collocation method with cubic B-splines for solving the MRLW equation. *Comput Appl Math* 212:406–418
- Lorentz GG (1986) *Bernstein polynomials*, 2nd edn. Chelsea Publishing Co., New York
- MATLAB Reference Guide Version 7.0 (2004) The Math Works Inc
- Nayfeh AH, Mook DT (2007) *Nonlinear oscillations*. Wiley, New York
- Süli E, Mayers D (2003) *An introduction to numerical analysis*. Cambridge University Press, Cambridge, UK
- Zerarka A, Nine B (2008) Solutions of the Von Kármán equations via the non-variational Galerkin-B-spline approach. *Commun Nonlinear Sci Numer Simulat* 13:2320–2327

# Chapter 6

## The Loss Tangent of Visco-Elastic Models

Franz Konstantin Fuss

**Keywords** Visco-elasticity • Elasticity • Viscosity • Loss tangent • Standard linear solid • Power model • Log model • Stress strain • Transient part • Steady state • Laplace transform • Inverse Laplace transform • Convolution integral

### 6.1 Introduction

Visco-elastic models are characterised by phenomena such as

- Velocity dependency: increase of stiffness with strain rate,
- Stress relaxation: decay of stress with time at constant strain,
- Creep: increase of strain with time at constant stress, and
- Loss of stored (elastic) energy due to inner friction resulting in unequal loading and unloading stress–strain curves, the area between the two curves corresponding to the dissipated energy (hysteresis).

Visco-elastic models can be classified in various ways, e.g.

#### (1) Linear or non-linear models

- Linear models with at least one elastic and one viscous element in parallel or in series;

---

F.K. Fuss (✉)

School of Aerospace, Mechanical, and Manufacturing Engineering, RMIT University,  
Melbourne, VIC, Australia

e-mail: [franz.fuss@rmit.edu.au](mailto:franz.fuss@rmit.edu.au)

- *QLV*/quasi-linear “viscous” models such as Prony series (Wiechert model; Wiechert 1889, 1893);
- Non-linear models such as logarithmic and power law models.

or

(2) According to the decay function of stress relaxation, which can be

- Exponential: linear three-element model (standard linear solid),
- Power: non-linear power law model, or
- Logarithmic: non-linear logarithmic law model

Examples for the latter two non-linear models are:

- Power law model: biological tissues such as ligaments (Provenzano et al. 2001), foams with non-negative stiffness (Fuss 2009), cork in cricket balls (Fuss 2008a, 2012);
- Logarithmic law model: solid polymers (Findley et al. 1989), polymer golf balls (Fuss 2008b, 2012).

The loss tangent,  $\tan \delta$ , is defined as the tangent of the phase angle  $\delta$ , which, in turn, is the ratio of loss modulus  $E''$  to storage modulus  $E'$ .

$$\tan \delta = \frac{E''}{E'} \quad (6.1)$$

where

$$E' = \frac{\sigma_0}{\varepsilon_0} \cos \delta \quad (6.2)$$

$$E'' = \frac{\sigma_0}{\varepsilon_0} \sin \delta \quad (6.3)$$

and  $\sigma_0$  and  $\varepsilon_0$  are the peak amplitudes of stress  $\sigma$  and strain  $\varepsilon$ , respectively.

The complex modulus  $E^*$  is defined as

$$E^* = \frac{\sigma_0}{\varepsilon_0} e^{i\delta} = \frac{\sigma_0}{\varepsilon_0} (\cos \delta + i \sin \delta) = E' + iE'' = |E^*| e^{i\delta} \quad (6.4)$$

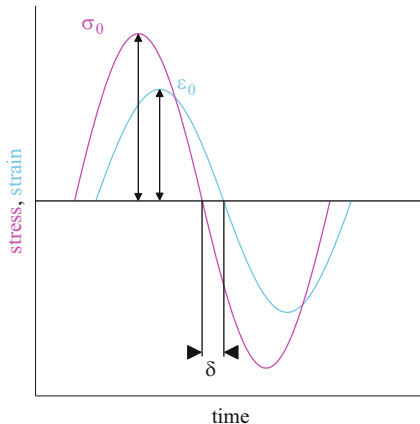
where  $i = \sqrt{-1}$ , and  $|E^*|$  is the dynamic modulus, the magnitude of  $E^*$ , i.e. the resultant of loss modulus  $E''$  and storage modulus  $E'$

$$|E^*| = \frac{\sigma_0}{\varepsilon_0} \quad (6.5)$$

The loss tangent,  $\tan \delta$ , is usually determined by subjecting a material or structure to sinusoidal strain  $\varepsilon$

$$\varepsilon = \varepsilon_0 \sin(2\pi f t) \quad (6.6)$$

**Fig. 6.1** Sinusoidal strain curve (blue) imposed on a visco-elastic material and resulting stress curve (pink);  $\sigma_0$ : stress amplitude, maximal stress;  $\epsilon_0$ : strain amplitude, maximal strain;  $\delta$ : phase shift



where  $f$  is the cyclic frequency (angular frequency  $\omega = 2\pi f$ ). The resulting reaction stress  $\sigma$  is equally sinusoidal, but out of phase with respect to the strain by the phase angle  $\delta$

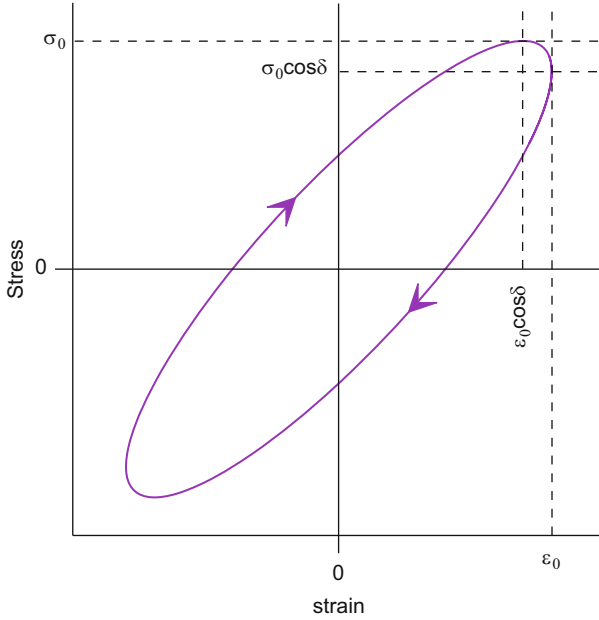
$$\sigma = \sigma_0 \sin(2\pi f t + \delta) \tag{6.7}$$

A positive phase angle  $\delta$  causes the stress peak to occur earlier than the strain peak (Fig. 6.1), resulting in the typical hysteresis effect of visco-elastic materials when plotting stress against strain (Fig. 6.2). The area of the hysteresis loop corresponds to the energy dissipated into the material as thermal energy.

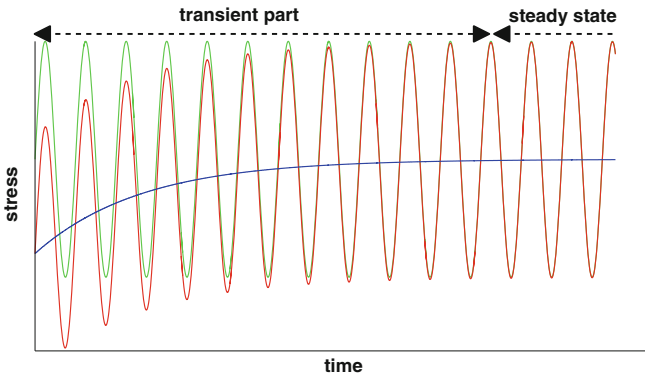
When subjecting a material to cyclic (sinusoidal) strain, the peak stress,  $\sigma_0$ , increases during the first cycles (*transient part*) until it reaches an equilibrium or *steady state* (Fig. 6.3). The phase angle  $\delta$  is measured once the steady state has set in.

It is evident, that the energy dissipated by inner friction depends on the viscosity parameter  $\eta$ . However, as the loss tangent is the ratio of loss to storage modulus, the strain rate independent elasticity parameter  $E$  is expected to influence the loss tangent too. Lastly, as the modulus (Young's and tangent) increases with strain rate and thus with frequency  $f$ , the latter could contribute to the loss tangent as well.

The objective of this *Book Chapter* is to explore in how far the viscosity parameter  $\eta$ , the strain rate independent elasticity parameter  $E$  and the strain frequency  $f$  affect the loss tangent. The aim is to derive the loss tangent at steady state of the three visco-elastic models mentioned above, in order to understand the interaction between elasticity, viscosity and frequency and their effect on energy loss. The function of stress relaxation with time is not the only difference between the three models mentioned above. A further objective of this paper is to analyse the basic differences of these models and to understand their applicability and constraints.



**Fig. 6.2** Hysteresis loop of stress–strain ellipse;  $\sigma_0$ : maximal stress;  $\epsilon_0$ : maximal strain;  $\delta$ : phase angle



**Fig. 6.3** Stress–time curve (*red*) during the first load cycles; *blue curve*: transient component; *green curve*: steady state component; the stress–time curve (*red*) is the sum of transient and steady state components

## 6.2 Analysis

### 6.2.1 The Standard Linear Solid (Zener Model)

The standard linear solid (SLS of Voigt form) consists of two Hookean springs and a viscous damper, where the spring with the modulus  $E_1$  is connected in series with a Kelvin–Voigt model, with spring of modulus  $E_2$  and damper of viscosity constant  $\eta$  connected in parallel (Fig. 6.2).

From Fig. 6.4

$$\varepsilon = \varepsilon_1 + \varepsilon_2 \quad (6.8)$$

$$\sigma = \varepsilon_1 E_1 \quad (6.9)$$

$$\sigma = \varepsilon_2 E_2 + \dot{\varepsilon}_2 \eta \quad (6.10)$$

Taking the Laplace transform of Eqs. (6.8)–(6.10), eliminating  $\varepsilon_1$  and  $\varepsilon_2$  by substitution, and solving for  $\hat{\sigma}$  yields the constitutive equation of the standard linear solid

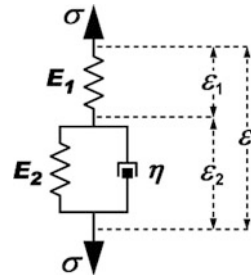
$$\hat{\sigma} = \hat{\varepsilon} \frac{E_1 E_2 + s \eta E_1}{E_1 + E_2 + s \eta} \quad (6.11)$$

where the caret ( $\hat{\phantom{x}}$ ) denotes the transformed parameter.

The equation for stress relaxation results from applying a constant strain  $\varepsilon_c$  to the model through a Heaviside function  $H(t)$

$$\varepsilon = \varepsilon_c H(t) \quad (6.12)$$

**Fig. 6.4** Standard linear solid of Voigt form;  $\sigma$ : stress;  $\varepsilon$ : strain;  $\eta$ : viscosity constant;  $E_1$ : modulus of series spring;  $E_2$ : modulus of parallel spring;  $\varepsilon_1$ : strain of series spring;  $\varepsilon_2$ : strain of Kelvin–Voigt model



the Laplace transform of which is

$$\widehat{\varepsilon} = \frac{\varepsilon_c}{s} \quad (6.13)$$

By substituting Eq. (6.13) into Eq. (6.11), we obtain

$$\widehat{\sigma} = \varepsilon_c \frac{E_1 E_2 + s\eta E_1}{s(E_1 + E_2 + s\eta)} \quad (6.14)$$

the inverse Laplace transform of which yields the function of stress relaxation

$$\frac{\sigma}{\varepsilon_c} = E_1 \frac{E_2 + E_1 e^{-t \frac{(E_1 + E_2)}{\eta}}}{E_1 + E_2} \quad (6.15)$$

where the stress  $\sigma$  is normalised to the constant strain  $\varepsilon_c$ . Equation (6.15) proves the exponential decay of stress with time, as mentioned above in the Introduction.

The loss tangent results from applying the sinusoidal strain of Eq. (6.6) to the constitutive Eq. (6.11). The Laplace transform of Eq. (6.6) is

$$\widehat{\varepsilon} = \varepsilon_0 \frac{2\pi f}{s^2 + (2\pi f)^2} \quad (6.16)$$

where  $\varepsilon_0$  is the peak amplitude of the strain.

By substituting Eq. (6.16) into the constitutive Eq. (6.11), we obtain

$$\widehat{\sigma} = \varepsilon_0 \left( \frac{2\pi f}{s^2 + (2\pi f)^2} \right) \left( \frac{E_1 E_2 + s\eta E_1}{E_1 + E_2 + s\eta} \right) \quad (6.17)$$

After rearranging

$$\widehat{\sigma} = \varepsilon_0 2\pi f E_1 \frac{E_2 + s\eta}{s^3 \eta + s^2 E_1 + s^2 E_2 + s\eta(2\pi f)^2 + E_1(2\pi f)^2 + E_2(2\pi f)^2} \quad (6.18)$$

the inverse Laplace transform of Eq. (6.18) yields

$$\sigma = \varepsilon_0 E_1 \frac{-E_1 \eta 2\pi f e^{-t \frac{E_1 + E_2}{\eta}} + E_1 \eta 2\pi f \cos(2\pi f t) + (E_1 E_2 + E_2^2 + \eta^2 (2\pi f)^2) \sin(2\pi f t)}{E_1^2 + 2E_1 E_2 + E_2^2 + \eta^2 (2\pi f)^2} \quad (6.19)$$

The numerator of Eq. (6.19) comprises of a *transient* part (exponential function) and the *steady state* part (sine and cosine functions). At large times, if  $t \rightarrow \infty$ , the transient part, i.e. the exponential term of the numerator, vanishes and the steady state sets in:

$$\sigma = \frac{\varepsilon_0 E_1}{(E_1 + E_2)^2 + \eta^2 (2\pi f)^2} \left[ E_1 \eta 2\pi f \cos(2\pi f t) + (E_1 E_2 + E_2^2 + \eta^2 (2\pi f)^2) \sin(2\pi f t) \right] \quad (6.20)$$

In order to obtain the loss tangent and the peak stress  $\sigma_0$ , we apply the addition rules to Eq. (6.7)

$$\sigma = \sigma_0 \sin(\delta) \cos(2\pi f t) + \sigma_0 \cos(\delta) \sin(2\pi f t) \quad (6.21)$$

where the unknowns are the two components of the peak stress  $\sigma_0$ :  $\sigma_0 \sin \delta$  and  $\sigma_0 \cos \delta$ .

The loss tangent thus results from the ratio of the two components

$$\tan \delta = \frac{\sigma_0 \sin \delta}{\sigma_0 \cos \delta} \quad (6.22)$$

and the resultant of the peak stress  $\sigma_0$  is obtained from

$$\sigma_0 = \sqrt{(\sigma_0 \sin \delta)^2 + (\sigma_0 \cos \delta)^2} \quad (6.23)$$

From Eq. (6.20) it follows that

$$\sigma_0 \sin \delta = \frac{\varepsilon_0 E_1}{(E_1 + E_2)^2 + \eta^2 (2\pi f)^2} E_1 \eta 2\pi f \quad (6.24)$$

and

$$\sigma_0 \cos \delta = \frac{\varepsilon_0 E_1}{(E_1 + E_2)^2 + \eta^2 (2\pi f)^2} (E_1 E_2 + E_2^2 + \eta^2 (2\pi f)^2) \quad (6.25)$$

The loss tangent thus is

$$\tan \delta = \frac{E_1 \eta 2\pi f}{E_1 E_2 + E_2^2 + \eta^2 (2\pi f)^2} \quad (6.26)$$

and the peak stress  $\sigma_0$  results from

$$\sigma_0 = \frac{\varepsilon_0 E_1}{(E_1 + E_2)^2 + \eta^2 (2\pi f)^2} \sqrt{E_1^2 \eta^2 (2\pi f)^2 + (E_1 E_2 + E_2^2 + \eta^2 (2\pi f)^2)^2} \quad (6.27)$$

In both Eqs. (6.26) and (6.27)  $f$  and  $\eta$  are linked together and always occur as the product  $f\eta$ .

As the maximal strain rate  $\dot{\varepsilon}_0$  of a sinusoidal strain function equals  $\varepsilon_0 2\pi f$ , the strain rate dependency of  $\sigma_0$  is given by



$$\sigma_0 = \frac{\varepsilon_0 E_1}{(E_1 + E_2)^2 + \eta^2 \left(\frac{\dot{\varepsilon}_0}{\varepsilon_0}\right)^2} \sqrt{E_1^2 \eta^2 \left(\frac{\dot{\varepsilon}_0}{\varepsilon_0}\right)^2 + \left[E_1 E_2 + E_2^2 + \eta^2 \left(\frac{\dot{\varepsilon}_0}{\varepsilon_0}\right)^2\right]^2} \quad (6.28)$$

The peak stress  $\sigma_0$  increases with  $f$  and  $\eta$ . If  $f$  or  $\eta \rightarrow \infty$ ,  $\sigma_0$  asymptotes to  $\varepsilon_0 E_1$ . If  $f$  or  $\eta \rightarrow 0$ ,  $\sigma_0 \rightarrow \varepsilon_0 (E_1^{-1} + E_2^{-1})^{-1}$ . The limits of  $\sigma_0$  are evident when considering that the peak strain rate changes with  $f$ . Thus, at zero strain rate or zero  $\eta$ , the standard linear solid reduces to two springs in series and the effective modulus of the model becomes  $(E_1^{-1} + E_2^{-1})^{-1}$ . At infinite strain rate or infinite  $\eta$ , the damper becomes rigid and the effective modulus of the model is just  $E_1$ .

The peak stress  $\sigma_0$  increases with  $E_1$ . If  $f$  or  $\eta \rightarrow \infty$  and  $E_1 \rightarrow 0$  or  $\infty$ , both the effective modulus and  $\sigma_0$  become equally 0 or  $\infty$ , respectively. If  $f$  or  $\eta \rightarrow 0$  and  $E_1 \rightarrow 0$  or  $\infty$ , the effective modulus and  $\sigma_0$  become zero in the former case, and  $E_2$  and  $\varepsilon_0 E_2$  in the latter.

If  $E_2 \rightarrow 0$ , the standard linear solid reduces to a Maxwell model (with a spring and a damper in series), the modulus of which is (1)  $\infty$  or (2)  $E_1$  or (3) 0, if (1)  $E_1$ , and  $\eta$  or  $f$ , are  $\infty$  or (2)  $\eta$  or  $f$  are  $\infty$  or (3)  $E_1$ , and  $\eta$  or  $f$  are 0, respectively.

At  $\eta > 0$  and  $f < \infty$ , and  $E_2 \rightarrow 0$ , the peak stress reaches

$$\sigma_0 = \frac{2\pi f \varepsilon_0 E_1 \eta}{\sqrt{E_1^2 + \eta^2 (2\pi f)^2}} \quad (6.29)$$

If  $E_2 \rightarrow \infty$ , the modulus of the standard linear solid reduces to  $E_1$ .  $\sigma_0$  shows a local minimum at a certain  $E_{2_0}$  (Fig. 6.5)

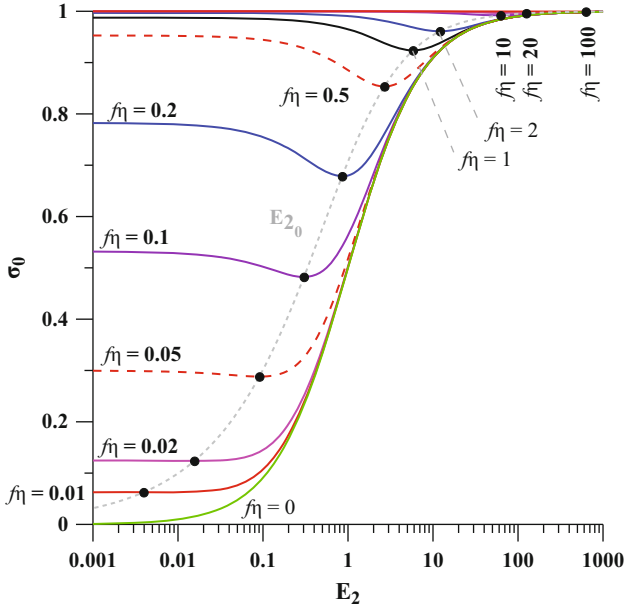
$$E_{2_0} = \frac{\sqrt{E_1^2 + 4\eta^2 (2\pi f)^2} - E_1}{2} \quad (6.30)$$

resulting from equating the first  $E_2$  derivative of Eq. (6.27) with zero and solving for  $E_2$ . Equation (6.30) is the only positive and real result of the fourth order polynomial nature of the first  $E_2$  derivative of Eq. (6.27).

The loss tangent, Eq. (6.26), reveals that the phase angle  $\delta$  depends on all four parameters,  $E_1$ ,  $E_2$ ,  $\eta$ , and  $f$ , where the latter two always occur as the product  $f\eta$ .

If the frequency  $f \rightarrow \infty$  or 0, the phase angle  $\delta \rightarrow 0$  in both cases. Thus, we expect a local maximum of  $\delta$  at a certain frequency  $f_0$  (actually shown by Findley et al. 1989). Taking the first derivative of  $\tan \delta$  with respect to  $f$  in Eq. (6.26), equating it with zero and solving for  $f$  provides

$$f_0 = \frac{\sqrt{E_1 E_2 + E_2^2}}{2\pi \eta} \quad (6.31)$$



**Fig. 6.5** Maximal stress  $\sigma_0$  against modulus  $E_2$  of parallel spring;  $f$ : frequency;  $\eta$ : viscosity constant;  $E_{2_0}$ :  $E_2$  at which a local minimum of  $\sigma_0$  is found

Replacing  $f$  by  $f_0$  in Eq. (6.26),

$$(\tan \delta)_{f_0} = \frac{E_1}{2\sqrt{E_1 E_2 + E_2^2}} \tag{6.32}$$

and equating its denominator with zero yields the scenario at which  $\delta = 0.5 \pi$ , which is  $E_1 = -E_2$  and  $E_2 = 0$ . In both cases, however,  $f_0 = 0$ , which means that a standard linear solid can never reach  $\delta = 0.5\pi$  or an infinite loss tangent.

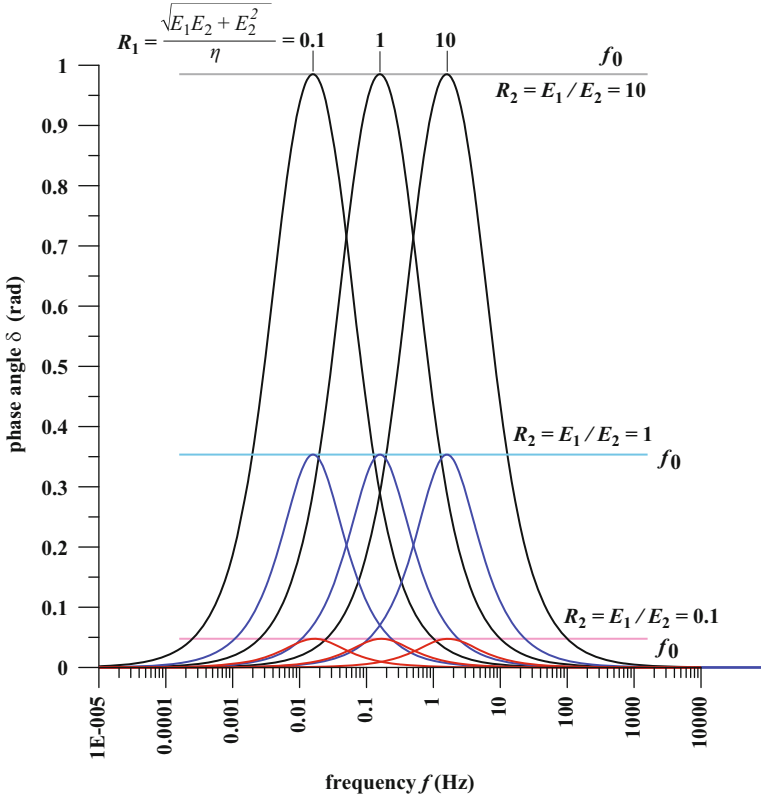
Rearranging Eqs. (6.31) and (6.32) shows that  $f_0$  depends on two ratios,  $R_1$  and  $R_2$ , whereas  $\tan \delta$  at  $f_0$  (but also any other  $f$  at the same  $R_1$ ) depends only on  $R_2$  (Fig. 6.6)

$$f_0 = \frac{R_1}{2\pi} \tag{6.33}$$

$$(\tan \delta)_{f_0} = \frac{R_2}{2\sqrt{1 + R_2}} \tag{6.34}$$

where

$$R_1 = \frac{E_2 \sqrt{1 + R_2}}{\eta} \tag{6.35}$$



**Fig. 6.6** Phase angle  $\delta$  against frequency  $f$ ;  $\eta$ : viscosity constant;  $E_1$  and  $E_2$ : moduli of series and parallel spring;  $f_0$ :  $f$  at which a local maximum of  $\delta$  is found

and

$$R_2 = \frac{E_1}{E_2} \tag{6.36}$$

Equation (6.34) shows that  $\tan \delta$  at  $f_0$  is a constant and independent of  $f_0$  (Fig. 6.6).

Changing  $\eta$  has the same effect as changing  $f$  has. If  $\eta \rightarrow \infty$  or  $0$ , the phase angle  $\delta \rightarrow 0$  in both cases and we obtain a local maximum of  $\delta$  at a certain  $\eta_0$

$$\eta_0 = \frac{\sqrt{E_1 E_2 + E_2^2}}{2\pi f} \tag{6.37}$$

Equation (6.37) is similar to Eq. (6.31) and the same principles apply.

The phase angle  $\delta$  and  $\tan \delta$  increase with  $E_1$ . If  $E_1 \rightarrow 0$ ,  $\tan \delta \rightarrow 0$ . If  $E_1 \rightarrow \infty$ ,  $\tan \delta$  asymptotes to

$$\lim_{E_1 \rightarrow \infty} (\tan \delta) = \frac{\eta 2\pi f}{E_2} \quad (6.38)$$

The phase angle  $\delta$  and  $\tan \delta$  decrease with increasing  $E_2$ , asymptoting to 0 if  $E_2 \rightarrow \infty$ . If  $E_2 \rightarrow 0$ ,  $\tan \delta$  reaches a limit of

$$\lim_{E_2 \rightarrow 0} (\tan \delta) = \frac{E_1}{\eta 2\pi f} \quad (6.39)$$

Lakes (2009) derived the loss tangent of the SLS of Maxwell form (with a spring connected in parallel with a Maxwell model).

## 6.2.2 The Power Law Model

The power law model is characterised by a power decay of stress  $\sigma$  with time  $t$ :

$$\frac{\sigma}{\varepsilon_c} = E t^{-\eta} \quad (6.40)$$

Stress  $\sigma$  is normalised to the constant strain  $\varepsilon_c$  applied by the Heaviside function  $H(t)$  of Eq. (6.12).

Taking Laplace transform of Eq. (6.40) yields

$$\hat{\sigma} = \varepsilon_c E \frac{\Gamma(-\eta + 1)}{s^{-\eta+1}} \quad (6.41)$$

where  $\Gamma$  denotes the Gamma function (Fuss 2008a, 2012).

By substituting Eq. (6.13) into Eq. (6.41), we obtain the constitutive equation of the power law of non-linear visco-elasticity (Fuss 2008a, 2012):

$$\hat{\sigma} = s^\eta \hat{\varepsilon} E \Gamma(-\eta + 1) \quad (6.42)$$

By substituting the sinusoidal strain of Eq. (6.16) into the constitutive Eq. (6.42), we obtain

$$\hat{\sigma} = \varepsilon_0 E \Gamma(-\eta + 1) s^\eta \frac{2\pi f}{s^2 + (2\pi f)^2} \quad (6.43)$$

Taking the inverse Laplace transform of Eq. (6.43), we obtain a general fractional derivative of the sine function

$$\sigma = \varepsilon_0 E \Gamma(-\eta + 1) \frac{d^\eta}{dt^\eta} \sin(2\pi f t) \quad (6.44)$$

where  $\sigma$  is the  $\eta$ th time derivative of the strain function of Eq. (6.6) times a constant.

The generalised solution of  $\frac{d^\eta}{dt^\eta} \sin(t)$  results from applying the inverse operation of the Riemann–Liouville fractional integration (with lower limit = 0),

$$\frac{d^\eta}{dt^\eta} \sin(t) = \sin\left(t + \eta \frac{\pi}{2}\right) + \left[ \frac{t^{-1-\eta}}{\Gamma(-\eta)} - \frac{t^{-3-\eta}}{\Gamma(-\eta-2)} + \dots \right], \quad (6.45)$$

i.e. Eq. 6.10.3 of Oldham and Spanier (1974)

consisting of *steady state* (sine function) and *transient* parts (Maclaurin series in square brackets). If  $\eta \rightarrow 0$  or 1, the denominators of the transient part approach  $\pm \infty$  (Gamma function of negative integers), the transient part reduces to 0, and a sine or cosine function, respectively, remains.

By replacing the lower limit of the inverse operation of the Riemann–Liouville fractional integration by  $-\infty$ , which equals the inverse operation of the Weyl integral (Weyl 1917), we obtain the steady state part of the fractional derivative of Eq. (6.44)

$$\sigma = \varepsilon_0 E \Gamma(-\eta + 1) (2\pi f)^\eta \sin\left(2\pi f t + \eta \frac{\pi}{2}\right) \quad (6.46)$$

Comparing Eq. (6.7) with the steady state Eq. (6.46), it becomes evident that the phase angle  $\delta$  is

$$\delta = \eta \frac{\pi}{2} \quad (6.47)$$

the loss tangent is

$$\tan \delta = \tan\left(\eta \frac{\pi}{2}\right) \quad (6.48)$$

and the peak stress  $\sigma_0$  is

$$\sigma_0 = \varepsilon_0 E \Gamma(-\eta + 1) (2\pi f)^\eta \quad (6.49)$$

The loss tangent depends solely on  $\eta$ , independent of  $E$  and  $f$ . If  $\eta \rightarrow 1$ , both  $\tan \delta$  and  $\sigma_0 \rightarrow \infty$ , as both  $\tan(0.5\pi)$  and  $\Gamma(0)$  are infinite. Thus,  $0 \leq \eta < 1$ .

The peak stress  $\sigma_0$  increases non-linearly with  $f$  and linearly with  $E$  and  $f^\eta$ . The peak stress  $\sigma_0$  decreases non-linearly with  $\eta$ , reaching a limit of  $\varepsilon_0 E$ , i.e. a Hookean spring, if  $\eta \rightarrow 0$ . If  $\eta \rightarrow 1$ ,  $\sigma_0$  reaches infinity.

As the maximal strain rate  $\dot{\varepsilon}_0$  of a sinusoidal strain function equals  $\varepsilon_0 2\pi f$ , the strain rate dependency of  $\sigma_0$  of Eq. (6.49) is given by

$$\sigma_0 = \varepsilon_0 \frac{E \Gamma(-\eta + 1)}{(2\pi f)^{1-\eta}} (2\pi f)^\eta (2\pi f)^{1-\eta} = \dot{\varepsilon}_0 \frac{E \Gamma(-\eta + 1) \varepsilon_0^{1-\eta}}{\dot{\varepsilon}_0^{1-\eta}} = \dot{\varepsilon}_0^\eta E \Gamma(-\eta + 1) \varepsilon_0^{1-\eta} \quad (6.50)$$

In contrast to the fractional derivative approach shown above, Lakes (2009) solved the loss tangent of the power model from the ratio of the constitutive equations of loss to storage modulus.

### 6.2.3 The Logarithmic Law Model

The logarithmic law model is characterised by a logarithmic decay of stress  $\sigma$  with time  $t$ :

$$\frac{\sigma}{\varepsilon_c} = E - \eta \log(t) \quad (6.51)$$

where “log” denotes the natural logarithm. Stress  $\sigma$  is normalised to the constant strain  $\varepsilon_c$  applied by the Heaviside function  $H(t)$  of Eq. (6.12).

Taking Laplace transform of Eq. (6.51) yields

$$\hat{\sigma} = \varepsilon_c \frac{E}{s} - \varepsilon_c \eta \left( -\frac{\gamma}{s} - \frac{\log s}{s} \right) \quad (6.52)$$

where  $\gamma$  denotes the Euler–Mascheroni constant, i.e. 0.577215665... (Fuss 2008b, 2012).

By substituting Eq. (6.13) into Eq. (6.52), we obtain the constitutive equation of the logarithmic law of non-linear visco-elasticity (Fuss 2008b, 2012):

$$\hat{\sigma} = \hat{\varepsilon} E + \hat{\varepsilon} \eta (\gamma + \log s) \quad (6.53)$$

By substituting the sinusoidal strain of Eq. (6.16) into the constitutive Eq. (6.53), we obtain

$$\hat{\sigma} = \frac{2\varepsilon_0 \pi f}{s^2 + (2\pi f)^2} (E + \gamma \eta + \eta \log s) \quad (6.54)$$

after rearranging

$$\hat{\sigma} = \varepsilon_0 E \frac{2\pi f}{s^2 + (2\pi f)^2} - \varepsilon_0 \eta \left( -\frac{\gamma}{s} - \frac{\log s}{s} \right) \left( \frac{s 2\pi f}{s^2 + (2\pi f)^2} \right) \quad (6.55)$$

and taking inverse Laplace transform, we obtain

$$\sigma = \varepsilon_0 E \sin(2\pi f t) - \varepsilon_0 \eta 2\pi f [\log t]^* [\cos(2\pi f t)] \quad (6.56)$$

where \* denotes a convolution.

Applying the convolution integral, the convolution  $\log(t) * \cos(\omega t)$ , where  $\omega = 2\pi f$ , is solved accordingly:

$$[\log t] * [\cos(\omega t)] = \int_0^t \log(\tau) \cos[\omega(t - \tau)] d\tau \quad (6.57)$$

where  $\tau$  is the dummy variable of the convolution integral.

Decomposition of  $\cos(\omega t - \omega\tau)$  according to the addition rules and subsequent partial integration yields:

$$[\log t] * [\cos(\omega t)] = \omega^{-1} [\text{Ci}(\omega\tau) \sin(\omega t) - \text{Si}(\omega\tau) \cos(\omega t) - \log(\tau) \sin(\omega t - \omega\tau)]_0^t \quad (6.58)$$

where Ci and Si denote cosine and sine integrals respectively (definition of Ci and Si according to Abramowitz and Stegun 1972).

Solving Eq. (6.58) from 0 to  $t$ :

$$[\log t] * [\cos(\omega t)] = \omega^{-1} \left[ \text{Ci}(\omega t) \sin(\omega t) - \text{Si}(\omega t) \cos(\omega t) - \log(t) \sin(0) - \text{Ci}(\omega 0) \sin(\omega t) + \text{Si}(\omega 0) \cos(\omega t) + \log(0) \sin(\omega t) \right] \quad (6.59)$$

Apparently, Eq. (6.59) contains an indeterminate form, as both  $\log(0)$  and  $\text{Ci}(0)$  are  $-\infty$ , and thus the term  $\log(0) \sin(\omega t) - \text{Ci}(\omega 0) \sin(\omega t)$ , or  $\sin(\omega t) [\log(0) - \text{Ci}(0)]$ , delivers  $\sin(\omega t) (\infty - \infty)$ .

However,

$$\text{Cin}(t) = \gamma + \log(t) - \text{Ci}(t) \quad (6.60)$$

where Cin denotes an alternative cosine integral (definition of Cin according to Schelkunoff 1944). As  $\text{Cin}(0) = 0$ ,  $\log(0) - \text{Ci}(0) = -\gamma$ .

Yet, the argument of the cosine integral in Eq. (6.58) is  $\omega\tau$ , in contrast to the one of the natural logarithm, which is just  $\tau$ . Thus we have to consider the multiplier  $\omega$ . This multiplier leads to a convergence constant other than simply  $-\gamma$ , if  $t \rightarrow 0$ .

$$\text{Ci}(\omega t) = \gamma + \log(\omega t) - \text{Cin}(\omega t) = \gamma + \log(t) + \log(\omega) - \text{Cin}(\omega t) \quad (6.61)$$

As  $\text{Cin}(0) = 0$ ,

$$\log(0) - \text{Ci}(0) = -\gamma - \log(\omega) \quad (6.62)$$

Thus,

$$\lim_{t \rightarrow 0} [\log(t) - \text{Ci}(\omega t)] = -\log(\omega) - \gamma \quad (6.63)$$

Equation (6.59) can now be solved, considering that  $\text{Si}(0) = 0$

$$[\log t]^* [\cos(\omega t)] = \omega^{-1} \{ \text{Ci}(\omega t) \sin(\omega t) - \text{Si}(\omega t) \cos(\omega t) - \sin(\omega t) [\log(\omega) + \gamma] \} \quad (6.64)$$

The solution of Eq. (6.56) after taking inverse Laplace follows

$$\sigma = \varepsilon_0 E \sin(\omega t) - \varepsilon_0 \eta \{ \text{Ci}(\omega t) \sin(\omega t) - \text{Si}(\omega t) \cos(\omega t) - \sin(\omega t) [(\log(\omega) + \gamma)] \} \quad (6.65)$$

The steady state of Eq. (6.65) sets in at large times, or  $t \rightarrow \infty$ . The *transient* part of Eqs. (6.64) and (6.65) comprises of the cosine and sine integrals. When considering the values at infinity of cosine and sine integrals, which are 0 and  $0.5 \pi$  respectively, the convolution of Eqs. (6.64) and (6.65) yields at large times (*steady state equation*)

$$\lim_{t \rightarrow \infty} [\log(t) * \cos(\omega t)] = \omega^{-1} \left\{ -\frac{\pi}{2} \cos(\omega t) - \sin(\omega t) [\log(\omega) + \gamma] \right\} \quad (6.66)$$

and

$$\sigma = \varepsilon_0 E \sin(\omega t) + \varepsilon_0 \eta \frac{\pi}{2} \cos(\omega t) + \varepsilon_0 \eta \sin(\omega t) [\log(\omega) + \gamma] \quad (6.67)$$

respectively.

After rearranging according to the procedure applied for determining the loss tangent of the standard linear solid, we obtain

$$\tan \delta = \left( \frac{\pi}{2} \right) \frac{\eta}{E + \eta [\log(2\pi f) + \gamma]} \quad (6.68)$$

and

$$\sigma_0 = \varepsilon_0 \sqrt{0.25\eta^2\pi^2 + [E + \eta \log(2\pi f) + \eta\gamma]^2} \quad (6.69)$$

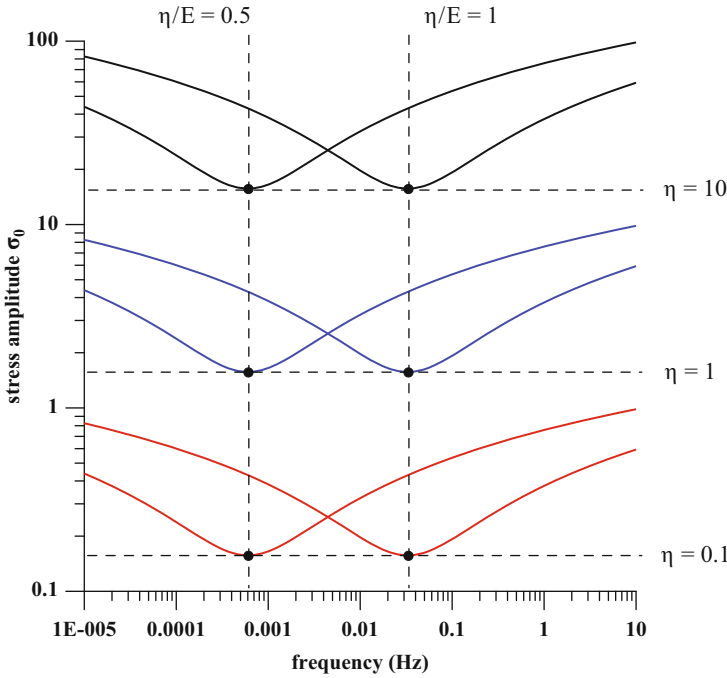
As the maximal strain rate  $\dot{\varepsilon}_0$  of a sinusoidal strain function equals  $\varepsilon_0 2\pi f$ , the strain rate dependency of  $\sigma_0$  after rearranging Eq. (6.69) is given by

$$\sigma_0 = \varepsilon_0 \sqrt{0.25\eta^2\pi^2 + \left[ E + \eta \log\left(\frac{\dot{\varepsilon}_0}{\varepsilon_0}\right) + \eta\gamma \right]^2} \quad (6.70)$$

The stress amplitude  $\sigma_0$  increases with  $E$ ,  $\eta$ , and  $f$ . This fact is obvious, as  $E$  is the strain rate independent elasticity parameter, i.e. the modulus or stiffness; the viscosity parameter  $\eta$  is linked to the strain rate, i.e. at a given strain rate,  $\sigma$  increases with  $\eta$ ; and the frequency  $f$  is linearly proportional to the strain rate applied periodically to the model, i.e. at a given  $\eta$ ,  $\sigma$  increases with  $f$ .

If  $\eta \rightarrow 0$ ,  $\sigma_0 \rightarrow \varepsilon_0 E$ , the stress equation of a Hookean spring. If  $\eta \rightarrow \infty$ ,  $\sigma_0 \rightarrow \infty$ .





**Fig. 6.7** Peak stress  $\sigma_0$  against frequency  $f$ ;  $\eta$ : viscosity constant;  $E$ : velocity independent elasticity parameter

If  $f \rightarrow 0$  or  $\infty$ ,  $\sigma_0 \rightarrow \infty$ . This means that  $\sigma_0$  has a local minimum at a certain frequency  $f_0$  (Fig. 6.7). As the first frequency derivative of Eq. (6.69) contains the arguments  $f$ ,  $\log(2\pi f)$ , and  $\log^2(2\pi f)$ , there is no closed-form analytical solution, and the frequency, at which  $\sigma_0$  is at a minimum, can only be obtained numerically. Figure 6.7 shows that  $\sigma_0$  increases with  $\eta$ , and  $f_0$  increases with the ratio  $\eta/E$ . The ratio  $\eta/E$  is identical to the viscosity of a log law model, not to be confused with the viscosity constant  $\eta$ .

The peak stress  $\sigma_0$  increases almost linearly with  $E$ .

If  $E \rightarrow \infty$ ,  $\sigma_0 \rightarrow \infty$ . If  $E \rightarrow 0$  (purely viscous material)

$$\lim_{E \rightarrow 0} \sigma_0 = \varepsilon_0 \eta \sqrt{0.25\pi^2 + [\log(2\pi f) + \gamma]^2} \tag{6.71}$$

and  $\sigma_0$  increases with  $\eta$  and  $f$ .

The loss tangent, Eq. (6.68), reveals that the phase angle  $\delta$  depends on  $E$ ,  $\eta$ , and  $f$ .

Increasing the elasticity parameter  $E$  reduces the phase angle  $\delta$ , thereby asymptoting to 0, when  $E$  approaches infinity. This result is evident, as a perfectly rigid solid ( $E = \infty$ ) does not deform and thus there are no losses ( $\delta = 0$ ).

Reducing  $E$  increases  $\tan \delta$ . If  $E$  approaches zero (purely viscous material),  $\tan \delta$  reaches a limit of

$$\lim_{E \rightarrow 0} (\tan \delta) = \left(\frac{\pi}{2}\right) \frac{1}{\log(2\pi f) + \gamma} \quad (6.72)$$

a constant, which is a function of  $f$  but independent of  $\eta$ . The magnitude of the viscosity  $\eta$  does not matter in this case, as the material is anyway purely viscous. The phase angle  $\delta$  becomes  $0.5 \pi$  if

$$f = \frac{1}{2\pi e^\gamma} \quad (6.73)$$

Thus, if  $E = 0$  and the loss tangent is positive,  $f$  must be  $\geq 0.08936 \dots$  Hz.

If  $f < 0.08936 \dots$  Hz,  $\delta > 0.5 \pi$ , and  $E'$  (storage modulus) and  $\tan \delta$  are negative, which is impossible, as the stored energy is zero (as  $E = 0$ ), and the energy dissipated cannot be negative.

If the frequency  $f \rightarrow \infty$  or  $0$ , the loss tangent  $\rightarrow 0$  in both cases, and the phase angle  $\delta \rightarrow 0$  or  $\pi$ , respectively. Thus,  $\delta$  crosses  $\pi/2$  at a certain frequency  $f_0$ . Equating the denominator of Eq. (3.20) with zero yields  $f_0$  at which  $\delta = 0.5 \pi$ :

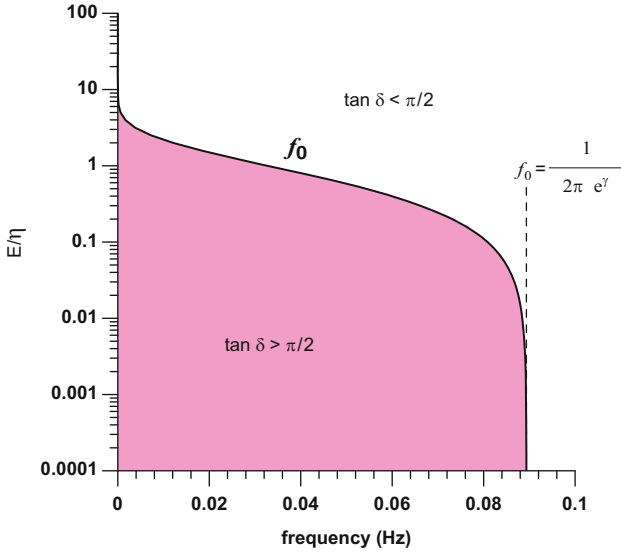
$$f_0 = \frac{e^{-\frac{E}{\eta} - \gamma}}{2\pi} \quad (6.74)$$

The ratio  $E/\eta$  in Eq. (6.68) equals the reciprocal value of the viscosity of a log law model. The higher the viscosity, the smaller is the ratio  $E/\eta$ . Figure 6.8 shows  $E/\eta$  against the frequency, the  $f_0$  curve at which  $\delta = 0.5 \pi$ , and the  $E/\eta$  and frequency ranges at which  $\delta$  is  $<$  or  $> 0.5 \pi$ . If  $E/\eta \rightarrow 0$ ,  $f_0$  approaches the value of Eq. (6.73), which is  $f_0 = 0.089359$  Hz, i.e. a cycle period of 11.2 s. From Fig. 6.8, the loss tangent, and this the storage modulus, is negative at small  $E/\eta$  (high viscosity) and small frequencies (large cycle periods with small strain rates). This is in accordance with a negative (elastic) modulus of the log law model at very small strain rates (Eq. 5.73 of Fuss 2012).

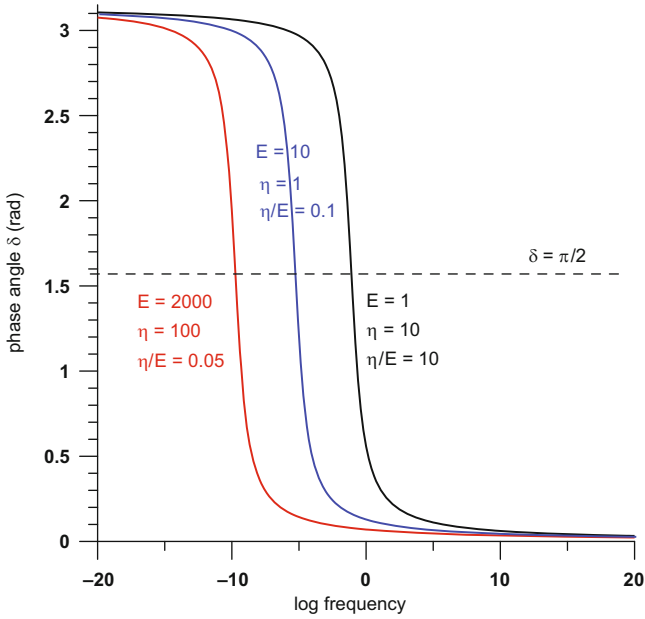
The reciprocal of Eq. (6.68)

$$\cot \delta = \left(\frac{2}{\pi}\right) \left(\frac{E}{\eta} + \log(2\pi f) + \gamma\right) \quad (6.75)$$

shows that the higher  $E/\eta$ , the higher is  $\cot \delta$ . Thus, the higher  $\eta/E$ , the higher is  $\tan \delta$ . This explains that the viscosity of a log law model corresponds to the ratio  $\eta/E$ , and not to the viscosity constant  $\eta$ . Figure 6.9 shows that the higher  $\eta/E$ , the higher is the phase angle  $\delta$ . The viscosity constant  $\eta$  alone has no influence on  $\delta$ .



**Fig. 6.8** Ratio of  $E/\eta$  against frequency  $f$ ;  $\eta$ : viscosity constant;  $E$ : velocity independent elasticity parameter;  $\tan \delta$ : loss tangent;  $f_0$ : frequency at which  $\tan \delta = \pi/2$



**Fig. 6.9** Phase angle  $\delta$  against  $\log$  frequency;  $\eta$ : viscosity constant;  $E$ : velocity independent elasticity parameter

### 6.3 Summary

Material models are always simplified descriptions to assist calculations and assessment of mechanical behaviour. This means that a material does not necessarily have to follow the behaviour of its model. For example, the negative storage modulus of a log law model at small  $E/\eta$  and small frequencies may well be a mathematical phenomenon but does not necessarily exist in reality. The standard linear solid is certainly oversimplified with three elements (2 springs, 1 damper) and is usually expanded to more elements for better linear material characterisation (Prony series; Wiechert model; Wiechert 1889, 1893; Fuss 2012).

#### 6.3.1 Loss Tangent and Viscosity

SLS:  $\tan \delta$  and  $\delta$  depend on  $E$ ,  $\eta$ , and  $f$ ; but at the same  $f$ ,  $\tan \delta$  and  $\delta$  at the same  $R_1$  depend only on  $R_2$  (Eqs. 6.35 and 6.36), i.e. the ratio of moduli of series to parallel spring, but not on the viscosity constant  $\eta$ .

Power law model:  $\tan \delta$  and  $\delta$  depend on  $\eta$  only;  $0 \leq \eta < 1$ .

Log law model:  $\tan \delta$  and  $\delta$  depend on  $E$ ,  $\eta$ , and  $f$ ; but at the same  $f$ , larger  $E/\eta$  have larger  $\tan \delta$  and  $\delta$ .

#### 6.3.2 Relationship Between Frequency and Viscosity Constant

SLS:  $f$  and  $\eta$  are linked together and always appear as the product  $f\eta$ . Equations (6.26) and (6.27).

Power law model:  $\eta$  has no relationship with  $f$  in  $\tan \delta$  (as  $f$  does not influence  $\tan \delta$ ), whereas for  $\sigma_0$ ,  $\eta$  appears in the gamma function, and is the exponent of  $f$ , i.e.  $f^\eta$ .

Log law model: the viscosity constant appears as a stand alone  $\eta$ , and as the product of  $\eta$  and  $\log 2\pi f$ .

#### 6.3.3 Transient and Steady State Parts

SLS: transient part: exponential function; steady state part: sine and cosine functions.

Power law model: transient part: Maclaurin series; steady state part: sine function with  $\eta\pi/2$  phase shift (resulting in sine and cosine functions after applying addition rules).

Log law model: transient part: cosine and sine integrals; steady state part: sine and cosine functions.

### 6.3.4 Negative Storage Modulus if $\tan \delta > \pi/2$

SLS:  $\tan \delta < \pi/2$

Power law model:  $\tan \delta < \pi/2$

Log law model:  $\tan \delta$  can be  $> \pi/2$  at small  $E/\eta$  (high viscosity) and small frequencies (large cycle periods with small strain rates).

## Key Symbols

Ci	Cosine integral function
Cin	Alternative cosine integral function
cos	Cosine function
cot	Cotangent function
d	Differential operator
$E$	Modulus (velocity independent elasticity parameter)
$E_1, E_2$	Moduli of springs of a standard linear solid
$E'$	Storage modulus
$E''$	Loss modulus
$E^*$	Complex modulus
$ E^* $	Dynamic modulus
e	Exponential function
$f$	Frequency
H	Heaviside (unit step) function
$i$	$\sqrt{-1}$
lim	Limit
log	Natural logarithm
$R_1, R_2$	Parameter ratios of a standard linear solid
$s$	Complex variable of transformed functions
Si	Sine integral function
sin	Sine function
t	Time
tan	Tangent function
$\tan \delta$	Loss tangent
$\Gamma$	Gamma function
$\gamma$	Euler–Mascheroni constant (0.577215665)
$\delta$	Phase shift angle
$\varepsilon$	Strain
$\varepsilon_c$	Constant strain

$\varepsilon_0$	Amplitude of strain, peak strain
$\widehat{\varepsilon}$	Transformed strain
$\dot{\varepsilon}_0$	Peak strain rate
$\eta$	Viscosity constant
$\pi$	pi (3.14159 . . . )
$\sigma$	Stress
$\widehat{\sigma}$	Transformed stress
$\sigma_0$	Amplitude of stress, peak stress
$\tau$	Dummy variable of convolution integral
$\omega$	Angular frequency $\omega = 2\pi f$
*	Convolution operator
$\infty$	Infinity

## References

- Abramowitz M, Stegun IA (1972) Handbook of mathematical functions, 9th edn. Dover Publishing, Mineola
- Findley WN, Lai JS, Onaran K (1989) Creep and relaxation of nonlinear viscoelastic materials. Dover Publications, New York
- Fuss FK (2008a) Cricket balls: construction, non-linear visco-elastic properties, quality control and implications for the game. *Sports Technol* 1(1):41–55. doi:10.1002/jst.8
- Fuss FK (2008b) Logarithmic visco-elastic impact modelling of golf balls. In: Estivalet M, Brisson B (eds) *The engineering of sport 7*. Springer, Paris, pp 45–51
- Fuss FK (2009) The collapse timing and rate of closed cell foams: revealing a conceptual misunderstanding of foam mechanics. In: Alam F, Smith LV, Subic A, Fuss FK, Ujihashi S (eds) *The impact of technology on sport III*. RMIT Press, Melbourne, pp 659–664
- Fuss FK (2012) Nonlinear visco-elastic materials: stress relaxation and strain rate dependency. In: Dai L, Jazar RN (eds) *Nonlinear approaches in engineering applications*. Springer, New York, pp 135–170. doi:10.1007/978-1-4614-1469-8\_5
- Lakes R (2009) *Viscoelastic materials*. Cambridge University Press, Cambridge
- Oldham KB, Spanier J (1974) *The fractional calculus*. Academic, New York
- Provenzano P, Lakes R, Keenan T, Vanderby R (2001) Nonlinear ligament viscoelasticity. *Ann Biomed Eng* 29(10):908–14
- Schelkunoff SA (1944) Proposed symbols for the modified cosine and exponential integral. *Q Appl Math* 2:90
- Weyl H (1917) Bemerkungen zum Begriff des Differentialquotienten gebrochener Ordnung. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* 62:296–302
- Wiechert E (1889) Ueber elastische Nachwirkung. Dissertation, Königsberg University, Königsberg
- Wiechert E (1893) Gesetze der elastischen Nachwirkung für constante Temperatur. *Ann Phys* 286:335–348, 546–570

# **Part II**

## **Practical Nonlinearity**

# Chapter 7

## Optimization of Hood Design to Minimise Pedestrian Head Injury in Impact

Revathi Krishnamoorthy, Monir Takla, Aleksandar Subic, and Derek Scott

**Keywords** Pedestrian protection • Head impact • Finite element analysis • Hood panels • Crash safety

### 7.1 Introduction

The term pedestrian generally refers to a person walking, running, skating, skateboarding and commuting using similar devices.

There is a recent trend for urban design to provide for more walkable communities. Present-day awareness of health, fitness and greenhouse gas emissions has increased the popularity of cycling and walking in urban areas. Therefore, the probability of pedestrian accidents has also increased.

Automobiles have evolved to be a successful medium for transportation due to substantial inventions and developments in various countries. The concept of large-scale production-line manufacturing enabled the affordable pricing of automobiles, which has increased the demand for them globally.

In 1900, the United States of America was the only country manufacturing cars and built only 4,192 passenger cars (Elert 2001). There were no buses or trucks at that time. As of 2010, there are more than 600 million passenger cars worldwide (Worldometers 2010). These numbers are increasing rapidly. It is estimated that, if the present trends continue, the number of cars in the world will double by 2030.

As the demand for automobiles has increased, automotive design is improving every day in many aspects of comfort, technology, efficiency, performance and

---

R. Krishnamoorthy (✉) • M. Takla • A. Subic • D. Scott  
School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University,  
Bundoora, PO Box 71, Melbourne, VIC 3083, Australia  
e-mail: [kara947@hotmail.com](mailto:kara947@hotmail.com); [monir.takla@rmit.edu.au](mailto:monir.takla@rmit.edu.au); [aleksandar.subic@rmit.edu.au](mailto:aleksandar.subic@rmit.edu.au);  
[derekemails@gmx.com](mailto:derekemails@gmx.com)





**Fig. 7.1** Unsafe conditions for pedestrians (World Health Organization 2011)

safety. However, as many countries and manufacturers have been participating in the manufacture of cars, there are significant differences in the quality and safety performances offered by the vehicles.

The World Health Organization photograph (Fig. 7.1) portrays the unsafe conditions for pedestrians, who risk their lives in order to fulfil their day-to-day responsibilities.

In 1965, Ryan and McLean (1965) described the sequence of events when a passenger car collides with an adult pedestrian standing erect.

Those sequences are as follows:

- Contact between the bumper and the lower legs of the pedestrian
- Contact between the leading edge of the hood and upper legs/pelvis of the pedestrian
- Contact between the pedestrian's head/upper torso and the top surface of the hood/windscreen/windscreen frame (Fig. 7.2)
- Contact between the pedestrian and the ground.

They also speculated that it would be possible to minimise injuries in a Pedestrian–Passenger Car Front–end Collision (PPCFC) by changing the frontal shape of a car. Their study however did not suggest any design solutions, as they did not have enough data.

In those days, the belief was that the only way to reduce pedestrian fatalities and injuries was to prevent pedestrian–vehicle collisions. Consideration of modification of vehicle design for pedestrian protection was not an option at that time. From this sequence of events, it can be stated that typically the colliding vehicle runs under the pedestrian and the severity of injuries vastly depend on the vehicle shape and certain characteristics such as energy absorption.



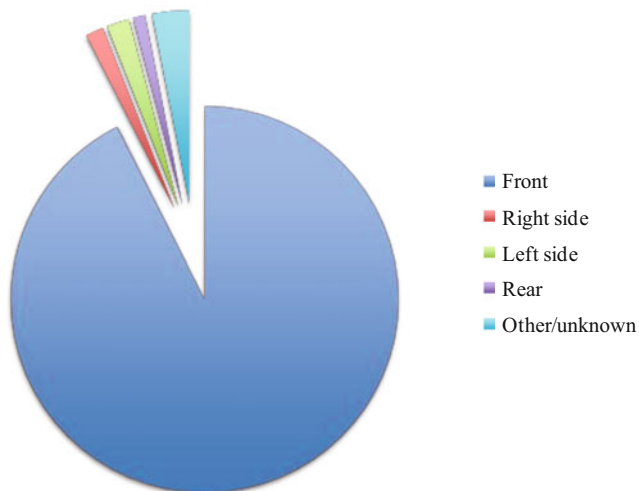
Fig. 7.2 Kinematics of a pedestrian in PPCFC ([Synaptic analysis consulting group inc](#))

## 7.2 Background Research and Benchmarking

A collision between a pedestrian and a passenger car causes injuries to the whole body. Recovery from head injuries however takes substantial time and these injuries are more likely to be fatal. In 2003, Iskander Farooq and his team compiled the data from the International Road Traffic and Accident Database (IRTAD) and the German In-Depth Accident Study (GIDAS). They found that 62 % of pedestrian fatalities occur due to head injuries.

The statistics show that about 92 % of pedestrian fatalities in the US are due to the impact with the front-end of a car (National highway traffic safety administration 2009) as presented in Fig. 7.3. The pedestrian's head contact with the vehicle front-end or the ground could cause significant head injuries. However, studies show that head impact with the vehicle front-end is likely to cause more damage than head impact with the ground.

Other studies on injuries in children pedestrians show that the injuries to different regions of the body vary based on a child's age. Head and chest impacts are the sources of fatal injuries for younger children. Contact between the pelvis and the abdomen with the leading edge of the hood can cause fatal injuries for an older age group.



**Fig. 7.3** Direction of impact for pedestrian fatalities due to collision with a passenger car in the US (National highway traffic safety administration 2009)

**Table 7.1** Distribution of pedestrian injury AIS 2–6 (Mizuno 2003)

Body region	USA (%)	Germany (%)	Japan (%)	Australia (%)	Total (%)
Head	32.7	29.9	28.9	39.3	31.4
Face	3.7	5.2	2.2	3.7	4.2
Neck	0.0	1.7	4.7	3.1	1.4
Chest	9.4	11.7	8.6	10.4	10.3
Abdomen	7.7	3.4	4.7	4.9	5.4
Pelvis	5.3	7.9	4.4	4.9	6.3
Arms	7.9	8.2	9.2	8.0	8.2
Legs	33.3	31.6	37.2	25.8	32.6
Unknown	0.0	0.4	0.0	0.0	0.2
Total	100	100	100	100	100

According to the International Harm Reduction Association (IHRA) accident data, head and leg injuries account for two thirds of pedestrian injuries. Approximately 84 % of AIS 2–6 injuries are caused by pedestrian contact with the vehicle front-end. In Australia (Table 7.1), 39.3 % of AIS 2–6 injuries occur due to head impact.

Head impact is the most prominent factor in pedestrian fatalities and injuries. Thus, it represents a key hurdle in the pedestrian protection assessment criteria. Due to these two critical features, it was decided to focus on improving vehicle design to reduce head injuries and to develop a methodology for optimising hood panels of passenger cars to ensure that the pedestrian Head Injury Criterion (HIC) falls below the threshold values specified by both the GTR-9 and the consumer metric,

the Australasian New Car Assessment Program (ANCAP). This chapter presents the development of a hood configuration that provides robust and homogeneous HIC for different impact positions in the central area of the hood of a large sedan, taking into consideration of the limited space available for deformation.

## 7.3 Vehicle Design Requirements

### 7.3.1 *Regulatory Vehicle Design Requirement for Pedestrian Protection*

The GTR-9 has adopted the test procedures developed by the IHRA. Active and passive components are included so as to ensure the same level of pedestrian protection as 2003/102/EC-Phase II. The active component mandates fitting of Brake Assist, a system designed to sense an emergency braking situation and assist the driver in achieving the maximum possible deceleration in the prevailing conditions, thus reducing the braking distance. The passive Component includes a series of subsystem tests where the resulting injury levels must be below the prescribed limits. This component of GTR-9 requires a careful vehicle design that reduces the probability of severe head and lower leg injury to pedestrians in a collision with the car.

Europe, Japan, China, United States of America, India, Korea and Australia are the few among the many countries plan to adopt these regulations as vehicle design regulatory requirements for pedestrian protection. Negotiations are in progress with automakers and governments globally regarding the timing for enforcing GTR-9.

The GTR-9 test protocol required four subsystem tests to evaluate the vehicle design for protection of pedestrians in a collision with a passenger car. They are:

- Lower leg form fired at 40 km h to bumper system
- Upper leg form fired at 20–40 km h (depending on vehicle shape) to the hood leading edge
- Child head form fired at 35 km h to the hood top, at an angle of 50° to the horizontal plane
- Adult head form fired onto the hood top at an angle of 65° to the horizontal plane, at 35 km h.

### **Head Impact**

The total head impact area is determined in a similar manner to the European Commission Directive. Two thirds of the total impact area zoned as 'HIC1000' and one-third zoned as 'HIC1700'. The HIC1000 zone must contain at least half of the child impact area. This regulation tests the protection offered to both the child and adult head in this impact area. The impact area for the child head is between 1000

WAD and 1700 WAD. The adult head impact area is between 1700 WAD and 2100 WAD. The impact area does not include the windscreen.

### **Lower Leg Impact**

The bottom face of the lower leg impactor is positioned 25 mm above the ground level. The area between the two bumper corners is the lower leg impact area. The resulting injury values for the lower leg within the lower leg impact area must be below a set of prescribed limits in the evaluation of vehicle design. The bounds for the injury values are:

- Tibia acceleration is less than or equal to 170 g
- Knee-bending angle is less than or equal to 19°
- Knee shear displacement is less than or equal to 6 mm.

### **Upper Leg Impact**

The upper leg impact is a guided impact and the impact energy, impact angle and impact velocity are calculated based on the Hood Leading Edge and Upper Bumper Reference Line positions. Currently, there are no legal limits for this test and it is conducted for monitoring purposes only.

## ***7.3.2 Consumer Metric Vehicle Design Requirement for Pedestrian Protection***

The various new car assessment programs, usually made up of a consortium of government and non-government safety organisations, have established testing programs in many countries to evaluate new passenger vehicles for safety performance, including pedestrian protection performance, and published the results for consumers.

The pedestrian protection performance rating reported by Euro NCAP is one of the most influential consumer metrics in the European market. The EEVC WG17 pedestrian protection test procedures serve as a base for these test and assessment procedures. The Euro NCAP protocol version 5.3.1, adopted from February 2010, has harmonised the test procedures with the GTR-9. Euro NCAP also proposes four tests through vehicle subsystems (Fig. 7.4) similar to those of the GTR-9.

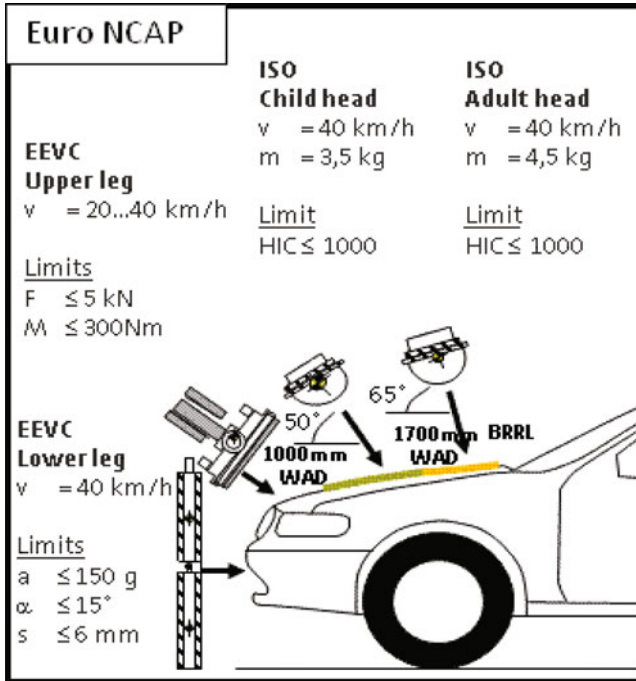


Fig. 7.4 Pedestrian protection technical requirements (Euro NCAP 2010)

### 7.4 Evaluating Pedestrian Protection Offered by a Vehicle

In 1989, EEVC WG10 (Hardy et al. 2007) proposed the impactor test method for vehicle pedestrian impacts. They developed four different types of subsystem test.

As stated earlier, so far all the regulatory requirements and consumer metrics tests that evaluate pedestrian protection use this subsystem test only (Fig. 7.5).

The Institut National de Recherche sur les Transports et leur Sécurité (INRETS) created the original leg form impactor (Hardy et al. 2007). However, EEVC WG 10 and EEVC WG 17 accepted the lower leg impactor developed by Transport Research Laboratory (TRL) (Transport Research Laboratory 2012) to be used in EEVC pedestrian protection testing, due to its high level of performance and repeatability. These TRL-PLI impactors represent the 50th percentile male leg designed for right side impact testing. Accelerometers and potentiometer in the leg form impactor collect the acceleration and the relative rotation to assess the severity of injuries. The TRL-PLI has a rigid femur and tibia covered with foam to represent soft tissues and skin (Fig. 7.6).

In 2000, to improve the accuracy of the leg form, JAMA and the Japan Automobile Research Institute (JARI) initiated the development of a Biofidelic



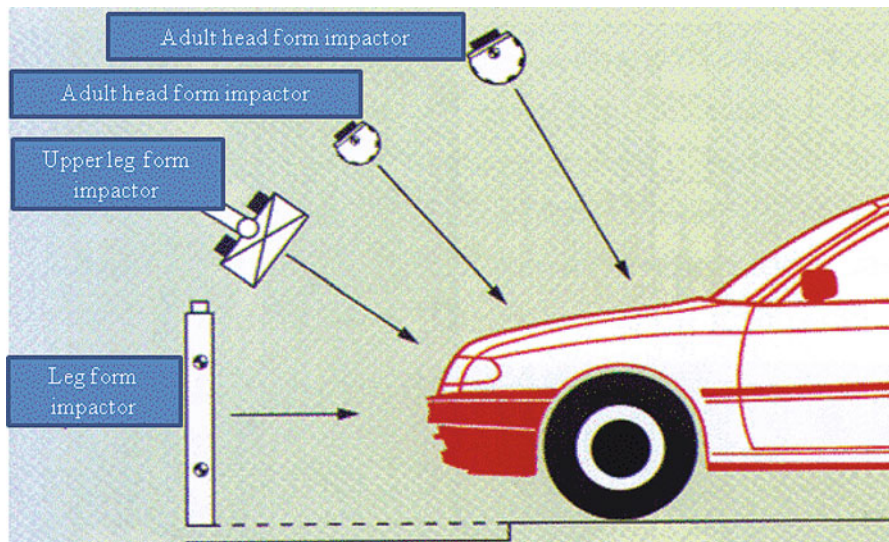


Fig. 7.5 Subsystem test method (Hardy et al. 2007)

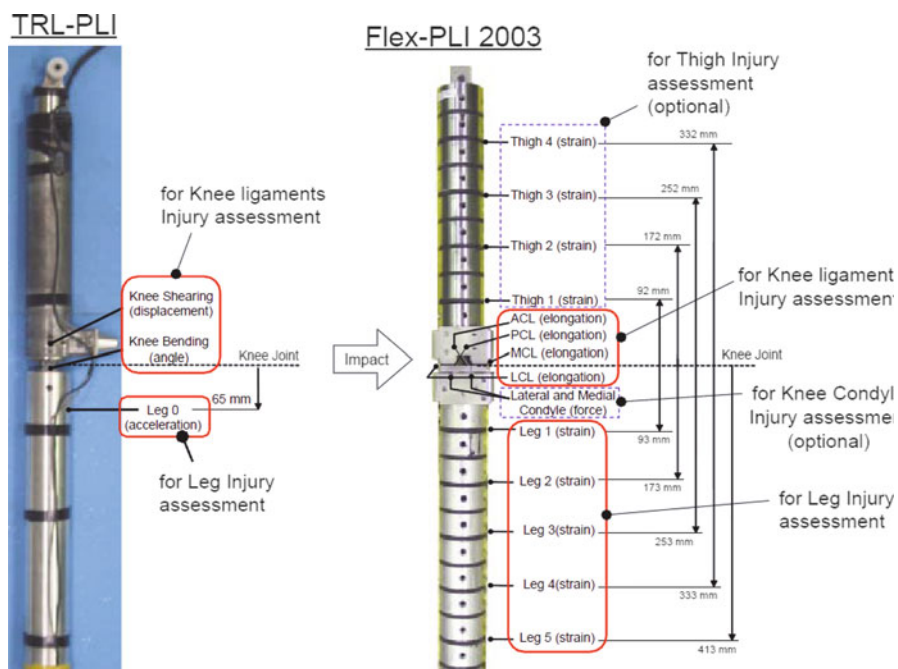


Fig. 7.6 Comparison between TRL-PLI and Flex-PLI (World Health Organization 2010)



**Fig. 7.7** TRL-PLI lower leg impactor



**Fig. 7.8** Leg impact and head impact test machines (Centre of automotive safety research 2007)

Flexible Pedestrian Leg form Impactor, Flex-PLI (Fig. 7.7). This impactor uses four femur joints and five tibia joints to improve the dynamic response.

Flex-TEG, a technical evaluation group consisting of government and industry members, is evaluating the possibility of using the Flex-PLI impactor for the GTR-9. The linear guidance propulsion system launches the lower leg impactor in a free flight mode when it has attained steady desirable velocity. This test simulates the human lower leg (tibia, knee joint and femur) in a passenger car front-end impact with a pedestrian.

In 1995, Germany's federal highway research institute (BAST) conducted upper leg form tests. Later TRL improved this upper leg form impactor. The EEVC accepted this upper leg form impactor in its tests, even though JARI reported that the measured forces are still lower than the inertia forces. During the test, a linear guided launcher fires the upper leg impactor onto a static vehicle (Fig. 7.8). This test simulates the upper leg covered with foam to replicate flesh and skin and rotates around a friction-loaded pivot to mimic femur and pelvis fractures.

Originally, production of the head form impactors used a plastic sphere covered by a rubber skin. EEVC WG17 updated the head form specifications to an aluminium sphere covered with a 12 mm PVC skin without changing the outer



diameter and mass for continuity of measurement. The head impactor is launched in the same manner as the lower leg impactor. The head form contains a tri-axial accelerometer that is used to measure the resultant acceleration throughout the impact. The HIC value is calculated using the measured resultant acceleration.

In 2005, a team from the University of Virginia and General Motors Corporation (Kam et al. 2005) conducted a study to develop full-scale pedestrian impact tests using Post Mortem Human Surrogates (PMHS). This study was conducted utilising existing test procedures, real-world data and MADYMO simulations to create an experimental test system with full-scale pedestrians. In this study, the PMHS was positioned laterally in mid-gait stance with the struck leg extended to the rear and both hands extended to the front. The PMHS was attached to an electromagnetic release mechanism, which was triggered to release the PMHS just before impact. After the impact, a catcher mechanism would absorb the kinetic energy from the PMHS to avoid any secondary injuries because of contact with the ground. This study did not lead to the development of a full-scale pedestrian dummy.

In contrast, Honda developed a crash test dummy to understand the kinematics of vehicle pedestrian impacts in 1998 with a second-generation released in 2000, the POLAR II. This dummy gathers measurements from the head, cervical spine, thorax, abdomen, pelvis, femur, knee and tibia. In 2008, Honda introduced the third-generation pedestrian dummy, POLAR III (Fig. 7.9). This dummy has an improved ability to evaluate bone fractures in the lower back and upper leg compared to the POLAR II. Lower back and upper leg injuries are common in pedestrian impact with higher vehicles like SUVs.

At present, both NCAP and regulatory tests use subsystem testing due to their simplicity, repeatability and the ability to predict injury values accurate enough for pedestrian protection. Since the number of tests required to assess pedestrian protection is large, testing with full-scale pedestrian dummies would be expensive and time consuming. The author is not aware of any current plans to examine pedestrian kinematics using full-scale pedestrian impact tests.

## **7.5 Problem Identification**

### ***7.5.1 Peak Acceleration***

Numerical tests were conducted in this work to obtain accelerations, time histories and impactor displacements in  $x$ ,  $y$  and  $z$  directions. The resultant accelerations calculated from these numerical tests were used in calculating the HIC values. As explained by Krishnamoorthy et al. (2013) in *Advanced Materials Research*, the HIC calculation involves an exponent of 2.5 applied to the average head acceleration measured over a window of 15 ms. The result of this formulation is that small changes in the average acceleration can lead to large changes in the final HIC value.



**Fig. 7.9** POLAR III pedestrian dummy (Honda Worldwide 2008)

The average acceleration of any given head acceleration curve can be increased by:

- Simply increasing the magnitude of a single peak acceleration
- Increasing the duration or width of a single peak
- Increasing the time difference between the first and an adjacent peak
- Increasing the magnitude of the acceleration of the first peak and the adjacent peaks.

The accelerations calculated from the numerical tests conducted in this work are shown in Figs. 7.10 and 7.11. It is evident from the curves that the HIC value increases with the magnitude of resultant acceleration and increases with the recording period for similar peak acceleration. Therefore, a low magnitude and small duration of peak acceleration are required to keep the severity of head injuries less than the limits specified by the GTR-9 and ANCAP.

Decreasing the structural stiffness of the hood panels decreases the magnitude of acceleration and in turn increases the intrusion of the hood panels, otherwise known as the deformation space. It is this interplay between the measured average acceleration (and hence HIC) and the structural characteristics of the hood that requires careful development if we are to minimise our requirement for the highly valuable deformation space. Thus, optimisation of the structural stiffness of the hood panels is essential to meet the requirements for head impact regulations.

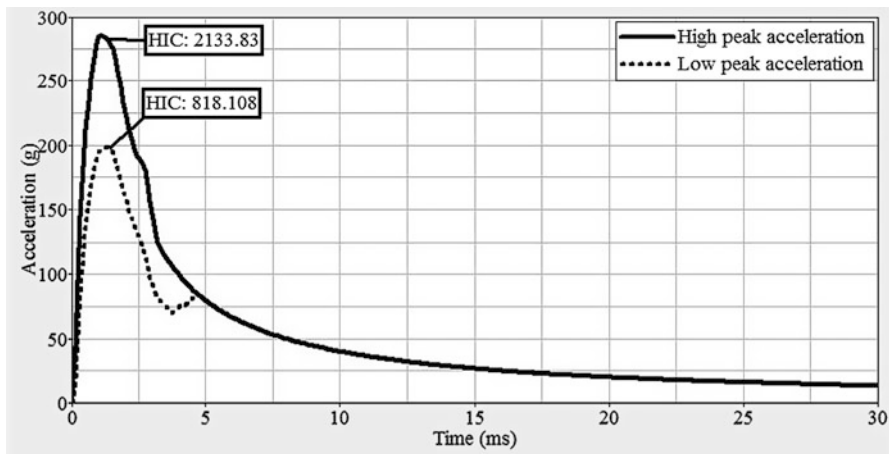


Fig. 7.10 Comparison of peak acceleration magnitude

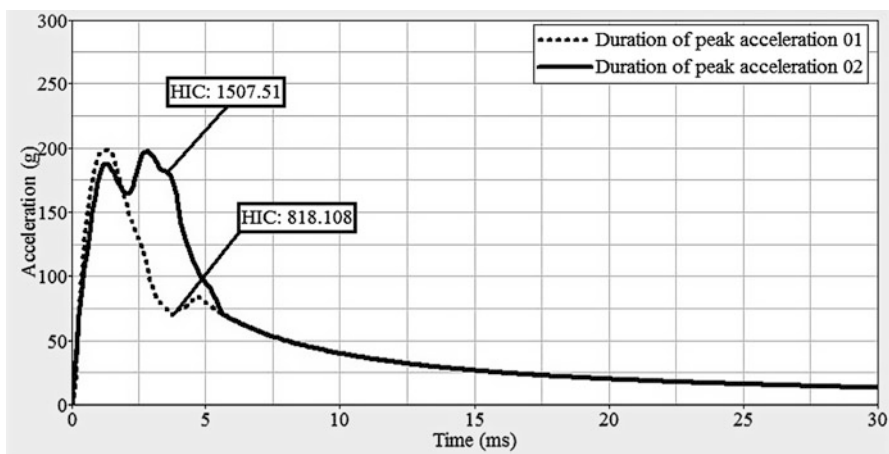


Fig. 7.11 Comparison of peak acceleration duration

### 7.5.2 Deformation Space

Pedestrian protection requires a soft, energy-absorbing hood assembly, whereas durability, noise and vibration require a hood assembly with higher structural stiffness. The properties of the direct and indirect contact surfaces of the head in a crash influence the severity of head injuries. The indirect contact surfaces are the surfaces under the hood. The three important parameters that influence the severity of pedestrian head impact injuries are the available deformation space, structural

stiffness and structural inertia. In this instance, the structural inertia is the resistance given by the hood panel to changing its state to the changes in velocity and is proportional to its mass.

Among these parameters, the available deformation space between the hood assembly and the hard components within the engine bay is critical. The intensity of the contact between the hood assembly and the under hood components with high structural stiffness (secondary impact) increases the severity of pedestrian head injuries. The deformation space required to absorb the energy varies with the head impact mass, impact speed and impact angle. Intensity of the secondary impact can be reduced either by increasing the structural stiffness and structural inertia of the inner and outer hood panels or by increasing the deformation space.

Increasing the structural stiffness of the hood increases the magnitude of acceleration during impact due to reducing the hood deformation. Therefore, excessive stiffness of the hood causes severe acceleration, which in turn increases the severity of head injuries. Accordingly, increasing the structural stiffness to reduce the intensity of the secondary impact is only possible up to a point, beyond which the severity of injuries due to the contact between the hood assembly and the head impactor (primary impact) becomes excessive. The simulations from the numerical tests conducted in this work with increased structural stiffness for the inner and outer hood panels are shown in Fig. 7.12. It shows that the deformation of the hood panels is reduced whereas; the structural inertia of the panels has increased.

The resultant acceleration calculated for the simulation shown in Fig. 7.12 is presented in Fig. 7.13, which exhibits a dominant initial peak resulting from high inertial loading due to the structural inertia of the inner and outer hood panels.

Similarly, the simulation and resultant acceleration from the numerical tests conducted in this work with the lack of deformation space are presented in Figs. 7.14 and 7.15. These figures show that the lack of deformation space leads to direct contact of the head form with the hard components approximately 12 ms from the start of the impact. This causes severe acceleration, hence increasing the risk of head injury.

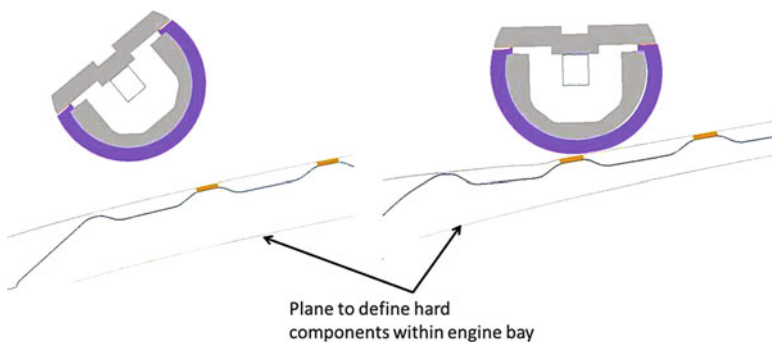


Fig. 7.12 Primary impact of head with hood panels with excessive stiffness

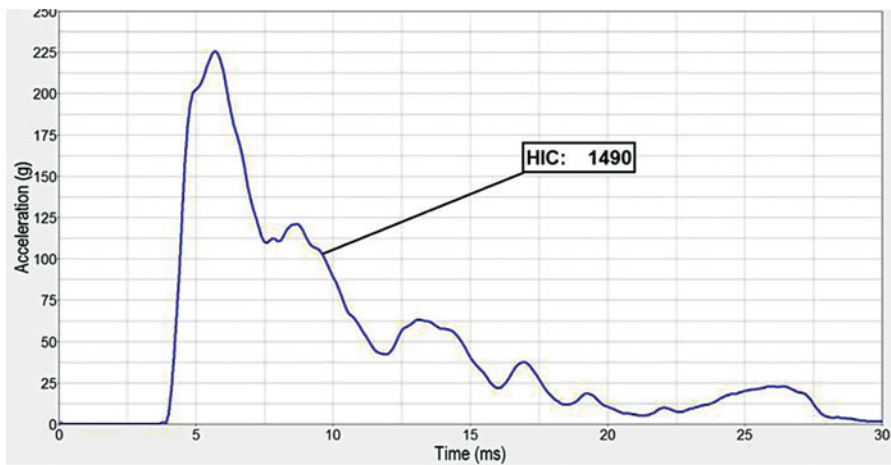


Fig. 7.13 Effect of inertial loading of the head form on HIC value

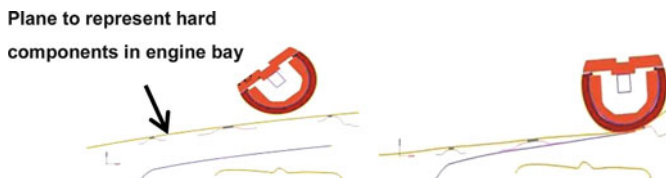


Fig. 7.14 Secondary impact of head with hard components within engine bay

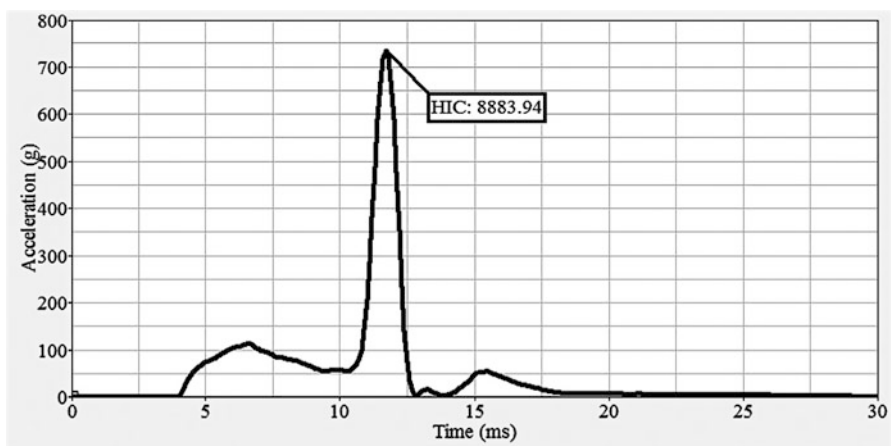


Fig. 7.15 Waveform with secondary impact

While increasing the deformation space improves pedestrian protection, it may conflict with other vehicle design requirements. For example, the positioning of

the hard components, such as the engine and the suspension, should meet ground clearance requirements. In addition, the relative height difference between the hood and the driver has to meet visibility requirements. Therefore, when the hood is elevated to meet the deformation space requirements for pedestrian protection, the driver also needs to be elevated to meet the visibility requirements. Thus, vehicle manufacturers need to increase the height of the car for occupant comfort, ergonomics and safety, which will compromise aerodynamics, vehicle stability and fuel economy. It might also have adverse effects on vehicle styling.

Therefore, in order to reduce the pedestrian head injury risk, it is important to optimise the hood assembly by balancing its structural stiffness for minimal deformation or risk a series of far reaching flow on effects into a whole range of areas such as engine bay packaging, occupant comfort and visibility, styling and even aerodynamics.

### 7.5.3 *Optimal Waveform*

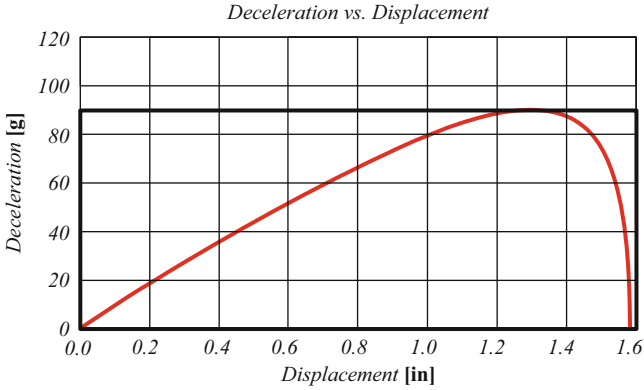
A waveform in this study is the shape of a curve describing the relationship between the resultant acceleration and time. Meeting the requirements of the HIC value with the lowest deformation space requires balancing hood stiffness and deformation space. The parameters that affect the HIC value are impact velocity, available deformation space, magnitude of peak acceleration and pulse duration. The resultant acceleration and pulse duration are the key factors that affect the HIC value. These factors are influenced by the impact velocity and the available deformation space. In regulatory and consumer metrics tests, the impact velocity is given as a constant. Therefore, the main factors that affect the HIC value, which is calculated from the wave form, are the structural stiffness and the deformation space.

In 1995, Lim et al. explored options to develop a methodology for estimating the minimum deformation space required in an impact to a pedestrian head. They used generic waveforms and idealised waveforms of square waves, sine waves and half-sine waves in formulating equations for deformation space requirements.

They used the ‘area ratio’ method to determine the waveform efficiency. In this method, the ratios of areas under the acceleration–intrusion curves, shown in Fig. 7.16, are used to determine the efficiency of the waveform, as shown in Eq. (7.1).

$$\eta = \frac{A_1}{A_2} \quad (7.1)$$

In which  $\eta$  = efficiency,  $A_1$  = area enclosed by the acceleration–displacement curve of the given waveform and the horizontal plane,  $A_2$  = area of the shape that encloses the geometrical shape of the acceleration–displacement curve of the given waveform.



**Fig. 7.16** Waveform efficiency ( $\eta$ ) of a half-sine waveform (Lim et al. 1995)

With this method, the square waveform was determined as the most efficient waveform, with 100 % efficiency (Fig. 7.16). Lim and his team calculated the efficiency of the other waveforms by benchmarking against the square waveform. Thus, the waveform efficiency of a sine waveform was determined as 63.66 % and of a half-sine waveform as 50 %.

In 2007, Wu and Beaudet (2007) challenged the optimality of aiming at the square waveform to achieve the minimal deformation space required to obtain a given HIC value. They derived a realistic optimal waveform that is more practical based on the theoretical optimal waveform, which is WSTC.

As shown in Eq. (7.2) below, HIC value calculation is the average of head acceleration  $a(t)$  over an interval  $t_1$  to  $t_2$ , followed by  $a(t)$  is raised to the 2.5th power and multiplied by the interval. The condition for selection of the interval is that it should provide the maximum result within the 15 ms window:

$$HIC = \max \left\{ \left[ \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} a(t) dt \right]^{2.5} (t_2 - t_1) \right\} \quad (7.2)$$

To achieve the maximum result for HIC value for a given magnitude of peak acceleration, the acceleration between  $t_1$  and  $t_2$  should be equal, which produces a square waveform. Therefore, assuming  $a(t_1) = a(t_2)$ , Wu and Beaudet derived the formula to calculate the maximum intrusion of the hood, which produces a square waveform in a pedestrian head impact as shown in Eq. (7.3):

$$d = 11.12 \times 10^{-3} \frac{v_0^{\frac{8}{3}}}{HIC^{\frac{2}{3}}} \quad (7.3)$$

Where  $d$  = maximum intrusion in metres,  $v_0$  = impact velocity in m/s.

**Table 7.2** Comparison of deformation space requirements for various waveforms

Waveform	Deformation space in impact direction (mm)
Optimal	50.50
Square	67.32
Sine	79.49
Half-sine	96.38
Isosceles triangle	83.97
Ramp-up triangle	111.94
Ramp-down triangle	55.98

Wu and Beaudet stated that numerous other waveforms show same HIC value with smaller deformation space requirements. They concluded that the square waveform is ideal when the criterion is either maximum head acceleration or force. Since HIC is the criterion for head impact, the function derived is directly related to WSTC, which proposes that the head can tolerate large head acceleration for a very short duration and smaller head acceleration for a longer duration. Therefore, they considered WSTC as the optimal waveform and derived the equation for calculating the head travel or deformation space as stated in Eq. (7.4):

$$d = \int v \times dt = 8.342 \times 10^{-3} \times \frac{v_0^{\frac{8}{3}}}{HIC^{\frac{2}{3}}} \quad (7.4)$$

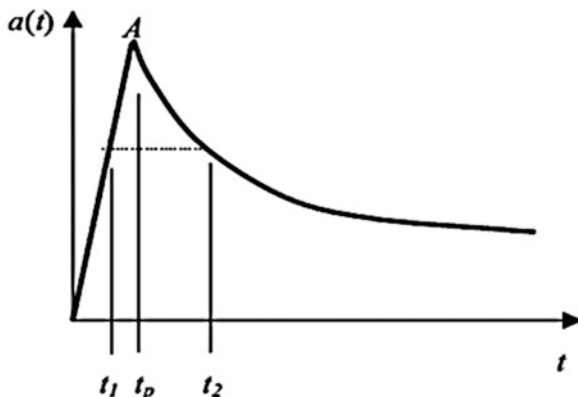
Comparison of the deformation space requirement for a HIC value of 1000 at 40 km/h impact velocity of the optimal waveform and other waveforms reveals that the optimal waveform requires the least amount of deformation space and the second best is the ramp-down triangle waveform. Ramp-up triangle waveform in Table 7.2 refers to a triangular waveform with late peak in which the resultant acceleration gradually increases over time and drops down over a short period. Ramp-down triangle waveform is the triangular waveform with early peak in which the resultant acceleration increases from zero to peak within a short period and decreases gradually. Wu and Beaudet did not consider rebound in their calculations.

The optimal waveform (WSTC) suggests that head acceleration could be infinite at time zero, but in reality, this is impossible. Therefore, Wu and Beaudet considered the optimal waveform as the theoretically optimal waveform. When  $t_p$  equal a 2 ms, delay is applied to obtain peak acceleration, the theoretically optimal waveform becomes realistic and efficient, thus realistic optimal waveform as shown in Fig. 7.17.

With the realistic optimal waveform, they stated that as design guidance, the deformation space required to obtain a HIC value less than or equal to 1000 is 60 mm for a child head at 40 km/h impact velocity. For adult head impact, they suggested that the required deformation space is 12 % more than that required for child head impact, due to the steeper impact angle.



**Fig. 7.17** Realistic optimal waveform for pedestrian head impact (Wu and Beaudet 2007)



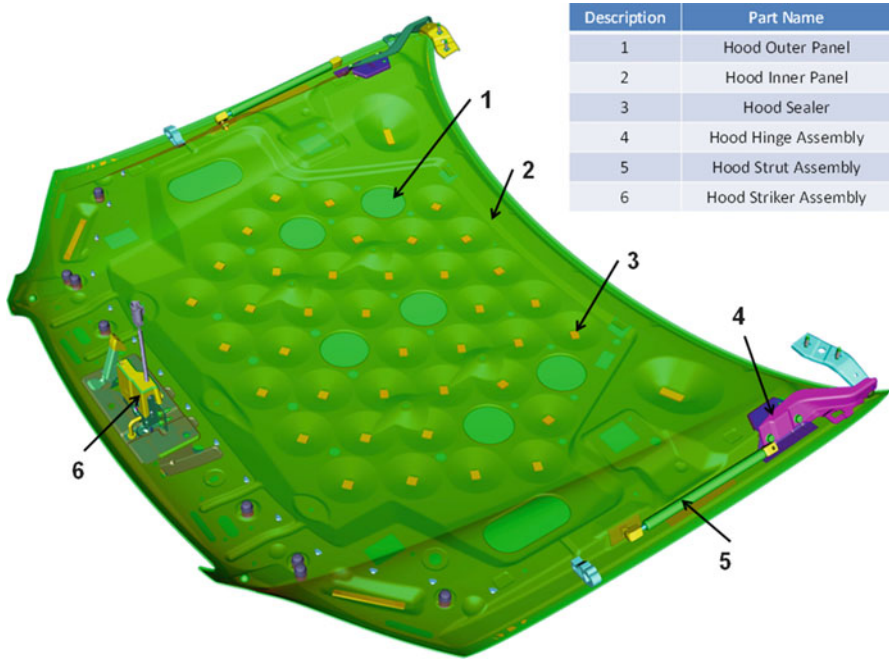
In the research presented in this chapter, the impact performance offered by a vehicle to the child head has been tested with 60 mm of deformation space as a starting point. Wu and Beaudet derived the value of the required deformation space using theoretical calculations ignoring rebound and other parameters that could influence the head impact performance. In practice, the value of the deformation space required to meet the compliance threshold homogeneously will be higher than 60 mm.

#### 7.5.4 Influence of Hood Design

The structure and stiffness of the components in a hood assembly (Fig. 7.18) of a car generally influence the severity of head injury in a pedestrian hood impact. To reduce the severity of head injuries closer to the edges of the hood, the hood assembly has to be weaker in this area. This will increase the deflection of the hood assembly in the middle, thus increasing the severity of head injury closer to the edges of the hood. The size of the hood also has an effect on the probability of severe head injury because, as the hood size is increased, the deflection in the middle of the hood also increases. Thus, more deformation space is required to reduce the intensity of the secondary impact.

The challenge of reducing the severity of head impact near the fender can be addressed by changing the hood design from an inlaid hood (Fig. 7.19) to a wraparound hood (Fig. 7.20). However, the wraparound hood design raises issues related to repair costs in low-speed impacts because it is more exposed to damage.

The length of the hood also plays a vital role in pedestrian protection in a passenger car front-end collision with a pedestrian. In the case of a shorter hood length, the pedestrian is likely to come into contact with the A-pillar, generally an area of very high stiffness, resulting in increased risk of head injury. As the length of the hood increases, the challenge to meet the durability, noise and vibration guidelines become more difficult.



**Fig. 7.18** Underside view of a hood assembly



**Fig. 7.19** Inlaid hood (A2Mac1 2010)

The choice of material and thickness of the outer hood panel is important in reducing the probability of severe head injuries. A combination of material and thickness with better energy absorption is favourable for pedestrian protection.



**Fig. 7.20** Wraparound hood (A2Mac1 2010)

The preferred material for pedestrian protection should also satisfy quality, durability, noise and vibration requirements. At present, the materials generally considered for the outer hood panel include steel, aluminium, plastic, carbon fibre and fibreglass.

The inner hood panel acts as an energy absorber and provides torsional and bending stiffness to the hood assembly. Therefore, the choice of structure, material and thickness for this component also plays a vital role in reducing the severity of head impact.

Many concepts for the structure of the inner hood panel have been considered by automakers around the world such as multi-cone, multi X pattern, frame, single skin, double skin etc.

The choice of glue that is used between the inner and outer hood panels also has an important role in reducing the severity of head injuries. The stiffness of the hood assembly also increases with the number of gluing points and glue length between the inner and outer hood panels.

Therefore, optimising the central area of a large hood poses several challenges. This work includes developing a hood configuration that provides a robust and homogeneous head impact performance for different impact positions in the central area of the hood of a large passenger car while minimising the deformation space required.

## 7.6 Methodology

### 7.6.1 *Parametric Geometric Model Creation*

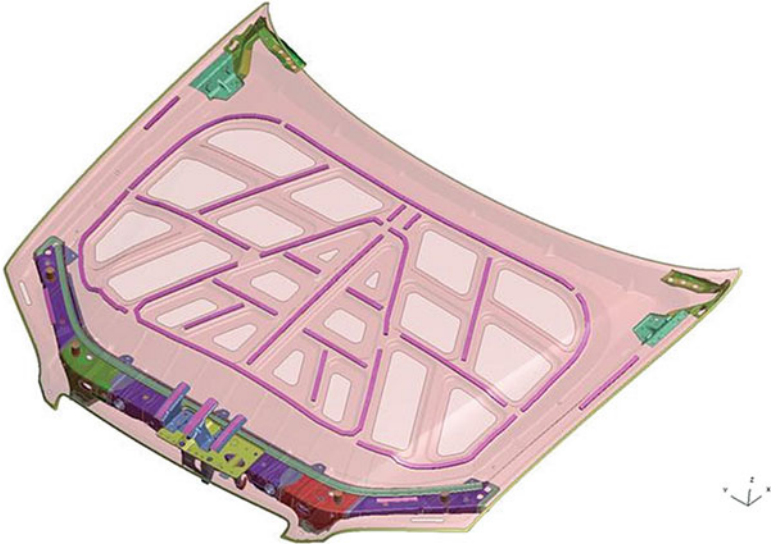
In this research, the designs of the hood assembly were created as three-dimensional parametric geometric models using the chunky solid method in Unigraphics NX5 CAD software. The design boundaries were set up using datum planes at set coordinates. This methodology enables faster design modifications in the hood assembly.

Once a design alternative was developed, checking for fitment and clearance was carried out by virtually assembling the hood to the original donor vehicle. The mid-surface was extracted from these geometric models to enable the FE model development to assess the head impact performance.

### 7.6.2 *Finite Element Model Creation*

The FE models were developed using ANSA V13 with the mid-surfaces extracted from the geometric models. PRIMER was utilised to replicate the hood interface conditions by assembling these FE models to the donor vehicle. The section, material properties, joints and constraints to the components of the hood assembly to the replicated design solution intensions were assigned using PRIMER as shown in Fig. 7.21. The FE model of the head impactor was positioned at the intended impact position and angle as shown in Fig. 7.22. The head impact condition of the pedestrian was simulated using LS-DYNA-3D Computer Aided Engineering (CAE) software as exhibited in Fig. 7.23. It shows different orientations of the head impactor at 5 ms intervals for one data point. The HYPERGRAPH and HYPERVIEW programs were used to post-process the results which are shown in Figs. 7.24, 7.25, and 7.26.

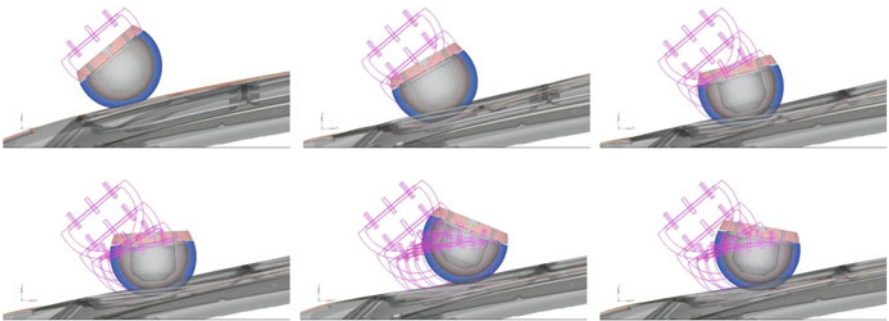
The energy data from the finite element analysis is presented in Fig. 7.27. In general, the energy absorption data is used as a quality check for the finite element model. In the finite element model constructed and utilised in this work, the energy input is from the mass and velocity of the impactor. There are no other components such as airbag is included in the model that can influence change in total energy within the duration of the analysis. Therefore, the total energy remains constant for the duration of the head impact analysis as shown in Fig. 7.27. The kinetic and internal energy plots demonstrate that the major portion of the kinetic energy is converted into internal energy. Thus, it can be stated that the finite element model does not contain errors either in contact definition or in constraints, which can influence energy fluctuations.



**Fig. 7.21** FEA model for optimisation



**Fig. 7.22** FEA model with head impactor for optimisation



**Fig. 7.23** FEA model simulation using LS-DYNA for optimisation

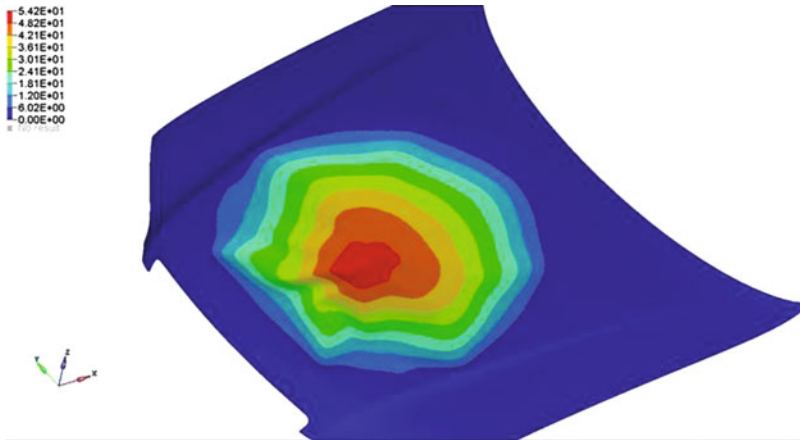


Fig. 7.24 Contour plot of hood displacement using HYPERVIEW

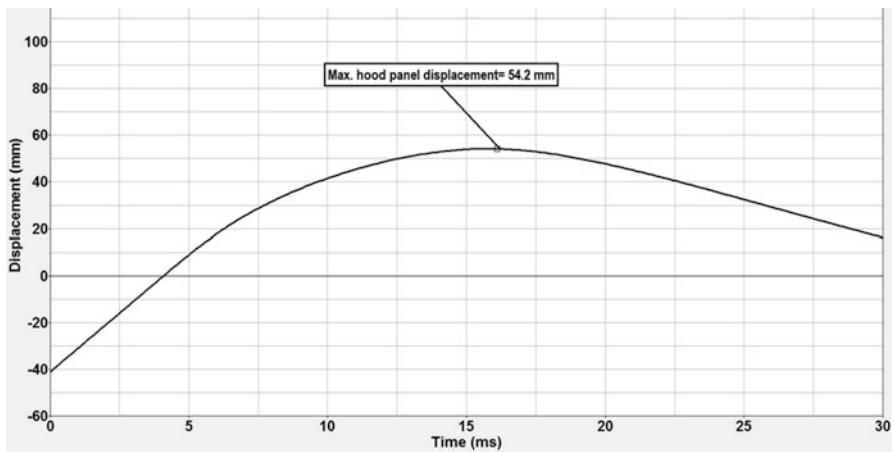


Fig. 7.25 Plot of hood displacement vs. time using HYPERGRAPH

### 7.6.3 Head Impact Model

Currently, finite element models of head impactors, complete anthropomorphic test dummies and complete human body models are being used for head impact analysis in the safety community. In this research, only the child head impactor considered as this is the impactor used in the GTR-9 and NCAP tests to assess the protection offered by a vehicle in a child head impact.

The FE model for simulating the head impact contains the hood assembly and the interfacing components to constrain the hood assembly. It contains a validated head impactor model and the prescribed requirements for the impactor such as mass,

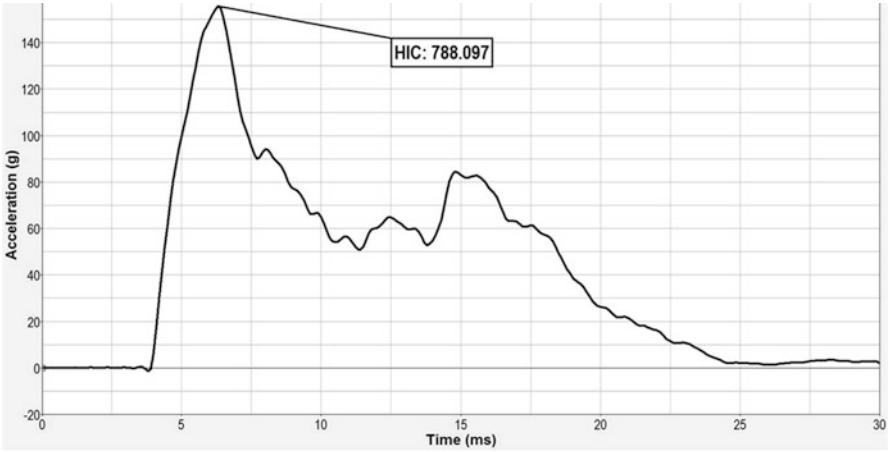


Fig. 7.26 Plot of resultant acceleration vs. time using HYPERGRAPH

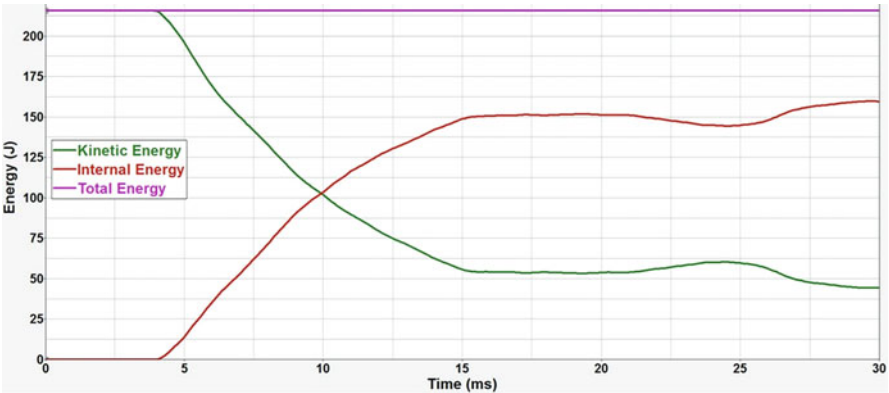


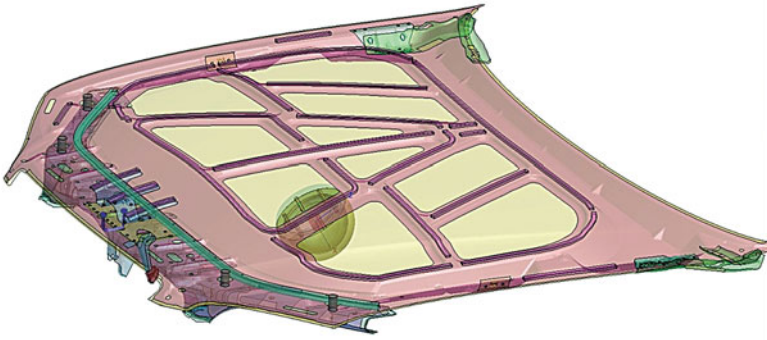
Fig. 7.27 Energy plot from the finite element model

diameter, angle, position and initial velocity of the impact to match the ANCAP pedestrian protection testing protocol.

To minimise the CPU time for solving the FE model, the FE model was developed for only the hood assembly and the front tie bar assembly. The front tie bar assembly has been clipped, as shown in Fig. 7.28. The interface between the modelled parts and the rest of the car body has been modelled by attaching it to the ground through applying appropriate boundary conditions.

The deforming components have been modelled accurately with the necessary mesh quality. This requires checking of the element length, aspect ratio, warpage and angles of the mesh in the FE model. The position and properties of the connections between the components such as the welds, joints, clinches and adhesives have been matched to the physical vehicle build. Validated material





**Fig. 7.28** Pedestrian head impact finite element model

models have been used to assign the material properties of various components used in this research. Stress–strain curves at specific strain rates have been included for the deforming components to improve the predictive accuracy of the model.

The interaction between the vehicle and impactor has been defined using the contact definitions in LS-DYNA. The coefficient of friction has been assigned to define the friction between the vehicle components and the impactor components as well as the friction between the vehicle and the impactor.

#### ***7.6.4 Correlating Finite Element Simulations with Experimental Results***

The experimental results of a large sedan already in the market and tested by ANCAP for pedestrian protection have been utilised for correlation. ANCAP tested this vehicle to assess the level of pedestrian protection offered using the Euro NCAP pedestrian subsystem testing protocol version 5.3. The pedestrian protection experimental tests were conducted at the Centre of Automotive Safety Research, Adelaide.

Figures 7.29 and 7.30 show good correlation between experimental test results and CAE simulations in both measured head acceleration and the deformation pattern of the hood outer surface.

The correlation was used to calibrate the FE models to obtain good predictive accuracy in the numerical tests for the optimal hood design solution.



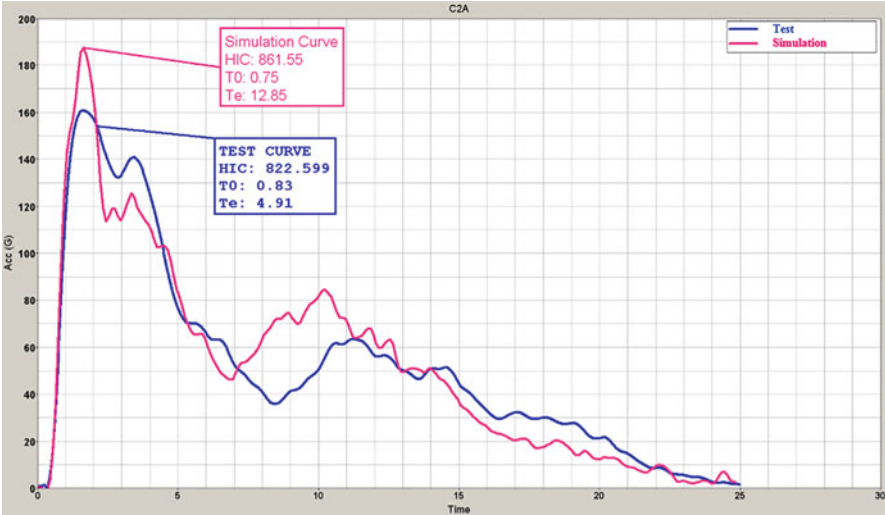


Fig. 7.29 Correlation between ANCAP test and FE simulation

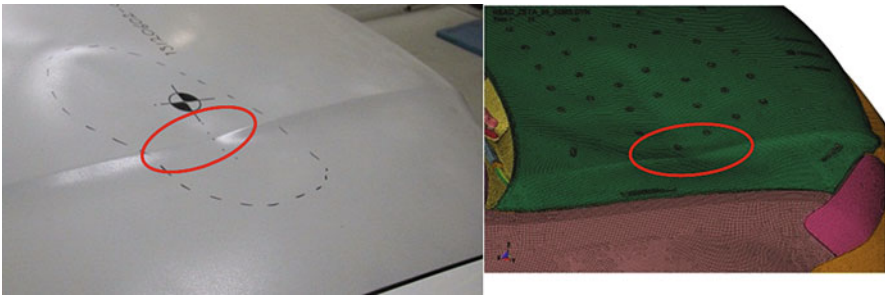


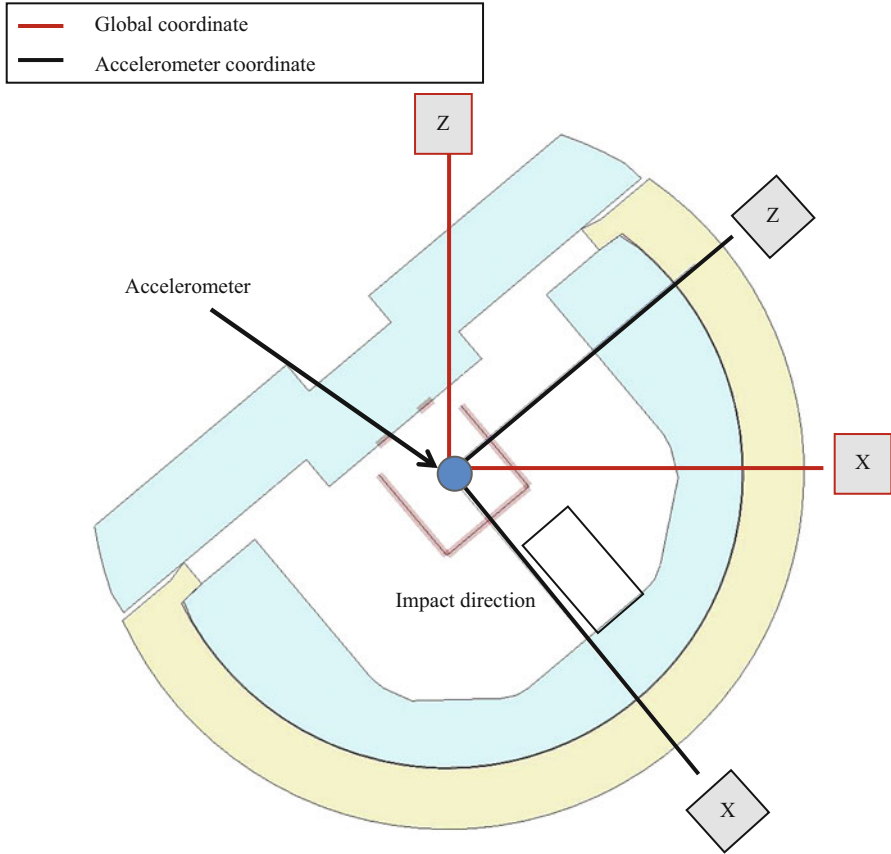
Fig. 7.30 Correlation of hood deformation between ANCAP test and FE simulation

### 7.6.5 Computation of Hood Intrusion

The accelerometer, which measures the acceleration in three directions, is located at the mass centre of the impactor. In the FE model, a node has been placed at this location to represent the accelerometer and enable computation of the impactor displacement in both local and global coordinates, as shown in Fig. 7.31.

When the displacement is determined in the impact axis using the node on the accelerometer, the displacement (Intrusion A) is more than the calculated intrusion in impact direction (Intrusion T), as shown Fig. 7.32.

This is due to the rotation of the accelerometer axis along with the impactor rotation in rebound as shown in Fig. 7.33. Therefore, the impactor intrusion in the vertical axis, which is the Z direction, was computed in global coordinates.



**Fig. 7.31** Coordinates for intrusion computation

The intrusion computed in the Z direction was multiplied by  $1/\sin(\theta)$ , which was calculated using trigonometry, to derive the intrusion in the impact direction (Intrusion T) as exhibited in Fig. 7.34.

Although 'Intrusion T' was calculated from the displacement of the centre of the head impactor in the global Z direction, there could be a variation between the actual intrusion and 'Intrusion T' if the first point of contact or impact point is different to the measured point. In addition, an assumption has been made that there is no compression of the head form in the impact. In this research, the 'Intrusion T' measurement has been utilised as the actual deformation space required in all the calculations and comparisons.



Fig. 7.32 Comparison of impactor intrusion



Fig. 7.33 Accelerometer rotation due to impactor rebound

## 7.7 Vehicle Design for Pedestrian Head Impact Protection

### 7.7.1 Design Space Definition

The design space is the available space into which the hood assembly should fit without affecting the down vision plane, according to the regulatory requirement ECE-R125 and positioning of the components within the engine bay.

Figure 7.35 illustrates the available packaging space for inner and outer hood panels. The inner and outer hood panels are the two main components in the hood assembly that affect pedestrian protection performance.

Therefore, the scope of this research is to optimise the inner and outer hood panels to meet the HIC value requirements with minimal deformation, as shown in

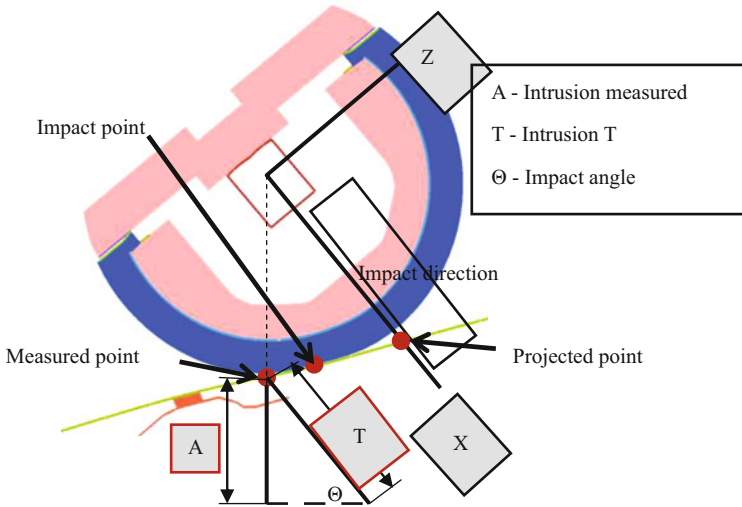


Fig. 7.34 ‘Intrusion T’ calculation method

Fig. 7.35 Design space for hood assembly

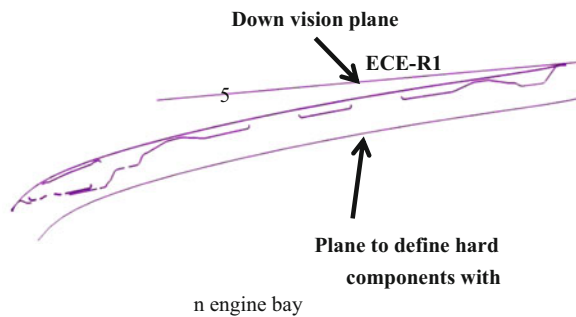


Fig. 7.36. However, modifications that may be required in interfacing components to improve the pedestrian safety performance offered by the vehicle are not included in the scope of this research.

### 7.7.2 Impact Area Definition

The head impact area on the hood was defined as per the GTR-9 test area definition. The Hood Leading Edge (HLE—Fig. 7.37) or 1000 Wrap Around Distance (WAD), Hood Rear Reference Line (HRRL—Fig. 7.38) and Side Reference Lines (Fig. 7.39) determine the boundaries for head impact area. WAD divides the head impact area into adult and child head impact zones. WAD is the distance measured from the ground using a flexible tape held taut on the outer surface of the vehicle front structure in a vertical longitudinal plane.

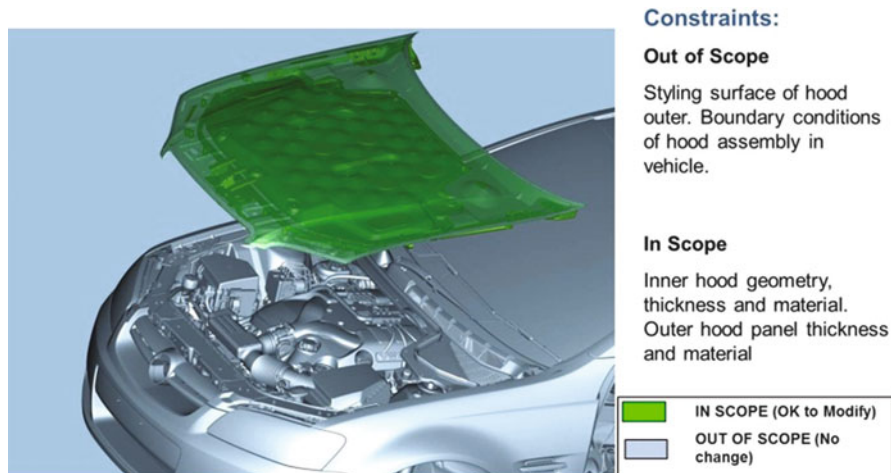


Fig. 7.36 Scope of the project

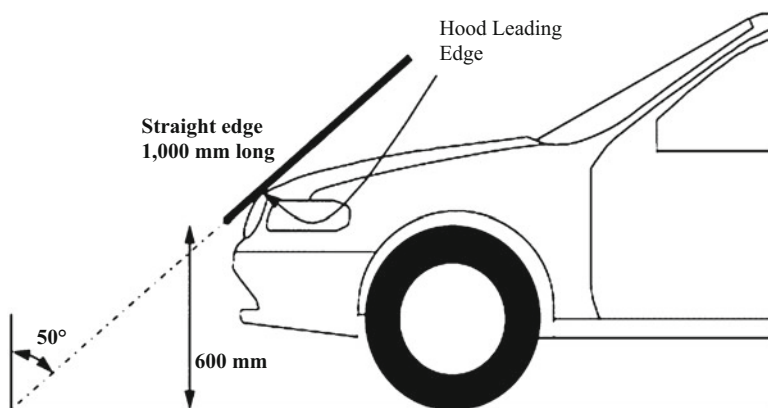


Fig. 7.37 Hood leading edge (United Nations Economic Commission for Europe 1998)

In the GTR-9 protocol, the accident data and technical feasibility of regulating within the test area are the basis for WAD boundaries. The child impact area is defined from 1000 WAD or HLE to 1700 WAD and the adult head impact area are defined from 1700 WAD to HRRL or 2100 WAD. The resulting test area is as shown in Fig. 7.40.

A range of impact points has also been selected for the numerical simulations, varying in fore-aft as well as in cross-car directions of the vehicle at regular intervals in order to analyse the central area of the hood, as shown in Fig. 7.41. These impact positions are within the child head impact zone. This area has been selected because

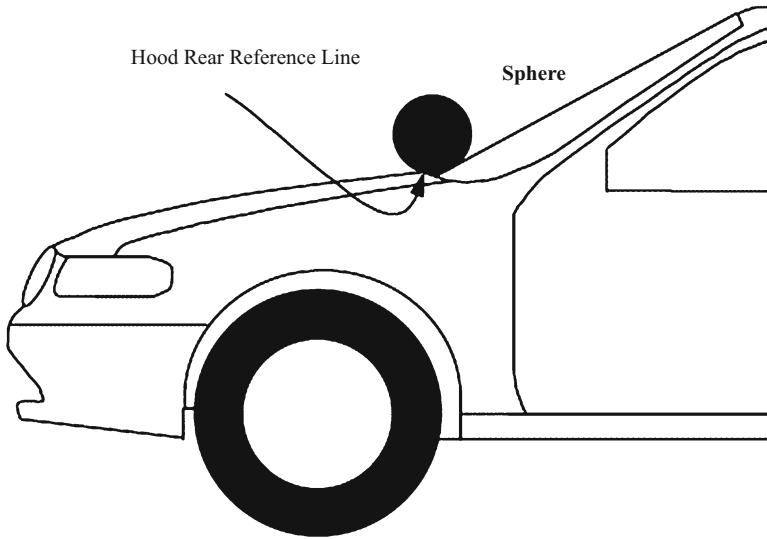


Fig. 7.38 Hood rear reference line (United Nations Economic Commission for Europe 1998)

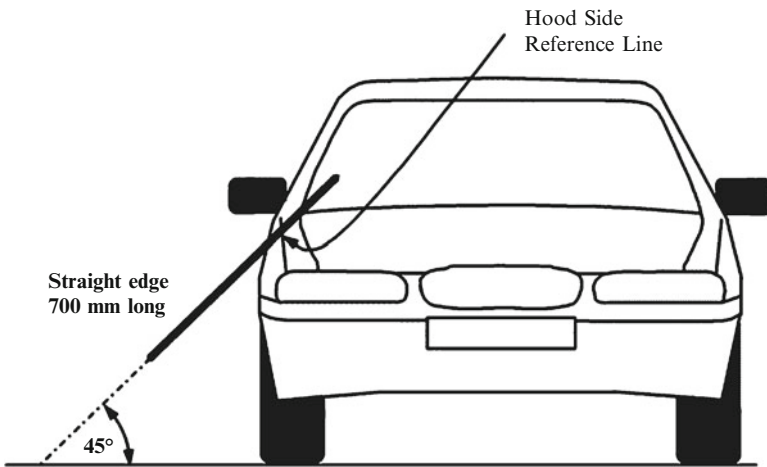
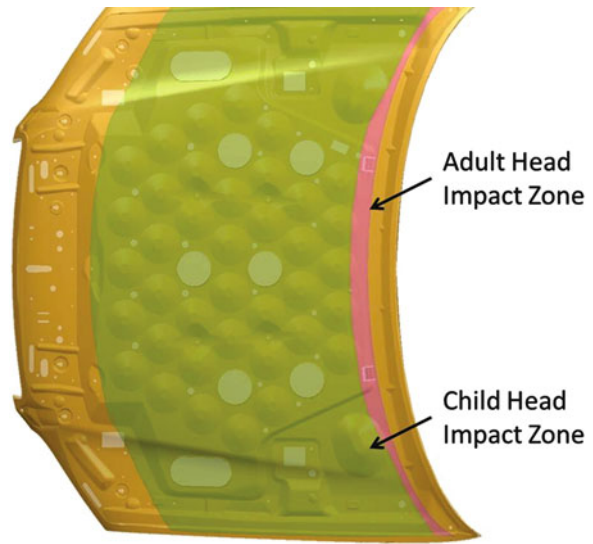


Fig. 7.39 Hood side reference line (United Nations Economic Commission for Europe 1998)

high hood intrusions occur in the central area of the hood and the interfaces around the edges prevent deformation.

The impactor speed has been set to the ANCAP requirement, 40 km/h that is higher than the GTR-9 test speed, which is 35 km/h, to analyse the most severe load case.

**Fig. 7.40** Head impact test area on hood



**Fig. 7.41** Impact positions considered for optimisation



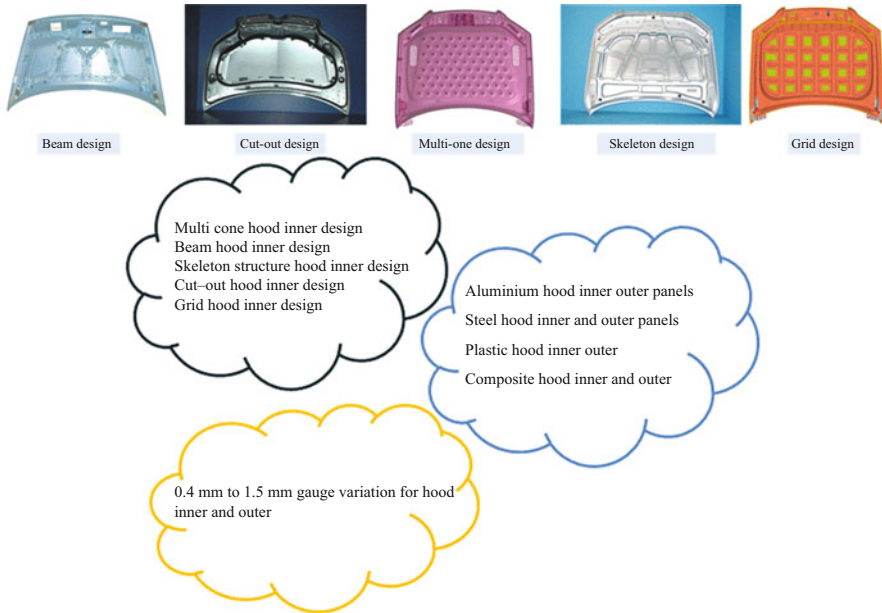


Fig. 7.42 Concepts for improving pedestrian protection performance

## 7.8 Parameters for Improving Pedestrian Protection

### 7.8.1 Concepts Generation

Various concepts for the inner hood structure were considered. The selection of concepts was done referring to literature and the hoods available in the market that could be considered in the hood design for a large sedan as shown in Fig. 7.42. The inner hood structures considered for improving pedestrian protection performance are the beam, cut-out, skeleton, grid and multi-cone patterns. The beam pattern has beams positioned in the fore–aft direction of the vehicle, which provides structural stiffness to the hood assembly. The cut-out pattern is a mass-saving and cost-saving option in which the centre portion of the inner hood has a large opening. The skeleton pattern is beams nested in an A-shape. This pattern allows improving the structural stiffness of the hood in the centre, but outboard openings reduce the mass and structural stiffness. The multi-cone pattern is cone shaped forms nested at constant distances. This design provides homogeneous structural stiffness. The grid pattern consists of longitudinal and lateral beams arranged at regular intervals. Thus, this pattern also should provide homogeneous structural stiffness and the cut-outs enable mass reduction.

Aluminium and steel were considered in the analysis for the inner and outer hood material because they generally meet the requirements for noise and vibration,



Concept	Steel (current)	Aluminium (concept1)	Plastic (concept2)	Composite (concept3)
Criteria				
Impact on cost	-	+	-	-
Impact on mass	+	+	+	+
Ease of assembly installation	S	+	S	S
Durability	S	-	-	-
Ability of with stand engine compartment heat	S	-	-	-
Hood stiffness consistency	S	-	-	-
Compatibility with high-speed crash conditions	S	-	-	-
Noise and vibration	S	-	-	-
Long-term reliability	S	-	-	-
Appearance	S	-	-	-
Σ+ (Sum of positives)	1	3	1	1
Σ- (Sum of negatives)	1	7	8	8
ΣS (Sum of "same")	8	0	1	1

Concept	Beam design	Cut-out design (concept1)	Multi-cone design (concept2)	Skeleton design (concept3)	Grid design (concept4)
Criteria					
Impact on cost		+	-	-	-
Impact on mass		+	-	-	-
Ease of assembly installation		S	S	S	S
Complexity of design (manufacturability)		S	S	S	S
Durability		S	+	+	+
Hood stiffness consistency		-	+	+	+
Compatibility with real-world impacts		-	+	+	+
Noise and vibration		-	+	+	+
Long-term reliability		-	+	+	+
Appearance		-	+	+	+
Σ+ (Sum of positives)		2	6	6	6
Σ- (Sum of negatives)		5	2	2	2
ΣS (Sum of "sames")		3	2	2	2

+	= Better than datum
-	= Worse than datum
S	= Same as datum

Fig. 7.43 Pugh matrix

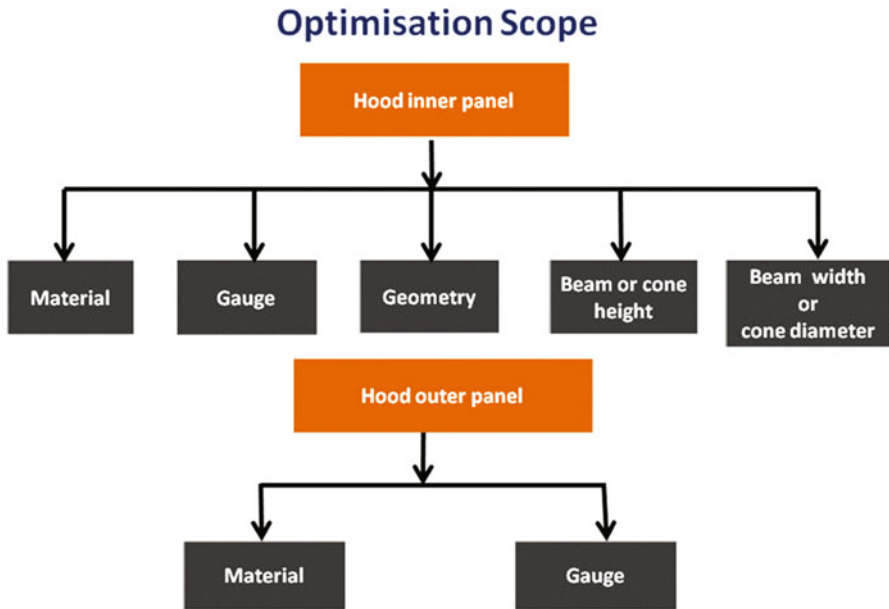


Fig. 7.44 Inner hood geometries chosen for developing an optimal solution

durability and high-speed crash performances. A thickness range of 0.4–1.5 mm was considered for the inner and outer hood thickness, values less than 0.4 mm will not meet the durability and forming requirements while those above 1.5 mm will result in high structural stiffness and structural inertia as well as increased mass.

### 7.8.2 Concepts Evaluation

The concepts mentioned in the previous section were analysed through the Pugh matrix as shown in the previous page in Fig. 7.43. The comparison of the concepts to various performance criteria, multi-cone, skeleton and grid patterns for inner hood structure (Fig. 7.44) ended up with the same merits, as well as aluminium and steel for the material. Therefore, these configurations were chosen for further evaluation. The hood attachment and boundary strategies, though may improve pedestrian protection performance, were not been taken into consideration throughout the analysis.



**Fig. 7.45** Control factors in scope for refinement

**Table 7.3** Variables considered for refinement

Variable	Values of variable	Type of variable
Inner hood geometry	Skeleton pattern, multi-cone pattern grid pattern	Discrete
Inner/outer hood material	Aluminium steel	Discrete
Inner hood gauge	0.4–1.5 mm	Continuous
Outer hood gauge	0.4–1.5 mm	Continuous
Impact position	1–8	Discrete

### 7.8.3 Control Factors in Scope

Design parameters such as thickness and material for the inner and outer hood panels as well as inner hood structure were considered for improving the pedestrian protection performance. Figure 7.45 shows the list of control factors that are in scope to modify the concepts selected through the Pugh matrix.

The variables considered in this study have been summarised in Table 7.3; tolerances otherwise known as noise factors that might have some influence on the resulting HIC value and the deformation of hood panels have not been considered in this research work for simplicity purposes.

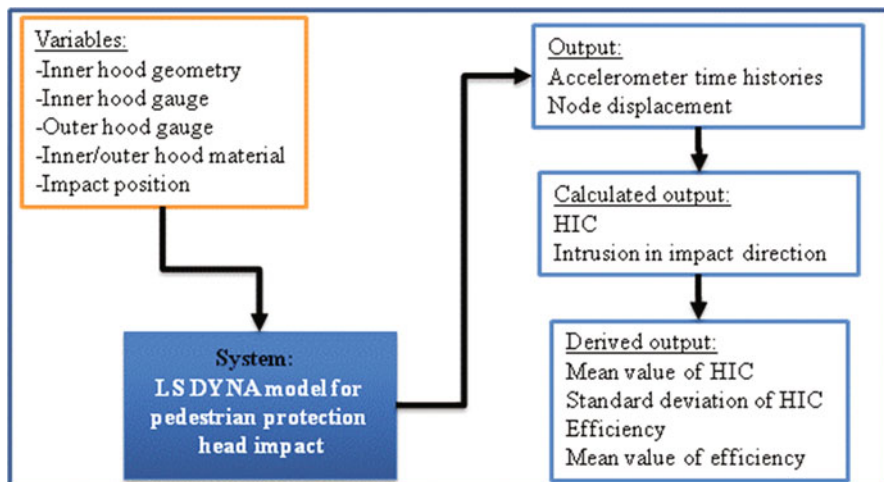


Fig. 7.46 Parameter diagram

### 7.8.4 Concepts Refinement

The parameter diagram as shown below in Fig. 7.46 clearly outlines the relationship between the input and output parameters.

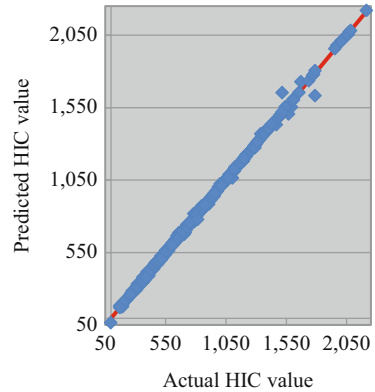
## 7.9 Optimisation of Hood Panels

### 7.9.1 Methodology

Strength-two orthogonal array design of experiments was used in this work to create a matrix with the variables mentioned in Table 7.3. This matrix was utilised to create a multitude of hood designs by varying the values of these design parameters. Non-linear LS DYNA models were created to replicate pedestrian protection head impact physical testing for all the configurations as per the Design of Experiments (DOE) matrix. The accelerometer time histories were gathered as output for all numerical tests, and from that, the HIC values were calculated using Eq. (7.2). The vertical intrusion of the head impactor from the node displacement was also collected to calculate the intrusion in impact direction. Accordingly, the efficiency was derived for all hood configurations. Finally, the mean value of efficiency ( $\mu\eta$ ), the mean value of HIC ( $\mu\text{HIC}$ ) and the standard deviation of HIC ( $\sigma\text{HIC}$ ) were derived for all the hood configurations while considering all impact positions.

The  $\mu\eta$  provides average efficiency between the impact positions for a given configuration of the hood assembly. Similarly,  $\sigma\text{HIC}$  from the calculated mean

**Fig. 7.47** Predictability of response surface



value for a given configuration of the hood assembly between the impact positions was utilised to quantify the homogeneous behaviour of the hood assembly. The  $\mu\text{HIC}$  provides the average HIC value between all the impact positions for a given configuration of the hood assembly. The aim was to obtain the lowest value of  $\sigma\text{HIC}$  thus achieving better homogeneous head impact performance with the highest value of  $\mu\eta$  for HIC values less than 1000.

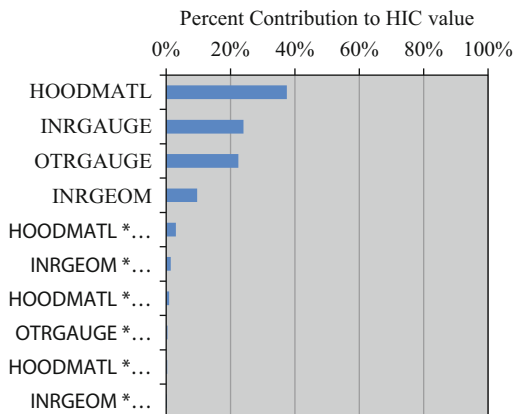
The Kriging response surface based approach was used as an interpolation method to predict the values for the variables considered in the improvement of pedestrian protection. The Kriging interpolation method was named after a South African mining engineer DG Krige and it has been used as a fundamental tool in the field of mining (Bohling 2005). In recent times, this method has been successfully implemented in a variety of applications. In the Kriging approach, the response surface is mapped with limited sample data in the design space followed by utilising this surface to estimate the values of variables at locations where the sample data is unavailable. In this research, response surfaces were mapped with the numerical results for  $\mu\eta$ ,  $\mu\text{HIC}$  and  $\sigma\text{HIC}$ .

Figure 7.47 shows the predictive accuracy of the  $\mu\text{HIC}$  response surface. The analysis of variance (ANOVA) technique was used to assess the relationship between the inputs and outputs as well as to identify the important variables in this study (Gelman 2004). The ANOVA charts showed that the material, thickness of the inner hood panel, thickness of the outer hood panel and the hood inner structure contribute a substantial percentage in determining the  $\mu\eta$ ,  $\mu\text{HIC}$  and  $\sigma\text{HIC}$  values. The ANOVA graph for the  $\mu\text{HIC}$  was presented in Fig. 7.48.

In Fig. 7.48, HOODMATL is the hood panel material, OTRGAUGE is the thickness assigned for the outer hood panel, INRGAUGE is the thickness assigned for the inner hood panel and INRGEOM is the type of structure of the inner hood.

The Monte Carlo method, which is a mathematical technique for statistical sampling was utilised to check whether any other configuration might provide a better output than the ones considered in the sample. This method generates a cloud of input variables that obey the defined properties for which the output was predicted from the existing response surfaces.

**Fig. 7.48** ANOVA chart for mean value of HIC



The sum of the mean value of HIC and its standard deviation ( $\mu_{\text{HIC}} + \sigma_{\text{HIC}}$ ) provides 67 % confidence to this research, assuming normal distribution (Narasimhan 1996). The aim of this study is to obtain 95 % confidence by setting the HIC threshold condition to the sum of the mean value of HIC and twice its standard deviation ( $\mu_{\text{HIC}} + 2\sigma_{\text{HIC}}$ ) to be less than or equal to 1000. Aiming for 99 %, robustness requires the HIC threshold condition to be the sum of the mean value of HIC and three times its standard deviation ( $\mu_{\text{HIC}} + 3\sigma_{\text{HIC}}$ ). For a production representative hood, it would be necessary to achieve all the impact positions less than the HIC value of 1000. This would be achieved through the detailed design work on a specific hood design and its interfacing conditions. Since this research focuses on a more general study looking at material, thickness and construction of hood panels, the 95 % confidence limit is used.

The severity of pedestrian head injuries could be influenced either by the primary impact or by the secondary impact. These concepts were refined in two stages. The first step was the selection of the type of the inner hood structure and inner and outer hood material first by considering only primary impact followed by considering both the primary and secondary impacts. The second step was the selection of the outer hood panel thickness and the inner hood panel thickness and structure. The two-stage optimisation approach was utilised in this research to reduce the number of simulations and modelling complexity.

### 7.9.2 Selection of Hood Structure and Material

The components within the engine bay were not considered in this analysis model; thus, only primary impact has been taken into consideration.

In this research, the efficiency of a hood design was calculated as the ratio of the actual deformation of hood assembly ('Intrusion T') and the optimal deformation calculated from the HIC value of the corresponding analytical test. The optimal

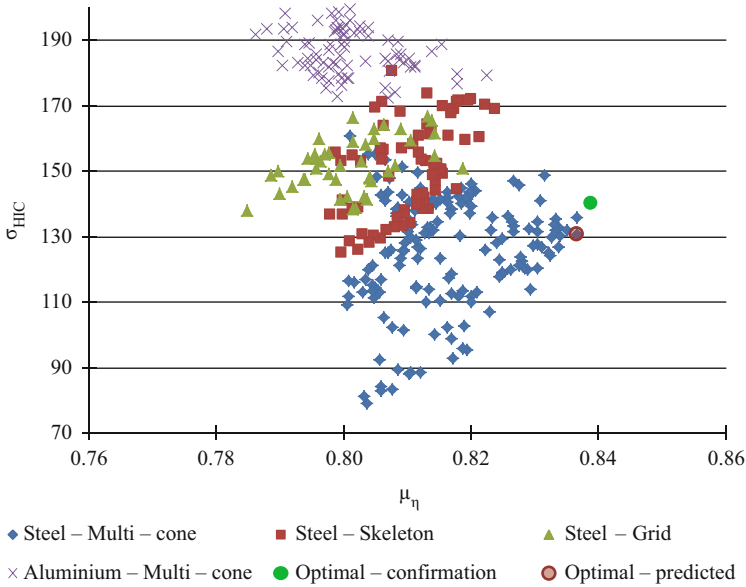


Fig. 7.49 Evaluation of efficiency

deformation is the theoretical minimum deformation needed for a given HIC value and it was calculated using Eq. (7.5) (Wu and Beudet 2007) as shown below:

$$d = 8.342 \times 10^{-3} \frac{v_0^{\frac{8}{3}}}{HIC^{\frac{2}{3}}} \tag{7.5}$$

Where, d is the deformation and  $v_0$  is the impact velocity.

In numerical testing, the actual deformation was computed in the Z direction and used to derive the ‘Intrusion T’ as discussed earlier. The equation to calculate the efficiency ( $\eta$ ) of the hood assembly is:

$$\eta = \frac{\text{Optimal deformation}}{\text{Intrusion T}} \tag{7.6}$$

Thus, a hood configuration with maximum efficiency for an impact had minimum ‘Intrusion T’ for a given HIC value.

Figure 7.49 shows the  $\mu_\eta$  against respective  $\sigma_{HIC}$ . Among these configurations, the optimal hood configuration that provides the highest  $\mu_\eta$  was identified. The numerical tests conducted to check the predictive accuracy of this optimal solution show that the results are comparable. This showed that the deviation between the predicted and numerical values is within 0.01 and therefore it could be stated that accurate predictions can be made from the response surfaces.

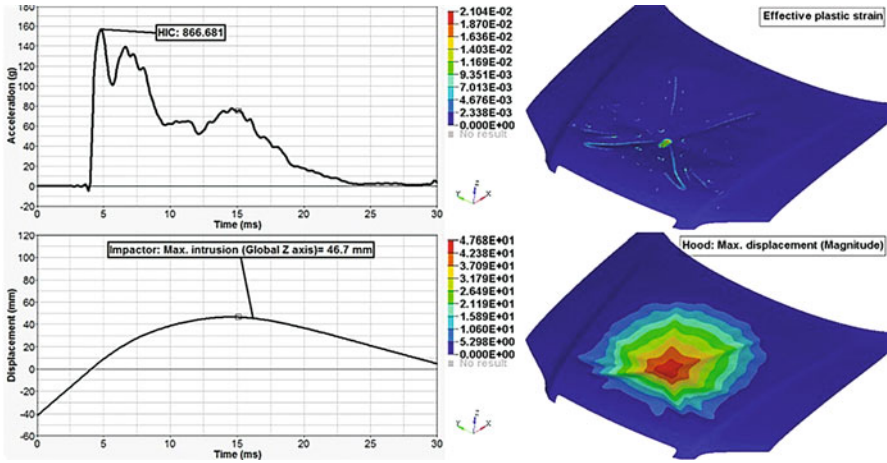


Fig. 7.50 HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point A

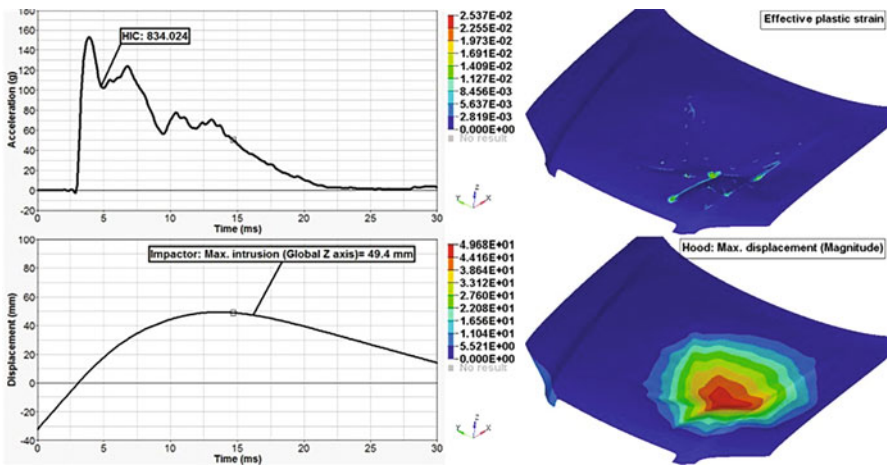


Fig. 7.51 HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point B

Numerical tests were conducted for eight impact locations (A–H) as shown in Fig. 7.41 to derive the values for this data point and the details of the results from the numerical tests are shown in figures ranging from Figs. 7.50, 7.51, 7.52, 7.53, 7.54, 7.55, 7.56, and 7.57. The details included in these figures are,

- Plot of resultant acceleration against time
- Plot of the displacement of the impactor in global Z direction against time
- Contour plot of the effective plastic strain
- Contour plot of the magnitude of the deformation of the hood panels.

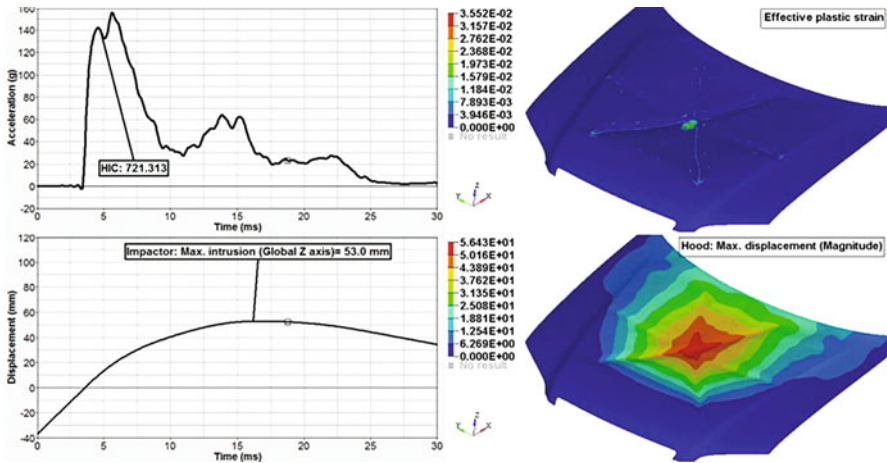


Fig. 7.52 HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point C

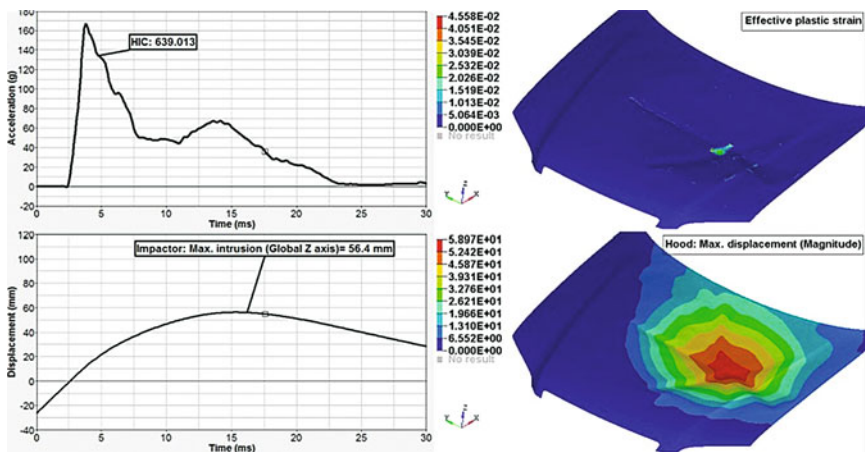


Fig. 7.53 HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point D

The results exhibited in figures ranging from Figs. 7.50, 7.51, 7.52, 7.53, 7.54, 7.55, 7.56, and 7.57 show that the deformation of the hood panels is not excessive, even while secondary impact is not considered, due to the optimisation of the structural stiffness of the hood panels.

HIC values less than the threshold value (1000) are feasible with various combinations of inner and outer hood thicknesses. The efficiency variation is only 3 % for various combinations of the inner and outer hood thickness between 0.5 and 1.4 mm.



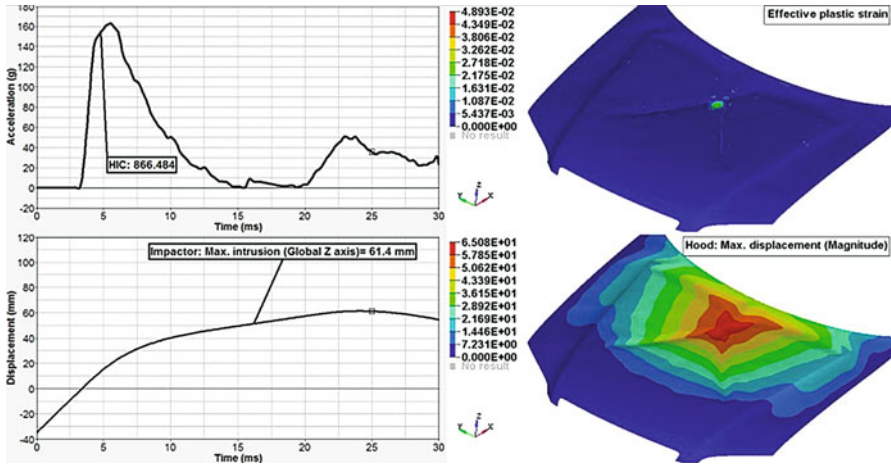


Fig. 7.54 HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point E

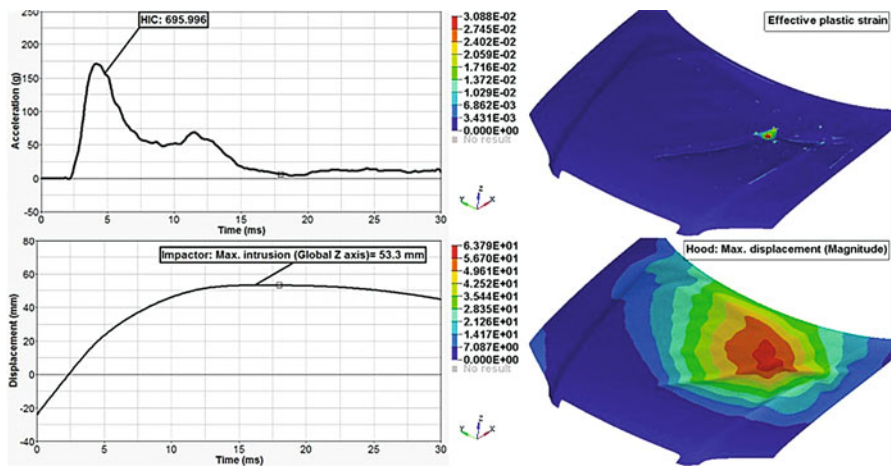
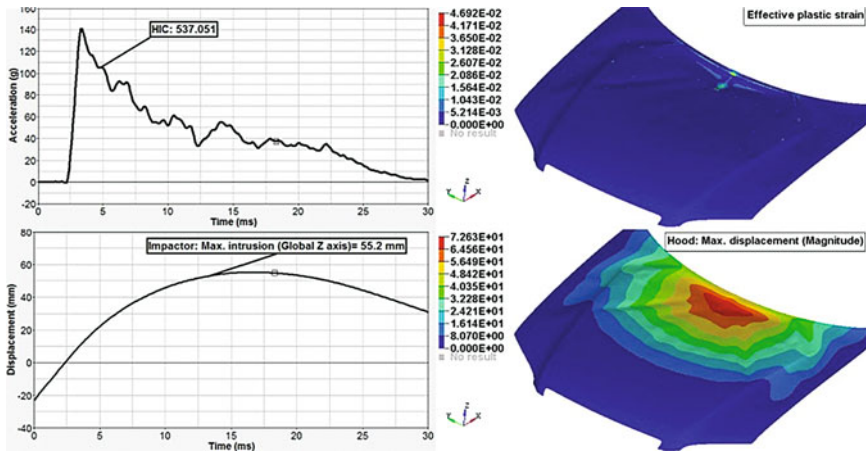


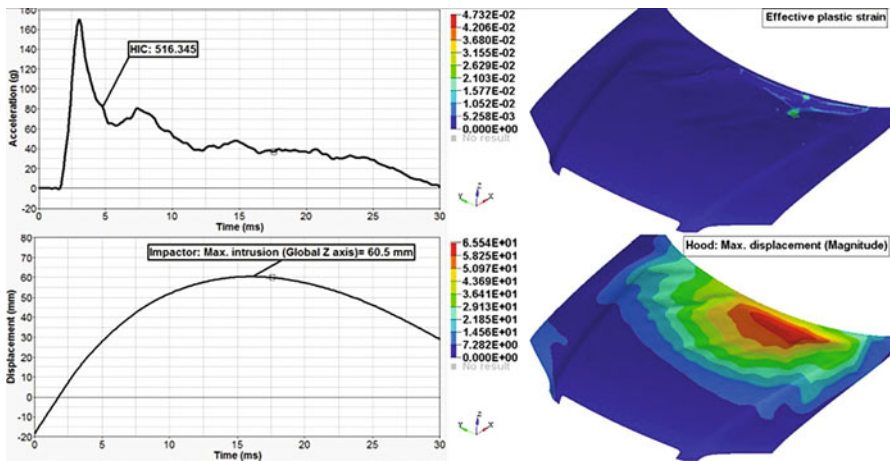
Fig. 7.55 HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point F

Thus, thinner inner and thicker outer hood thicknesses could be chosen within this range taking into consideration the torsional and bending stiffness, durability, noise and vibration performances.

Figure 7.58 shows the mean value of efficiency plotted against the combined inner and outer hood panel thickness. Although it could be concluded that similar efficiency is possible with various combinations of inner and outer hood thicknesses, efficiency varies with the combined thickness of inner and outer hood, as shown in Fig. 7.58. The optimal efficiency is achievable with a combined thickness of about



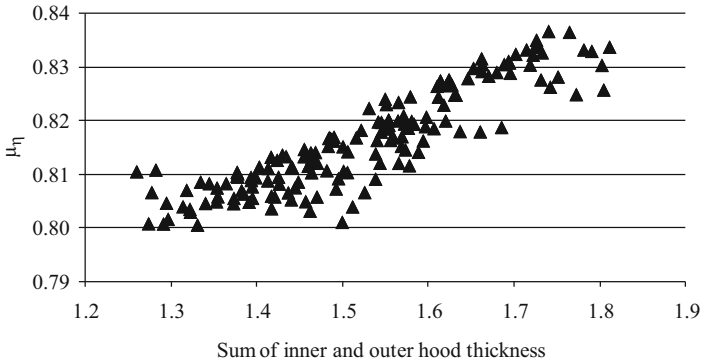
**Fig. 7.56** HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point G



**Fig. 7.57** HIC value, effective plastic strain, displacement of the impactor and hood with consideration to primary impact only for point H

1.76 mm. A trend line in this plot will show that the efficiency continues to increase with higher gauges but the HIC values will exceed the required limit.

The Eq. (7.5) that was utilised to derive the optimal deformation was derived from the theoretical optimal waveform, which is the WSTC where the resultant acceleration is infinite at zero millisecond and is impractical to obtain. In practice, it takes a minimum of 1–2 ms to reach the peak acceleration in a pedestrian head impact. Therefore, it could be concluded that in practice it is impossible to achieve 100 % efficiency with any hood configuration.



**Fig. 7.58** Efficiency and combined hood panel thickness relationship

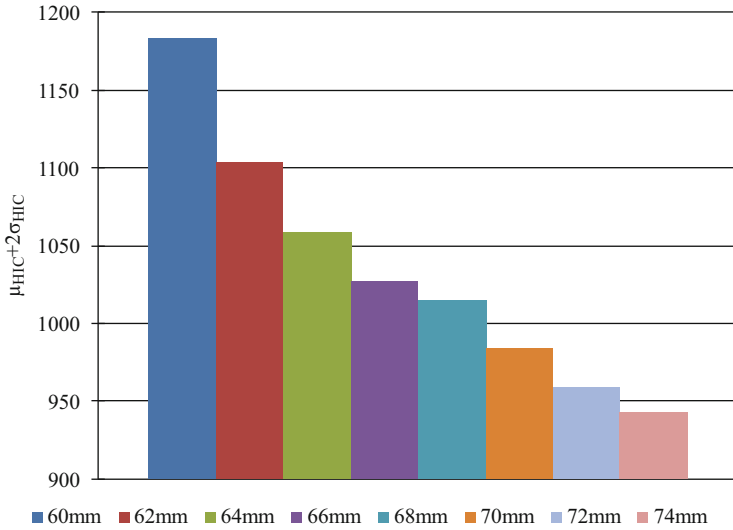
From the optimal acceleration curve, the calculated optimal deformation space required for HIC value that is less than 1000 is 50.5 mm. The results presented in Fig. 7.58 show that 84 % efficiency is possible with a steel multi-cone hood structure with a combined thickness of 1.76 mm for inner and outer hood panels. The deformation space required for a hood assembly with 84 % efficiency has been found to be approximately 60 mm. This is in agreement with suggestions given by Wu and Beaudet, who also derived equations to calculate the optimal deformation space as shown Eq. (7.5). Thus, a hood configuration with 84 % efficiency could be considered as the practical limit of efficiency as well as optimal hood configuration.

To check the effect of allowing 70 mm deformation space, impact has been simulated with different values of deformation space within the range from 60 to 74 mm. Same variables and qualifiers such as the mean value of HIC and its standard deviation have been used in the analysis. Comparison of  $\mu_{\text{HIC}} + 2\sigma_{\text{HIC}}$  for various deformation space as shown in Fig. 7.59 indicates that a minimum of about 69 mm deformation space is required to obtain robust and homogeneous head impact performance when considering both primary and secondary impacts.

Therefore, it can be stated that a multi-cone hood inner structure made of steel with approximately 1.6 mm of combined thickness of thinner inner and thicker outer hood panels requires a deformation space of about 69 mm to obtain robust and homogeneous head impact performance.

### 7.9.3 Optimisation of Panel Thickness for Aluminium Material

From the previous optimisation steps, it was identified that multi-cone structure for inner hood gave optimal head impact performance for steel and aluminium hood panels. However, it was also noticed that a 70 mm deformation space was not enough to obtain the HIC value less than 1000 when hood panels are made of



**Fig. 7.59** Influence of deformation space on HIC value (steel)

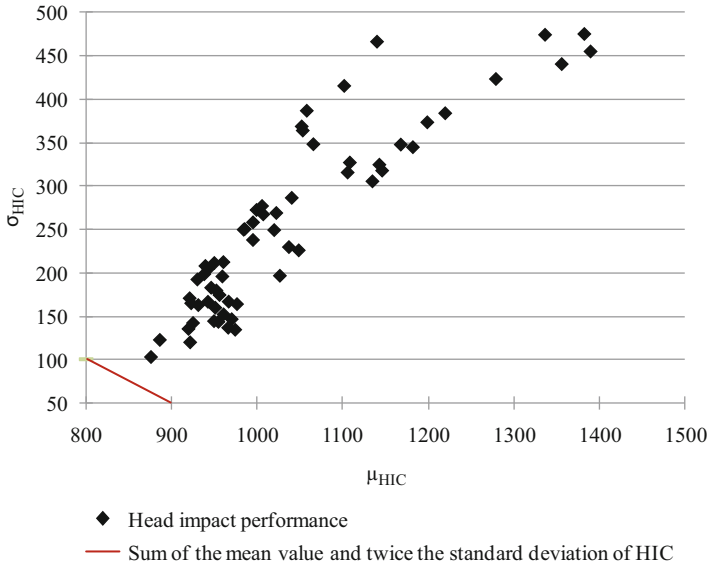
aluminium. Thus, another study was conducted with deformation space of 80 mm for aluminium hood panels with multi-cone inner hood structure.

Figure 7.60 shows the mean value of HIC plotted against its standard deviation. The head impact performance results from this study shows that 80 mm is enough to obtain the mean value of HIC less than 1000, although it is not enough to provide robust head impact performance.

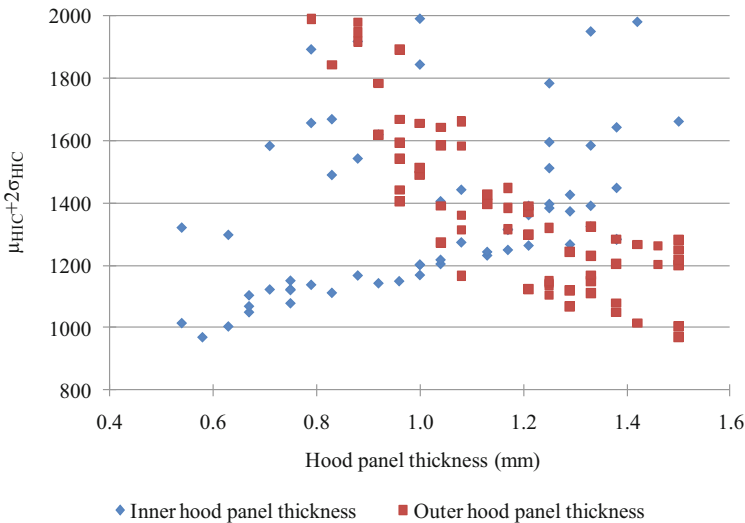
Therefore, analysis was carried out by increasing the deformation space from 80 to 85 mm. The plot of inner and outer hood panel thickness against the  $\mu_{HIC} + 2\sigma_{HIC}$  in Fig. 7.61 shows that it is possible to obtain robust head impact performance with HIC value less than the compliance threshold with a 85 mm deformation space. Inner hood thickness of 0.58 mm and outer hood thickness of 1.5 mm show optimal head impact performance with aluminium hood panels.

Figure 7.62 shows that the HIC value increases rapidly when the combined thickness of hood panels is below the optimal panel thickness, due to the intensity of the secondary impact. The HIC value slowly ramps up when the combined thickness is higher than the optimal thickness.

Figure 7.63 presents a close-up view of the optimal performance from Fig. 7.62. From this graph, it could be stated that combined hood panel thickness of 2.1 mm with aluminium material shows optimal head impact performance. Figure 7.64 presents the combined inner and outer hood panel thickness used in various vehicle models. As mentioned earlier, vehicle manufacturers balance the deformation space, mass, noise, vibration, durability and pedestrian protection requirements when selecting the inner and outer hood panel thicknesses.

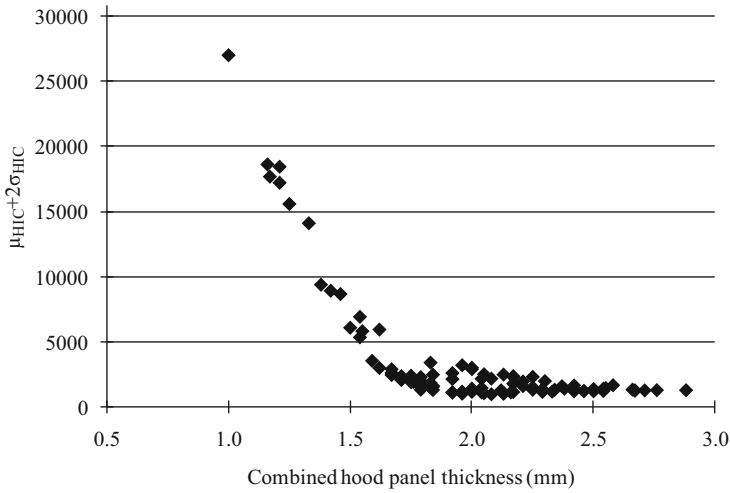


**Fig. 7.60** Head impact performance with 80 mm deformation space

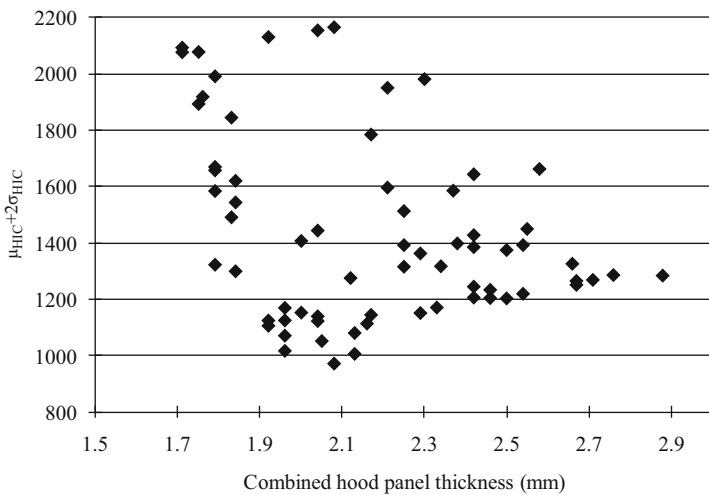


**Fig. 7.61** Influence of hood panel thickness on HIC value (aluminium)

To check the effect of allowing 85 mm deformation space, impact has been simulated with different values of deformation space within the range from 80 to 85 mm. Same variables and qualifiers such as the mean value of HIC and its standard deviation have been used in the analysis. Comparison of  $\mu_{HIC} + 2\sigma_{HIC}$



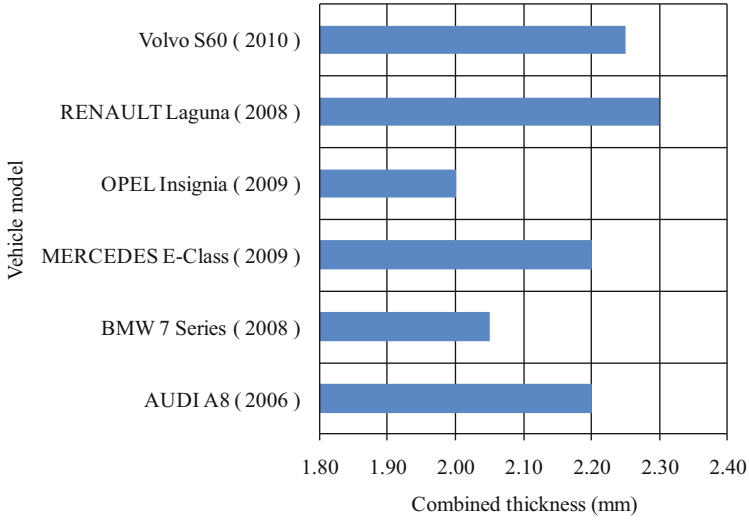
**Fig. 7.62** Influence of combined hood panel thickness on HIC value (aluminium)



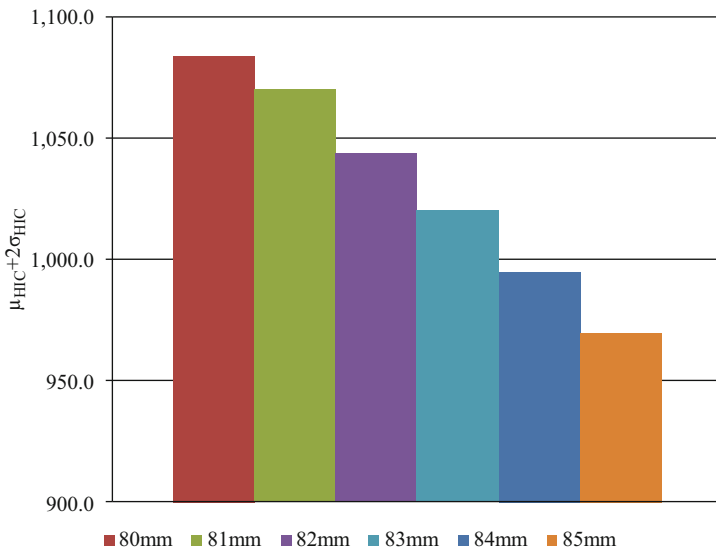
**Fig. 7.63** Combined hood panel thickness vs. the sum of the mean value of HIC and twice its standard deviation (aluminium)

for various deformation space as shown in Fig. 7.65 indicates that a minimum of about 84 mm deformation space is required to obtain robust and homogeneous head impact performance when considering both primary and secondary impacts.

Therefore, it can be stated that multi-cone inner hood structure made of aluminium with approximately 2.1 mm of combined thickness of thinner inner and thicker outer hood panels require a deformation space of about 84 mm to obtain robust and homogeneous head impact performance.



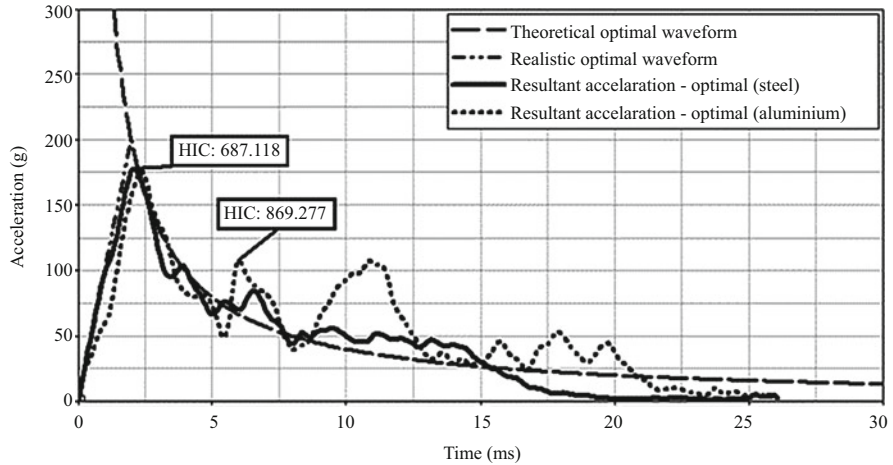
**Fig. 7.64** Combined thickness of aluminium outer and inner hood panels in various vehicles



**Fig. 7.65** Influence of deformation space on HIC value (aluminium)

### 7.9.4 Final Design Concept

The plots of the resultant acceleration against time curve, known as waveform, for steel and aluminium hood assemblies were compared to the theoretical and realistic optimal waveforms.



**Fig. 7.66** Waveform comparison-70 mm deformation space

The realistic optimal waveform is the combination of peak acceleration ramping up within 1 or 2 ms from the start of head impact followed by acceleration ramping down to a lower value that minimises the head injury. This waveform requires minimum deformation space for a given HIC value, in this instance 1000. The basis of the realistic optimal waveform is the theoretical optimal waveform in which the resultant acceleration reaches infinity at the beginning of the impact, but is impractical to obtain.

Figure 7.66 presents the head impact performance of a steel hood configuration with multi-cone inner hood structure with 0.5 mm for inner hood panel thickness and 1.1 mm for outer hood panel thickness. The aluminium hood assembly was analysed with a combined thickness of inner and outer panels of 2.7 mm. It shows that similar magnitudes of peak acceleration are possible with steel and aluminium by varying the panel thickness. Aluminium material however provides a higher HIC value compared to steel due to an increase in acceleration in the later part of the impact duration.

The peak acceleration of approximately 175 g within 2 ms is feasible due to the structural stiffness of the thicker outer hood panel, and the acceleration ramps down as the head contacts the thinner inner hood panel in the optimal hood configuration.

The plots in Fig. 7.67 show that the optimal hood configuration with aluminium material does not match the realistic optimal waveform, because it is not optimal for deformation space.



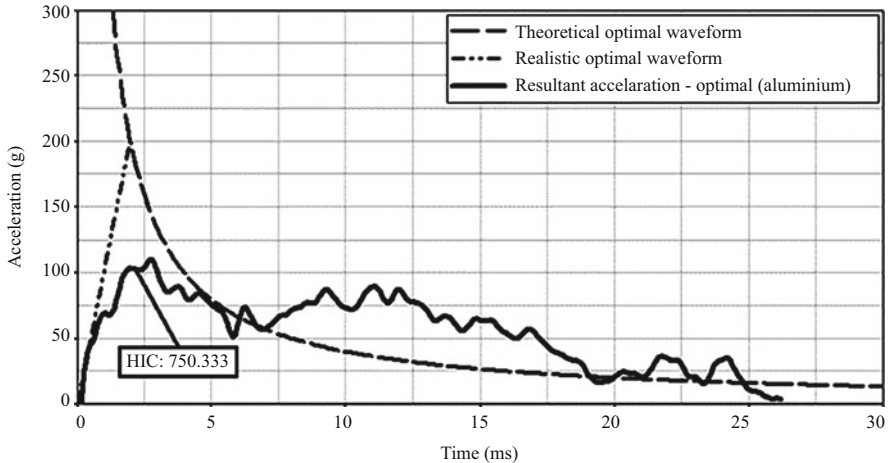


Fig. 7.67 Waveform comparison aluminium with 85 mm deformation space

## 7.10 Discussion

Since the magnitude of peak acceleration and rebound velocity have substantial influence on the HIC value of pedestrian when colliding with the front of a passenger car, the factors affecting the magnitude of peak acceleration were investigated by creating an ANOVA chart for peak acceleration.

In Fig. 7.68, HOODMATL is the hood panel material, OTRGAUGE is the thickness assigned for the outer hood panel, INRGAUGE is the thickness assigned for the inner hood panel and INRGEOM is the type of structure of the inner hood. The ANOVA chart showed that hood material had the highest effect on the magnitude of peak acceleration. Outer hood panel thickness was the next contributing factor, followed by inner hood panel thickness and inner hood structure respectively, as shown in Fig. 7.68. These findings support the above mentioned results, where increase in panel thickness was required for aluminium hood panels to match the peak acceleration of steel panels.

To understand the reason for the increase in HIC value of pedestrian when colliding with hood panels made of aluminium, the resultant acceleration was plotted against vertical intrusion, as shown in Fig. 7.69. Although both steel and aluminium result in similar values for peak acceleration and total vertical intrusion, the value of HIC varied between them.

This prompted a comparison between the average values of the rebound velocity of the hood for various impact positions for both materials as presented in Fig. 7.70. This shows that in general, the rebound velocity for aluminium is approximately 2.2 times higher than that for steel.

The integration of the resultant acceleration, which is velocity, is used in the calculation of HIC value. Thus, an increase in the rebound velocity increases the

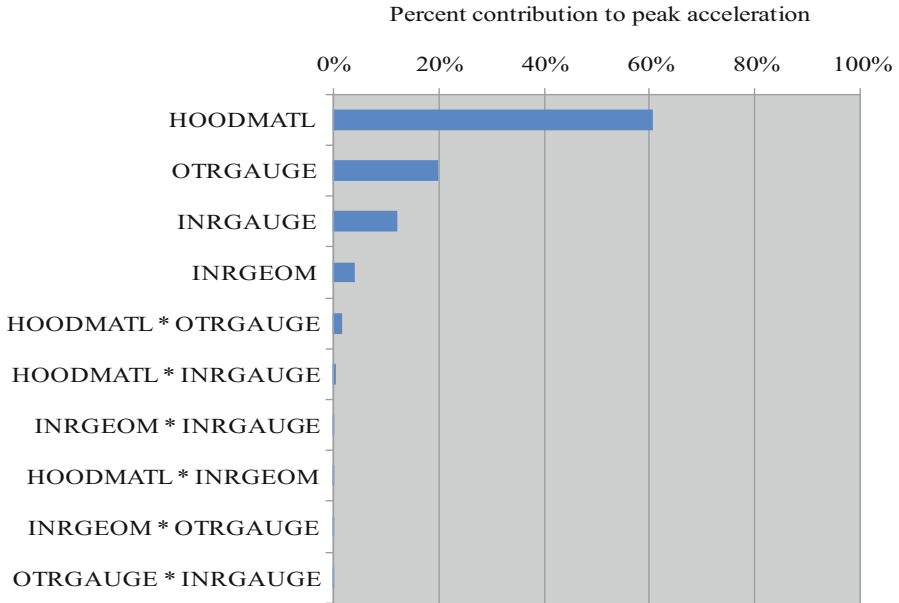


Fig. 7.68 ANOVA chart for peak acceleration

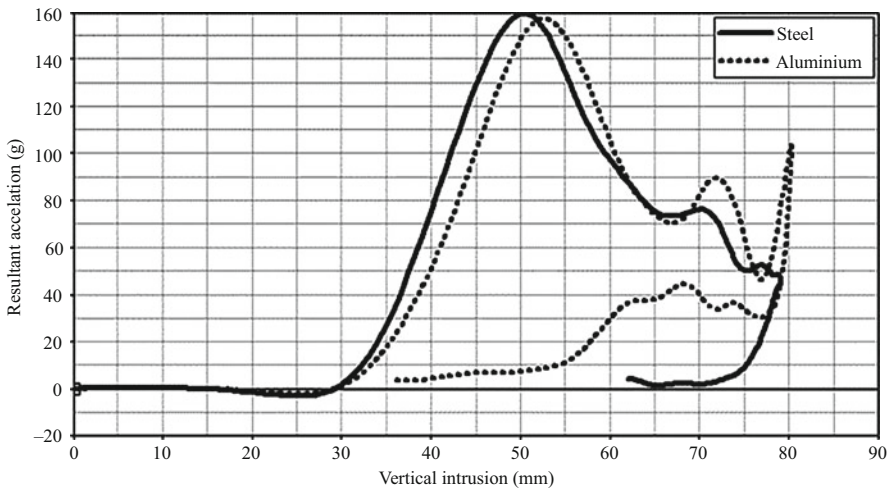
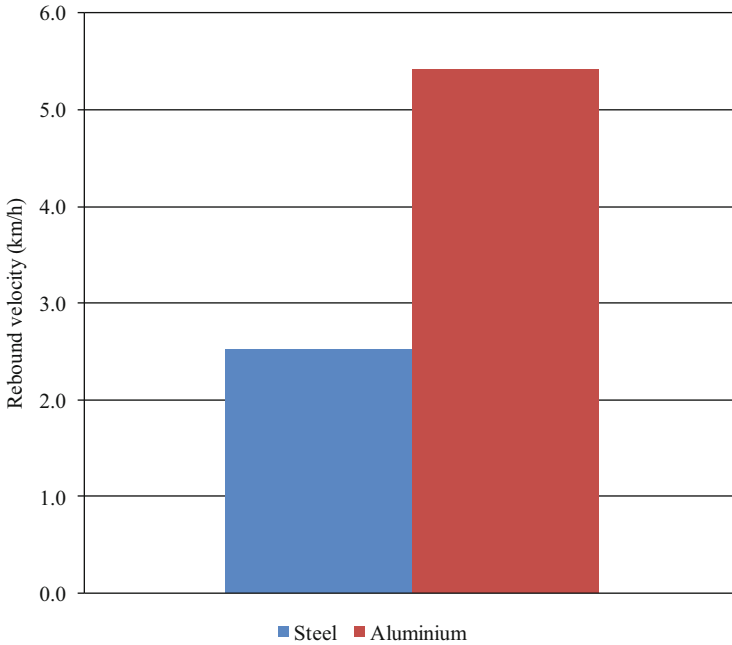


Fig. 7.69 Resultant acceleration against vertical intrusion

HIC value. Therefore, further exploration has been carried out to understand the exact reason for this increase in rebound velocity in aluminium hood panels.

To understand the basic concepts behind the difference in rebound velocities, a simple example has been studied theoretically. The strain energy for steel and



**Fig. 7.70** Comparison of rebound velocities

aluminium was calculated for a simplified beam with a rectangular cross section to demonstrate the difference in stored elastic energy. To simplify the calculation, the beam is assumed to be loaded to its elastic bending moment.

$$U_e = \int_0^L \frac{M^2}{2EI} dx \quad (7.7)$$

Where,  $U_e$  is the strain energy in the beam due to the load,  $M$  is the bending moment,  $E$  is the Young's modulus and  $I$  is the sectional modulus.

Equation (7.8) below shows the distribution of stress in a beam:

$$\sigma_y = \frac{My}{I} \quad (7.8)$$

Where,  $y$  is equal to half of the thickness of the beam, that is,  $t/2$ , assuming  $t$  is the thickness of the beam. Therefore, it can be written as follows:

$$M = \frac{\sigma_y bt^2}{6} \quad (7.9)$$

Where,  $\sigma_y$  is the stress at yield point and  $b$  is the width of the beam.

Substitution of Eq. (7.10) in Eq. (7.8) provides the following:

$$U_e = \frac{\sigma_y^2 b t l}{6E} \quad (7.10)$$

In which,  $l$  is the length of the beam.

If we assume same length and width of the beam for aluminium and steel, the calculation of the ratio of elastic strain energy between aluminium and steel is as follows:

$$\frac{U_e \text{ of aluminium}}{U_e \text{ of steel}} = \frac{\sigma_{y(al)}^2 t_{al}}{E_{al}} \times \frac{E_{st}}{\sigma_{y(st)}^2 t_{st}} \quad (7.11)$$

Where,  $t_{al}$  is the thickness value assigned for aluminium, in this case, 2.7 mm;  $t_{st}$  is the thickness value assigned for steel, in this case 1.6 mm,  $E_{al}$  is the Young's modulus value for aluminium that is 70,000 MPa and  $E_{st}$  is the young's modulus value for steel that is 200,000 MPa.

The yield strength of the aluminium material used in this example is assumed to have  $\sigma_y(al)$  while the yield strength of the steel material is assumed to be  $\sigma_y(st)$ . It was assumed in this work that  $\sigma_y(al) = \sigma_y(st)$  as the aluminium material selected for hood panels has yield strength close to that of steel grade usually used to manufacture hood panels.

When these values were substituted in Eq. (7.11), the ratio of the elastic strain energy of steel to aluminium was 1:4.8.

Therefore, it could be stated that when loaded to same level of stress, hood panels made with aluminium store 4.8 more times the elastic strain energy stored by steel. This causes an increase in rebound velocity by using the equation shown below:

$$U_e = \frac{mv^2}{2} \quad (7.12)$$

Where,  $m$  is the mass and  $v$  is the velocity of the head.

So, 4.8 times increase in elastic strain energy leads to 2.2 times increase in rebound velocity. This coincides with the average value of rebound velocities comparison, shown in Fig. 7.70.

In Eq. (7.10), the stored elastic strain energy is proportional to the thickness. Thus, the value of rebound velocity could be reduced by reducing the thickness. However, this change will increase the deformation space due to the reduction in structural stiffness. Since the structural stiffness of aluminium is lower than steel, reduction in panel thickness to reduce the rebound velocity will increase the deformation of hood made of aluminium. Therefore, it could be stated that aluminium requires more deformation space in comparison to steel to meet pedestrian head injury requirements.

Steel hood panels exhibit lower rebound velocity, which matches the realistic optimal waveform. Therefore, the preferred choice of material for hood panels is

steel if the main objective is to reduce the deformation space. The benefits of minimising the deformation space are significant. They include the freedom of styling, improved aerodynamics, and hence improvements in vehicle stability and fuel economy. If the main objective is to reduce mass, then the preferred choice for hood panels is aluminium; however, this will require more deformation space.

The area for aluminium and steel hood panels were the same in this study. Thus, when 2.7 mm for aluminium and 1.6 mm for steel was substituted for panel thickness value along with their respective density values, the result showed 42 % less mass for aluminium hood assembly compared to steel hood assembly.

## 7.11 Conclusion

The main objective of this work was to develop a methodology for optimising the hood assembly for minimal deformation with robust and homogeneous impact behaviour while meeting pedestrian head injury requirements. In practice, the design of the outer and inner hood panels must meet all general design requirements in addition to pedestrian head impact performance. Achievement of design criteria for homogeneous impact performance for the pedestrian head could aid in meeting the general design requirements.

In this research, the important design parameters were identified and a design of experiments matrix was created. Then, three-dimensional geometric models and FE models were created to the design parameters defined in the design of experiments matrix. The results from the numerical experiments were utilised to create the response surfaces for the mean value of efficiency ( $\mu\eta$ ), the mean value of HIC ( $\mu\text{HIC}$ ) and the standard deviation of HIC ( $\sigma\text{HIC}$ ). The Analysis of Variance (ANOVA) technique was used to quantify the contribution of design parameters in determining HIC value and efficiency. Then these data were filtered with the condition that the sum of the mean value of HIC and twice its standard deviation ( $\mu\text{HIC} + 2\sigma\text{HIC}$ ) should be less than or equal to 1000. This was done to match the head injury value specified by the GTR-9 and ANCAP and to produce 95 % confidence in the results of this optimisation, assuming normal distribution.

The filtered data was compared using  $\sigma\text{HIC}$  as a qualifier for homogeneous performance and  $\mu\eta$  as a qualifier for minimal deformation of outer and inner hood panels. This resulted in the identification of the multi-cone pattern for hood inner structure as the best structure for both steel and aluminium compared to other structures. Further refinement indicated that steel was the preferred choice of material for the hood assembly to obtain robust and homogeneous head impact performance with minimal deformation compared to other materials.

Findings from this study showed that it is possible to obtain HIC values less than the threshold value, while achieving similar efficiencies for various combinations of thinner inner and thicker outer hood panels. However, efficiency variation was noticed when changing the value of the combined thickness of inner and outer hood panels.

The results from a previous study showed that it is possible to obtain  $\mu$ HIC less than 1000 with 60 mm deformation space with a multi-cone structure made of steel of about 1.8 mm of combined inner and outer hood thickness, but this is not enough to obtain robust and homogeneous behaviour less than the HIC threshold value when secondary impact is considered.

Therefore, impact has been simulated with different values of deformation space within the range from 60 to 74 mm. Same variables and qualifiers such as the mean value of HIC and its standard deviation have been utilised in the analysis. The results indicated that a minimum of about 69 mm deformation space was required to obtain robust and homogeneous head impact performance when considering both primary and secondary impact.

In conclusion, multi-cone inner hood structure made of steel for inner and outer hood panels is the preferred choice for hood assembly when minimal deformation is the main requirement for the vehicle design team.

Another study was conducted to quantify the minimum deformation space required for the aluminium hood system. The multi-cone structure for the inner hood required the least amount of deformation space in comparison to other geometries. Impact simulation with 80 mm of deformation space showed that it is possible to obtain  $\mu$ HIC less than 1000, but this is not enough to obtain robust and homogeneous behaviour below the HIC threshold value when secondary impact is considered. Thus, impact has been simulated with different values of deformation space within the range from 80 to 85 mm. The same variables and qualifiers such as the mean value of HIC and its standard deviation have been utilised in this analysis. The results indicated that a minimum of about 84 mm deformation space was required to obtain robust and homogeneous head impact performance when considering both primary and secondary impact.

In conclusion, a robust and homogeneous pedestrian head impact performance is possible from the following design parameters for aluminium hood panels: multi-cone structure for inner hood, and approximately 2.1 mm combined thickness with thinner inner and thicker outer hood combination.

## 7.12 Recommendations for Further Work

Vehicle design for pedestrian protection requires improvements in many components in the vehicle front-end such as the lower bumper system, upper bumper system, interfaces of hood assembly to the vehicle and hood assembly. To optimise the vehicle front-ends to improve pedestrian protection performance while meeting all the other requirements of vehicle design is a very large scope of work. In this research, the development of optimisation methodology of large-size hoods for robust and homogeneous performance with minimal deformation was completed while meeting pedestrian performance requirements for head impact.

The analysis in this research was carried out with a 3.5 kg child head impactor fired at 50° impact angle from the horizontal plane. In summary, the hood

configuration was optimised for that given mass and angle of impact, and if an area required both a child head impactor (3.5 kg) and an adult head impactor (4.5 kg), compromises would need to be made to accommodate various masses and angles of impact. Further studies, however may be required to optimise medium and smaller size hoods. Similarly, work may be required to optimise components such as hood hinges, A-pillar, base of windscreen, wiper assembly, fender brackets, under hood components, glue between inner and outer hood panels and hood latch.

Although many key parameters were identified and optimised, further work would be required to maximise the head impact performance in pedestrian protection offered by vehicles without compromising all other requirements such as occupant safety, quality, fuel economy, durability and low repair costs. However, a foundation had been laid for continued research in improving the pedestrian protection offered by vehicle front-ends in a pedestrian collision with a passenger car.

## Key Symbols

A	Acceleration (g)
T	Duration of acceleration (s)
a(t)	Resultant acceleration (g)
$t_1$	Time at the beginning of the impact (s)
$t_2$	Time at the end of the impact (s)
$\eta$	Efficiency
$A_1$	Area enclosed by acceleration–displacement curve
$A_2$	Area of the shape that encloses acceleration–displacement curve
d	Maximum intrusion (m)
$v_0$	Initial impact velocity (m/s)
$\mu_\eta$	Mean value of efficiency
$\mu_{HIC}$	Mean value of HIC
$\sigma_{HIC}$	Standard deviation of HIC
$U_e$	Strain energy (J)
M	Bending moment (Nm)
E	Young's modulus (MPa)
I	Sectional modulus ( $\text{mm}^3$ )
t	Thickness of the beam (mm)
b	Width of the beam (mm)
l	Length of the beam (mm)
$\sigma_y$	Stress at yield point (MPa)
$t_{al}$	Thickness value assigned for aluminium (mm)
$t_{st}$	Thickness value assigned for steel (mm)
$E_{al}$	Young's modulus value for aluminium (MPa)
$E_{st}$	Young's modulus value for steel (MPa)
$\sigma_{y(al)}$	Stress at yield point for aluminium (MPa)

$\sigma_{y(st)}$	Stress at yield point for steel (MPa)
m	Mass of the head impactor (kg)
v	Velocity of the head impactor (m/s)

## References

- A2Mac1 (2010) Automotive benchmarking, viewed 22 March 2010. <http://a2mac1.com/home>
- Bohling G (2005) Kriging, University of Kansas, viewed 28 July 2011. <http://people.ku.edu/~gbohling/cpe940/Kriging.pdf>
- Centre of automotive safety research (2007) Head impact test machine, The University of Adelaide, viewed 17 December 2009. <http://casr.adelaide.edu.au/impact.html>
- Elert G (2001) Number of cars, Glenn Elert, viewed 16 June 2010. <http://hypertextbook.com/facts/2001/MarinaStasenko.shtml>
- European Commission (2010) Road safety, viewed 13 February 2010. [http://ec.europa.eu/transport/road\\_safety/specialist/statistics/trends](http://ec.europa.eu/transport/road_safety/specialist/statistics/trends)
- Gelman A (2004) Analysis of variance-why it is more important than ever, Columbia University, viewed 28 July 2011. <http://www.stat.columbia.edu/~gelman/research/published/banova7.pdf>
- Hardy BJ, Lawrence GJL, Knight IM, Simmons ICP, Carroll JA, Coley G, Bartlett RS (2007) A study of possible future developments of methods to protect pedestrians and other vulnerable road users, UPR/VE/061/07, TRL Limited
- Honda Worldwide (2008) Honda to begin using third generation pedestrian dummy in crash tests to enable evaluation of injuries to the lower back and upper leg, viewed 17 December 2009. <http://world.honda.com/news/2008/4080918POLAR-III/>
- Kam CY, Kerrigan J, Meissner M, Drinkwater C, Murphy D, Bolton J, Arregui C, Kendall R, Ivarsson J, Crandall J, Deng B, Wang JT, Kerkeling C, Hahn W (2005) Design of a full-scale impact system for analysis of vehicle pedestrian collisions, SAE Technical Paper 2005-01-1875, SAE World Congress, Detroit, Michigan, SAE International, Warrendale, Pennsylvania, USA
- Krishnamoorthy R, Takla M, Subic A, Scott D (2013) Design optimisation of passenger car hood panels for improved pedestrian protection. *Adv Mater Res* 633:62–76
- Lim GG, Chou CC, Patel RN, Shahab SA, Patel PJ (1995) Estimating the minimum space to meet federal interior head impact requirement, SAE Technical Paper 950333, SAE International Congress and Exposition, Detroit, Michigan, SAE International, Warrendale, Pennsylvania, USA
- Mizuno Y (2003) Summary of IHRA pedestrian safety WG activities (2003) – proposed test methods to evaluate pedestrian protection offered by passenger cars
- Narasimhan B (1996) The normal distribution, Stanford University, viewed 28 July 2011. <http://www-stat.stanford.edu/~naras/jsm/NormalDensity/NormalDensity.html>
- National highway traffic safety administration (2009) viewed 10 October 2011. <http://www-fars.nhtsa.dot.gov/People/PeoplePedestrians.aspx>
- Ryan GA, McLean AJ (1965) Pedestrian survival. Safe car guide 2009, Australian New Car Assessment Program, viewed 25 November 2009, <http://www.safecarguide.com/exp/ancap/ancap.htm>
- Synaptic analysis consulting group inc viewed 23 February (2011). [http://www.synapacg.com/casestudies\\_pedestrian.html](http://www.synapacg.com/casestudies_pedestrian.html)
- Transport Research Laboratory (2012) Pedestrian Safety, Transport Research Laboratory, viewed 15 October 2012. [http://www.trl.co.uk/research\\_development/improving\\_safety/user\\_safety/pedestrian\\_safety.htm](http://www.trl.co.uk/research_development/improving_safety/user_safety/pedestrian_safety.htm)



- United Nations Economic Commission for Europe (1998) Global technical regulation No. 9 (Pedestrian safety), United Nations viewed 12 July 2011. <http://live.unece.org/trans/main/wp29/wp29wgs/wp29gen/wp29registry/gtr9.html>
- World Health Organization (2010) Global status report on road safety: time for action, World Health Organization, viewed 12 August 2010. [http://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2009](http://www.who.int/violence_injury_prevention/road_safety_status/2009)
- World Health Organization (2011) Violence and injury prevention and disability, World Health Organization, viewed 10 October 2011. [http://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/multimedia/pictures/en/index.html](http://www.who.int/violence_injury_prevention/road_safety_status/multimedia/pictures/en/index.html)
- Worldometers (2010) World statistics updated in real time, Worldometers, viewed 15 July 2011. <http://www.worldometers.info/cars/>
- Wu J, Beaudet B (2007) Optimization of head impact waveform to minimize HIC. World Congress

# Chapter 8

## Nonlinear Finite Element and Post-Buckling of Large Diameter Thin Walled Tubes

Toh Yen Pang, Dale Waterson, Rees Veltjens, and Tristan Garcia

**Keywords** Finite element analysis • Thin wall structure • Axial crush testing • Post-buckling • Non-linear

### 8.1 Introduction

The post-buckling and crushing behaviour of structures is of great importance in many engineering disciplines. In particular, post-buckling behaviour is of interest when considering highly dynamic collision events; where both the integrity of a structure may be compromised and when predicting response to assess safety aspects (such as a motor vehicle collision). In these cases, the way in which a structural member fails can strongly influence the manner in which forces are distributed throughout the remainder of the structure and transferred to passengers involved in an impact (Alghamdi 2001). Virtually all modes of modern day transportation (including sea, land and air) consist of thin walled structural components. The benefits of these structures are clear; with the reduction in weight there is the potential to dramatically improve the performance and efficiency of such vehicles. However, with the increase in performance of each of these vehicles, as well as the increase in societal awareness with regards to safety and design standards, there is an increasingly large requirement for these structures to have the capacity to dissipate a large proportion of impact energy to reduce the trauma imposed on occupants during a collision.

---

T.Y. Pang (✉) • D. Waterson • R. Veltjens • T. Garcia  
School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University,  
Bundoora East Campus, Bundoora, Melbourne, VIC, Australia  
e-mail: [tohyen.pang@rmit.edu.au](mailto:tohyen.pang@rmit.edu.au); [dale.waterson@woodgroup.com](mailto:dale.waterson@woodgroup.com); [rees.veltjens@gmail.com](mailto:rees.veltjens@gmail.com);  
[tristan.garcia@hotmail.com](mailto:tristan.garcia@hotmail.com)

In recognition of the role that thin walled structures play in collision safety, there has been a plethora of research conducted with regard to this particular field. This includes (but is certainly not limited to) the manner in which thin wall structures may collapse based on their geometry (Song et al. 2012) to how thin walled structures may be utilised as impact energy absorbers (Abramowicz 2003). Ultimately, the survivability of motor vehicle collisions is not only a function of the amount of impact energy present, but also the manner in which that energy is dissipated. This impact energy needs to be distributed in a controlled manner such that the deceleration of the vehicle improves the chances of occupant survivability (Abramowicz 2003). In addition to this, there is the need to verify the functionality of designs in collision incidents, including both experimental and computational analysis of structures. In order to enhance the collision behaviour of thin walled structures, analysis into the collapse behaviour of individual thin walled structural components needs to be undertaken.

This study aims to investigate the crush behaviour of a large diameter, thin walled cylindrical tube under a compressive axial load. Experimental analysis will be conducted to determine the cylinder's modes of failure under a constant rate of compression and gain data relating to the forces required to crush the specimen. This will then be followed by a numerical finite element analysis, which will be used to compare and verify the ability to correctly and adequately predict the results of such an experiment.

## 8.2 Experimental Test

### 8.2.1 *Specimens and Test Machines*

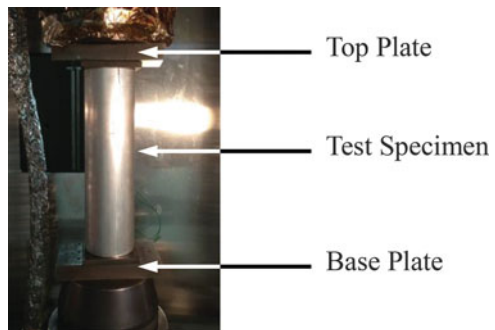
The specimens used during the experiment consisted of two thin walled aluminium cylindrical tubes (shown in Fig. 8.1), each with the length of 300 mm, outside diameter of 100 mm and wall thickness of 3.0 mm.

The experiment involved crushing each specimen in a vertical press which was capable of providing a constant rate of compression. Each thin walled cylindrical tube was placed between the top and base plates of the press. The test specimen was then axially compressed by 150 mm. The force required to maintain the constant rate of compression was measured using load cells and recorded by a computer. The setup of the load press is shown in Fig. 8.2. The tests were conducted at a compression rate of 20 mm/min.

**Fig. 8.1** Thin walled cylindrical tube used for the experiment



**Fig. 8.2** Experimental setup of test specimen in the load press



### **8.2.2 Results of Experimental Test**

During the compression both specimens exhibited similar modes of axisymmetric local buckling failure that involved the tube folding outwards and then back inward as shown in Fig. 8.3. The tube rippled and folded neatly across the centreline of the wall section.

Two out of the three tests displayed near identical buckling modes throughout the full range of compression. One other specimen exhibited a similar failure mode initially; however a non-symmetric failure mode became dominant for the remainder of the range.

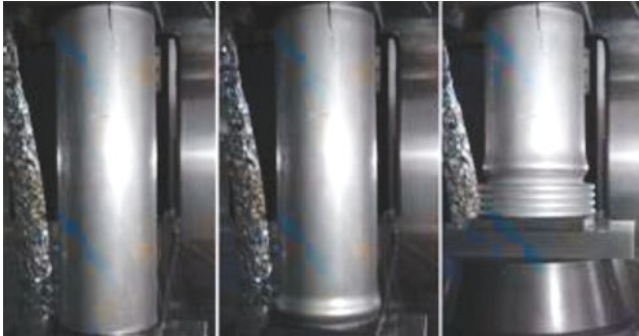


Fig. 8.3 Images of experiment demonstrating mode of failure

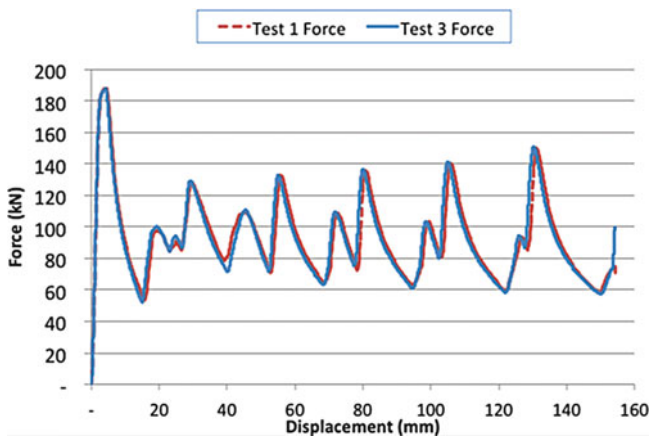


Fig. 8.4 Force vs. displacement graph of experiment

The force–displacement plots for the tests with axisymmetric failure modes are shown together for comparison in Fig. 8.4. There is an extremely high level of consistency in the experimental results. Both tests demonstrate a nearly identical pattern in terms of the force required to compress the tube throughout the range of the experiment. Each of the peaks corresponds neatly with the point at which each tube begins to buckle, with the peak stress of approximately 190 kN.

## 8.3 Numerical Calibration

### 8.3.1 Finite Element (FE) Models

The FE code Abaqus/Explicit v6.11 was adopted to simulate the crushing process of the cylindrical tubes. The cylindrical tube was meshed using four-node doubly curved thin or thick shell S4R elements with a shell thickness of 3 mm.

The boundary conditions of the simulation were setup according to the experimental buckling tests. The base plate was initially contacting the bottom edge of the tube, and then moved axially upward to allow the top edge of the tube to contact the top plate and then crush the tube. The upper and lower plates were modelled as analytical rigid bodies. Explicit procedure was used with crushing velocity of 3.1 m/s.

### 8.3.2 Material Model

The shell material was Aluminium alloy 6060 T5, which is a ductile homogenous metal and exhibits a very linear stress–strain relationship, i.e. elastic region, up to a yield point. The mechanical properties for the elastic region are as follows: density  $\rho = 270 \times 10^3 \text{ kg/m}^3$ , Young's modulus  $E = 70 \text{ GPa}$ , Poisson's ratio  $\nu = 0.3$ , yield stress  $\sigma_0 = 130 \text{ MPa}$ .

After the yield point is reached, the metal starts to deform plastically, i.e. large deformations that permanently deform the structure. The plastic stress–strain relationship was derived using the Ramberg–Osgood equation given by (Wei and Elgindi 2013):

$$\varepsilon_{plastic} = \frac{\sigma}{E} + \frac{\sigma_0}{B} \left( \frac{\sigma}{\sigma_0} - 1 \right)^H \quad (8.1)$$

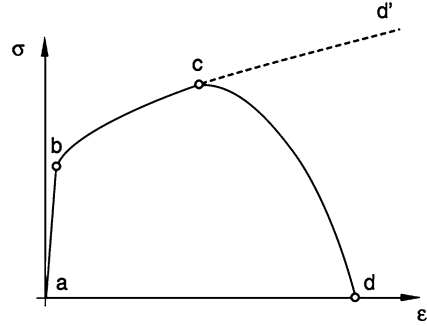
where B and H are the parameters given as  $B = 7 \text{ GPa}$ ,  $H = 3.0$ .

The elastic and plastic data was tabulated and input into the ABAQUS material model. A maximum plastic strain of 20 % was used to capture the material failure before the maximum strains could be reached. In order to realistically represent the experimental conditions, damage criteria was also included in the material model.

### 8.3.3 Model for Ductile and Shear Failure Initiation

Fracture damage can occur via two main mechanisms; due to the nucleation, growth and coalescence of voids, and shear fracture due to shear band localization. For the analysis models using damage, there were three criteria of relevance: ductility,

**Fig. 8.5** Stress–strain behaviour of material damage where softening occurs, as well as degradation of material



shear and Müschenborn–Sonne forming limit diagram (MSFLD) criteria (Dassault Systèmes 2011a). Inputs for these criteria determine the point at which softening occurs, as well as the rate of softening as the material is strained. This is illustrated in Fig. 8.5, where beyond point *c* the capability of the material to carry load is reduced and follows damage mechanism *c–d* until rupture occurs.

### 8.3.4 Ductile Damage and Failure Criterion

Abaqus offers a variety of choices of damage initiation criteria for ductile metals. The model assumes that the equivalent plastic strain at the onset of damage,  $\bar{\varepsilon}_D^{pl}$ , is a function of stress triaxiality and strain rate:

$$\bar{\varepsilon}_D^{pl} \left( \eta, \dot{\varepsilon}^{pl} \right) \quad (8.2)$$

where  $\eta = -p/q$  is the stress triaxiality,  $p$  is the pressure stress,  $q$  is the Mises equivalent stress, and  $\dot{\varepsilon}^{pl}$  is the equivalent plastic strain rate (Dassault Systèmes 2011b).

The ductility failure initiation criterion that is represented by a correlation to the equivalent plastic strain rate (Dassault Systèmes 2011a; Hooputra et al. 2004) is given by:

$$\bar{\varepsilon}_D^{pl} \left( \eta, \dot{\varepsilon}^{pl} \right) = \frac{\varepsilon_T^+ \sinh [k_0 (\eta^- - \eta)] + \varepsilon_T^- \sinh [k_0 (\eta - \eta^+)]}{\sinh [k_0 (\eta^- - \eta^+)]} \quad (8.3)$$

where  $\varepsilon_T^+$  and  $\varepsilon_T^-$  correspond to the equivalent plastic strain at ductile damage initiation for equibiaxial tensile and equibiaxial compressive deformation, respectively.  $\eta^+$  and  $\eta^-$  correspond to the stress triaxiality for isotropic materials in equibiaxial tensile deformation state and compressive deformation state, respectively.  $\eta$  is the ratio of the equivalent mean stress to the Mises equivalent stress.  $k_0$  is the coefficient.

### 8.3.5 Shear Failure Criterion

For shear failure, it is assumed that the equivalent strain at failure  $\bar{\varepsilon}_S^{pl}$  is a function of the shear stress ratio,  $\theta_S$ , given by

$$\theta_S = \frac{1 - k_s \eta}{\phi} \quad (8.4)$$

where  $k_s$  is a material parameter and  $\phi = \tau_{max}/\sigma_{eq}$ . The shear failure initiation criterion that is represented by a correlation to the equivalent plastic strain is given by:

$$\bar{\varepsilon}_S^{pl}(\theta_S, \dot{\bar{\varepsilon}}^{pl}) = \frac{\varepsilon_S^+ \sinh[f(\theta_S - \theta_S^-)] + \varepsilon_S^- \sinh[f(\theta_S^+ - \theta_S)]}{\sinh[f(\theta_S^+ - \theta_S^-)]} \quad (8.5)$$

where  $\theta_S^+$  and  $\theta_S^-$  correspond to the values  $\theta_s$ ,  $f = \text{constant}$ , and  $\varepsilon_S^+$  and  $\varepsilon_S^-$  correspond to the equivalent plastic strain at shear damage initiation for equibiaxial tensile and equibiaxial compressive deformation, respectively.

For this study, all the ductile and shear failure parameter data, in Eqs. (8.3) and (8.5), for aluminium was taken from the ABAQUS example manual (Dassault Systèmes 2011a). These parameters were manipulated to suit aluminium 6060 T5 and are covered further in the discussion section.

### 8.3.6 Damage Evolution and Element Removal

In Abaqus, damage evolution occurs once the damage initiation criteria are satisfied. Plastic displacement-based linear damage evolution law is used for all damage initiation criteria. The value of the plastic displacement at which the damage variable reaches 1 is taken as 0.1. The default maximum degradation rule is used, and the elements are removed from the mesh once the maximum degradation has occurred (Dassault Systèmes 2011a).

### 8.3.7 Solution Procedure Using ABAQUS/Explicit

The ABAQUS/Explicit performs dynamic analysis, using a Lagrangian formulation and integrating the equations of motion in time explicitly by means of central differences. In the dynamic analysis, the basic equations of motion of a cylindrical shell are expressed as:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{c}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{F}^{\text{ext}}(\mathbf{u}(t)) \quad (8.6)$$



where  $[\mathbf{M}]$  is the mass matrix,  $[\mathbf{c}]$  the damping matrix,  $[\mathbf{K}]$  the stiffness matrix including non-linearity,  $u$  the nodal displacement vector,  $\dot{u}$  the nodal velocity vector,  $\ddot{u}$  the nodal acceleration vector and  $[\mathbf{F}^{\text{ext}}]$  the external load vector (Bisagni 2005).

Following the standard equation of motion, we can write down the spatially discretised finite element equations of motion for the elastodynamic contact problem, such as buckling behaviour of the cylindrical shells under axial compression loads, as follows (Burlayenko and Sadowski 2014):

$$\mathbf{M}\ddot{u}(t) + \mathbf{c}\dot{u}(t) + \mathbf{K}u(t) = \mathbf{F}^{\text{ext}}(u(t)) - \mathbf{F}^{\text{cont}}(u(t)) \quad (8.7)$$

Where:  $\mathbf{F}^{\text{cont}}$  is the contact force vector.

When we define a contact/impact problem, which subjects to given initial conditions  $u_0 = u_0$  and  $\dot{u}_0 = v_0$  and a given load history  $[\mathbf{F}^{\text{ext}}(u(t))]$ , the solution of Eq. (8.7) is to hold for all time interval  $t \in [0, T]$  (Burlayenko and Sadowski 2014; Heinstejn et al. 2000; Strömberg 2005). The time interval is first partitioned into  $i$  subintervals  $[t_i, t_{i+1}]$  such that  $t_i < t_{i+1}$ , and  $t_1 = 0$  and  $t_{i+1} = T$ .

Let the time increment be  $\Delta t = t_{i+1} - t_i$ , and accelerations, velocities and displacements referring to this time increment are denoted by  $\ddot{u}_{i+1}$ ,  $\dot{u}_{i+1}$  and  $u_{i+1}$ , respectively. Hence, the discretized system of Eq. (8.7) at a certain time  $t_{i+1}$  can be written as:

$$\mathbf{M}\ddot{u}_{i+1} + \mathbf{c}\dot{u}_{i+1} + \mathbf{K}u_{i+1} = \mathbf{F}_{i+1}^{\text{ext}} - \mathbf{F}_{i+1}^{\text{cont}} \quad (8.8)$$

### 8.3.8 Central Finite Difference

In FE analysis of time-dependent problems, the equation of motion is mostly solved by using an explicit central difference integration rule (Heinstejn et al. 2000; Sun et al. 2000; Dassault Systèmes 2011c):

$$u_{i+1} = u_i + \Delta t_{i+1} \dot{u}_{i+1/2} \quad (8.9)$$

$$\dot{u}_{i+1/2} = \dot{u}_{i-1/2} + \frac{1}{2} (\Delta t_{i+1} + \Delta t_i) \ddot{u}_i \quad (8.10)$$

The acceleration at  $t_1 = 0$  from Eq. (8.7):

$$\ddot{u}_i = \mathbf{M}^{-1} (\mathbf{F}_i^{\text{ext}} - \mathbf{c}\dot{u}_i - \mathbf{K}u_i - \mathbf{F}_i^{\text{cont}}) \quad (8.11)$$

Where:  $i, i - 1/2$  and  $i + 1/2$  refer to the increment number and mid-increment numbers.

The explicit procedure is based on the implementation of an explicit integration rule along with the use of diagonal element mass matrices, hence no couple equations need be solved as the kinematic state is advanced using known values of  $\dot{u}_{i-1/2}$  and  $u_i$  from the previous increment.

### 8.3.9 Contact Constraint

The tube was compressed freely and axially and there were a number of surfaces which were identified as likely to contact during the simulation;

- The upper and lower edges of the tube, which would contact with the load press plates.
- The side walls of the tube and the load plate, which would occur as the thin walled tube buckled and folded over during the first failure.
- The self-contact of the tube, which would occur as each ripple of the wall folded back onto itself.

Penalty friction formulation, with the Coulomb friction coefficient, was adopted in the simulation to enforce the contact constraint between the master surface and the slave surface.

Whenever contact takes place, the kinematic variables at the contact interface between the two bodies can be split into (1) the normal,  $\mathbf{F}_N^c$ , and (2) the tangential,  $\mathbf{F}_T^c$ , directions for a single constraint to the global solution as (Sheng et al. 2006):

$$\mathbf{F}_i^{cont} = (\mathbf{F}_N^c)_i + (\mathbf{F}_T^c)_i \quad (8.12)$$

For the penalty method, the matrix form of the normal contact is expressed by (Heinstein et al. 2000; Wriggers 2006):

$$\mathbf{F}_N^c = \varepsilon_N G_N G_N^T \quad (8.13)$$

Where:  $\varepsilon_N$  is the penalty parameter,  $G_N$  is the gap function assembled into a matrix, which depends on the deformation.

Discretizing Eq. (8.13) and its associated tangent matrix will produce:

$$(\mathbf{F}_N^c)_i = \varepsilon_N A_i (G_N G_N^T)_i \quad (8.14)$$

In the case of the kinematical relations for the tangential motion,  $\mathbf{F}_T^c$ , in the contact area, it is necessary to distinguish between stick and slip motion, where only the classical Coulomb law is applied. Over a time increment  $\Delta t = t_{i+1} - t_i$ , the incremental tangential displacement is given by  $(\Delta G_{T\alpha})_i = \mathbf{T}_{\alpha i} \Delta u_i$ , where  $\mathbf{T}_{\alpha}$  is the expression for the tangent matrix (Sheng et al. 2006; Wriggers 2006). First consider the stick case, where

$$\left(\mathbf{F}_T^{c\ stick}\right)_i = \varepsilon_T A_i \mathbf{T}_{\alpha i} \mathbf{T}_{\alpha i}^T \quad (8.15)$$

and in the case of slip:

$$\left(\mathbf{F}_T^{c\ slip}\right)_i = \mu \varepsilon_N A_i \mathbf{T}_{\alpha i} \mathbf{G}_{N i}^T \quad (8.16)$$

Where:  $A_i$  is the contact area,  $\mu$  is the friction coefficient, and  $\varepsilon_T$  is the regularisation parameter which could be viewed as an elastic or spring constant applied in the tangential direction.

Refer to Wriggers (2006) for more discussion on the derivation and the explicit form of tangential stick and slip.

### 8.3.10 Results of FE Analysis

When comparing the numerical results to the experimental results, reasonable correlation between the results was obtained when the ductile and MSFLD damage criterion were used together. The ductility criteria, which were adopted from the ABAQUS example ‘Failure of a thin walled tubes’ (Dassault Systèmes 2011a), were scaled to suit the material properties. A scaling factor of 2/3 was found to yield results which closely matched the experimental response throughout the range of compression.

The force magnitude response of the post-buckling regime exhibits very similar behaviour to the experimental results for each peak value during folding (Fig. 8.6). The initial peak stress of 190 MPa (when buckling begins to occur) is almost identical to that which is exhibited by the experimental test specimens, in both magnitude and the displacement at which initial buckling of the structure occurs.

The mode of failure and layered folding pattern exhibited by the FE model (using shell elements) demonstrated reasonably good agreement with the experimental results (Fig. 8.7). The failure mode started as local denting followed by global folding of the thin wall, which is associated with large strains and extensive plastic deformation. The close-up deformation is shown in Fig. 8.8.

## 8.4 Discussion

The equations of equilibrium governing the dynamic response of the cylindrical shell were solved by an explicit integration operator, using the explicit central difference formula implemented in Abaqus/Explicit (Dassault Systèmes 2011c). Several researchers (Alghamdi 2001; Song et al. 2012; Heinsteinst et al. 2000; Bisagni 2000) have showed that the dynamic analysis using an explicit integration operator

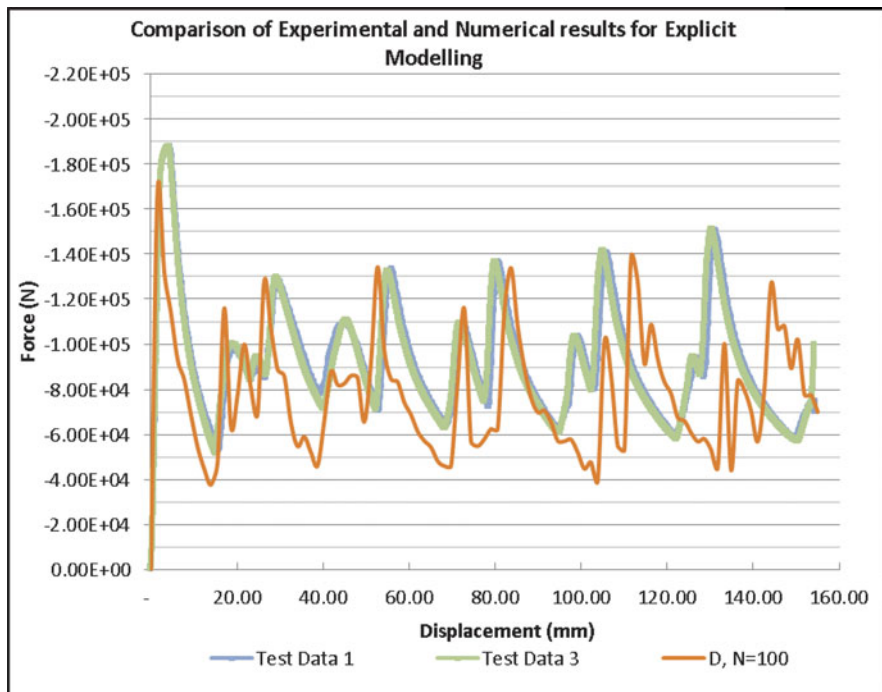


Fig. 8.6 Comparison of force–displacement plots for numerical and experimental analyses

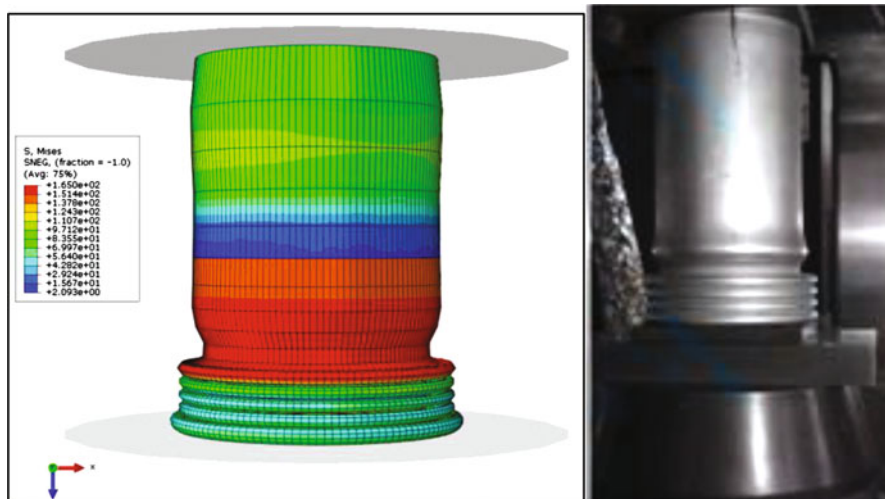
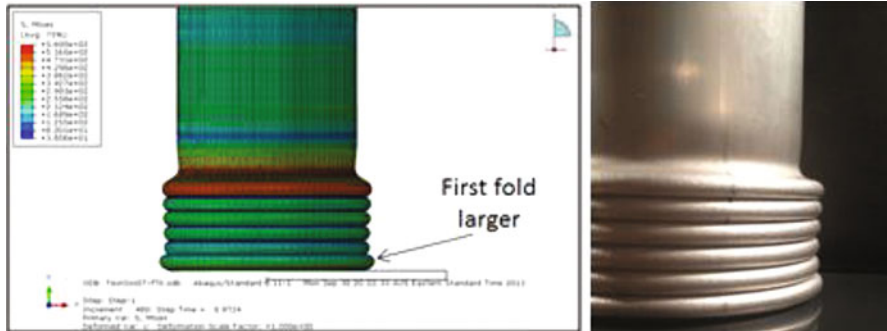


Fig. 8.7 Deformation of aluminium tubing using ductility and MSFLD damage criteria for Aluminium 6060 T5



**Fig. 8.8** Close-up deformation of aluminium tubing using ductility and MSFLD damage criteria for Aluminium 6060 T5

is efficient to study the buckling behaviour of cylindrical shells. In the study, only linear reduced integration elements were used, due to the inherent simplification of the mass matrix. Linear reduced elements have the mass lumped to the nodes (lumped mass matrix), which results in a simplified method of determining acceleration from the inverse diagonal mass matrix. Hence, the computational cost required is reduced.

Sensitivity checks were carried out for the cylindrical tube with global seed sizes of 10, 7.5, 5 and 2.5 mm around the collapsing end. The finest mesh size of 2.5 mm was found to yield satisfactory results; a difference of only 2 % in the plastic strain at the onset of the failure was achieved when compared with the coarser size of 5 mm. According to Giglio et al. (2012) a reasonable compromise between accuracy of results and duration of the analysis for sheet metals convergence is achieved when additional element division does not change the final plastic strain by more than 2 %. The final model of the cylindrical tube in this study was built with 9,394 elements.

The mesh refinement showed the best correlation to the experimental results using a biased mesh, especially at the end of the tube where the buckling initiated. The biased meshing method was found to be important in modelling the failure mode of the structure. Using biased mesh reduced the computational resources required for analysis, as the total number of elements (and consequently the number of equations) was reduced. Although using a biased mesh reduced the total number of elements in the analysis, there was still sufficiently density of mesh at the buckling end to produce accurate results.

The inclusion of damage criteria plays an important role in the collapse of structures in ABAQUS. Without the inclusion of such failure modes, significant issues arise with non-convergence and analysis error unless initial imperfections are included in the simulation. However as the experimental tubing was undamaged before the crushing occurred, it was considered that the model should be developed without initial imperfections. Hence the use of ductility damage criteria was included to initiate buckling failure. The ductility criteria were taken from the literature (Dassault Systèmes 2011a; Hooputra et al. 2004) and entered into the

material model for the aluminium. Analysis without ductility showed poor results where collapse was not present in the model. The best results for explicit modelling with damage criteria were produced when a  $2/3$  scaling factor was used on ductility criteria and used in conjunction with MSFLD criteria.

The inclusion of MSFLD criteria (Dassault Systèmes 2011a) had a significant effect on the failure mode of the tube. The force–displacement response for the model without MSFLD does not correlate well with the experiment for collapse forces or the buckling pattern. Figure 8.7 displays the collapse mode for models where the MSFLD criteria were included. The numerically simulated collapse occurs with the same failure mode as the experiments, where five circular folds occurred at one end of the tubing, hence the use of MSFLD criteria was considered necessary.

The general contact definition with hard contact relationship and penalty enforcement within the ABAQUS code (Dassault Systèmes 2011d) was used to model the contact between the press plates and the tube. It was also found that the FE model was sensitive to the coefficient of friction ( $\mu$ ) used for the contact surfaces. The crushing tube was studied for  $\mu = 0, 0.1, 0.2, 0.5$  and  $0.7$ . The final model required a lower coefficient of friction,  $\mu = 0.2$  to gain accurate results.

The dynamic buckling of cylindrical tube was challenging to simulate accurately due to the strong nonlinearities caused by frictional contact, large deformation (geometrical) and elasto-plasticity (materials) (Sheng et al. 2006). Extra efforts were needed to select appropriate material properties and contact modelling methods in order for the simulation to accurately represent the experimental results.

Overall, the finite element model was capable of accurately simulating the crushing behaviour of the thin walled tube. It was able to predict, with high accuracy, the force and displacement of the load press at which critical buckling would occur, as well as the relative magnitude of post-buckling loads required at each subsequent failure of the tube. It demonstrated peak forces corresponding to the preliminary and secondary folds, which occurred during each subsequent buckle. The model was able to adequately capture the behaviour of the collapsing structure for engineering purposes.

## 8.5 Conclusion

ABAQUS explicit modelling is suitable for post-buckling analysis with large deformations. Mesh refinement in the area of failure is required to allow the analysis to converge on a solution. Additionally, biased meshing is useful to control where the buckling collapse occurs in structures with identical ends, such as the specimens analysed. Biased meshing also reduces the number of elements and computations required to complete analysis. The inclusion of ductility damage criteria and MSFLD criteria had a significant effect on the collapse of the cylindrical tube without initial imperfections, where experimental and numerical results demonstrated good correlation.

## Key Symbols

$A_i$	Contact area
$\mathbf{c}$	Damping matrix
$E$	Young's modulus
$\Delta t$	Time increment
$\mathbf{F}$	Force vector
$\mathbf{F}_N^c$	Matrix form of the normal contact
$\mathbf{F}_T^c$	Matrix form of the tangential contact
$G_N$	Gap function
$\mathbf{K}$	Stiffness matrix
$\mathbf{M}$	Mass matrix
$\mathbf{T}_\alpha$	Tangent matrix
$u, \dot{u}, \ddot{u}$	Nodal displacement, velocity and acceleration vector
$\varepsilon_N$	Penalty parameter
$\varepsilon_T^+$	Equibiaxial tensile
$\varepsilon_T^-$	Equibiaxial compressive deformation
$\bar{\varepsilon}_D^{pl}$	Equivalent plastic strain at the onset of damage
$\eta$	Stress triaxiality
$\nu$	Poisson's ratio
$\mu$	Friction coefficient
$\theta_S$	Shear stress ratio
$\rho$	Density
$\sigma_o$	Yield stress
$i, i - 1/2, i + 1/2$	Increment number and mid-increment numbers

## References

- Abramowicz W (2003) Thin-walled structures as impact energy absorbers. *Thin Walled Struct* 41:91–107
- Alghamdi AAA (2001) Collapsible impact energy absorbers: an overview. *Thin Walled Struct* 39:189–213
- Bisagni C (2000) Numerical analysis and experimental correlation of composite shell buckling and post-buckling. *Compos Part B Eng* 31:655–67
- Bisagni C (2005) Dynamic buckling of fiber composite shells under impulsive axial compression. *Thin Walled Struct* 43:499–514
- Burlayenko VN, Sadowski T (2014) Nonlinear dynamic analysis of harmonically excited debonded sandwich plates using finite element modelling. *Compos Struct* 108:354–66
- Dassault Systèmes (2011a) Example problems manual, volume I: static and dynamic analyses. Dassault Systèmes Simulia Corp., Providence
- Dassault Systèmes (2011b) Analysis user's manual, volume III: materials. Dassault Systèmes Simulia Corp., Providence
- Dassault Systèmes (2011c) Analysis user's manual, volume II: analysis. Dassault Systèmes Simulia Corp., Providence

- Dassault Systèmes (2011d) Analysis user's manual, volume V: prescribed conditions, constraints & interaction. Dassault Systèmes Simulia Corp., Providence
- Giglio M, Manes A, Viganò F (2012) Ductile fracture locus of Ti–6Al–4V titanium alloy. *Int J Mech Sci* 54:121–35
- Heinstein MW, Mello FJ, Attaway SW, Laursen TA (2000) Contact—impact modeling in explicit transient dynamics. *Comput Methods Appl Mech Eng* 187:621–40
- Hooputra H, Gese H, Dell H, Werner H (2004) A comprehensive failure model for crashworthiness simulation of aluminium extrusions. *Int J Crashworthiness* 9:449–64
- Sheng D, Wriggers P, Sloan SW (2006) Improved numerical algorithms for frictional contact in pile penetration analysis. *Comput Geotech* 33:341–54
- Song J, Chen Y, Lu G (2012) Axial crushing of thin-walled structures with origami patterns. *Thin Walled Struct* 54:65–71
- Strömberg N (2005) An implicit method for frictional contact, impact and rolling. *Eur J Mech – A Solids* 24:1016–29
- Sun JS, Lee KH, Lee HP (2000) Comparison of implicit and explicit finite element methods for dynamic problems. *J Mater Process Technol* 105:110–8
- Wei D, Elgindi MBM (2013) Finite element analysis of the Ramberg–Osgood bar. *Am J Comput Math* 3:211–2116
- Wriggers P (2006) *Computational contact mechanics*. Springer, Berlin



# Chapter 9

## Exhaust System Acoustic Modeling

Milan Simic

**Keywords** Exhaust system • Sound • Acoustic filter • Physical networks  
• Transfer function • Power spectrum • Bode plot

### 9.1 Introduction

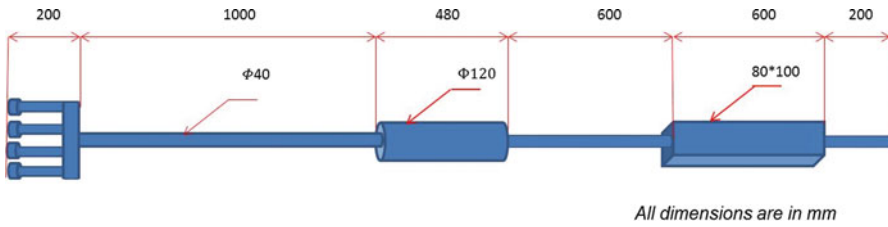
An exhaust system removes chemical and acoustic pollutions generated by a combustion engine. Like all other subsystems in automotive engineering, exhaust systems have evolved for all this years of automotive history, from simple cast iron manifold, with pipes and a silencer to complex systems that include catalytic converters, supertone silencer, muffler, and tailpipe tips. One of the key system components is a catalytic converter which converts poison gasses pollution by transforming it into carbon dioxide and water. System often has few exhaust pipes and silencers.

All components of the system, regardless of their primary role of chemical or noise pollution control, can be seen as a kind of acoustic filters. Mufflers are particularly designed in such a way that acoustic waves, travelling through, are maximally attenuated, obstructed, and subjected to self-interference. A study on acoustic modeling and testing of exhaust and intake system components is presented in Elnemr (2011). In addition to that, research of automotive exhaust system hanger location, for the best prevention of vibration transfer to the vehicle body, is conducted as presented in Jianwang et al. (2008). In our research, at this stage, we have concentrated on the acoustic side of the system.

---

M. Simic (✉)

School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University,  
Bundoora East Campus, Plenty Road, Bundoora, Melbourne, VIC 3083, Australia  
e-mail: [milan.simic@rmit.edu.au](mailto:milan.simic@rmit.edu.au)



**Fig. 9.1** An exhaust system structure

There is a large variety of system constructions for various vehicles. A generic exhausted system is shown in Fig. 9.1. Dimensions are given as they are used in our modeling. Some systems may have everything doubled, or only mufflers, or just tailpipe. This is used to increase power and torque, i.e., to introduce less attenuation of those key performance engine characteristics. For example, an investigation on exhaust system parameters for fuel economy improvement of small gasoline engine is presented in Peng et al. (2009). In addition to that there is still energy that can be extracted from the exhaust system as shown in Capel et al. (2013). We have measured the real signal, i.e., sound pressure, from the vehicle and designed exhaust system model starting from that.

Our investigation of the system acoustics is based on the Physical Networks' approach. Any engineering system can be expressed as a physical network (Sanford 1965). It can incorporate physical quantities of the same type, like mechanical system with translator motion, or just electrical system. Modern engineering systems are mainly comprehensive, including physical subsystems with different types of quantities, like electromechanical, or more than that, mechatronics (Silva 2005). Physical networks approach, modeling and simulations of acoustic, hydraulic, and mechanical systems, using Electronics Circuits Analysis Program (ECAP) are presented in Simic et al. (1978).

In this paper, the basic physical network principles are explained in Sect. 9.1. Components of the exhaust physical network are introduced in Sect. 9.2. The whole system model and simulation by a large ladder filter topology are presented in Sect. 9.3. Transfer function, i.e., bode diagram of the model is given in the next section. Finally frequency characteristic of the measured sound, i.e., power spectrum of the sound from the real vehicle is compared to the model performances and presented in the final section of this chapter.

## 9.2 Physical Networks

Long time ago, in the nineteenth century, scientists, physicists, and philosophers have stated that the unity of the nature is expressed in the amazing analogies of the differential equations used to represent various physical phenomena. The same

equations can be used for the study of hydrodynamics and for the electrical potential theory. The theory of turbulence in liquids and the theory of friction in gases show extreme analogies with the electromagnetic theory.

Our, nonlinear, exhaust system, as shown in Fig. 9.1, can be presented as a cascaded structure of linear, one-dimensional physical subsystems with constant parameters. Those building components, called physical networks, consist of ideal linear elements with two terminals. Network elements represent mathematical relationships between two dependent system variables in a physical system. Often, real systems are extremely complex, so that they cannot be easily simplified, but many systems are in that category, or have linear subsystems as their integral parts.

An equivalent physical network represents a sort of linear graph associated to the system equations. The basic definitions, principles, and rules for solving system equations represented by the network are independent of physical system that the particular network is representing.

### 9.2.1 Network Variables

Physical networks have two basic time dependent variables: flow,  $f$ , and potential,  $p$ .

- Flow is a variable that flows through network elements and connection lines
- Potential is a variable measured across a network element, or between two network points
- Potential of a point in the network depends on the chosen referent point.

Examples of physical variables are electrical *current* and *potential*. Current is directly related to the flow of electrons as  $i = \frac{dq}{dt}$ . Potential difference is called voltage. Then we have *force* and *velocity* in a mechanical system with translation, or *torque* and *angular* speed in a mechanical system with rotation. In a hydraulic system we have *flow* and *pressure*. All those systems can be modeled using the same type of ordinary differential equations (ODE).

### 9.2.2 Ordinary Differential Equations

We can use ODEs with constant coefficients for modeling various one-dimensional physical systems. An ODE is a relation between two variables,  $t$  and  $y = y(t)$ , and the derivatives of  $y$ ,  $\frac{dy}{dt}$ ,  $\frac{d^2y}{dt^2}$ ,  $\dots$ ,  $\frac{d^ny}{dt^n}$ .

$$A_n \frac{d^n y}{dt^n} + A_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + A_1 \frac{dy}{dt} + A_0 y = f(t) \quad (9.1)$$

Equation (9.1) is called ordinary because only one independent variable exists, which is usually time, and it is linear because only the first exponent of dependent variable, or its derivatives is present. General solution,  $y_G(t)$ , of an ODE is given as

$$y_G(t) = y_H(t) + y_P(t),$$

where  $y_H(t)$  is general solution of equation when  $f(t) = 0$ , and  $y_P(t)$  is any solution of the Eq. (9.1).

### 9.2.3 Network Elements

Physical network is a linear graph that includes network elements and connecting lines between them. There are two types of elements: active and passive. Active elements are flow, or potential sources, whose operations are expressed as functions of time:  $F = F(t)$  and  $P = P(t)$ . They supply energy to the system. Passive elements are two connection points elements with defined relationship between *flow* through and *potential* across elements' terminals. They cannot supply more energy to the system than what was already accumulated. That is expressed through initial conditions. There are three types of passive elements used to express three different relationships between network variables, *flow* and *potential*. Those basic relationships are *proportion*, *integration*, and *differentiation*.

Examples of all three types of passive elements in generic physical network, as well as in electrical and mechanical system with translation are given in the Table 9.1. Expressions for the accumulated energy and power dissipation in each of those systems are also given. In a generic physical system, two network quantities relations as integration, proportion, and differentiation are associated to elements labeled as  $A$ ,  $B$ , and  $C$ , as shown in Table 9.1. The general name for all elements,  $A$ ,  $B$ , and  $C$  is impedance,  $Z$ , or admittance  $Y$ . Similarly, analog to those quantities, in an electrical system we have inductivity,  $L$ , conductivity  $G$ , i.e., resistivity,  $R = 1/G$ , and capacity  $C$ . Network variables are current,  $i$ , and voltage,  $u$ , which is electrical potential difference. Finally for the translation we have  $k$  for stiffness,  $B$  for friction,  $m$  for mass of the object. Network variables here are force,  $F$ , and velocity  $v$ .

**Table 9.1** Analogies relationships

Relationship	Prototype	Electrical	Translation
Integration	$f = A \int p dt$	$i = \frac{1}{L} \int u dt$	$F = k \int v dt$
Accumulated energy	$\frac{f^2}{2A}$	$\frac{Li^2}{2}$	$\frac{F^2}{2k}$
Proportion	$f = Bp$	$i = Gu = \frac{1}{R}u$	$f = Bv$
Power dissipation	$fp = \frac{f^2}{B} = p^2B$	$iu = i^2R = \frac{u^2}{R}$	$Fv = \frac{F^2}{B} = v^2B$
Differentiation	$f = C \frac{dp}{dt}$	$i = C \frac{du}{dt}$	$F = m \frac{dv}{dt}$
Accumulated energy	$\frac{Cp^2}{2}$	$\frac{Cu^2}{2}$	$\frac{mv^2}{2}$

There are also other systems like thermal, fluids, or rotation where analogies could be easily established as given in Sanford (1965) and Silva (2005).

### 9.2.4 *Mono and Multi-dimensional Systems*

Apart from already presented mono-dimensional systems, real physical systems are often multi-dimensional. In that case, they have to be represented by partial differential equations including more than one independent variable. Systems that deal with the physical fields' effects in 3D space, like gravity, or electromagnetic field are multi-dimensional. Even when energy is distributed through one dimension, but that transfer is the function of time, we have a multi-dimensional system.

In mono-dimensional dynamic systems, where independent variable is time, we can assume that, the energy associated to the dependent variables is transferred instantly through the system, to all remote parts. Accordingly, dependent variables are just functions of time and not the distance from the energy source. We should mention here that, there are also static mono-dimensional systems, where independent variable is not time, but could be distance  $x$ . Examples of such systems could be found in Statics, a branch of Mechanics.

The issue of system multi-dimensionality, caused by the energy transition time, could appear in electrical systems of the small dimensions, when the emitting frequency of the energy source is high, or when we have long distance energy transmissions. While electrical energy travels relatively quickly many other forms of energy are for the order of magnitude slower than that. Examples are acoustic waves, forces, and waves in liquids, and especially heat energy transfer in solids. Those systems have to be treated as multi-dimensional even when low frequencies are involved and physical dimensions of the systems are small.

In order to get objective criteria for the best estimation of the system dimensionality, when harmonic energy sources are applied, we should look at the source signal, i.e., energy wavelength,  $\lambda$ . It is given as

$$\lambda = \frac{c}{f} \quad (9.2)$$

where  $c$  is the transfer speed and  $f$  is the frequency defined by the energy source. Wavelength must be larger than the longest dimension in the physical system so that the system can be treated as mono-dimensional. When that is not the case, then system has to be treated as multi-dimensional since impedance is not concentrated in one point only. It is distributed along the length of the system and following that, partial differential equations have to be used, with two independent variables, time, and the distance.

### 9.3 Acoustic Elements

Exhaust system is a nonlinear, multi-dimensional physical network. Since we are now concentrated on the acoustics of the system, only, it is for us an acoustic physical network. Nonlinearities come from the changes of the temperature and the gas pressure along the system, as well as the physical dimensions and the shape of the system structure.

#### 9.3.1 From Multi to Mono-Dimensional System Model

Let us start with the dimensionality problem first. The longest dimension in the systems is the length, which is 3,800 mm. The speed of sound, in the air, at 20 °C is equal to  $c_{20} = 343$  m/s. Temperature of the gasses is higher than the temperature of the environment, which influence the speed, but for the initial estimations we will take the speed as  $c = c_{20} = 343$  m/s. Now we need the frequencies of the emitted sound. In order to get that, sound is recorded and power spectrum is extracted using the program shown in Fig. 9.2.

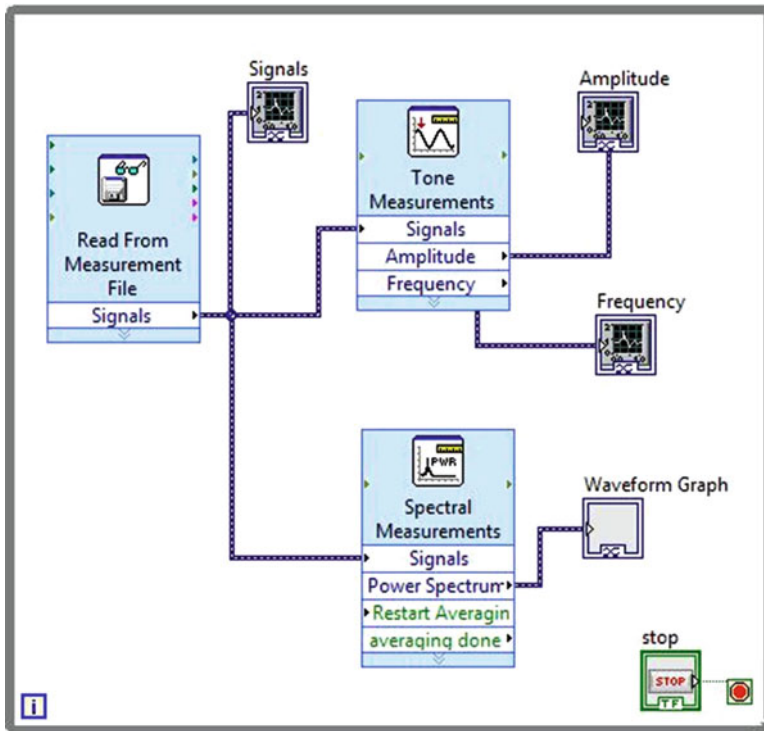
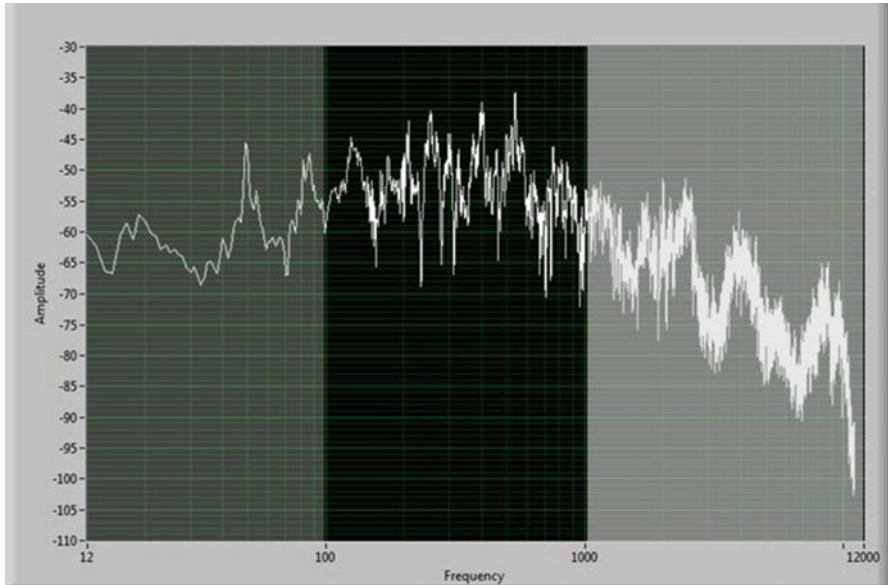


Fig. 9.2 Power spectrum measurement program



**Fig. 9.3** Exhaust system sound power spectrum

Powers spectrum of the sound is shown in Fig. 9.3. We can see that the majority of energy is transferred by the frequency components in the rage of 10–1,000 Hz. The maximum energy is carried by the components around 500 Hz.

Following that we can calculate wavelengths associated to those frequencies:

$$\lambda_1 = \frac{c}{f_1} = \frac{343}{10} = 34.3m \quad \lambda_2 = \frac{c}{f_2} = \frac{343}{1,000} = 0.343m \quad (9.3)$$

We will now split the system into the subsystems whose maximum linear dimensions are small compared to the minimum wavelength in the range of interest, which is  $\lambda_2$ . In the first simulation we will adapt the subsystems' lengths of 80, 160, and 200 mm as given by the vector  $V$ . Vector components, associated to the exhaust system from the Fig. 9.1, are all given in mm. After comparing simulation results by the real measurements, given in the Fig. 9.3, we will see if further subdivisions are needed.

$$V = (200, 200, 200, 200, 200, 200, 200, 80, 80, 80, 80, 80, 80, 200, 200, 200, 200, 200, 200, 200) \quad (9.4)$$

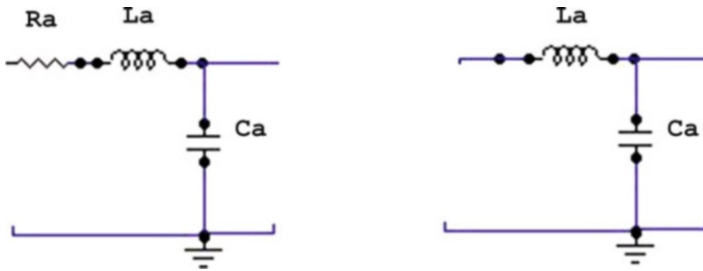


Fig. 9.4 Basic acoustic subsystem

### 9.3.2 Acoustic Network Elements Linearization

Exhaust system, shown in the Fig. 9.1, can be represented as a cascaded network of  $R_a L_a C_a$  subsystems, shown in Fig. 9.4, where  $R_a$  is a symbol for the acoustic resistance,  $L_a$  is a symbol for the acoustic inductance, and  $C_a$  represents acoustic capacity.

Acoustic resistivity is an element that dissipates acoustic energy in the system. It incorporates all sound attenuation components, as found along the whole exhaust, especially in catalytic converters and mufflers. For the simplicity of the initial modeling we will neglect effects of the resistivity as a separate component and represent basic subsystem without proportional element, i.e., resistivity, as shown in the Fig. 9.4. At the same time we are going to include attenuation through the effects of more rapid temperature decrease in the converter and mufflers chambers. The fast temperature change is caused by the gas molecules kinetic energy drop, when colliding with obstacles on the path through the various system components.

Acoustic capacity and inductivity are calculated using the following expressions:

$$C_a = \frac{Sl}{\rho c^2} \quad L_a = \rho \frac{l}{S} \quad (9.5)$$

where  $c$  is the transfer speed, as before,  $S$  is the cross section surface area,  $l$  is the length of the subsystem, and  $\rho$  is the gas density (Sanford 1965; Simic et al. 1978). Parameter  $l$  (length) takes the values from the components of the vector  $\mathbf{V}$ , given by Eq. (9.4).

The next nonlinearity issue to challenge is the temperature dependency of the whole system that can be represented as changes in network elements  $L_a$  and  $C_a$ , as the function of the temperature. Let us start with the ideal gas law equation

$$\frac{pV}{T} = Const \quad (9.6)$$

$$\frac{p_1 V_1}{T_1} = \frac{p_0 V_0}{T_0}$$



where

$$p_0 = 760 \text{ mmHg} = 0.1 * 10^6 \text{ N/m}^2 = 10^6 \mu\text{bar}$$

is normal atmospheric pressure, while  $V_0$  is gas volume at the temperature of  $T_0$ ,

$$T_0 = 273.15 \text{ K} = 0^\circ\text{C}.$$

Let us assume that the pressure difference at the borders of the subsystems is  $\Delta p = 200 \mu\text{bar}$ . Then we have the following:

$$p_1 = p_0 + \Delta p$$

$$\frac{p_1}{p_0} = \frac{10^6 + 200}{10^6} = 1.0002 \approx 1$$

Following that we have simplified Eq. (9.6) as given below,

$$\frac{V_1}{T_1} = \frac{V_0}{T_0} \quad \text{or} \quad \frac{V_0}{V_1} = \frac{T_0}{T_1} \quad (9.7)$$

Gas density is expressed as

$$\rho_0 = \frac{m}{V_0} \quad \text{for } T = T_0$$

$$\rho_1 = \frac{m}{V_1} \quad \text{for } T = T_1$$

Dividing last two equations we have

$$\frac{\rho_1}{\rho_0} = \frac{V_0}{V_1} \quad (9.8)$$

From the Eqs. (9.7) and (9.8) we derive

$$\rho_1 = \rho_0 \frac{V_0}{V_1} = \rho_0 \frac{T_0}{T_1} \quad (9.9)$$

Finally expression for the acoustic inductivity becomes

$$L_a = \rho \frac{l}{S} = \rho_0 \frac{l}{S} \frac{T_0}{T_1} \quad (9.10)$$

Specific gas density is labeled as  $\rho_0$  at the temperature of  $T_0 = 0^\circ\text{C} = 273.15 \text{ }^\circ\text{K}$ . Working temperature is labeled as  $T_1$ . Gas density is another nonlinear quantity, which depends on the engine cycle, but we will assume that it is constant and equal to the air density.

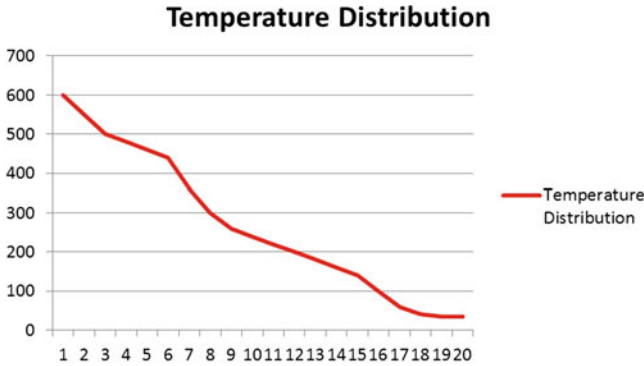


Fig. 9.5 Gas temperature distribution along the system

Finally expression for the acoustic capacity would be:

$$C_a = \frac{Sl}{\rho c^2} = \frac{Sl}{\rho_0 \frac{T_0}{T_1} \left( c_0 \sqrt{\frac{T_1}{T_0}} \right)^2} = \frac{Sl}{\rho_0 c_0^2} \tag{9.11}$$

We can see that the acoustic capacity is not the function of the temperature.

Gas burning temperature depends on the engine type but it is around 650 °C generally. At the exit, out of the tailpipe, temperature is still above ambient temperature and we will adapt the value of 35 °C. Temperature distribution is given in the Fig. 9.5.

We can see that there are 20 subsystems, with the length distribution as given by the vector **V**, Eq. (9.4), and with the temperature distribution as per Fig. 9.5, from 600 °C at the start of the cascade down to 35 °C at the end. In addition to that, each subsystem has its cross section surface area as specified by Fig. 9.1. We have circles with given diameters as  $\Phi 1 = 40$  and  $\Phi 2 = 120$  and a rectangle with  $80 \times 100$ , all in *mm*.

In the subsystems which have higher temperatures, sound speed is higher too, and accordingly the wavelength as per Eq. (9.2). As a consequence, the approximation of the physical system is more accurate because of the better relationship between wavelength and the maximum liner dimensions of the system.

### 9.4 Ladder Topology

We can now design the whole system. Figure 9.6 shows a segment of the ladder network that consists of the subsystems as given in the Fig. 9.4.

In this physical network we have *flow* and *pressure* as network variables.

An Excel program is used to calculate distributed parameters. A fragment from the calculations is demonstrated in Fig. 9.7.

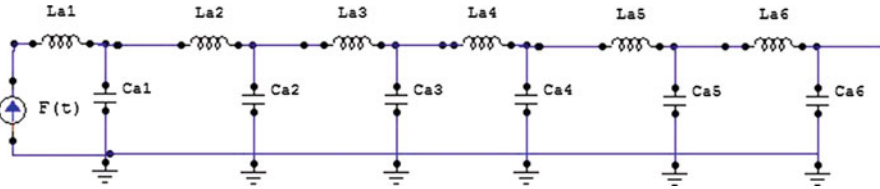


Fig. 9.6 A segment from the acoustic ladder network

Distributed Parameters

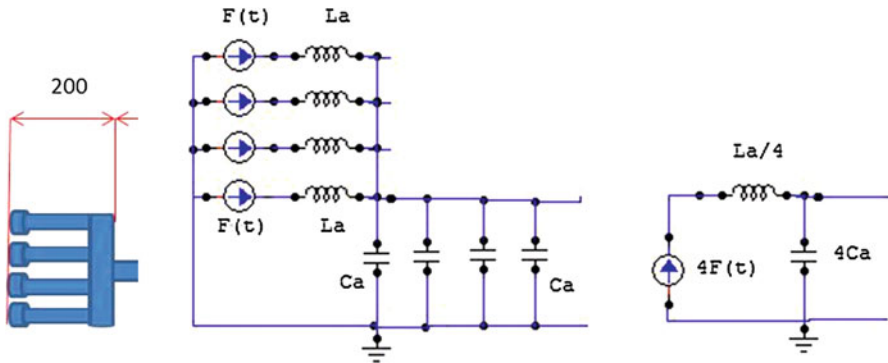
Distributed Parameters																	
Ti=To+Ti)																	
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
480	460	440	360	300	260	240	220	200	180	160	140	100	60	40	35		
753.15	733.15	713.15	633.15	573.15	513.15	493.15	473.15	453.15	433.15	413.15	373.15	333.15	313.15	308.15	306.15		
0.467853	0.480659	0.4940945	0.5585245	0.6147841	0.6600287	0.6886876	0.7145759	0.7447894	0.7775989	0.8134307	0.8528706	0.9442945	1.0578722	1.1252227	1.1434804		
0.2	0.2	0.2	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.2	0.2	0.2	0.16	0.16	0.16	0.2	0.2
0.001256	0.001256	0.001256	0.011304	0.011304	0.011304	0.011304	0.011304	0.011304	0.011304	0.001256	0.001256	0.001256	0.00	0.00	0.00	0.001256	0.001256
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
74.498888	76.531184	78.67747	3.3386024	4.3509126	4.6773443	4.8596436	5.0587294	5.2704768	123.81957	123.53674	135.80742	1.889583	2.1153444	2.2504455	182.06287	1777E-09	1777E-09
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
437.6018	431.75181	425.8221	1003.0636	954.3593	920.4548	903.02532	885.25274	867.11597	339.43666	331.86163	324.1094	385.0255	363.80417	352.71504	278.91027		
7552248.5	7351697.5	7151146.5	33680891	35920561	33413673	32160230	30906786	29653342	4543983.8	4343432.8	4142881.9	58465311	5218803.3	4906444.4	3089989.2	2748.1355	2711.4014
2674.1628	6293.2771	5393.3764	5780.4562	5670.959	5553.3872	5445.4883	2131.6622	2084.0904	2035.407	2417.9601	2284.6902	2215.0504	1757.8365				

Fig. 9.7 Distributed parameters calculations

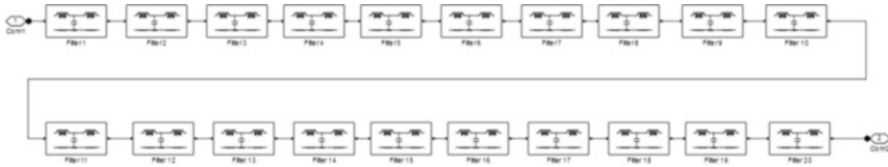
We have to clarify just one more design solution. Header, as shown in Figs. 9.1 and 9.8, consists of four extractors with the diameter of 40 mm and the length of 200 mm. Those four pipe segments can be represented in the physical network as four parallel connected acoustic inductors and four parallel connected acoustic capacitors, as shown in Fig. 9.8. Energy sources are presented as series connection with acoustic inductivities. Equivalent circuit or network structure is given in Fig. 9.8. This subsystem is presented as parallel connection of elements because we have about the same pressure difference on each of them and the total flow down the pipe, i.e., the whole exhaust system is equal to the sum of four flows. This qualifies such a system as a parallel structure in the physical networks. According to that, total equivalent inductivity  $L_{ae}$  and capacity  $C_{ae}$  of the first subsystem are given as

$$L_{ae} = \frac{L_a}{4} \quad \text{and} \quad C_{ae} = 4C_a \tag{9.12}$$

To finalize this discussion we have to mention that the exhaust system behaves as a sound source at the end of tailpipe. We need to know impedance of the open pipe. It can be show that it is predominantly inductive. We also have to know that the gas has a one-directional motion component, apart from oscillations, and that it moves outside of the exhaust forming a hot gas cylinder. By observation we concluded that this cylinder has the length of around 200 mm as the most of our subsystems. We put that in the calculations together with the simulation of the open space by the larger values for the last subsystem dimensions.



**Fig. 9.8** Header and its equivalent physical network: Flow sources  $F(t)$  are represented by arrows, acoustic inductors as  $L_a$  and capacitors as  $C_a$ . Equivalent circuit is given on the right where equivalent source flow is  $4F(t)$  and equivalent components are as given in the text and on the figure



**Fig. 9.9** Simulink model of the cascaded acoustic filter

We have used various modeling environments in the different stages of the design. Modern technology enables easy integration and combination of results from the different programming platforms. Our model in Simulink is shown in Fig. 9.9. Ladder network shown in Fig. 9.6 is built on  $L$  half sections, while the other one from the Fig. 9.9 is based on  $T$  section filter subsystems integration.

### 9.5 Transfer Function

Let us look at the basic  $L$  Half section of our ladder network as shown in Fig. 9.4. Transfer function in Laplace domain of this  $L_a C_a$  filter is given as

$$T = \frac{u_{out}}{u_{in}} = \frac{\frac{1}{C_s}}{\frac{1}{C_s} + Ls} = \frac{1}{1 + LCs^2} \tag{9.13}$$

Applying this to the whole ladder network we come up with the acoustic transfer function of the exhaust system.

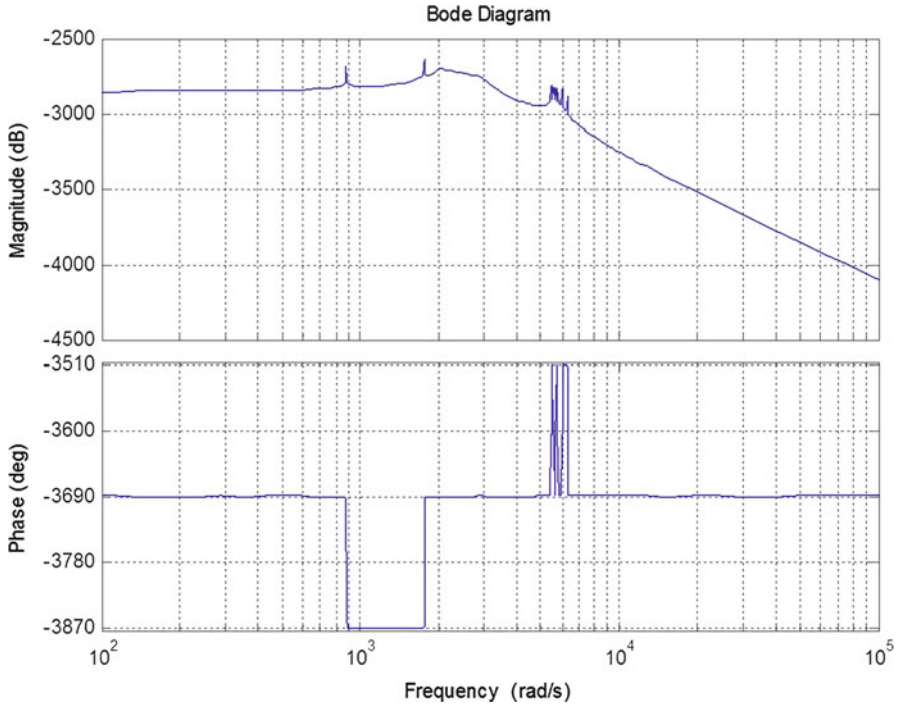


Fig. 9.10 Bode plots of the exhaust system

$$T_{all} = \frac{1}{\prod_{i=1}^{i=20} (1 + L_{ai} C_{ai} s^2)} \tag{9.14}$$

Bode plot of the transfer function is calculated using MATLAB and the magnitude and phase shift graphs are shown in the Fig. 9.10.

## 9.6 Conclusion

Comparing bode plot graph from the Fig. 9.9 and the power spectrum diagram from the real exhaust system, shown on the Fig. 9.3, we can see that our model simulate exhaust system with extremely high precision. Figure 9.9 also presents phase response of our acoustic ladder filter. Both transfer function diagrams have frequency expressed in *rad/s* shown on *x* axis in logarithmic scale. Phase shift plot is shown in *degrees*, while magnitude is presented in decibels of power. The range presented is from  $10^2$  to  $10^5$  rad/s. Since we know that  $w[\text{rad/s}] = 2\pi f[\text{Hz}]$ , the frequency range covered by the diagrams in Fig. 9.9, in *Hz* is from 16 to 1,600 Hz. This was our range of interest according to the real, measured data from the vehicle, Fig. 9.3.

This model is now a tool for the further investigation in acoustic properties and design of the new exhaust systems. We can easily increase number of subsystems and so achieve better relationship between wavelength and the maximum dimension in the subsystem. In addition to that we could place temperature sensors along the system to get more accurate temperature distribution. Microphones, pressure sensors, infrared, and other could be also used to collect more real data (Ba et al. 2010).

## Key Symbols

$A, B, C$	Network elements
$A_i$	Ordinary Differential Equation (ODE) Coefficient $i, i = 0, 1, \dots, n$
$c$	Wave transfer speed
$C$	Capacity
$C_a$	Acoustic capacity
$C_{ae}$	Equivalent acoustic capacity
$\frac{d^i y}{dt^i}$	$i$ th derivative of $y, i = 1, 2, \dots, n$
$dp$	An infinitesimal change in potential
$dq$	An infinitesimal change in electric charge
$\frac{dq}{dt}$	First derivative of charge with respect to time
$dt$	An infinitesimal change in time
$du$	An infinitesimal change in voltage
$dv$	An infinitesimal change in velocity
$E$	Equivalent sound source
$f$	Flow
$f$	Frequency
$f(t)$	Function of time
$F, F(t)$	Flow source
$F$	Force
$G$	Conductivity
$i$	Electric current
$k$	Spring stiffness
$l$	Subsystem length
$L$	Inductivity
$L_a$	Acoustic inductance
$L_{ae}$	Equivalent acoustic inductance
$m$	Mass
$p$	Potential
$P, P(t)$	Potential source
$q$	Electric charge
$R$	Resistivity
$R_a$	Acoustic resistivity
$s$	Complex argument

$S$	Cross section surface area
$t$	Time
$T, T_0$	Temperature
$T$	Transfer function
$u$	Voltage
$v$	Velocity
$w$	Angular frequency
$V, V_0$	Gas volume
$\mathbf{V}$	Lengths vector
$y$	Variable
$y_G(t)$	ODE general solution
$y_H(t)$	Homogenous ODE solution
$y_P(t)$	A particular solution of an ODE
$Y$	Admittance
$Z$	Impedance
$\Delta p$	Pressure difference
$\lambda$	Wavelength
$\pi$	pi constant
$\prod_{i=1}^{i=20}$	Product from $i = 1$ to $i = 20$
$\rho$	Gas density
$\Phi$	Diameter

## References

- Ba F, Wan L (2010) Five components automobile exhaust measurement system. In: International conference on educational and network technology (ICENT), pp 248–250
- Capel EM, Ibrahim T, Nor NM (2013) Hybrid energy from exhaust system. In: IEEE seventh international power engineering and optimization conference (PEOCO), pp 134–138
- Elnemr Y (2011) Acoustic modeling and testing of exhaust and intake system components. Licentiate Thesis, Department of Aeronautical and Vehicle Engineering, Royal Institute of Technology, Stockholm
- Jianwang S, Aimin D, Xiaoxiong J (2008) Research of automotive exhaust system hanger location. In: IEEE Vehicle power and propulsion conference (VPPC'08), pp 1–3
- Peng H, Yunqing L, Lifeng Z (2009) Study on exhaust system parameters for fuel economy improvement of small gasoline engine. In: International technology and innovation conference (ITIC 2009), pp 1–5
- Sanford RS (1965) Physical networks. Prentice-Hall, Englewood Cliffs
- Silva CW (2005) Mechatronics an integrated approach. CRC Press LLC, New York
- Simic M (1978) Physical systems analysis using ECAP. Masters Research, Electronics Engineering, University of Nis, Nis, Serbia

# Chapter 10

## Nonlinear Approaches in Three Dimensional Medical Image Registration

Ruwan Tennakoon, Alireza Bab-Hadiashar, and Zhenwei Cao

**Keywords** Non-rigid registration • Medical image analysis • Deformable registration

### 10.1 Introduction

Image registration is the task of finding a spatial transformation  $T$  that aligns the objects in two or more images capturing the same or related scene. It is one of the most crucial problems of computer vision and has been studied for over three decades. The underlying task is very general and has a wide range of applications in the fields of medical imaging, machine vision, remote sensing, cartography, etc. In this chapter we focus on the application of image registration in the fields of medical imaging.

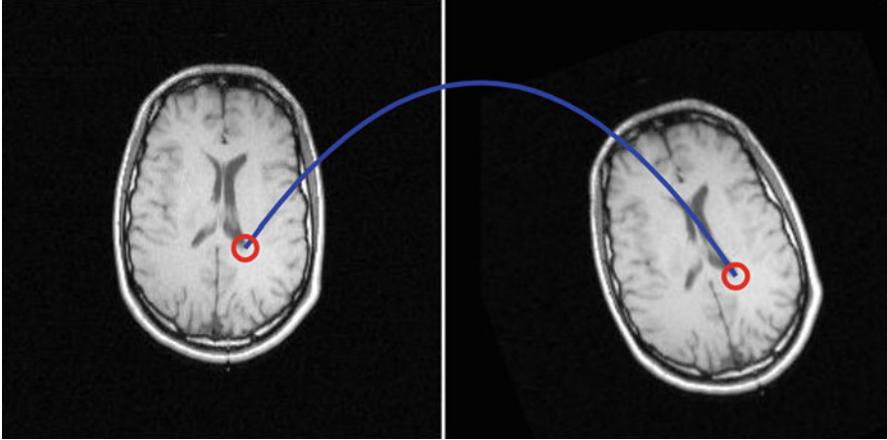
Although registration can involve two or more images, typical medical image registration problem comprises of two images. The two images to be registered are named fixed or target ( $I_F(x)$ ) and moving or source ( $I_M(x)$ ) images. The two images can be from different times, captured using different sensors or from different viewpoints (Zitova and Flusser 2003).

---

R. Tennakoon (✉) • Z. Cao  
Faculty of Science, Engineering and Technology, Swinburne University of Technology,  
Hawthorn, VIC 3122, Australia  
e-mail: [rtennakoon@swin.edu.au](mailto:rtennakoon@swin.edu.au)

A. Bab-Hadiashar  
School of Aerospace, Mechanical and Manufacturing, RMIT University,  
Melbourne, VIC 3001, Australia





**Fig. 10.1** In this illustration the moving image needs to be both translated and rotated to align with the fixed image

Mathematically the problem of image registration can be expressed as finding the spatial mapping

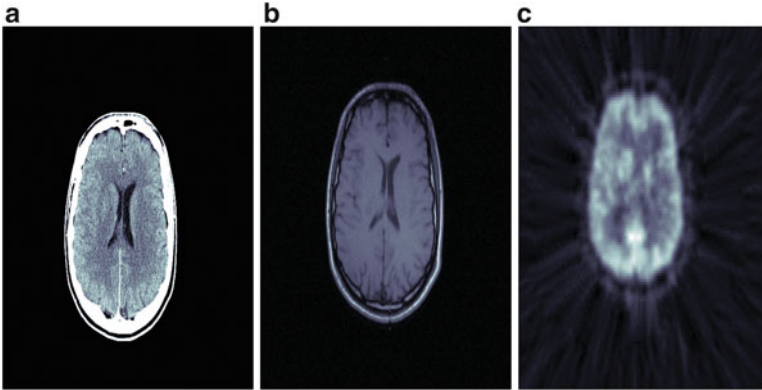
$$T : \Omega_F \rightarrow \Omega_M$$

such that the following correspondence holds:

$$I_F(x) \leftrightarrow I_M(T(x))$$

where  $\Omega_F$  is the fixed image domain with dimensionality  $d$  and  $\Omega_M$  the moving image domain. The image dimensionality  $d$  can be either two (2D) or three (3D). Hence, the transformation function can be a mapping from  $2D \rightarrow 2D$ ,  $3D \rightarrow 2D$  or  $3D \rightarrow 3D$ . Figure 10.1 shows an example of two images with overlap that can be registered in one coordinate system by translating and rotating one image (moving) to be aligned with the other (fixed) image.

Medical imaging consists of many modalities depending on the technique used to capture the images. The common modalities are X-ray, computed tomography (CT), magnetic resonant imaging (MRI), positron emission tomography (PET), single proton emission tomography (SPECT) and ultrasound (US) imaging. These modalities can be categorised into two main groups according to the purpose of the images (Oliveira and Tavares 2014): Anatomical images and functional images. CT, MRI, US and X-ray fall into the anatomical category where the purpose of the image is to study the anatomy, while functional MRI (fMRI), PET and SPECT fall into the functional category where the purpose is to study the functions performed by the imaged organ. Figure 10.2 shows brain images from the same subject captured using three different techniques (CT, MIT and PET). We can observe that each modality



**Fig. 10.2** (a) CT, (b), MRI, (c) PET images of the same subject. Images are from the RIRE dataset (<http://www.insight-journal.org/rire>)

provides a specific type of information. For example, MRI captures the soft tissue in detail whereas the bone structures are clearer in the CT image. A registration task may involve images from the same modality or different modalities. Depending on the number of modalities involved, the registration task can be classified as either mono-modal or multi-modal registration. In mono-modal registration both images are from the same modality (e.g. time series CT image registration for change detection), whereas in multi-modal registration the images are from two different modalities (e.g. fusion of CT with MRI images for image guided radiotherapy).

The solution to the registration problem generally consists of three main parts: transformation model, objective function and the optimisation procedure (Sotiras et al. 2013; Oliveira and Tavares 2014). Other than the key parts there are several other support modules such as the image interpolation, multi-resolution pyramids, pre-processing and feature extraction. The techniques used in each module of a given algorithm depend on the specific requirements of the given registration problem.

**Transformation Model.** Transformation model defines the nature of the geometric transformation between the two images. The selection of the model depends on the expected deformations in the imaged object and the selection critically affects the accuracy of the final registration outcome.

**Objective Function.** The objective function characterises the desired output of the registration process. Typically an objective function consists of two parts: the (dis-)similarity measure  $\mathcal{M}$ , which quantifies the alignment between images and the regularisation term  $\mathcal{R}$ , which constrains the solution domain. A typical objective function is as follows:

$$\mathcal{M}(I_F(x), I_M(T(x))) + \lambda \mathcal{R}(T) \quad (10.1)$$

where  $\lambda$  represents the trade-off between alignment and complying with the constraints. The choice of appropriate measure in objective function depends on the modalities involved in the given registration problem as well as the image quality, texture and deformation involved.

**Optimisation Algorithm.** The optimisation method tries to find the parameters of the transformation model that maximises (or minimises) the similarity (or dissimilarity) measure. In a very few cases using landmark based objective function the optimal parameters can be directly calculated but in general more sophisticated optimisation algorithms are needed to find the optimal transformation model.

In addition to the three key parts described above, there are several support modules involve in registration, such as image interpolation, hierarchical strategies and pre-processing.

**Interpolator.** The mapped point from one image domain to the other may not exactly coincide with grid position. Hence, an interpolator is used to calculate the intensity of the non-grid positions using grid values. The most common interpolation functions used in medical image registration are nearest neighbour, linear and B-Spline interpolators.

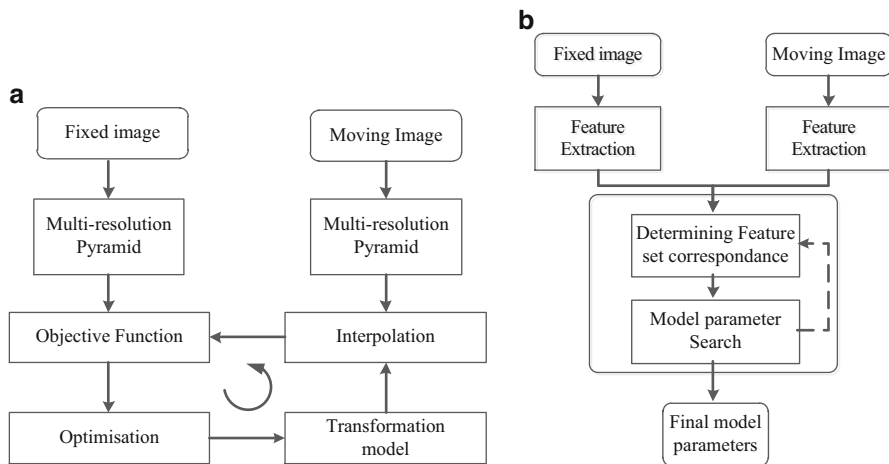
**Hierarchical Strategies.** The complex and nonlinear objective functions in image registration lead to many local minima. In order to escape from being trapped by one of these local minima, most registration algorithms use hierarchical approaches. These techniques can be classified into three main categories (Lester and Arridge 1999): increasing data complexity, increasing warp complexity and increasing model complexity.

**Image Pre-processing.** In order to get the best performance, the images for registration may need to be pre-processed before feeding them to the registration algorithm. Smoothing, motion correction, segmentation of relevant organ and normalisation are some of the pre-processing steps that can be performed. Although the registration algorithms vary widely, in general the flow of a registration algorithm can be summarised by flowcharts shown in Fig. 10.3.

### 10.1.1 Application of Image Registration in Medical Imaging

There are many applications of image registration in medical imaging field and some of those are listed below. A detailed description of those applications is outside the scope of this exposition and can be found in Holden (2008), Oliveira and Tavares (2014) and Rueckert and Aljabar (2010).

- *Disease Diagnosis:* The early detection is the most effective way to reduce mortality in many diseases. Computer Aided Diagnosis (CAD) systems are developed to enhance the capabilities of medical practitioners by attracting attention to certain areas in images, providing better comparisons and enabling



**Fig. 10.3** Registration flow of a typical intensity information based registration algorithm

medical practitioners to measure tumour sizes effectively. These systems can speed up the diagnostic process, reduce errors or misses and improve quantitative evaluation. Image registration plays a major role in CAD by enabling to combine information from multiple views, previous times, and multiple modalities (Giger et al. 2008). CAD is used to diagnose breast cancer (Wei et al. 2009), lung cancer (Guo et al. 2012; Shiraishi et al. 2006; Van Ginneken et al. 2001), Parkinson's disease (Illán et al. 2012), emphysema (Sluimer et al. 2006), brain tumour (Jaya and Thanushkodi 2011) and many other common diseases.

- *Progress monitoring*: Detecting and measuring changes of an anomalous region over time is vital to evaluate the success and to plan the future course of medication. But due to orientation changes, organ motions and anatomical changes during successive image acquisitions, direct comparisons are not possible. Image registration helps to reduce the effects of these deformations in order to make a realistic comparison (Ko and Betke 2001; Liang et al. 2013; Arzhaeva et al. 2010; Timp et al. 2007; Yuanjie et al. 2009).
- *Image guided surgery*: Image guidance provides valuable information about the target position and optimal probe manipulations in minimal invasive treatment procedures such as neurosurgery (Huang et al. 2012; Mostayed et al. 2013), endoscopy (Burschka et al. 2004; Mirota et al. 2012), interventional cardiology (Linte et al. 2010; Xishi et al. 2009), minimal invasive lung surgery (Sadeghi Naini et al. 2010; Uneri et al. 2013) and robotic microsurgery (Gerber et al. 2013). Here registration plays an essential role in combining and augmenting the high quality pre-operative images (CT, MRI, PET) with the intra-operative images from patient on the operating table (X-ray, US, interventional cone beam CT). A comprehensive review of registration techniques that are applied specifically for minimal invasive therapy can be found in Rui et al. (2013).

- *Image guided radiotherapy*: In radiotherapy generous safety margins have to be set around the target due to organ movements and target localisation errors. These margins lead to a compromise in the dose delivered to the tumour hence, accurate target localisation (defining the tumour volume) and alignment is very important in order to maximise the radiation dose to the target while minimising the effects on the surrounding organs. Image guided radiotherapy techniques use on-line or off-line image registration. Off-line registration of images from different modalities (CT/MRI, CT/PET) can improve target localisation whereas organ motion can be determined by on-line registration of repeatedly acquired images. This information is then used to adapt the treatment (patient positioning, directing radiation beam) (van Herk 2007; Robertson et al. 2013; Worm et al. 2010).

In medical imaging applications the deformations involved in the images are often highly nonlinear. Also the relationships between the voxel intensities of the same structure in two different images can be nonlinear, especially for multi-modal registration. Due to these reasons the resulting optimisation is highly nonlinear and the main objective of this chapter is to explore the use of nonlinear methods in medical image registration applications.

Section 10.2 explores the nonlinear similarity measures used in medical image registration while Sect. 10.3 investigates the transformation models and Sect. 10.4 the solution to nonlinear registration problem. Finally Sect. 10.5 concludes the chapter.

## 10.2 Similarity Measures

A similarity measure is part of the objective function which quantifies the alignment between the images being registered. Registration algorithms can be broadly classified into two main groups, depending on the nature of the information used to calculate the similarity measure: *extrinsic* or *intrinsic*. Extrinsic methods use artificial markers attached to the patient, prior to being imaged. These markers are visible in all imaged modalities and will lead to efficient registration methods where parameters can be calculated explicitly. However, due to the invasiveness of attaching markers, restrictions on the number of markers that can be attached and the difficulty in capturing non-rigid motions, these methods are limited to rigid registration and pose estimation (Maintz and Viergever 1998). In contrast, intrinsic methods rely only on information readily available in the image. Depending on the feature space used in the similarity measure, registration algorithms are broadly classified into three main categories (Sotiras et al. 2013): image feature based methods, voxel information based methods and hybrid methods. Feature based methods use landmarks or features extracted from the images in the registration process while the voxel information based methods use the voxel information directly. Hybrid methods are a combination of the two techniques.

### 10.2.1 Voxel Information Based Similarity Measures

In voxel information based registration, the similarity measure takes the intensity information of the two images as input and outputs a value that relates to how well the two images are aligned. It is important to note that in medical image registration the images can be from the same modality (mono-modal registration) or they can be from different modalities (multi-modal registration).

Firstly, we consider the mono-modal registration case where the images are acquired using the same imaging technique. Examples of this type of registration would be MRI–MRI registration or CT–CT registration. The simplest similarity measure is based on the assumption that corresponding anatomical structures in both images have similar voxel values except for Gaussian noise and this measure is named sum of square difference (SSD) similarity measure. The mathematical representation of this measure is as follows:

$$\mathcal{M}_{SSD}(\theta) = \frac{1}{N} \sum_{x \in \Omega_F \cap \Omega_M} (I_F(x) - I_M(T(x)))^2 \quad (10.2)$$

where  $N$  is the number of voxels in the overlapping region of these images. In cases when the images contain some voxels with large errors (outliers) SSD would fail, as these high error values skew the objective function. This can be common in medical imaging due to the injection of contrasting agents or position changes of the patient in two time-lapse images (Hill et al. 2001). This problem can be mitigated by using the sum of absolute difference (SAD) similarity measure instead, which is more robust to outliers.

$$\mathcal{M}_{SAD}(\theta) = \frac{1}{N} \sum_{x \in \Omega_F \cap \Omega_M} |I_F(x) - I_M(T(x))| \quad (10.3)$$

The assumption that the voxel values are identical except for Gaussian noise is very restrictive and a more realistic assumption would be to hypothesise a linear relationship between intensities. For this assumption the optimal similarity measure is cross correlation or normalised cross correlation of image intensities which can be represented as:

$$\mathcal{M}_{CC} = \frac{\sum_x \{(I_F(x) - \mu_F)(I_M(T(x)) - \mu_M)\}}{\sqrt{\sum_x (I_F(x) - \mu_F)^2 \sum_x (I_M(T(x)) - \mu_M)^2}} ; x \in (\Omega_F \cap \Omega_M) \quad (10.4)$$

In the above measure  $\mu_F$  and  $\mu_M$  are the mean intensities of the fixed and the moving images, respectively. It is important to note here that even though the relationship between the intensity values of the two images is assumed to be linear, the resulting cost functions will be nonlinear.

In some medical image registration applications, where the organ of interest is elastic, the assumption of linear relationship between corresponding voxels is unrealistic. Examples of this scenario would be the lung CT images. The density of tissue, which is measured by CT, may vary depending on the volume of the air inside the lung. Hence the intensity of the same anatomical structure would change during the breathing cycle. In order to overcome this problem specifically in lung image registration there have been methods developed that assume a nonlinear relationship between voxel values, using the conservation of mass principle. Gorbunova et al. (2012) assumed that the mass of parenchyma remains constant and the density of the lung tissue is inversely proportional to the amount of air. The change in volume was calculated using the determinant of the Jacobian (Yin et al. 2009) and this was integrated to the model to obtain the following cost function.

$$\mathcal{M}_{MP} = \frac{1}{N} \sum_{x \in (\Omega_F \cap \Omega_M)} (I_F(x) - |J_T(x)| \cdot I_M(T(x)))^2 \quad (10.5)$$

where  $|J_T(x)|$  is the determinant of the Jacobian of the transform model.

Castillo et al. (2009, 2012) used the principle of mass preservation in an optical flow setting. They incorporated the mass preservation assumption, which is,  $\int_{\Omega} I_F(x) dx = \int_{\Omega} I_M(x) dx$ , in an optical flow equation:

$$I_F(x) + \nabla \cdot I_M(x) \cdot v + I_M(x) \text{div}(v) = 0 \quad (10.6)$$

where  $v$  is the velocity field in the image. The integrated form of this equation (Corpetti et al. 2002) leads to the following compressible flow similarity measure:

$$\mathcal{M}_{CF} = \frac{1}{N} \sum_{x \in N(x)} (\ln(I_F(x)) - \ln(I_M(T(x))) + \text{div}(T(x) - x))^2. \quad (10.7)$$

It is important to note that most similarity measures such as SSD and SAD assume independence between intensities from voxel to voxel. This means that if the same permutation of voxel indices is applied to both images the resulting similarity measure will be unchanged. To account for the spatially varying intensity distortions, Myronenko and Xubo (2010) developed a new similarity measure named residual complexity. To arrive at the correct form of this measure, they first introduced an intensity correction field that represents the differences of intensities of the registered images. By analytically solving for this correction field and using adaptive regularisation, they derived the following similarity measure:

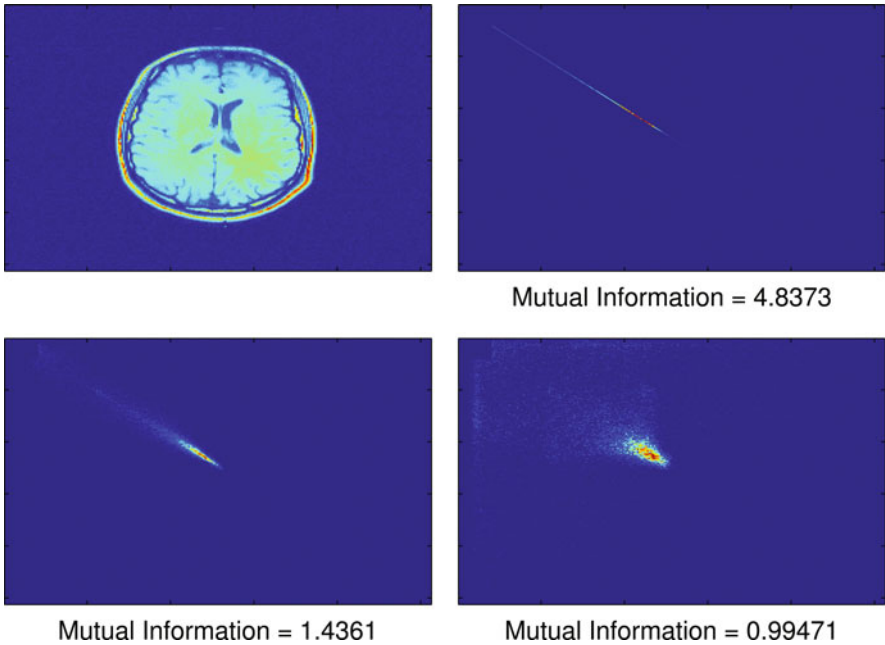
$$\mathcal{M}_{RC} = \sum_{n=1}^N \log \left( \frac{(q_n^T r)^2}{\alpha} + 1 \right) \quad (10.8)$$

where  $r = (I_F(x) - I_M(T(x)))$  and  $q_i$  is an eigenvector of  $Q$  where  $Q\Lambda Q^T$  is the spectral decomposition of  $P^T P$  where  $P$  is the regularisation operator of the correlation field.

Secondly, we will consider the multi-modal image registration. Here, the voxel information of the two images (to be registered) is generated through different physical processes and as such the relationship between the voxel intensities of the same structure can be complex. For this kind of applications the use of information theory measures such as Mutual Information (MI) proposed by Viola and Wells III (1997) and Collignon et al. (1995) has gained a lot of attention. A detailed review of these techniques was provided by Pluim et al. (2003). In these methods only a probabilistic relationship between the voxel values is assumed and the mutual information is designed to measure how well one image explains the other image. The similarity measure based on mutual information can be expressed as:

$$\mathcal{M}_{MI} = H(I_F) + H(I_M^T) - H(I_F, I_M^T) \tag{10.9}$$

where  $H(I)$  is the Shannon entropy of the image and  $H(I_F, I_M^T)$  is the joint entropy of two images. Figure 10.4 shows the joint histograms between a brain MRI image and a rotated version of itself. These plots show that when the images are perfectly registered the joint histogram is concentrated whereas when the mis-registered



**Fig. 10.4** Joint histograms between the original image (top-left) and rotated versions of it. (Top-right) No rotation, (bottom-left) rotated by  $2^\circ$ , (bottom-right) rotated by  $15^\circ$



increases they become dispersed. The mutual information value for each case is also shown in this figure and we can see that the more the images aligned, the larger the mutual information values.

The above measure does not work well when the overlap of two images is low. To solve this problem, the normalised mutual information (10.10) was proposed (Studholme et al. 1999).

$$\mathcal{M}_{NMI} = \frac{H(I_F) + H(I_M^T)}{H(I_F, I_M^T)} \quad (10.10)$$

Cahill et al. (2008) further analysed the overlap invariance of several information theory based similarity measures including NMI and found that in some cases those methods would still fail. To make the NMI measure overlap invariant, they introduced a modified form of NMI.

Another drawback of mutual information based methods is that it does not consider the relevant spatial information. As a result, MI based similarity measures may face problems in presence of local structures or spatially varying intensity distortions. Various techniques have been proposed that integrate spatial information in the MI cost function. For instance, Plum et al. (2000) proposed a new measure that combined mutual information with spatial information given by gradient magnitude and orientation. The new measure was constructed by simply multiplying the NMI cost function with proposed gradient measure. In contrast, Rueckert et al. (2000) incorporate the spatial information using the second order entropy and joint entropy in the calculation of NMI.

Another approach to include spatial information in the MI cost function is to use local estimation of the mutual information by progressively subdividing the images (Likar and Pernuš 2001; Andronache et al. 2006). Studholme et al. (2006) combined these regional terms into one global cost function called Regional Mutual Information (RMI). RMI incorporates the region information describing the global relationship between intensities, as a third channel of information (in addition to the two intensity channels) in MI calculation. Using additional channels in MI calculation is computationally expensive and Klein et al. (2008) proposed to reduce the computational complexity by using random sampling. Loeckx et al. (2010) also incorporated a third spatial dimension to create the conditional mutual information (cMI) similarity measure.

Local structural information can also be added to MI computation using the graph based  $\alpha$ -MI techniques (Hero et al. 2002; Staring et al. 2009). Recently, Rivaz and Collins (2012) used self-similarity, which measures the similarity of one image patch to another patch in the same image in a weighted graph based  $\alpha$ -MI implementation for non-rigid image registration.

Calculation of mutual information involves some nontrivial and sensitive smoothing parameters for density approximations and the resulting cost function is highly non-convex and contains many local minima. In order to overcome these limitations, (Haber and Modersitzki 2006; Modersitzki 2008) proposed the use of

Normalised Gradient Fields (NGFs). The rationale behind this is that if the images are mis-registered the angle between the two gradient fields at a given point will be non-zero. The resulting similarity measure can be expressed as:

$$\mathcal{M}_{NFG} = \frac{1}{2} \sum_{x \in \Omega} \|n_\epsilon(I_F, x) \times n_\epsilon(I_M, T(x))\| \quad (10.11)$$

where

$$n_\epsilon(I, x) = \frac{\nabla I(x)}{\sqrt{\nabla I(x)' \nabla I(x) + \epsilon^2}} \quad (10.12)$$

Ruhaak et al. (2013) used this similarity measure in registration of full-body CT and PET images.

## 10.2.2 Image Feature Based Similarity Measures

The next class of image registration algorithms is based on matching image features. These methods first identify anatomically important point, curve or surface landmarks of the images. Once the correspondences are established between the landmarks, a similarity measure (SSD) that minimises the Euclidian distance between points can be used to calculate the model parameters. A detailed review of feature based registration methods was provided by Audette et al. (2000). The interest in the use of pure feature based methods has gradually diminished since the landmark or feature extraction processes introduce some errors that would propagate through the algorithm. Moreover, the need for extrapolation/interpolation from a sparse set of landmarks to recover the dense deformation field means that the estimation of those fields away from the landmarks will be unreliable (Sotiras et al. 2013).

Intensity based methods only account for intensity patterns and do not care about the structure while feature based methods only use sparse landmarks. The hybrid algorithms combine the two in order to achieve more robust registration results (Rueckert and Aljabar 2010).

## 10.3 Transformation Model

The transformation model in image registration context is a spatial mapping from one image domain to another. The transformation model that best matches a given registration application depends on the underlying deformations that are to be recovered. There are many transformation models that are used in medical image

processing. A transformation model or function can be either linear or nonlinear. The definition of a linear transformation function is as follows:

*Definition* Let  $V$  and  $W$  be two vector spaces over the same scalar field  $\mathbb{F}$ . A function

$$T : V \rightarrow W$$

is said to be a linear transformation if the following two properties hold for any two points  $x, y \in V$  and scalar  $\lambda \in \mathbb{F}$  (Rynne and Youngson 2008):

$$T(x + y) = T(x) + T(y)$$

$$T(\lambda x) = \lambda T(x).$$

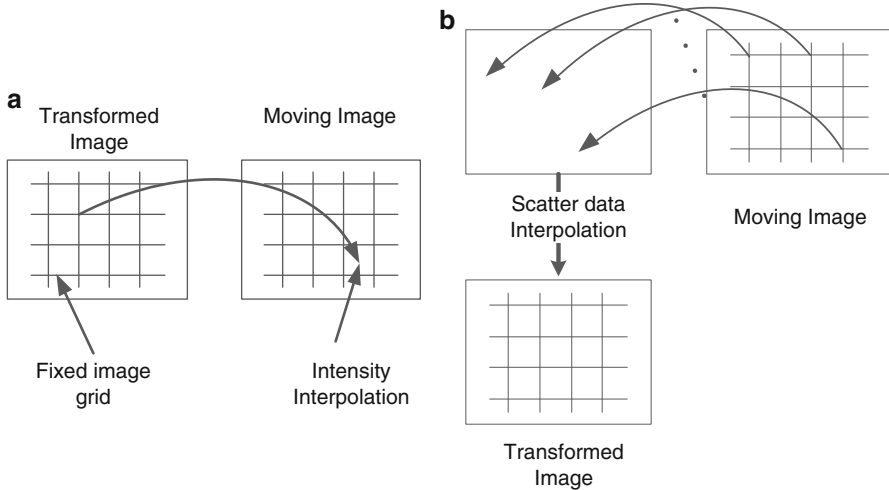
In general linear transformations map a line to a line or a point. In contrast nonlinear transformations map lines to curves. In most cases of medical image registration the nature of the deformations is nonlinear due to the non-rigid nature of organs and muscles. Hence, transformation functions that are used are mostly nonlinear.

In image registration, the objective is to map the moving image to the coordinate system of the fixed image which is called the forward transformation. However, most algorithms due to practical advantages calculate the inverse transformation (transformation from fixed image to moving image). The rationale behind this choice (Sotiras et al. 2013) is that if a forward transformation is calculated, in order to generate the transformed image each grid point in  $\Omega_M$  needs to be transformed using the calculated transformation function followed by a scattered data interpolation and due to the large number of points the computational cost of the interpolation would be huge. On the other hand, if the inverse transformation is calculated, the values of transformed image can be calculated by transforming each point in the new grid to  $\Omega_M$  and picking the relevant value from the moving image (when the transformation leads to a non-grid position in the moving image, intensity interpolation can be done using neighbouring voxels), which has far less computational workload. Figure 10.5 further elaborates this point.

Transformation models can be first grouped into two main categories: rigid and non-rigid. The simplest type of transformation function used in medical imaging is the rigid transformation. Rigid Transformation is a combination of a rotation and a translation. In three-dimensions, rigid motion can be represented by three rotation and three translation parameters. Rigid registration can be used to match images that are of bone structures or enclosed by bony structures.

The rigid transformation falls within the broader category of affine transformations. This can be expressed using the following equation:

$$x'_{[d \times 1]} = A_{[d \times d]} \cdot x_{[d \times 1]} + b_{[d \times 1]} \quad (10.13)$$



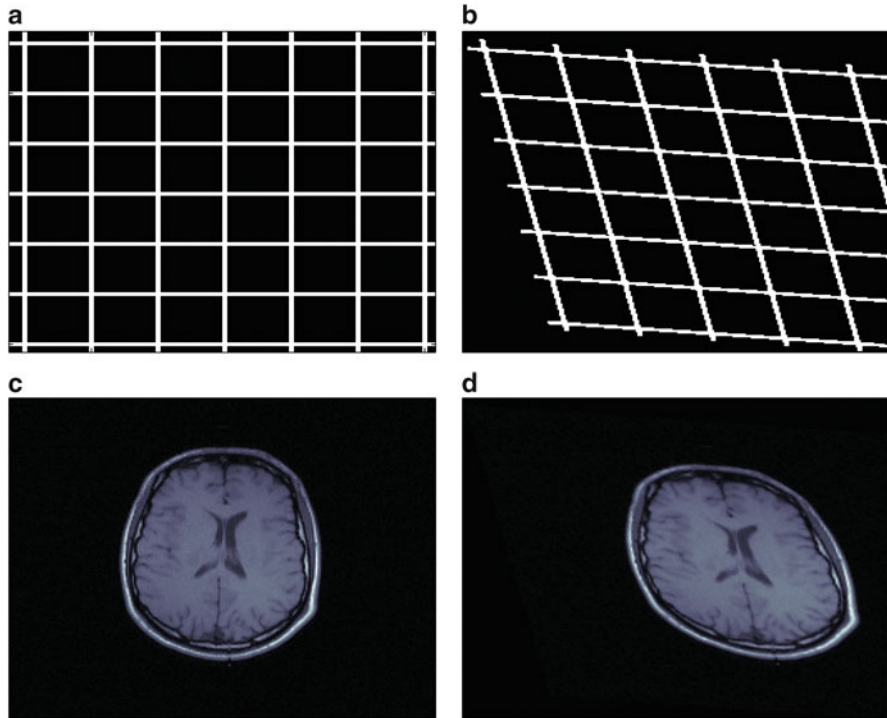
**Fig. 10.5** Inverse transformation vs. forward transformation. The figure shows that using inverse transform is computationally advantageous. (a) Inverse transform. (b) Forward transform

Affine transformation consists of a linear transformation ( $A \cdot x$ ) representing rotation, scaling and shear and a translation vector ( $b$ ). Except for the special case where the translation part is zero the affine transformation in general is nonlinear, contrary to its appearance that has caused some authors to classify it as linear transform (Glocker et al. 2011; Lester and Arridge 1999). Due to the scaling and shear components the general affine transformation is also categorised as non-rigid. Affine registration is sometimes used in cases where images are distorted due to physical characteristics of the acquisition system. It is also extensively used as a pre-registration technique in complex registration algorithms with large degrees of freedom (Fig. 10.6).

The limited degrees of freedom allowed by affine or the rigid transformation is not adequate to represent the complex deformations involved in medical image registration. Hence, non-rigid registration models with much higher degrees of freedom are introduced. Holden (2008) and Sotiras et al. (2013) categorised those non-rigid registration algorithms into three main groups: transformations derived from physical models, transformations derived from interpolation theory and constraints on transformation.

### 10.3.1 Transformations Derived from Physical Models

These models use the characteristics of physical phenomena such as elastic material and viscous fluid flows to model the deformations in medical images.



**Fig. 10.6** Original image (c) and the image after an affine transform (d). The transformation of both the grid (a, b) and the intensities are shown

### Elastic Model

These models assume that the images behave similar to elastic bodies under deformation. Here, the external forces due to the mismatch in images (driving forces in registration are calculated using the similarity measure) are balanced with the internal forces that impose smoothness in the deformation field.

The elastic deformations of an object can be represented using Navier–Cauchy equation:

$$\mu \nabla^2 u(x) + (\lambda + \mu) \nabla (\nabla^T u(x)) + f(x) = 0 \quad (10.14)$$

where  $f(x)$  is the external force that drives the registration, which is the similarity measure,  $u(x)$  is the deformation at point  $x$ ,  $\mu$  quantifies the stiffness of the material and  $\lambda$  is the Lamé's first coefficient.

This Navier–Cauchy partial differential equation leads to an optimisation problem which can be solved in many ways such as: variational, finite difference, finite element, basis function expansions and Fourier transform methods (Modersitzki 2004; Broit 1981; Gee et al. 1994; Christensen et al. 1994).

One of the drawbacks in using this model is that the resulting transformation is not inversely consistent. This can be corrected either by adding a constraint that penalises the inverse inconsistency (after calculating forward and backward transformations) or by coupling the forward and backward transformations in deriving the final transformation. Detailed description of this is provided later when constraint on transformation is explained.

In the linear elastic model explained above, the second order terms of the displacement field gradients are ignored. This means that the relationship between stress and strain is assumed to be linear. This makes this model incapable of handling large deformations. To overcome these issues nonlinear elastic models were developed. These models were based on hyperelastic material properties (Rabbitt et al. 1995) and St Venant–Kirchoff elasticity energy (Pennec et al. 2005).

## Fluid Flow

In fluid flow models the images are assumed to behave as viscous fluids under deformation. The advantage of these models is that they do not assume small deformations.

The fluid flow transformation model can be represented by the following Navier–Stokes partial differential equation.

$$\mu \nabla^2 v(x) + (\lambda + \mu) \nabla (\nabla^T v(x)) + f(x, u) = 0. \quad (10.15)$$

This equation is very similar to the elastic model equation in the previous section. The only difference is that in fluid flow model a velocity field is used instead of the deformation field. In the above equation the first term imposes spatial smoothness of the velocity field whereas the second term allows for expansions and contractions.  $\lambda$  and  $\mu$  are the viscosity coefficients.

Finding the solution to fluid flow models involves solving a large set of partial differential equations. The earliest implementations used successive over-relaxation which is computationally expensive (Christensen et al. 1994). Christensen et al. (1996) used parallel computations to reduce the computational time. Some faster ways of solving differential equations such as scale space convolution with a filter (impulse response of the regularisation operator) (Bro-Nielsen and Gramkow 1996) have also been proposed.

### 10.3.2 Transformations Derived from Interpolation Theory

In the transformation models that belong to this category the displacement field is modelled using functions derived from interpolation and approximation theories. The requirement to intercept sample values is relaxed in approximation assuming errors in sample values (Holden 2008). In these methods firstly the basis function

centre points (or control points) are selected and then the parameters of each basis function are adjusted (or calculated) to fit the deformation field.

### Radial Basis Functions

Radial basis functions (RBFs) are the most commonly used interpolation theory based transformation model in landmark based image registration. The value of an RBF only depends on the distance from its control point. Several such basis functions placed on the spatial domain of the image can be used to model the deformation field of the image effectively. The transformation function based on RBF can be defined as follows:

$$T(x) = x + \sum_{k=1}^N \theta_k R(\|x - x_k\|) \tag{10.16}$$

where  $\theta_k$  are the parameters that control the shape of the deformation field,  $x_k, k \in 1 \dots N$  are the RBF control points and  $R(\cdot)$  is the RBF kernel. One of the advantages of RBFs is the ability to place the control points in an irregular grid which enables modelling non-uniform deformations with less parameters than a grid-based control point placement scheme.

In landmark based registration, the RBF is combined with a polynomial basis. The polynomial basis  $\phi(x)$  is added to take care of the more global deformations while the RBF part highlights the local changes. This transformation function can be as follows:

$$T(x) = x + \sum_{k=1}^N \theta_k R(\|x - x_k\|) + \sum_{j=1}^m \theta_{N+j} \phi_j(x) \tag{10.17}$$

Once the problem is in this setting, if the landmarks and their correspondences are available a closed form solution can be found as the RBFs are positive definite. Some of the RBFs commonly used in literature are given in Table 10.1. Here we note that these RBFs are nonlinear.

The RBF described above has a global support region or in other words a given control point effects the entire deformation field. This property has adverse effects when there are outliers or the desired deformation field is local.

**Table 10.1** Commonly used radial basis function kernels where  $r = \|x - x_i\|$

Thin plate spline	$r_i^2 \log r_i$
Multi-quadratic	$\sqrt{r_i^2 + d^2}$
Inverse multi-quadratic	$(r_i^2 + d^2)^{-1/2}$
Generalised multi-quadratic	$(r_i + c_1^2)^{c_2}$
Gaussian	$e^{-(r_i/c_1)^2}$

In medical image registration a given image may contain several local motions due to the presence of different anatomical structures. In order to recover these types of local motions RBFs that have compact spatial support are proposed. Examples of these functions are Wendland Function and Wu's compact support RBF (Rohde et al. 2003).

### Free Form Deformation

The main idea behind free form deformation (FFD) based image registration is to match two images by deforming one image via manipulating a grid of control points superimposed over that image. FFD is the most widely used transformation model in medical image registration. Typically the control point grid is a rectangular grid with arbitrary resolution which is coarser than that of the image. The control points are moved in such a way that a similarity measure is maximised and the deformations at each voxel are then calculated by interpolation using local supported kernels such as B-Splines.

A B-Spline function is a compact support kernel which makes it suitable for modelling deformation fields with local motions. The B-Spline kernel of order zero is a rectangle function and the higher order kernels are calculated by convolution of the zero order kernel with itself  $N$  times ( $N$  is the degree of the kernel). The first four B-Spline kernels are shown in Fig. 10.7.

These spline kernels are stitched together over the image domain in a rectangular grid in order to model the deformations of the image. Given the compact support of the B-Spline kernel, the deformation of a particular point represented by a B-Spline would only depend on several control points in the immediate neighbourhood of that point. Using this property, for 3D control point grid with spacing  $[\delta_x, \delta_y, \delta_z]$  the

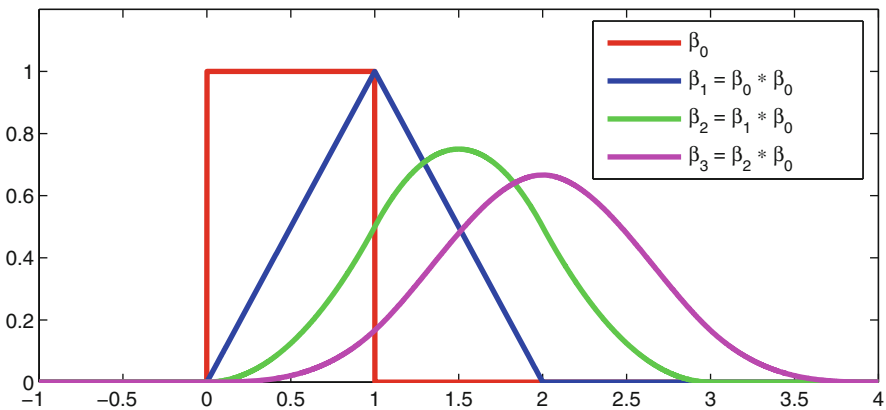


Fig. 10.7 B-Spline kernels



multivariate extension of cubic B-Splines based FFD transformation model can be written as follows (Rueckert et al. 1999):

$$T(x) = x + \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_{l,3}(u) B_{m,3}(v) B_{n,3}(w) d_{i+l, j+m, k+n} \quad (10.18)$$

where  $i = \lfloor x/\delta_x \rfloor - 1$ ,  $j = \lfloor y/\delta_y \rfloor - 1$ ,  $k = \lfloor z/\delta_z \rfloor - 1$ ,  $u = x/\delta_x - \lfloor x/\delta_x \rfloor$ ,  $v = y/\delta_y - \lfloor y/\delta_y \rfloor$  and  $w = z/\delta_z - \lfloor z/\delta_z \rfloor$ .

The cubic B-Spline functions in the above equation are:

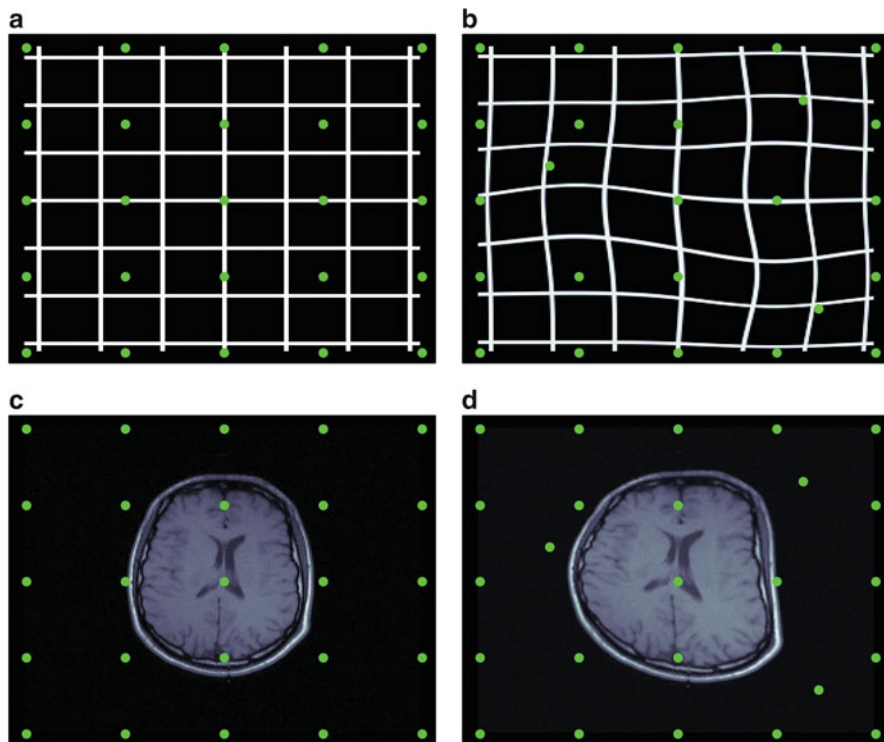
$$\begin{aligned} B_{0,3}(u) &= \frac{(1-u)^3}{6} \\ B_{1,3}(u) &= \frac{(3u^3+6u^2+4)}{6} \\ B_{2,3}(u) &= \frac{(-3u^3+3u^2+3u+1)}{6} \\ B_{3,3}(u) &= \frac{(u^3)}{6} \end{aligned} \quad (10.19)$$

The B-Spline kernels are also  $\mathcal{C}^2$  smooth at the control points which provide continuity when the control points are moved. This enables the B-Spline transformation model to generate smooth deformations. An example of a B-Spline transformation is shown in Fig. 10.8 where the control points are placed in a regular grid.

Rueckert et al. (1999) used the B-Spline motion model after an affine pre-registration step for non-rigid registration of contrast enhanced breast MRI images. In their work NMI was used as the similarity measure. Kybic and Unser (2003) used the B-Spline deformation model coupled with SSD similarity measure to register images from MRI, SPECT, CT and ultrasound modalities. Their method also included a soft landmark based constraint to guide registration in difficult scenarios.

The control node displacements act as parameters in an FFD. Hence, the details captured by the deformation field depend on the resolution of the control point grid. Large spacing enables the modelling of global deformations whereas small spacing represents highly local deformations. The number of control points also directly affects the computational complexity of the resulting optimisation problem. The optimal control point grid is decided based on the model flexibility and computational complexity.

The B-Spline based FFD framework does not guarantee the preservation of topology and there are several methods that can be employed to rectify this issue. Noblet et al. (2005) enforced topology preservation by controlling the Jacobean of the transformation and Kim (2004) suggested a sufficient condition for local invariability. Chun and Fessler (2008) extended this and proposed a new penalty based approach for topology preservation B-Spline transformation.



**Fig. 10.8** Original image and the image after a B-Spline transform. The transformation of both the grid and the intensities are shown together with the control point locations

Although in the basic implementation of B-Spline based FFD the control point grid is uniform there has been some work that employs a non-uniform grid (Wang and Jiang 2007; Jacobson and Murphy 2011).

### Wavelet and Fourier

Fourier series and wavelet functions are extensively used to represent signals. In Fourier representation the signal is characterised as a combination of sinusoidal signals which localise the signal in frequency domain. On the other hand, the wavelet representation localises the signal in both frequency and spatial domain making it ideal to represent localised deformations (Wu et al. 2000).

One of the major attractions of these techniques in modelling deformation fields is their ability to decompose in to multi-resolution.

## Local Affine

In locally affine transformation model, the image domain is partitioned into independent blocks and for each block a rigid or affine transformation is calculated by maximising a voxel based similarity measure (Pitiot et al. 2003). An issue with this local transformation is that it may not be continuous at the block edges. In order to overcome this issue regularisation can be performed using either nonlinear interpolation or filtering. Another possible solution is to use fuzzy regions instead of the definitive independent blocks (Arsigny et al. 2005).

One of the parameters needed to be selected is the size of the block. This parameter controls the degrees of freedom in the registration and if a small block size is selected there will be inadequate information inside the block to find the registration while a large block size may not capture local deformations. There are several algorithms present in literature that can automatically calculate the appropriate block size using image information such as structural information, residual error and motion.

### 10.3.3 Constraints on Transformation

Constraints on transformation model limit the available solution space by imposing constraints based on prior knowledge of the physical properties of the objects in the registration problem.

#### Inverse Consistency

Given two images  $I_F$  and  $I_M$  to be registered, a basic assumption would be that the transformations that map moving to fixed images  $T_{MF} : (I_M \rightarrow I_F)$  and fixed to moving images  $T_{FM} : (I_F \rightarrow I_M)$  are directly related:

$$T_{MF} = T_{FM}^{-1} \quad (10.20)$$

However, most of the existing algorithms are asymmetric, which violate the above condition. Hence, the accuracy of these algorithms may vary depending on which image is selected as the fixed image.

One option in overcoming this issue is to add an additional term to the objective function that penalises the difference between the forward and reverse transformation (Christensen and He 2001).

$$\mathcal{R}_{IC} = \sum_{x \in \Omega} (T_{MF}(x) - T_{FM}^{-1}(x)) \quad (10.21)$$

Another solution is to transform both images to the same domain and construct the final transform by taking the inverse of one part and adding it to the other.

### Topology Preservation

The aim of topology preservation is to keep the connected structures together and maintain the neighbourhood relationship between structures. This will prevent the disappearance of existing structures and the appearance of new structures in the image (Musse et al. 2001). In order to preserve topology the transformation function must be continuous, bijective and the determinant of the Jacobean must be positive.

### Diffeomorphic Transformation

A diffeomorphic transformation maps a differentiable manifold to another differentiable manifold such that both the function and its inverse are differentiable. This means that a diffeomorphic transformation has a unique inverse which maps that fixed image back to the moving image. The diffeomorphic assumption is violated in some cases such as images before a surgery and after where some anatomical structures are changed (Christensen 1999). Diffeomorphic mapping also guarantees topology preservation.

Continuum mechanical methods such as linear elasticity and viscous fluid were used to enforce diffeomorphism by Christensen et al. (Christensen 1999; Christensen et al. 1996).

Vercauteren et al. (2007) used the demons algorithm combined with Lie group framework on diffeomorphism and an optimisation procedure on Lie groups to provide non-parametric free-form diffeomorphic transformation.

## 10.4 Optimisation

The optimisation procedure tries to find the transformation model that best aligns the images according to an objective function. The objective function  $\Psi$  employed by most image registration scenarios is nonlinear as well as non-convex.

$$\Psi(\theta) = \mathcal{M}(I_F(x), I_M(T_\theta(x))) + \mathcal{R}(T_\theta) \quad (10.22)$$

Here  $\mathcal{M}$  is the similarity measure which quantifies the alignment between the two images and  $\mathcal{R}$  is a regularisation term that incorporates a prior known property of the deformation. The similarity form for most problems is organised such that when the alignment is at the highest the value of the similarity measure is the minimum (used like a dissimilarity measure).

This leads to a nonlinear optimisation problem of the form:

$$\hat{\theta} = \arg \min_{\theta} [\Psi(\theta)] \quad (10.23)$$

where  $\theta$  are the parameters of the transformation and  $\hat{\theta}$  is the optimal parameters that best align the images.

In some cases the parameters of the registration can be calculated directly by solving a set of linear equations. An example of such a scenario is landmark (with correspondences) based registration using RBF.

Other than these specific cases most registration problems need an iterative approach to find the optimum transformation model. These optimisation algorithms can be categorised into two main parts: continuous optimisation and discrete optimisation.

### 10.4.1 Continuous Optimisation

The continuous optimisation methods can be used in problems where the parameters are real valued and the objective function is differentiable. These methods update the parameters using an iterative method that follows the equation:

$$\theta_{n+1} = \theta_n + \alpha_n g_n(\theta_n) \quad (10.24)$$

where  $\theta$  are the parameters,  $n$  is the index of the current iteration,  $\alpha_n$  is the step length and  $g_n(\theta_n)$  is the search direction. There are several methods that belong to this category and they differ based on the method used to calculate step length and search direction.

In gradient descent (or steepest descent) the cost function gradients at the current position is selected as search direction. Once the direction is set, a line search method or a gradually decreasing function can be used as the step length (Klein et al. 2007). The gradient decent descent has a low convergence rate. The nonlinear conjugate gradient method provides a better convergence where the search for direction is along the gradient component from which previous gradient direction is excluded. In both of the above algorithms the knowledge of a mathematical model is needed in order to calculate the gradients. However, in Powell's conjugate gradient method no such information is needed as it does not calculate gradients.

In Quasi-Newton methods Hessian of the objective function is set as the step length. Because the calculation and inversion of the Hessian is computationally expensive an approximation of the Hessian is used commonly. Some well-known methods to approximate the Hessian are Symmetric-Rank-1, Davidon-Fletcher-Powell and Broyden-Fletcher-Goldfarb-Shanno.

Gauss-Newton method is a gradient based optimisation method that can be employed when the objective function is in the form of sum of squares. Calculation

of the Hessian, which is computationally expensive, is not needed for this method. Here Hessian is approximated using the Jacobean and the search direction is given by  $(J^T(\theta)J(\theta))^{-1} \nabla_{\theta}$ . Levenberg–Marquardt method is also used to optimise a function defines as sum of squares but here the parameter updates are adaptively changed between gradient descent update and Gauss–Newton update. The search direction is given by  $(J^T(\theta)J(\theta) + \lambda \text{diag}(J^T(\theta)J(\theta)))^{-1} \nabla_{\theta}$ . The advantage of this method is that it can converge to the solution starting form a location far away from the final solution.

Another very popular method in medical image is the stochastic optimisation where an approximation of the gradient is used as the search direction together with a diminishing step size. The gradient approximation can be calculated either using analytical gradients using only a subsample of points, Kiefer–Wolfowitz method or Simultaneous Perturbation. Klein et al. (2007) showed that out of the above methods analytical gradients calculated using only a subset of points selected by uniform sampling is the most efficient. Uniform sampling gives equal significance to all the voxels in an image and since the true function in this case is not distributed uniformly, using only a subset of voxels produces a biased estimate of the gradients. To reduce the sampling bias they proposed to renew the set of chosen samples at every iteration of the optimisation routine giving equal significance to all the voxels. To improve the registration accuracy Bhagalia et al. (2009) introduced the idea of using the importance sampling (IS) technique to select a subset of voxels. Tennakoon et al. (2014) proposed the use of rank ordered statistics based robust segmentation technique on the intensity difference image to identify the relevant voxels for registration and used these in calculating the gradients.

## 10.4.2 Discrete Optimisation

Discrete optimisation schemes are used when the parameters of the optimisation problem are in discrete forms. For instance, the image registration problem can be seen as a discrete optimisation problem when posed as a metric labelling problem. However generally this will result in a non-convex energy with many local minima and the optimisation would involve a large set of parameters in which finding a solution is computationally expensive. Kleinberg and Tardos (2002) have shown that original registration problem when posed as metric labelling problem is equivalent to a discrete Markov Random Fields (MRFs) minimisation problem, where the aim is to assign each node (voxel or parameter) a label  $l \in \mathcal{L}$  (where  $\mathcal{L}$  is a discrete label set) which minimises an energy function  $E(l)$  of the following form.

$$E(l) = \underbrace{\sum_{p \in P} D_p(l_p)}_{\text{Data Term}} + \lambda \underbrace{\sum_{\{p,q\} \in N} V_{p,q}(l_p, l_q)}_{\text{Smoothness Term}} \quad (10.25)$$

In this equation  $D_p(l_p)$  is the cost of assigning label  $l_p$  to node  $p$  and  $V_{p,q}(l_p, l_q)$  is the cost of assigning labels  $l_p, l_q$  to neighbouring pixels  $p, q$ . Recently many techniques emerged that can solve this problem efficiently and these methods can be segmented into three main groups: Graph-cut based methods, message passing methods and linear programming methods.

Graph-cut based methods use the max-flow min-cut principle (Ford and Fulkerson 2010). Here a graph is constructed where each node is connected to its neighbours and two special nodes called the source and the sink. Once the graph is constructed the minimum cut that segments this graph into two partitions is given by the saturated edges at the maximum flow from source to sink. Boykov et al. (2001) extended this idea to multiple labels. They introduced two algorithms, the  $\alpha$ - $\beta$ -swap and  $\alpha$ -expansion for energies where  $V_{p,q}(\cdot)$  is semi metric and metric, respectively. These methods use large moves compared to standard algorithms making them computationally efficient and also there is a theoretical guarantee of convergence within a known factor (vicinity) of the global minimum. So and Chung (2009) and So et al. (2011) formulated the image registration problem as a discrete labelling problem and assigned each voxel a discrete displacement label. They used an energy function which comprised of SAD as the data term and a smoothness term that penalises the difference in displacement in adjacent voxels. They used  $\alpha$ -expansion algorithm to derive the optimal labelling.

Graph-cut methods can only be applied to a limited set of energy functions (Kolmogorov and Zabini 2004). However, the MRFs problem discussed above can be at least approximately solved using the max-product Belief Propagation (BP) technique introduced by Pearl (1988). BP uses local messages passed along the edges of a graphical model to derive the solution. The solution will only be exact in case the graph is a tree, but if the graph contains loops only an approximate solution can also be found. Generalised BP algorithm proposed by Yedidia et al. (2000) achieves better convergence than the ordinary BP. Max-product tree-reweighted (TRW) message passing introduced by Wainwright et al. (2005) is a more generalised version that efficiently finds the solution to any type of graph and this approach was further modified by Kolmogorov (2006). A detailed description of the BP algorithm is provided in Szeliski et al. (2008) and Kschischang et al. (2001).

Shekhovtsov et al. (2008) and Lee et al. (2008) used TRW algorithm in 2D and 3D intensity based image registrations, respectively. They decomposed the graph into single dimensional graphs in order to reduce the computations. Kwon et al. (2011) used TRW to optimise an energy that contains a dense local descriptor as data term and higher order smoothness prior.

Linear programming based methods try to improve the efficiency by solving a linear programming relaxation of the original labelling problem. These methods are generally faster than graph-cut based methods and the regularisation term does not have to be metric as in  $\alpha$ -expansion. Glocker et al. (2008) used the FFD model with B-Splines to reduce the dimensionality of the problem and used an energy function that is projected onto the control points (which comprised of an arbitrary similarity measure and a smoothness measure that penalises the difference in adjacent parameter labels). Then the problem is modelled as an equivalent integer

program which was subjected to linear program relaxation. The optimal labelling is then found by using the fast primal-dual algorithm (Komodakis et al. 2007). Additionally they used a hierarchical warping scheme to account for the large deformation and imposed a hard constraint to preserve the diffeomorphism. In Glocker et al. (2011) the authors extended this work to integrate both iconic and geometric registration into a unified formulation and proposed two algorithms. The first one is a computationally efficient implementation while the second uses a tight relaxation.

## 10.5 Summary

Image registration of medical images is a fundamental task in image processing and it has been studied for over three decades. Due to the non-rigid motions associated with soft tissue, different types of modalities involved in capturing those images and large number of parameters needed to model the associated deformation image registration problem, the associated objective function, transformation models and optimisation procedures are highly nonlinear. This chapter outlines the nonlinear methods used in the key segments of solving the medical image registration problem.

### Key Symbols

$T$	Spatial transformation
$x$	Voxel index
$I_F(x)$	Fixed image
$I_M(x)$	Moving image
$\Omega_F$	Fixed image domain
$\Omega_M$	Moving image domain
$\theta$	Transformation parameter vector
$u(x)$	Deformation at voxel $x$
$v(x)$	Velocity at voxel $x$
$\Psi(\theta)$	Objective function
$\mathcal{M}(\cdot)$	Similarity measure
$\mathcal{R}(\cdot)$	Regularisation term
$J(\theta)$	Jacobian of the transformation



## References

- Andronache A, Cattin P, Székely G (2006) Local intensity mapping for hierarchical non-rigid registration of multi-modal images using the cross-correlation coefficient. *Biomed Image Regist* 4057:26–33 [Springer]
- Arsigny V, Pennec X, Ayache N (2005) Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations—application to the registration of histological slices. *Med Image Anal* 9:507–523
- Arzhaeva Y, Prokop M, Murphy K, Van Rikxoort EM, de Jong PA, Gietema HA, Viergever MA, Van Ginneken B (2010) Automated estimation of progression of interstitial lung disease in CT images. *Med Phys* 37:63–73
- Audette MA, Ferrie FP, Peters TM (2000) An algorithmic overview of surface registration techniques for medical imaging. *Med Image Anal* 4:201–217
- Bhagalia R, Fessler JA, Boklye K (2009) Accelerated nonrigid intensity-based image registration using importance sampling. *IEEE Trans Med Imaging* 28:1208–1216
- Boykov Y, Veksler O, Zabih R (2001) Fast approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Mach Intell* 23:1222–1239
- Broit C (1981) Optimal registration of deformed images. University of Pennsylvania
- Bro-Nielsen M, Gramkow C (1996) Fast fluid registration of medical images. In: Höhne K, Kikinis R (eds) *Visualization in biomedical computing*. Springer, Berlin
- Burschka D, Li M, Taylor R, Hager GD (2004) Scale-invariant registration of monocular endoscopic images to CT-scans for sinus surgery. In: *Medical image computing and computer-assisted intervention—MICCAI 2004*. Springer
- Cahill ND, Schnabel JA, Noble JA, Hawkes DJ (2008) Revisiting overlap invariance in medical image alignment. In: *IEEE computer society conference on computer vision and pattern recognition workshops, 2008 (CVPRW'08)*. IEEE, pp 1–8
- Castillo E, Castillo R, Zhang Y, Guerrero T (2009) Compressible image registration for thoracic computed tomography images. *J Med Biol Eng* 29:222–233
- Castillo E, Castillo R, White B, Rojo J, Guerrero T (2012) Least median of squares filtering of locally optimal point matches for compressible flow image registration. *Phys Med Biol* 57:4827–4833
- Christensen GE (1999) Consistent linear-elastic transformations for image matching. In: Kuba A, Šámal M, Todd-Pokropek A (eds) *Information processing in medical imaging*. Springer, Berlin
- Christensen G, He J (2001) Consistent nonlinear elastic image registration. In: *IEEE workshop on mathematical methods in biomedical image analysis, 2001 (MMBIA 2001)*. IEEE, pp 37–43
- Christensen GE, Rabbitt RD, Miller MI (1994) 3D brain mapping using a deformable neuroanatomy. *Phys Med Biol* 39:609
- Christensen GE, Rabbitt RD, Miller MI (1996) Deformable templates using large deformation kinematics. *IEEE Trans Image Process* 5:1435–1447
- Chun SY, Fessler JA (2008) Regularized methods for topology-preserving smooth nonrigid image registration using b-spline basis. In: *5th IEEE international symposium on biomedical imaging: from nano to macro, 2008 (ISBI 2008)*. IEEE, pp 1099–1102
- Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marchal G (1995) Automated multi-modality image registration based on information theory. *Inf Process Med Imaging* 3:263–274
- Corpetti T, Mémin É, Pérez P (2002) Dense estimation of fluid flows. *IEEE Trans Pattern Anal Mach Intell* 24:365–380
- Ford D, Fulkerson DR (2010) *Flows in networks*. Princeton University Press, Princeton
- Gee JC, Haynor DR, Reivich M, Bajcsy RK (1994) Finite element approach to warping of brain images. In: *Medical imaging 1994. International Society for Optics and Photonics*, pp 327–337
- Gerber N, Gavaghan KA, Bell BJ, Williamson TM, Weisstanner C, Caversaccio MD, Weber S (2013) High-accuracy patient-to-image registration for the facilitation of image-guided robotic microsurgery on the head. *IEEE Trans Biomed Eng* 60:960–968

- Giger ML, Chan H-P, Boone J (2008) Anniversary paper: history and status of CAD and quantitative image analysis: the role of medical physics and AAPM. *Med Phys* 35:5799–5820
- Glocker B, Komodakis N, Tziritis G, NAVAB N, Paragios N (2008) Dense image registration through MRFs and efficient linear programming. *Med Image Anal* 12:731–741
- Glocker B, Sotiras A, Komodakis N, Paragios N (2011) Deformable medical image registration: setting the state of the art with discrete methods. *Annu Rev Biomed Eng* 13:219–244
- Gorbunova V, Sporning J, Lo P, Loeve M, Tiddens HA, Nielsen M, Dirksen A, De Bruijne M (2012) Mass preserving image registration for lung CT. *Med Image Anal* 16:786–795
- Guo W, Li Q, Boyce SJ, Mcadams HP, Shiraishi J, Doi K, Samei E (2012) A computerized scheme for lung nodule detection in multiprojection chest radiography. *Med Phys* 39:2001–2012
- Haber E, Modersitzki J (2006) Intensity gradient based registration and fusion of multi-modal images. In: *Medical image computing and computer-assisted intervention—MICCAI 2006*. Springer
- Hero A, Ma B, Michel OJ, Gorman J (2002) Applications of entropic spanning graphs. *IEEE Signal Process Mag* 19:85–95
- Hill DL, Batchelor PG, Holden M, Hawkes DJ (2001) Medical image registration. *Phys Med Biol* 46:R1
- Holden M (2008) A review of geometric transformations for nonrigid body registration. *IEEE Trans Med Imaging* 27:111–128
- Huang CH, Hsieh CH, Lee JD, Huang WC, Lee ST, Wu CT, Sun YN, Wu YT (2012) A CT-ultrasound-coregistered augmented reality image-guided surgery system and its preliminary study on brain-shift estimation. *J Instrum* 7, P08016
- Illán IA, Góriz JM, Ramírez J, Segovia F, Jiménez-Hoyuela JM, Ortega Lozano SJ (2012) Automatic assistance to Parkinsons disease diagnosis in DaTSCAN SPECT imaging. *Med Phys* 39:5971–5980
- Jacobson TJ, Murphy MJ (2011) Optimized knot placement for B-splines in deformable image registration. *Med Phys* 38:4579–4582
- Jaya J, Thanushkodi K (2011) Implementation of computer aided diagnosis system based on parallel approach of ant based medical image segmentation. *J Comput Sci* 7:291–297
- Kim J (2004) Intensity based image registration using robust similarity measure and constrained optimization: applications for radiation therapy. *Citeseer*
- Klein S, Staring M, Pluim JPW (2007) Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. *IEEE Trans Image Process* 16:2879–2890
- Klein S, Van der Heide UA, Lips IM, Van Vulpen M, Staring M, Pluim JP (2008) Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med Phys* 35:1407–1417
- Kleinberg J, Tardos E (2002) Approximation algorithms for classification problems with pairwise relationships: metric labeling and Markov random fields. *J ACM* 49:616–639
- Ko JP, Betke M (2001) Chest CT: automated nodule detection and assessment of change over time—preliminary experience. *Radiology* 218(1):267–273
- Kolmogorov V (2006) Convergent tree-reweighted message passing for energy minimization. *IEEE Trans Pattern Anal Mach Intell* 28:1568–1583
- Kolmogorov V, Zabini R (2004) What energy functions can be minimized via graph cuts? *IEEE Trans Pattern Anal Mach Intell* 26:147–159
- Komodakis N, Tziritis G, Paragios N (2007) Fast, approximately optimal solutions for single and dynamic MRFs. In: *IEEE conference on computer vision and pattern recognition, 2007 (CVPR'07)*, 17–22 June 2007, pp 1–8
- Kschischang FR, Frey BJ, Loeliger H-A (2001) Factor graphs and the sum-product algorithm. *IEEE Trans Inf Theory* 47:498–519
- Kwon D, Lee KJ, Yun ID, Lee SU (2011) Nonrigid image registration using higher-order MRF model with dense local descriptor. In: *Computer vision and pattern recognition workshop, 2011*
- Kybic J, Unser M (2003) Fast parametric elastic image registration. *IEEE Trans Image Process* 12:1427–1442

- Lee KJ, Kwon D, Yun ID, Lee SU (2008) Deformable 3D volume registration using efficient MRFs model with decomposed nodes. In: *BMVC 2008*, pp 1–10
- Lester H, Arridge SR (1999) A survey of hierarchical non-linear medical image registration. *Pattern Recogn* 32:129–149
- Liang Y, Ryan NS, Schott JM, Fox NC (2013) Imaging the onset and progression of Alzheimer's disease: implications for prevention trials. *J Alzheimers Dis* 33:S305–S312
- Likar B, Pernuš F (2001) A hierarchical approach to elastic registration based on mutual information. *Image Vis Comput* 19:33–44
- Linte CA, White J, Eagleson R, Guiraudon GM, Peters TM (2010) Virtual and augmented medical imaging environments: enabling technology for minimally invasive cardiac interventional guidance. *IEEE Rev Biomed Eng* 3:25–47
- Loeckx D, Slagmolen P, Maes F, Vandermeulen D, Suetens P (2010) Nonrigid image registration using conditional mutual information. *IEEE Trans Med Imaging* 29:19–29
- Maintz JBA, Viergever MA (1998) A survey of medical image registration. *Med Image Anal* 2: 1–36
- Mirota DJ, Hanzi W, Taylor RH, Ishii M, Gallia GL, Hager GD (2012) A system for video-based navigation for endoscopic endonasal skull base surgery. *IEEE Trans Med Imaging* 31:963–976
- Modersitzki J (2004) *Numerical methods for image registration (numerical mathematics and scientific computation)*. Oxford University Press, Oxford
- Modersitzki J (2008) FLIRT with rigidity—image registration with a local non-rigidity penalty. *Int J Comput Vis* 76:153–163
- Mostayed A, Garlapati RR, Joldes GR, Wittek A, Roy A, Kikinis R, Warfield SK, Miller K (2013) Biomechanical model as a registration tool for image-guided neurosurgery: evaluation against bspline registration. *Ann Biomed Eng* 41:2409–2425
- Musse O, Heitz F, Armpach J-P (2001) Topology preserving deformable image matching using constrained hierarchical parametric models. *IEEE Trans Image Process* 10:1081–1093
- Myronenko A, Xubo S (2010) Intensity-based image registration by minimizing residual complexity. *IEEE Trans Med Imaging* 29:1882–1891
- Noblet V, Heinrich C, Heitz F, Armpach J-P (2005) 3-D deformable image registration: a topology preservation scheme based on hierarchical deformation models and interval analysis optimization. *IEEE Trans Image Process* 14:553–566
- Oliveira FP, Tavares JM (2014) Medical image registration: a review. *Comput Methods Biomech Biomed Eng* 17:73–93
- Pearl J (1988) *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, San Francisco
- Pennec X, Stefanescu R, Arsigny V, Fillard P, Ayache N (2005) Riemannian elasticity: a statistical regularization framework for non-linear registration. In: *Medical image computing and computer-assisted intervention—MICCAI 2005*. Springer
- Pitiot A, Malandain G, Bardinet E, Thompson PM (2003) Piecewise affine registration of biological images. In: *Biomedical image registration*. Springer
- Pluim JP, Maintz JA, Viergever MA (2000) Image registration by maximization of combined mutual information and gradient information. In: *Medical image computing and computer-assisted intervention—MICCAI 2000*. Springer, pp 452–461
- Pluim JP, Maintz JA, Viergever MA (2003) Mutual-information-based registration of medical images: a survey. *IEEE Trans Med Imaging* 22:986–1004
- Rabbitt RD, Weiss JA, Christensen GE, Miller MI (1995) Mapping of hyperelastic deformable templates using the finite element method. In: *SPIE's 1995 international symposium on optical science, engineering, and instrumentation*. International Society for Optics and Photonics, pp 252–265
- Rivaz H, Collins DL (2012) Self-similarity weighted mutual information: a new nonrigid image registration metric. *Med Image Comput Comput Assist Interv* 15:91–98
- Robertson S, Weiss E, Hugo GD (2013) Deformable mesh registration for the validation of automatic target localization algorithms. *Med Phys* 40(7):071721

- Rohde GK, Aldroubi A, Dawant BM (2003) The adaptive bases algorithm for intensity-based nonrigid image registration. *IEEE Trans Med Imaging* 22:1470–1479
- Rueckert D, Aljabar P (2010) Nonrigid registration of medical images: theory, methods, and applications [applications corner]. *IEEE Signal Process Mag* 27:113–119
- Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ (1999) Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging* 18:712–721
- Rueckert D, Clarkson M, Hill D, Hawkes DJ (2000) Non-rigid registration using higher-order mutual information. In: *Medical imaging 2000*. International Society for Optics and Photonics, pp 438–447
- Ruhaak J, Konig L, Hallmann M, Papenberg N, Heldmann S, Schumacher H, Fischer B (2013) A fully parallel algorithm for multimodal image registration using normalized gradient fields. In: *2013 IEEE 10th international symposium on biomedical imaging (ISBI)*, 7–11 April 2013, pp 572–575
- Rui L, Li Z, Ying S, Shun M, Chefid'hotel C (2013) A review of recent advances in registration techniques applied to minimally invasive therapy. *IEEE Trans Multimed* 15:983–1000
- Rynne BP, Youngson MA (2008) *Linear functional analysis*. Springer, London
- Sadeghi Naini A, Patel RV, Samani A (2010) CT-enhanced ultrasound image of a totally deflated lung for image-guided minimally invasive tumor ablative procedures. *IEEE Trans Biomed Eng* 57:2627–2630
- Shekhovtsov A, Kovtun I, Hlaváč V (2008) Efficient MRF deformation model for non-rigid image matching. *Comput Vis Image Underst* 112:91–99
- Shiraishi J, Abe H, Li F, Engelmann R, Macmahon H, Doi K (2006) Computer-aided diagnosis for the detection and classification of lung cancers on chest radiographs. ROC analysis of radiologists' performance. *Acad Radiol* 13:995–1003
- Sluimer I, Schilham A, Prokop M, Van Ginneken B (2006) Computer analysis of computed tomography scans of the lung: a survey. *IEEE Trans Med Imaging* 25:385–405
- So R, Chung A (2009) Multi-level non-rigid image registration using graph-cuts. In: *IEEE international conference on acoustics, speech and signal processing, 2009 (ICASSP 2009)*, 19–24 April 2009, pp 397–400
- So RWK, Tang TWH, Chung ACS (2011) Non-rigid image registration of brain magnetic resonance images using graph-cuts. *Pattern Recogn* 44:2450–2467
- Sotiras A, Davatzikos C, Paragios N (2013) Deformable medical image registration: a survey. *IEEE Trans Med Imaging* 32:1153–1190
- Staring M, van der Heide UA, Klein S, Viergever MA, Pluim JP (2009) Registration of cervical MRI using multifeature mutual information. *IEEE Trans Med Imaging* 28:1412–1421
- Studholme C, Hill DL, Hawkes DJ (1999) An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn* 32:71–86
- Studholme C, Drapaca C, Iordanova B, Cardenas V (2006) Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. *IEEE Trans Med Imaging* 25:626–639
- Szeliski R, Zabih R, Scharstein D, Veksler O, Kolmogorov V, Agarwala A, Tappen M, Rother C (2008) A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. *IEEE Trans Pattern Anal Mach Intell* 30:1068–1080
- Tennakoon RB, Bab-Hadiashar A, Cao Z, De Bruijne M (2014) Nonrigid registration of volumetric images using ranked order statistics. *IEEE Trans Med Imaging* 33:422–432
- Timp S, Varela C, Karssemeijer N (2007) Temporal change analysis for characterization of mass lesions in mammography. *IEEE Trans Med Imaging* 26:945–953
- Uneri A, Nithiananthan S, Schafer S, Otake Y, Stayman JW, Kleinszig G, Sussman MS, Prince JL, Siewerdsen JH (2013) Deformable registration of the inflated and deflated lung in cone-beam CT-guided thoracic surgery: initial investigation of a combined model- and image-driven approach. *Med Phys* 40
- Van Ginneken B, Ter Haar Romeny BM, Viergever MA (2001) Computer-aided diagnosis in chest radiography: a survey. *IEEE Trans Med Imaging* 20:1228–1241

- van Herk M (2007) Different styles of image-guided radiotherapy. *Semin Radiat Oncol* 17: 258–267
- Vercouteren T, Pennec X, Perchant A, Ayache N (2007) Non-parametric diffeomorphic image registration with the demons algorithm. In: *Medical image computing and computer-assisted intervention—MICCAI 2007*. Springer
- Viola P, Wells WM III (1997) Alignment by maximization of mutual information. *Int J Comput Vis* 24:137–154
- Wainwright MJ, Jaakkola TS, Willsky AS (2005) MAP estimation via agreement on trees: message-passing and linear programming. *IEEE Trans Inf Theory* 51:3697–3717
- Wang J, Jiang T (2007) Nonrigid registration of brain MRI using NURBS. *Pattern Recogn Lett* 28:214–223
- Wei J, Chan HP, Sahiner B, Zhou C, Hadjiiski LM, Roubidoux MA, Helvie MA (2009) Computer-aided detection of breast masses on mammograms: dual system approach with two-view analysis. *Med Phys* 36:4451–4460
- Worm ES, Hansen AT, Petersen JB, Muren LP, Præstegaard LH, Høyer M (2010) Inter- and intrafractional localisation errors in cone-beam CT guided stereotactic radiation therapy of tumours in the liver and lung. *Acta Oncol* 49:1177–1183
- Wu Y-T, Kanade T, Li C-C, Cohn J (2000) Image registration using wavelet-based motion model. *Int J Comput Vis* 38:129–152
- Xishi H, Jing R, Guiraudon G, Boughner D, Peters TM (2009) Rapid dynamic image registration of the beating heart for diagnosis and surgical navigation. *IEEE Trans Med Imaging* 28: 1802–1814
- Yedidia JS, Freeman WT, Weiss Y (2000) Generalized belief propagation. In: *NIPS 2000*, pp 689–695
- Yin Y, Hoffman EA, Lin C-L (2009) Mass preserving nonrigid registration of CT lung images using cubic B-spline. *Med Phys* 36:4213
- Yuanjie Z, Kambhamettu C, Bauer T, Steiner K (2009) Accurate estimation of pulmonary nodule's growth rate in CT images with nonrigid registration and precise nodule detection and segmentation. In: *IEEE computer society conference on computer vision and pattern recognition workshops 2009 (CVPR workshops 2009)*, 20–25 June 2009, pp 101–108
- Zitova B, Flusser J (2003) Image registration methods: a survey. *Image Vis Comput* 21:977–1000

# Chapter 11

## Pose Estimation of a Dioptric Imaging Sensor with a Circle-Projecting Collimated Laser Moving Inside a Pipeline

Abbasali Dehghan Tezerjani, Mehran Mehrandezh, and Raman Paranjape

### 11.1 Introduction

Pipelines are the primary source for the transport of material such as water, oil, and gas in the world. Therefore, quality control at the manufacturing stage and/or pro-active timely inspection of them, when in service, are critical. Failure in pipes can have severe adversarial environmental and economic shocks. It is estimated that over 50 % of the pipeline infrastructure in North America are older than 60 years, Cohen (2012). Development of non-destructive testing methods for quality control of pipes, at the manufacturing stage, and inspection, when pipes in use, is paramount to the industry. This research is a timely response to a worldwide need for tools that can do consistent reporting on pipe's condition with minimum man-hour required. Robotic inspection systems would be a perfect fit to this task.

An image-based optics laser sensor, targeted towards the inspection of pipes, was developed in our group, Tezerjani et al. (2013). The information obtained via images can be used to (1) detect defects, (2) classify them, (3) size them (i.e., making metric measurements on defects via images), and (4) map them (i.e., localize them inside the pipe with respect to a reference point). The focus of this research paper is to provide a guideline for precise pose estimation of an optics laser imaging sensor required for sizing the defects and providing metric measurements from images. In this context, precise image-based metric measurements would require: (1) an accurate estimation of the pose of the imaging sensor with respect to the detected defects, and (2) an accurate projection model to convert highly distorted 2D images to a scaled Axonometric projection space.

---

A.D. Tezerjani (✉) • M. Mehrandezh • R. Paranjape  
Faculty of Engineering and Applied Science, University of Regina,  
3737 Wascana Pky, Regina, SK, Canada S4S 0A2  
e-mail: [dehghana@uregina.ca](mailto:dehghana@uregina.ca); [mehran.mehrandezh@uregina.ca](mailto:mehran.mehrandezh@uregina.ca); [raman.paranjape@uregina.ca](mailto:raman.paranjape@uregina.ca)

Vision-based pose estimation of imaging sensors has been studied for the past three decades. The seminal work of Dementhon and Davis (1995) on real-time pose estimation of a perspective camera pioneers the research. However, there is a huge body of literature on improving the image-based pose-estimation algorithms with the objective to make them less sensitive to the image noise, and lens imperfection/distortion, e.g. Simo-Serra et al. (2012) and Josephson and Byrod (2009).

In general, the pose estimation of imaging sensors was applied primarily in three areas: (1) visual servoing of unmanned systems, e.g. Tahri et al. (2010), (2) 3D digital reconstruction of objects from sequential images and/or video streams, e.g. Spica et al. (2014) and Chang (2002) and (3) Augmented reality, e.g. Hagbi et al. (2011). In this research, the pose estimation is targeted towards the 3D reconstruction of an object (i.e., pipe's surface) from sequential images obtained via an optics laser (i.e., a laser in conjunction with a camera). The technical term widely used in the literature is: structure from motion (SFM), Spica et al. (2014) and Chang (2002). A super-resolution 3D reconstruction of an object can be achieved using the optics laser. The accuracy can be comparable to that in military-grade inertial measurement units (IMUs).

Image-based pose estimation in unstructured environments through an SFM paradigm remains a challenge Daniilidis and Spetsakis (1997). Feature detection, and feature matching/adjustment are paramount in the accuracy of the pose estimations. Usually, one would require significant image overlaps between sequential images to identify false detection or outliers Iketani et al. (2007). Dynamic features can also affect the results, therefore, intensive statistical analysis would be needed to identify and bundle static features and to separate them from dynamic features, Alvarez et al. (2007). An alternative to the SFM in unstructured environment is to use structured light patterns, such as a grid-dot or circular laser patterns projected on objects Andrews et al. (2013). M. Vincez et al. use a laser projector with a grid-dot pattern to estimate the relative pose between a spray-painting nozzle attached to a robot's end-effector and the surface to be painted with applications in automotive industry Vincze et al. (2002). Similar work has been done for robotic welding Pipeline and Gas Journal (1999).

S. Winkelbach et al. use optics laser for 3D reconstruction of objects Winkelbach et al. (2006). However, they assume that the imaging tool is calibrated offline and that it remains stationary. Also, these imaging sensors can be used to scan small objects only. For scanning large objects, such as that in a pipeline, the imaging tool needs to move long distances. Any disturbance in the imaging sensor's motion would affect the quality of the digital reconstruction of the objects. Further continuous pose estimation techniques would be needed for better accuracy.

More specifically, image-based inspection of pipelines is gaining significant attention by the research institutes and industry, e.g. Duran et al. (2007, 2003) and Matsui et al. (2010). Duran et al. used a laser projecting device with a circular pattern and a perspective camera to detect defects Duran et al. (2007). Their problem

formulation is based on the assumption that the laser/camera would remain at the center point of the pipe all the time. Besides, their work revolved around defect detection and classification only.

A large number of companies making pipe inspection tools use optics laser to detect and classify defects in pipes as well, e.g. Maverick Inspection Ltd. (2009) and Raush Electronics USA (2014). However, the precise sizing of the detected defects would require additional time-consuming a priori system setup, such as alignment of the scanning device in the pipe. A similar concept was used by Unnikrishnan et al. (2009) to estimate the pose of an optics laser scanning device for the inspection of underwater infrastructure. The pose estimations were used only to navigate their underwater vehicle system. Therefore, a low-resolution image would suffice. In their work, a look-up table, containing the mapping between the laser ring image and the pose of the robot, was generated through a training step. Usually, the training is carried out on a set of representative tests and then the results are extrapolated to un-rehearsed scenarios. This can be highly susceptible to the image noise, and the quality and diversity of the data used for training. A much higher accuracy would be needed in estimating the pose of the imaging sensor for 3D reconstruction of a surface in high resolution.

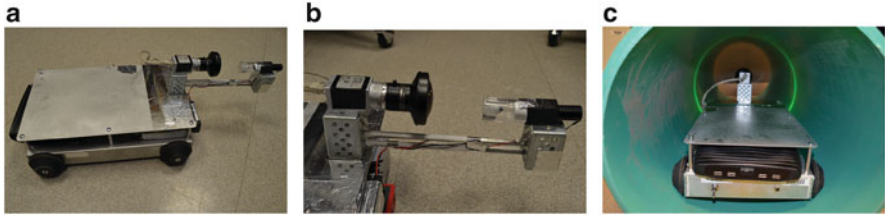
To the best of our knowledge, this is the first time a systematic algorithm is developed for pose estimation of a panoramic optics laser imaging sensor inside a pipe. This would minimize the man-hour needed to set up the device, simply because no alignment would be required. Also, the algorithm accounts for unwanted disturbances in the sensor's motion. Therefore, the pose estimation can be done more accurately and, consequently, this will improve the accuracy of the image-based metric measurements of the defects.

The layout of the paper is as follows: Sect. 11.2 formulates the projection geometry of the proposed imaging platform, and introduces the required components for the pose estimation. Section 11.3 presents a novel approach for mapping the pose estimation problem into the numerical optimization problem. Performance of the proposed pose estimation technique is tested using the synthesized images and the experimental results are presented in Sect. 11.4. In this section, the performance of the proposed algorithm is discussed, and effects of different parameters on the performance of the pose estimation algorithm are studied.

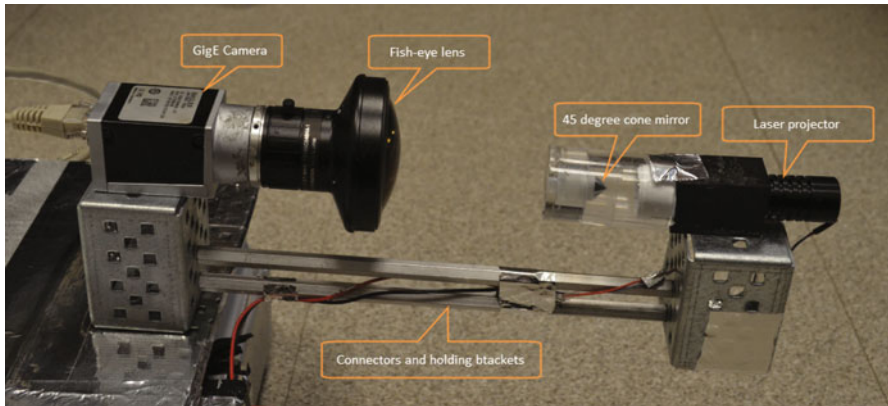
## 11.2 Problem Formulation

Figure 11.1a–c show the robotic inspection system, the imaging sensor mounted on the robot, and a snapshot of the robot inside a PVC pipe, respectively. The main components of the imaging unit, namely (1) the camera with a fish-eye lens, (2) a laser projector, (3) a 45° conic mirror, and (4) connectors and holding brackets are shown in Fig. 11.2. The laser dot reflected on the mirror will be projected as a perfect circle on the pipe's surface as long as the laser/camera is centered inside the pipe. The image of the circle projected on the pipe's wall not be necessarily





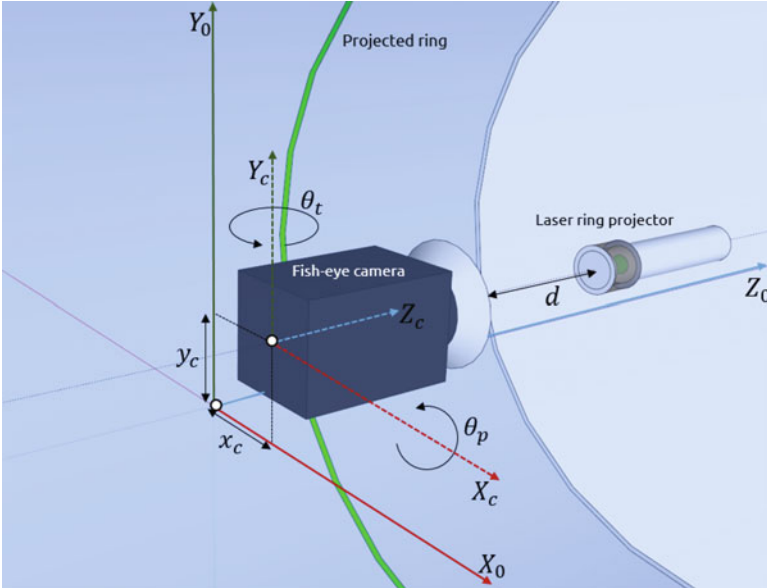
**Fig. 11.1** Robotic pipe inspection system. (a) Robotic inspection system. (b) Imaging sensor mounted on the robot. (c) Snapshot of the robot inside a PVC pipe



**Fig. 11.2** Main components of the imaging unit

a conic because of the lens distortion. Its overall shape would depend on the radial/tangential distortions in the fish-eye lens and also the aspect ratio of the camera's CCD array and/or the aspect ratio of the display unit (i.e., the computer screen). The image of the projected laser would be also skewed depending on the lateral and vertical sway motions in the laser/camera and also its pan and tilt angles. However, it is noteworthy that the axial movement of the imaging sensor inside the pipe and its rolling motion will not affect the image. The objective is to estimate four degrees of motion, namely two translational sway motions along horizontal and vertical axes and two pan and tilt angles from a set of image moments of the laser image seen by the camera. The selection of the image moments will be discussed below.

First, let us define the notations used for representing the pose of the sensor and also the image moments. For this purpose, two rectangular coordinate systems are introduced. The "0" frame denotes a fixed frame whose z-axis is along the longitudinal axis of the pipe (see Fig. 11.3). The "C" frame, also seen in Fig. 11.3, denotes a frame whose origin is attached to the center of projection of the camera. The pose of the imaging sensor is represented by a quadruplet;  $\langle x_c, y_c, \theta_p, \theta_t \rangle$ , where  $x_c$  and  $y_c$  denote the lateral and vertical sway in camera's motion with respect



**Fig. 11.3** Proposed imaging platform and its configuration inside the pipe

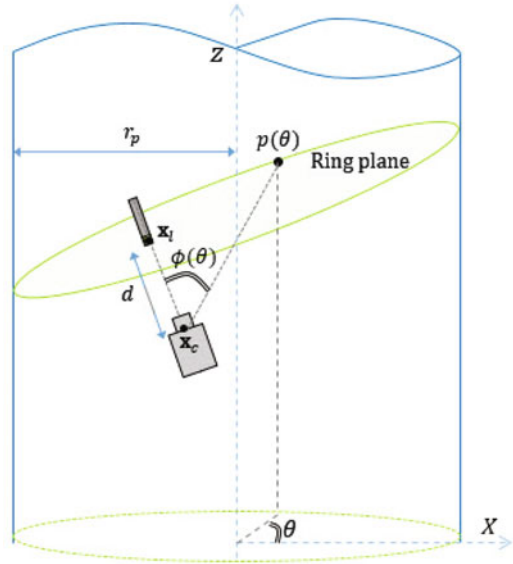
to the fixed frame, respectively. Also,  $\theta_p$  and  $\theta_t$  denote the pan and tilt rotational angles around the  $x$  and  $y$  axes of the fixed frame, respectively (see Fig. 11.3).

The image moments are defined as the coefficients of a high-order polynomial optimally fit to the laser image. It is noteworthy that the laser image seen by the camera would not be a conic in general, therefore, it cannot be represented by a quadratic polynomial. Extra care must be taken to make sure that the functional relationship between the pose quadruplet and the image moments is of full rank. Furthermore, due to the radially symmetric shape of the pipe, finite number of disperse solutions to the pose estimation problem may exist. A binary filter would be needed to select the right solution.

*Model reduction by the separation of variables:* The relationship between the pose quadruplet and the image moments will be highly nonlinear and coupled, therefore, nonlinear optimization techniques must be adopted. However, some variables can be decoupled through a separation of variable technique. This model reduction can speed up the convergence of the optimization process while assuring that a local optimum solution will be reached within the initial simplex defined by the user. This is in particular important to avoid jumping between finite possible solutions that may exist.

In the proposed model reduction process, first the relationship between the parametric equation of the 3D curve, which is the result of the intersection of the laser cone with the pipe's wall and the pose quadruplet, is derived. In the second step, the relationship between the parametric equations of the same 3D curve with the image moments is derived. Solutions obtained via optimization techniques from

**Fig. 11.4** Projection geometry of the imaging platform



one stage are traversed to the second and a back-propagation correcting mechanism is used, based on a least square error of a fitting function, iteratively till a solution is converged.

Stage I—Calculating the intersection curve between the laser cone and the pipe’s wall using the pose quadruplet:

As seen in Fig. 11.4, the omnidirectional laser projects a ring perpendicular to its central axis on the pipe’s wall. The image of the ring collinear with the laser projector at the distance  $d$  is captured by the fish-eye camera. This image is used to estimate the relative pose of the camera,  $\langle x_c, y_c, \theta_p, \theta_t \rangle$  inside the pipe. Figure 11.4 demonstrates the projection geometry of the imaging platform. Here,  $\mathbf{x}_c$ , and  $\mathbf{x}_l$  represent the coordinates of the camera pinhole and the laser projector head, respectively. Given the pose quadruplet  $\langle x_c, y_c, \theta_p, \theta_t \rangle$ , the coordinate of the laser projector head is calculated as follows:

$$\mathbf{x}_l = \mathbf{x}_c + R_x(\theta_p) * R_y(\theta_t) * \begin{pmatrix} 0 \\ 0 \\ d \end{pmatrix} = \begin{pmatrix} x_c - d \sin \theta_t \\ y_c + d \cos \theta_t \sin \theta_p \\ d \cos \theta_p \cos \theta_t \end{pmatrix} \quad (11.1)$$

where  $R_x$  and  $R_y$  denote the rotation matrices about the X and Y axes. The equation of the ring plane is calculated as following:

$$\begin{aligned} \overrightarrow{\mathbf{x}_l - \mathbf{x}_c} \cdot [x, y, z]^T - \mathbf{x}_l &= 0 \Rightarrow \\ \mathbf{x} \sin \theta_t - y \cos \theta_t \sin \theta_p - z \cos \theta_p \cos \theta_t + \xi &= 0, \end{aligned} \quad (11.2)$$

where  $\xi = d - x_c \sin \theta_t + y_c \cos \theta_t \sin \theta_p$ . We define the 3D curve  $\mathbf{p}(\theta)$ , as the intersection of the ring plane and the pipe, where  $\theta \in [0, 2\pi]$  refers to the point angle with respect to the  $X$  axis in the  $XY$  plane (see Fig. 11.4). Given the equation of the ring plane, parametric 3D coordinates of the intersection curve  $\mathbf{p}(\theta)$  are calculated as:

$$\mathbf{p}(\theta) = \begin{pmatrix} r_p \cos \theta \\ r_p \sin \theta \\ \frac{\xi + r_p \cos \theta \sin \theta_t - r_p \cos \theta_t \sin \theta_p \sin \theta}{\cos \theta_p \cos \theta_t} \end{pmatrix}, \theta \in [0, 2\pi] \quad (11.3)$$

Given the 3D coordinates of the intersection curve,  $\mathbf{p}(\theta)$ , the incident ray angle of the fish-eye lens corresponding to each point on this curve, is calculated as follows:

$$\phi(\theta) = \angle(\overrightarrow{p(\theta) - \mathbf{x}_c}, \overrightarrow{\mathbf{x}_l - \mathbf{x}_c}) \quad (11.4)$$

It should be added that the angle  $\theta$  on the fixed frame maps to the angle  $\psi$  on the camera's frame by an offset angle  $\delta$ . Equation (11.5) shows the formulation of the offset angle based on the pose quadruplet. Further details about the offset angle and the way it is calculated can be found in Appendix 1.

$$\psi = \theta + \delta,$$

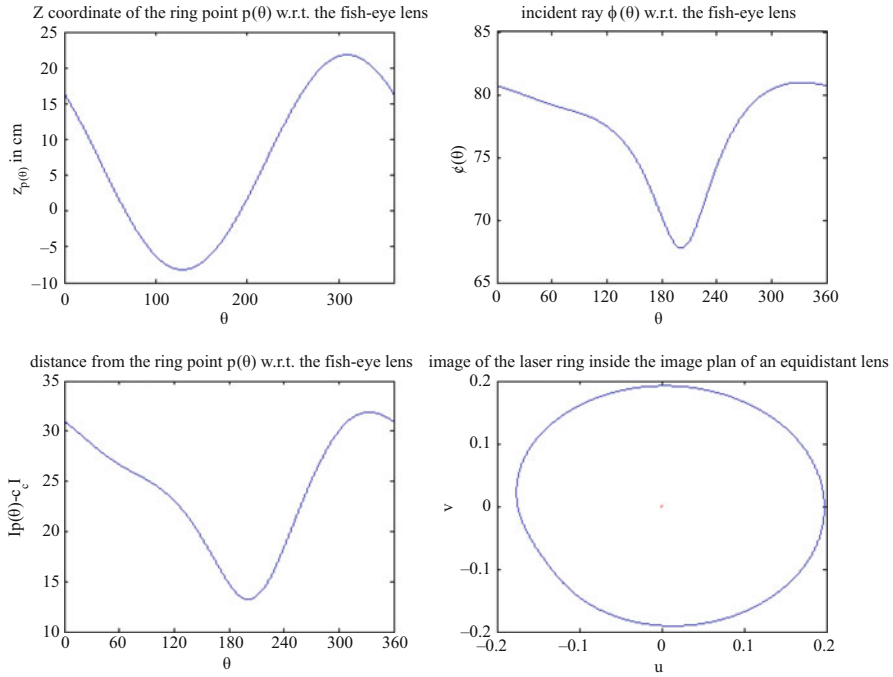
$$\delta = \tan^{-1} \left( -\frac{d \sin \theta_p + y_c + \cos \theta_t + r_p \sin \theta_p \sin \theta_t - x_c \sin \theta_p \sin \theta_t}{\cos \theta_p (r_p - x_c + d \sin \theta_t)} \right) \quad (11.5)$$

Stage II: Calculating the incident ray angle of the intersection points with respect to the fish-eye lens, using the image moments:

Employing the widely used equi-distant projection model for the fish-eye lens, one can relate the pixel coordinates of the image of the intersection curve,  $\mathbf{p}(\theta)$  to the lens focal length,  $f$ , and angles  $\phi$  and  $\psi$  as in Eq. (11.6). Further details on this can be found in Fleck (1995).

$$\begin{bmatrix} u(\theta) \\ v(\theta) \end{bmatrix} = f \cdot \phi(\theta) \cdot \begin{bmatrix} \cos \psi \\ \sin \psi \end{bmatrix} \quad (11.6)$$

This projection model was numerically simulated for a particular case, using a real camera with the equidistant fish-eye lens with a focal length of  $f = 0.14$  cm, and the following configurations:  $x_c = -6$ ,  $y_c = -4$ ,  $\theta_p = \pi/6$ ,  $\theta_t = \pi/8$ ,  $d = 5$  cm, and  $r_p = 20$  cm. Figure 11.5 shows the  $z$ -coordinates of  $\mathbf{p}(\theta)$ , incident ray angle,  $\phi(\theta)$ , the distance from  $\mathbf{p}(\theta)$  to the fish-eye lens, and the image of the laser ring in the image plane. The negative values of the  $z$ -coordinates of the



**Fig. 11.5** Projection of the laser ring on the image plane of an equidistant fish-eye lens for a sample camera pose and configuration  $f = 0.14$  cm,  $t_x = -6$ ,  $t_y = -4$ ,  $\alpha_x = \pi/6$ ,  $\alpha_y = \pi/8$ ,  $d = 5$  cm, and  $r_p = 20$  cm

intersection curve,  $\mathbf{p}(\theta)$ , correspond to the points behind the camera pinhole. As this figure shows, the laser’s image is not always conic or even congruent.

In the next section, we formulate a nonlinear optimization algorithm for estimating the pose quadruplet using the ring’s image.

### 11.3 Solution Methodology

To estimate the relative pose of the camera inside the pipe, we employ a curve fitting technique, and convert the pose estimation problem to an optimization problem using the following steps.

Step 1—Defining the reference signal:

Based on the formulated projection geometry of the proposed imaging platform, we define the reference signal  $ref(\theta)$  as follows:

$$ref(\theta) = \| p(\theta) - \mathbf{x}_c \|^2 = d^2 / \cos^2 \phi(\theta), \quad \theta \in [0 \ 2\pi] \quad (11.7)$$

### Step 2—Deriving the analytic form of the reference signal

The expanded analytical form of the reference signal,  $ref_a(\theta)$ , using Eq. (11.3) is:

$$ref_a(\theta) = (r_p \cos \theta - x_c)^2 + (r_p \sin \theta - y_c)^2 + \left( \frac{\xi + r_p \cos \theta \sin \theta_t - r_p \cos \theta_t \sin \theta_p \sin \theta}{\cos \theta_p \cos \theta_t} \right)^2 \quad (11.8)$$

With some mathematical manipulation it can be shown that the analytical reference signal can be also written as:

$$ref_a(\theta) = a_1 + a_2 \sin \theta + a_3 \cos \theta + a_4 \sin \theta \cos \theta + a_5 \sin^2 \theta \quad (11.9)$$

One may refer to Appendix 2 for the details of the derivation of the coefficients  $a_1, \dots, a_5$ .

### Step 3—Calculating the reference observation signal from the image moments

Having the calibrated fish-eye camera, the direction of incident ray  $\phi(\theta)$  corresponding to the image coordinates  $(u, v)^T$  is calculated using the intrinsic parameters of the camera, regardless of the calibration technique used (Gennery 2006; Kannala and Brandt 2004; Scaramuzza et al. 2006). Here,  $\theta$  denotes the orientation of the incident ray in the image plane. Kannala and Brandt (2004) formulated the mapping from the image coordinate to the direction of incident ray angle for the calibrated fish-eye lenses. Assuming the fish-eye camera calibrated, this mapping will be as follows:

$$\phi(\theta) = \begin{bmatrix} \cos \psi \\ \sin \psi \end{bmatrix}^{-1} f^{-1} \begin{bmatrix} u(\theta) \\ v(\theta) \end{bmatrix} \quad (11.10)$$

From Eq. (11.10), another reference signal, this time obtained from the image, denoted by  $ref_o(\theta)$  is calculated as follows:

$$ref_o(\theta) = d^2 / \cos^2 \phi(\theta - \delta) \quad (11.11)$$

where the offset angle  $\delta$  is calculated using Eq. (11.5). The signal  $ref_o$  is called the reference observation signal.

### Step 4—Fitting a trigonometric function to the observation reference signal

As Eq. (11.9) shows, the reference signal is a trigonometric function of  $\theta$ . It means that the reference observation signal can also be represented in this form. In order to compare the analytical and observation forms of the reference signal, we fit the

reference observation signal to the function given in Eq. (11.9) using the least square fitting algorithm. The fitted signal can be then represented as:

$$ref_o(\theta) = o_1 + o_2 \sin \theta + o_3 \cos \theta + o_4 \sin \theta \cos \theta + o_5 \sin^2 \theta \quad (11.12)$$

Step 5—Calculating the objective function

Given Eqs. (11.9) and (11.12), an objective function is defined as follows:

$$J(x_c, y_c, \theta_p, \theta_t, \zeta) = \int_0^{2\pi} [ref_o(\theta) - ref_a(\theta)]^2 d\theta, \quad (11.13)$$

where  $\zeta$  represents all constant parameters in the equations (i.e.,  $d$  and  $r_p$ ). By substituting the analytical and observation reference signals from Eqs. (11.9) and (11.12) into Eq. (11.13) and doing some mathematical manipulation, the objective function can be simplified to:

$$J(x_c, y_c, \theta_p, \theta_t, \zeta) = \frac{\pi * (8 e_1^2 + 8 e_1 e_5 + 4 e_2^2 + 4 e_3^2 + e_4^2 + 3 e_5^2)}{4}, \quad (11.14)$$

where  $e_i = a_i - o_i$ . This objective function is used in the optimization process to estimate the optimal value of the pose quadruplet. In a short form, the optimization problem can be mathematically written as:  $\min_{Q = \langle x_c, y_c, \theta_p, \theta_t \rangle} \{J(x_c, y_c, \theta_p, \theta_t, \zeta)\}$ .

Step 6—Minimizing the performance index,  $J$  using Nelder–Mead method

The optimization problem is defined as finding the optimal values of the pose quadruplet,  $Q = \langle x_c, y_c, \theta_p, \theta_t \rangle$ , which minimizes the value of the objective function  $J(x_c, y_c, \theta_p, \theta_t, \zeta)$ . The optimization algorithm starts with an initial seed for the pose quadruplet. The optimal value of the pose is calculated by minimizing the value of the objective function through iterations using the *Nelder–Mead* simplex method (Lagarias et al. 1998).

Algorithm 1 represents the optimization process, in a nutshell:

---

**Algorithm 1** Estimating the camera pose quadruplet

---

- 1: Select an initial seed for the current camera pose  $Q = \langle x_c, y_c, \theta_p, \theta_t \rangle$ .
  - 2: Calculate the offset angle  $\delta$  using equation (11.5).
  - 3: Calculate the reference observation signal  $ref_o(\theta)$  using the equations (11.10) and (11.12) and the offset angle  $\delta$  calculated in step 2.
  - 4: Fit the trigonometric function to the reference observation signal  $ref_o(\theta)$ .
  - 5: Calculate the value of the objective function  $J$  using the equation (11.14).
  - 6: Find the optimal pose quadruplet  $\min_{Q = \langle x_c, y_c, \theta_p, \theta_t \rangle} \{J(x_c, y_c, \theta_p, \theta_t, \zeta)\}$  using the NM’s simplex method.
-

Analysis of the proposed method shows that the backward mapping from the laser ring image to the camera pose is not a one-to-one mapping. In fact, each specific image of the laser ring is obtainable through two different camera poses.

For the numerical optimization technique used in this work, the initial seed value plays a major role in convergence to the right solution. Selecting an appropriate value for the initial seed not only improves the success rate of the optimization algorithm but also affects the computation time of the algorithm. In the next section, we study the effect of the initial seed through some experiments and present an initial seed selection method to improve the success rate of the algorithm.

## 11.4 Experiment Results and Discussion

In order to evaluate the performance of the proposed algorithm for estimating the camera pose inside the pipe, a simple GUI which generates the synthetic images based on the camera pose, the pipe radius, the camera–laser distance, and the projection model of the fish-eye lens was developed. Figure 11.6 shows a snapshot of the developed GUI. Each image generated by this tool is passed to the pose estimation algorithm, and the relative pose of the camera inside the pipe is estimated. The numerically estimated camera pose is compared to its real value and the difference is reported.

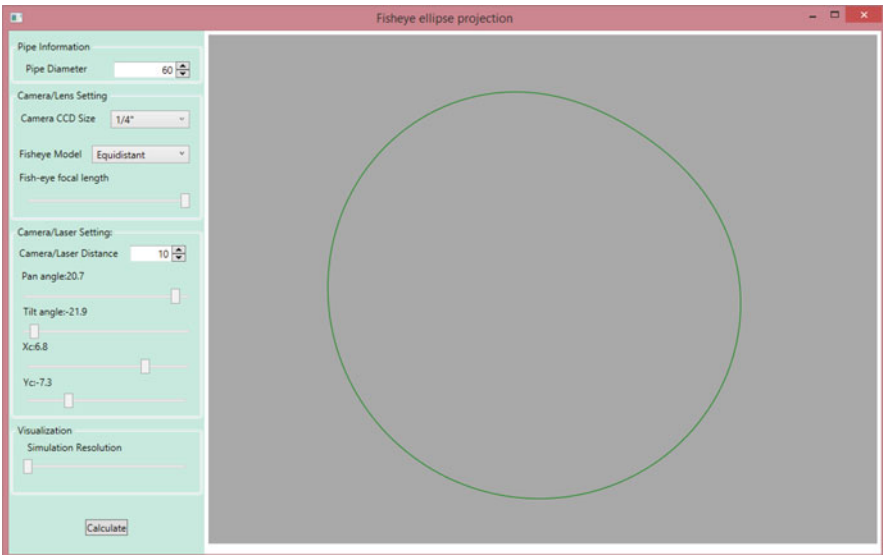
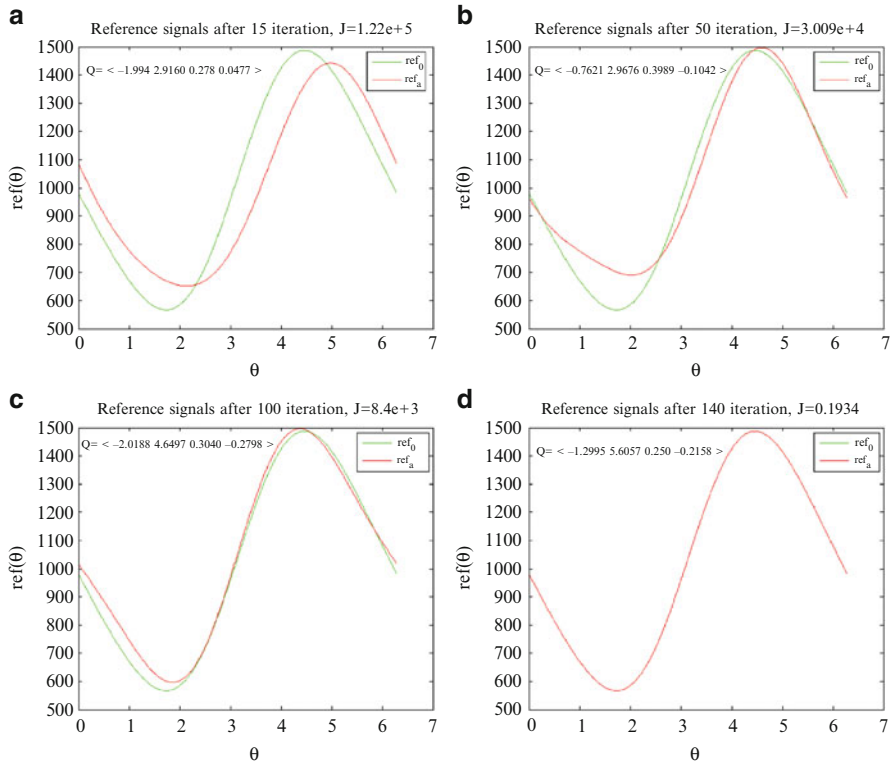


Fig. 11.6 Snapshot of developed GUI for generating sample data



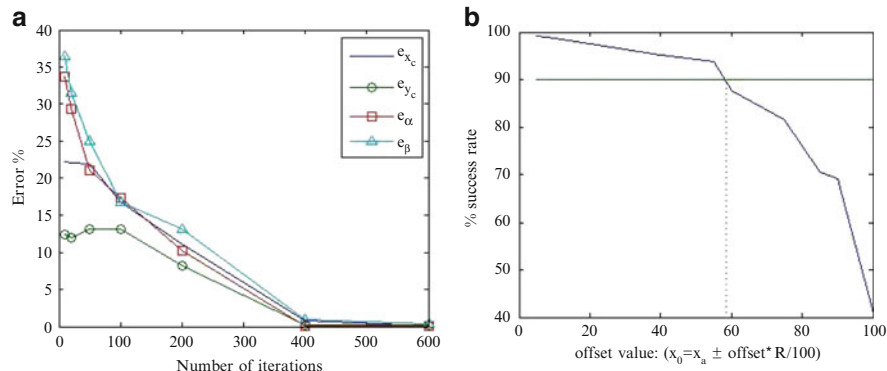


**Fig. 11.7** Observation and analytical reference signals during the optimization process, where the actual pose is  $Q_{actual} = \langle x_c = -1.2988 \ y_c = 5.6033 \ \theta_p = 0.2496 \ \theta_t = -0.2153 \rangle$ : (a) after 15 iterations, (b) after 50 iterations, (c) after 100 iterations, and (d) after 140 iterations

The nonlinear functions used in the proposed pose estimation technique were cyclic with respect to the pan/tilt angles with a period of  $2\pi$  radians. However, they are not cyclic in terms of the lateral/vertical sway motion in the camera. In fact, there are only two possible identical solutions to the sway motions that would yield exactly the same image. In other words, the solutions to the sway motion estimation can have dual jumps. These jumps, and also the cyclic solutions of the pan/tilt estimates, were detected numerically using a binary filter. No false return was ever perceived in the large number of numerical simulations carried out in this research.

Figure 11.7 shows the time-elapsed results of the optimization process. Insets show the variation of the reference signal with  $\theta$  at different iterations. After about 140 iterations the optimization process was terminated meeting the stop criterion.

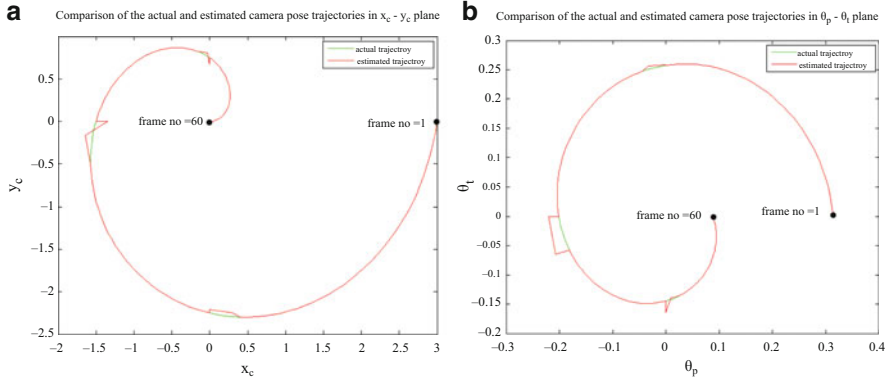
Figure 11.8a shows the error of each component of the estimated camera pose algorithm with respect to the number of iterations. As this figure shows, the rate of change of error in all components of the estimated camera pose declines below 1% after 500 iterations.



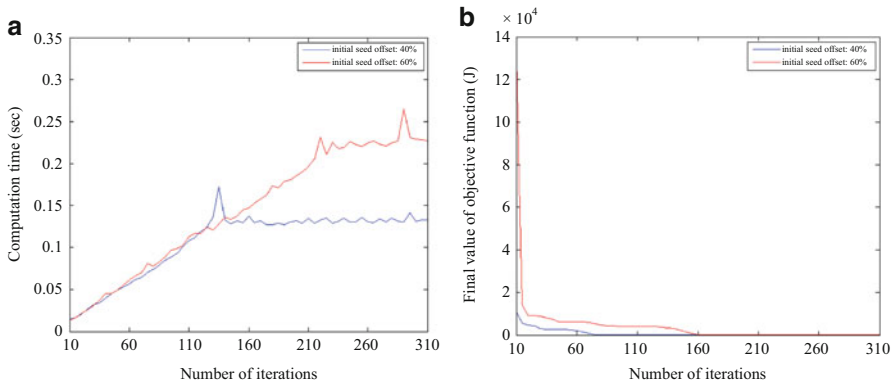
**Fig. 11.8** Performance of the pose estimation algorithm vs. number of iterations and initial seed value. (a) Pose estimation error vs. the number of iterations. (b) Success rate of the pose estimation algorithm vs. the offset value of the initial seed for the camera pose

Figure 11.8b shows the success rate of the pose estimation algorithm with respect to the initial seed value. Here,  $R$  denotes the range of the corresponding camera pose component. As this figure shows, choosing the initial seed with less than 57 % offset still gives more than 90 % success rate.

One immediate solution to the initial seed selection is to use the last-known camera pose as the initial seed. Assuming the slow movement of the robot between two successive frames captured by the camera, the last-known configuration of the robot would remain reasonably close to the current actual camera pose. In this experiment, all the components of the camera pose are changing simultaneously in a spiral model. To examine this, the actual camera pose in 60 consecutive frames is synthetically generated based on a spiral model. Then, the actual value of the camera pose in each frame is used for generating the image moments. These image moments are used as the input of the pose estimation algorithm. Figure 11.9 compares the actual and estimated trajectories of the camera pose using the proposed algorithm, where the last estimated pose is used as the initial seed for pose estimation. This figure shows the actual trajectory that camera has traveled and its corresponding estimated trajectory. Here, the estimated pose does not converge to the actual pose in a few cases. Further elaboration of this result reveals that one of the camera pose components on these spots is close to zero, and consequently, its rate of change is too high. It means that in these spots, the last-known camera pose is not a good initial seed for estimating the camera pose in the current frame. Except these spots, the pose estimation error is negligible and the trajectory of the estimated camera pose is the same as the trajectory of the actual camera pose. This experiment indicates that selecting the last-known camera pose yields a high success rate of the pose estimation algorithm.



**Fig. 11.9** Comparison of the actual and estimated camera pose for 60 consecutive frames of the camera moving in a spiral path. In this experiment the last estimated pose (last-known configuration) is used as the initial seed for the pose estimation on the next frame. (a) Trajectories in X-Y plane. (b) Trajectories in pan-tilt plane



**Fig. 11.10** Computation time and optimization error vs. number of iterations for two initial seeds with 40 and 60% offset. (a) Computation time vs. number of iterations. (b) Value of the objective function  $J$  vs. number of iterations

The computation time is one of the major challenges for the iteration-based optimization algorithms. In the proposed algorithm, the computation time depends on the number of iteration, and the initial seed value. Figure 11.10 shows the computation time and the optimization error vs. number of iterations for two different initial seeds with different offsets. Here, the computation time increases linearly by the number of iterations. However, when the algorithm converges to the actual pose, the rising trend of the computation time stops. It means that increasing the number of iterations has no significant effect on the computation time, after that the optimization algorithm converges to the actual pose. Also, as seen in this figure,

the lower offset of the initial seed with respect to the actual pose yields in lower computation time of the algorithm in converging to the actual pose.

For real-time pose estimation, the analytical calculation of the derivatives would speed up the Nelder–Mead’s convergence. However, in real-world applications, there will also be a trade-off between the speed of convergence and the accuracy of results when having noise in the image. This can be verified only experimentally. However, a numerical calculation of derivative was adopted to have a more realistic optimization approach encountered in real scenarios.

## 11.5 Conclusions and Future Work

In this work, a new optimization method for estimating the 4-DOF pose of a sensor module comprised of a conical laser projector and fish-eye camera was proposed. We formulated the projection geometry of our proposed imaging platform, and developed a GUI for generating the synthesized image set, to evaluate our method.

We did different experiments for finding the correlation between the initial seed value and the success rate of the algorithms. Our experiments reveal high success rates of the proposed pose estimation approach, when the initial seed is not too far from the actual pose. Our experiments show that the initial seed value and the success rates of the algorithms are highly correlated. Regardless of this correlation, our experiments imply that the proposed approach has high performance when the initial seed is selected with low offset from the actual pose. Based on this result, we proposed an efficient way of choosing the initial seed, using the last known pose of the robot, in order to minimize the effect of the initial seed value on the performance of the algorithms.

A learning-based method was also adopted to solve the backward mapping for the pose estimation problem, however the results were not satisfactory. The culprit in ANN-based approach to solve an optimization problem with multiple possible solutions of this caliber is that the “random-walk” nature of the ANN can intermittently converge to multiple non-local optimal solutions. Even after setting a small simplex, the random walks associated with the ANN may go beyond a local search. Further rudimentary attractors/penalty functions would be required to force the ANN-based search take place within the initial simplex.

The main contribution of this work comparing with similar techniques is that no prerequisite calibration or setup is required. As the future step, we will do a sensitivity analysis on un-modeled errors in the system, in a simulation environment first and then verify the results via experimentation.

**Acknowledgement** The partial funding provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) is greatly acknowledged.

## Appendix 1: Calculating the Offset Angle

In order to calculate the offset angle between the fixed frame and the camera frame, one needs to rotate a sample intersection point  $p(\theta_0)$  on the pipe's inner wall, with respect to the camera orientation,  $\langle \theta_p \ \theta_d \rangle$ , in opposite direction. To ease the calculation, we chose the point  $p(\theta = 0)$  as the sample point. The coordinates of the point are calculated using Eq. (11.3) as following:

$$p(\theta = 0) = \left( r_p \quad , \quad 0 \quad , \quad \frac{d + r_p \sin(\theta_t) - x_c \sin(\theta_t) + y_c \cos(\theta_t) \sin(\theta_p)}{\cos(\theta_p) \cos(\theta_t)} \right)^T \quad (11.15)$$

Now, by rotating this point around the pipe fixed frame in opposite direction, the coordinate of the corresponding point on the pipe wall,  $cp(\theta = 0)$  is calculated as following.

$$cp(\theta = 0) = R_X(-\theta_p) * R_Y(-\theta_t) * p(0), \quad (11.16)$$

where  $R_X$  and  $R_Y$  are the rotation matrices along the X and Y axes. The offset angle is calculated as projected angle of the point  $cp_0$  in the  $X - Y$  plane as following:

$$\begin{aligned} \delta &= \text{atan} \left( \frac{y_{cp(\theta=0)}}{x_{cp(\theta=0)}} \right) \\ &= \text{atan} \left( -\frac{d \sin \theta_p + y_c + \cos \theta_t + r_p \sin \theta_p \sin \theta_t - x_c \sin \theta_p \sin \theta_t}{\cos \theta_p (r_p - x_c + d \sin \theta_t)} \right) \end{aligned} \quad (11.17)$$

Having the offset angle  $\delta$ , the intersection point  $p(\theta)$  on the pipe's inner wall is mapped to the corresponding point on the image plain with the relative orientation along the x axis equal to  $\theta + \delta$ .

## Appendix 2: Calculating the Analytic Form of the Reference Signal

In order to represent the analytic form of the reference signal, we expand Eq. (11.8), and separate the terms as the trigonometric functions of the variable  $\theta$ . By doing this, the reference function is represented by the following form:

$$ref_a(\theta) = f_1 + f_2 \sin \theta + f_3 \cos \theta + f_4 \sin \theta \cos \theta + f_5 \sin^2 \theta \quad \theta \in [0 \ 2\pi] \quad (11.18)$$

where  $f_1, f_2, \dots$ , and  $f_5$  are the parametric functions of the camera pose  $\langle x_c, y_c, \theta_p, \theta_t \rangle$ . These coefficients are calculated as following:

$$\begin{aligned} f_1 &= x_c^2 + y_c^2 + r_p^2 + \frac{2m_1^2 + m_2^2 + m_3^2}{2m_4} \\ f_2 &= -2y_c r_p + \frac{2m_1 m_3}{m_4} & f_3 &= -2x_c r_p + \frac{2m_1 m_2}{m_4} \\ f_4 &= \frac{m_2 m_3}{m_4} & f_5 &= \frac{m_2^2 - m_3^2}{2m_4} \end{aligned}$$

where

$$\begin{aligned} m_1 &= d - x_c \sin \theta_t + y_c \cos \theta_t \sin \theta_p \\ m_2 &= r_p \sin \theta_t \\ m_3 &= -r_p \cos \theta_t \sin \theta_p \\ m_4 &= \cos^2 \theta_p \cos^2 \theta_t \end{aligned} \quad (11.19)$$

It shows that the  $ref_a$  is the parametric trigonometric function of the variable  $\theta$ .

## References

- Alvarez L, Deriche R, Papadopoulos T, Sánchez J (2007) Symmetrical dense optical flow estimation with occlusions detection. *Int J Comput Vis* 75(3):371–385
- Andrews DP, Bedford J, Bryan PG (2013) A comparison of laser scanning and structure from motion as applied to the great barn at Harmondsworth, UK. *ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci XL-5/W2:31–36* [Online]. Available: <http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-5-W2/31/2013/>
- Chang P (2002) Robust tracking and structure from motion with sampling method. Ph.D. dissertation, Robotic Institute, Carnegie Mellon University
- Cohen BR (2012) Fixing America's crumbling underground water infrastructure. Last viewed: March-2014 [Online]. Available: <http://cei.org/issue-analysis/fixing-americas-crumbling-underground-water-infrastructure>
- Daniilidis K, Spetsakis ME (1997) Understanding noise sensitivity in structure from motion. In: *Visual navigation: from biological systems to unmanned ground vehicles*. Citeseer [Online]. Available: [citeseer.ist.psu.edu/daniilidis97understanding.html](http://citeseer.ist.psu.edu/daniilidis97understanding.html)
- Dementhon DF, Davis LS (1995) Model-based object pose in 25 lines of code. *Int J Comput Vis* 15(1–2):123–141
- Duran O, Althoefer K, Seneviratne LD (2003) Pipe inspection using a laser-based transducer and automated analysis techniques. *IEEE/ASME Trans Mechatron* 8(3):401–409
- Duran O, Althoefer K, Seneviratne LD (2007) Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network. *IEEE Trans Autom Sci Eng* 4(1):118–126
- Fleck MM (1994) Perspective projection, the wrong imaging model. Technical report 95-01. Computer Science, University of Iowa
- Gennery D (2006) Generalized camera calibration including fish-eye lenses. *Int J Comput Vis* 68(3), 239–266 [Online]. Available: <http://dx.doi.org/10.1007/s11263-006-5168-1>
- Hagbi N, Bergig O, El-Sana J, Billingham M (2011) Shape recognition and pose estimation for mobile augmented reality. *IEEE Trans Vis Comput Grap* 17(10):1369–1379

- Iketani A, Sato T, Ikeda S, Kanbara M, Nakajima N, Yokoya N (2007) Video mosaicing based on structure from motion for distortion-free document digitization. In: *Computer Vision—ACCV 2007*. Springer, Berlin, pp 73–84
- Josephson K, Byrod M (2009) Pose estimation with radial distortion and unknown focal length. In: *IEEE Conference on Computer Vision and Pattern Recognition, 2009 (CVPR 2009)*. IEEE, New York, pp 2419–2426
- Kannala J, Brandt S (2004) A generic camera calibration method for fish-eye lenses. In: *Proceedings of the 17th International Conference on Pattern Recognition, 2004 (ICPR 2004)* 1:10–13
- Lagarias JC, Reeds JA, Wright MH, Wright PE (1998) Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J Optim* 9:112–147
- Matsui K, Yamashita A, Kaneko T (2010) 3-d shape measurement of pipe by range finder constructed with omni-directional laser and omni-directional camera. In: *2010 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, New York, pp 2537–2542
- Maverick Inspection Ltd. (2009) Laser pipe profiling. <http://www.maverickinspection.com/Services/LaserPipeProfiling/tabid/135/Default.aspx>. Accessed Feb 2014
- Pipeline and Gas Journal (1999) Crc-evans automatic welding introduces laser vision. <http://www.pipelineandgasjournal.com/crc-evans-automatic-welding-introduces-laser-vision-welding-system>. Accessed 1 March 2014
- Rausch Electronics USA (2014) Rausch laser pipe profiling system. <http://rauschusa.com/products/laserprofile>. Accessed Feb 2014
- Scaramuzza D, Martinelli A, Siegwart R (2006) A flexible technique for accurate omnidirectional camera calibration and structure from motion. In: *IEEE International Conference on Computer Vision Systems, 2006 (ICVS '06)*, pp 45–45
- Simo-Serra E, Ramisa A, Alenya G, Torras C, Moreno-Noguer F (2012) Single image 3d human pose estimation from noisy observations. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, New York, pp 2673–2680
- Spica R, Giordano PR, Chaumette F (2014) Active structure from motion for spherical and cylindrical targets. In: *Proceeding of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, Hong Kong
- Tahri O, Mezouar Y, Chaumette F, Araujo H (2010) Visual servoing and pose estimation with cameras obeying the unified model. In: *Visual Servoing via Advanced Numerical Methods*. Springer, Berlin, pp 231–250
- Tezerjani AD, Mehrandezh M, Paranjape R (2013) Design and development of an optics laser sensor for quality control and inspection of pipes. Faculty of Engineering and Applied Science, University of Regina, internal document, September 2013
- Unnikrishnan P, Thornton B, Ura T, Nose Y, A conical laser light-sectioning method for navigation of autonomous underwater vehicles for internal inspection of pipelines. In: *OCEANS 2009-EUROPE*. IEEE, New York, pp 1–9
- Vincze M, Pichler A, Biegelbauer G, Hausler K, Andersen H, Madsen O, Kristiansen M (2002) Automatic robotic spray painting of low volume high variant parts. In: *Proceedings of the 33rd ISR (International Symposium on Robotics)*, 2002
- Winkelbach S, Molkenstruck S, Wahl FM (2006) Low-cost laser range scanner and fast surface registration approach. In: *Pattern Recognition*. Springer, New York, pp 718–728

# Chapter 12

## Analysis of Tide Gauge Sea Level Time Series

Albert Parker

**Keywords** Data mining • Statistical analysis • Curve fitting • Forecast

### 12.1 Introduction

Global warming refers to a continuing rise in the average temperature of Earth's oceans and air and sea at Earth's surface as a result of the increasing anthropogenic carbon dioxide emission. According to the IPCC (<http://www.climatechange2013.org/report/>), '*Human influence has been detected in warming of the atmosphere and the ocean, in changes in the global water cycle, in reductions in snow and ice, in global mean sea level rise, and in changes in some climate extremes. This evidence for human influence has grown since AR4. It is extremely likely (95–100%) that human influence has been the dominant cause of the observed warming since the mid-20th century.*' Because of this strong political focus on the climate, the analysis of time series of measured parameters linked to the climate has become a central point in the new science of climate change mostly made of modelling results.

Linear and nonlinear regressions and curve fittings have been used to analyse the time series of climate parameters. Classic statistical tools may determine the values of parameters for an equation whose form may be specified to better represent the experimental evidence with a best fit of a set of data values. Linear analyses are widespread, but polynomial, exponential, periodic, and general multiple nonlinear functions are similarly or even more helpful. In addition to computing the optimal values of the parameters to best fit the function to the data, it is also important to

---

A. Parker (✉)  
School of Aerospace, Mechanical and Manufacturing Engineering,  
RMIT University, Melbourne, VIC, Australia  
e-mail: [albertparker@y7mail.com](mailto:albertparker@y7mail.com)



compare the data points with the fitted equation. The distribution of residual values and the computation of the auxiliary statistic help in understanding the quality of the fitting.

Worth of mention, statistic is applied to a set of data generally assumed to be a truly accurate representation of the tracked phenomena. However, every measurement has his error, but random and systematic, and there may be significant biases due to many factors not under control, for example the modified heat release and heat storage close to a thermometer location or the land motion where a tide gauge is located.

Unfortunately, the time series of climate parameters are often characterised by scattered measurements of poor quality. Statistic is not helpful to address these issues, and the reliability of the results from the analysis drastically reduces. Many '*observational*' parameters such as the global air and sea surface temperature or the global sea level are not truly measured products, but the result of a complex computational procedure from supporting experimental information is particularly poor especially for the past. What has not been measured cannot be guessed with the accuracy needed in the climate analyses.

This chapter provides some simple tools common to the most part of the statistical packages and their sample use to time series of sea levels. According to the global warming theory, the increased carbon dioxide emission produces an increased heat uptake, and the most part of this heat uptake goes to warm the oceans that consequently increase in volume mostly by thermal dilatation and in less extent by the addition of fresh water from the melting of ices. By applying the proposed technique, the reader may understand if there is any effect of global warming on the recorded time series.

## 12.2 A Technique for Analysing Sea Levels

Sea levels are very well known to oscillate in time following many periodicities, with the detectable longest periodicity obviously linked to the length of the measured record. The sea level measured by tide gauges is not the absolute sea level, as the land locally may move upwards or downwards because of isostasy or subsidy. Other perturbing factors like the change of the instrument or the relocation of the instrument or works about the tide gauge location may all reduce the quality of the record. At the best, tide gauge records are available from the mid-late 1800s and in very few selected locations worldwide. Usually, monthly average mean sea levels are considered for each individual location.

Generally speaking ([http://www.fig.net/pub/monthly\\_articles/july\\_2010/july\\_2010\\_hannah.pdf](http://www.fig.net/pub/monthly_articles/july_2010/july_2010_hannah.pdf)), the issues with the analysis of tide gauges include tide gauge errors, the differences between the reading of the tide gauge and the actual sea level; error in the datum of the tide gauge; analysis errors, as the consideration of records that are short or incomplete; geophysical effects, including land motion and harbour construction. This chapter is only dealing with the analysis the data recorded by the tide gauge as a true representation of the monthly mean sea level.

## 12.3 Tide Gauge Measurements

Tide gauges as the one at Fort Denison in Sydney, NSW, Australia or the one in Fremantle, Western Australia, Australia are measuring tides since the 1800s. Tides have been measured at Fort Denison since 1857 on completion of the fort, initially using a bench mark cut into the stonework. From 1867 consecutive measurements with different instruments were used as tide measuring technology developed. Measurements in Fremantle are more recent.

Sydney has two tide gauge records in Fort Denison, SYDNEY, FORT DENISON [PSMSL; <http://www.psmsl.org/data/obtaining/>, ID 65, latitude  $-33.850$ , longitude  $151.233$ , Time span of data 1886–1993, Completeness (%) 100] and SYDNEY, FORT DENISON 2 [PSMSL; <http://www.psmsl.org/data/obtaining/>, ID 196, latitude  $-33.855$ , longitude  $151.226$ , Time span of data 1914–2012, Completeness (%) 98]. The two tide gauges have almost 80 years of successful overlapping and the integration of the two records in a composite record spanning 1886–2012 without any gaps does not pose any quality issue.

Fremantle has only one tide gauge, FREMANTLE [PSMSL; <http://www.psmsl.org/data/obtaining/>, ID 111, latitude  $-32.065556$ , longitude  $115.748139$ , Time span of data 1897–2012, Completeness (%) 92]. This tide gauge has significant gaps, and the way the gaps are filled interpolating neighbouring years may change the result of a statistical analysis. In a first instance the missed data are simply neglected.

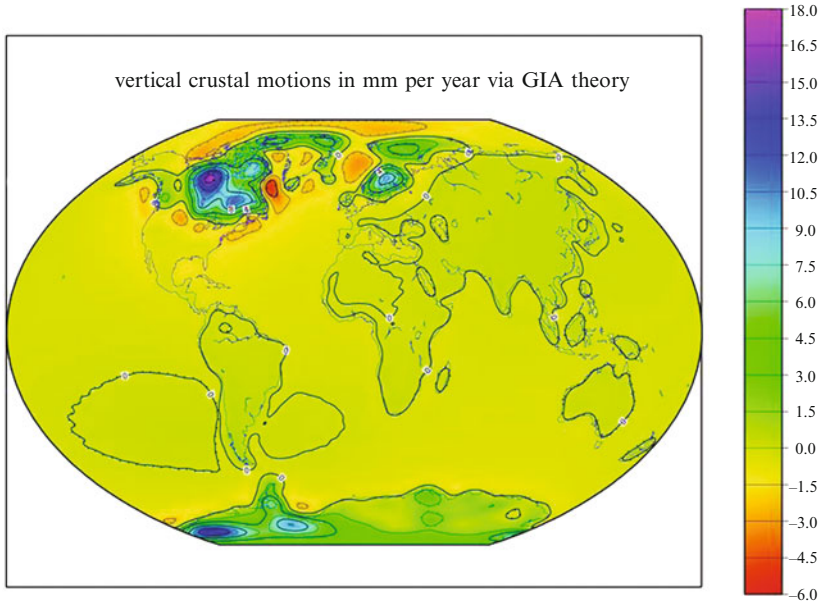
Tide gauges do not measure the absolute sea level but the sea level relative to the land. Being the land subject to isostasy or subsidy, sea levels measured by tide gauges may rise or fall simply because the land is moving downwards or upwards.

Figure 12.1 (from [http://en.wikipedia.org/wiki/Glacial\\_isostatic\\_adjustment](http://en.wikipedia.org/wiki/Glacial_isostatic_adjustment)) presents computations for the glacial isostatic adjustment. The land may clearly be affected by velocities comparable with the velocity of rise or fall of the oceans. The figure presents a modelling result of present mass change due to post-glacial rebound and the reloading of the ocean basins with seawater. Blue and purple areas indicate rising due to the removal of the ice sheets. Yellow and red areas indicate falling as mantle material moved away from these areas in order to supply the rising areas, and because of the collapse of the fore bulges around the ice sheets. Models are absolutely not accurate but they give an idea of a general trend.

Additional land motions exist because of more local phenomena, both natural and man-made, as it is the case of the large subsidence along the coast of the US states of Texas or Louisiana or in the Pacific atolls.

GPS monitoring (<http://www.sonel.org/>) and local surveying may help understanding the land motion contribution to the tide gauge signal, but in the limit of the accuracy of the measurement of the vertical position, the stability of the positioning of the tide gauge vs. the GPS datum, the recalibration issues and the reduced length of the record.

The present distribution of the PSMSL stations (<http://www.psmsl.org/data/obtaining/>) is apparently sufficient with about 2,000 stations in the PSMSL database for which monthly and annual mean sea level information is available. However,



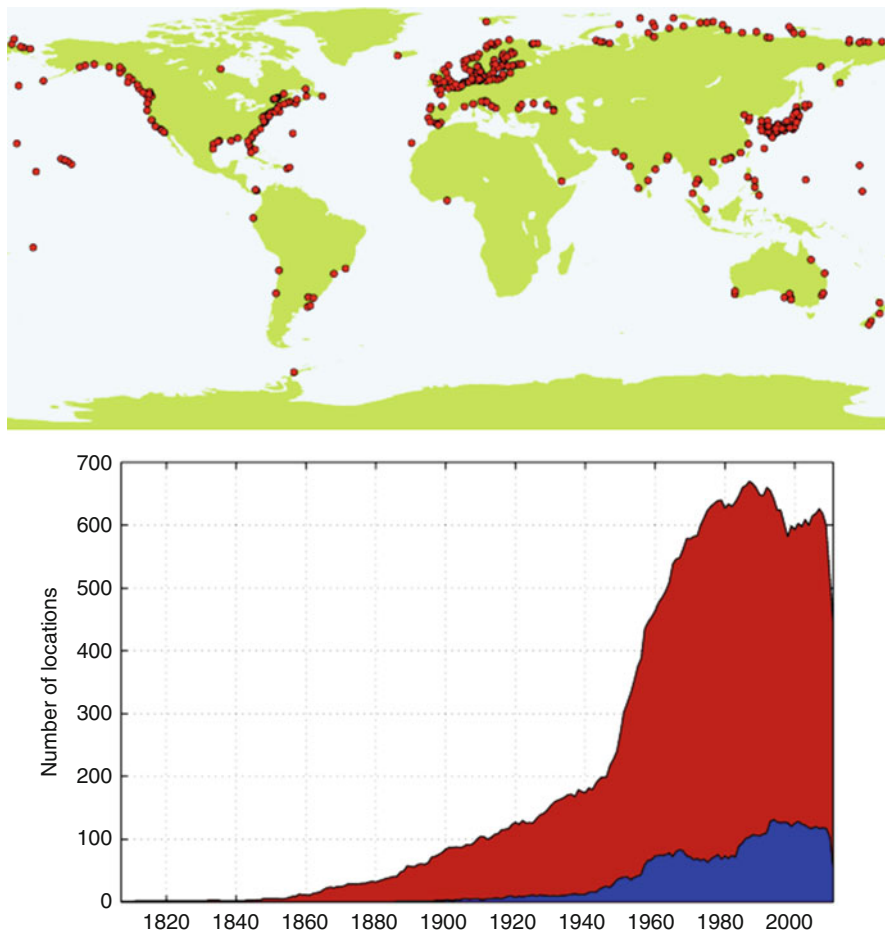
GMT 2010 Oct 26 18:37:21 deg ord 70 paulson GIA appx. uplift rates in mm per yr

**Fig. 12.1** Glacial isostatic adjustment (from [http://en.wikipedia.org/wiki/Glacial\\_isostatic\\_adjustment](http://en.wikipedia.org/wiki/Glacial_isostatic_adjustment))

many stations have historically only been measured for some months or years and the longer records are very scattered and located only in few selected areas. The geographical distribution of the longer records contains significant geographical bias towards the Northern Hemisphere (Europe, Japan and USA). The tide gauge records do not cover the same time period of time, and many of them have quality issues. Only the long-term time series may be used to infer sea level trends if there are no quality issues.

Figure 12.2 (from <http://www.psmsl.org/data/obtaining/>) shows the locations of sites with not less than 40 years of data. The coverage of the map is everything but good. The coverage drops dramatically requiring at least 60 years of data that is the minimum necessary to understand a trend (Parker 2013a, b, c, d, e, f, g, 2014; Parker et al. 2013). The figure also presents the number of PSMSL stations with data available from Northern and Southern Hemisphere (NH and SH respectively). Brown colour corresponds to the mean sea level data available from NH and blue colour corresponds to the mean sea level data available from SH. The most part of the world oceans is definitively not covered enough.

Sydney is the only tide gauge covering the South Pacific since more than 100 years. Sydney and Fremantle are the two tide gauges covering the southern hemisphere extending more than 100 years.



**Fig. 12.2** Demography of PSMSL stations (from <http://www.psmsl.org/data/obtaining/>, updated: 22 August 2012): stations with at least 40 years of data (*top*) and stations from Northern and Southern Hemisphere (*bottom*)

## 12.4 Linear Model

Traditionally, tide gauge data have been analysed by using a linear model (Parker 2013a, b, c, d, e, f, g, 2014; Parker et al. 2013). This basic linear model is the simple linear regression which has only one independent variable:

$$y = \beta_0 + \beta_1 \cdot x + \varepsilon \quad (12.1)$$

where  $\beta_0$ ,  $\beta_1$  are coefficients and  $\varepsilon$  is the random error. Linear regression fits a straight line to a given data set. The fitting model is written as:

$$y = \beta_0 + \beta_1 \cdot x + \varepsilon$$

$$\varepsilon \approx N(0, \sigma^2)$$

where  $\beta_0$  is the  $y$  intercept,  $\beta_1$  is the slope, and  $\varepsilon$  is the error term. The error term represents the unexplained variation in the dependent variable and is usually assumed to have a mean value of zero.

To estimate the parameters, the chi-square minimization or “*weighted least-square*” method is used. The sum of the squares of the deviations between the theoretical curve and the experimental points for a range of independent variables is minimised. After fitting, the model may then be evaluated using hypothesis tests and plotting of residuals.

Figure 12.3 presents the sea levels results for the Sydney and Fremantle tide gauges from the basic linear analysis. The figure presents the sea levels in Sydney on the left and Fremantle on the right. Figure 12.3a shows the monthly sea levels (from <http://www.psmsl.org/data/obtaining/>), Fig. 12.3b the linear trend, Fig. 12.3c the oscillations of the monthly sea levels about the linear trend and Fig. 12.3d the land motion in nearby GPS stations (from <http://www.sonel.org/>).

For what concerns Fremantle, the nearby GNSS Stations from SONEL are located at a distance from the tide gauge of 32,074 m. For what concerns Sydney, the nearby GNSS Stations from SONEL are located at a distance from the tide gauge of 10,775 m. In addition to the vertical motion of the tide gauge, other factors as harbour constructions or the use of successive instruments as tide measuring technology developed may have an influence on the recorded signal.

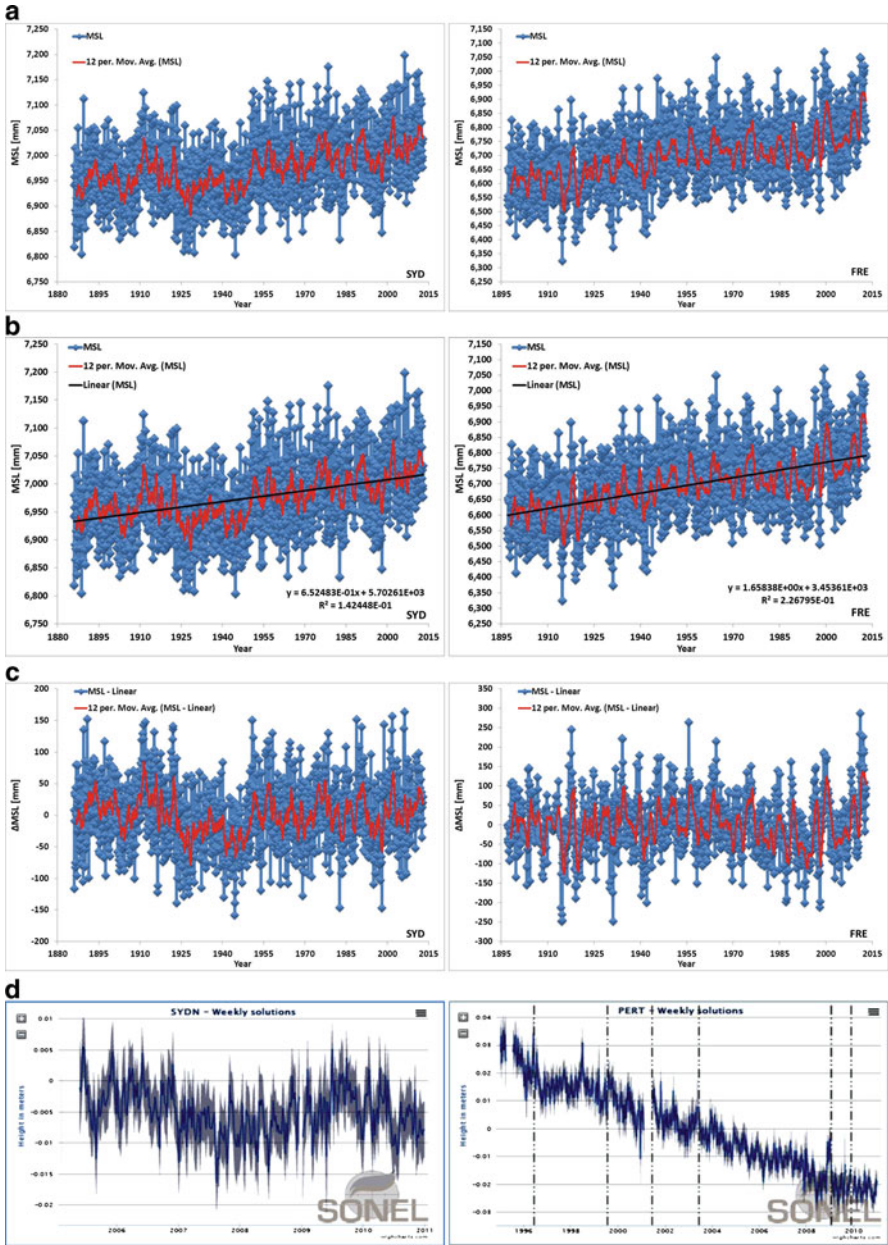
The classic approach to analyse sea level data is to use a linear fit. In Eq. (12.1),  $x$  is the time,  $y$  is the sea level, and  $\beta_0$  and  $\beta_1$  are constants to compute the average sea level rise over the period of observation.

The R-squared value of the fitting is not a measure of the accuracy of the trend line but only a parameter related to the legitimate oscillations about the linear trend. The coefficient of determination compares the estimated and actual  $y$ -values, and ranges in value from 0 to 1. If it is 1, there is no difference between the estimated  $y$ -value and the actual  $y$ -value. If it is 0, the regression equation is not helpful in predicting a  $y$ -value. Being very well known that the sea levels oscillate, it should not be a surprise that  $R^2$  is not unity.

The present SLR is 0.652 mm/year for Sydney and 1.658 mm/year for Fremantle. The vertical motion of the Sydney and Fremantle tide gauges is unknown, but the nearby GPS stations show significant subsidy over the last decade with rates of  $-0.89 \pm 0.65$  mm/year and  $-2.99 \pm 0.34$  mm/year respectively.

The presence or absence of acceleration in the measured signal is detectable by observing the graph of the monthly deviations vs. this linear trend. The monthly sea levels oscillates about the linear trend, and these oscillations are periodical. Therefore, there is no visual evidence of acceleration in the recorded signal.

The accuracy and reliability of each measurement is not an issue that may be addressed with a statistical analysis. The only accuracy parameter that may be considered in the statistical analysis is the percentage of gaps (missed data) in



**Fig. 12.3** Sea levels in Sydney (*left*) and Fremantle (*right*): **(a)** monthly sea levels; **(b)** linear trend; **(c)** oscillations of monthly sea levels about the linear trend; **(d)** land motion in nearby GPS stations



the time series and the length of the record. Biasing to the instrument reading for multiple reasons is difficult to be detected by simply considering the recorded data.

Because temperatures and sea levels are very well known to oscillate with up to quasi-60 years detected in the available instrumental data, the length of the record and the time when the measurement starts play a significant role in determining the longer term sea level rate of rise. For example, in a location with a perfect 60 years periodical oscillation, while a record of 60 years length may include a complete oscillation, with less than 60 years the rate of rise may be overestimated or underestimated depending on the start of the measurements about a peak or a valley of the oscillation. Worth of mention, also records longer than 60 years have a result that depends on a lesser extent on the starting point.

This classic approach requires long term records to avoid extrapolation of the sea level rate. The linear approximation of the variation of sea level from only a multi-decadal oscillation may lead to wrong results. classic approach requires records of enough length to avoid computing as a longer term sea level rate of rise the variation of sea level from a valley to a peak of a multi-decadal oscillation, therefore using as the longer term linear trend a wrong curve (Parker 2013a, b, c, d, e, f, g, 2014; Parker et al. 2013). The time histories of linear fitting parameters are an additional tool that may help clarifying the influence of the natural oscillations on the result of the basic linear model.

## 12.5 Time Histories of Linear Fitting Parameters

A fully satisfying monitoring technique may be developed by using linear fittings to compute reference velocity and acceleration of sea levels and detect the periodicities of the natural oscillations. The dependent  $y$ -values are the monthly average sea levels and the independent  $x$ -values are the time in years. According to Eq. (12.1), the rate of rise of sea level SLR is equal to  $\beta_1$ . From a recorded time series  $x_i, y_i, i = 1, \dots, N$  with  $i = 1$  the oldest data point and  $N$  the latest data point, it is possible to compute by linear fitting many rates of rise  $SLR_{j,k}$  depending on the initial and final points  $j, k$  selected in the distribution.

The calculations for the sea level rate of rise  $SLR_{j,k}$  is based on the formula:

$$SLR_{j,k} = \frac{\sum_{i=j}^k (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=j}^k (x_i - \bar{x})^2} \quad (12.2)$$

In this equation  $\bar{x}$  and  $\bar{y}$  are the sample means and  $j$  and  $k$  are the indices of the first and last record of the measured distribution considered for the particular SLR estimation.

The LINEST function in Microsoft Excel gives the desired statistics for a line by using the “least squares” method to calculate a straight line that best fits the data and returns an array that describes the line.

$$SLR_{j,k} = \text{LINEST}(y_j : y_k, x_j : x_k)$$

If at a certain time  $x_k$ ,  $x_j$  is taken as  $(x_k - 20)$  to compute the  $SLR_{20}$  with time windows of 20 years, it is taken as  $(x_k - 30)$  to compute the  $SLR_{30}$  with time windows of 30 years, and it is taken as  $(x_k - 60)$  when computing the  $SLR_{60}$  with a time window of 60 years.  $x_j$  is then simply taken as  $x_1$  when computing the  $SLR_A$  by using all the data available at any time. This way, from a measured distribution  $x_i, y_i$  for  $i = 1, \dots, N$  it is possible to estimate the time histories of  $SLR_{20}$ ,  $SLR_{30}$ ,  $SLR_{60}$ , and  $SLR_A$  (Parker 2013a, b, c, d, e, f, g, 2014; Parker et al. 2013).

Providing that more than 60–70 years of continuously recorded data, without any quality issues, are available in a given location, the  $SLR_{A,m}$  returns a reasonable estimation of the velocity of sea level change at the present time  $x_m$ . The acceleration may then be computed as

$$SLA_m = \frac{SLR_{A,m} - SLR_{A,m-1}}{x_m - x_{m-1}} \quad (12.3)$$

This conventional velocity and acceleration will clearly oscillate, and their time history, rather than a single value, will be a parameter of interest (Parker 2013a, b, c, d, e, f, g, 2014; Parker et al. 2013).

Alternative acceleration parameters may be obtained by using a polynomial fitting of the data in the recorded time series  $x_i, y_i, i = 1, \dots, N$ . A second order polynomial fitting (parabolic fitting) returns the average acceleration over the length of the considered time window as twice the second order coefficient. The LINEST function may be combined with other functions to calculate the statistics for other types of models that are linear in the unknown parameters, including polynomial, logarithmic, exponential, and power series. In the specific of the parabolic fitting, it is possible to define an acceleration parameter.

$$SLA_{j,k} = 2 \cdot \text{LINEST}(y_j : y_k, x_j : x_k \wedge \{1, 2\})$$

More than one value of SLA is indeed the time histories of  $SLA_{20}$ ,  $SLA_{30}$ ,  $SLA_{60}$ , and  $SLA_A$  that may be of interest.

The availability of a long record permits to understand which conclusions could have been inferred from measurements in the same location if the measurements could have started much later. At a certain time  $x_N$ , the graph of  $SLR_{k,N}$  vs. the record length  $x_N - x_k$  gives an idea of the errors in evaluating the rate of rise in shorter records.

In a case with non-accelerating tide gauge records as the norm,  $SLR_{1,N}$  returns the present sea level rate of rise, and the graphs of  $SLR_{j,k}$  and  $SLA_k$  help to confirm the lack of any acceleration and understand the longer periodicities in the oscillations. In



a case of accelerating tide gauge records, the approach would confirm the presence of acceleration in the form of a constantly increasing  $SLR_{j,k}$  and a constantly positive  $SLA_k$  rather than oscillating values about the longer term trend and the zero.

Figure 12.4 presents the time histories for Sydney and Fremantle of the above velocity and acceleration parameters based on the linear model. The figure presents the sea levels in Sydney (left) and Fremantle (right). Figure 12.4a is the time history of SLR computed with 20, 30 and 60 years windows. Figure 12.4b is the time history of SLR computed by using all the data available at any time—this is the conventional sea level velocity. Figure 12.4c is the time history of SLA computed as the time rate of change of the above SLR. Figure 12.4d is the apparent present velocity that could have been computed with the linear model in case the measuring record could have started in different times.

Figure 12.4a evidences the presence of multi-decadal oscillations of different periodicities in Sydney and Fremantle. Figure 12.4b shows the oscillating conventional sea level velocity, and Fig. 12.4c the conventional sea level acceleration oscillating between positive and negative values. Figure 12.4d shows the huge errors that could have been made using short records to assess the rate of rise of sea levels.

Figure 12.5 presents for Sydney (left) and Fremantle (right) in Fig. 12.5a the time history of SLA computed with 20, 30 and 60 years windows, in Fig. 12.5b the time history of SLA computed by using all the data available at any time—this is the average sea level acceleration since the start of the record and in Fig. 12.5c the apparent present acceleration that could have been computed as twice the second order coefficient of a polynomial fitting in case the measuring record could have started in different times. The accelerations clearly evidence the many oscillations similarly to the velocities and the positions.

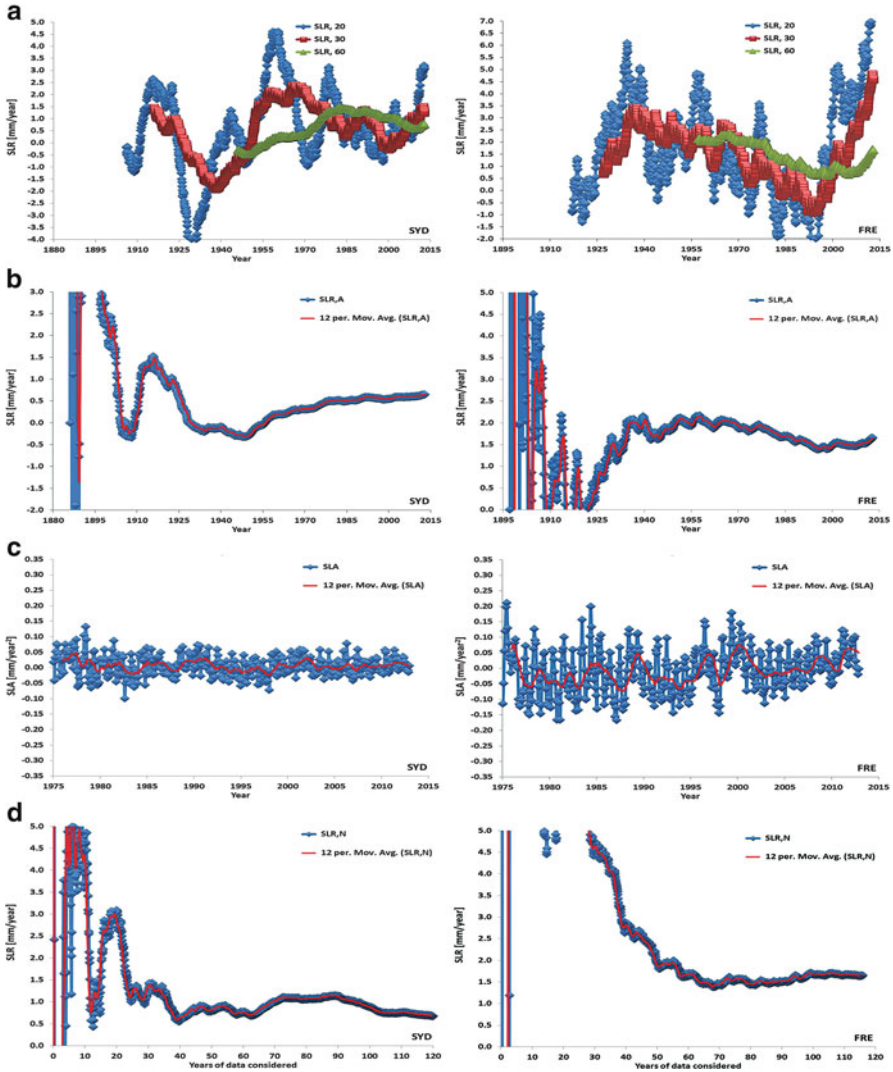
## 12.6 Nonlinear Model

The linear model may be further refined by introducing a nonlinear model based on multiple fittings. This model may be used to forecast future sea levels including the effect of the periodic oscillations, obviously on the basis of what has been measured up to a certain time. The nonlinear regression model is

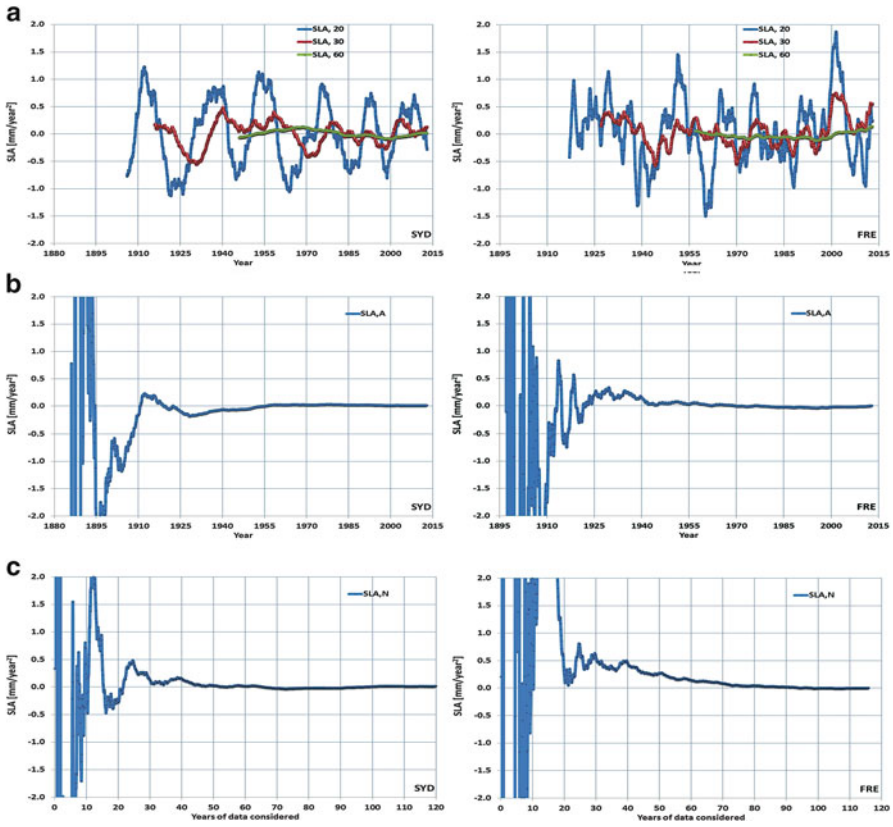
$$y = f(x, \vec{\theta}) + \varepsilon$$

Where  $x$  is the independent variable and  $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_j)$  are  $j$  parameters. In this equation, fit parameters are related to the independent variable in a nonlinear fashion.

When the relationship between dependent and independent variable is nonlinear there are no explicit solutions for the equation parameters and parameter values are estimated by using an iterative process.



**Fig. 12.4** Sea levels in Sydney (*left*) and Fremantle (*right*): (a) time history of SLR computed with 20, 30 and 60 years windows; (b) time history of SLR computed by using all the data available at any time—this is the conventional sea level velocity; (c) time history of SLA computed as the time rate of change of the above SLR; (d) apparent present velocity that could have been computed with the linear model in case the measuring record could have started in different times



**Fig. 12.5** Sea levels in Sydney (*left*) and Fremantle (*right*): (a) time history of SLA computed with 20, 30 and 60 years windows; (b) time history of SLA computed by using all the data available at any time—this is the average sea level acceleration since the start of the record; (c) apparent present acceleration that could have been computed as twice the second order coefficient of a polynomial fitting in case the measuring record could have started in different times

A common measure of the quality of nonlinear fitting parameters is the chi-square value:

$$\chi^2 = \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2$$

Here  $w_i$  is the weighting coefficient,  $y_i$  are the experimental data points and  $\hat{y}_i$  are the theoretical points. The “best-fit” parameter values are obtained by minimising chi-square.

We propose here a method to analyse tide gauge records by fitting with the available data with a line and multiple sinusoidal functions and use these data to fill the gaps eventually available in the record. We apply these methods to the tide gauge records of Sydney and Fremantle.

We do use a line and multiple sinusoidal functions fitting to better describe the tide gauge behaviour. The fitting curve is:

$$y = (y_0 + a \cdot x) + \sum_{i=1}^n \left[ y_{0,i} + A_i \cdot \sin \left( \pi \cdot \frac{x - x_{c,i}}{w_i} \right) \right] \quad (12.4)$$

In this equation,  $n$  is the number of sinus functions and  $y_0$ ,  $a$ ,  $y_{0,i}$ ,  $A_i$ ,  $x_{c,i}$ ,  $w_i$  are the fitting coefficients.

The composite tide gauge record for Sydney does not pose any quality issue having no gaps. Fremantle has significant gaps, and the way the gaps are filled interpolating neighbouring years may change the result of a spectral analysis of the time series and affect the periodical oscillations detected.

In the present approach, the parameters of Eq. (12.4) are computed first from the measured data, then the gaps in the measured data are filled with the estimation from Eq. (12.4) and new parameters of Eq. (12.4) are computed. This iterative procedure is the best approach to avoid introducing unrealistic periodicities of oscillations.

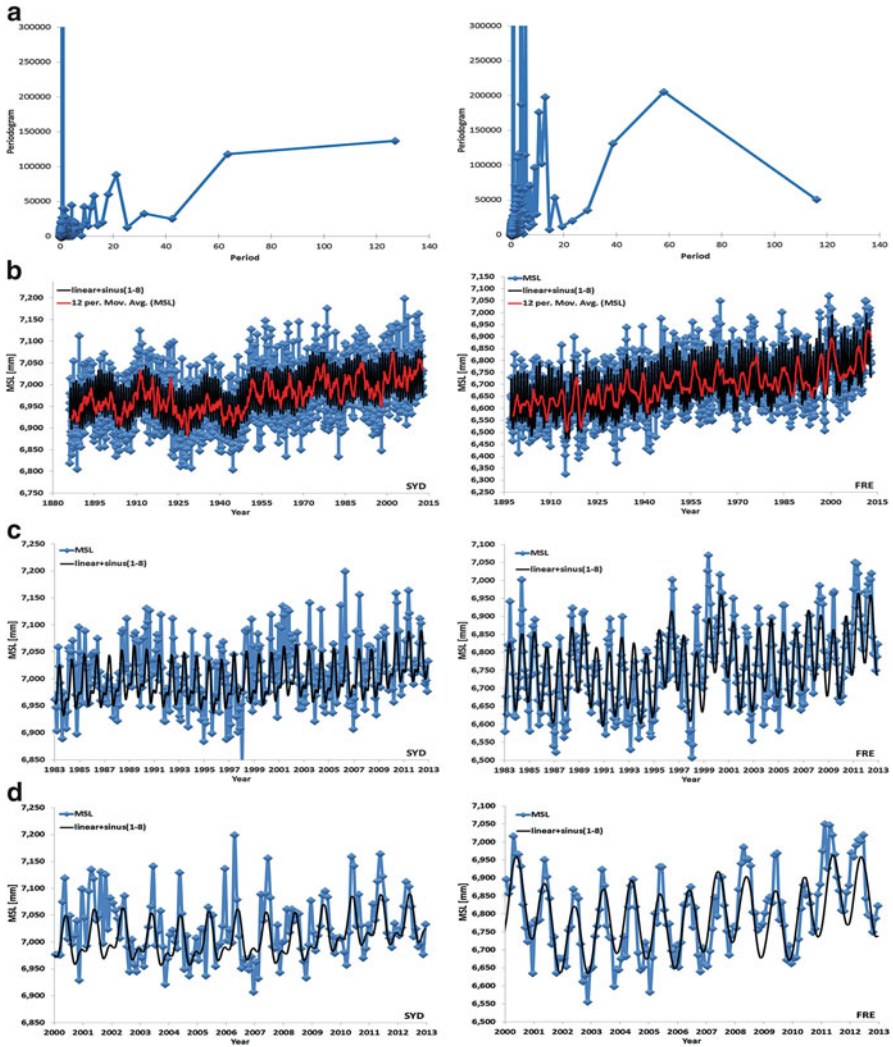
The Solver suite of commands in Microsoft Excel may be used to compute the fitting coefficients  $y_{0,i}$ ,  $A_i$ ,  $x_{c,i}$ ,  $w_i$  by minimising the chi-square value. Solver is a process of changing the values in cells to see how those changes affect the outcome of formulas on the worksheet. Solver works with a group of variable cells that participate in computing the formulas in the objective and constraint cells. However, add-in as XLSTAT (Parker 2013e) may certainly make easier to perform the nonlinear fitting or run a spectral analysis to compute the spectral density or the periodogram of the time series.

Figure 12.6 presents the results for Sydney and Fremantle. The figure presents the periodogram from the spectral analysis of the monthly departures vs. the linear trend of Fig. 12.3 then the measured monthly sea levels and the values computed with the fitting with a line and multiple sinusoidal functions of Eq. (12.4).

The spectral analyses of the Sydney and Fremantle time series are significantly different. The spectral analysis allows to transform a time series into its coordinates in the space of frequencies, and then to analyse its characteristics in this space. From these coordinates it is possible to extract the magnitude and the phase and build representations such as the periodogram or the spectral density.

The spectral representation of a time series  $\{X_t\}$ , ( $t = 1, \dots, n$ ), decomposes  $\{X_t\}$  into a sum of sinusoidal components with uncorrelated random coefficients. From there we can obtain decomposition the auto-covariance and autocorrelation functions into sinusoids. The spectral density corresponds to the transform of a continuous time series. However, we usually have only access to a limited number of equally spaced data, and therefore, we need to obtain first the discrete Fourier coordinates (cosine and sine transforms), and then the periodogram. From the periodogram, using a smoothing function, we can obtain a spectral density estimate which is a better estimator of the spectrum.

The iterative process to compute the parameters of the sinus fitting in Eq. (12.4) may return periodicities different from the periodogram especially in the longer periodicities because of lack of data and noise and the different algorithm. Noise



**Fig. 12.6** Sea levels in Sydney (*left*) and Fremantle (*right*): (a) periodogram of the monthly oscillations about the linear trend; (b), (c) and (d) measured and computed monthly sea levels by the proposed multiple nonlinear fitting over different time windows—better accuracy may be obtained by using other periodic functions

in the measurements and length of the record limit the accuracy of the detection especially of the longer periodicities.

Table 12.1 presents the coefficients of Eq. (12.4) for Sydney and Fremantle.

The periodicities of the sinusoidal oscillations of (12.4) considered in Sydney are 1.00, 0.50, 58.07, 20.04, 12.93, 4.19, 30.85 and 18.39 years, with decreasing amplitudes 37.54, 21.56, -13.23, 12.68, 10.58, 6.76, 6.56 and -3.96 mm.

**Table 12.1** Coefficients of Eq. (12.4) for Sydney and Fremantle

<i>Sydney</i>											
$y_0$	5703.6	$y_{0i}$	0.0392	-0.0241	-0.1997	0.0267	0.0001	-0.0544	-0.003	-0.0092	
$a$	0.652	$x_{ci}$	-1.2521	0.102	1509.7	120.9778	-756.3	79.2851	-31.4733	19.3755	
		$W_i$	0.5003	0.25	15.4263	29.0358	10.0214	9.1953	6.4644	2.0947	
		$A_i$	37.5435	21.5583	6.5628	-13.2312	12.6756	-3.9577	10.5844	6.7593	
<i>Fremantle</i>											
$y_0$	3457.3	$y_{0i}$	-0.0081	0.2473	-0.0081	-0.008	-0.0084	2.7812	-0.0081	-0.3093	
$a$	1.6565	$x_{ci}$	0.6338	4.7161	-17.0552	12.3597	2.3347	-312.3075	70.4469	-13.3771	
		$W_i$	0.4999	0.2301	2.793	1.8029	2.0694	34.0483	6.2119	1.8847	
		$A_i$	101.4107	3.5149	25.5496	21.2954	-21.8308	23.2554	19.25	13.468	

Especially the longer periodicities are sensitive to the parameters of the iterative nonlinear fitting procedure. The periodicities of the sinusoidal oscillations of (12.4) considered in Fremantle are 1.00, 5.59, 68.10, 4.14, 3.61, 12.42, 3.77 and 0.46 years, with decreasing amplitudes 101.41, 25.55, 23.26, -21.83, 21.30, 19.25, 13.47 and 3.51 mm.

The oscillations are mostly different in amplitude and period (but also in phase) in Fremantle (Indian Ocean) and Sydney (Pacific Ocean). Apart from the common 1 and 0.5 years periodicities and the close quasi-60 years oscillation, then Sydney has also strong oscillations of periodicity about 20 years that are not detectable in Fremantle. The longer periodicity of Sydney is very close to the record length and clearly indicates the need of more data to detect the periodicities above the quasi-60 years that is already difficult to clarify when less than two full cycles available.

The shape of the periodic function may be optimised for a better accuracy (reduction of the residual error).

The reconstruction of the tide gauge record for small gaps as in the case of Fremantle does not pose particular issues. However, when the missed information exceeds the recorded information, the reliability of the infilling techniques drastically reduces.

Breakpoint alignment may be necessary in cases of abrupt changes in a record resulting from perturbing events as relocation of the tide gauge, destruction of the tide gauge, construction works close to the tide gauge and any other biasing issue that may be identified in the record.

The extension of the tide gauge record to cover more than the length of the measurements is not recommended.

Testing of the local statistic vs. a regional expectation is similarly not recommended, being the sea level behaviour still very far from being completely understood locally and globally.

## 12.7 Conclusions

A simple but reliable methodology has been shown to compute the rate of rise of sea levels and assess the presence or absence of accelerations or decelerations.

The methodology builds on the linear fittings of the available data over different time windows and the analysis of the time histories of velocity and acceleration parameters.

Multiple nonlinear fittings may help to infill gaps in incomplete records, detect the periodicities of the natural oscillations about a longer term trend and eventually predict the future trends on the basis of what has been recorded as to date.

This forecasting procedure is legitimate if a specific tide gauge of enough quality and length and all the neighbouring tide gauge of about same quality and length are all acceleration free.

## Key Symbols

$y$	Sea level
$x$	Time
$\beta$	Fitting coefficient
$\varepsilon$	Error
$\sigma$	Standard deviation
SLR	Sea level velocity
SLA	Sea level acceleration
$w$	Weight or fitting coefficient
$a$	Fitting coefficient
$A$	Fitting coefficient
$x_c$	Fitting coefficient

## References

- Parker A (2013a) Natural oscillations and trends in long-term tide gauge records from the Pacific. *Pattern Recognit Phys* 1:1–13. doi:[10.5194/prp-1-1-2013](https://doi.org/10.5194/prp-1-1-2013)
- Parker A (2013b) Oscillations of sea level rise along the Atlantic coast of North America north of Cape Hatteras. *Nat Hazards* 65(1):991–997. doi:[10.1007/s11069-012-0354-7](https://doi.org/10.1007/s11069-012-0354-7)
- Parker A (2013c) Sea level trends at locations of the United States with more than 100 years of recording. *Nat Hazards* 65(1):1011–1021. doi:[10.1007/s11069-012-0400-5](https://doi.org/10.1007/s11069-012-0400-5)
- Parker A (2013d) The non-linear, naturally oscillating pattern of sea-levels in the Chesapeake Bay, East Coast, USA. *Nonlinear Eng* 2(1–2):1–10. doi:[10.1515/nleng-2013-0008](https://doi.org/10.1515/nleng-2013-0008)
- Parker A (2013e) Apparent hot and cold spots of acceleration along the Atlantic and Pacific coasts of the United States. *Nonlinear Eng*. doi:[10.1515/nleng-2013-0012](https://doi.org/10.1515/nleng-2013-0012)
- Parker A (2013f) Why global warming went missing since the year 2000. *Nonlinear Eng*. doi:[10.1515/nleng-2013-0017](https://doi.org/10.1515/nleng-2013-0017)
- Parker A (2013g) Minimum 60 years of recording are needed to compute the sea level rate of rise in the western south Pacific. *Nonlinear Eng*. doi:[10.1515/nleng-2013-0011](https://doi.org/10.1515/nleng-2013-0011)
- Parker A (2014) Confirming the lack of any sea level acceleration around the Australian coastline. *Nonlinear Engineering* 3(2):99–105, doi: [10.1515/nleng-2013-0025](https://doi.org/10.1515/nleng-2013-0025)
- Parker A, Saad Saleem M, Lawson M (2013) Sea-level trend analysis for coastal management. *Ocean Coastal Manage* 73:63–81. doi:[10.1016/j.ocecoaman.2012.12.005](https://doi.org/10.1016/j.ocecoaman.2012.12.005)



# Index

## A

- Acoustic filter, 235, 246
- Active vibration control. *See* Axially translating cable systems
- Amplitude synchronization
  - application, control strategy, 98–100
  - chaotic motion, 100–103
  - control strategy, 95, 97, 98, 102
  - input control, 97, 99, 101, 102, 104
  - reference signal, 97, 99, 101, 102, 104
- Autodriver algorithm, 10, 29
- Autonomous vehicle steering dynamics
  - changes, transient and steady-state responses
    - differences, instantaneous and steady-state values, 11–13, 15, 16
    - equations of motion, 10–11
    - global frame, angular orientation, 13–14
    - instantaneous and steady-state paths, 16, 17
    - integration, 15
    - magnification, stage depicts, 16–18
    - proximity, 29
    - realistic step change, 10
    - rotation center, 14–15
    - steady-state time history, 11, 12
    - time history, 11
    - understeer passenger car, 9
    - velocity components, 15
  - SSDS (*see* Steady-state dynamic steering (SSDS))
- Axial crush testing, 220–221
- Axially translating cable systems
  - aerospace and aeronautical structures, 81
  - control strategy development, 91–93

- deployment process, 82
- description, 81
- dynamic behavior and stabilities, 82
- equations of motion
  - constant velocity, 87–88
  - displacement field, 85
  - Hamilton's principle, 86
  - kinetic energy, 85
  - Lagrangian function, 86
  - model, 84, 85
  - nonlinear differential governing equation, 89
  - strain energy, 86
  - von Karman-type equations, 85
- FSMC, 83, 84
- Galerkin method, discretization, 90
- non-dimensionalization, 89
- pinned–pinned boundary, 90–91
- SMC, 83
- spacecraft antenna, 82
- vibration control (*see* Vibration control)
- volume control, 83

## B

- Bézier curves
  - Bernstein polynomials, 121, 122
  - CAD software, 121
  - 2D Cartesian coordinate system, 121
  - linear, quadratic and cubic, 121, 122
  - Newton interpolation, 121, 123
  - parametric form, 123
  - step formulas, 124
  - wood/plastic, protrusion, 120
- Bode plot, 247

**C**

## Chaos

- motion, 94–96
- piecewise linear systems, 61
- Timoshenko beams, 82

## Circle-projecting collimated laser

- analytic form, reference signal, 292–293
- offset angle calculation, 292
- pipelines (*see* Pipelines)

## Control strategy development

- desired/reference vibration, 92
- fuzzy rule, 92
- Galerkin method, 91–92
- governing equation, 91, 93
- third-order Galerkin discretization, 93

## Convolution integral, 150

## Coordinate transformation, 36

## Crash safety

- high-speed, 194
- severity, head injuries, 172
- testing, 170

## Curve fittings

- linear and nonlinear regressions, 295
- pipeline, 284

**D**

## Data mining, 197

## Deformable registration

- HIC value requirements, 188–189
- hood assembly, 198
- optimal waveform, 177

Dioptric imaging sensor. *See* Pipelines

## Dynamic systems of multi-dimensions, 83, 84, 95

**E**Elasticity, loss tangent. *See* Visco-elasticity

## Error analysis

- multi-step methods, 120
- and stability, 132

## Exact methods

- continuity and compatibility, 63–64
- linear equation, 67–68
- one-to-one relationship, 64
- secondary suspension, 63
- steady state periodic response, 63
- transcendental equations, 65–68

## Exhaust system acoustic modeling

- control, chemical pollutions, 235
- ECAP, 236
- intake system components, 235
- ladder topology, 244–246

## linearization, network elements

- capacity, 242
- gas density, 243, 244
- pressure difference, 243
- relationship, 244
- resistivity, 242
- subsystem, 242
- temperature dependency, 242–243
- multi to mono-dimensional system, 240–241
- physical networks (*see* Physical networks)
- structure, 236
- transfer function (*see* Transfer function, exhaust system)

**F**

## Feature based methods, 260

## Finite element analysis

- Abaqus/Explicit v6.11, 223, 225–226
- ANCAP test and FE simulation, 185, 186
- central finite difference, 226–227
- close-up deformation, 228, 230
- contact surfaces, 231
- correlation, hood deformation, 185, 186
- ductility and MSFLD damage criteria, Aluminium 6060 T5, 228, 229
- energy plot, 181, 184
- force–displacement plots, 228, 229
- head impactor, 181, 182
- HYPERVIEW programs
  - acceleration vs. time, 181, 184
  - hood displacement vs. time, 181, 183
- implementation, 228
- LS-DYNA-3D, 181, 182
- MSFLD damage criterion, 228
- NCAP pedestrian subsystem testing
  - protocol, 185
  - optimisation, 181, 182
- PRIMER, 181

## Forecast

- sea levels, 304
- tide gauge, 310

## Fourier series and wavelet functions, 268

## Free form deformation (FFD)

- B-Spline transformation, 266–268
- control node displacements, 268
- medical image registration, 265

## Frequency response

- bilinear mass-spring vibrator, 60
- nonlinear vibrating systems, 71
- primary suspension, 68
- stiffness values, 59
- transcendental equations, 59

- Frequency synchronization  
 application, control strategy, 105, 106, 108, 109, 111, 112  
 control input, 107, 110, 113  
 reference signal, 105, 107, 108, 110
- FSMC. *See* Fuzzy sliding mode control (FSMC)
- Fuzzy sliding mode control (FSMC), 83, 84
- G**
- Galerkin method, 90–92
- H**
- Head impact model  
 ANCAP pedestrian protection testing protocol, 184  
 child head impactor, 183  
 finite element model, 184–185  
 GTR-9 and NCAP tests, 183  
 hood assembly, 183  
 requirements, 183–184
- Hood design  
 assembly, 178, 179  
 central area, 180  
 inlaid, 178, 179  
 inner panel, 180  
 length, 178  
 material and thickness, outer panel, 179–180  
 passenger car front-end collision, 178  
 structure, 180  
 wraparound, 178, 180
- Hood intrusion computation  
 accelerometer rotation, 186, 188  
 coordination, 186, 187  
 direction, 186  
 displacement, impactor, 186, 188  
 ‘Intrusion T’ calculation method, 187, 189  
 measurements, acceleration, 186
- Hood leading edge (HEL), 189, 190
- Hood panels  
 aluminium and steel, 213–214  
 ANOVA chart, HIC mean value, 197, 198  
 final design  
 aluminium with 85 mm deformation space, 209–210  
 head impact performance, steel hood configuration, 209  
 optimal hood configuration, 209  
 peak acceleration, 209  
 waveforms, 208–209  
 HIC values, 196–197
- HOODMATL, 197  
 and INRGEOM, 197  
 Kriging approach, 197  
 Monte Carlo method, 197  
 nonlinear LS DYNA models, 196  
 predictability, response surface, 197  
 primary impact, head injuries, 198  
 selection, structure and material  
 actual and optimal deformation, 198–199  
 evaluation, efficiency, 199  
 HIC values, 201, 204  
 hood panel thickness relationship, 202–204  
 multi-cone hood inner structure, 204  
 numerical tests, 200–203  
 optimal deformation space, 204  
 primary impact, 198, 200  
 structural stiffness, 201  
 theoretical optimal waveform, 203  
 WSTC, 203  
 standard deviation, 198  
 strength-two orthogonal array design, 196  
 thickness, aluminium material  
 deformation space, 207, 208  
 HIC value, 205–207  
 80 mm deformation space, 205, 206  
 multi-cone structure, 204–205  
 outer and inner, 204, 208
- Hood rear reference line (HRRL), 189, 191
- I**
- Initial value problem (IVP), 117, 118, 127  
 Intensity based methods, 260  
 Interpolation theory  
 FFD, 265–268  
 Fourier series and wavelet functions, 268  
 locally affine, 268  
 RBFs, 264–265
- Inverse Laplace transform  
 logarithmic law model, 149, 151  
 SLS, 142  
 transfer function, 246
- L**
- Ladder topology, acoustic modeling  
 energy sources, 245, 246  
 header and equivalent physical network, 245, 246  
 parameters calculations, 244, 245  
 pipe segments, 245  
 Simulink model, 246

- Laplace transform  
 power law model, 147–149  
 SLS, 141  
 stress relaxation, 142
- Logarithmic law model  
 addition rules, 150  
 Hookean spring, 151  
 of non-linear visco-elasticity, 149  
 storage modulus, 153  
 viscosity constant, 153, 154
- Loss tangent  
 visco-elastic models (*see* Visco-elasticity)  
 and viscosity, 155
- M**
- Medical image registration  
 applications, 254–256  
 computer vision, 251  
 CT, MRI and PET, 252, 253  
 image interpolation, 253  
 intensity information based registration  
 algorithm, 254, 255  
 interpolator, 254  
 mono-modal registration, 253  
 multi resolution pyramids, 253  
 objective function, 253–254  
 optimization method, 254, 255  
 pre-processing and feature extraction, 253  
 similarity measures  
 feature based methods, 260  
 voxel information based registration,  
 256–259  
 spatial mapping, 251  
 transformation model (*see* Transformation  
 model)  
 translation and rotation, image, 252
- Multiple coordinate system kinematics  
 Euler derivative transformation formula  
 classical mechanics, 36  
 global-fixed and rotating body-fixed  
 frame, 34–35  
 kinematics, 36  
 local/simple, 34  
 time-derivative vector, 34
- kinematic chain rule  
 angular acceleration, 40, 41  
 magnitudinal and convective rate of  
 change, 41  
 multiple angular velocity vectors, 38–40  
 three angular velocities, 41
- Razi acceleration, 32  
 rigid body motion, 33  
 second derivative, 36–37  
 three frames, 37–38
- Multi-step method  
 Adams–Moulton method, 120, 132  
 Runge–Kutta, 118  
 stability and error analysis, 120
- N**
- Nonlinear behaviors, 81
- Nonlinear differential equation  
 Adams–Moulton method, 120, 128  
 Bézier curves (*see* Bézier curves)  
 B-spline polynomials, 120  
 BVPs and IVP, 117  
 consistency, stability and convergence  
 characteristics polynomials, 124–125  
 Dahlquist's equivalence theorem,  
 127  
 $k$ -step method, 126  
 $m+1$ -step formula, 125–126  
 multi-step method, 125, 127  
 Euler vs. Runge–Kutta method, 118, 119  
 harmonic motion, 132  
 $k$ -step method, 119  
 Lotka–Volterra equation, 129, 131  
 mass–spring–damper system, 128–130  
 multi-stage methods, 118  
 ODE45, 128  
 ODEs (*see* Ordinary differential equations  
 (ODEs))  
 Runge–Kutta methods, 118  
 Van der Pol oscillator, 129, 133–134
- Nonlinear finite element and post-buckling  
 ABAQUS code, 230, 231  
 axisymmetric local buckling failure, 221,  
 222  
 collision safety, 220  
 and crushing behaviour, 219  
 cylindrical tube, 229–231  
 damage criteria, 230–231  
 dynamic response, cylindrical shell, 228,  
 230  
 failure mode, 221, 222  
 force vs. displacement graph, 222  
 identical buckling modes, 221  
 MSFLD, 231  
 numerical calibration  
 central finite difference, 226–227  
 contact constraint, 227–228  
 damage evolution and element removal,  
 225  
 ductile damage and failure, 224  
 FE models (*see* Finite element analysis)  
 initiation, ductile and shear failure,  
 223–224

- material, 223
- shear failure criterion, 225
- solution procedure, ABAQUS/explicit, 225–226
- sheet metals convergence, 230
- specimens and test machines
  - aluminium cylindrical tubes, 220, 221
  - load press, 220–221
  - structure, thin wall (*see* Thin wall structure)
  - transportation, 219
- Nonlinear vibration isolator
  - blades and magnetics, 82
  - cable system, 84
  - frequency response, 71
  - FSMC strategy, 84
  - multi-dimensional systems, 84
- Non-rigid registration
  - B-Spline motion model, 267
  - interpolation and approximation theories, 264–268
  - physical phenomena, 263–264
  - scaling and shear components, 262
  - transformation constrains, 269–270
- Numerical methods
  - multi-step methods, 124
  - Runge–Kutta, 118
- O**
- Ordinary differential equations (ODEs)
  - first order system, 128
  - ODE45, 128
  - physical networks, 237–238
- P**
- Pedestrian protection, head injury
  - AIS 2–6 injuries, 164
  - ANCAP, 164–165
  - ANOVA chart, peak acceleration, 210, 211
  - automobiles, 161–162
  - car manufacturing, 161
  - and chest impacts, 163
  - control factors, refinement scope, 195
  - cycling and walking, urban areas, 161
  - deformation space
    - direct and indirect contact surfaces, 172
    - hood panels, excessive stiffness, 173
    - inertial loading, 173, 174
    - secondary impact, 173, 174
    - structural stiffness, 172–173
    - vehicle design requirements, 174–175
  - design, hood, 178–180
  - differences, rebound velocities, 211–212
  - EEVC WG10, 167
  - EEVC WG17, 169–170
  - elastic strain energy, aluminium material, 213
  - fatalities and injuries, 162–164
  - filtered data, 214
  - finite element model, 181–183
  - Flex-PLI and TRL-PLI, 167, 168
  - GIDAS, 163
  - HIC values, 164, 214–215
  - hood intrusion, 186–189
  - HOODMATL, 210
  - inner hood geometries, 194
  - integration, 210–211
  - IRTAD, 163
  - leg and head impact test machines, 169
  - lower leg impactor, TRL-PLI, 169
  - minimum deformation space, 215
  - multi-cone inner hood structure, 215
  - NCAP and regulatory tests, 170
  - optimal waveform, 175–178
  - optimisation, hood panels (*see* Hood panels)
  - panel thickness, 210
  - parametric geometric model, 181
  - passenger car collides, 162
  - peak acceleration
    - duration, 171, 172
    - HIC calculation, 170–171
    - magnitude, 171, 172
    - and rebound velocity, 210
  - performance, 193–194
  - plastic sphere coverage, 169
  - PMHS, 170
  - POLAR II and III, 170
  - PPCFC, 162, 163
  - Pugh matrix, 194
  - recommendations, 215–216
  - refinement, 196
  - requirements, vehicle design (*see* Vehicle design, pedestrian protection)
  - research, design parameters, 214
  - resultant acceleration vs. vertical intrusion, 210, 211
  - standard deviation, HIC, 214
  - stress distribution, beam, 212–213
  - subsystem test method, 167, 168
  - technical, 169
  - tri-axial accelerometer, 170
  - unsafe conditions, 162
  - variables, 195
  - vehicle front-end, 163

- Physical networks
  - approach, engineering system, 236
  - ECAP, 236
  - elements, 238–239
  - mono and multi-dimensional systems, 239
  - ODE, 237–238
  - structure, 236, 237
  - variables, 237
- Piecewise linear systems
  - asymmetric systems, 75–77
  - DOF vibrating system, 60
  - double/multiple hitting, 72
  - dual rate bilinear vibrating system, 59, 60
  - exact solution (*see* Exact methods)
  - gap-damper engagement, 59, 60
  - implicit plot, 73
  - mechanical model, 61
  - motion, system, 62
  - nondimensional form, 62
  - nonlinearity range, 73–75
  - secondary suspension, 61
  - stability analysis, 61
  - steady response, 73, 75
  - time response
    - excitation frequency, 69
    - frequency, 68, 69
    - implicit plot, 69
    - primary suspension, 68
    - steady response, 70
  - transcendental equations, 59
  - vibration isolator, 70
- Pipelines
  - computation time and optimization error
    - vs.* number of iterations, 290
  - imaging platform and configuration, 280, 281
  - imaging unit components, 279, 280
  - inspection tools, 279
  - laser ring projection, 283, 284
  - learning-based method, 291
  - metric measurements, image-based, 277
  - model reduction process, 281
  - Nelder–Mead method, 286, 291
  - non-destructive testing methods, 277
  - numerical optimization technique, 287
  - optimization techniques, nonlinear, 281
  - pose estimation algorithm *vs.* initial seed value, 289
  - projection geometry, imaging, 282
  - reference observation signal, 284, 285
  - robotic inspection systems, 277, 279, 280
  - time-elapsed results, optimization process, 288
  - trigonometric function, 285–286
  - underwater infrastructure inspection, 279
  - variables separation, model reduction, 281
  - vision-based pose estimation, 278
  - water, oil and gas transport, 277
- Post-buckling. *See* Nonlinear finite element and post-buckling
- Power law model
  - fractional derivative approach, 149
  - Riemann–Liouville fractional integration, 148
- Power spectrum
  - measurement program, 240
  - system sound exhaustion, 241
- R**
- Radial basis functions (RBFs), 264–265, 270
- Razi acceleration
  - classical expression, 33
  - coordinate frames, 41–42
  - inertial, 41, 43–45
  - mechanical interpretation
    - angular velocity vectors, 50–52
    - compound rotation motion, 49, 50
    - derivative transformation method, 48
    - inertial and non-inertial frame, 48, 49
    - Newton’s second law of motion, 48
    - three coordinate frames system, 48–49
    - total acceleration vector, 49–51
  - mixed derivative transformation formula, 31–33
  - transformation, multiple coordinate frames, 31, 45–47
- Rigid body kinematics
  - angular velocity, 48
  - application, 52
  - compound rotation motion, 33, 44, 45
  - coordinate transformation, 36
  - inertial acceleration, 46, 47
  - linear acceleration vector, 46
  - mechanics, Razi acceleration, 33
  - non-inertial rotating coordinate frames, 33
- S**
- Sliding mode control (SMC), 83
- SLS. *See* Standard linear solid (SLS)
- Sound
  - attenuation components, 242
  - measures, 236
  - and power spectrum, 240, 241
  - pressure, 236
  - real vehicle, 236

- source, 245
  - speed of, 240, 244
  - SSDS. *See* Steady-state dynamic steering (SSDS)
  - Standard linear solid (SLS)
    - Kelvin–Voight model, 141
    - loss tangent, 142, 143
    - Maxell model, 147
    - maximal stress vs. modulus, parallel spring, 144, 145
    - phase angle vs. frequency, 145, 146
  - Statistical analysis
    - accuracy and reliability, 300
    - dynamic features, 278
    - tide gauge, 297
  - Steady state
    - fractional derivative, 148
    - SLS, 142
    - stress–time curve, 139, 140
    - visco-elasticity, 155–156
  - Steady-state dynamic steering (SSDS)
    - angles, front-wheel-steering vehicle, 20–21
    - conditions, 24
    - coordination, road curvature center, 21–22
    - cycloidal motion lane change
      - maneuver-moving, 19
      - radius of curvature, 19, 20
    - description, 16
    - final position, vehicles, 28–29
    - forward velocity, 19
    - kinematics angles
      - actual road vs. steady-state road, 27, 28
      - average, 23, 24
      - body frame, 23, 25
      - differences, 27, 28
      - global frame, 27
      - motion expression, body frame, 23, 25–26
      - outer and inner wheel, 22–23
      - rate of change, vehicle heading angle, 23, 25
      - side slip velocity, 23, 24
      - vs. steady-state, 27
      - steering road, 24, 26
      - velocity, global frame, 23, 26
    - measurements, 29
    - output–input relationships, 17–18
    - road curvature center, 19–20
  - Steering dynamics. *See* Steady-state dynamic steering (SSDS)
  - Stress strain
    - loss tangent, 139
    - Riemann–Liouville fractional integration, 148
- T**
- Thin wall structure
    - aluminium cylindrical tubes, 220, 221
    - collision safety, 220
    - crushing behaviour, 220, 231
    - failure, 228
    - global folding, 228
    - impact, energy absorbers, 220
  - Tide gauge sea level time series
    - classic statistical tools, 295
    - global warming, 295
    - linear model
      - multi-decadal oscillation, 302
      - in Sydney, 300, 301
    - measurement technology
      - glacial isostatic adjustment, 297, 298
      - PSMSL station demography, 298, 299
    - nonlinear model
      - breakpoint alignment, 310
      - chisquare value, 304, 306
      - coefficients, Sydney and Fremantle, 307, 309
      - fitting curve, 306–307
      - in Sydney and Fremantle, 307, 308
    - statistical packages, 296
    - technique, 296
    - thermometer location/land motion, 296
    - time histories, linear fitting parameters, 302–304
  - Time histories, linear fitting parameters
    - acceleration parameters, 303
    - LINEST function, Microsoft Excel, 302
    - monitoring technique, 302
    - non-accelerating tide gauge records, 303
    - sea levels in Sydney, 304–306
  - Transfer function, exhaust system
    - acoustic filter, 246
    - Bode plots, 247
    - ladder network, 246
  - Transformation model
    - continuous optimization methods, 270–271
    - diffeomorphic, 269–270
    - discrete optimization, 272
    - forward vs. inverse, 261
    - hierarchical methods, 272
    - interpolation theory
      - FFD, 265–268
      - Fourier series and wavelet functions, 268
      - locally affine, 268
      - RBFs, 264–265
    - inverse consistency, 269

- Transformation model (*cont.*)
    - physical characteristics
      - elastic model, 263
      - fluid flow, 264
    - pre-registration technique, 262
    - topology preservation, 269
  - Transient part
    - power law model, 148
    - SLS, 142
    - visco-elasticity, 155–156
- V**
- Vehicle design, pedestrian protection
    - components, GTR-9, 165
    - consumer metric, 166–167
    - design space
      - definition, 188
      - hood assembly, 188–189
      - modifications, interfacing components, 189
      - project scope, 189, 190
      - regulatory requirement, 188
    - head impact area
      - adult and child, 165–166
      - child, 190–191
      - GTR-9 protocol, 189, 190
      - HEL, 189, 190
      - hood side reference line, 189, 191
      - HRRL, 189, 191
      - impactor speed, 191
      - optimisation, 190, 192
      - WAD divides, 189
    - lower leg, 166
    - passenger car, 165
    - upper leg, 166
  - Vehicle dynamics
    - equations of motion
      - coefficients, 4, 5
      - forward velocity, 5
      - planar bicycle car model, 3–4
      - steady-state conditions, 4
      - steer angle, 4–5
      - rotation center, 8–9
      - steady-state responses, 6–7
      - steering angle
        - changes, 9–16
        - steady-state, 16–29
- Vibration control
- amplitude synchronization (*see* Amplitude synchronization)
  - axially translating cable, 94
  - chaotic motion, 94–96
  - frequency synchronization (*see* Frequency synchronization)
  - Galerkin method, 94
  - numerical simulations, 93
- Visco-elasticity
- characteristics, 137
  - decay function, stress relaxation, 138
  - hysteresis loop, stress–strain ellipse, 139, 140
  - linear/non-linear models, 137–138
  - negative storage modulus, 156
  - sinusoidal strain, 139
  - stress–time curve, 139, 140
  - transient and steady state parts, 155–156
- Viscosity, loss tangent. *see* Visco-elasticity
- Voxel information based registration
- Gaussian noise, 257
  - image intensities, 257
  - joint histograms, brain MRI image and rotated version, 259
  - mono-model registration, 256
  - optical flow setting, 258
  - residual complexity, 258
  - sum of absolute difference (SAD), 257
  - sum of square difference (SSD), 256
- Z**
- Zener model. *See* Standard linear solid (SLS)