Liming Dai · Reza N. Jazar *Editors*

Nonlinear Approaches in Engineering Application

Design Engineering Problems

Springer

Nonlinear Approaches in Engineering Application

Liming Dai • Reza N. Jazar
Editors

Nonlinear Approaches in Engineering Application

Design Engineering Problems

Springer

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

Reza N. Jazar
Mechanical and Automotive Engineering
RMIT University
Bundoora, VIC, Australia

ISBN 978-3-030-82718-2 ISBN 978-3-030-82719-9 (eBook)
https://doi.org/10.1007/978-3-030-82719-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The beauty of nonlinear approaches to engineering problems is that you must be talented to get even a wrong result.

Dedicated to Xinming and Mojgan

Preface

This book is the seventh volume in the series of Nonlinear Approaches in Engineering Applications, organized by the editors. This series collects individual applications on engineering problems in which the nonlinearity is quite important. Those system have been introduced, modeled mathematically, and the nonlinearity in their equations have been used to make the system optimize, stable, analyzed, etc.

This book is a collection of 15 important problems set in two parts: Modeling of Engineering Design Problems, and Applied Design of Engineering Problems. Both parts are more or less focused on applications of engineering problems. Chapter 1 is on numerical approaches for solving linear and nonlinear dynamic systems, showing better method than traditional numerical methods of solving nonlinear differential equations. Other chapters of Part 1 are on predictor/corrector formulations for solving non-linear initial value problems, control of nonhyperbolic dynamical systems, space charge phenomena, attitude control of acrobatic autonomous spacecraft, and form-finding design of space cable structures. The second part begins with application of genetic algorithm weld process design, then other applications will be presented in order of motion sickness in land vehicles, vehicle vibrations considering tire-road separation, predictive control optimizers for adaptive cruise control, lateral asymmetry on car's lateral dynamics, roll model control of autonomous vehicle, active anti-roll bar system of heavy vehicles, thermal comfort and game theory, and, finally wind resource assessment.

The nonlinear analysis, techniques, and applications have been developed in the past two to three centuries when the linear mathematical modeling of natural dynamical phenomena appeared not to be exact enough for some practical applications. The positive aspects of linear approximation of dynamic phenomena are simplicity and solvability. Linear approximation of a system provides us with the simplest model acting as the base and standard for which other nonlinear models should approach when the nonlinearities become negligible. Solvability is another positive characteristic of all linear systems. These two characteristics provide us with great ability and desire to model dynamic systems linearly. However, there exists many systems whose linear model and solution cannot provide exact enough

approximation of the real system behavior. For such systems, considering nonlinearities of the phenomena are unavoidable. Although the nonlinear approximation of a system provides us with a better and more accurate model, it also provides us with several complications. One of them is that it makes us search for indirect methods to gain some information on the possible solutions. Due to the nonlinearity and complexity of nonlinear systems, usually, it is very difficult or impossible to derive any analytical and closed-loop solutions for the systems. In solving or simulating the nonlinear systems, we have to rely on approximate or numerical methods, which may only provide approximate results while errors are unavoidable during the processes of generating the approximate results.

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 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 modern applications of nonlinearities.

The primary audience of this book are researchers, graduate students, and engineers in mechanical engineering, engineering mechanics, electrical engineering, civil engineering, aerospace engineering, mathematics, and science disciplines. In particular, the book can be used for training 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 systems, structure dynamics, and engineering applications of nonlinear science. It can also be utilized as a guide to the readers' fulfillment 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 book consists of two parts: Modeling of Engineering Design Problems, and Applied Design of Engineering Problems, including 15 chapters. Each chapter 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 details. The book and each of the chapters are 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 grasp the main concepts of each chapter with as less as possible the need to refer to the concepts of the other chapters and references. 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 college 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.

Acknowledgments

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 a book, their zeal 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 through the development of this book. The assistance provided by the staff members has been valuable and efficient. We thank Spinger-Verlag for their production of an elegant book.

Regina, SK, Canada — Liming Dai
Bundoora, VIC, Australia — Reza N. Jazar

Contents

List of Figures

Part I
Modeling of Engineering Design Problems

Chapter 1
Improved Theoretical and Numerical Approaches for Solving Linear and Nonlinear Dynamic Systems

Fang Pan, Dai Liming, Wang Kexin, and Wang Luyao

Abstract Complex dynamic systems are described by differential dynamic equations, mostly nonlinear without closed form analytic solution. To solve them numerically, there are many methods such as Euler's method, Taylor-series method and Runge-Kutta method, etc., each with advantages and disadvantages. In this chapter, a novel analytical and numerical methodology for the solutions of linear and nonlinear dynamical systems is introduced. The piecewise constant argument method combined with the Laplace transform, makes the new method called Piecewise constant argument-Laplace transform (*PL*). This method provides better reliability and efficiency for solving coupled dynamic systems. In addition, the numerical solutions of linear and nonlinear dynamic systems can be obtained smoothly and continuously on the entire time range from zero to *t*. The numerical results of the analytical solution of the method are given and compared with the results of the 4th-order Runge-Kutta (*RK*4) method, and the accuracy and reliability of the *PL* method are verified.

Keywords Complex dynamic systems · Numerical solutions · Argument-Laplace method · Vehicle laziness · Nonlinear differential equations

F. Pan (✉)
School of Mechanical Engineering, Southwest Petroleum University, Chengdu, China

Petroleum Engineering School, Southwest Petroleum University, Chengdu, China

Engineering and the Applied Sciences, University of Regina, Regina, Canada

W. Kexin
School of Mechanical Engineering, Southwest Petroleum University, Chengdu, China

D. Liming · W. Luyao
Engineering and the Applied Sciences, University of Regina, Regina, Canada

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_1

1.1 Introduction

Almost all the complex dynamic systems can be described by differential dynamic equations [1–3]. For exploring the dynamic characteristics of systems, obtaining accurate solutions of the dynamic equations is crucial in the engineering design. Of course, Euler's method, Taylor series method, and Runge-Kutta method are relatively mature methods existing in numerical analysis for vibration systems [4, 5]. Euler's method is the most traditional method to solve differential equations. However, the numerical accuracy of Euler's method should be decreased as integral curve is replaced with approximated polyline in the calculation process [6]. To perfect numerical accuracy in computing principle, trapezoidal method employs trapezoid to match integral curve in calculation process, and so the solution precision of this method is much higher than Euler's method [7–9]; but the computational efficiency is restricted on account of its own algorithm complexity and large calculating quantity. In order to improve the computation efficiency, Carl Runge and Martin Kutta proposed Runge-Kutta method, and the theoretical basis of this method is derived from Taylor's expansion and slope approximation [10]; in the solving process, the weighted average of the slope of multiple points in an integrating interval is calculated; the accuracy of the Runge-Kutta method is depended on the number of points in the integrating interval: the more points, the more accurate the calculation, and the more time should be needed to be taken in the calculation [11]. Since almost all terms are linearized and discontinuous when solving equations with Runge-Kutta method, calculative deviation is inevitable, which has attracted much more attention of numerous researchers. Shampine and Watts proposed a local error estimator, which seems to be of broad applicability [12]; according to this estimator, an efficient computational code (Ode45) is compiled [13, 14], and comparing the results of traditional Runge-Kutta, solution accuracy of Runge-Kutta is defective as the obtained results are not continuous, and the inherent characteristics of the system are lost in the second step of the calculation by using the Runge-Kutta method [15].

In this chapter, a novel analytical and numerical methodology for the solutions of linear and nonlinear dynamical systems is further developed. According to the piecewise constant argument mentioned in "Nonlinear Dynamics of Piecewise Constant Systems and Implementation of Piecewise Constant Arguments," combined with the Laplace transform, a method called piecewise constant argument-Laplace transform (PL) was established. On account of the characteristics of the Laplace transform, this method with better reliability and efficiency is for solving the coupled dynamic systems [16–18]. In addition, because of the continuity of solutions with the piecewise constant argument $[Nt]/N$ and the corresponding recurrence relations, the numerical solutions of linear and nonlinear dynamic systems can be obtained smoothly. Unlike the discrete solutions produced by existing numerical methods, the solutions given by the PL method to be presented are actually continuous on the entire time range from zero to t. Formulae for numerical computation in solving various dynamic systems are to be provided and discussed in the present chapter. Meanwhile, the numerical results of the analytical solution

of the method are given and compared with the results of the fourth-order Runge-Kutta (RK4) method [19, 20], and the accuracy and reliability of the PL method are verified.

With the development of science and technology, the study to the dynamics of low-dimensional nonlinear dynamic systems can hardly satisfy the requirements of actual engineering [21, 22]. To explore the exact dynamics in nonlinear systems, it is necessary to study the identification method of dynamic behavior for high-dimensional nonlinear systems. Nowadays, the principal method for identifying the dynamic behavior of the nonlinear system include Lyapunov exponent, Poincaré mapping method, and bifurcation theory. Lyapunov exponent is an important method for identifying chaotic signals of nonlinear dynamical systems [21, 23]. The important characteristic of the nonlinear dynamical systems is that the final value of the system is sensitively depended on the initial value; therefore, Lyapunov exponent represents the average exponential rates of divergence or convergence of closed orbits of the vibrating object in phase space of a dynamic system. The Lyapunov exponent is an efficient tool for estimating whether a dynamic system is periodic or chaotic. However, the Lyapunov exponent is not suitable for determining the dynamic behavior that is neither periodic nor chaotic. Poincaré mapping method can diagnose the dynamic characteristics of the nonlinear system based on the fixed points in the Poincaré section [24, 25]; however, it is unsuitable to determine the global dynamics and periodicity. Although global bifurcation and local bifurcation theory can discern the dynamic characteristics of higher-dimensional dynamical systems [26, 27], the prerequisite for this approach is BP normalization for the dynamic systems. Theoretically speaking, we can calculate the canonical form of any order from a given dynamical system. Actually, it is very difficult to calculate the high order normal form for the higher-dimensional dynamical system, because the process of the normalization computation is complex.

For the limitations of identifying characteristics of the dynamic system based on the methods above, Dai and Singh proposed a periodicity-ratio method to distinguish the dynamic characteristics of one-dimensional nonlinear dynamic system [25]. However, the identification method of the dynamic characteristics for high-dimensional nonlinear system is reported in the recent years. Mahmoud introduced a new theorem used to construct approximate analytical solutions for n-dimensional strongly nonlinear dynamical system, and then passive control method is also used to control n-dimensional chaotic complex nonlinear systems [28]. Here, considering the principle of Poincaré mapping, the periodicity-ratio (PR) method for diagnosing the dynamic characteristics of the high-dimensional nonlinear systems is proposed [29]. This method is employed to determine the dynamics and periodicity of the hull and Rössler system, respectively. The research findings will develop the evaluation method of dynamic characteristics for the high-dimensional nonlinear dynamic system.

1.2 Fundamental Theory

To develop the numerical solutions of the dynamic system in this chapter, the numerical approach named the piecewise constant argument method is used. By introducing a piecewise constant argument $[Nt]/N$, the original continuous governing equations are divided into many segments, denoted as N, and linearized in each time interval, $i/N \leq t < (i+1)/N$, in which a linear dynamic system is developed. In the expression, N is a parameter that controls the length of the time interval and the accuracy of the calculation. When N becomes large enough, the numerical solutions can be sufficiently accurate. To preferably utilize the PL and the PR method, some theorems must be understood.

1.2.1 Piecewise Constant Argument

Theorem 1.1 Suppose argument $[Nt]$ is the integer-valued function of product of time t and parameter N, where N is a positive integer [29]. When N is approached to infinity, the value of the argument $[Nt]/N$ tends to t, i.e.:

$$\lim_{N\to\infty} \frac{[Nt]}{N} = t. \tag{1.1}$$

Theorem 1.2 Considering a vibrating system:

$$f\left(x'', x', x\right) = g(t), \tag{1.2}$$

with initial conditions:

$$x(0) = d_0, x'(0) = v_0. \tag{1.3}$$

Then, in time interval $i/N \leq t < (i+1)/N$, Eq. (1.2) and Eq. (1.3) are converted to:

$$f\left(x_i'', x_i', x_i\right) = g\left(\frac{i}{N}\right), \tag{1.4}$$

and:

$$x_i\left(\frac{i}{N}\right) = d_i, x_i'\left(\frac{i}{N}\right) = v_i, \tag{1.5}$$

where $i = 0, 1, 2, \ldots, [Nt]$.

1.2.2 Laplace Transformation and Residues Principle

The coupling system can be decoupled with the Laplace transformation. Some significant theorems should be stressed.

Theorem 1.3 Supposing functions $x_i''(t)$, $x_i'(t)$, and $x_i(t)$ can be rewritten as Laplace style, i.e. [2, 30]:

$$L\left[x_i(t)\right] = X_i(s), \tag{1.6}$$

$$L\left[x_i'(t)\right] = sX_i(s) - x_i\left(\frac{i}{N}\right), \tag{1.7}$$

$$L\left[x_i''(t)\right] = s^2X_i(s) - sx_i\left(\frac{i}{N}\right) - x_i'\left(\frac{i}{N}\right), \tag{1.8}$$

where $x_i''(t) = \frac{d^2x_i}{dt^2}$, $x_i'(t) = \frac{dx_i}{dt}$.

Theorem 1.4 $F(s)$ is assumed to the Laplace style in complex field. $F(s)$ corresponding to time-domain solution can be determined with inverse Laplace transformation, i.e.:

$$f(t) = L^{-1}\left[F(s)\right] = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} F(s)e^{st}ds. \tag{1.9}$$

Theorem 1.5 Through the definition of residues, the integral in Eq. (1.9) can be expressed by the summing of all the residuals in the definition domain, i.e.:

$$\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} F(s)e^{st}ds = \sum_{k=1}^{n}\mathrm{Res}\left[F(s)e^{st}, s_k\right], \tag{1.10}$$

where s_k represents value of the kth pole point of $F(s)$. Here, considering $F(s)$ to be a rational function, the fraction style of $F(s)$ can be given by $F(s) = \frac{A(s)}{B(s)}$, where $A(s)$ and $B(s)$ are mutually irreducible, and the order of $A(s)$ is lower than $B(s)$ [30]. According to the exponent number of zero point in $B(s)$, Res$[F(s)e^{st}, s_k]$ can be calculated with following rules, i.e.:

A. If s_k is the first-order zero point of $B(s)$, then $\mathrm{Res}\left[\frac{A(s)}{B(s)}e^{st}, s_k\right] = \left.\frac{A(s)}{B'(s)}e^{st}\right|_{s=s_k}$.

B. If s_k is the n-order zero point of $B(s)$, then $\mathrm{Res}\left[\frac{A(s)}{B(s)}e^{st}, s_k\right]$ $= \frac{1}{(n-1)!}\lim_{s\to s_k}\frac{d^{n-1}}{ds^{n-1}}(s-s_k)^n\frac{A(s)}{B(s)}e^{st}$.

1.2.3 The Periodicity Ratio of Nonautonomous Systems

Consider the following second order of the n-dimensional nonautonomous system [31]:

$$\ddot{\mathbf{x}} = \mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, t), \quad \mathbf{x} \in \mathbf{R^n}. \tag{1.11}$$

Suppose that the system is subjected to an external excitation with period time T; meanwhile, $\mathbf{x}$ is the periodic solution of the differential system above, i.e., $\mathbf{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$. Usually, $\mathbf{x}$ is the solution of the harmonic vibration related to multiple period of T satisfying the following relationship:

$$\mathbf{x}(t_0) = \mathbf{x}(t_0 + jT). \tag{1.12}$$

where t_0 is the reference time and j is the number of period points of the system in the Poincaré section. For a completely periodic nonautonomous dynamic system, no matter how long the vibration is sustained, only j finite points appear in Poincaré section $\mathbf{x} - \dot{\mathbf{x}}$ ($x_r - \dot{x}_r$, $(r = 1, 2, \ldots, n)$). If the phase points are infinite in Poincaré section **x**—$\dot{\mathbf{x}}$, the n-dimensional nonautonomous system is aperiodic.

According to Eq. (1.12), the overlapping points in n Poincaré sections can describe the periodicity of n-dimensional nonautonomous dynamic system. The number of overlapping points in the rth Poincaré section can be determined by:

$$X_{r,ki} = x_r(t_0 + kT) - x_r(t_0 + iT), \tag{1.13}$$

$$\dot{X}_{r,ki} = \dot{x}_r(t_0 + kT) - \dot{x}_r(t_0 + iT). \tag{1.14}$$

where k and i are integers: $k \in [1, j]$, $i \in [1, m]$. Equation (1.13) represents the displacement difference between phase point i and k in the Poincaré section when the phase trajectories pass through the rth section. Equation (1.14) describes the velocity difference between phase point i and k in the Poincaré section when the phase trajectories pass through the rth section. m is the sum of the phase points in the Poincaré section, including the overlapping and nonoverlapping points. Therefore, the total number of phase points in the Poincaré section can be denoted as:

$$S_{\mathrm{a}} = n \cdot m \tag{1.15}$$

According to the above definition, the following conclusions are stressed:

1. In the Poincaré section, the so-called overlapping phase points i $(x_i, \dot{x}_i)$ and j $\left(x_j, \dot{x}_j\right)$ represent $x_i = x_j$ and $\dot{x}_i = \dot{x}_j$.
2. According to the overlapping property of the phase points, the phase points of the nonlinear periodic system in the rth Poincaré section should satisfy the following condition:

$$\begin{cases} X_{r,ki} = 0 \\ \dot{X}_{r,ki} = 0 \end{cases} \tag{1.16}$$

3. If the phase point in arbitrary Poincaré sections cannot satisfy Eq. (1.16), the phase points are nonoverlapping.

Applying Eq. (1.16), the total number $\boldsymbol{\zeta}(k)$ of the kth overlapping phase pion in the rth Poincaré section can be expressed as follows:

$$\zeta_r(k) = \left\{ \sum_{i=k}^{n} Q\left(X_{r,ki}\right) Q\left(\dot{X}_{rki}\right) \right\} P\left(\sum_{i=k}^{n} \left[Q\left(X_{r,ki}\right) Q\left(\dot{X}_{r,ki}\right) \right] - 1 \right) \tag{1.17}$$

where $\zeta_r(k)$ is applied to calculate the number of all phase points overlapping to the kth phase point. Q and P are step functions as follows:

$$Q(y) = \begin{cases} 1 & \text{if } y = 0 \\ 0 & \text{if } y \neq 0 \end{cases} \tag{1.18}$$

$$p(z) = \begin{cases} 1 & \text{if } z = 0 \\ 0 & \text{if } z \neq 0 \end{cases} \tag{1.19}$$

Considering Eq. (1.17), after the total number k of the overlapping phase points is determined, the number of the jth visible point corresponding to overlapping point can be calculated. Assign N_r as overlapping points in the rth Poincaré section; thus, N_r can be expressed by:

$$N_r = N_r(1) + \sum_{k=2}^{n} \zeta_r(k) P\left(\prod_{l=1}^{k-1} \left(X_{r,kl} + \dot{X}_{r,kl} \right) \right) \tag{1.20}$$

in which $\prod$ is the symbol for multiplication and $P(\cdot)$ is the step function as defined in Eq. (1.19). This equation ensures that the duplicate included in the calculations for N_r or missing in any overlapping point is prevented. If the response of a dynamic system is completely periodic, all the points in the Poincaré map must be overlapping, and the corresponding N_r can be simply expressed in the following form:

$$N_r = \sum_{k=1}^{j} \boldsymbol{\zeta}_r(k) \tag{1.21}$$

For periodicity of the nonlinear dynamic system, N_r represents the overlapping phase points in the rth Poincaré section. Therefore, the total number of overlapping points and phase points in n-dimensional space can be independently denoted by:

$$S = \sum_{r=1}^{n} N_r \tag{1.22}$$

$$S_a = \sum_{r=1}^{n} S_r, \tag{1.23}$$

where S_r is the number of phase points in the rth Poincaré section. Therefore, the periodicity ratio of the nonlinear nonautonomous system with n-dimensional space can be denoted by:

$$\gamma = \lim_{S_a \to \infty} \frac{S}{S_a}. \tag{1.24}$$

It can be known that the number of the overlapping points is less than or equal to the all phase points, i.e., $0 \leq S \leq S_a$; in this case, $0 \leq \gamma \leq 1$. If the dynamic responses of the nonlinear system are periodic, thus, all the phase points in the Poincaré section must be overlapping, and periodicity ratio γ is equal to 1. If the dynamic responses of the nonlinear system are chaotic, thus, all the phase points in the Poincaré section must be nonoverlapping, and periodicity ratio γ is equal to zero. Through the definition for the periodicity ratio γ of the nonlinear dynamic system, it is easy to find that periodicity ratio γ can describe the periodicity of the nonlinear dynamic system.

1.2.4 The Periodicity Ratio of Autonomous Systems

Consider the following second order of the n-dimensional autonomous system [32]:

$$\dot{\mathbf{x}} = f(\mathbf{x}) \qquad \mathbf{x} \in \mathbf{R^n} \tag{1.25}$$

As the system is an autonomous system without external excitation, the phase points in the Poincaré section cannot be determined by Eq. (1.12). For the n-dimensional nonlinear autonomous system, no matter how long the vibration of the system is sustained, only j finite phase points appeared in Poincaré section $x_{r-1} - x_r (r = 2, 3, \ldots, n)$. Therefore, the periodicity of the nonlinear system can be described by n–1 Poincaré sections. If there are infinite phase points in the n–1 Poincaré sections, the nonlinear dynamic system is nonperiodic.

If the vibration behavior of the autonomous system is periodic, period T_1 of x_1 can be estimated with the maximum method, as numerical solution of x_1 is a series of points related to time, i.e.,$x_{1(\tau)} = x_1(t_\tau)$, $\tau = 1, 2, 3 \ldots, \infty$. Thus, the search procedure can be employed to determine the maximum value of x_1. If $x_{1(\tau)}$ satisfy the following condition:

$$\begin{cases} x_{1(\tau+1)} > \frac{1}{2}\left(x_{1(\tau)} + x_{1(\tau+2)}\right) \\ x_{1(\tau)} > x_{1(\tau+1)} \\ x_{1(\tau)} > x_{1(\tau-1)} \end{cases} \tag{1.26}$$

thus $x_{1(\tau)}$ is the maximum point of x_1, and the point of x_2 corresponding to $x_{1(\tau)}$ in this time is $x_{2(\tau)}$. Thus, the phase point of the first Poincaré section is expressed by $(x_{1(\tau)}, x_{2(\tau)})$. The amount of the overlapping phase points in the first Poincaré section is written by:

$$N_1(k) = \zeta(1) + \sum_{k=2}^{n} \zeta_1(k) P\left(\prod_{l=1}^{k-1} x_{1(\tau),kl}\right), \tag{1.27}$$

in which $x_{1(\tau),\,kl} = |x_{1(\tau),\,k} - x_{1(\tau),\,l}|$. Therefore, the period ratio can be determined through the phase point in the first Poincaré section:

$$\gamma_1 = \lim_{S_a \to \infty} \frac{N_1}{S_1}. \tag{1.28}$$

If $\gamma_1 = 1$, the period T of the system can be confirmed as the following:

1. Assume that the number of the visible phase points in the first Poincaré section is j. Choosing the kth visible point of j, the number of the overlapping points is q, and so the span of average time between two arbitrary adjacent points can be expressed as:

$$\eta = \frac{t_{k,q} - t_{k,1}}{q-1}, \tag{1.29}$$

where $t_{k,\,q}$ represents the time of the qth overlapping points in the kth visible points and $t_{k,\,1}$ represents the time of the qth overlapping points in the first visible points. If all the overlapping points are periodic, parameter η is equivalent to the vibration period of the dynamic system.

2. To improve the accuracy and reliability for identifying the periodicity of the dynamic system, identification parameter ρ is defined as:

$$\rho^2 = \frac{\sum_{1}^{q-1}\left(t_{k,i+1} - t_{k,i} - \eta\right)}{q-1}, \quad i = 1, 2, 3, \ldots, q \tag{1.30}$$

From the equation above, it follows that when the identifying parameter $\rho = 0$, the vibration behavior of the dynamic system is periodic. In this case, the period T_1 of the system can be expressed by:

$$T_1 = t_{k,i+1} - t_{k,i}. \tag{1.31}$$

Therefore, the number of overlapping points in the rth Poincaré section can be determined by:

$$\begin{cases} X_{r,ki} = x_r\left(t_0 + kT_1\right) - x_r\left(t_0 + iT_1\right), \\ X_{r+1,ki} = x_{r+1}\left(t_0 + kT_1\right) - x_{r+1}\left(t_0 + iT_1\right), \end{cases} \tag{1.32}$$

where k and i are integers, $k \in [1, j]$, $i \in [1, m]$. Equation (1.32) represents the displacement difference between phase points i and k. m is the number of the phase points in an arbitrary Poincaré sections. In this case, the total number of the phase points for the n-dimensional system can be denoted by:

$$S_{\mathrm{a}} = (n-1) \cdot m. \tag{1.33}$$

According to the definition above, the following conclusions can be stressed:

1. The so-called overlapping phase points represent $x_i = x_j$ and $x_{i+1} = x_{j+1}$.
2. According to the characteristics of the overlapping points, the phase points of the periodic system satisfy the following condition:

$$\begin{cases} X_{r,ki} = 0, \\ X_{r+1,ki} = 0. \end{cases} \tag{1.34}$$

3. If Eq. (1.34) is not satisfied, these phase points cannot be overlapping.

Employing Eq. (1.34), the amount $\boldsymbol{\zeta}(k)$ of the phase points overlapping with the kth phase point can be determined with:

$$\zeta_r(k) = \left\{ \sum_{i=k}^{n} Q\left(X_{r,ki}\right) Q\left(X_{r+1,ki}\right) \right\} P\left(\sum_{i=k}^{n} \left[Q\left(X_{r,ki}\right) Q\left(X_{r+1,ki}\right) \right] - 1 \right). \tag{1.35}$$

According to Eq. (1.35), the overlapping points in the rth Poincaré sections can be defined by:

$$N_r = N_r(1) + \sum_{k=2}^{n} \zeta_r(k) P\left(\prod_{l=1}^{k-1} \left(X_{r,kl} + X_{r+1,kl} \right) \right). \tag{1.36}$$

where $\prod$ is multiplication and $P(\cdot)$ is step function defined previously. If the responses of the nonlinear dynamic system are periodic, N_r can be simply expressed by:

$$N_r = \sum_{k=1}^{j} \boldsymbol{\zeta}_r(k) \tag{1.37}$$

Therefore, the amount of the overlapping point and all the points in the Poincaré section can be represented by:

$$S = \sum_{r=1}^{n-1} N_r \tag{1.38}$$

$$S_a = \sum_{r=1}^{n-1} S_r \tag{1.39}$$

where S_r represents the amount of the phase points in the rth Poincaré section. Therefore, the periodicity ratio of the n-dimensional nonlinear autonomous system is written by:

$$\gamma = \lim_{S_a \to \infty} \frac{S}{S_a} \tag{1.40}$$

1.3 Analytical and Numerical Solutions of Stiffness Coupling Systems

To explore the PL method for solving the linear stiffness coupling system, consider the following research subjects: (1) stiffness coupling system, (2) stiffness and damping coupling system, and (3) stiffness and damping coupling system encountered with external excitation.

1.3.1 Stiffness Coupling System

A simple stiffness coupling system (SCS) is firstly shown in Fig. 1.1, where $m_n (n = 1, 2)$ is the mass of the vibrator, $k_n (n = 1, 2)$ represents the stiffness coefficient of the spring, and $x_n (n = 1, 2)$ denotes the displacement response of the vibrator. By applying an initial velocity and initial displacement to the vibrator, the response of the vibrator changes with time.

According to the physical model, the following differential equation can be established:

$$\begin{cases} m_1 x_1'' + k_1 (x_1 - x_2) = 0, \\ m_2 x_2'' + k_1 (x_2 - x_1) + k_2 x_2 = 0. \end{cases} \tag{1.41}$$

The simplified form of Eq. (1.41) can be expressed as:

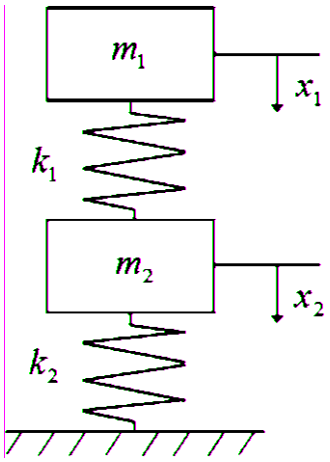


Fig. 1.1 The stiffness coupling system

$$\begin{cases} x_1'' + \omega_{1,1}^2 (x_1 - x_2) = 0, \\ x_2'' + \omega_{1,2}^2 (x_2 - x_1) + \omega_{2,2}^2 x_2 = 0, \end{cases} \tag{1.42}$$

where $\omega_{1,1} = \sqrt{k_1/m_1}$, $\omega_{1,2} = \sqrt{k_1/m_2}$, $\omega_{2,2} = \sqrt{k_2/m_2}$. The initial value of the system can be assumed by:

$$\begin{cases} x_1(0) = d_{1,0}, x_1'(0) = v_{1,0}, \\ x_2(0) = d_{2,0}, x_2'(0) = v_{2,0}. \end{cases} \tag{1.43}$$

Subsequently, the piecewise constant method should be employed to rearrange the continuous Eq. (1.42) as piecewise constant system. Replace terms $x_1(t)$ and $x_2(t)$ with piecewise constant functions $x_{1,i}\left(\frac{i}{N}\right)$ and $x_{2,i}\left(\frac{i}{N}\right)$ at arbitrary intervals of $i/N \leq t < (i+1)/N$ $(i = [Nt])$, and it can be expressed as:

$$\begin{cases} x_{1,i}'' + \omega_{1,1}^2 x_{1,i}\left(\frac{i}{N}\right) - \omega_{1,1}^2 x_{2,i}\left(\frac{i}{N}\right) = 0, \\ x_{2,i}'' + \omega_{1,2}^2 x_{2,i}\left(\frac{i}{N}\right) - \omega_{1,2}^2 x_{1,i}\left(\frac{i}{N}\right) + \omega_{2,2}^2 x_{2,i}\left(\frac{i}{N}\right) = 0. \end{cases} \tag{1.44}$$

In this case, the initial conditions in Eq. (1.43) can be transformed into the following form:

$$\begin{cases} x_{1,i}\left(\frac{i}{N}\right) = d_{1,i}, x_{1,i}'\left(\frac{i}{N}\right) = v_{1,i}, \\ x_{2,i}\left(\frac{i}{N}\right) = d_{2,i}, x_{2,i}'\left(\frac{i}{N}\right) = v_{2,i}. \end{cases} \tag{1.45}$$

It can be found from Eq. (1.44) that the piecewise style is coupled with respect to $x_{1,i}$ and $x_{2,i}$. Therefore, the Laplace transformation can be used to separate the coupling variables, and the plural form of Eq. (1.44) in the interval $i/N \leq t < (i+1)/N$ is calculated as:

$$\begin{cases} s^2 X_{1,i}(s) - s x_{1,i}\left(\frac{i}{N}\right) - x'_{1,i}\left(\frac{i}{N}\right) + \frac{\omega_{1,1}^2}{s} x_{1,i}\left(\frac{i}{N}\right) - \frac{\omega_{1,1}^2}{s} x_{2,i}\left(\frac{i}{N}\right) = 0, \\ s^2 X_{2,i}(s) - s x_{2,i}\left(\frac{i}{N}\right) - x'_{2,i}\left(\frac{i}{N}\right) + \frac{\omega_{1,2}^2}{s} x_{2,i}\left(\frac{i}{N}\right) - \frac{\omega_{1,2}^2}{s} x_{1,i}\left(\frac{i}{N}\right) + \frac{\omega_{2,2}^2}{s} x_{2,i}\left(\frac{i}{N}\right) = 0. \end{cases} \tag{1.46}$$

Then, $X_{1,\,i}(s)$ and $X_{2,\,i}(s)$ can be expressed as:

$$X_{1,i}(s) = \frac{1}{s} x_{1,i}\left(\frac{i}{N}\right) + \frac{1}{s^2} x'_{1,i}\left(\frac{i}{N}\right) - \frac{\omega_{1,1}^2}{s^3} x_{1,i}\left(\frac{i}{N}\right) + \frac{\omega_{1,1}^2}{s^3} x_{2,i}\left(\frac{i}{N}\right), \tag{1.47}$$

$$X_{2,i}(s) = \frac{1}{s} x_{2,i}\left(\frac{i}{N}\right) + \frac{1}{s^2} x'_{2,i}\left(\frac{i}{N}\right) - \frac{\omega_{1,2}^2}{s^3} x_{2,i}\left(\frac{i}{N}\right) + \frac{\omega_{1,2}^2}{s^3} x_{1,i}\left(\frac{i}{N}\right) - \frac{\omega_{2,2}^2}{s^3} x_{2,i}\left(\frac{i}{N}\right). \tag{1.48}$$

In this situation, $X_{1,\,i}(s)$ and $X_{2,\,i}(s)$ are the plural form of $x_{1,\,i}(t)$ and $x_{2,\,i}(t)$ in interval $i/N \le t < (i+1)/N$, respectively. Through using the reverse Laplace transformation, the expressions of $x_{1,\,i}(t)$ and $x_{2,\,i}(t)$ can be obtained by the replacement of i with $[Nt]$:

$$x_{1,i}(t) = \mathbf{A}_1 \mathbf{v}_i, \tag{1.49}$$

$$x_{2,i}(t) = \mathbf{A}_2 \mathbf{v}_i, \tag{1.50}$$

where:

$$\mathbf{A}_1 = \left[1 - \tfrac{1}{2}\omega_{1,1}^2\left(t - \tfrac{[Nt]}{N}\right)^2, t - \tfrac{[Nt]}{N}, \tfrac{1}{2}\omega_{1,1}^2\left(t - \tfrac{[Nt]}{N}\right)^2, 0\right],$$

$$\mathbf{A}_2 = \left[\tfrac{1}{2}\omega_{1,2}^2\left(t - \tfrac{[Nt]}{N}\right)^2, 0, 1 - \tfrac{1}{2}\omega_{1,2}^2\left(t - \tfrac{[Nt]}{N}\right)^2 - \tfrac{1}{2}\omega_{2,2}^2\left(t - \tfrac{[Nt]}{N}\right)^2, t - \tfrac{[Nt]}{N}\right],$$

$$\mathbf{v}_i = [d_{1,\,i}, v_{1,\,i}, d_{2,\,i}, v_{2,\,i}]^T.$$

Therefore, the velocities $x'_{1,i}(t)$ and $x'_{2,i}(t)$ should be rewritten as follows:

$$x'_{1,i}(t) = \frac{\partial \mathbf{A}_1}{\partial t} \mathbf{v}_i, \tag{1.51}$$

$$x'_{2,i}(t) = \frac{\partial \mathbf{A}_2}{\partial t} \mathbf{v}_i. \tag{1.52}$$

As mentioned previously, the displacements ($x_{1,\,i}(t)$, $x_{2,\,i}(t)$) and the velocities ($x'_{1,i}(t)$, $x'_{2,i}(t)$) are continuous in $t \in [0, +\infty)$, which is physically implied that there is no jump (break) or discontinuity in the displacement and velocity over $t \in [0, +\infty)$. Thus, the displacement and velocity in two adjacent points must satisfy the following continuity condition:

$$\begin{cases} x_{1,i}\left(\frac{i}{N}\right) = x_{1,i-1}\left(\frac{i}{N}\right), x'_{1,i}\left(\frac{i}{N}\right) = x'_{1,i-1}\left(\frac{i}{N}\right), \\ x_{2,i}\left(\frac{i}{N}\right) = x_{2,i-1}\left(\frac{i}{N}\right), x'_{2,i}\left(\frac{i}{N}\right) = x'_{2,i-1}\left(\frac{i}{N}\right). \end{cases} \tag{1.53}$$

According to the continuity condition in Eq. (1.53), the relation of displacement and velocity between the two adjacent truncation points can be represented as:

$$\mathbf{v}_i = \mathbf{p}\mathbf{v}_{i-1}, \tag{1.54}$$

where $\mathbf{v}_i = [d_{1,i}, v_{1,i}, d_{2,i}, v_{2,i}]^T$, $\mathbf{v}_{i-1} = [d_{1,i-1}, v_{1,i-1}, d_{2,i-1}, v_{2,i-1}]^T$, $\mathbf{p} = [p_{ij}]_{4\times 4}(i = 1, \ldots, 4, \ j = 1, \ldots, 4)$, and the elements of the matrix $\mathbf{p}$ are shown as follows:

$$p_{11} = 1 - \frac{\omega_{1,1}^2}{2N^2}, p_{12} = \frac{1}{N}, p_{13} = \frac{\omega_{1,1}^2}{2N^2}, p_{14} = 0, p_{21} = -\frac{\omega_{1,1}^2}{N}, p_{22} = 1, p_{23} = \frac{\omega_{1,1}^2}{N}, p_{24} = 0,$$
$$p_{31} = \frac{\omega_{1,2}^2}{2N^2}, p_{32} = 0, p_{33} = 1 - \frac{\omega_{1,2}^2}{2N^2} - \frac{\omega_{2,2}^2}{2N^2}, p_{34} = \frac{1}{N}, p_{41} = \frac{\omega_{1,2}^2}{N}, p_{42} = 0, p_{43} = -\frac{\omega_{1,2}^2}{N} - \frac{\omega_{2,2}^2}{N}, p_{44} = 1.$$

In light of Eq. (1.54), $\mathbf{v}_i$ can be expressed via initial condition $\mathbf{v}_0$ with iterative computations of i times, i.e.:

$$\mathbf{v}_i = \mathbf{p}^i\mathbf{v}_0, \tag{1.55}$$

where $\mathbf{v}_0 = [d_{1,0}, v_{1,0}, d_{2,0}, v_{2,0}]^T$.

When the initial condition $\mathbf{v}_0$ is determined, the displacement and velocity of the system at any time can be calculated by utilizing Eq. (1.55). The semi-analytical solution of the system in ith interval can be rewritten as the style of piecewise constant, i.e.:

$$x_{1,i}(t) = \mathbf{A}_1\mathbf{p}^{[Nt]}\mathbf{v}_0, \tag{1.56}$$

$$x_{2,i}(t) = \mathbf{A}_2\mathbf{p}^{[Nt]}\mathbf{v}_0. \tag{1.57}$$

The calculation program can be compiled by using the above expressions. The partial results of solving SCS for each method are obtained, as shown in Table 1.1.

With Table 1.1, take the results from RK4 with step 0.001 s as accurate values. By comparing the calculation results of each method with same step (0.01 s), the maximum relative errors of each method can be seen: Euler method, 15.28%; trapezoidal method, 9.54%; Ode45, 4.21%; RK4, 6.15%; and PL, 1.54%. Therefore, the most accurate method is the PL method, the second is ode45, the third is RK4, the fourth is the trapezoidal method, and the last is the Euler method. Figure 1.2 is the displacement responses of SCS. It also follows that the accuracy of numerical results calculated by the RK4 method is rougher than PL within time step 0.01 s; however, the PL solutions within time step 0.01 s are consistent with RK4 within time step 0.001 s. To evaluate the computational efficiency of the PL method, the CPU time for solving SCS with PL and RK4 methods in this same time domain

Table 1.1 Numerical results (x_1) in solving SCS

Times (s)	RK4 (step:0.01 s)	PL (step:0.01 s)	Euler method (step:0.01 s)	Trapezoidal method (step:0.01 s)	Ode45 (step:0.01 s)	RK4 (step:0.001 s)
3.0	−0.02015543916	−0.02070171306	−0.01723785663	−0.01904587966	−0.02045379097	−0.02069823574
3.5	−0.01980681766	−0.02119341174	−0.01864093145	−0.01975432560	−0.02059361094	−0.02118516786
4.0	−0.01121570163	−0.01279856223	−0.00936174295	−0.01026048735	−0.01179407431	−0.01278902287

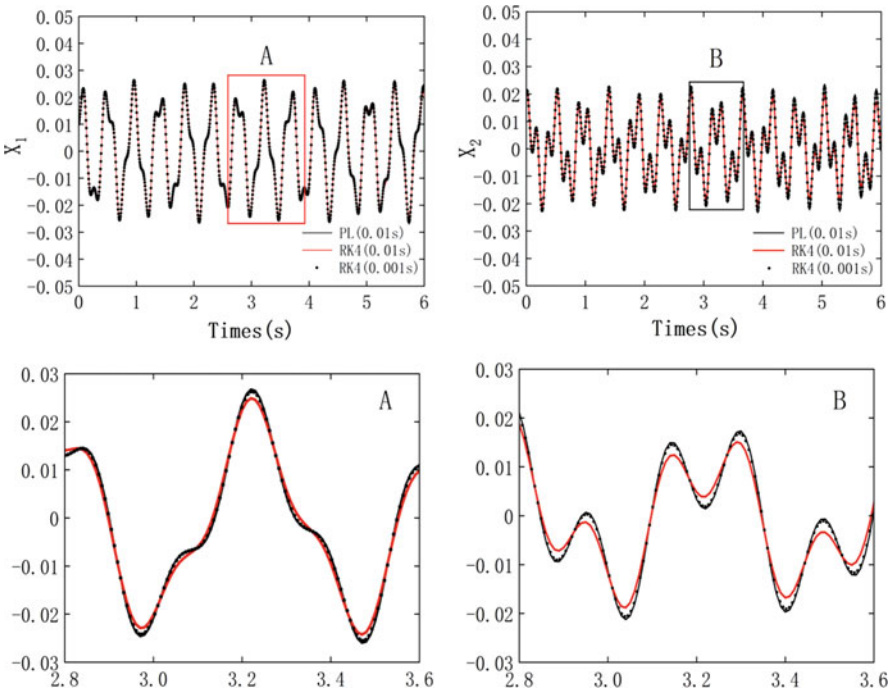


Fig. 1.2 Displacement response of SCS ($\omega_{1,1} = 22.361$, $\omega_{1,2} = 20$, $\omega_{2,2} = 17.321$, $x_1(0) = 0.01$, $x_1'(0) = 0.1$,$x_2(0) = 0.02$, $x_2'(0) = 0.2$.)

Table 1.2 SCS computation time in time history 30s

Numerical method	Time step(s)	Iterations	CPU time
RK4	0.001	30,000	4.100625
RK4	0.01	3000	0.431250
PL	0.001	30,000	2.093750
PL	0.01	3000	0.228125

is considered. As shown in Table 1.2, the CPU time spent by the PL method is approximately half of the RK4 method in time step 0.01 s, which reveals that the PL method is more efficient than RK4 in numerical calculation of SCS. Obviously, the same kind of computation method with short time step will take more time than long time step, but reliability result can be obtained with short time step.

1.3.2 Stiffness and Damping Coupling System

In actual engineering, damping usually existed in coupling dynamic system. Therefore, a dynamic system coupled with the terms of stiffness and damping (SDCS) is considered, as shown in Fig. 1.3, where $m_n (n = 1, 2)$ is the mass of the

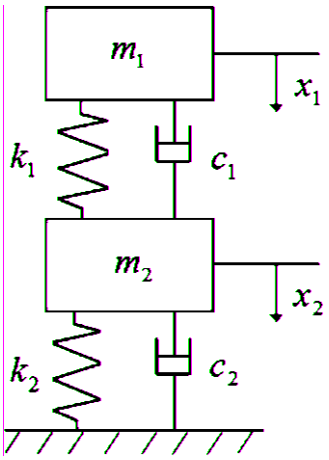


Fig. 1.3 The stiffness and damping coupling system

vibrator, $k_n(n = 1, 2)$ represents the stiffness coefficient of the spring, $c_n(n = 1, 2)$ is the damping coefficient of the damper, and $x_n(n = 1, 2)$ denotes the displacement response of the vibrator. Likewise, give the vibrator an initial velocity and initial displacement, and the analytical expression of the response of the vibrator can be obtained by the PL method.

Through this model, the following equation can be obtained:

$$\begin{cases} m_1 x_1'' + c_1 \left(x_1' - x_2'\right) + k_1 \left(x_1 - x_2\right) = 0, \\ m_2 x_2'' + c_1 \left(x_2' - x_1'\right) + k_1 \left(x_2 - x_1\right) + k_2 x_2 + c_2 x_2' = 0. \end{cases} \tag{1.58}$$

The simplified form of Eq. (1.58) can be expressed as:

$$\begin{cases} x_1'' + \xi_{1,1} \left(x_1' - x_2'\right) + \omega_{1,1}^2 \left(x_1 - x_2\right) = 0, \\ x_2'' + \xi_{1,2} \left(x_2' - x_1'\right) + \omega_{1,2}^2 \left(x_2 - x_1\right) + \omega_{2,2}^2 x_2 + \xi_{2,2} x_2' = 0, \end{cases} \tag{1.59}$$

where $\omega_{1,1} = \sqrt{k_1/m_1}, \omega_{1,2} = \sqrt{k_1/m_2}, \omega_{2,2} = \sqrt{k_2/m_2}, \xi_{1,1} = c_1/m_1, \xi_{1,2} = c_1/m_2, \xi_{2,2} = c_2/m_2$.

In arbitrary interval $i/N \le t < (i + 1)/N$, Eq. (1.59) can be rearranged as a style of piecewise constant. Replace terms $x_1(t)$, $x_1'(t)$, $x_2(t)$, and $x_2'(t)$ with piecewise constant functions $x_{1,i}\left(\frac{i}{N}\right)$, $x_{1,i}'\left(\frac{i}{N}\right)$, $x_{2,i}\left(\frac{i}{N}\right)$, and $x_{2,i}'\left(\frac{i}{N}\right)$, and so Eq. (1.59) can expressed as:

$$\begin{cases} x_1'' + \xi_{1,1}\left[x_{1,i}'\left(\frac{i}{N}\right) - x_{2,i}'\left(\frac{i}{N}\right)\right] + \omega_{1,1}^2\left[x_{1,i}\left(\frac{i}{N}\right) - x_{2,i}\left(\frac{i}{N}\right)\right] = 0, \\ x_2'' + \xi_{1,2}\left[x_{2,i}'\left(\frac{i}{N}\right) - x_{1,i}'\left(\frac{i}{N}\right)\right] + \omega_{1,2}^2\left[x_{2,i}\left(\frac{i}{N}\right) - x_{1,i}\left(\frac{i}{N}\right)\right] + \omega_{2,2}^2 x_{2,i}\left(\frac{i}{N}\right) + \xi_{2,2} x_{2,i}'\left(\frac{i}{N}\right) = 0. \end{cases} \tag{1.60}$$

The Laplace transformation can be used to separate the coupling variables, and the transformed plural form of Eq. (1.60) in the interval $i/N \le t < (i + 1)/N$ is calculated as:

$$\begin{cases} s^2X_{1,i}(s) - sx_{1,i}\left(\frac{i}{N}\right) - x'_{1,i}\left(\frac{i}{N}\right) = -\frac{\xi_{1,1}}{s}\left[x'_{1,i}\left(\frac{i}{N}\right) - x'_{2,i}\left(\frac{i}{N}\right)\right] - \frac{\omega^2_{1,1}}{s}\left[x_{1,i}\left(\frac{i}{N}\right) - x_{2,i}\left(\frac{i}{N}\right)\right], \\ s^2X_{2,i}(s) - sx_{2,i}\left(\frac{i}{N}\right) - x'_{2,i}\left(\frac{i}{N}\right) = -\frac{\xi_{1,2}}{s}\left[x'_{2,i}\left(\frac{i}{N}\right) - x'_{1,i}\left(\frac{i}{N}\right)\right] - \frac{\omega^2_{1,2}}{s}\left[x_{2,i}\left(\frac{i}{N}\right) - x_{1,i}\left(\frac{i}{N}\right)\right] - \frac{\omega^2_{2,2}}{s}x_{2,i}\left(\frac{i}{N}\right) \\ -\frac{\xi_{2,2}}{s}x'_{2,i}\left(\frac{i}{N}\right). \end{cases} \tag{1.61}$$

Then, $X_{1,\,i}(s)$ and $X_{2,\,i}(s)$ can be expressed as:

$$X_{1,i}(s) = -\frac{\xi_{1,1}}{s^3}\left[x'_{1,i}\left(\frac{i}{N}\right) - x'_{2,i}\left(\frac{i}{N}\right)\right] - \frac{\omega^2_{1,1}}{s^3}\left[x_{1,i}\left(\frac{i}{N}\right) - x_{2,i}\left(\frac{i}{N}\right)\right] + \frac{1}{s}x_{1,i}\left(\frac{i}{N}\right) + \frac{1}{s^2}x'_{1,i}\left(\frac{i}{N}\right), \tag{1.62}$$

$$X_{2,i}(s) = -\frac{\xi_{1,2}}{s^3}\left[x'_{2,i}\left(\frac{i}{N}\right) - x'_{1,i}\left(\frac{i}{N}\right)\right] - \frac{\omega^2_{1,2}}{s^3}\left[x_{2,i}\left(\frac{i}{N}\right) - x_{1,i}\left(\frac{i}{N}\right)\right] - \frac{\omega^2_{2,2}}{s^3}x_{2,i}\left(\frac{i}{N}\right) - \frac{\xi_{2,2}}{s^3}x'_{2,i}\left(\frac{i}{N}\right) + \frac{1}{s}x_{2,i}\left(\frac{i}{N}\right) + \frac{1}{s^2}x'_{2,i}\left(\frac{i}{N}\right). \tag{1.63}$$

Through using the inverse Laplace transformation, the expressions of $x_{1,\,i}(t)$ and $x_{2,\,i}(t)$ can be obtained by the replacement of i with $[Nt]$:

$$x_{1,i}(t) = \mathbf{B}_1\mathbf{v}_i, \tag{1.64}$$

$$x_{2,i}(t) = \mathbf{B}_2\mathbf{v}_i, \tag{1.65}$$

where:

$\mathbf{v}_i = \left[d_{1,i}, v_{1,i}, d_{2,i}, v_{2,i}\right]^T$,
$\mathbf{B}_1 = \left[1 - \frac{1}{2}\omega^2_{1,1}\left(t - \frac{[Nt]}{N}\right)^2, t - \frac{[Nt]}{N} - \frac{1}{2}\xi_{1,1}\left(t - \frac{[Nt]}{N}\right)^2, \frac{1}{2}\omega^2_{1,1}\left(t - \frac{[Nt]}{N}\right)^2, \frac{1}{2}\xi_{1,1}\left(t - \frac{[Nt]}{N}\right)^2\right]$,
$\mathbf{B}_2 = \left[\frac{1}{2}\omega^2_{1,2}\left(t - \frac{[Nt]}{N}\right)^2, \frac{1}{2}\xi_{1,2}\left(t - \frac{[Nt]}{N}\right)^2, 1 - \frac{1}{2}\omega^2_{1,2}\left(t - \frac{[Nt]}{N}\right)^2 - \frac{1}{2}\omega^2_{2,2}\left(t - \frac{[Nt]}{N}\right)^2, \left(t - \frac{[Nt]}{N}\right) - \frac{1}{2}\xi_{1,2}\left(t - \frac{[Nt]}{N}\right)^2 - \frac{1}{2}\xi_{2,2}\left(t - \frac{[Nt]}{N}\right)^2\right]$.

Therefore, the velocities $x'_{1,i}(t)$ and $x'_{2,i}(t)$ should be rewritten as follows:

$$x'_{1,i} = \frac{\partial \mathbf{B}_1}{\partial t}\mathbf{v}_i, \tag{1.66}$$

$$x'_{2,i} = \frac{\partial \mathbf{B}_2}{\partial t}\mathbf{v}_i. \tag{1.67}$$

In the light of the continuity condition in Eq. (1.53), the relation of displacement and velocity between the two adjacent truncation points can be represented as:

$$\mathbf{v}_i = \mathbf{q}\mathbf{v}_{i-1}, \tag{1.68}$$

where $\mathbf{v}_{i-1} = [d_{1,i-1}, v_{1,i-1}, d_{2,i-1}, v_{2,i-1}]^T$, $\mathbf{q} = [q_{ij}]_{4\times 4}$ $(i = 1, \ldots, 4, j = 1, \ldots, 4)$, and the elements in matrix $\mathbf{q}$ are given by:

$$\begin{aligned}
&q_{11} = 1 - \frac{\omega_{1,1}^2}{2N^2}, q_{12} = \frac{1}{N} - \frac{\xi_{1,1}}{2N^2}, q_{13} = \frac{\omega_{1,1}^2}{2N^2}, q_{14} = \frac{\xi_{1,1}}{2N^2},\\
&q_{21} = -\frac{\omega_{1,1}^2}{N}, q_{22} = 1 - \frac{\xi_{1,1}}{N}, q_{23} = \frac{\omega_{1,1}^2}{N}, q_{24} = \frac{\xi_{1,1}}{N},\\
&q_{31} = \frac{\omega_{1,2}^2}{2N^2}, q_{32} = \frac{\xi_{1,2}}{2N^2}, q_{33} = 1 - \frac{\omega_{1,2}^2}{2N^2} - \frac{\omega_{2,2}^2}{2N^2}, q_{34} = \frac{1}{N} - \frac{\xi_{1,2}}{2N^2} - \frac{\xi_{2,2}}{2N^2},\\
&q_{41} = \frac{\omega_{1,2}^2}{N}, q_{42} = \frac{\xi_{1,2}}{N}, q_{43} = -\frac{\omega_{1,2}^2}{N} - \frac{\omega_{2,2}^2}{N}, q_{44} = 1 - \frac{\xi_{1,2}}{N} - \frac{\xi_{2,2}}{N}.
\end{aligned}$$

According to Eq. (1.68), $\mathbf{v}_i$ can be expressed via initial condition $\mathbf{v}_0$ with iterative computations of i times, i.e.:

$$\mathbf{v}_i = \mathbf{q}^i\mathbf{v}_0, \tag{1.69}$$

where $\mathbf{v}_0 = [d_{1,0}, v_{1,0}, d_{2,0}, v_{2,0}]^T$.

When initial condition $\mathbf{v}_0$ is determined, the displacement and velocity of the system at any time can be calculated by utilizing Eq. (1.69). The semi-analytical solution of the system in ith interval can be rewritten as the style of piecewise constant, i.e.:

$$x_{1,i}(t) = \mathbf{B}_1\mathbf{q}^{[Nt]}\mathbf{v}_0, \tag{1.70}$$

$$x_{2,i}(t) = \mathbf{B}_2\mathbf{q}^{[Nt]}\mathbf{v}_0. \tag{1.71}$$

The partial numerical results of solving SDCS for each method are shown in Table 1.3. Similarly, the maximum relative errors of each method can be seen: Euler method, 23.12%; trapezoidal method, 11.97%; Ode45, 6.78%; RK4, 10.18%; and PL, 1.52%. Likewise, the most accurate method is the PL method. As can be seen from the maximum relative error, compared to solving SCS, the relative error of each method has increased. This is due to the added damping term of the SDCS, which makes the calculation more complicated.

The displacement responses of SDCS are shown in Fig. 1.4, and it can be seen that the accuracy of numerical results calculated by RK4 is rougher than PL within time step 0.01 s; however, on account of continuity of the PL method, the numerical results within time step 0.01 s are coincident with the RK4 method within time step 0.001 s. The value of amplitude of the SDCS is gradually damped with the increase of time t. In order to evaluate the computational efficiency of the PL method, the CPU time for solving SDCS with PL and RK4 methods is considered. As shown in

Table 1.3 Numerical results (x_1) in solving SDCS

Times (s)	RK4 (step:0.01 s)	PL (step:0.01 s)	Euler method (step:0.01 s)	Trapezoidal method (step:0.01 s)	Ode45 (step:0.01 s)	RK4 (step:0.001 s)
0.2	−0.00639401932	−0.00617951731	−0.00567427061	−0.00633640952	−0.00605843095	−0.00617905128
0.4	−0.00051551486	−0.00064252979	−0.00050319658	−0.00060136485	−0.00059791509	−0.00064272665
0.6	0.00334637075	0.00372683985	0.00286306634	0.00327953860	0.00320153875	0.00372569651

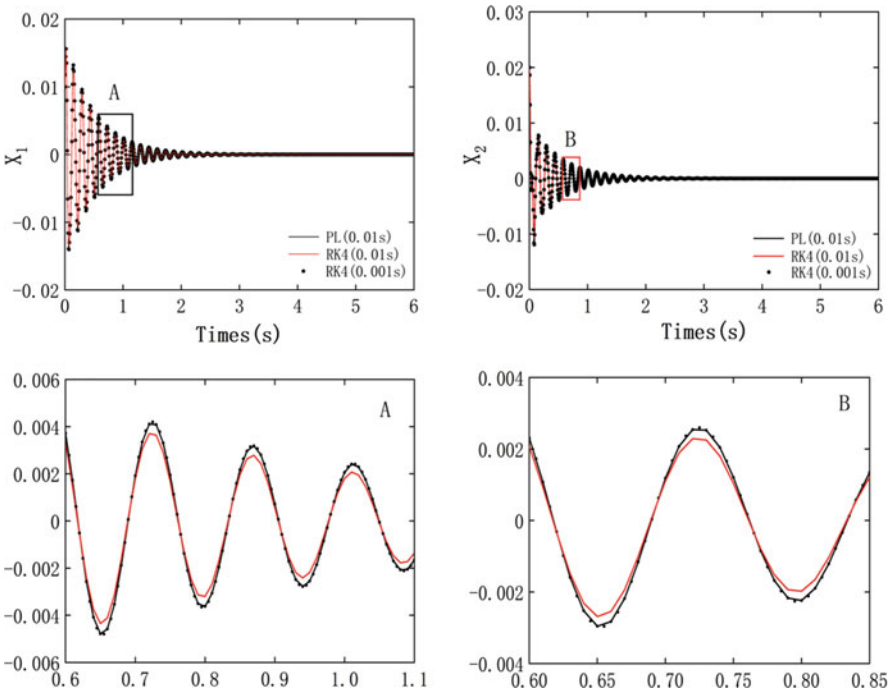


Fig. 1.4 Displacement response of SDCS ($\omega_{1,1} = 22.361$, $\omega_{1,2} = 20$, $\omega_{2,2} = 17.321$, $\xi_{1,1} = 10$, $\xi_{1,2} = 5$, $\xi_{2,2} = 2$, $x_1(0) = 0.01$, $x_1'(0) = 0.1$,$x_2(0) = 0.02$, $x_2'(0) = 0.2$.)

Table 1.4 SDCS computation time in time history 30s

Numerical method	Time step(s)	Iterations	CPU times
RK4	0.001	30,000	4.533125
RK4	0.01	3000	0.412515
PL	0.001	30,000	2.572750
PL	0.01	3000	0.296875

Table 1.4, the CPU time taken by the PL method is less than RK within the same time steps, which illustrates that PL calculating SDCS is faster than RK4 method. Comparing with Table 1.2, because of the existence of damping term, the CPU time taken by SDCS is longer than SCS.

1.3.3 Stiffness and Damping Coupling System with External Excitation

In practical engineering, the stiffness and damping coupling system is encountered with external excitation. Consider a cosine force related to time t acted in the system

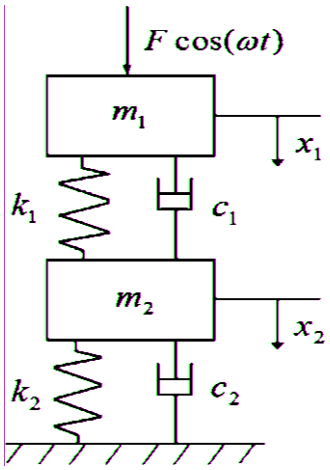


Fig. 1.5 The stiffness and damping coupling system with external excitation

(SDCSF), i.e. (Fig. 1.5): where $m_n(n = 1, 2)$ is the mass of the vibrator, $k_n(n = 1, 2)$ represents the stiffness coefficient of the spring, $c_n(n = 1, 2)$ is the damping coefficient of the damper, F is the magnitude of the cosine force, ω represents the frequency of the cosine force, and $x_n(n = 1, 2)$ denotes the displacement response of the vibrator. When the external excitation acting on the vibrator changes periodically with time, the vibrator will respond accordingly.

According to the physical model, the following equation is established:

$$\begin{cases} m_1 x_1'' + c_1 \left(x_1' - x_2'\right) + k_1 \left(x_1 - x_2\right) = F \cos\left(\omega t\right), \\ m_2 x_2'' + c_1 \left(x_2' - x_1'\right) + k_1 \left(x_2 - x_1\right) + k_2 x_2 + c_2 x_2' = 0. \end{cases} \tag{1.72}$$

The simplified form of Eq. (1.72) can be expressed as:

$$\begin{cases} x_1'' + \xi_{1,1} \left(x_1' - x_2'\right) + \omega_{1,1}^2 \left(x_1 - x_2\right) = a \cos\left(\omega t\right), \\ x_2'' + \xi_{1,2} \left(x_2' - x_1'\right) + \omega_{1,2}^2 \left(x_2 - x_1\right) + \omega_{2,2}^2 x_2 + \xi_{2,2} x_2' = 0, \end{cases} \tag{1.73}$$

where $a = F/m_1, \omega_{1,1} = \sqrt{k_1/m_1}, \omega_{1,2} = \sqrt{k_1/m_2}, \omega_{2,2} = \sqrt{k_2/m_2}, \xi_{1,1} = c_1/m_1, \xi_{1,2} = c_1/m_2, \xi_{2,2} = c_2/m_2$.

At time interval $i/N \le t < (i + 1)/N$, the piecewise constant method is employed to rearrange Eq. (1.73) as the style of piecewise constant. Replace terms $x_1'(t)$, $x_2'(t)$, and $a \cos(\omega t)$ with the piecewise constant functions $x_{1,i}'\left(\frac{i}{N}\right)$, $x_{2,i}'\left(\frac{i}{N}\right)$, and $a \cos\left(\omega \frac{i}{N}\right)$, respectively. Equation (1.73) can be expressed as:

$$\begin{cases} x_1'' + \xi_{1,1} \left[x_{1,i}'\left(\frac{i}{N}\right) - x_{2,i}'\left(\frac{i}{N}\right)\right] + \omega_{1,1}^2 \left(x_1 - x_2\right) = a \cos\left(\omega \frac{i}{N}\right), \\ x_2'' + \omega_{1,2}^2 \left(x_2 - x_1\right) + \xi_{1,2} \left[x_{2,i}'\left(\frac{i}{N}\right) - x_{1,i}'\left(\frac{i}{N}\right)\right] + \omega_{2,2}^2 x_2 + \xi_{2,2} x_{2,i}'\left(\frac{i}{N}\right) = 0. \end{cases} \tag{1.74}$$

The Laplace transformation can be used to separate the coupling variables, and the plural form of Eq. (1.74) in the interval $i/N \le t < (i+1)/N$ is calculated as:

$$\begin{cases} s^2 X_{1,i}(s) + \omega_{1,1}^2 \left[X_{1,i}(s) - X_{2,i}(s)\right] = a\ \cos\left(\omega \frac{i}{N}\right) - \xi_{1,1}\left[x'_{1,i}\left(\frac{i}{N}\right) - x'_{2,i}\left(\frac{i}{N}\right)\right] \\ \qquad\qquad + s x_{1,i}\left(\frac{i}{N}\right) + x'_{1,i}\left(\frac{i}{N}\right), \\ s^2 X_{2,i}(s) + \omega_{1,2}^2 \left[X_{2,i}(s) - X_{1,i}(s)\right] + \omega_{2,2}^2 X_{2,i}(s) = -\xi_{1,2}\left[x'_{2,i}\left(\frac{i}{N}\right) - x'_{1,i}\left(\frac{i}{N}\right)\right] \\ \qquad\qquad - \xi_{2,2} x'_{2,i}\left(\frac{i}{N}\right) + s x_{2,i}\left(\frac{i}{N}\right) + x'_{2,i}\left(\frac{i}{N}\right). \end{cases} \tag{1.75}$$

Then, $X_{1,i}(s)$ and $X_{2,i}(s)$ can be decoupled and expressed as:

$$\begin{aligned} X_{1,i}(s) = {} & \frac{sW}{QW - \omega_{1,1}^2\omega_{1,2}^2} x_{1,i}\left(\frac{i}{N}\right) + \left[\frac{W}{QW - \omega_{1,1}^2\omega_{1,2}^2} + \frac{\xi_{1,2}\omega_{1,1}^2 - \xi_{1,1}W}{s\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)}\right] x'_{1,i}\left(\frac{i}{N}\right) \\ & + \frac{s\omega_{1,1}^2}{QW - \omega_{1,1}^2\omega_{1,2}^2} x_{2,i}\left(\frac{i}{N}\right) + \left[\frac{\xi_{1,1}W - \xi_{1,2}\omega_{1,1}^2 - \xi_{2,2}\omega_{1,1}^2}{s\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)} + \frac{\omega_{1,1}^2}{QW - \omega_{1,1}^2\omega_{1,2}^2}\right] x'_{2,i}\left(\frac{i}{N}\right) \\ & + \frac{W}{s\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)} a\ \cos\left(\omega \frac{i}{N}\right), \end{aligned} \tag{1.76}$$

$$\begin{aligned} X_{2,i}(s) = {} & \frac{2s\omega_{1,2}^2}{QW - \omega_{1,1}^2\omega_{1,2}^2} x_{1,i}\left(\frac{i}{N}\right) + \left[\frac{\xi_{1,2}}{sW} + \frac{\omega_{1,2}^2}{QW - \omega_{1,1}^2\omega_{1,2}^2} - \frac{\xi_{1,1}\omega_{1,2}^2}{s\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)} + \frac{\xi_{1,2}\omega_{1,1}^2\omega_{1,2}^2}{sW\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)}\right] x'_{1,i}\left(\frac{i}{N}\right) \\ & + \left[\frac{1}{W} - \frac{\xi_{1,2} + \xi_{2,2}}{sW} + \frac{\xi_{1,1}\omega_{1,2}^2}{s\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)} + \frac{\omega_{1,1}^2\omega_{1,2}^2}{W\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)} - \frac{\xi_{1,2}\omega_{1,1}^2\omega_{1,2}^2 + \xi_{2,2}\omega_{1,1}^2\omega_{1,2}^2}{sW\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)}\right] x'_{2,i}\left(\frac{i}{N}\right) \\ & + \left[\frac{s}{W} + \frac{s\omega_{1,1}^2\omega_{1,2}^2}{W\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)}\right] x_{2,i}\left(\frac{i}{N}\right) + \frac{\omega_{1,2}^2}{s\left(QW - \omega_{1,1}^2\omega_{1,2}^2\right)} a\ \cos\left(\omega \frac{i}{N}\right), \end{aligned} \tag{1.77}$$

where $Q = s^2 + \omega_{1,1}^2$, $W = s^2 + \omega_{1,2}^2 + \omega_{2,2}^2$.

It is difficult to transform plural solution above into time-domain solution with the general inverse Laplace transformation. However, through Theorem 1.5 in Eq. (1.10), the expressions of $x_{1,i}(t)$ and $x_{2,i}(t)$ can be obtained by the replacement of i with $[Nt]$:

$$x_{1,i}(t) = \mathbf{C}_1 \mathbf{v}_i + \sum_{k=1}^{5} \frac{s_k^2 + \omega_{1,2}^2 + \omega_{2,2}^2}{5s_k^4 + 3s_k^2 L + \omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t - \frac{i}{N}\right)} a\ \cos\left(\omega \frac{i}{N}\right), \tag{1.78}$$

$$x_{2,i}(t) = \mathbf{C}_2 \mathbf{v}_i + \sum_{k=1}^{5} \frac{\omega_{1,2}^2}{5s_k^4 + 3s_k^2 L + \omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t - \frac{i}{N}\right)} a\ \cos\left(\omega \frac{i}{N}\right), \tag{1.79}$$

where $\mathbf{v}_i = \left[d_{1,i}, v_{1,i}, d_{2,i}, v_{2,i}\right]^T$, $L = \omega_{1,1}^2 + \omega_{1,2}^2 + \omega_{2,2}^2$,

$$\mathbf{C}_1 = \begin{bmatrix} \sum\limits_{k=1}^{4} \frac{s_k^2+\omega_{1,2}^2+\omega_{2,2}^2}{4s_k^2+2L} e^{s_k\left(t-\frac{i}{N}\right)} \\ \sum\limits_{k=1}^{4} \frac{s_k^2+\omega_{1,2}^2+\omega_{2,2}^2}{4s_k^3+2s_kL} e^{s_k\left(t-\frac{i}{N}\right)} + \sum\limits_{k=1}^{5} \frac{\xi_{1,2}\omega_{1,1}^2-\xi_{1,1}\left(s_k^2+\omega_{1,2}^2+\omega_{2,2}^2\right)}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} \\ \sum\limits_{k=1}^{4} \frac{\omega_{1,1}^2}{4s_k^2+2L} e^{s_k\left(t-\frac{i}{N}\right)} \\ \sum\limits_{k=1}^{4} \frac{\omega_{1,1}^2}{4s_k^3+2s_kL} e^{s_k\left(t-\frac{i}{N}\right)} + \sum\limits_{k=1}^{5} \frac{-\xi_{2,2}\omega_{1,1}^2-\xi_{1,2}\omega_{1,1}^2+\xi_{1,1}\left(s_k^2+\omega_{1,2}^2+\omega_{2,2}^2\right)}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} \end{bmatrix}^{\mathrm{T}},$$

$$\mathbf{C}_2 = \begin{bmatrix} \sum\limits_{k=1}^{4} \frac{\omega_{1,2}^2}{2s_k^2+L} e^{s_k\left(t-\frac{i}{N}\right)} \\ \sum\limits_{k=5}^{7} \frac{\xi_{1,2}}{3s_k^2+\omega_{1,2}^2+\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} + \sum\limits_{k=1}^{4} \frac{\omega_{1,2}^2}{4s_k^3+2s_kL} e^{s_k\left(t-\frac{i}{N}\right)} - \sum\limits_{k=1}^{5} \frac{\xi_{1,1}\omega_{1,2}^2}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} \\ + \sum\limits_{k=1}^{7} \frac{\xi_{1,2}\omega_{1,1}^2\omega_{1,2}^2}{7s_k^6+5s_k^4\left(\omega_{1,1}^2+2\omega_{1,2}^2+2\omega_{2,2}^2\right)+3s_k^2\left(2\omega_{1,1}^2\omega_{2,2}^2+\omega_{1,2}^4+\omega_{1,1}^2\omega_{1,2}^2+2\omega_{1,2}^2\omega_{2,2}^2+\omega_{2,2}^4\right)+\left(\omega_{1,1}^2\omega_{1,2}^2\omega_{2,2}^2+\omega_{1,1}^2\omega_{2,2}^4\right)} e^{s_k\left(t-\frac{i}{N}\right)} \\ \sum\limits_{k=5}^{7} \frac{1}{2} e^{s_k\left(t-\frac{i}{N}\right)} + \sum\limits_{k=1}^{4} \frac{\omega_{1,2}^2}{6s_k^4+4s_k^2\left(\omega_{1,1}^2+2\omega_{1,2}^2+2\omega_{2,2}^2\right)+2\left(2\omega_{1,1}^2\omega_{2,2}^2+\omega_{1,2}^4+\omega_{1,1}^2\omega_{1,2}^2+2\omega_{1,2}^2\omega_{2,2}^2+\omega_{2,2}^4\right)} e^{s_k\left(t-\frac{i}{N}\right)} \\ + \sum\limits_{k=6}^{7} \frac{\omega_{1,1}^2\omega_{1,2}^2}{6s_k^4+4s_k^2\left(\omega_{1,1}^2+2\omega_{1,2}^2+2\omega_{2,2}^2\right)+2\left(2\omega_{1,1}^2\omega_{2,2}^2+\omega_{1,2}^4+\omega_{1,1}^2\omega_{1,2}^2+2\omega_{1,2}^2\omega_{2,2}^2+\omega_{2,2}^4\right)} e^{s_k\left(t-\frac{i}{N}\right)} \\ \sum\limits_{k=5}^{7} \frac{1}{2s_k} e^{s_k\left(t-\frac{i}{N}\right)} - \sum\limits_{k=5}^{7} \frac{\xi_{1,2}+\xi_{2,2}}{3s_k^2+\omega_{1,2}^2+\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} + \sum\limits_{k=1}^{5} \frac{\xi_{1,1}\omega_{1,2}^2}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} \\ + \sum\limits_{k=1}^{5} \frac{\xi_{1,1}\omega_{1,2}^2}{6s_k^5+4s_k^3\left(\omega_{1,1}^2+2\omega_{1,2}^2+2\omega_{2,2}^2\right)+2s_k\left(2\omega_{1,1}^2\omega_{2,2}^2+\omega_{1,2}^4+\omega_{1,1}^2\omega_{1,2}^2+2\omega_{1,2}^2\omega_{2,2}^2+\omega_{2,2}^4\right)} e^{s_k\left(t-\frac{i}{N}\right)} \\ + \sum\limits_{k=6}^{7} \frac{\omega_{1,1}^2\omega_{1,2}^2}{6s_k^5+4s_k^3\left(\omega_{1,1}^2+2\omega_{1,2}^2+2\omega_{2,2}^2\right)+2s_k\left(2\omega_{1,1}^2\omega_{2,2}^2+\omega_{1,2}^4+\omega_{1,1}^2\omega_{1,2}^2+2\omega_{1,2}^2\omega_{2,2}^2+\omega_{2,2}^4\right)} e^{s_k\left(t-\frac{i}{N}\right)} \\ - \sum\limits_{k=1}^{7} \frac{\xi_{1,2}\omega_{1,1}^2\omega_{1,2}^2+\xi_{2,2}\omega_{1,1}^2\omega_{1,2}^2}{7s_k^6+5s_k^4\left(\omega_{1,1}^2+2\omega_{1,2}^2+2\omega_{2,2}^2\right)+3s_k^2\left(2\omega_{1,1}^2\omega_{2,2}^2+\omega_{1,2}^4+\omega_{1,1}^2\omega_{1,2}^2+2\omega_{1,2}^2\omega_{2,2}^2+\omega_{2,2}^4\right)+\left(\omega_{1,1}^2\omega_{1,2}^2\omega_{2,2}^2+\omega_{1,1}^2\omega_{2,2}^4\right)} e^{s_k\left(t-\frac{i}{N}\right)} \end{bmatrix}^{\mathrm{T}}.$$

Therefore, the velocities $x'_{1,i}(t)$ and $x'_{2,i}(t)$ should be rewritten as follows:

$$x'_{1,i}(t) = \frac{\partial \mathbf{C}_1}{\partial t}\mathbf{v}_i + \sum_{k=1}^{5} \frac{s_k^4+\omega_{1,2}^2s_k^2+\omega_{2,2}^2s_k^2}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} a\ \cos\left(\omega\frac{i}{N}\right), \tag{1.80}$$

$$x'_{2,i}(t) = \frac{\partial \mathbf{C}_2}{\partial t}\mathbf{v}_i + \sum_{k=1}^{5} \frac{\omega_{1,2}^2s_k}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k\left(t-\frac{i}{N}\right)} a\ \cos\left(\omega\frac{i}{N}\right). \tag{1.81}$$

The first-order zero in expression $5s_k^4 + 3s_k^2 L + \omega_{1,1}^2 \omega_{2,2}^2$ are shown below:

$$s_1 = -\left(-\frac{L}{2} - \frac{1}{2}M^{\frac{1}{2}}\right)^{\frac{1}{2}}, s_2 = -\left(\frac{1}{2}M^{\frac{1}{2}} - \frac{L}{2}\right)^{\frac{1}{2}},$$
$$s_3 = \left(-\frac{L}{2} - \frac{1}{2}M^{\frac{1}{2}}\right)^{\frac{1}{2}}, s_4 = \left(\frac{1}{2}M^{\frac{1}{2}} - \frac{L}{2}\right)^{\frac{1}{2}}, s_5 = 0,$$

where $M = \omega_{1,1}^2 + \omega_{1,2}^2 + \omega_{2,2}^2 + 2\omega_{1,1}^2\omega_{1,2}^2 + 2\omega_{1,2}^2\omega_{2,2}^2 - 2\omega_{1,1}^2\omega_{2,2}^2$.

According to the continuity condition in Eq. (1.53), the relation of displacement and velocity between the two adjacent truncation points can be represented as:

$$\mathbf{v}_i = \boldsymbol{\gamma}\mathbf{v}_{i\text{-}1} + \mathbf{g}\, a\, \cos\left(\omega\frac{i-1}{N}\right), \tag{1.82}$$

where: $\mathbf{g} = \left[\sum_{k=1}^{5} \frac{s_k^2+\omega_{1,2}^2+\omega_{2,2}^2}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k t},\ \sum_{k=1}^{5} \frac{s_k^3+\omega_{1,2}^2 s_k+\omega_{2,2}^2 s_k}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k t},\ \sum_{k=1}^{5} \frac{\omega_{1,2}^2}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k t},\ \sum_{k=1}^{5} \frac{\omega_{1,2}^2 s_k}{5s_k^4+3s_k^2L+\omega_{1,1}^2\omega_{2,2}^2} e^{s_k t}\right]^{\mathrm{T}}$, $\mathbf{v}_i = [d_{1,i}, v_{1,i}, d_{2,i}, v_{2,i}]^T$, $\mathbf{v}_i - 1$ {=} $[d_{1,\,i-1}, v_{1,\,i-1}, d_{2,\,i-1}, v_{2,\,i-1}]^T$, $\boldsymbol{\gamma} = [\gamma_{ij}]_{4\times 4}$ $(i = 1, \ldots, 4, \quad j = 1, \ldots, 4)$. The elements of matrix $\boldsymbol{\gamma}$ are shown as follows:

$$\gamma_{11} = \sum_{k=1}^{4} \frac{s_k^3+2\omega_{1,1}^2 s_k}{4s_k^3+6\omega_{1,1}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{12} = \sum_{k=1}^{5} \frac{-\xi_{1,1}s_k^2-\xi_{1,1}\omega_{1,1}^2+s_k^3+2\omega_{1,1}^2 s_k}{5s_k^4+9\omega_{1,1}^2 s_k^2+\omega_{1,1}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{13} = \sum_{k=1}^{4} \frac{\omega_{1,1}^2 s_k}{4s_k^3+6\omega_{1,1}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{14} = \sum_{k=1}^{5} \frac{\xi_{1,1}s_k^2+\xi_{1,1}\omega_{1,1}^2+\omega_{1,1}^2 s_k-\xi_{1,2}\omega_{1,1}^2}{5s_k^4+9\omega_{1,2}^2 s_k^2+\omega_{1,2}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{21} = \sum_{k=1}^{4} \frac{s_k^4+2\omega_{1,2}^2 s_k^2}{4s_k^3+6\omega_{1,2}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{22} = \sum_{k=1}^{5} \frac{-\xi_{1,2}s_k^3-\xi_{1,2}\omega_{1,2}^2 s_k+s_k^4+2\omega_{1,2}^2 s_k^2}{5s_k^4+9\omega_{1,2}^2 s_k^2+\omega_{1,2}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{23} = \sum_{k=1}^{4} \frac{\omega_{1,2}^2 s_k^2}{4s_k^3+6\omega_{1,2}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{24} = \sum_{k=1}^{5} \frac{\xi_{1,2}s_k^3+\omega_{1,2}^2 s_k^2}{5s_k^4+9\omega_{1,2}^2 s_k^2+\omega_{1,2}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{31} = \sum_{k=1}^{4} \frac{\omega_{1,2}^2 s_k^2}{4s_k^3+6\omega_{1,2}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{32} = \sum_{k=1}^{5} \frac{\omega_{1,2}^2 s_k+\xi_{1,2}s_k^2}{5s_k^4+9\omega_{1,2}^2 s_k^2+\omega_{1,2}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{33} = \sum_{k=1}^{4} \frac{s_k^3+\omega_{1,2}^2 s_k}{4s_k^3+6\omega_{1,2}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{34} = \sum_{k=1}^{5} \frac{s_k^3+\omega_{1,2}^2 s_k-\xi_{1,2}s_k^2-\xi_{1,2}\omega_{1,2}^2-\xi_{1,2}s_k^2}{5s_k^4+9\omega_{1,2}^2 s_k^2+\omega_{1,2}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{41} = \sum_{k=1}^{4} \frac{\omega_{2,2}^2 s_k^2}{4s_k^3+6\omega_{2,2}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{42} = \sum_{k=1}^{5} \frac{\omega_{2,2}^2 s_k^2+\xi_{2,2}s_k^3}{5s_k^4+9\omega_{2,2}^2 s_k^2+\omega_{2,2}^4} e^{\frac{s_k}{N}},$$
$$\gamma_{43} = \sum_{k=1}^{4} \frac{s_k^4+\omega_{2,2}^2 s_k^2}{4s_k^3+6\omega_{2,2}^2 s_k} e^{\frac{s_k}{N}}, \gamma_{44} = \sum_{k=1}^{5} \frac{s_k^4+\omega_{2,2}^2 s_k^2-\xi_{2,2}s_k^3-\xi_{2,2}\omega_{2,2}^2 s_k-\xi_{2,2}s_k^3}{5s_k^4+9\omega_{2,2}^2 s_k^2+\omega_{2,2}^4} e^{\frac{s_k}{N}}.$$

According to Eq. (1.82), $\mathbf{v}_i$ can be expressed via initial condition $\mathbf{v}_0$ with iterative computations of i times, i.e.:

$$\mathbf{v}_i = \boldsymbol{\gamma}^i \mathbf{v}_0 + \sum_{k=1}^{i} \boldsymbol{\gamma}^{k-1}\mathbf{g}\, a\, \cos\left(\omega\frac{i-1}{N}\right). \tag{1.83}$$

When the initial condition $\mathbf{v}_0$ is determined, the displacement and velocity of the system at any time can be calculated by using Eq. (1.83). The approximate solution of the system in ith interval can be rewritten as the style of piecewise constant, i.e.:

$$x_{1,i}(t) = \mathbf{C}_1 \left[\boldsymbol{\gamma}^{[Nt]} \mathbf{v}_0 + \sum_{k=1}^{[Nt]} \boldsymbol{\gamma}^{k-1} \mathbf{g}\, a\, \cos\left(\omega \frac{[Nt]-1}{N}\right) \right] + \sum_{k=1}^{5} \frac{s_k^2 + \omega_{1,2}^2 + \omega_{2,2}^2}{5s_k^4 + 3s_k^2 L + \omega_{1,1}^2 \omega_{2,2}^2} e^{s_k\left(t - \frac{[Nt]}{N}\right)} a\, \cos\left(\omega \frac{[Nt]}{N}\right), \tag{1.84}$$

$$x_{2,i}(t) = \mathbf{C}_2 \left[\boldsymbol{\gamma}^{[Nt]} \mathbf{v}_0 + \sum_{k=1}^{[Nt]} \boldsymbol{\gamma}^{k-1} \mathbf{g}\, a\, \cos\left(\omega \frac{[Nt]-1}{N}\right) \right] + \sum_{k=1}^{5} \frac{\omega_{1,2}^2}{5s_k^4 + 3s_k^2 L + \omega_{1,1}^2 \omega_{2,2}^2} e^{s_k\left(t - \frac{[Nt]}{N}\right)} a\, \cos\left(\omega \frac{[Nt]}{N}\right). \tag{1.85}$$

The partial numerical results of solving SDCSF for each method are as shown in Table 1.5. The maximum relative errors of each method can be found: Euler method, 38.47%; trapezoidal method, 18.21%; Ode45, 8.75%; RK4, 16.44%; and PL, 3.41%. The PL method still has the highest accuracy. Compared to solving SDCS, the relative error of each method is increasing. This is due to the added external excitation related to time in SDCSF, which makes the calculation process more complicated and leads to a reduction in accuracy.

Figure 1.6 shows the displacement responses of the SDCSF. It follows that the accuracy of PL results is more accurate than RK4 within same time step; but PL solutions with time step 0.01 s are more approximately approached to RK4 with time step 0.001 s. As shown in Table 1.6, computational speed of the PL method is faster than RK4 when calculating SDCSF, and the CPU time required in solving the SDCSF is longer than that of SCS and SDCS in time step 0.01 s because of the presence of damping and external excitation.

1.3.4 Convergence Analysis of the PL Method

The convergence of the PL method in calculating the above three systems is discussed, as illustrated in Figs. 1.7, 1.8 and 1.9. It can be seen that when $N = 1000$, the numerical results obtained by the PL method are deviated from the accurate value (AV). When $N = 2000$, the deviation of the numerical solution is reduced. Finally, when $N = 20{,}000$, the curve obtained by the PL method is coincident with that of AV. It is indicated that the numerical value obtained by the PL method is approached to the AV as the value of N is gradually increased. Obviously, parameter N is an important factor to control the precision of numerical results and speed

Table 1.5 Numerical results (x_2) in solving SDCSF

Times (s)	RK4 (step:0.01 s)	PL (step:0.01 s)	Euler method (step:0.01 s)	Trapezoidal method (step:0.01 s)	Ode45 (step:0.01 s)	RK4 (step:0.001 s)
1.2	0.00020859293	0.00013342414	0.00008574526	0.00011408736	0.00013016784	0.00013948698
1.4	0.00040135417	0.00037231668	0.00028704714	0.00034702754	0.00037051934	0.00037181098
1.6	0.00030952914	0.00033551885	0.00025139344	0.00030593157	0.00031803561	0.00032923365

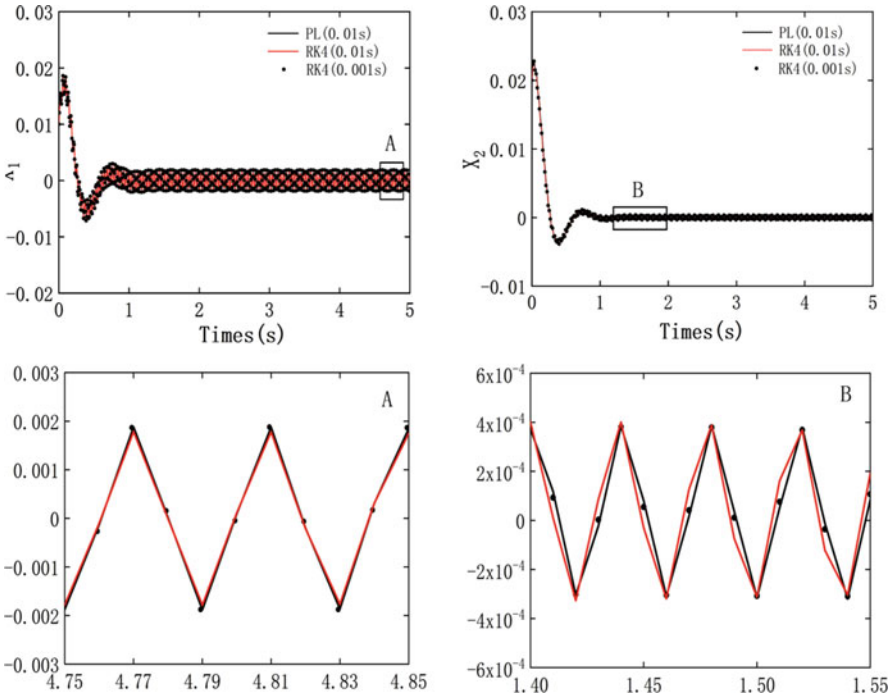


Fig. 1.6 Displacement response of SDCSF ($\omega_{1,1} = 15.811$, $\omega_{1,2} = 17.321$, $\omega_{2,2} = 20$, $\xi_{1,1} = 30$, $\xi_{1,2} = 20$, $\xi_{2,2} = 10$, $a = 50$, $\omega = 160$, $x_1(0) = 0.01$, $x_1'(0) = 0.1$,$x_2(0) = 0.02$, $x_2'(0) = 0.2$.)

Table 1.6 SDCSF computation time in time history 30s

Numerical method	Time step(s)	Iterations	CPU times
RK4	0.001	30,000	4.943750
RK4	0.01	3000	0.457751
PL	0.001	30,000	3.062512
PL	0.01	3000	0.371920

of convergence. Therefore, when the value of N is increased, the calculation can converge faster and the precision of the PL method is improved. In order to obtain the high-precision results and faster convergence for solving the coupling dynamic systems, it is better to select a sufficiently large value of N.

The above are the general processes of using PL method to solve the stiffness coupling systems, and some important conclusions should be stressed as the following:

1. When solving the coupling system by using the traditional RK method, high order differential equations of dynamic system are usually descended into multiple first-order differential equations. In this way, crucial dynamic property of the original system may be neglected or simplified, which leads to the computational error when searching solution for dynamic model encountered in

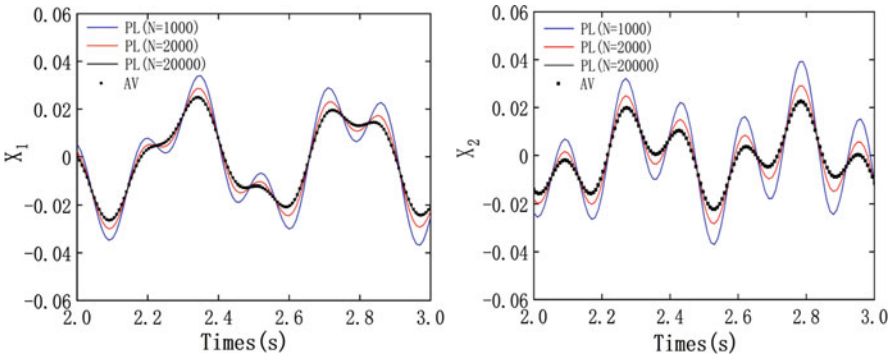


Fig. 1.7 Convergence of the solution of SCS by the PL method ($\omega_{1,1} = 22.361$, $\omega_{1,2} = 20$, $\omega_{2,2} = 17.321$, $x_1(0) = 0.01$, $x_1'(0) = 0.1$,$x_2(0) = 0.02$, $x_2'(0) = 0.2$.)

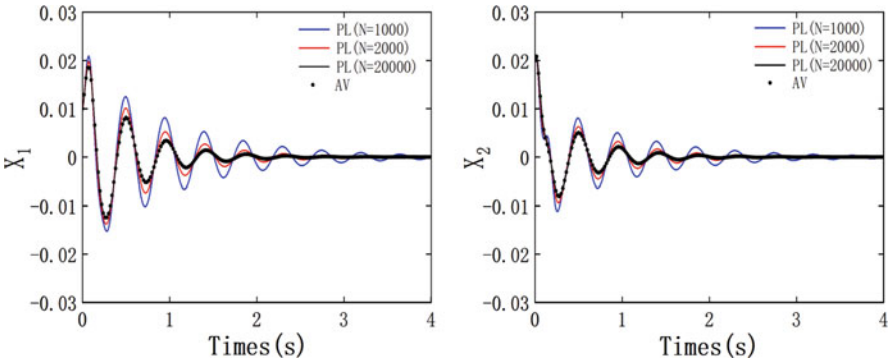


Fig. 1.8 Convergence of the solution of SCDS by the PL method ($\omega_{1,1} = 22.361$, $\omega_{1,2} = 20$, $\omega_{2,2} = 17.321$, $\xi_{1,1} = 10$, $\xi_{1,2} = 5$, $\xi_{2,2} = 2$, $x_1(0) = 0.01$, $x_1'(0) = 0.1$,$x_2(0) = 0.02$, $x_2'(0) = 0.2$.)

actual engineering. However, solving differential equations above with the PL method, the semi-analytical solution of systems is obtained directly through the derivation of equation. Therefore, the physical properties of systems are well preserved. In addition, the whole-time interval is divided into many tiny intervals by the PL method, and the solution is continuous on each interval. In this case, the accuracy of numerical solution of the PL method is better than the RK method, which is demonstrated by numerical analysis.

2. The computed efficiency of the PL and RK methods are compared through statistics of operation time of CPU during computing process. The CPU time taken by the PL method is less than the RK method in the same time steps, which indicates that the PL method is more efficient than the RK method. The main reason is that value continuity of two adjacent truncation points is maintained by using the PL method, which simplifies the solving process and saves the

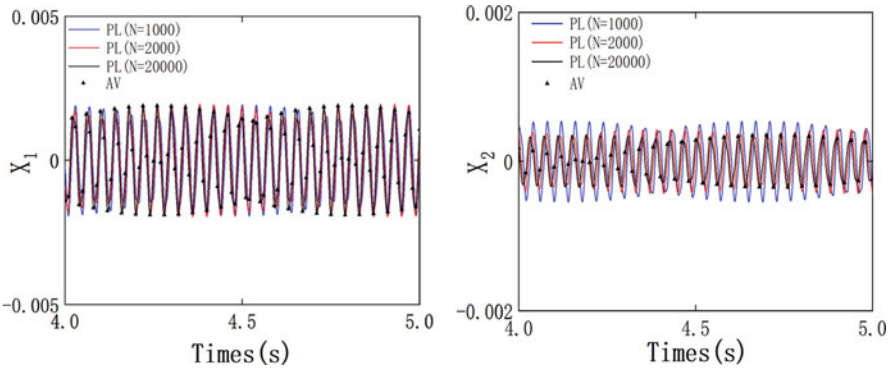


Fig. 1.9 Convergence of the solution of SCDSF by the PL method ($\omega_{1,1} = 15.811$, $\omega_{1,2} = 17.321$, $\omega_{2,2} = 20$, $\xi_{1,1} = 30$, $\xi_{1,2} = 20$, $\xi_{2,2} = 10$, $a = 50$, $\omega = 160$, $x_1(0) = 0.01$, $x_1'(0) = 0.1$,$x_2(0) = 0.02$, $x_2'(0) = 0.2$.)

calculated time, while the RK method combines iterative and averaged slope to search solutions, which makes the solution process more complicated.

3. The numerical results of the PL method reflect the essence of dynamic system, since the PL method keeps the physical characteristics of dynamic system. The precision of the solution obtained by the PL method is related to the value of *N*. This article is an exploratory study for implementation of the PL method in coupling systems. Therefore, the classic two-degree freedom systems are considered, and whether the PL method is suitable to solve problems of multi-degree freedom system should be further verified in our next work.

1.4 Analytical and Numerical Solutions of Inertial Coupling Systems

There is no direct solution for the dynamic equations, which are fundamentally mutual coupling with inertia terms. Development of solutions for such equation is therefore unique in comparing with conventional dynamic equation. For purpose of simplification and demonstration of PL method, the solutions for the following equation of inertial coupling dynamic systems are primarily considered.

1.4.1 Undamped Inertial Coupling System

Consider an undamped inertial coupling equation such as:

$$\begin{cases} m_1\ddot{x} + k_x x + m_2\ddot{y} = 0, \\ m_3\ddot{y} + k_y y + m_4\ddot{x} = 0. \end{cases} \tag{1.86}$$

where $m_i (i = 1, 2, 3, 4)$ is constant related to masses. x and y represent the displacement response in x and y directions, respectively. k_x and k_y represent the spring stiffness in x and y directions in the vibrating system, respectively.

The simplification style of the equation above can be expressed by:

$$\begin{cases} \ddot{x} + \omega_x^2 x + m\ddot{y} = 0, \\ \ddot{y} + \omega_y^2 y + n\ddot{x} = 0. \end{cases} \tag{1.87}$$

where $\omega_x = \sqrt{k_x/m_1}, \omega_y = \sqrt{k_y/m_3}, m = m_2/m_1, n = m_4/m_3$. The initial condition for the system may be assumed as follows:

$$\begin{cases} x(0) = d_0, \dot{x}(0) = v_0, \\ y(0) = \underset{\sim 0}{d}, \dot{y}(0) = \underset{\sim 0}{v}. \end{cases} \tag{1.88}$$

Then, the piecewise constant argument is applied to transform the continuous Eq. (1.87) into many piecewise constant systems. The system corresponding to that governed by Eq. (1.87) can be constructed by replacing the terms $x(t)$ and $y(t)$ with the piecewise constant function over an arbitrary time interval $i/N \le t < (i + 1)/N$ $(i = [Nt]/N)$. The corresponding equation of motion is expressible in the following form:

$$\begin{cases} \ddot{x}_i + m\ddot{y}_i = -\omega_x^2 x\left(\frac{[Nt]}{N}\right), \\ \ddot{y}_i + n\ddot{x}_i = -\omega_y^2 y\left(\frac{[Nt]}{N}\right). \end{cases} \tag{1.89}$$

In this case, in arbitrary time interval $i/N \le t < (i + 1)/N$, the local initial conditions can be considered as:

$$\begin{cases} x_i\left(\frac{i}{N}\right) = d_i, \dot{x}_i\left(\frac{i}{N}\right) = v_i, \\ y_i\left(\frac{i}{N}\right) = \underset{\sim i}{d}, y_i\left(\frac{i}{N}\right) = \underset{\sim i}{v}. \end{cases} \tag{1.90}$$

With the consideration of the Laplace transformation, Eq. (1.89) in interval $i/N \le t < (i + 1)/N$ can be expressed by plural form:

$$\begin{cases} s^2 X_i(s) + ms^2 Y_i(s) = -\frac{1}{s}\omega_x^2 x_i\left(\frac{i}{N}\right) + s x_i\left(\frac{i}{N}\right) + \dot{x}_i\left(\frac{i}{N}\right) + m s y_i\left(\frac{i}{N}\right) + m\dot{y}_i\left(\frac{i}{N}\right) \\ s^2 Y_i(s) + ns^2 X_i(s) = -\frac{1}{s}\omega_y^2 y_i\left(\frac{i}{N}\right) + ns x_i\left(\frac{i}{N}\right) + n\dot{x}_i\left(\frac{i}{N}\right) + s y_i\left(\frac{i}{N}\right) + \dot{y}_i\left(\frac{i}{N}\right) \end{cases} \tag{1.91}$$

Then, X_i and Y_i can be solved by:

$$X_i(s) = \left[\frac{\omega_x^2}{(mn-1)\,s^3} + \frac{1}{s}\right] x_i\left(\frac{i}{N}\right) + \frac{1}{s^2}\dot{x}_i\left(\frac{i}{N}\right) - \frac{m\omega_y^2}{(mn-1)\,s^3} y_i\left(\frac{i}{N}\right), \tag{1.92}$$

$$Y_i(s) = \frac{-n\omega_x^2}{(mn-1)\,s^3} x_i\left(\frac{i}{N}\right) + \left[\frac{1}{s} + \frac{\omega_y^2}{(mn-1)\,s^3}\right] y_i\left(\frac{i}{N}\right) + \frac{1}{s^2}\dot{y}_i\left(\frac{i}{N}\right). \tag{1.93}$$

$X_i(s)$ and $Y_i(s)$ are pluralities corresponding to $x_i(t)$ and $y_i(t)$ in interval $i/N \le t < (i+1)/N$, respectively. Based on inverse Laplace transformation and initial condition in Eq. (1.90), $x_i(t)$ and $y_i(t)$ is expressible in arbitrary time interval $i/N \le t < (i+1)/N$ with the substitution of i by $[Nt]$:

$$x_i(t) = \mathbf{A}_1 \mathbf{v}_i \tag{1.94}$$

$$y_i(t) = \mathbf{B}_1 \mathbf{v}_i \tag{1.95}$$

where:

$$\mathbf{A}_1 = \left[1 - \tfrac{\omega_x^2}{2(1-mn)}\left(t - \tfrac{[Nt]}{N}\right)^2,\ t - \tfrac{[Nt]}{N},\ \tfrac{m\omega_y^2}{2(1-mn)}\left(t - \tfrac{[Nt]}{N}\right)^2,\ 0\right],$$
$$\mathbf{B}_1 = \left[\tfrac{n\omega_x^2}{2(1-mn)}\left(t - \tfrac{[Nt]}{N}\right)^2,\ 0,\ \left[1 - \tfrac{\omega_y^2}{2(1-mn)}\right]\left(t - \tfrac{[Nt]}{N}\right)^2,\ t - \tfrac{[Nt]}{N}\right],$$
$$\mathbf{v}_i = \left[d_i\ \ v_i\ \ \underset{\sim i}{d}\ \ \underset{\sim i}{v}\right]^{\mathrm{T}}.$$

Consequently, the velocities of the system in x- and y- directions can be expressed by:

$$\dot{x}_i(t) = \frac{\partial \mathbf{A}_1}{\partial t}\mathbf{v}_i, \tag{1.96}$$

$$\dot{y}_i(t) = \frac{\partial \mathbf{B}_1}{\partial t}\mathbf{v}_i. \tag{1.97}$$

As a rotational hypothesis of the actual response of a piecewise constant system in practice, there should be no jump or discontinuity of the displacements $x_i(t)$ and $y_i(t)$ and velocities $\dot{x}_i(t)$ and $\dot{y}_i(t)$ in the time of $t \in [0, +\infty)$. This implies the kinematic parameters $x_i(t)$, $\dot{x}_i(t)$, $y_i(t)$, and $\dot{y}_i(t)$ are continuous within time t. The following condition of continuity should therefore be satisfied for the solutions of piecewise constant system over all of time intervals, i.e.:

$$\begin{cases} x_i\left(\frac{i}{N}\right) = x_{i-1}\left(\frac{i}{N}\right) \ , \ \dot{x}_i\left(\frac{i}{N}\right) = \dot{x}_{i-1}\left(\frac{i}{N}\right), \\ y_i\left(\frac{i}{N}\right) = y_{i-1}\left(\frac{i}{N}\right) \ , \ \dot{y}_i\left(\frac{i}{N}\right) = \dot{y}_{i-1}\left(\frac{i}{N}\right). \end{cases} \tag{1.98}$$

With the condition of continuity, a recursive relation is obtained with the consideration of Eqs. (1.94), (1.95), (1.96) and (1.97), such that:

$$\mathbf{v}_i = \boldsymbol{\alpha}\mathbf{v}_{i-1} \tag{1.99}$$

where $\boldsymbol{\alpha} = [\alpha_{ij}]_{4\times 4} \quad (i = 1, \ldots, 4, \ \ j = 1, \ldots, 4)$, $\mathbf{v}_{i-1} = \left[d_{i-1} \ v_{i-1} \ \underset{\sim}{d}_{i-1} \ \underset{\sim}{v}_{i-1} \right]^{\mathrm{T}}$. $\boldsymbol{\alpha}$ is a fourth-order matrix, and the elements of the matrix are shown as the following:

$$\begin{aligned}
&\alpha_{11} = 1 - \tfrac{\omega_x^2}{2(1-mn)N^2}, \alpha_{12} = \tfrac{1}{N}, \alpha_{13} = \tfrac{m\omega_y^2}{2(1-mn)N^2}, \ \ \alpha_{14} = 0, \\
&\alpha_{21} = -\tfrac{\omega_x^2}{(1-mn)N}, \ \ \alpha_{22} = 1, \quad \alpha_{23} = \tfrac{m\omega_y^2}{(1-mn)N}, \ \ \alpha_{24} = 0, \\
&\alpha_{31} = \tfrac{n\omega_x^2}{2(1-mn)N^2}, \ \ \alpha_{32} = 0, \quad \alpha_{33} = 1 - \tfrac{\omega_y^2}{2(1-mn)N^2}, \ \ \alpha_{34} = \tfrac{1}{N}, \\
&\alpha_{41} = \tfrac{n\omega_x^2}{(1-mn)N}, \ \ \alpha_{42} = 0, \quad \alpha_{43} = -\tfrac{\omega_y^2}{(1-mn)N}, \ \ \alpha_{44} = 1.
\end{aligned}$$

Because of an iterative procedure, $\mathbf{v}_i$ can be expressed by the initial displacements $\mathbf{v}_0$ in the following form:

$$\mathbf{v}_i = \boldsymbol{\alpha}^{[Nt]}\mathbf{v}_0 \tag{1.100}$$

where $\mathbf{v}_0 = \left[d_0 \ v_0 \ \underset{\sim}{d}_0 \ \underset{\sim}{v}_0 \right]^{\mathrm{T}}$.

One can see that the displacement and velocity of the system at any given point of time $[Nt]/N$ can be calculated by using Eq. (1.100) on the condition that the initial values of displacement and velocity are known. Considering that ith time interval is arbitrarily chosen, the complete solution of the system is obtained by:

$$x_i(t) = \mathbf{A_1}\boldsymbol{\alpha}^{[Nt]}\mathbf{v}_0 \tag{1.101}$$

$$y_i(t) = \mathbf{B_1}\boldsymbol{\alpha}^{[Nt]}\mathbf{v}_0 \tag{1.102}$$

The calculation program can be compiled by using the above expressions. To express the difference of the RK4 method and the PL method for computing the solution of the dynamic system, the comparison of the two methods with different time step is treated under the same initial value. Figure 1.10 shows the difference of the computation results for the inertial coupling system with undamping term (ICSUT). By using longer time step (0.01 s) with the PL method, one can obtain the same accurate solution as classical fourth-order RK4 method with shorter time step (0.001 s). However, the numerical results of RK4 method within time step 0.01 s are less accurate than that of the PL method within time step 0.01 and that of the

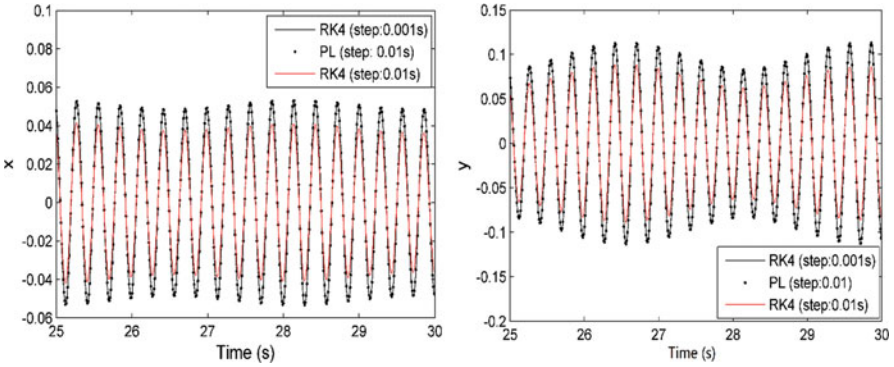


Fig. 1.10 Comparison of the numerical results with the RK4 method and the PL method for Eq. (1.86) ($x_0 = 0.1$, $\dot{x}_0 = 0.2$,$y_0 = 0$, $\dot{y}_0 = 0.1$.)

Table 1.7 Comparison of the numerical result for solving the ICSUT

Time (s)	RK4 (Step:0.01 s)	RK4 (Step:0.001 s)	PL (Step:0.01 s)
25.5	0.016473797231490	0.0190415455200814	0.0190694960305990
26.0	0.036368197172967	−0.0472301678349378	−0.0472486938586073
26.5	−0.013064521870364	−0.0147807819379706	−0.0147876980845763
27.0	0.035980540467921	0.0472258341798137	0.0472256879462556
27.5	0.012015075793593	0.0130923119300392	0.0130233245487664
28.0	−0.040012192731396	−0.0523579818551743	−0.0523989783656799

Table 1.8 Comparison of the CPU time computed for solving the ICSUT in time history 50s

Numerical method	Time step (s)	Iterations	CPU times
RK4	0.001	50,000	4.7813
RK4	0.01	5000	0.4063
PL	0.001	50,000	2.4219
PL	0.01	5000	0.2656

RK4 method within time step 0.001 s. This is mainly due to the cumulative error of iteration computation of the RK4 method in longer time step. Table 1.7 shows the computation values of the PL method and the RK4 method. It is indicated that the PL method is more accurate for solving the problem of dynamic system, because numerical solutions produced by the PL method are continuous everywhere in whole time history. Therefore, if a fixed accuracy of the numerical solution is required, the PL method with long time step can obtain the same or even more accuracy solution than the RK4 method. Table 1.8 shows the CPU time for solving ICSUT in the same time history. One can see that in the same time step, the PL method needs shorter CPU time to finish to the numerical computations than the RK4 method. This implies that the PL method is more efficient for computation of ICSUT than the RK4 method.

1.4.2 Damped Inertial Coupling System

In general, damping and resistance forces against motion may be existed in the inertia coupling system. For solving for the motion of a damped inertia coupling system, the dynamic equation in the following specific system is taken into consideration:

$$\begin{cases} m_1\ddot{x} + f_x\dot{x} + k_x x + m_2\ddot{y} = 0, \\ m_3\ddot{y} + f_y\dot{y} + k_y y + m_4\ddot{x} = 0. \end{cases} \tag{1.103}$$

With simplification, the dynamic equation can be also rewritten as:

$$\begin{cases} \ddot{x} + \xi_x\dot{x} + \omega_x^2 x + m\ddot{y} = 0, \\ \ddot{y} + \xi_y\dot{y} + \omega_y^2 y + n\ddot{x} = 0. \end{cases} \tag{1.104}$$

where $\omega_x = \sqrt{k_x/m_1}, \omega_y = \sqrt{k_y/m_3}, m = m_2/m_1, n = m_4/m_3, \xi_x = f_x/m_1, \xi_y = f_y/m_3$.

Then, the piecewise constant system corresponding to that governed by Eq. (1.104) can be constructed by replacing the terms $\dot{x}(t)$ and $\dot{y}(t)$ with piecewise constant function over an arbitrary time interval $i/N \le t < (i+1)/N$. The corresponding equation of motion is expressible in the following form:

$$\begin{cases} \ddot{x}_i + m\ddot{y}_i = -\xi_x\dot{x}_i\left(\frac{i}{N}\right) - \omega_x^2 x_i\left(\frac{i}{N}\right), \\ \ddot{y}_i + n\ddot{x}_i = -\xi_y\dot{y}_i\left(\frac{i}{N}\right) - \omega_y^2 y_i\left(\frac{i}{N}\right). \end{cases} \tag{1.105}$$

With the Laplace transformation, Eq. (1.105) can be also expressed as plural form:

$$\begin{cases} s^2X_i(s) + ms^2Y_i(s) = -\frac{1}{s}\omega_x^2 x_i\left(\frac{i}{N}\right) - \frac{1}{s}\xi_x\dot{x}_i\left(\frac{i}{N}\right) + sx_i\left(\frac{i}{N}\right) + \dot{x}_i\left(\frac{i}{N}\right) + msy_i\left(\frac{i}{N}\right) + m\dot{y}_i\left(\frac{i}{N}\right) \\ s^2Y_i(s) + ns^2X_i(s) = -\frac{1}{s}\omega_y^2 y_i\left(\frac{i}{N}\right) - \frac{1}{s}\xi_y\dot{y}_i\left(\frac{i}{N}\right) + nsx_i\left(\frac{i}{N}\right) + n\dot{x}_i\left(\frac{i}{N}\right) + sy_i\left(\frac{i}{N}\right) + \dot{y}_i\left(\frac{i}{N}\right) \end{cases} \tag{1.106}$$

Then, X_i and Y_i can be solved by:

$$X_i(s) = \left[\frac{1}{s} + \frac{\omega_x^2}{(mn-1)s^3}\right]x_i\left(\frac{i}{N}\right) + \left[\frac{1}{s^2} + \frac{\xi_x}{(mn-1)s^3}\right]\dot{x}_i\left(\frac{i}{N}\right) - \frac{m\omega_y^2}{(mn-1)s^3}y_i\left(\frac{i}{N}\right) - \frac{m\xi_y}{(mn-1)s^3}\dot{y}_i\left(\frac{i}{N}\right) \tag{1.107}$$

$$Y_i(s) = -\frac{n\omega_x^2}{(mn-1)s^3}x_i\left(\frac{i}{N}\right) - \frac{n\xi_x}{(mn-1)s^3}\dot{x}_i\left(\frac{i}{N}\right)$$

$$+\left[\frac{1}{s}+\frac{\omega_y^2}{(mn-1)\,s^3}\right]y_i\left(\frac{i}{N}\right)+\left[\frac{1}{s^2}+\frac{\xi_y}{(mn-1)\,s^3}\right]\dot{y}_i\left(\frac{i}{N}\right) \tag{1.108}$$

Based on inverse Laplace transformation, $x_i(t)$ and $y_i(t)$ can be obtained in arbitrary time interval $[Nt]/N \le t < ([Nt]+1)/N$:

$$x_i = \mathbf{A_2}\mathbf{v}_i \tag{1.109}$$

$$y_i = \mathbf{B_2}\mathbf{v}_i \tag{1.110}$$

with:

$$\mathbf{A}_2 = \begin{bmatrix} 1+\frac{\omega_x^2}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \\ \left(t-\frac{[Nt]}{N}\right)+\frac{\xi_x}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \\ -\frac{m\omega_y^2}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \\ -\frac{m\xi_y}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \end{bmatrix}^T,$$

$$\mathbf{B}_2 = \begin{bmatrix} -\frac{n\omega_x^2}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \\ -\frac{n\xi_x}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \\ 1+\frac{\omega_y^2}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \\ \left(t-\frac{[Nt]}{N}\right)+\frac{\xi_y}{2(mn-1)}\left(t-\frac{[Nt]}{N}\right)^2 \end{bmatrix}^T.$$

Then, the velocities of the system in x- and y- directions can be expressed by:

$$\dot{x}_i = \frac{\partial \mathbf{A_2}}{\partial t}\mathbf{v}_i \tag{1.111}$$

$$\dot{y}_i = \frac{\partial \mathbf{B_2}}{\partial t}\mathbf{v}_i \tag{1.112}$$

With the condition of continuity, the recursive relation is obtained by:

$$\mathbf{v}_i = \boldsymbol{\beta}\mathbf{v}_{i-1} \tag{1.113}$$

where $\boldsymbol{\beta} = [\beta_{ij}]_{4\times 4} \quad (i = 1, \ldots, 4, \quad j = 1, \ldots, 4)$, and the elements of matrix $\boldsymbol{\beta}$ are shown as the following:

$$\beta_{11} = 1 - \frac{\omega_x^2}{2(1-mn)N^2}, \beta_{12} = \frac{1}{N} - \frac{\xi_x}{2(1-mn)N^2}, \beta_{13} = \frac{m\omega_y^2}{2(1-mn)N^2}, \beta_{14} = \frac{m\xi_y}{2(1-mn)N^2},$$
$$\beta_{21} = -\frac{\omega_x^2}{(1-mn)N}, \beta_{22} = 1 - \frac{\xi_x}{(1-mn)N}, \beta_{23} = \frac{m\omega_y^2}{(1-mn)N}\beta_{24} = \frac{m\xi_y}{(1-mn)N},$$
$$\beta_{31} = \frac{n\omega_x^2}{2(1-mn)N^2}, \beta_{32} = \frac{n\xi_x}{2(1-mn)N^2}, \beta_{33} = 1 - \frac{\omega_y^2}{2(1-mn)N^2}, \beta_{34} = \frac{1}{N} - \frac{\xi_y}{2(1-mn)N^2},$$
$$\beta_{41} = \frac{n\omega_x^2}{(1-mn)N}, \beta_{42} = \frac{n\xi_x}{(1-mn)N}, \beta_{43} = -\frac{\omega_y^2}{(1-mn)N}, \beta_{44} = 1 - \frac{\xi_y}{(1-mn)N}.$$

Because of the iterative procedure, d_i, v_i, $\underset{\sim i}{d}$ and $\underset{\sim i}{v}$ can be expressed by the initial displacements d_0, $\underset{\sim 0}{d}$, and velocities v_0, $\underset{\sim 0}{v}$ in the following form:

$$\mathbf{v}_i = \boldsymbol{\beta}^i \mathbf{v}_0 \tag{1.114}$$

The displacement and velocity of the system at any given point of time $i/N (i = [Nt])$ can be calculated by using Eq. (1.114). Considering that ith time interval is arbitrarily chosen, the complete solution is as follows:

$$x_i(t) = \mathbf{A}_2 \boldsymbol{\beta}^{[Nt]} \mathbf{v}_0 \tag{1.115}$$

$$y_i(t) = \mathbf{B}_2 \boldsymbol{\beta}^{[Nt]} \mathbf{v}_0 \tag{1.116}$$

For mastering the accuracy and efficiency of the PL method for computation of the inertial coupling system with damping term (ICSDT), the numerical results of displacements in x- and y- directions are shown in Fig. 1.11. Just like the numerical result of the ICSUT, the data points obtained by the PL method in time step 0.01 s are overlapped with that of the RK4 method in time step 0.001 s. On the contrary, one can see that the RK4 method in time step 0.01 s reduces the calculation accuracy of the ICSDT. Therefore, in the same time step, the numerical result of PL method for solving the ICSDT is more accurate than the RK4 method. The reason is that the solution of every piecewise system obtained with the PL method is continuous, unlike the RK4 method missing the physical characteristics of original system. In speaking of vibrating characteristics, the amplitude of the inertial coupling system is damping with time, and value of amplitudes will be stabilized at zero. Table 1.9 shows comparison of the numerical result for solving ICSDT with the two methods. One can see the numerical value with the PL method is approached to that of the RK4 method, although the time step of the former method is tenfold as the latter. Table 1.10 shows the CPU time for solving ICSDT in time history 50s. Comparing with Table 1.7, the CPU time of ICSDT needs longer than that of ICSUT as the existence of the damping terms. This implies that the PL method is more efficient for computation of dynamic systems than the ICSDT in the same time history.

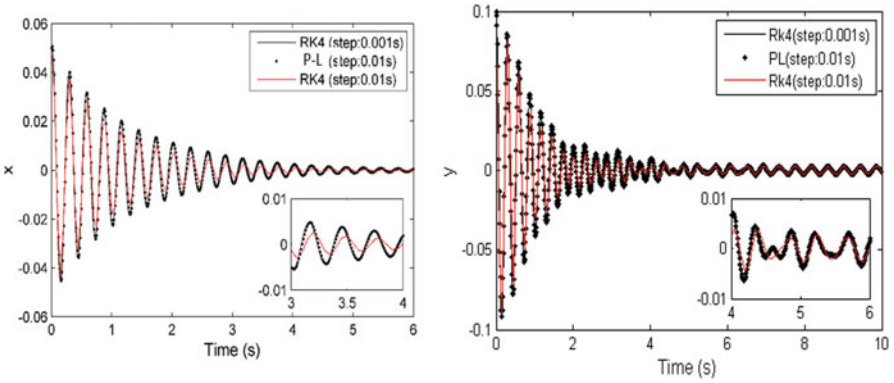


Fig. 1.11 Comparison of the numerical results with the RK4 method and the PL method for Eq. (1.103) ($\xi_x = 1.36$, $\xi_y = 1.52$, $x_0 = 0.05$, $\dot{x}_0 = 0.2$, $y_0 = 0.1$, $\dot{y}_0 = 0.$)

Table 1.9 Comparison of the numerical result for solving ICSDT

Time (s)	RK4 (Step:0.01)	RK4 (Step:0.001)	P-L (Step:0.01)
0.5	−0.008058698944357	−0.00896109624079123	−0.0089467489044681
1.0	−0.0208986465721304	−0.0224128402743939	−0.0224768947658039
1.5	0.0040899042076537	0.00538415167707660	0.0053284894060684
2.0	0.0096873942118691	0.0106810776987243	0.0106765545344556
2.5	−0.0021873664629323	−0.00373881916250538	−0.0037684737489403
3.0	−0.0035768493032044	−0.00458828573192142	−0.0045982333947589

Table 1.10 Comparison of the CPU time computed for solving ICSDT in time history 50s

Numerical method	Time step (s)	Iterations	CPU times
RK4	0.001	50,000	4.9674
RK4	0.01	5000	0.4565
PL	0.001	50,000	2.9531
PL	0.01	5000	0.3281

1.4.3 Forced and Damped Inertial Coupling System

In practical engineering, external forces always may be existed in the damped inertia coupling system. Suppose that the dynamic system is acted with a linear cosine force $F\cos\omega t$ related to time t, therefore, the dynamic equation in the following specific system is undertaken:

$$\begin{cases} m_1\ddot{x} + f_x\dot{x} + k_x x + m_2\ddot{y} = F\ \cos(\omega t), \\ m_3\ddot{y} + f_y\dot{y} + k_y y + m_4\ddot{x} = 0. \end{cases} \tag{1.117}$$

Simplification of Eq. (1.117) can be written by:

$$\begin{cases} \ddot{x} + \xi_x \dot{x} + \omega_x^2 x + m\ddot{y} = A\ \cos(\omega t), \\ \ddot{y} + \xi_y \dot{y} + \omega_y^2 y + n\ddot{x} = 0. \end{cases} \tag{1.118}$$

where $\omega_x = \sqrt{k_x/m_1}, \omega_y = \sqrt{k_y/m_3}, m = m_2/m_1, n = m_4/m_3, \xi_x = f_x/m_1,\ \xi_y = f_y/m_3, A = F/m_1$.

In this case, the piecewise constant system corresponding to that governed by Eq. (1.118) can be constructed by replacing damped terms $\dot{x}(t)$ and $\dot{y}(t)$ and external force $A\cos(\omega t)$ with piecewise constant function over an arbitrary time interval $i/N \le t < (i+1)/N$. The corresponding dynamics equation of the system is expressible in the following form:

$$\begin{cases} \ddot{x}_i + m\ddot{y}_i + \omega_x^2 x_i = -\xi_x \dot{x}_i\left(\frac{i}{N}\right) + A\cos\left(\omega\frac{i}{N}\right) \\ \ddot{y}_i + \omega_y^2 y_i + n\ddot{x} = -\xi_y \dot{y}_i\left(\frac{i}{N}\right). \end{cases} \tag{1.119}$$

With the Laplace transform, Eq. (1.119) can be also expressed as plural form:

$$\begin{cases} \left(s^2+\omega_x^2\right)X_i(s) + ms^2 Y_i(s) = -\frac{\xi_x}{s}\dot{x}_i\left(\frac{i}{N}\right) + \frac{A}{s}\cos\left(\omega\frac{i}{N}\right) + sx_i\left(\frac{i}{N}\right) + \dot{x}_i\left(\frac{i}{N}\right) + msy_i\left(\frac{i}{N}\right) + m\dot{y}_i\left(\frac{i}{N}\right) \\ \left(s^2+\omega_y^2\right)Y_i(s) + ns^2 X_i(s) = -\frac{\xi_y}{s}\dot{y}_i\left(\frac{i}{N}\right) + nsx_i\left(\frac{i}{N}\right) + n\dot{x}_i\left(\frac{i}{N}\right) + sy_i\left(\frac{i}{N}\right) + \dot{y}_i\left(\frac{i}{N}\right). \end{cases} \tag{1.120}$$

Further separating variables, X_i and Y_i can be obtained by:

$$\begin{aligned} X_i(s) = {} & \frac{s^3 - s\sigma\omega_y^2}{\upsilon(s)}x_i\left(\frac{i}{N}\right) + \frac{s^3 + \sigma\xi_x s^2 - \sigma\omega_y^2 s + \sigma\xi_x\omega_y^2}{s\upsilon(s)}\dot{x}_i\left(\frac{i}{N}\right) - \frac{\sigma m\omega_y^2 s}{\upsilon(s)}y_i\left(\frac{i}{N}\right) \\ & - \frac{\sigma m\xi_y s + \sigma m\omega_y^2}{\upsilon(s)}\dot{y}\left(\frac{i}{N}\right) - \frac{\sigma As^2 + \sigma A\omega_y^2}{s\upsilon(s)}\cos\left(\omega\frac{i}{N}\right) \end{aligned} \tag{1.121}$$

$$\begin{aligned} Y_i(s) = {} & -\frac{n\omega_x^2 s}{\upsilon(s)}x_i\left(\frac{i}{N}\right) - \frac{n\xi_x s + n\omega_x^2}{\upsilon(s)}\dot{x}_i\left(\frac{i}{N}\right) + \frac{(mn-1)s^3 - \omega_x^2 s}{\upsilon(s)}y_i\left(\frac{i}{N}\right) \\ & + \frac{(mn-1)s^3 + \xi_y s^2 - \omega_x^2 s + \xi_y\omega_x^2}{s\upsilon(s)}\dot{y}_i\left(\frac{i}{N}\right) + \frac{Ans}{\upsilon(s)}\cos\left(\omega\frac{i}{N}\right) \end{aligned} \tag{1.122}$$

where $\upsilon(s) = s^4 - \sigma\left(\omega_x^2 + \omega_y^2\right)s^2 - \sigma\omega_x^2\omega_y^2,\ \ \sigma = (mn-1)^{-1}$.

With the Theorem 1.5 in Eq. (1.10), substituting i with $[Nt]$, displacement responses $x_i(t)$ and $y_i(t)$ in arbitrary time interval $i/N \le t < (i+1)/N$ can be expressed by:

$$x_i = \mathbf{A_3}\mathbf{v}_i - \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{\sigma s_i^2 + \sigma\omega_y^2}{s_i - s_j}e^{s_i\left(t - \frac{[Nt]}{N}\right)}A\ \cos\left(\omega\frac{[Nt]}{N}\right) \tag{1.123}$$

$$y_i = \mathbf{B_3}\mathbf{v}_i - \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{ns_i}{s_i - s_j}e^{s_i\left(t - \frac{[Nt]}{N}\right)}A\ \cos\left(\omega\frac{[Nt]}{N}\right) \tag{1.124}$$

with:

$$\mathbf{A_3} = \begin{bmatrix} \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{s_i^3-\sigma\omega_y^2 s_i}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \\ \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{s_i^3+\sigma\xi_x s_i^2-\sigma\omega_y^2 s_i+\sigma\xi_x\omega_y^2}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \\ -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma m\omega_y^2 s_i}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \\ -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma m\xi_y s_i+\sigma m\omega_y^2}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \end{bmatrix}^{\mathrm{T}},$$

$$\mathbf{B_3} = \begin{bmatrix} -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma n\omega_x^2 s_i}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \\ -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma n\xi_x s_i+\sigma n\omega_x^2}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \\ \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{s_i^3-\sigma\omega_x^2 s_i}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \\ \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{s^3+\sigma\xi_y s_i^2-\sigma\omega_x^2 s_i+\sigma\xi_y\omega_x^2}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)} \end{bmatrix}^{\mathrm{T}}.$$

where s_j and s_k are zeros of function $\upsilon(s)$ shown as:

$$\begin{aligned} s_1 &= -\frac{\sqrt{2}}{2}\sqrt{\sigma\left(\omega_x^2+\omega_y^2\right)-\sqrt{\sigma^2\left(\omega_x^2+\omega_y^2\right)^2+4\sigma\omega_x^2\omega_y^2}}, \\ s_2 &= \frac{\sqrt{2}}{2}\sqrt{\sigma\left(\omega_x^2+\omega_y^2\right)-\sqrt{\sigma^2\left(\omega_x^2+\omega_y^2\right)^2+4\sigma\omega_x^2\omega_y^2}}, \\ s_3 &= -\frac{\sqrt{2}}{2}\sqrt{\sigma\left(\omega_x^2+\omega_y^2\right)+\sqrt{\sigma^2\left(\omega_x^2+\omega_y^2\right)^2+4\sigma\omega_x^2\omega_y^2}}, \\ s_4 &= \frac{\sqrt{2}}{2}\sqrt{\sigma\left(\omega_x^2+\omega_y^2\right)+\sqrt{\sigma^2\left(\omega_x^2+\omega_y^2\right)^2+4\sigma\omega_x^2\omega_y^2}}, \\ s_5 &= 0. \end{aligned}$$

It should be noted that $s_j - s_k = 1$ in the whole process of the computations when $j = k$.

Then, the velocities of the system in x- and y- directions can be expressed by:

$$\dot{x}_i = \frac{\partial \mathbf{A}_3}{\partial t}\mathbf{v}_i - \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{\sigma s_i^3+\sigma\omega_y^2 s_i}{s_i-s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)}A\ \cos\left(\omega\frac{[Nt]}{N}\right) \tag{1.125}$$

$$\dot{y}_i = \frac{\partial \mathbf{B_3}}{\partial t}\mathbf{v}_i - \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{ns_i^2}{s_i - s_j}e^{s_i\left(t-\frac{[Nt]}{N}\right)}A\ \cos\left(\omega\frac{[Nt]}{N}\right) \tag{1.126}$$

With the continuity condition, a recursive relation is obtained by combining Eqs. (1.123). (1.124), (1.125) and (1.126), such that:

$$\mathbf{v}_i = \boldsymbol{\gamma}\mathbf{v}_{i-1} + \mathbf{g}\,A\ \cos\left(\omega\frac{([Nt]-1)}{N}\right) \tag{1.127}$$

where $\mathbf{g} = \left[-\sum_{i=1}^{5}\prod_{j=1}^{5}\frac{\sigma s_i^2+\sigma\omega_y^2}{s_i-s_j}e^{\frac{s_i}{N}}, -\sum_{i=1}^{5}\prod_{j=1}^{5}\frac{\sigma s_i^3+\sigma\omega_y^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}}, -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{ns_i}{s_i-s_j}e^{\frac{s_i}{N}}, \ -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{ns_i^2}{s_i-s_j}e^{\frac{s_i}{N}}\right]^{\mathrm{T}}$, $\boldsymbol{\gamma} = [\gamma_{ij}]_{4\times 4}$, and the elements of matrixes $\boldsymbol{\gamma}$ are shown as:

$$\gamma_{11} = \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{s_i^3-\sigma\omega_y^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{12} = \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{s_i^3+\sigma\xi_x s_i^2-\sigma\omega_y^2 s_i+\sigma\xi_x\omega_y^2}{s_i-s_j}e^{\frac{s_i}{N}},$$
$$\gamma_{13} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma m\omega_y^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{14} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma m\xi_y s_i+\sigma m\omega_y^2}{s_i-s_j}e^{\frac{s_i}{N}},$$
$$\gamma_{21} = \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{s_i^4-\sigma\omega_y^2 s_i^2}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{22} = \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{s_i^4+\sigma\xi_x s_i^3-\sigma\omega_y^2 s_i^2+\sigma\xi_x\omega_y^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},$$
$$\gamma_{23} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma m\omega_y^2 s_i^2}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{24} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma m\xi_y s_i^2+\sigma m\omega_y^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},$$
$$\gamma_{31} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma n\omega_x^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{32} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma n\xi_x s_i+\sigma n\omega_x^2}{s_i-s_j}e^{\frac{s_i}{N}},$$
$$\gamma_{33} = \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{s_i^3-\sigma\omega_x^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{34} = \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{s^3+\sigma\xi_y s_i^2-\sigma\omega_x^2 s_i+\sigma\xi_y\omega_x^2}{s_i-s_j}e^{\frac{s_i}{N}},$$
$$\gamma_{41} = -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma n\omega_x^2 s_i^2}{s_i-s_j}e^{\frac{s_i}{N}},\ -\sum_{i=1}^{4}\prod_{j=1}^{4}\frac{\sigma n\xi_x s_i^2+\sigma n\omega_x^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}},$$

$$\gamma_{33} = \sum_{i=1}^{5}\prod_{j=1}^{5}\frac{s_i^4-\sigma\omega_x^2 s_i^2}{s_i-s_j}e^{\frac{s_i}{N}},\ \gamma_{34} = \sum_{i=1}^{4}\prod_{j=1}^{4}\frac{s^4+\sigma\xi_y s_i^3-\sigma\omega_x^2 s_i^2+\sigma\xi_y\omega_x^2 s_i}{s_i-s_j}e^{\frac{s_i}{N}}.$$

With the recursive relation in Eq. (1.127), the recursive relation with respect to initial conditions is obtained, such that:

$$\mathbf{v}_i = \boldsymbol{\gamma}^i\mathbf{v}_0 + \sum_{r=1}^{i}\boldsymbol{\gamma}^{r-1}\mathbf{g}\,A\ \cos\left(\frac{\omega\left([Nt]-r\right)}{N}\right) \tag{1.128}$$

The displacement responses of the system at any given point of time *i*/*N* can be calculated by using Eq. (1.128). Considering that *i*th time interval is arbitrarily chosen, thus, the complete solution is obtained by:

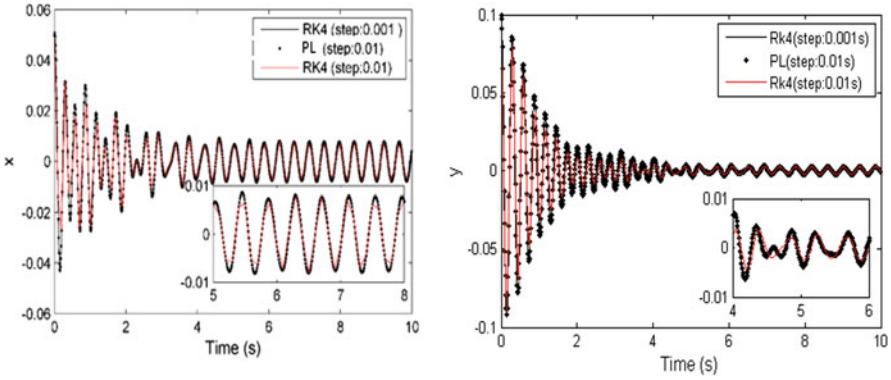


Fig. 1.12 Comparison of the numerical results with the RK4 method and the PL method for Eq. (1.117) ($\xi_x = 1.36, \xi_y = 1.52, A = 1.89, \omega = 15, x_0 = 0.05, \dot{x}_0 = 0.2, y_0 = 0.1, \dot{y}_0 = 0.$)

$$x_i = \mathbf{A}_3 \left\{ \boldsymbol{\gamma}^i \mathbf{v}_0 + \sum_{r=1}^{i} \boldsymbol{\gamma}^{r-1} \mathbf{g}\, A\, \cos\left(\frac{\omega\,([Nt]-r)}{N} \right) \right\} - \sum_{i=1}^{5} \prod_{j=1}^{5} \frac{\sigma s_i^2 + \sigma \omega_y^2}{s_i - s_j} e^{s_i \left(t - \frac{[Nt]}{N}\right)} A\, \cos\left(\omega \frac{[Nt]}{N} \right), \tag{1.129}$$

$$y_i = \mathbf{B_3} \left\{ \boldsymbol{\gamma}^i \mathbf{v}_0 + \sum_{r=1}^{i} \boldsymbol{\gamma}^{r-1} \mathbf{g}\, A\, \cos\left(\frac{\omega\,([Nt]-r)}{N} \right) \right\} - \sum_{i=1}^{4} \prod_{j=1}^{4} \frac{n s_i}{s_i - s_j} e^{s_i \left(t - \frac{[Nt]}{N}\right)} A\, \cos\left(\omega \frac{[Nt]}{N} \right), \tag{1.130}$$

In the engineering application, the damping term and the external force term are considered. Finally, the numerical computation is implemented with the PL method and the RK4 method for computation of the inertial coupling system with damping and external force term (ICSDEFT). Figure 1.12 shows the displacement responses in x and y directions with the two method. It is also evident that the RK4 method for solving dynamic characteristics of the ICSDEFT is less accurate than the PL method in same time step 0.01 s. In summary, the solution process of RK4 method, employing the slope iteration related to an arithmetic mean value, damages the physical meaning contained in the original dynamic system and causes undesirable influence on the accuracy and reliability of numerical results. And the comparison of the numerical result for solving the ICSDEFT is shown in Table 1.11. In the speaking of computational efficiency, the computation time with the PL method is also shorter than the RK4 method in the ICSDEFT, as shown in Table 1.12. Because of the existence of the damping term and the external force term, the CPU time

Table 1.11 Comparison of the numerical response in x direction computed by PL and RK4 methods

Time (s)	RK4 (Step:0.01 s)	RK4 (Step:0.001 s)	P-L (Step:0.01 s)
0.5	−0.00309940302934400	−0.00357349548568479	−0.00358645375869738
1.0	−0.0237709803733659	−0.0257734136575020	−0.02579745265768369
1.5	−0.00198463439997458	−0.00237332971024369	−0.00237746434566889
2.0	0.00928862918350417	0.0102924435591604	0.01029465738970809
2.5	0.00320229632300177	0.00380853448101097	0.00380573696026353
3.0	0.00080383120714058	0.00091033743407989	0.00091038769476465

Table 1.12 Comparison of the CPU time computed for solving the ICSDEFT in time history 50s

Numerical method	Time step (s)	Iterations	CPU times
RK4	0.001	50,000	5.8354
RK4	0.01	5000	0.5476
PL	0.001	50,000	5.1703
PL	0.01	5000	0.4985

of the ICSDEFT needed is longer than that of the ICSUT and ICSDT in the same iterations.

1.4.4 Convergence Analysis of the PL Method

Taking undamped inertial coupling system in Eq. (1.86) as a sample, the convergence of the PL method is discussed. Starting with the initial state, Eq. (1.101) and Eq. (1.102) are used in a computer program to obtain solution in each interval. The numerical results obtained by the application equations are illustrated in Fig. 1.13. As shown in the figure, the error of the numerical results from the accurate value (AV) is already small as N is given a value of 10. The numerical results are getting closer and closer to the AV as the value of N increases. The parameter N obviously acts as a factor controlling the accuracy of the numerical solution, since it is directly related to the time interval $i/N \leq t \leq (i+1)/N$. In order to have a numerical solution of high accuracy, one may choose a sufficiently large N.

With the finds for computations of the inertial coupling system, the following can be concluded:

1. When applying the RK4 method to solve the inertial coupling system, the dynamic equations are primarily rewritten into the first-order differential equations, which leads to the loss of the essential characteristics of the dynamic system. In addition, the solution process of RK4 method, employing the slope iteration related to an arithmetic mean value, damages the physical meaning contained in the original dynamic system and causes undesirable influence on the accuracy and reliability of numerical results.

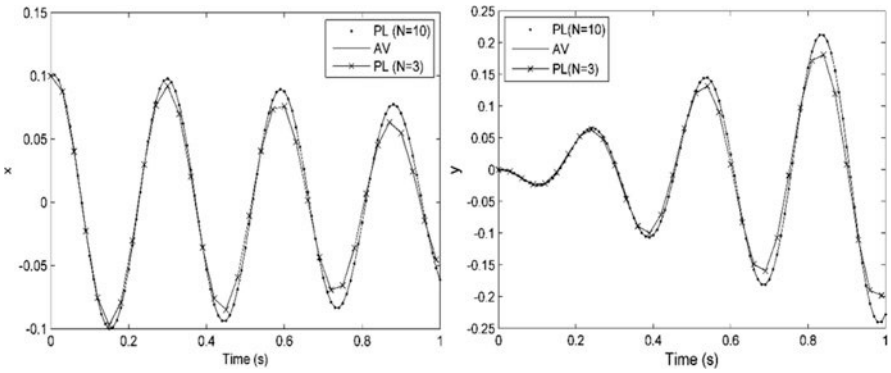


Fig. 1.13 Convergence of the solution of Eq. (1.59) ($x_0 = 0.1$, $\dot{x}_0 = 0.2$,$y_0 = 0$, $\dot{y}_0 = 0.1$.)

2. The solutions of the PL method are directly obtained from the original dynamic system and continuous over the whole-time domain considered, which can obtain more accurate solution for solving the inertial coupling system than RK4 method in the fixed time step. And by using the PL method, the semi-analytical solutions of the system are determined, which is also verified with numerical computations.
3. The PL method generates more accurate and reliable solutions in the comparison of RK4 method within shorter CPU time. The reason is that the solution of every piecewise system obtained with the PL method is continuous, unlike the RK4 method missing the physical characteristics of original system.
4. The inertial coupling system presented in this section is linear. The computation results with PL method reflect the nature of the inertial coupling system and are close to the exact solution, since the PL method maintains the original physics characteristics of dynamic system. However, for solving the nonlinear dynamic system, the solutions are sensitive to the initial conditions and system parameters. Whether the PL method is suitable to handle the nonlinear dynamic system is our next work to purse.

1.5 Diagnosing Irregularities of Nonlinear Systems

1.5.1 Nonlinear Nonautonomous System

Here, the fourth-order Runge-Kutta method is employed to determine the periodicity ratio of the hull system, as shown in Eq. (1.131). This model is applied to describe the nonlinear coupling characteristics of the pitching and rolling of the hull.

$$\begin{cases} \ddot{x}_1 + 2\mu_1\dot{x}_1 + \omega_1^2 x_1 = \alpha_1 x_1 x_2 + F\ \cos(\omega t) \\ \ddot{x}_2 + 2\mu_2\dot{x}_2 + \omega_2^2 x_2 = \alpha_2 x_1^2 + F\ \cos(\omega t) \end{cases} \tag{1.131}$$

Consider F and ω as control parameters, and the initial value of the system is assumed to be $x_1(0) = 0.1$, $\dot{x}_1(0) = 0.2$, $x_2(0) = 0.3$, and $\dot{x}_2(0) = 0.4$. The other parameters of the system are defined by $\mu_1 = 0.1$, $\mu_2 = 0.1$, $\alpha_1 = 0.5$, $\omega_1 = 5.5$, and $\omega_2 = 5.5$. Figure 1.14 shows the periodicity ratio when the parameters of the system satisfy that $\alpha_2 = 1.0$, $\alpha_2 = 1.5$, $\alpha_2 = 2.0$, and $\alpha_2 = 2.5$. The red region in this figure represents that the dynamic characteristics of the system are periodic, i.e., $\gamma = 1$; the blue region denotes that the dynamic characteristics of the system are chaotic, i.e., $\gamma = 0$; and the other color region signifies that the dynamic characteristics of the system are neither periodic nor chaotic, i.e., $0 < \gamma < 1$. The figure reveals the dynamic behavior of the system with the change of system parameters. As shown in this figure, when ω is located in region of $(0, 2]$, the vibration behavior of the system is transferred from chaos to periodicity with the increase of the external excitations; when ω is located in region of $(5.2, 5.8]$, the probability of nonperiodic vibration of the system is increased with the increase of coupling coefficient α_2; when ω is located in the other region, the vibration behavior of the system is periodic. It can be seen that the system is super near resonance, near resonance, sharp resonance, and far resonance when $\omega \in (0, 2]$, $\omega \in (2, 5.2]$, $\omega \in (5.2, 5.8]$, and $\omega \in (5.8, 10]$, respectively. It can be concluded that the dynamic characteristics are chaotic when the system is sharp resonance or super near resonance.

As shown in Fig. 1.14d, the system parameters are located in the blue region when $\omega = 0.2$, $\alpha_2 = 2.5$, and $F = 0.9$, and the vibration behavior is chaotic because of $\gamma = 0$; the system parameters are located in the red region when $\omega = 8$, $\alpha_2 = 2.5$, and $F = 20$, and the vibration behavior is chaotic because of $\gamma = 1$. The vibration characteristics of the hull are shown in Fig. 1.15 when $\omega = 0.2$, $\alpha_2 = 2.5$, and $F = 0.9$. Figure 1.15 (a) and (c) follow that the trajectory of the system is periodic in x_1-direction; and the trajectory is chaotic in x_2-direction as shown in Fig. 1.15b, d. Therefore, the vibration characteristics of the system are chaotic in the condition of $\omega = 0.2$, $\alpha_2 = 2.5$ and $F = 0.9$. As a result, the vibration characteristics of the system are chaotic in the condition of $\omega = 0.2$, $\alpha_2 = 2.5$ and $F = 0.9$. Figure 1.16 shows the vibration characteristics of the hull system considering $\omega = 8$, $\alpha_2 = 2.5$, and $F = 20$. It can be seen that the trajectory of the system is periodic in x_1- and x_2- directions. And so the vibration characteristics of the system are periodic in the condition of $\omega = 8$,$\alpha_2 = 2.5$, and $F = 0.9$. According to the analysis above, if the vibration characteristics in all dimensionality are periodic, the system is periodic or deterministic. Conversely, it is an uncertain system.

1.5.2 Nonlinear Autonomous System

Consider the famous Rössler system, shown in Eq. (1.132):

$$\begin{cases} \dot{x}_1 = -x_2 - x_3 \\ \dot{x}_2 = x_1 + a x_2 \\ \dot{x}_3 = b + (x_1 - c)\, x_3 \end{cases} \tag{1.132}$$

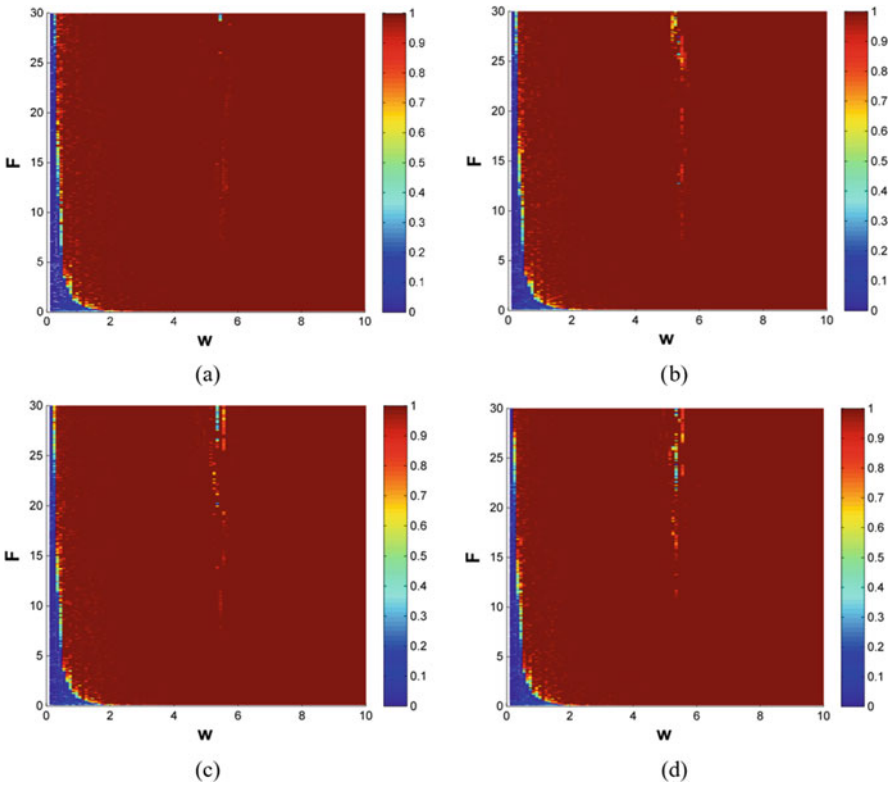


Fig. 1.14 Periodicity of hull dynamic system. (**a**) $\alpha_2 = 1.0$, (**b**) $\alpha_2 = 1.5$, (**c**) $\alpha_2 = 2.0$, (**d**) $\alpha_2 = 2.5$

Assuming b and c to be control parameters, the periodicity of Rössler system can be determined with the periodicity-ratio method in Sect. 1.2.4. Considering the initial value of the system as $x_1(0) = 0.1$, $x_2(0) = 0.1$, and $x_3(0) = 0.3$, Fig. 1.17 shows the periodicity of Rössler system with different values of a. The red region in this figure represents that the dynamic characteristics of the system are periodic, i.e., $\gamma = 1$; the blue region denotes that the dynamic characteristics of the system are chaotic, i.e., $\gamma = 0$; and the other color region signifies that the dynamic characteristics of the system are neither periodic nor chaotic, i.e., $0 < \gamma < 1$. The figure can intuitively determine the process of the dynamic behavior changed with parameters of the system. Because of the nonlinear characteristics of Rössler system, the dynamic characteristics of the system are sensitive to the system parameters.

As shown in Fig. 1.17d, the parameters are located in the red region when $a = 0.20$, $b = 0.20$, and $c = 3.50$; in this case, the vibration characteristics of the system are periodic; when $a = 0.20$, $b = 0.20$, and $c = 5.50$, the parameters are located in the blue region; in this case, the vibration characteristics of the system

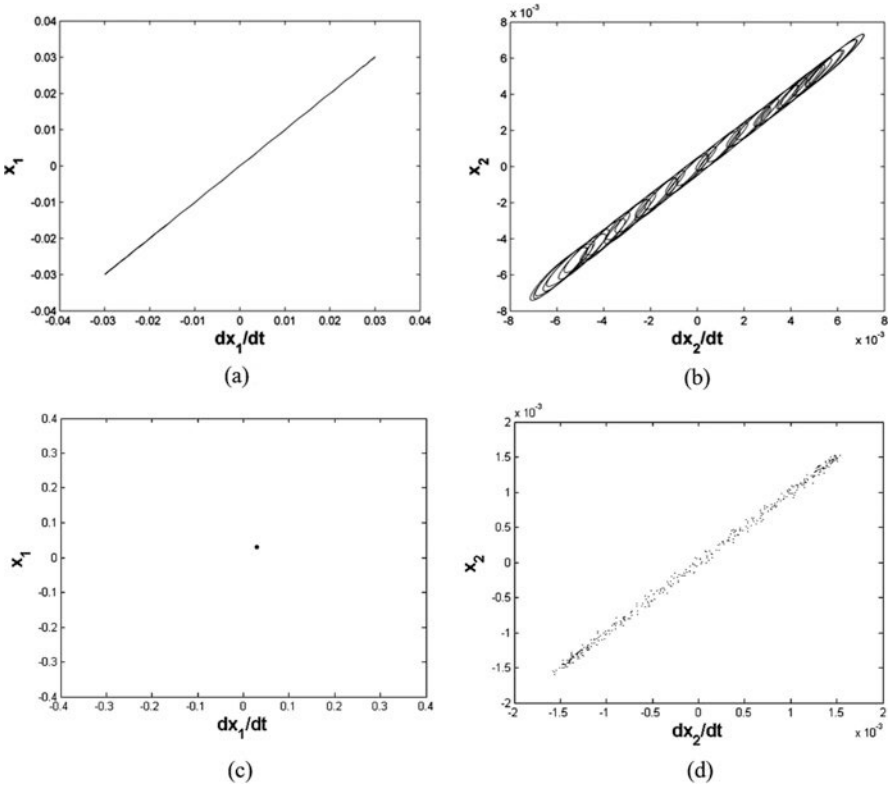


Fig. 1.15 Vibration characteristics of hull dynamic system for $\omega = 0.2\ \alpha_2 = 2.5\ F = 0.9$. (**a**) Phase diagram in x_1- direction, (**b**) phase diagram in x_2- direction, (**c**) Poincaré map in x_1- direction, and (**d**) Poincaré map in x_2- direction

are chaotic. Figure 1.18 shows the periodic vibration of the Rössler system. It can be found that the phase trajectory is periodic when $a = 0.20$, $b = 0.20$, and $c = 3.50$, as shown in Fig. 1.18e, and the number of the visible points in the Poincaré sections of x_1–x_2 and x_2–x_3 is 2, respectively. However, the phase trajectory is chaotic when $a = 0.20$, $b = 0.20$, and $c = 5.50$; therefore, the number of the visible points in the sections of x_1–x_2 and x_2–x_3 is infinite, as shown in Fig. 1.19.

Through the discussions of theoretical research and numerical analysis, it can be found that periodicity ratio is an effective tool to identify the dynamic behavior of high-dimensional nonlinear systems. The conclusions of the study on the periodicity ratio are as follows:

1. If the dynamic responses of the nonlinear system are periodic, the phase points in the Poincaré sections are overlapping. In this case, the value of the periodicity ratio is equal to 1.

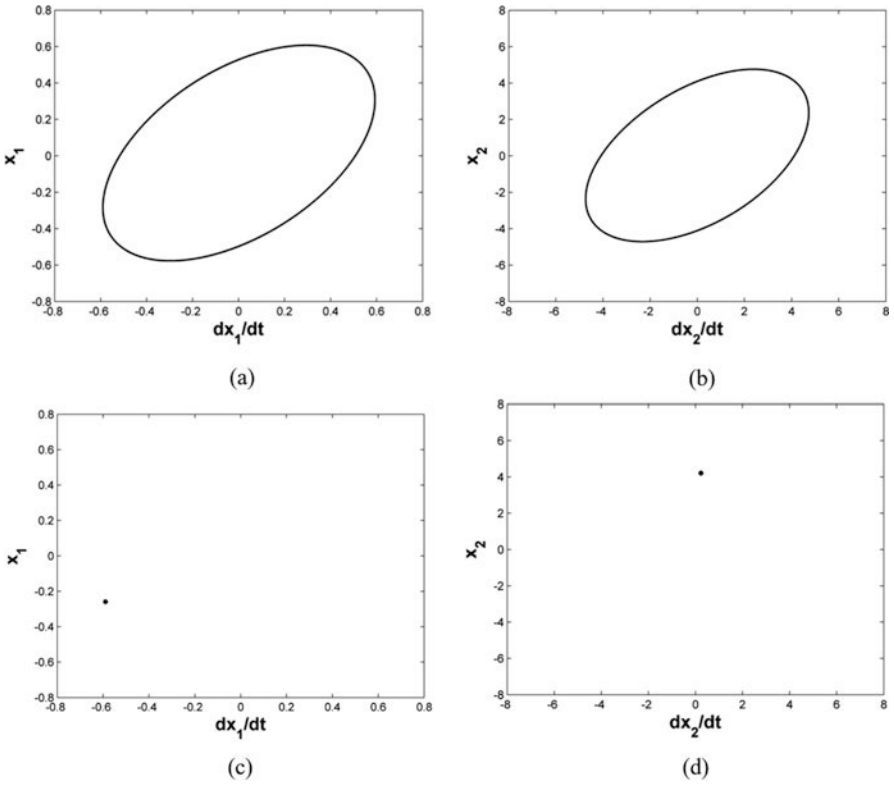


Fig. 1.16 Vibration characteristics of hull dynamic system for $\omega = 8$ $\alpha_2 = 2.5$ $F = 20$. (**a**) Phase diagram in x_1- direction, (**b**) phase diagram in x_2- direction, (**c**) Poincaré map in x_1- direction, and (**d**) Poincaré map in x_2- direction

2. If the dynamic responses of the nonlinear system are chaotic, the phase points in the Poincaré sections are nonoverlapping. In this case, the value of the periodicity ratio is equal to 0.
3. For a nonlinear dynamic system, there may exist an infinite number of nonperiodic solutions, which are neither periodic nor chaotic. For these nonperiodic solutions, the corresponding periodicity-ratio values are in the range of $0 < \gamma < 1$. The larger value of the periodicity ratio represents the dynamic characteristics closed to periodic motion, and the smaller value of the periodicity ratio represents the dynamic characteristics closed to chaotic motion.

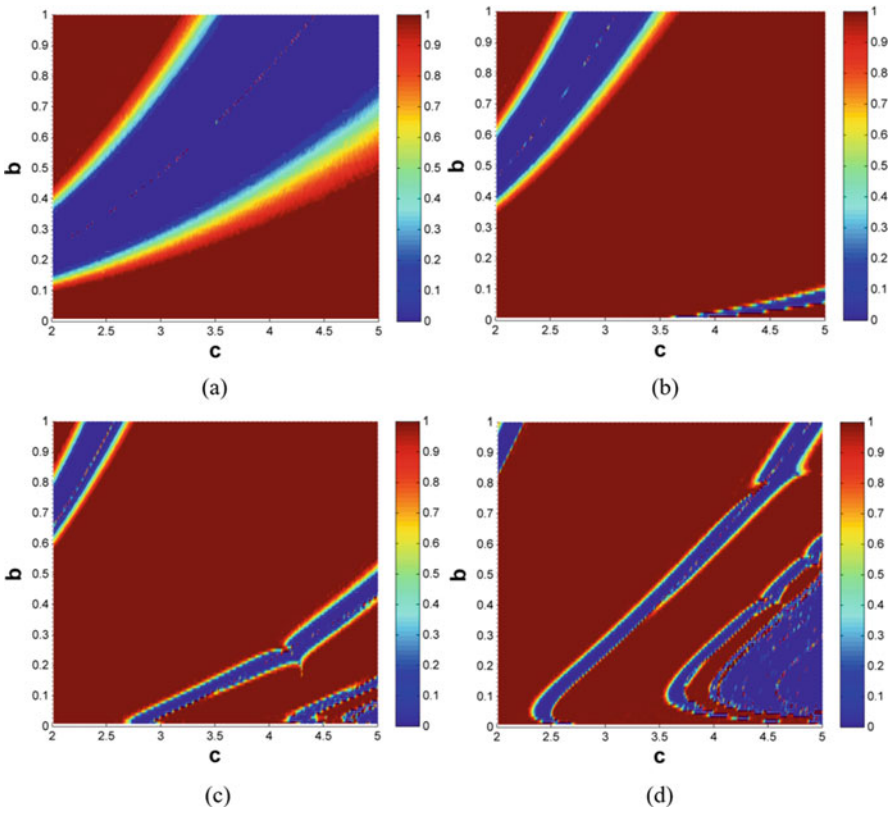


Fig. 1.17 Periodicity of Rössler dynamic system. (**a**) $a = 0.05$, (**b**) $a = 0.10$, (**c**) $a = 0.15$, (**d**) $a = 0.20$

1.6 Conclusion

Using the traditional RK method for solving the dynamic system, high order differential equations of dynamic system are usually descended into multiple first-order differential equations. In this way, crucial dynamic property of the original system may be neglected or simplified, which leads to the computational error when searching solution for dynamic model encountered in actual engineering. However, solving differential equations above with the PL method proposed in the chapter, the semi-analytical solution of systems is obtained directly through the derivation of equation. Therefore, the physical properties of systems are well preserved. In addition, the whole-time interval is divided into many tiny intervals by the PL method, and the solution is continuous on each interval. In this case, the accuracy of numerical solution of the PL method is better than the RK method, which is demonstrated by numerical analysis. In addition, the computed efficiency of the PL and RK methods are compared through statistics of operation time of CPU during

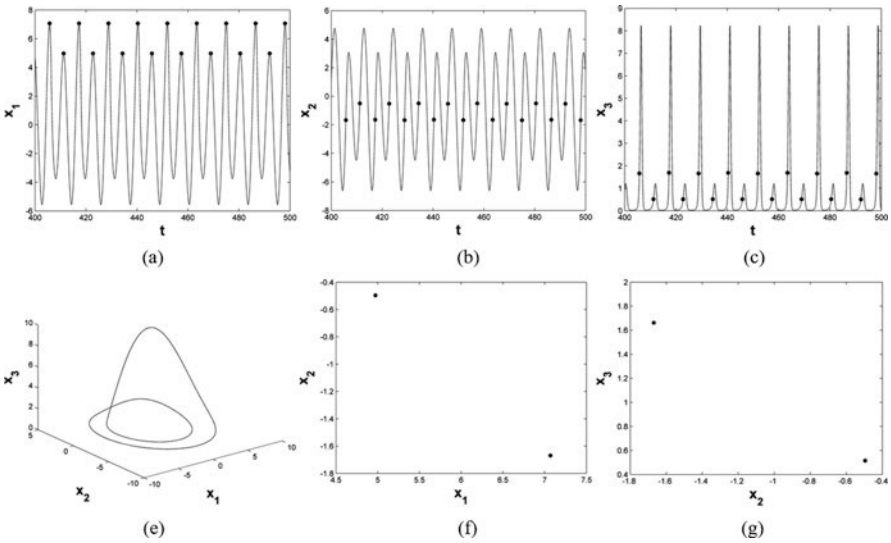


Fig. 1.18 Vibration characteristics of Rössler dynamic system for $a = 0.20$, $b = 0.20$, and $c = 3.50$

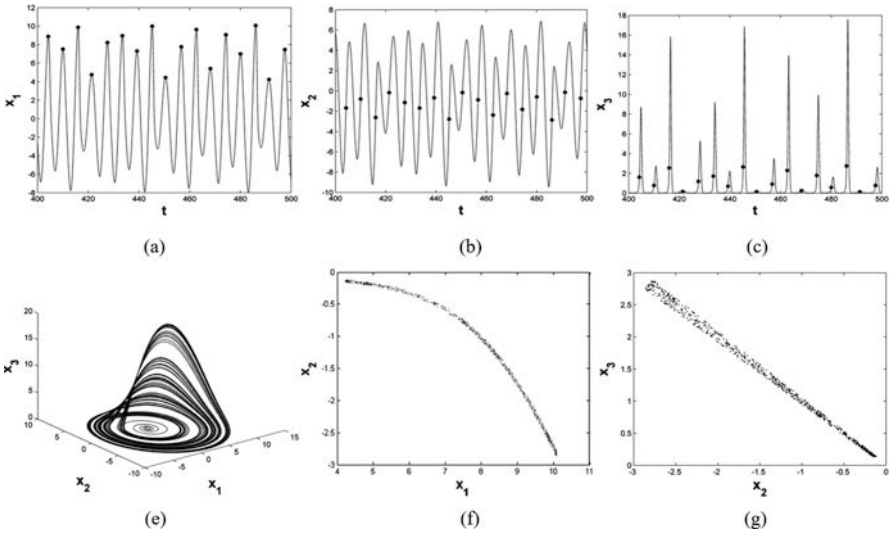


Fig. 1.19 Vibration characteristics of Rössler dynamic system for $a = 0.20$, $b = 0.20$, and $c = 5.50$

computing process. The CPU time taken by the PL method is less than the RK method in the same time steps, which indicates that the PL method is more efficient than the RK method. The main reason is that value continuity of two adjacent truncation points is maintained by using the PL method, which simplifies the solving process and saves the calculated time, while the RK method combines iterative

and averaged slope to search solutions, which makes the solution process more complicated. Therefore, the numerical results of the PL method reflect the essence of dynamic system, since the PL method keeps the physical characteristics of dynamic system. The precision of the solution obtained by the PL method is related to the value of N. This article is an exploratory study for implementation of the PL method in coupling systems. Therefore, the classic two-degree freedom systems are considered, and whether the PL method is suitable to solve problems of multi-degree freedom system should be further verified in our next work. Based on the PL method, the periodicity ratio is proposed to explore the dynamic characteristics of the nonlinear system. If the dynamic responses of the nonlinear system are periodic, the phase points in the Poincaré sections are overlapping. In this case, the value of the periodicity ratio is equal to 1. If the dynamic responses of the nonlinear system are chaotic, the phase points in the Poincaré sections are nonoverlapping. In this case, the value of the periodicity ratio is equal to 0. For a nonlinear dynamic system, there may exist an infinite number of nonperiodic solutions, which are neither periodic nor chaotic. For these nonperiodic solutions, the corresponding periodicity-ratio values are in the range of $0 < \gamma < 1$. The larger value of the periodicity ratio represents the dynamic characteristics closed to periodic motion, and the smaller value of the periodicity ratio represents the dynamic characteristics closed to chaotic motion.

References

1. Abukhaled MI, Allen EJ. A class of second-order Runge-Kutta methods for numerical solution of stochastic differential equations. Stochastic Analysis & Applications. 1998, 16(6):977–991.
2. Cai ZF, Kou KI. Laplace transform: a new approach in solving linear quaternion differential equations. Mathematical Methods in the Applied Sciences. 2018, 41(11):4033–48.
3. Chen Z, Qiu Z, Li J. Two-derivative Runge-Kutta-Nyström methods for second-order ordinary differential equations. 2015, 70(4):897–927.
4. Butcher, JC. Numerical Methods for Ordinary Differential Equations. Ussr Computational Mathematics & Mathematical Physics. 2010, 53:153–170.
5. Dai LM. Nonlinear Dynamics of Piecewise Constant Systems and Implementation of Piecewise Constant Arguments: World Scientific; 2008.
6. Iserles A, Ramaswami G and Mathematics M. Runge-Kutta methods for quadratic ordinary differential equations. Bit Numerical Mathematics. 1998, 38(2):315–46.
7. Hussain, Kasim A, Ismail F. Fourth-Order Improved Runge-Kutta Method for Directly Solving Special Third-Order Ordinary Differential Equations. Iranian journal of ence and technology. transaction a, 2017,41:429–437.
8. Ji XY, Zhou J. Solving High-Order Uncertain Differential Equations via Runge-Kutta Method. IEEE Transactions on Fuzzy Systems 2017:1–11.
9. Kanagarajan K, Suresh R, Mathematics A. Runge-Kutta method for solving fuzzy differential equations under generalized differentiability. 2016, 14(14):1–12.
10. Forcrand PD, Jäger B. Taylor expansion and the Cauchy Residue Theorem for finite-density QCD. The 36th Annual International Symposium on Lattice Field Theory. 2018.
11. Zingg DW, Chisholm TT. Runge-Kutta methods for linear ordinary differential equations. Applied Numerical Mathematics 1999, 31:227–238.

12. Zhang CY, Chen L. A symplectic partitioned Runge-Kutta method using the eighth-order NAD operator for solving the 2D elastic wave equation. Journal of Seismic Exploration. 2015, 24:205–230.
13. State Locus Drawing of RLC Circuit Based on Ode45 of MATLAB. China Science and Technology Information, 2008.
14. Immler, Fabian, Traut C. The Flow of ODEs: Formalization of Variational Equation and Poincaré Map. Journal of Automated Reasoning, 2019, 62(2):215–236.
15. Hartman, Philip. Ordinary differential equations. Mathematics of Computation.1982, 20:82–122.
16. Dai LM, Xia DD, Chen CP, An algorithm for diagnosing nonlinear characteristics of dynamic systems with the integrated periodicity ratio and lyapunov exponent methods, Communications in Nonlinear Science and Numerical Simulation, 2019, 73: 92–109.
17. Dai LM, Wang X, Chen C. Accuracy and Reliability of Piecewise-Constant Method in Studying the Responses of Nonlinear Dynamic Systems. Journal of Computational and Nonlinear Dynamics. 2015, 10(2):021009–10.
18. Dai LM, Singh MC, Structures. An analytical and numerical method for solving linear and nonlinear vibration problems. International Journal of Solids & Structures. 1997, 34(21):2709–2731.
19. Lu X, Sun J, Li G, Wang Q, Zhang D. Dynamic analysis of vibration stability in tandem cold rolling mill. Journal of Materials Processing Technology. 2019, 272:47–57.
20. Reyhanoglu M, van der Schaft A, Mcclamroch NH, Kolmanovsky I. Dynamics and control of a class of underactuated mechanical systems. IEEE Transactions on Automatic Control, 2017, 44(9):1663–1671.
21. Bryant P, Brown R, Abarbanel H. Lyapunov Exponents from Observed Time Series. Phys. rev. lett. 1990, 65(13):1523–1526.
22. Zhang W, Ye M. Local and global bifurcations of valve mechanism. Nonlinear Dynamics. 1994, 6(3):301–316.
23. Wolf A, Swift JB, Swinney HL, Vastano JA. Determining Lyapounov exponents from a time series. Springer Berlin Heidelberg. 1985, 16(3):285–317.
24. Jung C. Poincare map for scattering states. Journal of Physics A General Physics 19.8(1999):1345.
25. Dai, LM, Singh MC. Periodic, quasiperiodic and chaotic behavior of a driven Froude pendulum. International Journal of Nonlinear Mechanics. 1998, 33(6):947–965.
26. Ma H, Ho DWC, Lai YC, Lin W. Detection meeting control: Unstable steady states in high-dimensional nonlinear dynamical systems. Physical Review E. 2015.92.
27. Zhang W, Wang FX, Zu JW. Local bifurcations and codimension-3 degenerate bifurcations of a quintic nonlinear beam under parametric excitation. Chaos, Solitons & Fractals. 2005, 24(4):977–98.
28. Fang P, Dai LM, Hou YJ, Du MJ, Wang LY, Xi ZJ. Numerical Computation for the Inertial Coupling Vibration System Using PL Method. Journal of Vibration Engineering & Technologies, 2019, 7(2):139–148.
29. Dai L, Assessment of Solutions to Vibration Problems Involving Piecewise Constant Exertions, Symposium on Dynamics, Acoustic and Simulations (DAS2000) in ASME IMECE 2000, Orlando, 2000.
30. Arendt W, Batty CJ, Hieber M, Neubrander F. Vector-Valued Laplace Transforms and Cauchy Problems. Israel Journal of Mathematics. 1987, 59(3):327–352.
31. Fang P, Dai L, Hou Y, Du M, Wang L. The Study of Identification Method for Dynamic Behavior of High-Dimensional Nonlinear System. Shock and vibration. 2019.
32. Fang P, Hou YJ, Zhang LP, Du MJ, Zhang MY. Synchronous behavior of a rotor-pendulum system. Acta Physica Sinica. 2016, 65.

Chapter 2
Novel Predictor-Corrector Formulations for Solving Nonlinear Initial Value Problems

Vahid Ahmadi Kalkhorani and Mohammad Mohammadi Aghdam

Abstract In this chapter, a novel predictor-corrector formulation is introduced based on the Bézier curve interpolation. The proposed formulation can be employed to determine numerical integration and solve nonlinear initial value problems with low computational cost while having larger stability region. The stability region of the method is examined in solving benchmark initial value problems, which showed a larger stability region both on the imaginary and real axis compared to some other well-known methods. The results of applying this formulation in solving Duffing and Van der Pol equations as numerical experiments showed that the proposed method is more stable and accurate for a wide range of step size compared to other well-known methods. Considering performance of the method in terms of accuracy, stability, and cost, it is expected that the method to be used is more complicated and applied nonlinear science and engineering problems.

Keywords Predictor-corrector · Multistep method · Bézier curve · Nonlinear initial value problem

2.1 Introduction

In this chapter, the main focus is on presenting a new predictor-corrector technique to solve a single or a system of nonlinear initial value problems (IVPs). Performance

F. Pan (✉)
School of Mechanical Engineering, Southwest Petroleum University, Chengdu, China

Petroleum Engineering School, Southwest Petroleum University, Chengdu, China

Engineering and the Applied Sciences, University of Regina, Regina, Canada

W. Kexin
School of Mechanical Engineering, Southwest Petroleum University, Chengdu, China

D. Liming · W. Luyao
Engineering and the Applied Sciences, University of Regina, Regina, Canada

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_2

of the method is examined in terms of accuracy, stability, and computational cost for various applied engineering problems.

A general form of a single IVP of order n can be considered as:

$$y^{(n)} = F\left(x, y, y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}\right) \tag{2.1}$$

in which x is the independent variable, $y^{(i)}$ is the i^{th} derivation of y with respect to x, and F is any linear or nonlinear function of independent and dependent variables and their derivatives.

In many cases, there is no analytical solution for differential equations, in particular for nonlinear systems. Thus, different numerical methods including one-step [4] and multistep methods [8, 10] have been developed to solve these equations. One-step methods such as Euler technique require information about the previous point only, while multistep methods such as Adams method [9] require information in more than one previous point. There are many problems in applied engineering sciences, which require solving a system of initial value problems. In addition, in order to solve a set of differential equations of order n, one may convert it into a system of n first-order differential equations. The differential equation of order n (2.1) can be converted to a set of n first-order differential equations (2.2) with the initial values (2.3).

$$\begin{cases} \frac{dy_1}{dx} = f_1\left(x, y_1, y_2, \ldots, y_k\right) \\ \frac{dy_2}{dx} = f_2\left(x, y_1, y_2, \ldots, y_k\right) \\ \qquad\qquad \vdots \\ \frac{dy_n}{dx} = f_n\left(x, y_1, y_2, \ldots, y_k\right) \end{cases} \tag{2.2}$$

$$y_1(a) = Y_1, y_2(a) = Y_2, \ldots, y_n(a) = Y_n \tag{2.3}$$

By integrating the differential Eqs. (2.1, 2.2, and 2.3), the system of differential equations is converted to a numerical integral problem.

The integral form of (2.2) is:

$$\begin{cases} y_1\left(x_{i+1}\right) = y_1\left(x_i\right) + \int_{x_i}^{x_{(i+1)}} f_1 dx \\ y_2\left(x_{i+1}\right) = y_2\left(x_n\right) + \int_{x_i}^{x_{(i+1)}} f_2 dx \\ \qquad\qquad \vdots \\ y_n\left(x_{i+1}\right) = y_n\left(x_n\right) + \int_{x_i}^{x_{(i+1)}} f_n dx \end{cases} \tag{2.4}$$

Since the integrands in most practical problems include nonlinear functions, various methods for calculating the integrals are proposed. Most of these methods are based on approximating the integrand with simple polynomials that can be

easily integrated. In this chapter, Bézier curves, which are weighted combination of Bernstein polynomials, are employed to determine integrals in (2.4).

2.2 Bezier Curves

One of the efficient methods to solve IVPs is implementing Bezier curves for approximating the integrands [1]. Using multipoint information, the Bezier provides a curve, which passes through the beginning and end points of the interval while not necessarily crossing the midpoints. In fact, the Bezier curve forms a set of parametric points, whose weight functions are obtained based on the Bernstein polynomials. The general form of the Bezier curve for $m + 1$ given points is:

$$\begin{cases} x_{(u)} = \sum_{i=0}^{m} B_i x_i \\ f_{(u)} = \sum_{i=0}^{m} \mathrm{B_i} f_i \end{cases} \tag{2.5}$$

in which B_i is the Bernstein basis polynomials of degree m defined as:

$$B_i = \binom{m}{i} (1-u)^{m-i} u^i \quad 0 \le u \le 1 \tag{2.6}$$

The seventh degree Bernstein basis polynomials are illustrated in Fig. 2.1 as an example.

Consider a cubic Bezier curve, which is defined by four different points as:

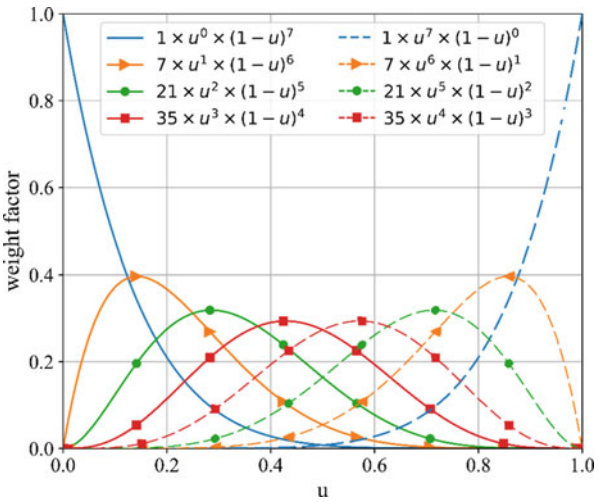


Fig. 2.1 Seventh degree Bernstein basis polynomials

$$\mathrm{P}_0 = \begin{cases} x_0 \\ f_0 \end{cases} \quad \mathrm{P}_1 = \begin{cases} x_1 \\ f_1 \end{cases}$$
$$\mathrm{P}_2 = \begin{cases} x_2 \\ f_2 \end{cases} \quad \mathrm{P}_3 = \begin{cases} x_3 \\ f_3 \end{cases} \tag{2.7}$$

Using Eqs. (2.6) and (2.7), a cubic Bezier curve can be determined in a parametric form as:

$$\begin{cases} X_{(u)} = u^3 x_3 + 3\,(1-u)\,u^2 x_2 + 3(1-u)^2 u x_1 + (1-u)^3 x_0 \\ F_{(u)} = u^3 f_3 + 3\,(1-u)\,u^2 f_2 + 3(1-u)^2 u f_1 + (1-u)^3 f_0 \end{cases} \quad 0 \le u \le 1 \tag{2.8}$$

It is clear that for $u = 0$ the curve passes through P_0, while in $u = 1$ the curve passes through the end point P_3.

This interpolation idea can be employed to estimate the nonlinear integrands in the system of ODEs and solve the resulting linear problem analytically.

2.3 Multistep Method

The general form of a multistep method can be defined as (2.9). This method is called $p + 1$ step when α_p and/or β_p are nonzero.

$$y_{n+1} = \sum_{i=0}^{p} \alpha_i y_{n-i} + h \sum_{i=-1}^{p} \beta_i f\,(x_{n-i}, y_{n-i}) \quad n \ge p \tag{2.9}$$

This equation is explicit when $\beta_{-1} = 0$ and is implicit otherwise. For a $p + 1$ step method, the initial values of the first $p + 1$ points need to be determined using other one-step methods such as Euler or Runge-Kutta [12].

Assuming $\hat{y}_{n+1}$ be the true value of y at x_{n+1}, the truncation error (TE) of the multistep method (2.9) can be defined as:

$$TE_n = \frac{1}{h}\left[\hat{y}_{n+1} - y_{n+1}\right] \tag{2.10}$$

where h is the step size and is equal to the difference between x_{n+1} and x_n. A multistep method is called consistent if the maximum truncation error approaches zero as the step size decreases.

2.4 Methodology

Consider an ordinary differential equation subjected to the initial condition as:

$$\frac{\mathrm{dy}}{\mathrm{dx}} = f_{(x,y)} \tag{2.11}$$

$$\mathrm{y}_{(\mathrm{x=a})} = y_a \tag{2.12}$$

In cases where f is a nonlinear complicated function and there is no exact solution for the problem, the nonlinear function should be approximated with a linear function. If f is only a function of one independent variable, then it can be approximated in the whole domain in advance. However, in cases where f is a function of both dependent and independent variables, it is impossible to approximate f in the whole span. In such cases, the differential equation should be first solved using one-step methods such as classical Runge-Kutta for a few steps, and then the problem is solved using the values calculated in the one-step method.

Using Bézier interpolation approach, one may rewrite the integral form of differential Eq. (2.11) as:

$$\int_{y_n}^{y_{n+1}} dy = \int_{x_n}^{x_{n+1}} f_{(x,y)} dx = \int_{u_1}^{u_2} F_{(u)} \mathrm{X}'_{(\mathrm{u})} du \tag{2.13}$$

Substituting and integrating a four-point Bézier interpolation scheme introduced in (2.8), the value of y for the next step, which is called the predictor, is determined as:

$$y_{n+1}^{P} = y_n + \frac{h}{108}\left[175 f_n - 81 f_{n-1} + 15 f_{n-2} - f_{n-3}\right] \tag{2.14}$$

By interpolating the integrand function using Bézier curves with different number of control points, various formulations can be obtained. Resulting equations of several modes based on their number of points are shown in Table 2.1 where β_i is the coefficient of f_i. Using equations presented in the table, it is possible to calculate the information of the next point in each step based on the information of

Table 2.1 Bézier predictor coefficients for different number of steps

# of points	β_n	β_{n-1}	β_{n-2}	β_{n-3}	β_{n-4}	β_{n-5}
3	19/12	−2/3	1/12			
4	175/108	−3/4	5/36	−1/108		
5	2101/1280	−4/5	113/640	−3/160	1/1280	
6	31,031/18750	−5/6	763/3750	−101/3750	7/3750	−1/18750

the previous points by extrapolation. Values obtained from this table are called the *predictors*.

Using the predicted data in the first step, a new Bezier curve can pass through the previous points and this new point. By integrating the function approximated using this new Bezier curve, the obtained value in the previous step can be corrected by interpolation. The value obtained from this step is called the *corrector*. The results of using correctors for solving differential equations with the help of Bezier curves indicate that a wider step could be chosen for the problem, which can reduce computational cost. This method also has a better stability and accuracy compared to the one solved without the help of correctors.

In order to calculate the corrector, it is possible to use four or five points as the control points. For instance, if all previous points except the first one and the predicted value are used to calculate the control points, the four-control point corrector formula becomes:

$$\begin{cases} X_{(u)} = u^3 x_{n+1} + 3\,(1-u)\,u^2 x_n + 3(1-u)^2 u x_{n-1} + (1-u)^3 x_{n-2} \\ F_{(u)} = u^3 \mathrm{f}_{\left(\mathrm{x_{n+1}}, \mathrm{y}^{\mathrm{P}}_{\mathrm{n+1}}\right)} + 3\,(1-u)\,u^2 f_{(x_n, y_n)} + 3(1-u)^2 u f_{(x_{n-1}, y_{n-1})} + (1-u)^3 f_{(x_{n-2}, y_{n-2})} \end{cases} \tag{2.15}$$

Considering these values, the first correction for y_{n+1} can be derived as:

$$\left[y^C_{n+1}\right]^1 = y_n + \frac{h}{108}\left[65 f_{(x_{n+1}, y^P_{n+1})} + 33 f_{(x_n, y_n)} + 9 f_{(x_{n-1}, y_{n-1})} + f_{(x_{n-2}, y_{n-2})}\right] \tag{2.16}$$

Consequently, the general corrector relation for the K^{th} iteration can be obtained as:

$$\left[y^C_{n+1}\right]^K = y_n + \frac{h}{108}\left[65 f_{\left(x_{n+1}, \left[y^C_{n+1}\right]^{K-1}\right)} + 33 f_{(x_n, y_n)} + 9 f_{(x_{n-1}, y_{n-1})} + f_{(x_{n-2}, y_{n-2})}\right] \tag{2.17}$$

Various types of corrector relations can be obtained depending on the number of nodes used to approximate the function. Furthermore, the general forms of Bézier predictor and corrector equations are presented in Eqs. (2.18) and (2.19), respectively.

$$y^P_{n+1} = y_n + h\sum_{i=0}^{m} \beta_{n-i} f_{n-i} = y_n + h\sum_{i=0}^{m}\left[\left(m\int_1^{\frac{m+1}{m}} \binom{m}{i} (1-u)^{m-i} u^i du\right) f_{n-m+i}\right] \tag{2.18}$$

Table 2.2 Bézier corrector coefficients for different number of steps

# of points	β_{n+1}	β_n	β_{n-1}	β_{n-2}	β_{n-3}	β_{n-4}
3	7/12	1/3	1/12			
4	65/108	11/36	1/12	1/108		
5	781/1280	47/160	53/640	1/80	1/1280	
6	3843/6250	359/1250	103/1250	53/3750	1/750	1/18750

$$y_{n+1}^{k} = y_n + h\sum_{i=-1}^{m}\beta_{n-i}\,f_{n-i} = y_n + h\sum_{i=0}^{m}\left[\left(m\int_{\frac{m-1}{m}}^{1}\binom{m}{i}(1-u)^{m-i}u^i du\right)f_{n-m+i+1}^{k}\right] \tag{2.19}$$

The coefficients of the predictor and corrector equations for different number of steps are shown in Tables 2.1 and 2.2, respectively. All the presented formulations in these tables have the error of $\mathcal{O}\left(h^2\right)$.

2.5 Stability Analysis

The general form of a $p+1$ step predictor and corrector Bézier method can be, respectively, defined as:

$$y_{n+1}^{P} = y_n + h\sum_{i=0}^{p}\beta_{n-i}^{*}\,f\left(x_{n-i}, y_{n-i}\right) \tag{2.20}$$

$$y_{n+1}^{C} = y_n + h\left[\beta_{n+1}\,f\left(x_{n+1}, y_{n+1}^{P}\right) + \sum_{i=0}^{p-1}\beta_{n-i}\,f\left(x_{n-i}, y_{n-i}\right)\right] \tag{2.21}$$

To assess the stability region of a multistep linear method, one may employ the method to solve the test equation [3, 7] defined as:

$$\frac{dy}{dx} = \lambda y \;\mathrm{Re}\left(\lambda\right) \leq 0 \tag{2.22}$$

A multistep method with step size h is called stable for all set of $h\lambda$ values when it has a bounded solution as x increases for all initial values. The region of the stability for the predictor and corrector Bézier formulations is the region in the $h\lambda$ plane where all roots of the stability polynomials (2.23) are less than or equal to 1 [6].

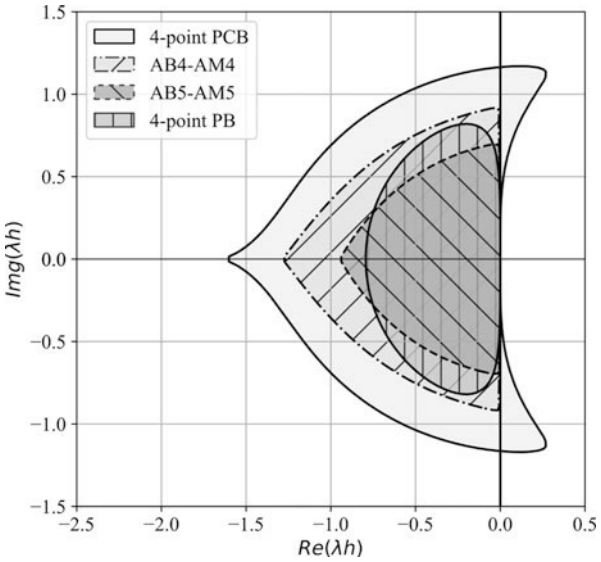


Fig. 2.2 The stability region of the four-point predictor-corrector, predictor-corrector Adams methods, and Bézier predictor

$$\begin{aligned} \Pi^{\mathrm{P}}(\xi) &= \left(\xi^{p+1} - \xi^{p}\right) - h\lambda \sum_{i=0}^{p} \beta^{*}_{n-i} \xi^{p-i} \\ \Pi^{C}(\xi) &= \xi^{p+1} - (1 + \beta_{n+1} h\lambda)\, \xi^{p} - h\lambda \beta_{n+1} \sum_{i=0}^{p} \left(h\lambda \beta^{*}_{n-i} + \frac{\beta_{n-1}}{\beta_{n+1}}\right) \xi^{p-i} \end{aligned} \tag{2.23}$$

Figure 2.2 illustrates the stability region of the four-point formulas of the PCB method and four-point and five-point predictor-corrector Adams (PCA) methods. For the PCA method, the Adams-Bashforth method is employed as predictor, and the Adams-Moulton method is used as corrector. The shaded areas in this figure are the stability regions of each method. As shown in this figure, the stability region of the presented PCB method is almost three times greater than the PB method implying that it is stable for a larger step size. Also, the stability domain of PCB method is larger than the Adams predictor-corrector methods with four and five points.

2.6 Numerical Experiments

In this section, two well-known initial value problems are solved using the developed *predictor-corrector Bezier (PCB)* method. For comparison purposes, other well-known techniques such as Runge-Kutta, Adams-Moulton, and four-point Bezier method are used to examine performance of the method in terms of accuracy, convergence, and stability. In these examples for the small enough step sizes, all methods are closely related to the physical solution. However, as the step size

increases, the modified Bezier method is more stable and closer to the physical solution while other methods diverge.

2.6.1 *Duffing Equation*

One of the well-known examples of nonlinear ODEs in engineering is the Duffing equation, named after George Duffing (1861–1944), which describes an oscillating system with cubic nonlinearity. This equation has been used in various forms to describe different nonlinear systems in dynamics, including the chaos and bifurcation of Duffing oscillator. Furthermore, the Duffing equation has been introduced with or without driving force, which normally appears in the right hand side of the equation.

Here, the Duffing equation which describes governing equation for nonlinear oscillation of a mass, damper, and spring system with no driven force [11] is considered as:

$$\frac{d^2y}{dt^2} + \frac{c}{m}\frac{dy}{dt} + \frac{K_L}{m}y + \frac{K_{NL}}{m}y^3 = 0 \tag{2.24}$$

Subjected to the initial conditions as:

$$\begin{cases} y_{(0)} = 1 \\ y'_{(0)} = 0 \end{cases} \tag{2.25}$$

The Duffing equation has been solved for $\frac{c}{m} = 0.5$, $\frac{K_L}{m} = 1.0$, and $\frac{K_{NL}}{m} = 1.0$ using the Bezier curve, and it has been observed that using Bezier curve leads to a more stable solution with larger step sizes compared to Adams-Moulton method requiring less computational cost compared to classical Runge-Kutta method.

Although using Bezier curve decreases computational cost while increasing the accuracy and stability of the solution, increasing time steps, similar to all other methods, causes loss of accuracy and precision. In such a way, for a step size equal to 1 s, the Bezier curve also diverges and does not provide any solution to the problem. Nevertheless, implementing the correction for the Bezier curve at each step not only increases the accuracy of the solution but also allows larger step sizes to apply. This is particularly important for problems requiring a great marching in time or problems with a large domain, since in these cases if a larger step can be used, the computational cost will be significantly reduced.

In Fig. 2.3, results of the solution for the Duffing equation are shown using different number of corrections. As shown in this figure, the corrected Bezier method still has a stable solution for a time step as large as 0.7. The solution of Duffing equation using classical fourth-order Runge-Kutta method is also included. As shown, while for smaller step size all methods are stable and close to the ground

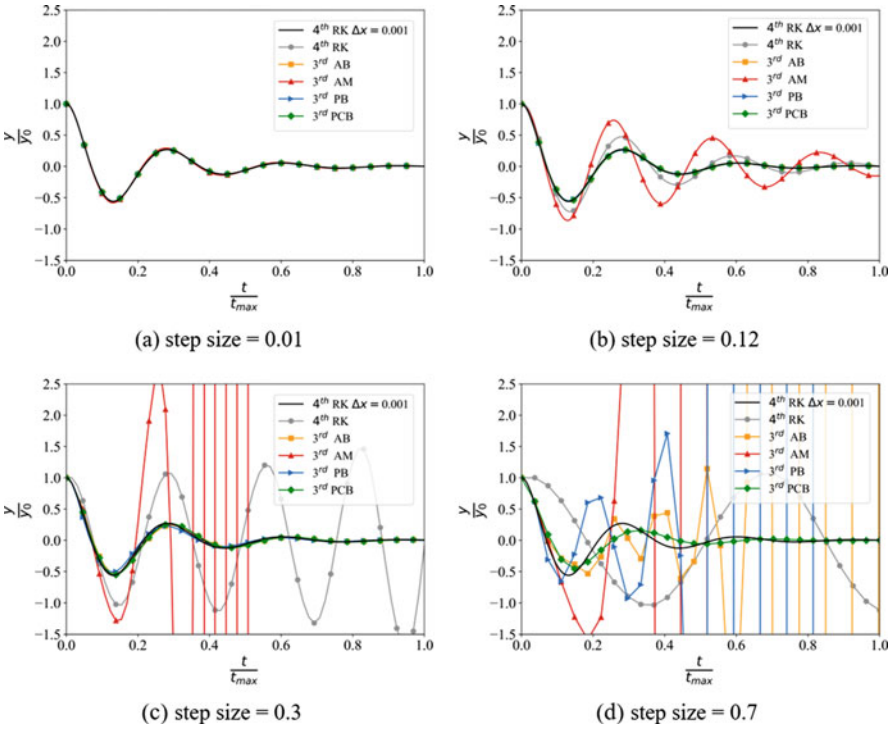


Fig. 2.3 Effect of the step size on the result of different methods for solving Duffing equation

true solution, increasing the step size leads to an unstable result for three-point Adams-Moulton (AM) and Runge-Kutta methods.

To assess the error of the methods, the mean absolute error (MAE) of the methods is plotted for different time steps in Fig. 2.4. As shown in this figure, for all step sizes, the PCB method has the lowest error compared to others.

2.6.2 Van Der Pol Equation

The next benchmark study is related to another well-known equation in the field of nonlinear dynamics and vibration. In the Van der Pol equation, oscillation of a nonconservative system with nonlinear damping is designated. Although the equation was first introduced to examine stability conditions for electrical circuits by Van Der Pol as an electrical engineer in Philips, it has also been used to describe various physical, biological [5] and seismological [2] systems.

The van der Pol equation for an oscillator with nonlinear damping which is governed by the second-order differential equation is:

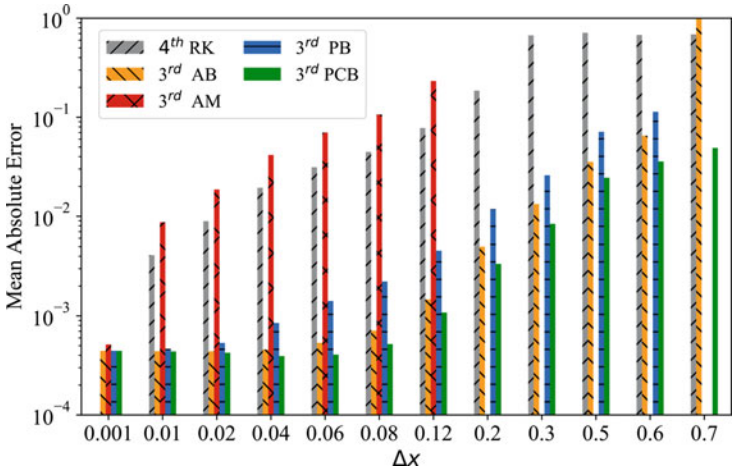


Fig. 2.4 The mean absolute error of solving Duffing equation for different step sizes

$$\begin{aligned} &\frac{d^2y}{dt^2} - \delta\left(1 - y^2\right)\frac{dy}{dt} + y = 0 \\ &\text{Initial conditions}: \begin{cases} y|_{t=0} = 1 \\ \frac{dy}{dt}|_{t=0} = 0 \end{cases} \end{aligned} \tag{2.26}$$

in which y is the position of dynamical variable and δ is a scalar parameter to introduce the nonlinearity of the damping.

This initial value problem is also solved with five different methods and different time steps. The time series results for the time steps of 0.01 and 0.2 are shown in Fig. 2.5. As shown in this figure, the PB and PCB methods show more stable and accurate results in solving this equation.

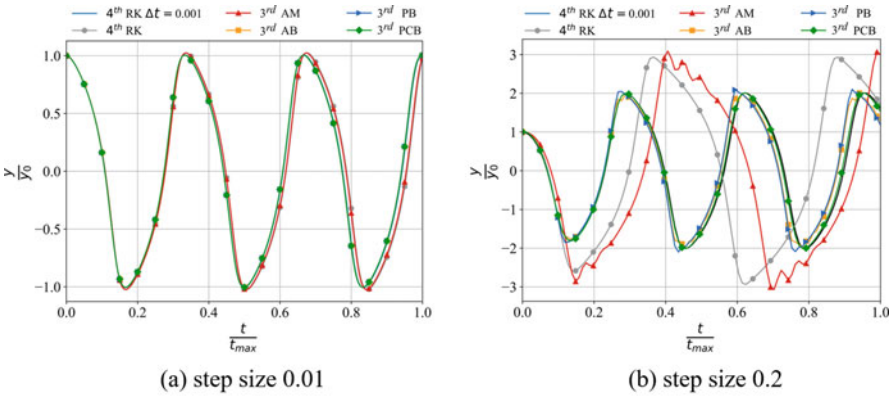


Fig. 2.5 The result of solving Van Der Pol equation with five different methods

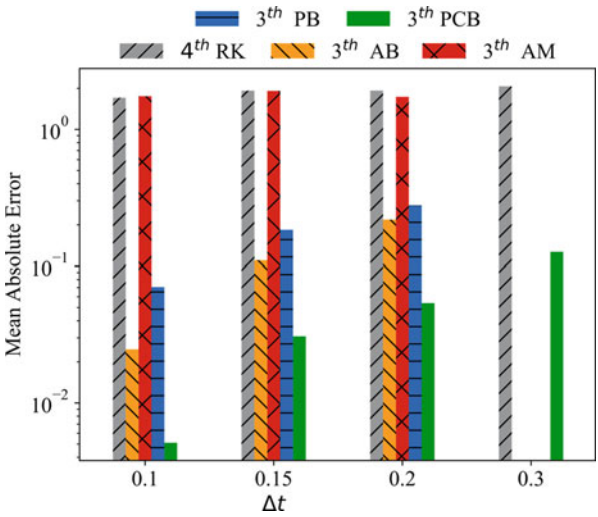


Fig. 2.6 The mean absolute error of Van Der Pol equation for five different methods and step sizes

The mean absolute error of different methods and various step sizes are plotted in Fig. 2.6. Although increasing the step size leads to a higher error for all methods, the PCB method has lowest MAEs for all step sizes and it is still stable for the step size of 0.25.

2.7 Conclusion

In this chapter, a predictor-corrector method for solving initial value problems is proposed. To this end, first, predictor and corrector formulations are developed in general form, and then the truncation error for both predictor and corrector is calculated for some cases using Taylor series. In addition, some nonlinear initial value problems are included as examples to illustrate the effectiveness and reliability of the presented method comparing to the classical Runge-Kutta and Adams-Moulton methods. Finally, the effect of step size on the results of the proposed method is examined for the examples. Results indicate that the presented approach is more accurate and stable having less computational costs and lower iteration steps, suggesting the method as a new alternative solution technique with more efficiency in solving different linear and nonlinear problems in various science and engineering applications.

References

1. Aghdam, M. M., Haghi, P., & Fallah, A. (2015). Nonlinear initial value ordinary differential equations. In *Nonlinear Approaches in Engineering Applications: Applied Mechanics, Vibration Control, and Numerical Analysis* (pp. 117–136). Springer International Publishing. https://doi.org/10.1007/978-3-319-09462-5_5
2. Cartwright, J., Eguiluz, V., Hernandez-Garcia, E. and Piro, O., (1999). "Dynamics of elastic excitable media", *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 9, 2197–2202, https://doi.org/10.1142/S0218127499001620
3. Dahlquist, G. G., Liniger, W., & Nevanlinna, O. (1983). Stability of Two-Step Methods for Variable Integration Steps. *SIAM Journal on Numerical Analysis*, *20*(5), 1071–1085. https://doi.org/10.1137/0720076
4. Euler, L. (1845). *Institutionum calculi integralis*. https://books.google.com/books?hl=en&lr=&id=fLNBAQAAMAAJ&oi=fnd&pg=PA33&dq=Institutionum+calculi+integralis+is+a+three-volume+textbook+written+by+Leonhard+Euler+and+published+in+1768&ots=RuQNpP5bEp&sig=CvJkhG6KYGcnKrdFI2JyTmHUGmE
5. FitzHugh, R., (1961). "Impulses and physiological states in theoretical models of nerve membranes", *Biophysics J*, *1*(6), 445–466. https://doi.org/10.1016/S0006-3495(61)86902-6
6. Garrappa, R. (2010). On linear stability of predictor-corrector algorithms for fractional differential equations. *International Journal of Computer Mathematics*, *87*(10), 2281–2290. https://doi.org/10.1080/00207160802624331
7. Ghrist, M. L., Fornberg, B., & Reeger, J. A. (2015). Stability ordinates of Adams predictor-corrector methods. *BIT Numerical Mathematics*, *55*(3), 733–750. https://doi.org/10.1007/s10543-014-0528-7
8. Kutta, W. (1901). Beitrag zur näherungsweisen Integration totaler Differentialgleichungen. *Zeitschrift Für Mathematik Und Physik*, *46*, 435–453. https://ci.nii.ac.jp/naid/10004015913/
9. Misirli, E., & Gurefe, Y. (2011). Multiplicative Adams Bashforth-Moulton methods. *Numerical Algorithms*, *57*(4), 425–439. https://doi.org/10.1007/s11075-010-9437-2
10. Runge, C. (1895). Ueber die numerische Auflösung von Differentialgleichungen. In *Mathematische Annalen* (Vol. 46, Issue 2). https://doi.org/10.1007/BF01446807
11. Yusufoğlu (Agadjanov), E. (2006). Numerical solution of Duffing equation by the Laplace decomposition algorithm. *Applied Mathematics and Computation*, *177*(2), 572–580. https://doi.org/10.1016/j.amc.2005.07.072
12. Zingg, D. W., & Chisholm, T. T. (1999). Runge-Kutta methods for linear ordinary differential equations. *Applied Numerical Mathematics*, *31*(2), 227–238. https://doi.org/10.1016/S0168-9274(98)00129-9

Chapter 3
Control of Nonhyperbolic Dynamical Systems Through Center Manifold Control

Firdaus E. Udwadia

Abstract This chapter proposes a simple approach for the control of nonlinear dynamical systems with nonhyperbolic equilibrium points. Such equilibrium points are generally much more difficult to analyze dynamically, and correspondingly the control of nonlinear systems in the vicinity of such points can often become more difficult. The aim is to bring about asymptotically stable behavior of the controlled system in the vicinity of the nonhyperbolic equilibrium point. A new way to control such systems is proposed here through control of their local center manifolds. A simple and effective methodology for doing this is provided, and its advantages are illustrated through several examples.

Keywords Dynamical Systems · Center manifold control · Nonhyperbolic dynamical systems · Asymptotic stability · Examples

3.1 Introduction

In local analysis of nonlinear dynamical systems, one of the most useful and powerful results is the Grobman-Hartman (GH) result [1, 2] that proves topological conjugacy in the vicinity of a hyperbolic equilibrium (fixed) point between a nonlinear dynamical system and its linearization at that equilibrium point. By a hyperbolic equilibrium point is meant one at which the Jacobian of the system has eigenvalues whose real parts are nonzero, and by topological conjugacy is meant that there is a continuous invertible bijective mapping that preserves the direction of time and that maps the phase portrait of the nonlinear system to that of its linearization in the vicinity of the equilibrium point.

When an equilibrium point of a nonlinear system is not hyperbolic, the Grobman-Hartman (GH) result can no longer be used, and the dynamics that ensues in

F. E. Udwadia (✉)
Departments of Aerospace and Mechanical Engineering, Civil and Environmental Engineering, Mathematics, Information and Operations Management, University of Southern California, Los Angeles, CA, USA

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_3

its vicinity can be considerably more complex and therefore more challenging to control. One of the approaches to understanding the consequent dynamical behavior in this situation is the use of results that utilize the concept of center manifolds [3–7]. The theory of center manifolds is a rigorous development of the theory of nonlinear differential equations that includes systems with multiple timescales, especially systems that have so-called slow and fast variables.

In this chapter, we consider an approach which is simple and straightforward for the control of such systems and obtain controls that ensure asymptotic stability at the nonhyperbolic equilibrium point along with a region of attraction that is often reasonable in "size" (in phase space) from a practical engineering viewpoint.

Consider the nonlinear dynamical system described by the equations:

$$\begin{aligned} \dot{x} &= Ax + f\,(x, y) \\ \dot{y} &= By + g\,(x, y) \end{aligned} \tag{3.1}$$

in which $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, A and B are constant matrices and the dots denote differentiation with respect to time. The matrix A is assumed to have eigenvalues, λ_A, whose real parts are zero, and the matrix B is assumed to have eigenvalues, λ_B, whose real parts are negative. The matrix B is taken to be a stabile matrix because our interest is in stability/instability of the nonhyperbolic equilibrium points of systems. The functions $f(x, y)$ and $g(x, y)$ are assumed to be C^2 with $f(0, 0)) = 0$, $Df(0, 0) = 0$, $g(0, 0) = 0$, and $Dg(0, 0) = 0$, where $Df(0, 0)$ denotes the Jacobian of f evaluated at $(0, 0)$. The set $S \subset \mathbb{R}^{n+m}$ is a local invariant manifold of Eq. (3.1) if for any solution $(x(t), y(t))$ with $(x(0), y(0)) \subset S$ there is a positive time T such that $(x(t), y(t)) \subset S$ for $t \in [0, T]$.

One could also consider a more general system that is topologically conjugate to Eq. (3.1), described by $\dot{z} = w(z), z \in \mathbb{R}^{n+m}$ whose linearization about the equilibrium point $z = 0$ yields a Jacobian $Dw(0)$ that has n eigenvalues whose real parts are zero and m eigenvalues whose real parts are negative. However, in this chapter we will continue to use Eq. (3.1) since it is more explicit.

Were the functions f and g in Eq. (3.1) to be identically zero, it would simplify to $\dot{x} = Ax, \dot{y} = By$. Then $y = 0$ would be an invariant manifold, meaning that for $y(t = 0) = 0$, $y(t) = 0$, for $\forall t$, and the flow on this subspace would be given by the simpler equation $\dot{u} = Au$. Were the initial condition to be $y(t = 0) = \delta$, for "small" δ, $y(t)$ would exponentially go to zero, that is, to the invariant manifold $y = 0$; the long-term behavior of system (3.1) would again then be provided by the equation $\dot{u} = Au$. This simpler self-contained equation with $u \in \mathbb{R}^n$ can be viewed as a kind of "reduction" of the $k := n + m$ dimensional system (3.1) to one of lower dimension, namely, n.

The generalization of this idea of decoupling the dynamics and, in essence, reducing the order of the system's dimension when the functions f and g are not identically zero is provided by the central results obtained in center manifold theory. It can be shown that the system described by Eq. (3.1) has the following properties [2–7]:

Existence
There exists a smooth (C^k) m-dimensional center manifold of the form $y = h(x)$, $\|x\| < \delta$, such that $y(0) = h(0) = 0$ and $Dh(0) = 0$. The dynamics on this center manifold are governed by the equation:

$$\dot{u} = Au + f\left(u, h(u)\right). \tag{3.2}$$

In other words, system (3.1) whose dimension is $k := n + m$ possesses a lower-dimensional center manifold of dimension m, and the dynamics on this manifold is self-contained and given by the n-dimensional system (3.2).

Stability
Suppose that the zero solution of Eq. (3.2) is stable (asymptotically stable) (unstable), then the zero solution of Eq. (3.1) is stable (asymptotically stable) (unstable).

Asymptotic Behavior of Trajectories
Suppose that the zero solution is stable. Let $(x(t), y(t))$ be a solution of Eq. (3.1) with the initial condition $((x(0), y(0))$ sufficiently small (though in practice, sufficiently small may be quite substantial); then there exists a solution of Eq. (3.2), $u(t)$, such that as $t \to \infty$:

$$x(t) = u(t) + \mathrm{O}\left(\mathrm{e}^{-\gamma t}\right), \; y(t) = h\left(u(t)\right) + \mathrm{O}\left(\mathrm{e}^{-\gamma t}\right), \tag{3.3}$$

where $\gamma > 0$ is some constant.

In other words, from a sizable range of initial conditions near the origin, all solutions of Eq. (3.1) tend exponentially in time to a solution on the center manifold. The n-dimensional reduced system given by Eq. (3.2) on the m-dimensional center manifold $y = h(x)$ faithfully models the original system (3.1) as $t \to \infty$.

Approximation
To get an approximation of the center manifold, $y = h(x)$ is substituted in the second relation is Eq. (3.1) to obtain, using the chain rule,

$$Dh(x)\left[Ax + f\left(x, h(x)\right)\right] - Bh(x) - g\left(x, h(x)\right) = 0 \tag{3.4}$$

which is required to be solved along with the conditions that $h(0) = 0$ and $Dh(0) = 0$. An approximation to $h(x)$ is obtained by defining:

$$M\tilde{h}(x) := D\tilde{h}(x)\left[Ax + f\left(x, \tilde{h}(x)\right)\right] - b\tilde{h} - g\left(x, \tilde{h}(x)\right) \tag{3.5}$$

and attempting to solve approximately $M\tilde{h} = 0$. Here, let $\tilde{h}$ be a map from $\mathbb{R}^n$ into $\mathbb{R}^m$ with $\tilde{h}(0) = 0$ and $D\tilde{h}(0) = 0$. If as $x \to 0$, $M\tilde{h} = \mathrm{O}\left(\|x\|^q\right)$ then $h(x) = \tilde{h}(x) + \mathrm{O}\left(\|x\|^q\right)$. In other words, if Eq. (3.4) is satisfied to some order of accuracy, then the center manifold will have been found to the same order of accuracy.

Often, in bifurcation studies, a p-vector, ε, consisting of constant parameters, is present in Eq. (3.1). Then, if [1]:

$$M\tilde{h} = \mathrm{O}\left(\|x\|^q, \|\varepsilon\|^s\right), \text{ then } h(x) = \tilde{h}(x) + \mathrm{O}\left(\|x\|^q, \|\varepsilon\|^s\right). \tag{3.6}$$

Remark 1 The center manifold, $y = h(x)$, is generated by our desire that it satisfies Eq. (3.4), but finding a suitable $h(x)$ so that this equation is exactly satisfied is usually not possible, because it would be tantamount to solving the nonlinear Eq. (3.1). Hence the approximate solution correct to some order is sought by setting $M\tilde{h} = 0$. Often a power series approximation is used.

Assuming now that Eq. (3.1) is a model of a naturally occurring system or an engineered system, we could envisage controlling the system by controlling its center manifold so in the presence of the control: (1) instead of getting an approximation to $h(x)$ that is correct to some $\mathrm{O}(\|x\|^q, \|\varepsilon\|^s)$ as in Eq. (3.6), we can obtain the function exactly, and (2) we can make a system that is unstable at a hyperbolic equilibrium point asymptotically stable by suitable choices of center manifolds that are cognizant of practical needs, limitations, and requirement, for controlling the system. It is these aspects of control that are related to center manifold theory that this chapter is centrally concerned with.

3.2 Control of Dynamical Systems with Nonhyperbolic Equilibrium Points

Our intent is to provide a method to control the unstable dynamics at a nonhyperbolic equilibrium point (which is taken, with no loss of generality, to occur at the origin) and make this nonhyperbolic equilibrium point asymptotically stable.

To illustrate the central ideas addressed in this chapter, we begin with a simple example. Consider the system of equations:

$$\begin{aligned} \dot{x} &= xy + ax^3 + bxy^2 = 0x + f\,(x, y) \\ \dot{y} &= -y + cx^2 + dx^2 y = -y + g\,(x, y) \end{aligned} \tag{3.7}$$

in which a, b, c, and $d \in \mathbb{R}$ are constants and form the parameter 4-vector $\varepsilon = [a, b, c, d]^{\mathrm{T}}$. The matrix $A = 0$, and $B = -1$ here. By the Existence property above, a center manifold $y = h(x)$ exists, which can be obtained by solving Eq. (3.4), which here is:

$$h'(x)\left[xh(x) + ax^2 + bxh(x)^2\right] + y - cx^2 - dx^2 y = 0. \tag{3.8}$$

Using Eq. (3.6), an approximation to the center manifold is obtained by setting $M\tilde{h}(x) = 0$. Assuming that $M\tilde{h}(x) = O\left(x^2\right)$ this gives (see [1] for details):

$$M\tilde{h}(x) = \tilde{h}(x) - cx^2 + O\left(x^4\right) = 0, \tag{3.9}$$

so that

$$h(x) = \tilde{h}(x) + O\left(x^4\right) = cx^2 + O\left(x^4\right). \tag{3.10}$$

Eq. (3.2) then yields, from the Existence property given in Sect. 3.1 above, that

$$\dot{u} = uh(u) + au^3 + buh^2(u) = (a+c)\,u^3 + O\left(x^5\right). \tag{3.11}$$

When $a + c < 0$, Eq. (3.11) is asymptotically stable at the origin, and by the Stability property of the previous section, so is Eq. (3.7). When $a + c > 0$, Eq. (3.11) is unstable at the origin, and therefore so is Eq. (3.7). The situation when $a + c = 0$ poses a difficulty and a higher order approximation is required [1].

While this might be satisfactory in certain situations, from an engineering controls perspective, it leaves open the following practical questions:

1. What if the parameters a and c describing the system were such that $a + c > 0$? How would one control such a system with a nonhyperbolic equilibrium point and make the nonhyperbolic equilibrium point asymptotically stable?
2. Often one is interested in controlling a system to follow a given trajectory or in controlling a system so that *all* its orbits in phase space lie on a prespecified curve (surface, in higher dimensions) as they approach an equilibrium point, so how does one control the system?
3. For engineering applications, one is often interested in the exact description of a local center manifold, instead of obtaining a center manifold that, in this example, is locally approximated by $h(x) = cx^2 + O(x^4)$. How can one ensure that the center manifold is exactly given by $h(x) = cx^2$?
4. More generally, what if one wanted to create, as we shall see below, a center manifold for a given system described by $y = s(x)$, where s is a preferred function of x?

It is the answers to questions like these, which arise mainly from a "controls" perspective, that this chapter deals with.

We note that: (i) Eq. (3.10) gives only an approximation to the center manifold and (ii) Eq. (3.11) that gives the dynamics on the center manifold is also approximately and qualitatively obtained, with our understanding pinned to what happens as $\|x\| \to 0$, though it might be applicable to a region considerably larger.

From an engineering viewpoint, we may want that the system has a desired center manifold and that the dynamics on it evolve in a definite manner. Noting that the center manifold relies on having $y = h(x)$ satisfy the second equation of system (3.7), from an engineering viewpoint, one can add a control to this equation so that we now have the controlled system:

$$\begin{aligned} \dot{x} &= xy + ax^3 + bxy^2 \qquad\qquad = 0x + f\,(x, y) \\ \dot{y} &= -y + cx^2 + dx^2 y + w(x) = -y + g_c\,(x, y)) \end{aligned} \tag{3.12}$$

where $w(x)$ is the nonlinear control that is added, with $w(0) = \dot{w}(0) = 0$; the subscript "c" on g denotes the controlled system.

One could perhaps then determine the control $w(x)$ and, instead of having a somewhat vaguely defined approximation of the center manifold given in Eq. (3.10), demand that the center manifold of the controlled system be exactly $y = h(x) = c_0 x^\beta$ for a suitable (integer) value of β. We note that since we require $h(0) = Dh(0) = 0$, we must have $\beta \geq 2$. Substituting $y = c_0 x^\beta$ in the second equation in (3.12), we get

$$\begin{aligned} w(x) &= \beta c_0 x^{\beta-1}\left[c_0 x^{\beta+1} + ax^3 + bc_0^2 x^{2\beta+1}\right] + c_0 x^\beta - cx^2 - c_0 d x^{\beta+2} \\ &= c_0^2 \beta x^{2\beta}\left(1 + bc_0 x^\beta\right) + c_0\,(a\beta - d)\,x^{\beta+2} + c_0 x^\beta - cx^2. \end{aligned} \tag{3.13}$$

This provides the explicit control $w(x)$ to be applied to Eq. (3.7) so that the center manifold of the controlled system is exactly given by $y = c_0 x^\beta$, and the dynamics unfurls on this center manifold according to the equation:

$$\dot{u} = uh(u) + au^3 + buh^2(u) = c_0 u^{\beta+1} + au^3 + c_0^2 b u^{2\beta+1}. \tag{3.14}$$

Notice that the dynamics only depends on the parameters a and b of the uncontrolled system that appear in the (x-equation) and the parameter c_0 that defines the "controlled" center manifold. By the Stability property in the previous section, asymptotic stability of the origin in system (3.12) is assured when the origin in Eq. (3.14) is asymptotically stable.

From a controls perspective, in order to ensure stability of the nonhyperbolic origin, the "controlled" center manifold needs to be selected with an eye to making the reduced system, which is described by Eq. (3.2), stable at the origin. As seen from Eq. (3.14), β would thus need to be an even integer.

Numerical Example 1 The following numerical example captures the usefulness of this simple approach to the control of nonlinear systems in the vicinity of unstable nonhyperbolic equilibrium points rendering them asymptotically stable.

(i) When $a + c < 0$, as suggested by Eq. (3.10), a suitable controlled center manifold can be taken as $y = c_0 x^\beta$ with $c_0 = c = -2$ and $\beta = 2$. This is brought about by controlling the second equation in (3.12) with the control $w(x)$ given in Eq. (3.13)

The phase plot of the controlled system that employs the control described in Eq. (3.13) with the parameter vector $\varepsilon = [0.5, 1, -2, 1]^{\mathrm{T}}$ is shown in Fig. 3.1. As expected, the equilibrium point at the origin is seen to be asymptotically stable since $a + c_0 = -1.5$. The thin solid line, convex upward, shows the manifold which has exactly the equation $y = cx^2$, and the local dynamics on this manifold is described using Eq. (3.14) by the equation

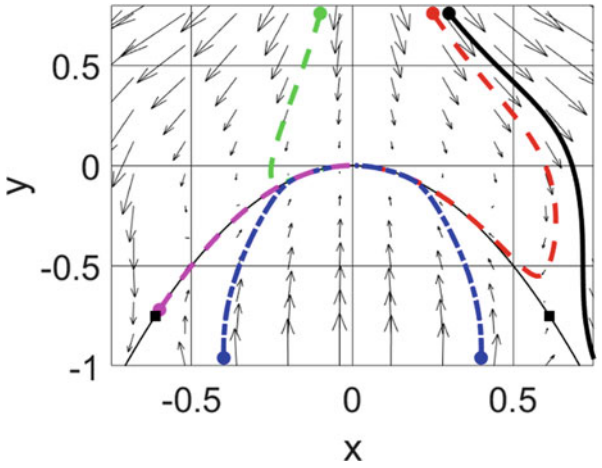


Fig. 3.1 System 3.12 is controlled so that its center manifold is $y = -2x^2$. $a = 0.5$, $b = 1$, $c = -2$, $d = 1$

$$\dot{u} = -1.5u^3 + 4u^5.$$

which shows that the equilibrium point $u = 0$ is asymptotically stable for orbits for which $|u|$ is sufficiently small. The control $w(x)$ provided in Eq. (3.12) is

$$w(x) = 8x^4\left(1 - 2x^2\right).$$

Orbits starting from different initial conditions are shown by dash lines. The initial conditions for each orbit are shown by a solid circle. As seen, the origin is asymptotically stable over a generous region of initial conditions around it. Trajectories move toward the center manifold and move along it to reach the origin.

However, the plotted vector field and the orbit shown by the solid line show that the origin is not globally asymptotically stable; in fact, a pair of unstable equilibrium points exit at $(\pm 0.6123, -0.75)$ shown by solid squares in the figure showing that the Stability property, though it yields asymptotic stability of the equilibrium point, is, of course, local. Determining the basin of attraction of the nonhyperbolic equilibrium point is generally difficult, and it depends both on the parameters of the system and the choice of the desired center manifold.

Remark 2 Where the parameter $a < 0$, one could also have controlled the stability of the nonhyperbolic equilibrium point at the origin and the behavior of nearby orbits by controlling the center manifold of the controlled system to be $y = h(x) = c_0x^4$, $c_0 = \pm 1$, through the use of a suitable control $w(x)$ (as given in Eq. (3.13)). For then, Eq. (3.14) would become:

$$\dot{u} = uh(u) + au^3 + buh^2(u) = au^3 \pm u^5 + bu^9, a < 0$$

whose origin is asymptotically stable for $|u|$ sufficiently small.

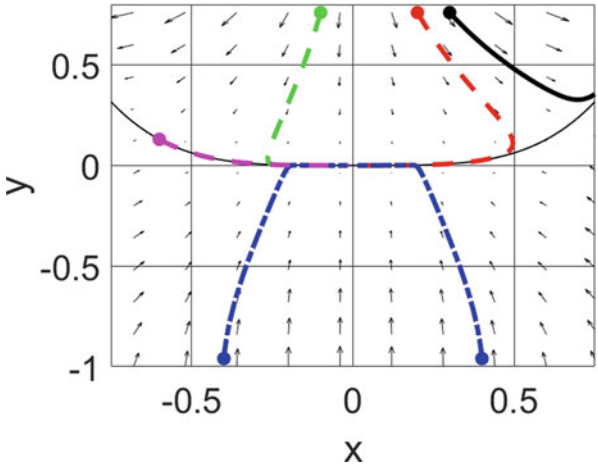


Fig. 3.2 System (3.12) is controlled so that its center manifold is $y = x^4$. $a = -0.5$, $b = 1$, $c = -2$, $d = 1$

This is illustrated for the parameter vector $\varepsilon = [-0.5, 1, -2, 1]^{\mathrm{T}}$, now with $a = -0.5$ and $c_0 = 1$ Using the control $w(x) = 4x^8(1 + x^4) + 2x^2 + x^4 - 3x^6$ obtained from Eq. (3.13), Fig. 3.2 shows the phase portrait and the vector field of the controlled system. The center manifold is now controlled to lie on the curve $y = x^4$ shown by the thin line, which is concave upward now. The equilibrium point at the origin is asymptotically stable. Representative orbits that start in the vicinity of the origin move toward it exponentially fast. (Note the orbit shown by the solid line that does *not* go the origin, illustrating the local nature of the stability.)

In the vicinity of the origin, the system closely follows the trajectory given by the equation $y = x^4$, and the local dynamics (Eq. (3.14)) on this manifold is given by $\dot{u} = -0.5u^3 + u^5 + u^9$.

(ii) When $a + c > 0$, using the Stability property of the previous section, we see from Eq. (3.11) that the uncontrolled system (3.7) is unstable at the origin. However, asymptotic stability of the origin can still be achieved by controlling the center manifold of system through the addition of a control $w(x)$ as in Eq. (3.12). For $\beta = 2$, the center manifold of the controlled system becomes $y = c_0x^2$. With this as the controlled center manifold, by using Eq. (3.14) with $\beta = 2$, we see that asymptotic stability of the origin of the controlled system can then be guaranteed by choosing $c_0 < -a$.

Numerical Example 2 Consider the controlled system with the parameter vector $\varepsilon = [0.5, 1, 0.5, 1]^{\mathrm{T}}$ for which $a + c = 1 > 0$. The phase portrait of the uncontrolled system (3.7) in the vicinity of its unstable nonhyperbolic equilibrium point at the origin is shown in Fig. 3.3a. As expected from Eq. (3.11), this nonhyperbolic equilibrium point is unstable.

We now use the control:

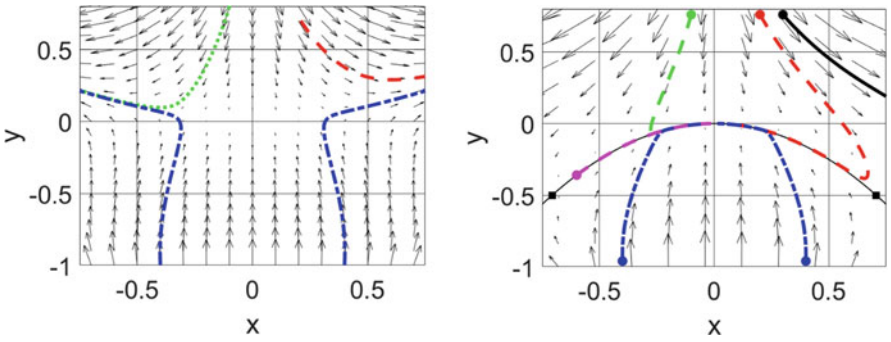


Fig. 3.3 (**a**) Unstable nonhyperbolic equilibrium point of system (Eq. (3.7)) when $a + c > 0$ of uncontrolled system. (**b**) Exponentially stable nonhyperbolic equilibrium point of system (Eq. (3.12)) whose center manifold is controlled

$$w(x) = 2x^4\left(1 - x^2\right) - 1.5x^2$$

to effectively control the center manifold to lie on the curve $y = c_0x^2 = -x^2$.

Since $-1 = c_0 < -a = -0.5$, the unstable nonhyperbolic equilibrium point at the origin in the uncontrolled system is converted to an asymptotically stable equilibrium point. Figure 3.3b shows the behavior of the controlled system. One notes that though the statements in the previous section are true for "sufficiently small" initial conditions, the basin of attraction of the origin controlled in this manner is quite generous. More importantly, (a) asymptotic stability is guaranteed by the first three center manifold properties of the previous section without the need to appeal to the usual Lyapunov direct method that entails finding an appropriate Lyapunov function, and (b) the system has an explicitly described center manifold, and trajectories starting on or close to it (and from a generous enough region around the nonhyperbolic equilibrium point) will very closely track the trajectory given by:

$$x(t) = u(t),\ y = c_0x^2 = -x^2$$

where $\dot{u} = (c_0 + a)\,u^3 + u^5 = -0.5u^3 + u^5$. Evidently, any value of $c_0 < -a$ can be chosen depending on practical considerations that may arise when implementing the control, thereby controlling the center manifold.

Finally, there is no difficulty posed when $a + c = 0$ in order to ensure asymptotic stability of the origin as was previously encountered with the uncontrolled system (see Eq. (3.11)). If $a > 0$, we see from Eq. (3.14) that the system can be controlled to have a center manifold $y = -c_0x^2$, $c_0 > a$, so that the nonhyperbolic equilibrium point at the origin is asymptotically stable (in the vicinity of the origin). Similarly, when $a < 0$, the center manifold can be controlled to be $y = -c_0x^2$, $c_0 > 0$.

This leads us to the following result.

Result Given the system described by Eq. (3.1), and a C^1 function $y = h(x)$ from $\mathbb{R}^n \to \mathbb{R}^m$, the system:

$$\begin{aligned} \dot{x} &= Ax + f\,(x, y) \\ \dot{y} &= By + g\,(x, y) + w(x) \end{aligned} \tag{3.15}$$

The feedback control $w(x) \in C^2$, $w(0) = 0$, $Dw(0) = 0$ where:

$$w(x) = Dh(x)\,[Ax + f\,(x, h(x))] - Bh(x) - g\,(x, h(x)) \tag{3.16}$$

causes the system (3.15) to have the center manifold $y = h(x)$. The dynamics on this center manifold is described by the n-dimensional equation:

$$\dot{u} = Au + f\,(u, h(u))\,. \tag{3.17}$$

Proof: We simply apply the Existence property to the system described by Eq. (3.15). We note that the control is not full-state and is applied to the m-dimensional subsystem that is described by the y-equation in Eq. (3.15). □

Remark 3 The other three properties given in Sect. 3.1 above follow in a similar manner. The Asymptotic Behavior of Trajectories property assures us that for each trajectory of the full dynamical system (3.15) (provided it stays sufficiently close to the nonhyperbolic equilibrium point), there is a particular solution of the lower-dimensional system (3.17) on the center manifold that is approached exponentially fast.

The central idea behind the Result is to provide a control $w(x)$ so that system (3.15), which has a nonhyperbolic equilibrium point at the origin, has a suitable center manifold that attracts nearby orbits in a region around the origin and makes it (the origin) asymptotically stable.

We now use this methodology to the area of mechanics that deals with stability of the nonhyperbolic equilibrium point that arises at the origin when the angular velocity of a rigid body is controlled.

Consider the Euler equations for the rotation of a rigid body in the absence of external torques in the angular velocity space $(\omega_1, \omega_2, \omega_3)$, a problem related to satellite attitude control in deep space. The aim is to ensure that the system remains asymptotically stable for small perturbations around the origin. Assuming that the moments of inertia (MI) about the three principal axes are I_1, I_2, I_3, we obtain the equations:

$$\begin{aligned} \dot{\omega}_1 &= a\omega_2\omega_3 - \alpha_1\omega_1^3 \\ \dot{\omega}_2 &= b\omega_3\omega_1 - \alpha_2\omega_2^3 \\ \dot{\omega}_3 &= c\omega_1\omega_2 - \alpha_3\omega_3 \end{aligned} \tag{3.18}$$

where:

$$a = \frac{(I_2 - I_3)}{I_1}, b = \frac{(I_3 - I_1)}{I_2}, \text{ and } c = \frac{(I_1 - I_2)}{I_3}. \tag{3.19}$$

Our aim is to have the origin $(0, 0, 0)$ asymptotically stable. To achieve stability, we take the parameters $\alpha_1, \alpha_2, \alpha_3 > 0$. Thus, the coordinate ω_3 is provided linear negative feedback, while the coordinates ω_1 and ω_2 are provided negative cubic feedbacks. For small perturbations from the origin, the control torques resulting from the cubic feedback terms would be negligible compared to those resulting from the use of a linear term, resulting in reduced control costs, hence the motivation to use nonlinear cubic feedback for the evolution of ω_1 and ω_2. We consider different cases depending on the values of the three principal moments of inertia.

The central idea is to provide a suitable control $w(x)$ so that the controlled system has a suitable center manifold that attracts nearby orbits in a region around the nonhyperbolic origin and makes it (the origin) asymptotically stable.

Case 1 $I_1 < I_3 < I_2$. I_3 is the intermediate principal moment of inertia (MI); the linear feedback is provided to the coordinate that corresponds to this intermediate moment of inertia I_3. From Eq. (3.19) we see that $a, b > 0$ and $c < 0$. Note that the linear feedback is provided to the coordinate ω_3 which corresponds to the intermediate moment of inertia, I_3.

The system can be rewritten as:

$$\begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\omega}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha_3 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} a\omega_2\omega_3 - \alpha_1\omega_1^3 \\ b\omega_3\omega_1 - \alpha_2\omega_2^3 \\ c\omega_1\omega_2 \end{bmatrix} := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha_3 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} f_1(\omega_1, \omega_2, \omega_3) \\ f_2(\omega_1, \omega_2, \omega_3) \\ g(\omega_1, \omega_2, \omega_3) \end{bmatrix} \tag{3.20}$$

showing that the equilibrium point at the origin, $\omega_1 = \omega_2 = \omega_3 = 0$, is nonhyperbolic. From the nature of the feedback control used here, one might intuit that the equilibrium point is stable. By the Existence property in Sect. 3.1, the system has a center manifold, described by the equation $\omega_3 = h(\omega_1, \omega_2)$. The center manifold of the system can be obtained by using Eq. (3.4) which yields

$$\frac{\partial h}{\partial \omega_1}\left[a\omega_2 h - \alpha_1\omega_1^3\right] + \frac{\partial h}{\partial \omega_2}\left[a\omega_1 h - \alpha_2\omega_2^3\right] + \alpha_3 h - \omega_1\omega_2 = 0.$$

An approximate solution of this equation (Eq. (3.5)) assuming that

$$\tilde{h} = \frac{c}{\alpha_3}\omega_1\omega_2 + O\left(\|\omega\|^4\right) \tag{3.21}$$

gives the center manifold, according to the Approximation property given in Sect. 3.1, as

$$\omega_3 = \frac{c}{\alpha_3}\omega_1\omega_2 + O\left(\|\omega\|^4\right) \tag{3.22}$$

and on this manifold, by Eq. (3.2) the dynamical system evolves according to the relations

$$\begin{aligned} \dot{u}_1 &= \tfrac{ac}{\alpha_3}u_1u_2^2 - \alpha_1u_1^3 + O\left(\|u\|^6\right) = u_1\left(\gamma_1u_2^2 - \alpha_1u_1^2\right) + O\left(\|u\|^6\right) \\ \dot{u}_2 &= \tfrac{bc}{\alpha_3}u_1^2u_2 - \alpha_2u_2^3 + O\left(\|u\|^6\right) = u_2\left(\gamma_2u_1^2 - \alpha_2u_2^2\right) + O\left(\|u\|^6\right) \end{aligned} \tag{3.23}$$

where $\gamma_1 = \frac{ac}{\alpha_3} < 0$ and $\gamma_2 = \frac{bc}{\alpha_3} < 0$. Using the Lyapunov function $V(u_1, u_2) = \left(u_1^2 + u_2^2\right)/2$ we find that

$$\dot{V} = u_1\dot{u}_1 + u_2\dot{u}_2 = -\alpha_1u_1^4 - \alpha_2u_2^4 + (\gamma_1 + \gamma_2)\,u_1^2u_2^2 < 0$$

so that Eq. (3.23) is asymptotically stable at the origin. It should be noted that the truncated Eq. (3.23) has only one equilibrium point, except when $\gamma_1\gamma_2/(\alpha_1\alpha_2) = 1$, which is a situation that can be excluded from consideration because the feedback gains α_1, α_2 and α_3 can always be appropriately selected.

Instead of thinking of the local center manifold as given approximately by the relation $w_3 = \frac{c}{\alpha_3}\omega_1\omega_2 + O\left(\|\omega\|^4\right)$, we can start by controlling the system so that its center manifold is exactly $w_3 = c_0\omega_1\omega_2$. In fact, one can consider making the local center manifold of the controlled system to be:

$$\omega_3 = c_0\omega_1\omega_2 + c_1\omega_1\omega_2^3 = h(\omega_1, \omega_2) \tag{3.24}$$

in which the constant coefficients c_0 and c_1 will be chosen to ensure asymptotic stability of the nonhyperbolic equilibrium point at the origin, accompanied by a generous neighborhood around it that the equilibrium point attracts.

The controlled system can be written as:

$$\begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\omega}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha_3 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} a\omega_2\omega_3 - \alpha_1\omega_1^3 \\ b\omega_3\omega_1 - \alpha_2\omega_2^3 \\ c\omega_1\omega_2 + w(\omega_1, \omega_2) \end{bmatrix} \tag{3.25}$$

and it is identical to that shown in Eq. (3.20), except for the addition of a nonlinear control torque in the last equation. Substituting for $h(\omega_1, \omega_2)$ the expression given in Eq. (3.16), we obtain explicitly:

$$w(\omega_1,\omega_2)=\omega_1\omega_2\Big(c_0^2b\omega_1^2+c_0^2a\omega_2^2+4c_0c_1b\omega_1^2\omega_2^2+2c_0c_1a\omega_2^4-c_0\alpha_1\omega_1^2-c_0\alpha_2\omega_2^2 \\ +c_0\alpha_3+3c_1^2b\omega_1^2\omega_2^4+c_1^2a\omega_2^6-c_1\alpha_1\omega_1^2\omega_2^2-3c_1\alpha_2\omega_2^4+c_1\alpha_3\omega_2^2-c\Big) \tag{3.26}$$

Using Eq. (3.24) in Eq. (3.17), we get:

$$\begin{aligned}\dot{u}_1&=u_1\left(c_0au_2^2-\alpha_1u_1^2\right)+c_1au_1u_2^4\\ \dot{u}_2&=u_2\left(c_0bu_1^2-\alpha_2u_2^2\right)+c_1bu_1^2u_2^3.\end{aligned} \tag{3.27}$$

The Stability property in Sect. 3.1 tells us that if Eq. (3.27) is stable at the origin, then the controlled system will also be stable at the origin. Consider the candidate Lyapunov function:

$$V(u_1,u_2)=\frac{1}{2}\left(u_1^2+u_2^2\right).$$

Then the time derivative of V along the orbit of the dynamical system (3.27) is simply:

$$\dot{V}=u_1\dot{u}_1+u_2\dot{u}_2=-\alpha_1u_1^4-\alpha_2u_2^4+c_0(a+b)u_1^2u_2^2+c_1(a+b)u_1^2u_2^4, \tag{3.28}$$

which is guaranteed to be negative definite for all $c_0, c_1 < 0$, since $a, b > 0$. Thus, by controlling the center manifold to satisfy Eq. (3.24), the nonhyperbolic equilibrium point of the controlled system (3.25) is guaranteed to be stable for values of c_0, $c_1 < 0$. Notice the absence of the parameter c in Eq. (3.28) indicating that the stability of the origin has been made independent of this parameter.

Numerical Example 3 Consider the control of the center manifold of the system described by Eq. (3.18) so that the center manifold is given exactly by $y = -\omega_1\omega_2$. Then the control torque required in Eq. (3.25) is

$$w(\omega_1,\omega_2)=-\omega_1\omega_2\left(-c_0^2b\omega_1^2-c_0^2a\omega_2^2+c_0\alpha_1\omega_1^2+c_0\alpha_2\omega_2^2-c_0\alpha_3+c\right).$$

On this center manifold, the trajectories are given by Eq. (3.27) with $c_0 = -1$, and $c_1 = 0$.

Figure 3.4a shows the behavior of system (3.25) in which the moments of inertia are $I_1 = 1$, $I_2 = 2.5$ and $I_3 = 2$ in consistent units, so that $a = 0.5$, $b = 0.4$, and $c = -0.75$. The parameters $\alpha_1 = \alpha_2 = \alpha_3 = 0.5$. These parameter values for α_1, α_2, and α_3 will be used throughout for all the cases except in Figs. 3.7 and 3.8 that look at the influence of these parameters on the controlled phase portrait of the system. The center manifold is shown by the shaded surface, and the start of each trajectory is shown by a small sphere. Exponential convergence to the center

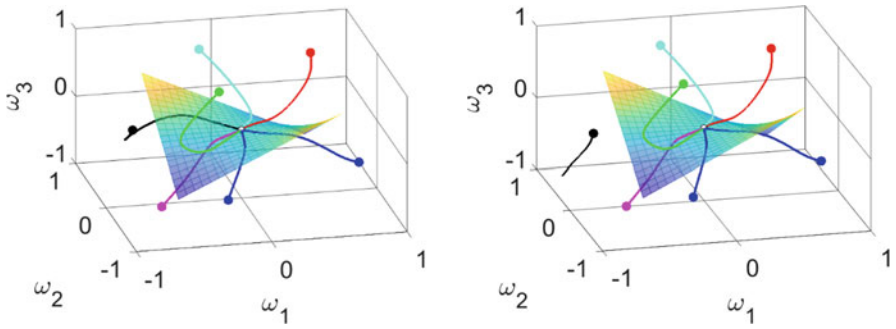


Fig. 3.4 (**a**) System (3.18) is controlled so that its center manifold is $\omega_3 = -\omega_1\omega_2$. (**b**) System (3.18) is controlled so that its center manifold is $\omega_3 = -\omega_1\omega_2 - 0.2\omega_1\omega_2^2$

manifold occurs, and asymptotic stability of the nonhyperbolic equilibrium point at the origin is seen. The representative trajectories are tangent to the center manifold at the origin and are exponentially attracted to the nonhyperbolic equilibrium point at the origin over a generous region of initial conditions around the origin. Note how all the trajectories are "guided" to lie on the local center manifold as they approach the equilibrium point.

Figure 3.4b shows the behavior of the system controlled to have a center manifold defined by $\omega_3 = -\omega_1\omega_2 - 0.2\omega_1\omega_2^3$ using the same initial conditions for the representative trajectories as in Fig. 3.4a. The other parameter values chosen are the same as above. While the trajectories remain, as expected, tangent at the origin to the center manifold, and the origin is asymptotically stable, the domain of attraction of the nonhyperbolic equilibrium point has been somewhat diminished as seen from the one trajectory that is not attracted any more to the origin.

Case 2

$I_2 < I_3 < I_1$. I_3 is again the intermediate moment of inertia (MI), and from Eq. (3.19) $a, b < 0$ and $a, b < 0$. To ensure that the nonhyperbolic equilibrium point at the origin is asymptotically stable, we need $\dot{V}$ given in Eq. (3.28) to be negative definite; hence, we control the system to achieve this objective, by controlling it to have a center manifold given by Eq. (3.24) with $c_0, c_1 > 0$; this ensures that $\dot{V}$ is negative definite. A simple way to see that a control manifold can be used to make the origin stable is simply to rename the second principal axis as the first principal axis in Case 1 and vice versa.

Case 3

$I_1 < I_2 < I_3$. I_2 is now the intermediate principal moment of inertia (MI), and I_3 is the largest principal MI. Now, $a, c < 0$ and $b > 0$. To ensure asymptotic stability of the origin when $(a + b) > 0$, from Eq. (3.28) we see that we can control the center manifold to be that given by Eq. (3.24) with $c_0, c_1 < 0$. When $(a + b) < 0$, from Eq. (3.28) we obviously require $c_0, c_1 > 0$.

Numerical Example 4 Consider the rigid body with $I_1 = 1, I_2 = 2, I_3 = 2.5$ so that $a = -0.5$, $b = 0.75$, and $c = -0.4$. Here $a + b > 0$ and we control the local

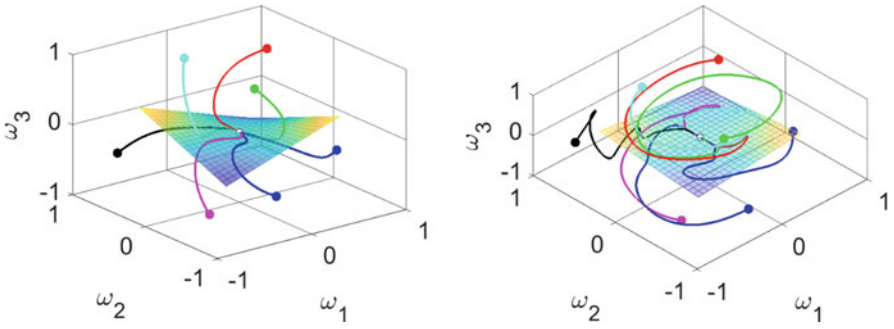


Fig. 3.5 (**a**) Stability of origin when $I_1 < I_2 < I_3$ and $a + b > 0$. (**b**) Stability of origin when $I_1 < I_2 < I_3$ and $a + b < 0$

center manifold of the system so that it is given by $y = -\omega_1\omega_2$ (Eq. (3.24) with $c_0 = -1$ and $c_1 = 0$). The parameters $\alpha_1 = \alpha_2 = \alpha_3 = 0.5$. The phase portrait of the controlled system is shown in Fig. 3.5a. As seen, the origin is asymptotically stable with a generous region in phase space around it that it attracts.

With $I_1 = 1$, $I_2 = 2$, $I_3 = 6$, we have $a = -4$, $b = 2.5$, and $c = -1/6$ so that $(a + b) < 0$. By controlling the center manifold to be as given in Eq. (3.24) with $c_0 = 1.2$, and $c_1 = 0.2$ the nonhyperbolic origin is made asymptotically stable as seen in Fig. 3.5b. Asymptotic stability and trajectories tangential to the center manifold are again seen.

Using similar reasoning, the center manifold when $I_2 < I_1 < I_3$ can also be obtained. Now $a < 0$ and $b, c > 0$. The parameter c does not influence our choice of the parameters that describe the desired center manifold, only the parameters a and b matter is assuring asymptotic stability of the origin (Eq. (3.28)). Hence, this is the same as Case 3 discussed earlier, and stability can be assured by a proper choice of a center manifold.

Case 4

$I_3 < I_1 < I_2$. Here I_1 is the intermediate principal MI and hence the parameters $a > 0$ and $b, c < 0$, which is the same as Case 3, which has already been discussed. When $(a + b) > 0$, stability of the equilibrium point at the origin is assured by controlling the system to have Eq. (3.24) be its center manifold with $c_0, c_1 < 0$.

Numerical Example 5 Considering the system $I_1 = 2$, $I_2 = 3.5$, $I_3 = 1$, with the parameters $\alpha_1 = \alpha_2 = \alpha_3 = 0.5$ so that $a = 1.25$, $b = -0.286$, $c = -1.5$. The controlled local manifold is taken to be Eq. (3.24) with $c_0 = -1.2$, and $c_1 = -0.2$. The phase portrait is shown on Fig. 3.6. As before, the spheres show the initial conditions from which the representative trajectories start. Asymptotic convergence to the stable nonhyperbolic equilibrium point at the origin is seen from a sizable region of phase space. All the trajectories are guided by the control to lie on the local center manifold $\omega_3 = -1.2\omega_1\omega_2 - 0.2\omega_1\omega_2^3$ as they approach the equilibrium point.

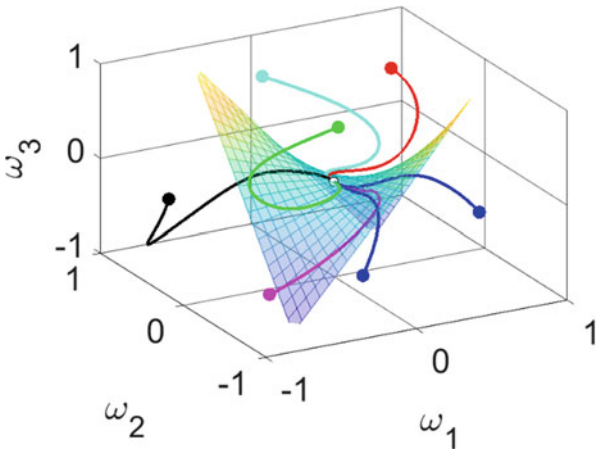


Fig. 3.6 $I_1 < I_2 < I_3$ and $a + b > 0$

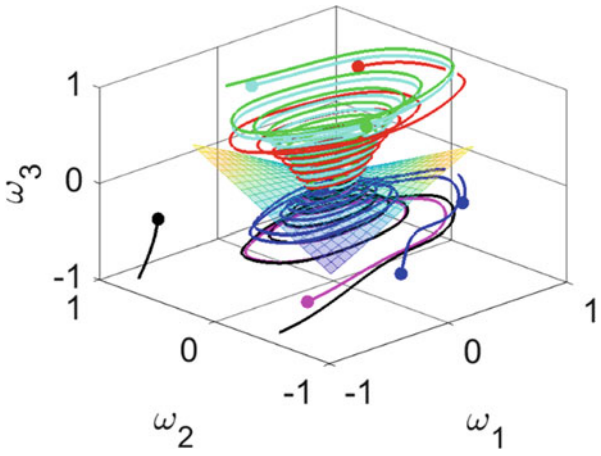


Fig. 3.7 $\alpha_1 = 0, \alpha_2 = 0,$ and $\alpha_3 = 0.01$

Lastly, we consider the effect of the system's parameters on the efficacy of the control. With values of I_1, I_2, and I_3 the same as above we consider a system with greatly reduced feedback gains (see Eq. (3.18)) of $\alpha_1 = 0$, $\alpha_2 = 0$, and $\alpha_3 = 0.01$. Thus, there is no feedback to the first two equations, and the feedback to the third equation is drastically reduced. The nonlinear system is controlled in the vicinity of its nonhyperbolic fixed point by controlling the local center manifold to be $\omega_3 = -1.2\omega_1\omega_2 - 0.2\omega_1\omega_2^3$. Figure 3.7 shows the phase portrait of the system and the center manifold, and Fig. 3.8 shows the projection of the phase portrait on the $\omega_1 - \omega_2$ plane. As seen, trajectories starting from the same points as those in Fig. 3.6 again approach the equilibrium point along the center manifold. The trajectory starting at the leftmost point in Fig. 3.7 appears to go out a distance beyond the limits of the plot shown and returns to the center manifold.

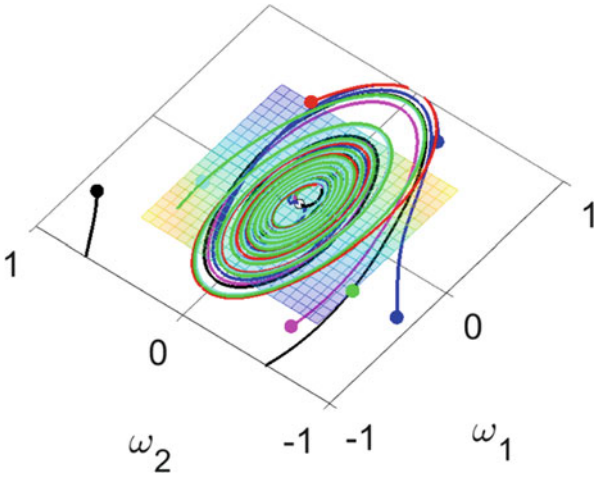


Fig. 3.8 $\alpha_1 = 0, \alpha_2 = 0,$ and $\alpha_3 = 0.01$

3.3 Conclusions

The analysis of nonlinear systems with hyperbolic equilibrium points is greatly facilitated by the Grobman-Hartman (GH) result that states that in the vicinity of such an equilibrium point, the linearized system is topologically equivalent to the nonlinear system. This result makes possible the control of such systems in a straightforward and simple manner. However, the presence of nonhyperbolic equilibrium points makes the GH result inapplicable, and this often poses challenges in our understanding, our analysis, and our ability to control such systems when operating in the vicinity of such nonhyperbolic equilibrium points. This chapter explores a new approach to the control of such systems using results from the theory of center manifolds.

The central idea is to put in place a control on an easily distinguishable part (subsystem) of the dynamical system that results in a desirable local center manifold which ensures asymptotic stability of the nonhyperbolic equilibrium point. The desired center manifold would depend, in general, on practical considerations such as actuator requirements, the region of attraction required around the nonhyperbolic equilibrium point in phase space, etc. A straightforward methodology for doing this is developed. The system is controlled to have a desired local center manifold upon which the evolving dynamics is assured to be asymptotically stable.

The nonlinear dynamical system is divided into two parts. The first part constitutes the subsystem that has the nonhyperbolic equilibrium point, and the remaining hyperbolic subsystem constitutes the second part. As seen, the control effort needs to be applied only to this remaining subsystem. Therefore, the methodology has the advantage of not requiring full-state control while guaranteeing asymptotic stability.

While one could use Lyapunov's direct method in some instances as done here, a disadvantage of the approach presented here is the difficulty in exactly finding the region of attraction of the asymptotically stable nonhyperbolic equilibrium point

after a desirable local center manifold is chosen. As illustrated here, the size of this region of attraction in phase space depends on the structure of the nonlinear system, on the parameters of the system, and on the choice of the desired local center manifold.

Though it has been assumed here that the second remaining subsystem, which is hyperbolic, is stable, extensions of the approach where it could contain both stable and unstable hyperbolic equilibrium points can be easily made by using control methods standardly used when dealing with hyperbolic equilibrium points.

The results obtained here appear to provide a new additional quiver in our arsenal of methods to bring about control of nonlinear dynamical systems, the focus in this chapter being on those systems that have nonhyperbolic equilibrium points in whose vicinity: (i) the nonlinear dynamics is usually more complex, and (ii) linearization methods do not work. As often done when dealing with the theory of center manifolds, examples are provided, one of which has considerable relevance to rigid body dynamics and deals with the control of the angular velocity of a rotating rigid body. The advantages of using the approach, as illustrated by these examples, are its simplicity, efficacy, and ease of use.

References

1. Hartman, P., Ordinary Differential Equations, SIAM Classics in Applied Mathematics, SIAM, 2002.
2. Chicone, C., Ordinary Differential Equations with Applications, Springer, 2006.
3. Carr, J., Applications of Center Manifold Theory, Springer-Verlag, NY, 1981.
4. Carr, J., and Muncaster, R.G., The Application of Center Manifolds to Amplitude Expansions I. Ordinary Differential Equations, Journal of Differential Equations, Vol. 50, pp. 260-279, 1983.
5. Sijbrand, J., Properties of Center Manifolds, Trans. Amer. Math. Soc. Vol. 268, pp. 431-469, 1985.
6. Roberts, A. J., Low Dimensional Modeling Of Dynamical Systems, arXiv:chao-dyn/9705010, 1997.
7. Kim, D., Udwadia, F. E., and Proskurowski, W., A Perturbation Approach for Approximate Inertial Manifolds, Intl. Journ. Comp. and Appl. Math., Vol. 2(1), pp. 15-41, 2007.

Chapter 4
Linear and Nonlinear Aspects of Space Charge Phenomena

Sayavur I. Bakhtiyarov and Dale C. Ferguson

Abstract *Space charge* has a destructive and hazardous impact on electronics, navigation systems, and instrumentation onboard spacecraft, as well as the spacecraft itself. Therefore, it is required to screen and predict the spacecraft charging levels to provide a safe space mission. This chapter reviews the theoretical and experimental studies of linear and nonlinear aspects of space charge phenomena as well as existing methods its mitigation.

Keywords Space charge phenomena · Spacecraft navigation · Surface charging · Polytetrafluoroethylene · Dielectric nanocomposites

Abbreviations

ASTROD	Astrodynamical Space Test of Relativity using Optical Devices
DSCS	Defense Satellite Communications Systems
EDT	Electrodynamic tether
EDX	Energy-dispersive X-ray spectroscopy
ELF	Electron-emitting film
EMU	Extravehicular mobility unit (spacesuit)
EPR	Electron paramagnetic resonance
ESD	Electrostatic discharge
EVA	Extravehicular activity
EWB	Environments WorkBench
FEP	Fluorinated ethylene propylene
FPMU	Floating potential measurement unit
FPP	Floating potential probe
FTIR	Fourier-transform infrared spectroscopy
GCR	Galactic cosmic radiation
GEO	Geosynchronous (geostationary) Earth orbit
IM	Ionic materials

F. E. Udwadia (✉)
Departments of Aerospace and Mechanical Engineering, Civil and Environmental Engineering, Mathematics, Information and Operations Management, University of Southern California, Los Angeles, CA, USA

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_4

ISS	International Space Station
LEO	Low Earth orbit
MEO	Medium Earth orbit
MNSP	Defense Meteorological Satellite Program
MUSCAT	Multi-utility spacecraft charging analysis tool
NASCAP	NASA Air Force Spacecraft Charging Analyzer Program
NEA	Negative electron affinity
PASEC	Passive anode surface emission cathode
PCU	Plasma contacting unit
PEO	Polar Earth orbit
PFPE	Perfluoropolyether
PI	Polyimide
PU	Polyurethane
RCE	Rapid charging event
RIC	Radiation-induced conductivity
SCATHA	Spacecraft Charging at High Altitudes (P78-2 satellite)
SEC	Surface emission cathode
SEM	Scanning electron microscopy
SEP	Solar energetic particle
SPIS	Spacecraft Plasma Interaction Software
SWNT	Single-walled carbon nanotube
TEM	Transmission electron microscopy
UV	Ultraviolet
XPS	*X*-ray photoelectron spectroscopy

Nomenclature

A	Surface area
A_m	Atomic mass
A_N	Atomic number
C	Capacitance
c	Speed of light
d	Thickness
E	Electron (electric field) energy
e	Elementary (electron) charge (1.602×10^{-19}C)
h	Plank's constant (4.136×10^{-15} eV s)
I	Excitation potential
J	Incident electron flux; emission-current density
M	Molar mass
m	Electron mass (9.109×10^{-31} kg)
N_A	Avogadro number
n	Density of free electrons
p	Density of free holes for electrons
R	Maximum distance an electron can penetrate through the material
r	Radius
t	Time
U	Potential
V	Voltage (potential); Nordheim's elliptic function
v	Velocity
β	Surface leakage parameter
δ	Secondary emission coefficient

ε	Permittivity (vacuum permittivity $\varepsilon_0 = 8.854 \bullet 10^{-12}$ Fm^{-1})
ρ	Material density
σ	Electric charge surface density
Φ	Work function of material
Ω	Electrical conductivity

4.1 Introduction

Space charge is a continuum of excess electric charge distributed over an area or a volume. In a conductive environment, space charge is rapidly neutralized. Therefore, space charge occurs in vacuum or dielectric media. Galactic cosmic rays, solar flares, coronal mass ejections on Sun and the Earth's magnetosphere create highly energetic particles (electrons, protons, solar ultraviolet photons, ions, etc.) in space whose potentials may reach millions electron volts [3, 115]. These high-energy charged particles may deposit in or on spacecraft materials and create an electrical field. This electrical field may build up and eventually create an electrical discharge.

The charging is usually classified into the *surface charging (aircraft charging)* and *deep dielectric (internal, bulk) charging*. The surface charging is triggered by low-energy electrons (<100 keV). These electrons cannot breach the material's surface due to their energy deficiency. The mass of the charged particles strongly affects the energy accumulation process. Heavy particles (e.g., protons) convey the energy just below the surface. They may displace nuclei, create phonon excitation and vibrational transitions, and ultimately breach molecular bonds [88]. Lighter particles (e.g., electrons, *UV* photons, ions, etc.) (>200 keV) deliver energy via electronic excitation deeper in the material. Particles with sufficient energy can rupture chemical bonds due to certain reactions [89, 119]. A significant potential difference will be accumulated if the charge deposited in the material is greater than the leaked charges. The level of the electrostatic discharge (ESD) is defined by the amount of the energy E stored in the capacitance C of the object. These two parameters are related as:

$$E = \frac{CV^2}{2}, \tag{4.1}$$

where V is the voltage (potential). The capacitance is defined as:

$$C = \frac{\varepsilon A}{d}, \tag{4.2}$$

where A is the surface area, d is the material thickness between the surfaces, and ε is the permittivity of the material. Material damage takes place when the energy exceeds some critical value, which is in some cases is 10 mJ [76].

Space charge has a destructive and hazardous impact on electronics, navigation systems, and instrumentation onboard spacecraft, as well as the spacecraft itself.

Therefore, it is required to screen and predict the spacecraft charging levels to provide a safe space mission. *Space weather* (temperature, temperature difference, energy levels of charged particles and UV photons, etc.) is mainly controlled by the Sun and depends on the distance from the Earth's center (R_E). For example, geosynchronous (or geostationary) Earth orbit (*GEO*) (6.6 R_E) is characterized with high vacuum, high-energy (5–20 keV) electron and proton fluxes, and *UV* photons. MEO (medium Earth orbit, the domain of navigation satellites) is in the Van Allen radiation belts and can have very high-energy electron (~2 MeV) and proton (~10 MeV) fluxes. And in low Earth orbit (LEO) at altitudes between 200 and 1000 km, the speed of the spacecraft exceeds the ion velocity and creates a wake behind the spacecraft. Paillous [110] distinguished the following major components of the space radiation environment:

- *Cosmic galactic radiation* consists of 90% protons, He^{2+} and heavy ions; energies varying between 10^{-2} and 10^{10} GeV; very low fluxes.
- *Solar wind* consists of 96% protons, He^{2+}, oxygen ions and electrons; energies ~1 keV for protons, 1 eV for electrons; 2×10^8 protons $cm^{-2}s^{-1}$.
- *Solar flare ions* consist of 95% protons, ions; energies varying between 1 and 100 MeV; highly variable fluxes.
- *Earth-trapped protons and electrons*; proton energies <30 MeV in inner belt and <1 MeV in outer belt; electron energies <5 MeV; fluxes depend on altitude, magnetic latitude, solar activity and longitude.
- *Auroral fluxes* consist of protons and electrons; proton energies varying between 80 and 800 keV; electron energies varying between 2 and 20 keV; fluxes depend on solar activity.

Table 4.1 shows spacecraft orbits and space environment conditions on those orbits.

The first empirical models were developed in the 1960s to predict the average space radiation environment during the maximum and minimum phase of the solar cycle. Later, these models were modified and updated [39, 124, 140]. However, the updated models cannot predict the radiation doses for low inclination orbits between 6000 and 12,000 km, as well as the radiation due to the higher energy electrons (>1 MeV) and protons (>40 MeV) created during the geomagnetic storms. Ginet et al. [43] developed a set of models to predict energetic electrons, energetic protons, and the space plasma environment. Although the new models quantify uncertainties due to the space weather, they cannot predict the space weather variations for more than a day.

Cho et al. [17] conducted a statistical analysis of the environmental conditions in polar Earth orbit (PEO) using the data obtained by Defense Meteorological Satellite Program (MNSP) satellites orbiting the Earth at ~840-km altitude. The authors estimated the probability of current densities of aurora electrons and thermal ions assuming an isotropic distribution for the aurora electron velocity.

An overview of the energetic particle radiation hazard on human health was provided by McKenna-Lawlor [90] and McKenna-Lawlor et al. [91]. It is shown that beyond LEO, the major radiation hazards for human health are posed by

Table 4.1 Spacecraft orbits and space environment conditions on space orbits

Orbit	Altitude (km)	Inclination to equator (deg)	Space environment conditions
Polar Earth orbit (PEO)	100–200	>65	Solar ultraviolet
			Ionosphere
			Orbital debris
			Solar flares
Low Earth orbit (LEO)	200–1000	<65	Ionospheric plasma
			Solar ultraviolet
			Orbital debris
Medium Earth orbit (MEO)	1000–36,000	<65	Solar ultraviolet
			Plasmasphere
			Van Allen radiation belts
Geostationary (Geosynchronous equatorial) Earth orbit (GEO)	35,786	0	UV radiation
			Van Allen radiation belts
			Solar flares
Interplanetary orbits	>65,000	–	Solar flares
			Cosmic radiation
			Solar plasma

galactic cosmic radiation (GCR) and by solar energetic particles (SEPs). Because of the rapid passage of lunar astronauts through the Van Allen Radiation Belts, they present a lesser radiation hazard. The authors concluded that reliable forecasts of dangerous particle radiation levels are required to manage human health risks during interplanetary space missions.

Engelhart et al. [26, 27] identified the following effects due to the electrostatic discharges (arcs) of spacecraft:

- Huge and fast current spikes, which can destroy electronics
- Transients in spacecraft power
- Electromagnetic interference with communication systems
- Deterioration of physical (e.g., optical, electrical, etc.) properties
- Power loss due to short-circuits between solar cells
- Total power loss due to arcing from solar array strings to spacecraft ground

McKnight [92] analyzed manmade aspects of the spacecraft failures. The aspects include design, manufacture, integration/installation, parts quality, testing completeness, and operations. The author provides instructions based on some space operational examples.

4.2 Spacecraft Charging Effects on Dielectric Materials

Near future space exploration requires much more space power than it is presently available. Dielectrics and electrical insulations are key components of space electrical systems. Polymers are the most generally used dielectrics due to their reliability, availability, simplicity of fabrication, and cost [18]. Usually, dielectrics are non-conducting coatings on the conducting substrate. The electric field inside this thin dielectric coating layer is considered as homogeneous due to its small thickness as compared to the surface size. Then a normal component of the electric field can be expressed as [109]:

$$E_n = \frac{\sigma_d}{\varepsilon_0} + \frac{\varepsilon}{d}\left(U_s - U_d\right), \tag{4.3}$$

where σ_d is the electric charge surface density on the dielectric coating; U_s and U_d are the substrate and the dielectric coating potentials, respectively; ε the permittivity of the dielectric coating; and ε_0 the vacuum permittivity ($8.854 \cdot 10^{-12}$ Fm^{-1}).

If high-energy charged particles (electrons, ions) penetrate dielectric materials, they may remain inside for a long time due to the low conductivity of the material. A discharge takes place once the charge reaches some level. The critical electric potential for dielectric failure depends on the material geometry and properties (bulk and surface electric conductivity, dielectric permittivity, secondary electron emission yield, radiation induced conductivity, aging), as well as the space weather (electron energy, solar activity, etc.). The critical electric field for dielectric failure is $10^6 - 10^8$ V/m [77]. The energy loss per unit depth x of penetration of charged particle (ions, protons, alpha particles) with speed v, charge Z in multiples of the electron charge, and energy E into the substance is defined by the Bethe-Bloch equation which was derived using quantum mechanical perturbation theory [152]:

$$\frac{dE}{dx} = -\frac{4\pi Z^2 n_e}{mv^2}\alpha^2 \cdot \left[\ln\left(\frac{2mv^2}{I \cdot \left(1-\beta^2\right)}\right) - \beta^2\right], \tag{4.4}$$

where $\beta = \frac{v}{c}$; $\alpha = \frac{e^2}{4\pi\varepsilon_0}$; $n_e = \frac{N_A \cdot A_N \cdot \rho}{A_m \cdot M}$ is the electron number density; c is the speed of light; $I = (10\text{ eV}) \cdot A_N$ is the mean excitation potential; e is the electron charge; m is the invariant (rest) mass; ε_0 is the vacuum permittivity; N_A is the Avogadro number; A_N is the atomic number; ρ is the material density; A_m is the relative atomic mass; and M is the molar mass constant. NIST [107] maintains tables of penetration depth for energetic charged particles into materials.

Electron irradiation can generate substantial potential differences at material surfaces, which can cause anomalies, such as electromagnetic disturbances, damage of electronics, and modification of electric properties of the dielectric material. The Bethe-Bloch Eq. (4.3) is not valid for the charged atomic particles with electrons due to the small mass of the electrons. Therefore, for smaller energies, Eq. (4.3) needs to be modified. One of the modifications was made by Barkas and Anderson

who proposed Z^3 in formula (4.3). Another correction was proposed by Bloch who suggested Z^4 in that formula. The corrections are required also at very high energies as well (e.g., Fermi's density correction). Most dielectric materials (e.g., polymers) become conductive (*radiation-induced conductivity – RIC*) for some time after irradiation due to the ionization process. Therefore, some corrections need to be done in Eq. (4.3) for electrons which pass through the material (e.g., "shell" correction).

Latham [78] showed that a dielectric impurity (with surface area between 10^{-9} and 10^{-13} m^2) on the metal surface may increase the electric field several thousand times as a result of the transition of electrons from the metal surface to the conduction band of the dielectric inclusion where electrons are accelerated and create holes due to the ionization. These holes migrate to the metal-dielectric interface and enhance the electric field. Finally, the electrons emit from the dielectric-vacuum interface. The point where the dielectric, conductor, and adjacent plasma come together, a favored place for electrical discharges, is called the "triple junction."

Due to the combined effects of radiation and thermal cycling on the dimensional stability of polymer composites, the latter will plasticize at high temperatures and cause permanent residual strains. Low temperatures will develop micro-cracks and additional residual strains in the polymer composites [120].

Czepiela et al. [22] developed a simple one-dimensional finite-difference model of deep dielectric charging in composite materials. The model predicts the electric field in space and time as a function of the material (organic polymer) properties and environmental conditions (7000-km altitude orbit). The model predicted a steady-state electric field ~4400 V/m and a development time ~ 1.3 s. Varying charged particle fluxes and material properties, the authors concluded that typical flux levels will not form significant deep electric charging in usual composite materials. However, many commonly used spacecraft materials (such as printed circuit board materials like FR4 fiberglass) store charge for days, weeks, or months. The ability of Teflon to store charge for long periods of time led to NASA-TM 2361 [118] to recommend that it not be used on spacecraft surfaces or in the interior where charges may accumulate.

In the United States, the most commonly used code for deep dielectric charging (NUMIT) was developed by A.R. Frederickson [35–37], and variants of it have been produced for Air Force and JPL use.

Some researchers characterize the material properties of human-made space objects (e.g., satellites, rockets, debris, etc.) using astronomical reflectance spectroscopy [1, 12, 102]. They collect light reflected from those objects illuminated by the Sun. The shape and absorption features of the spectrum emitted by reflected light provide the information on material's composition.

Novikov et al. [108] analyzed the results of magnetospheric plasma fluxes obtained Russian "Electro" (GEO; 1995–1997), "Meteor" (LEO; 2003), and "Universitetskiy-Tatyana" (LEO; 2005) spacecraft. In their analyses, the electrons spectra were divided into four groups by mean energies: (1) below 2 keV, (2) between 2 keV and 4 keV, (3) between 4 keV and 5 keV, and (4) above

5 keV. Using the Coulomb mathematical model, the authors showed that the data obtained at "Electro" spacecraft can be described by a two-temperature Maxwellian approximation function.

Craven et al. [21] reported a new type of event – a high negative amplitude rapid charging event (RCE) detected on the International Space Station (ISS) in LEO altitudes by the floating potential measurement unit (FPMU). RCEs have very short rise time for charging ($t < 10$ s), which could not be detected by the original floating potential probe (FPP), which took data every 0.1 s and analyzed them in 20-s increments. RCEs survive for less than 30 s and have a faster rise time than decline time. Ferguson et al. [30] proposed a model to describe RCE ignoring the effects of the Earth's magnetic field. The authors compared the results of simulations using the Environments WorkBench code (EWB) to the events observed on ISS. It was concluded that RCEs are related to charging caused by the ISS high-voltage solar array coming out of eclipse.

Liu et al. [80] simulated the space charging process of the Astrodynamical Space Test of Relativity using Optical Devices (ASTROD I) test mass due to solar energetic particles, interplanetary electrons, and heavy nuclei using the GEANT 4 simulation toolkit, which employs Monte Carlo particle ray-tracing model [2, 46]. ASTROD I utilizes two-way laser interferometric ranging and laser pulse ranging between a drag-free spacecraft in a solar orbit and the deep space laser stations on Earth [106]. The test mass is a 1.75-kg rectangular parallelepiped Au-Pt alloy with a very low magnetic susceptibility to minimize magnetic disturbances. It is located in the center of spacecraft to protect it from external dynamic impacts. According to the model predictions, a charging rate due to solar energetic particles (SEPs) is much greater than that due to the galactic cosmic radiation (GCR). The authors acknowledged $\pm 30\%$ error in the net charging rates due to the spectra, physics models, and geometry applications in addition to the Monte Carlo model uncertainties.

Bao [11] used the GEANT 4 simulation toolkit to predict the charging process of the ASTROD I test mass by SEPs. The simulation results show that the net charge accumulated on the test mass will affect its drag-free motion through imitation Coulomb and Lorentz forces.

Davis et al. [23] compared the true 3D computer codes NASCAP-2K (US), MUSCAT (Japan) and SPIS (ESA) for the low temperature, dense plasma environment in LEO. They demonstrated a good agreement between the different approaches of the program codes.

Ferguson and Wimberly [31] compared a variation of potential of the metal ground in both Single and Double Maxwellian hot plasmax with the plasma temperature and electron current density obtained using Coulomb-2, NASCAP-2K and MUSCAT program codes. The obtained results showed an increase of the potential with both temperature and electron current density. An adequate agreement for Coulomb-2 and NASCAP-2K codes was obtained. However, the MUSCAT program code gave lower values of potential as compared to two other codes.

Ferguson et al. [32] originated a "spacecraft charging index," the total electron flux above about 9-keV energy, as a shorthand way of characterizing surface charging environments. It is based on the fact that low-energy electrons discharge surfaces through secondary electron emission.

The SEM analysis conducted by Mundari et al. [101] showed that the exposure to atomic oxygen changes the surface morphology and the surface structure of silicon doped polyimide. The XPS analysis showed an increase of the oxygen concentration and a decrease of carbon on the surface of the test material, which consequently would affect its electrical conductivity.

Novikov et al. [109] discussed numerical modeling of spacecraft charging in GEO using various program codes (Coulomb-2 (Russia), NASCAP-2K, MUSCAT, SPIS) and conducted simulations for similar spacecraft models with identical surface materials (graphite, NPaint) and space plasma environments. Some discrepancies in obtained results using various program codes are explained by the differences in application of different procedures for solution of the equations describing the charging process.

Paulmier et al. [114] conducted experimental studies of charging behavior of some polymers used in space: polyetherimide (PEI) sheets, polyetheretherketone (PEEK) sheets, ethylene tetrafluoroethylene (ETFE) wires, polyimide Kapton® wires, polyurethane (PU) (Solithane®) films, silicone (MAPSIL®) films and polyurethane PU1 (black paint) films. The test results showed that except for PU1 (black paint), all other materials charge under 20-keV electron bombardment. PEI, PEEK, ETFE, and Kapton® demonstrated a very low conductivity. At higher radiation doses, Kapton® wires became more conductive. Both the radiation dose and the duration of the exposure to the radiation affect the conductivity of tested materials. The authors also developed a physical/numerical model to describe a generation of surface potential and RIC of materials exposed to high-energy electrons. The irradiated sample is considered as a system of a capacitance and an electrical resistance connected in parallel. Then a change of the surface potential in time can be expressed as:

$$\frac{dV}{dt} = \frac{dJ\,(1-\delta-\beta) - e\left(\mu_n n + \mu_p p\right) V}{\varepsilon} \overrightarrow{AB}, \tag{4.5}$$

where d is the material thickness, J the incident flux of low energy trapped electrons (20 keV), δ the secondary electron emission yield, β the surface leakage parameter, V the surface potential, e the elementary charge, ε the material permittivity, n density of free electrons, and p density of free holes for electrons. The authors found that the numerical results are in good agreement with the experimental data.

A charge-discharge process of electron-irradiated polyimide (PI) is divided into three stages [27]:

I. *Charging* – electron deposition, secondary electron generation, radiation-induced conductivity (RIC), thermally assisted hopping (dark current) conductivity

II. *Pre-transit discharge* – RIC, thermally assisted hopping (dark current) conductivity
III. *Post-transit discharge* – thermally assisted hopping (dark current) conductivity

Sim [128] proposed the following unified model to predict the variation of the surface potential with time at stage I of the charging-discharging process in polymers:

$$V(t) = \frac{edn}{\varepsilon_0 \varepsilon_r}\left(1 - \frac{R}{d}\right) R\left\{1 - \exp\left[\frac{A_c J \tau_{\text{onset}}(1 - n_s)}{e(1-m)}\right]\left[1 - \left(1 + \frac{t}{\tau_{\text{onset}}}\right)^{1-m}\right]\right\}, \tag{4.6}$$

where ε_0 and ε_r are the permittivity of vacuum and relative permittivity of the material, respectively; J is the electron beam flux; n_s is the number of secondary electrons emitted per incident electron; R is the maximum distance an electron can penetrate through the material; A_c is the capture cross-section; τ_{onset} is the characteristic onset time for the current decay; and m is the power parameter ($0 < m < 1$). Song et al. [133] provided a model for the number of secondary electrons emitted per incident electron.

One of the mathematical models to describe the stage II of the charging-discharging process was proposed by Toomer and Lewis [137]:

$$V(t) = V_0\left\{1 - \frac{V_0 \mu_0}{2d^2 \varrho}\left[r_t t + \frac{r_t}{\varrho}\left(1 - e^{-\varrho t}\right) + \frac{r_r a_1^2}{2\alpha}\left(1 - e^{-2\alpha t}\right) - \frac{r_t a_1^2}{\varrho + 2\alpha}\left(1 - e^{-(\varrho + 2\alpha)t}\right)\right] - \lambda t^{a_2}\right\}, \tag{4.7}$$

where V_0 is the initial surface potential; ϱ is the charge transport dynamics parameter; r_r and r_t are the probabilities of released and trapped charging rates, respectively; λ is the dispersive time; μ_0 is the carrier mobility between traps; a_1 and a_2 are fitting parameters; and α is the rate of charge moved from the surface to the bulk.

The surface potential at the stage III of the charging-discharging process was predicted by Dennison et al. [24] as:

$$V(t_d) = b_1 t_d^{b_2}, \tag{4.8}$$

where b_1 and b_2 are fitting parameters and t_d is the surface potential decay time. Then the dark resistivity of the material can be defined as:

$$\sigma_{\text{dark}} = \frac{t_d}{\varepsilon_0 \varepsilon_r}. \tag{4.9}$$

The most common materials used for the exterior of spacecraft to control temperature in the presence of sunlight are multi-layers of aluminized (or silvered) polymers, e.g., polyimide (PI, Kapton-H®). These so-called MLIs (Multi-Layer Insulations) are in the "preferred class" according to the classification of polymers by their radiation stability provided by Frederickson et al. [37].

Engelhart et al. [27] investigated the chemical and physical properties of PI during and after irradiation with high-energy electrons (90 keV) using a space weather simulation facility. The following simultaneous kinetic processes causing chemical structure modifications due to the interaction of PI with space radiation have been identified:

- Damage: broken chemical bonds due to the interaction of PI with highly energetic particles
- Healing: creation of new bonds similar to damaged ones
- Scarring: formation of new chemical bonds different from those in the original material

The analysis of the transmission spectra revealed a damage of the optical properties of PI samples due to the radiation. The chemistry of the radiation-induced damage was evaluated by using Fourier-transform infrared (FTIR) spectroscopy. The measurements of the bulk conductivities of irradiated PI samples showed a rapid (~3 h.) recovery of material's conductivity due to the air exposure, while the recovery in vacuum takes more than 3 weeks [25]. The authors concluded that air exposure governs the post-irradiation chemistry of materials.

Plis et al. [117] provided an extensive review of the published information over the 40 years on the effect of electron, proton, and UV photon irradiation at geosynchronous altitudes on mechanical, absorption, and charge transport properties, as well as morphology of polyimide (PI, Kapton-H®).

Polytetrafluoroethylene (PTFE) (Teflon®) – low conductivity polymer. The post-transition region of the surface potential decay curve (stage III) starts when a charged body has reached the grounded back plane, and it takes years for PTFE [27]. It is shown that a decay time and the dark conductivity of the material are well described by the empirical expressions (10 and (11) proposed by Dennison et al. [24].

PTFE has one of the lowest coefficients of friction of any solid matter. Some thermophysical properties of PTFE are given in Table 4.2.

Table 4.2 Thermophysical properties of PTFE

Density	2200 kg/m^3
Melting point	327°C
Thermal expansion	$112 - 125 \times 10^{-6}$ K^{-1}
Thermal diffusivity	0.124 mm^2/s
Young's modulus	0.5 GPa
Yield strength	23 MPa
Bulk resistivity	$10^{18}\ \Omega \cdot$ cm
Coefficient of friction	0.05–0.10
Dielectric constant	$\varepsilon = 2.1$; $\tan(\delta) < 5 \times 10^{-4}$
Dielectric strength (1 MHz)	60 MV/m
Magnetic susceptibility (SI, 22°C)	-10.28×10^{-6}

Frederickson et al. [37] provided a classification of polymers by their radiation stability. According to their classification, PTFE (Teflon®) is in the "not recommended class" due to the properties in the presence of oxygen and electrical loss properties in air and vacuum. However, it may be used in a vacuum if it is not subjected to physical stresses and in some application where dielectric loss is not important. However, PTFE is not recommended in applications where charge storage and subsequent deep dielectric discharges may be a problem [118].

Laghari and Hammoud [67] reviewed radiation effects on polymer dielectrics most commonly used for power transmission and storage (e.g., PI Kapton®, PTFE Teflon®, PMMA Plexiglass, Polyamide Nylon, PVC, etc.). The authors analyzed the effect of the type of radiation, dose rate, and exposure time on the electrical (dielectric strength, dielectric constant, dielectric loss, radiation-induced conductivity, volume and surface resistivity), mechanical (elastic modulus, tensile strength), and physical (density, weight loss, optical transmission, color, melting point) properties of polymer dielectrics. A threshold level of deterioration upon irradiation is indicated as one of the material characteristics. It is defined as the radiation dose required to cause a 25% deterioration of material property.

Hanna et al. [47] suggested the following mechanism for radiation process in Teflon® fluorinated ethylene propylene (FEP), which demonstrates high charging potentials due to a steep rise of electric induced resistivity with increasing radiation dose:

- Generation of electron-hole pairs through ionization by the high energy electrons
- Deep trapping of electrons
- Long life time for free holes
- High recombination rate between electron-hole

Arnaout et al. [5] experimentally confirmed these suggestions on radiation-induced conductivity processes in Teflon® FEP under GEO conditions using a multi-energetic irradiation vacuum chamber SIRENE facility installed at the French Aerospace Lab ONERA. The facility was equipped with a non-contact pulsed electro acoustic (PEA) device to analyze the evolution of charge distribution.

4.3 Space Charge Behavior of Dielectric Nanocomposites

Early studies contributed the specific behavior of dielectric materials to the existence of spots (e.g., impurities) of higher conductivity than the matrix [127]. A presence of semi-conducting particles increases the apparent conductivity of the dielectric material in an alternating electric field from that in a steady field, and it will be a function of the field frequency. One of the earliest models was proposed by Wagner [141], assuming that the conducting particles embedded into the dielectric matrix are spheres sporadically dispersed in the dielectric material. Experiments conducted with a suspension of water droplets in wax [127] did not agree with Wagner's model

predictions. However, a good agreement was found at temperatures near to the melting point of the matrix (wax). An analytical model developed by Sillars [127] for an inhomogeneous matrix with spheroidal particles showed that a small amount of conducting particles in the shape of fine needles will cause a bigger potential loss than spherical shape particles of the same quantity.

Although the embedded particles can improve thermal and electrical conductivities, enhance discharge resistance, as well as control thermal expansion and glass transition temperature of the dielectric material, they do not affect significantly its electric strength. It is assumed that the embedded particles generate a layer of "immobilized" polymer which controls the electrical, mechanical, and thermal properties of the composite [105]. Due to the created interfaces, the embedded particles may increase the space-charge buildup and the Maxwell-Wagner polarization and allow automatic and predictable self-assembly of the nanocomposite.

The idea of improving the electrical properties of dielectric polymers by dispersing CNTs into polymer matrix has attracted significant attentions of recent studies [81, 135, 143, 149]. Also, by monitoring the electrical conductivity of the polymer nanocomposite, it is possible to measure its strain, which is very useful in health monitoring of spacecraft structures [16, 112, 148].

The pulsed electro-acoustic space charge measurements conducted by Nelson and Fothergill [105] showed that addition of TiO_2 nanoparticles (~23 nm) into a bisphenol-A epoxy significantly changes the behavior of the nanocomposite as compared to the micron size (~1.5 μm) TiO_2 particles imbedded into the same epoxy matrix. The maximum electric strength of epoxy/TiO_2 nanocomposite was achieved at 10 wt.% of TiO_2 nanoparticles.

Roy et al. [122] experimentally studied dielectric properties of both SiO_2-polyethylene micro- and nanocomposites. Embedding silica nanoparticles into polyethylene increased the breakdown strength and voltage durability considerably as compared to the micron size silica fillers. The dielectric spectroscopy showed a decrease in dielectric permittivity for the SiO_2-polyethylene nanocomposite over the matrix polymer. Also, embedded silica nanoparticles impact the space charge distribution and dynamics. The bonding between the silica and polyethylene was studied using Fourier-transform infrared (FTIR) spectroscopy, electron paramagnetic resonance (EPR), and X-ray photoelectron spectroscopy (XPS). It is concluded that both the enhanced interfacial area and particle-matrix bonding are significant in the dielectric behavior of nanocomposites.

Huang et al. [52] conducted SEM and TEM analyses, as well as dielectric and mechanical properties of copper phthalocyanine oligomer (*o-CuPc*) nanoparticles embedded into the polyurethane (*PU*) matrix by the combination of *top-down* and *bottom-up* self-assembly approaches and creating three-dimensional networks. The results of the experimental tests showed that the self-assembled nanocomposite with low volume fraction of the nanofillers (~3.5 vol.%) exhibits much higher strain and elastic energy density, and much lower dielectric and electromechanical responses as compared to the simple blend of *o-CuPc* and *PU* composite with higher content of nanofillers (~16 vol.%).

Lopez-Pamies et al. [82] developed an analytical model for the homogenized (i.e., macroscopic) dielectric behavior of nanoparticulate composites having interphasial space charges between nanoparticles and matrix. The model predicts a microscopic isotropic permittivity of the composite which involve the matrix material filled with polydisperse spherical particles containing space charges at the interphase. In the absence of particles, the proposed model reduces to the Maxwell-Garnett model. In the absence of charges, the model reduces to the model proposed by Milton [94].

Ghosh et al. [42] proposed the homogenized equations for the macroscopic response of time-dependent polymer composites with space charges varying spatially under alternating electric fields. The model considers three-phase dielectrics with isotropic behavior that consists of a matrix filled with spherical particles. These particles are connected to the matrix via active space charged interphases, and each element has different homogeneous isotropic permittivity. The predicted effective complex permittivities (with and without interphases and space charges) were compared to those experimentally obtained earlier for various polymer matrices (*polyetherimide, epoxy, polyurethane*) with different types of fillers (*copper phthalocyanine oligomer o-CuPC*, Al_2O_3, TiO_2). It is shown that the predictions with consideration of both interphases and space charges are in good agreement with experimental data published earlier.

Kutelia et al. [64, 65] proposed a modification of the PTFE matrix with micro-/nanoparticles which may allow developing new composite materials with good space charge behavior, as well as outstanding physical, mechanical, and functional characteristics. The reviews of the studies in the last decade [4, 40, 150] show that a friction and wear performances of PTFE may be improved by micro- and nanoscale inorganic fillers. The shape, size, aspect ratio, composition, volume fraction, and specific area of added particles/nanoparticles are important factors that affect the complex mechanical properties of the newly synthesized polymer-based composites. Along with the development of nanoparticles production technology during the last decade, the interest in them as the PTFE matrix modifiers have also considerably raised. Review of data regarding the modification of PTFE matrix by ceramic [61, 136, 151] and carbon nanoparticles [15, 59, 63, 66] showed that the optimal concentrations of the nano-size filler particles of both types are 2–10 vol.%, while an optimum amount of micron size fillers in most typical polymer composites is ~30 vol.%. The previous studies also showed that in developing the high-performance tribological material (e.g., advanced PTFE-based antifrictional/frictional nanocomposite materials) with high wear resistance the following two practically conflicting requirements need to be satisfied. First, the material must be soft with extremely low shear strength, and second, it must be strong enough to support substantial static and dynamic loads. Among the existing polymers, the PTFE is considered as a most promising basic material for developing the composites addressing both requirements due to its low friction coefficient (<0.1) and very big value of the ratio of the bulk shear strength ($\tau = 20$ MPa) to the surface shear strength ($s = 4$ MPa) that is equal to 5, while the same ratio for the polyethylene does not exceed 1.4 and less than 1 for all other existing polymers

[98]. Therefore, one would expect that the PTFE-based carbon nanocomposites would demonstrate good heat and electric conductivities with high strength, which is impossible for composites with ceramic nanoparticles. The studies [15, 59, 63, 66] have already revealed a very high potential of carbon nanoparticles (CNTs, CNPs, CNBs, CNWs) as fillers of the PTFE matrix for the production of multi-functional nanocomposites for the broad range of applications. The advantage of these nano-size fillers consists in the possibility of their incorporation in the polymer matrix along with the micro-size fillers without any chemical modifications [4, 79]. However, there is a lack of information concerning the application of CNF-based hybrid nanoparticles for modifying the PTFE in order to develop high-performance nanocomposite materials which cannot be obtained by using of single CNF fillers. Therefore, the core-shell-type nanoparticles of CNFs (CNTs, CNPs, CNBs) doped with the ferromagnetic atom (Fe, Co, Ni) clusters [62, 63, 66, 123] as the hybrid fillers for the PTFE matrix is quite attractive. A comparative study of inelastic/elastic behavior and physico-mechanical properties of PTFE and PTFE-based nanocomposite material filled with different amounts (2.5, 5.0, 7.5 and 10.0 wt.%) of monomodal and bimodal nanopowders (Fe atom cluster doped CNTs) has been performed by Kutelia et al. [64, 65] using a low-frequency amplitude-independent and amplitude-dependent internal friction measurement techniques. Variations of activation energies, frequency factors of β (crystalline) and α (amorphous) relaxation processes and critical amplitudes of micro-plastic deformation beginning of the PTF and PTFE-based nanocomposite materials with the concentration and mode of the Fe atom cluster-doped CNT fillers were determined from experimentally measured $Q^{-1}(T, \varepsilon, f)$ and G (T, ε, f) spectra. It is shown that the bimodal Fe atom cluster-doped CNT filler provides a more significant and complex effect on the PTFE matrix structure (ratio of β- and α- phases) and thermodynamic parameters of β- and α-phases (activation energies) that consequently lead to ~25% additional increase in the critical amplitude of micro-plastic deformation beginning (ε_c) in the synthesized PTFE-based new nanocomposite material.

Analytical model of the dielectric material with embedded nanoparticles doped with conductive atom clusters containing the interphase space charges has been developed by Bakhtiyarov and Ferguson [9]. A combined solution of the permittivity equation of the heterogeneous system with the Gauss-Codazzi and the electric potential continuity equations is allowed to simulate the relative permittivity of the composite dielectric material as a function of the concentration of the embedded nanoparticles, as well as the space charge content. The simulations for the polytetrafluoroethylene (*PTFE*) dielectric material with embedded carbon nanoparticles doped with *Fe* atom clusters showed that the relative permittivity of the *PTFE* nanocomposite increases with increasing the amount of the embedded CNPs doped with Fe clusters and the charge content. In the absence of embedded CNPs, the solution reduces to the Maxwell-Garnett equation.

Improvements in dielectric properties of polymeric nanocomposites are attributed to various factors, such as [122]:

- The high surface area-to-volume ratio of nanoparticles creates a larger interfacial area (*interaction zone*) for the same particle loading [6].
- The nanoparticles' surfaces change in the morphology of semi-crystalline polymers [84, 85].
- The small size of particles reduces the internal field.
- Changes in the space charge distribution [58, 97, 104].
- Scattering mechanism, i.e., an increase in path length of the carriers controlling the breakdown processes [38].

When measuring the bulk resistivity of inhomogeneous materials, it makes a difference whether one is interested in resistivity at a constant voltage (circuit resistance) or in charge storage time (spacecraft charging). Plis et al. [116] have shown that in carbon-loaded Kapton, the two methods can yield answers that disagree by orders of magnitude, depending on the amount of suspended carbon.

4.4 Mitigation Methods

Several surface charging mitigation techniques were proposed in the last decades [75, 103]. All the mitigation methods can be divided into two categories: active and passive.

4.4.1 Active Mitigation Methods

The active type of mitigation can be controlled by command. One of the active surface charging mitigating techniques proposes an emission of positive ions, plasma, or polar molecules [68]. Emission of plasma (mixture of low-energy ions and electrons) is an effective active mitigation technique, which combines the benefits of both ion and electron emission. The technique was tested on the Defense Satellite Communications Systems (DSCS) at GEO altitudes [100]. In these experiments, the plasma (ionized xenon gas) of energy less than 10 eV was tested on two dielectric specimens (Kapton and quartz) and promising results were obtained. This method is recommended to control the spacecraft potential [138] and has been implemented on the ISS [14].

Grard [45] developed an electron field emission probe to transport electrons from the spacecraft into space, thus significantly decreasing the negative charge of the spacecraft ground. However, it cannot reduce the dielectric surface potential, which develops substantial differential charging between the conducting aircraft ground and the dielectric surfaces.

Lai and Murad [71, 73, 74] proposed a *vaporization* technique to mitigate surface charging. In this method, the surface charging can be controlled by spraying polar molecule liquid (e.g., ionized water) droplets over spacecraft surfaces. The droplets

evaporate after attaching the electrons on the spacecraft surfaces and remove the excess electrons. However, this technique is applicable for deep dielectric charging. It also may cause contamination problems.

In the *hot-filament emission* mitigation method, the hot filaments emit electrons according to Richardson's thermionic emission principle [121] below the melting points of the materials. However, above the melting points of the materials, the hot filaments will radiate both neutrals and ions. The method may result in differential charging since electron emission can lessen the charging level of the spacecraft ground but not the dielectric surface. Also, due to the low energy of thermal electrons, the emitted current is limited by space charge saturation close to the filament [75].

Emission of low-energy positive ions was tested for mitigation of negative surface potential on the *Spacecraft Charging at High Altitudes* (SCATHA) (P78-2 satellite) at GEO altitudes [19, 68]. The potential of the spacecraft body can be affected by photoelectrons from neighboring surfaces secluded from it. These photoelectrons will return to the spacecraft body. The emitted ions neutralize the secondary electrons. They are effective for both conductive and dielectric surfaces and depend on the material properties (e.g., secondary electron emission coefficient). However, this mitigation technique can only work until the spacecraft potential energy is greater than the initial energy of the ions emitted from the spacecraft, and this was proved by numerical simulations [72, 139, 142]. Therefore, a low-energy plasma beam emission is considered a more efficient mitigation method than electron or ion beam techniques.

4.4.2 Passive Mitigation Methods

The passive type of mitigation does not require any control. Some passive methods are based on using some surface materials with high secondary emission coefficient. Again, it is limited by the critical value of the electron temperature. One of the passive mitigation techniques is a *sharp spike* method where an intense electric field can be created by sharp spikes located on charged surfaces [75]. The generated field energy is proportional to the tip's radius of curvature as: $E \sim r^{-2}$. Consequently, the field electron emission will reduce the negative potential of the conductive surfaces joined to the spike. However, in this mitigation method, the electrons will be removed only from the conducting spacecraft ground and differential charging will take a place. Also, the tip's high electrostatic field can attract the surrounding positively charged ions and blunt the tip [75].

Manners [86, 87] proposed metalized dielectrics to mitigate deep dielectric charging. In order to achieve homogeneous distribution of the conductivity of the composite material, the author introduced metal atoms in the molecular level inserting metal atoms by opening the rings of dielectric polymer molecules, rather than in the lattice level. The experimental studies conducted by Balmain

[10] demonstrated an effectiveness of the metallized polymers in mitigating deep dielectric charging.

Another passive mitigation method is the application of surface materials (coatings) with high secondary emission coefficient ($\delta_{\max} \gg 1$) (e.g., copper-beryllium alloy). A secondary emission coefficient is defined as a ratio of the number of departing secondary electrons per unit arriving electrons. It depends on the surface material type and the roughness and increases with the arriving electron energy until it reaches a maximum value. The method also was tested on SCATHA mission and the results are reported by Lai [69, 70].

To remove excess electrons from spacecraft surfaces, a field emission method was proposed by several investigators. One of them is a *diamond surface-emission cathode* (SEC) [41]. The model is based on the combination of the enhanced electric field of a triple junction at the intersection of metal and diamond interfaces in vacuum with the negative electron affinity (NEA) of the diamond surface. From this surface, the accelerated electrons will be emitted into the vacuum. The authors found many similarities in electrical properties of diamond and non-diamond (organic salts, lithium fluoride, aluminum nitride, calcium fluoride, boron nitride, cesium-doped glass, etc.) cathodes:

- Electrons are emitted with a narrow energy spread.
- Addition of cesium (or cesium salts) increases the emission.
- Fraction of the current across the surface is emitted into vacuum.
- Relationship between the emission current and the voltage across the cathode can be described by the Fowler-Nordheim equation:

$$J = \frac{eE^2}{8\pi h\Phi} e^{-\frac{4}{3}\sqrt{\frac{8\pi^2 m}{h^2}}\Phi^{\frac{3}{2}}\frac{V(E,\Phi)}{e^{\frac{3}{2}}E}}, \tag{4.10}$$

where J is the emission-current density ($A\ \mathrm{m}^{-2}$); E the local electric field energy ($V\ \mathrm{m}^{-1}$); Φ the work function of the material (eV); e the electron charge (1.602×10^{-19}C); h Plank's constant (4.136×10^{-15} eV s); m the electron mass (9.109×10^{-31} kg); and $V(E, \Phi)$ Nordheim's elliptic function.

Cooke and Geis [20] proposed the passive anode surface emission cathode (PASEC) where emission ensues over a larger surface area. The PASEC uses a triple junction between metal electrode, insulator, and vacuum. It generates high electric fields that transfer electrons from metal to vacuum. The authors propose to use the new technology in space applications such as electrodynamic tether (EDT), passive control of spacecraft charging, and the neutralization of ion propulsion engines.

The experimental studies showed that at potentials as low as −60 V extravehicular mobility units (EMU) (spacesuits) may arc [14, 125] and some parts of ISS may float to these potentials under certain plasma environment [29, 99]. Therefore, plasma contacting units (PCUs) were used in ISS design in order to lessen the floating potentials by electron emission. Additionally, some passive potential control

methods also were applied during ISS missions. Ferguson [28] discussed both advantages and disadvantages of possible alternative passive and active potential control techniques to the ISS PCUs:

- Passive methods:
  - Shunting
  - Turning the arrays into their own wakes
  - Adding grounded ion-collecting area
  - Preventing arcs by ameliorating the hazardous conditions
  - Grounding positive end of the solar arrays
- Active methods:
  - Other hollow-cathode electron emitters
  - Other electron-emitting devices
  - Charging-monitored array shunting

In *conduction grid* passive mitigation technique [75], a dielectric surface (e.g., solar cell) is covered with a conducting wire mesh to provide a uniform potential distribution all over the area. However, between the wires and the dielectric surface area, the potential differences may be generated periodically.

Tahara et al. [134] conducted ground-based experiments on a plasma contactor with a hollow cathode which is proposed for both electron emission and collection [113, 147] in LEO plasma environment. The experimental results demonstrated that the presence of the simulating LEO plasma intensively raised the collection current compared with cases without the simulating LEO plasma and hollow cathode discharge. The current emission characteristics of a hollow cathode are important as a plasma contactor with a hollow cathode in space.

Iwata et al. [54] proposed an electron-emitting polyimide-copper laminate film (ELF) to mitigate surface charging. In this technique, the device is connected to the conducting surface emitting electrons in order to lessen the negative voltage of the surface. This method does not require any power source as the power is provided by the electric field generated by the surface charging and charging of spacecraft capacitance with respect to the ambient plasma. Also, the surface charging automatically starts the field emission and there is no need for any sensor to initiate its operation. ELF cathode can be easily installed as it does not involve any wire connectors and working gas to create electrons. Its total mass is about 10 g.

Deep dielectric charging can be mitigated by application of finite conductivity materials. It is suggested to use some filters to reduce electromagnetic wave interference. According to some studies [48], the partially conducting (finite conductivity) materials can mitigate internal charging. Shielding is another way to inhibit the internal charging. High-energy charged particles can entirely pass through the very thin (micro- or nano-) device without any buildup. And, high bulk-conductive materials can dissipate surface and dielectric charging before it reaches an arc threshold. Purvis et al. [118] recommend bulk resistivities of 10^{10} ohm-cm or less.

According to some concepts, an application of the partially conductive coatings (e.g., indium oxide, indium-tin oxide, zinc ortho-titanate, alodyne, etc., [118]) or polymers [37] can prevent a spacecraft from charging without any control. However, at the same time the applied coating will isolate the instrumentation inside the spacecraft from the space environment. Most dielectric films are essentially non-conductive materials, and they will be charged like a capacitor in space environment. The film can discharge in certain circumstances and damage nearby equipment and materials. According to the estimations, in order to avoid electrostatic charge build-up, the dielectric film must have a surface resistivity in the range of $10^6 - 10^{10}\Omega$/sq [132] without weakening other physical and mechanical properties of the material.

Application of partially conductive paints or coatings (e.g., indium oxide, indium-tin oxide, zinc ortho-titanate, alodyne, etc., [118]) allows maintaining a low solar absorptivity and high optical transparency of the substrate dielectric material. However, these conductive coatings may be brittle and in the space environment can be easily fractured. An alternative to conductive coatings is an assimilation of homogeneously dispersed conductive additives in the dielectric matrix. Over the last two decades, an application of single-walled carbon nanotubes (SWNTs) due to their high aspect ratio and transport properties received special attention of many researchers as conductive additives to polymeric materials used in space applications. Nevertheless, due to their insolubility, it was challenging to obtain homogeneous dispersion of SWNTs in the solution. Different approaches were initially reported to achieve a homogeneous dispersion of SWNTs in polymers:

- In situ polymerization with simultaneous ultrasonic treatment approach [111, 129, 144].
- Addition to amide acid polymers terminated with alkoxysilane moieties [129–131].
- Spray coating of SWNTs onto the surface [44, 145].
- Combination of SWNTs with a second conducting agent [126, 132].

The first two approaches provided volume conductivity of the polymer. Addition of 0.5 wt.% of SWNTs allowed to achieve a desired bulk conductivity in both approaches. However, addition of SWNTs deteriorated the optical transparency and increased solar absorptivity of the dielectric material. The third approach provided just surface conductivity to the dielectric material and it did not affect the optical properties of the polymer. The fourth approach suggested to generate bulk electric conductivity with SWNTs combined with second conducting agent (e.g., inorganic salts $CuSO_4$, $CuCl_2$) in order to preserve thermal and optical properties of the polymer. However, optical micrographs and electrical resistivity tests showed that addition of inorganic salts increased agglomeration of SWNTs without any increase in surface resistivity and improvement of thermal and optical properties [132]. Watson et al. [146] reported the test results of transparent films prepared with three different polyimide (LaRC™ CP2, TOR-NC and PPO-6FDA) containing SWNTs. It is shown that the optical properties were improved for SWNTs spray-coated films as compared to the films containing SWNTs throughout the volume. Also, SWNTs coated films demonstrated higher flexibility and strength.

4.5 Impact of Space Radiation on Ionic Materials

Previous studies on deep dielectric charging mostly have been focused on spacecraft components such as wire insulations and printed circuit boards. During the last two decades, *ionic materials (IM)* are of particular interest to the space industry due to their extremely low vapor pressure. The traditional lubricating materials used in space (mineral oils, polyol ester, PFPE, Pennzane, etc.) have limited lifetimes in vacuum due to the catalytic degradation on metal surfaces, high vaporization at high temperatures, dewetting, and other disadvantages. The lubricants for space applications must have vacuum stability (i.e., low vapor pressure), high viscosity index (wide liquid range), low creep tendency, good elastohydrodynamic and boundary lubrication properties, radiation, atomic oxygen resistance, and sometimes optical or infrared transparency. This distinguishing factor makes these lubricants functional in ultra-high vacuum environments, where conventional lubricants would quickly vaporize leaving critical mechanical components without lubrication. The properties such as base oil and additive volatility, creep tendency, surface tension, viscosity, chemical composition, weight loss, density, vapor pressure, etc. are important in order to predict a resistance of the ILs to space radiation. Various studies have been conducted to determine the tribological parameters of Ionic fluids [55].

Unfortunately, properties such as heat flow, heat capacity, thermogravimetric weight loss, and non-linearity in the rheological behavior of the lubricants are not studied well for newly developed systems. These properties are crucial to analyzing thermodynamic and energy dissipative aspects of the lubrication process.

4.6 Nonlinear Phenomena in Space Charge

The phenomena related to the space charge are quite complex. According to Hofmann and Franchetti [51], from a nonlinear dynamics point of view, the main reason of this complexity is the self-consistency problem. Conventionally, a linear perturbation in space charge phenomena is considered as a linear combination of the Eigen wave modes, which can be obtained from the linear wave dispersion relations of the space plasma [83]. It is assumed that there is no interaction between these Eigen wave modes, and the coefficients of the linear combinations are time independent. Therefore, using these linear Eigen wave modes characteristics, the linear evolution of a linear perturbation can be predicted. However, this assumption cannot be applied to nonlinear waves because the coefficients of the linear combinations vary with time. Hence, one cannot predict the nonlinear waves based on the characteristics of the linear Eigen modes. There are several methods to study the nonlinear waves' phenomena in the space charges, such as quasi-linear approximation technique, pseudo potential method, numerical simulation techniques, nonlinear jump conditions concept, probability approach, etc.

Quasi-linear approximation can be used to study nonlinear waves of small amplitudes. It is based on Taylor expansion of nonlinear equations. Pseudo potential method is used to solve analytically one-dimensional and steady-state nonlinear wave problems using the conservation of energy flux equation. Numerical simulation methods are considered as a powerful instrument to predict and describe nonlinear waves' behavior. In nonlinear jump conditions concept, nonlinear jump conditions of shock waves, instabilities, and discontinuities in collisionless dissipation process can be derived using the fundamental equations, such as conservation of mass flux equation, momentum flux equation, energy flux equation, and Maxwell's equations. Probability (or statistic) approach can describe nonlinear phenomena using the theories such as chaos, fractal, and turbulence.

There are several models to describe three-dimensional space charges. For example, Vlasov-Maxwell plasma model, two-fluid Maxwell model, one-fluid MHD, quasi-MHD plasma model, etc. In these models, the gravitational force is neglected as it is much smaller than the Lorentz force.

The linear theory of space-charge oscillations in an electron beam passing through a region with trapped electrons was considered by Koshunov [60] in the geometrical-optics approximation. It was shown that the presence of trapped particles leads to instability in the beam. With increasing oscillation amplitude, the interaction of the beam with the trapped electrons becomes basically nonlinear and cannot be described analytically. Ivanov et al. [53] used a partial numerical simulation (Runge-Kutta) technique to solve the nonlinear problem of the interaction between the space-charge waves of an intense electron beam and a relatively small number of electrons trapped in a potential well. The initial distribution of the trapped electrons was homogeneous. It was shown that if the well is not very deep, a monochromatic space charge wave is unstable with respect to excitation of satellites with shorter wavelength. However, when the potential well depth is increased the dominant role in the space-charge wave packet was assumed by the wave with the maximal increment. According to the authors, the trapped electrons draw from the electric field of the excited wave an additional energy. At a small potential-well depth, the energy acquired by the trapped electrons from the field is much greater than their initial energy. The behavior of the trapped electrons is determined in this case mostly by their interaction with the field of the excited wave.

Moehl and H. Schoenauer [95] proposed a theory to solve the dispersion relation, including the influence of external and space-charge induced nonlinearities. The preliminary showed that the combined effect of external and space-charge nonlinearities in the two transverse planes can considerably enhance Landau damping in low-energy machines (up to 10 GeV). It is also shown that the sign of the octupole current is important especially for a flat beam and should be chosen in such a way that tune increases with amplitude.

Using Vlasov-Poisson equations Hofmann [49] derived coherent frequencies and stability characteristics of anisotropic beams with different constants and emittances. The obtained dispersion equations of multipole oscillations with different polar symmetries have been solved numerically. The simulation results showed that for large energy anisotropy, some of the eigenmodes become unstable in the

space charge governed regimes. The results have been applied to high-current linear accelerators and it is found that non-equipartitioned beams may exist in large regions of space parameter under stable conditions.

Franchetti et al. [34] presented a micro map technique to include space charge defocusing forces for a coasting beam. The authors developed a macro-particle model was developed to calculate the space charge forces under the assumption of elliptic beam symmetry. The obtained results prove a coherent effect of space charge which improves the exchange with respect the single particle exchange in the absence of space charge.

Blaskiewicz [13] predicted the effects of the direct space charge force on transverse stability using a momentum and energy conservation model to describe this force. The author solves one-dimensional equations of motion which include nonlinear interparticle forces. To obtain the dispersion relation the author used a first-order perturbation theory on the Vlasov equation. It is shown that the nonlinearity in the space charge force can modify dispersion diagrams without Landau damping (exponential decrease of longitudinal space charge waves in plasma or a similar environment as a function of time). Nonlinear elements also reduce stability for a reasonable machine impedance and cause the betatron (cyclic particle accelerator) tune to increase in amplitude.

Hofmann and Boine-Frankenheim [50] compared Kapchinsky-Vladimirsky analytical theory [56, 57] with their 2D simulations. This study was focused on an extended region around the case of symmetric focusing constants. It is shown that a solution of the linearized Vlasov equation leads to dispersion relations for eigenmodes with space charge potentials. The authors also discuss the applicability of the linearized Kapchinsky-Vladimirsky theory.

Aslaninejad and Hofmann [7] studied the response of coasting beams in the presence of space charge to linear coupling as well as gradient errors, based on linearizing the self-consistent Vlasov-Poisson equation. The coherent shift of the resonance conditions as well as the response on lattice errors in smooth approximation has been determined. A characteristic feature is cancellation between the external forces due to the lattice errors and the beam induced space-charge forces, if the resonance condition is satisfied for the (space-charge shifted) single-particle tunes. For the linear coupling mode and small tune split, we also identify a parameter region, where spontaneous instability exists in addition to the resonance.

Franchetti et al. [33] reported a pronounced collective nonlinear behavior in a beam with linear coupling and space charge described by the complete second order momentum equations. It is shown that the space charge has the effect so the resonance condition becomes a dynamically varying quantity depending on the instantaneous emittances and saturation effects caused by space charge. In a linearized Vlasov theory, collective nonlinearity through a space charge is absent and resonance effects lead to infinite response.

Montague [96] for the first time showed that the space-charge potential could excite the fourth-order coupling resonance, now known as the Montague resonance. This mechanism leads to the oscillations between the horizontal and vertical planes for the single-particle motion, resulting in an apparent increase in emittance in

the plane of smaller emittance. This phenomenon is very fast as growths in few turns for a synchrotron at the space-charge limit. The later it was revealed that the Montague mechanism is in fact a combination of instability and resonance: it develops as a pure instability for a Kapchinsky – Vladimirsky (KV) distribution [57], which has no initial driving nonlinearity. The initial growth for the KV case is found to be exponential, starting from noise level. Metral et al. [93] conducted experimental studies in the CERN (Organisation européenne pour la recherche nucléaire) proton synchrotron (PS) to analyze the following space charge effects: (i) transverse emittance blow-up due to space-charge induced crossing of the integer or half-integer stop-band, (ii) space-charge and octupole-driven resonance trapping, and (iii) intensity-dependent emittance transfer between the two transverse planes. The experimental results have been compared to the numerical predictions (3D PIC code IMPACT). Asvesta et al. [8] presented a new optics for the CERN PS-Booster (PSB). They conducted emittance measurements crossing the coupling resonance for the nominal and the new optics. Also the fourth-order coupling space charge-driven resonances have been studied analytically. Experimental studies confirmed that the proposed new optics with unequal integer tunes is much less affected by coupling, as the Montague resonance is not excited and the linear coupling can be compensated to a large extend. Nevertheless, if the beam interacts with the resonance for an extended period of time, some emittance exchange is observed but much less pronounced compared to the existing optics.

Hofmann and Franchetti [51] discussed the effect of space charge on nonlinear dynamics issues of high-intensity beams. In this study, the octupole effect (driving term for the resonance) was considered with and without space charge. It is shown that space charge in combination with nonlinear dynamics can create other phenomena, such as nonlinear detuning, coherent shift of resonance conditions, saturation effects, etc.

4.7 Conclusions

The review of the research studies demonstrates that the phenomena introduced by space charge are relatively complex, and its connection with nonlinear dynamics may lead to additional phenomena. The main reason of complexity is the issue of self-consistency of the space charge electric field. Analytical solutions of the problem are practically impossible. It is understandable that the computer simulations based on the analytical guidelines for the two- and three-dimensional cases and more experimental works will be required to fully understand the space charge phenomena.

Acknowledgments This research was supported in part by the Air Force Research Laboratory (*Spacecraft Charging and Instrument Calibration Laboratory*), Space Vehicles Directorate, through the Air Force Office of Scientific Research Summer Faculty Fellowship Program®, Contract Numbers FA8750-15-3-6003 and FA9550-15-0001.

References

1. Aberkromby, K. J., Hamada, K., Guyote, M., Okada, J. and Barker, E. Remote and Ground Truth Spectral Measdurements Comparisons of FORMOSAT III. Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, USA, 2007.
2. Agostinelli, S., Allison, J., Amako, K. et al. 2003. GEANT 4 – A Simulation Toolkit. Nucl. Instrum. Meth. Phys. Rev., A 506, pp. 250–303.
3. Allahdadi, F. A., Bakhtiyarov, S. I., Wyss, G. D., Polansky, G. F., Sholtis, J. A. and Botts, C. D. 2013,"Nuclear-Powered Payload Safety". In F. A. Allahdadi (Ed.), Safety Design for Space Operations (pp. 255–370). Butterworth-Heinemann: Elsevier Ltd.
4. A.A. Aly, E.-S.B. Zeidan, A.A. Alshennawy, A.A. El-Masry and W.A. Wasel: *World J. Nano Sci. and Eng.*, 2012, vol. 2, pp. 32–39.
5. Arnaout, M., Paulmier, T., Dirassen, B. and Payan, D. Study of Radiation Induced Conductivity and Photoconduction Phenomenon for Materials Used in Space Environment. Journal of Electrostatics, 2016, Vol. 84, pp. 48–53.
6. Ash, B. J., Siegel, R. W. and Schadler, L. S. "Glass Transition Temperature Behavior of Alumina/PMMA Nanocomposites", J. Polymer Sci. B, Vol. 42, pp. 4371–4383, 2004.
7. M. Aslaninejad and I. Hofmann. Effect of Space Charge on Linear Coupling and Gradient Errors in High-Intensity Rings. Physical Review Special Topics – Accelerators and Beams, Vol. 6, 2003, 124202, pp. 1–11.
8. F. Asvesta, F. Antoniou, H. Bartosik, G. P. Di Giovanni, Y. Papaphilippou. Coupling and Space Charge Studies at the CERN PSB. 110th International Particle Accelerator Conference. Journal of Physics: Conference Series, Vol. 1350, 2019, pp. 1–6.
9. Bakhtiyarov, S. I. and Ferguson, D. C., 2020, "Modeling of Dielectric Behavior of Polymer Materials Embedded with Metal-Doped Nanoparticles that Contain Interphasial Space Charges", i-manager's Journal on Future Engineering and Technology, Vol. 15, No. 2, pp. 1–7 doi: https://doi.org/10.26634/jfet.15.2.16615.
10. Balmain, K., "Metal Based Polymers for Reduced Discharge Occurrence", in Proc. 6th Spacecraft Charging Technology Conf., Hanscom AFB, MA, Nov. 1998.
11. Bao, G. 2012, The ASTROD I charging simulation and disturbances due to solar energetic particles at 0.5AU, Acta Astronautica, Volume 77, Pages 29–33.
12. Bedard, M. D. Using a Physics-Based Reflection Model to Study the Reddening Effect Observed in Spectrometric Measurements of Artificial Space Objects. Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, USA, 2007.
13. M. Blaskiewicz. Transverse stability with nonlinear space charge. Physical Review Special Topics: Accelerators and Beams, 2001, Vol. 4, 044202, pp. 1–9.
14. Carruth, M. R. et al. 2001. ISS and Space Environment Interactions without Operating Plasma Contactor. AIAA Paper 2001–0401.
15. W.X. Chen, F. Li, G. Han, J.B. Xia, L.Y. Wang, J.P. Tu and Z.D. Xu, *Tribology Letters*, 2003, vol. 15(3), pp. 275–78.
16. Chugg, A. and Randall, G. 1999, Modeling the charge and dose deposition at electronic devices within spacecraft. Radiation Physics and Chemistry, Vol. 56, pp. 259–266.
17. Cho, M., Saito, K. and Hamanaga T. 2012. Data analysis of the polar plasma environment for spacecraft charging analysis, Acta Astronautica, Volume 81, Issue 1, 160–173.
18. Clark, F., Insulating Materials for Design and Engineering Practice. New York: Wiley, 1962.
19. Cohen, H.A. and Lai, S. T., "Discharging the P78-2 Satellite Using Ions and Electrons", in Proc. AIAA 20th Aerospace Science Meeting, Orlando, FL, 1982, Paper # AAA-82-0266.
20. Cooke, D. L. and Geiss, M. 2002. Introducing the Passive Anode Surface Emission Cathode. Proc. AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, 7–10 July 2002, Indianapolis, Indiana, USA, AIAA Paper 2002–4049.
21. Craven, P. D., Wright, K. H., Jr., Minow, J. I., Coffey, V. N., Schneider, T. A., Vaughn, J. A., Ferguson, D. C. and Parker, L. N. 2009. Survey of International Space Station Charging Events. 47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition, 5–8 January 2009, Orlando, Florida, USA, AIAA 2009–119.

22. Czepiela, S. A.; McManus, H. L.; Hastings, D. Charging of Composites in the Space Environment. Journal of Spacecraft and Rockets. 2000, Vol. 37, No. 5, pp. 556–560.
23. Davis, V. A., Mandell, M. J., Cooke, D. C., Wheelock, A., Mateo Velez A., Roussel, J. C., Payan, J. F., Cho, D. and Koga, K. 2013. Comparison of low earth orbit wake current collection simulations using Nascap-2k, SPIS, and MUSCAT computer codes. Trans. Plasma Sci., 41(12), 3303–3309.
24. Dennison, J. R., Brunson, J., Swaminathan, P., Green, N. W. and Frederickson, A. R., 2006, "Methods for High Resistivity Measurements Related to Spacecraft-Charging", IEE Transactions on Plasma Science. Vol. 34(5), pp. 2191-2203.
25. Engelhart, D. P., Plis, E. A., Ferguson, D., Cooper, R. and Hoffmann, R. C. Optical and Chemical Characterization of Polyimide in a GEO-Like Environment. Technical Paper. Advanced Maui Optical and Space Surveillance Technologies Conference. Maui, USA, 2016.
26. Engelhart, D. P., Plis, E. A., Humagain, S., Greenbaum, S., Ferguson, D., Cooper, R. et al., 2017. Chemical and Electrical Dynamics of Polyimide Film Damaged by Electron Radiation. IEEE Transactions on Plasma Science. Vol. 45(9), pp. 2573–2577.
27. Engelhart, D. P., Plis, E. A., Ferguson, D., Johnson, W.R., Cooper, R. and Hoffmann, R. C. 2018, Space Plasma Interactions with Spacecraft Materials, Chapter 12, In: Plasma Science and Technology – Basic Fundamentals and Modern Application. IntechOpen, pp. 225–245.
28. Ferguson, D. C. Alternatives to the ISS Plasma Contacting Units. Proc. 40th Aerospace Sciences Meeting & Exhibit, 14–17 January 2002, Reno, Nevada, USA, AIAA Paper 2002-0934.
29. Ferguson, D. C. and Gardner, B. 2002. Modeling International Space Station (ISS) Floating Potentials. AIAA Paper 2002-0933.
30. Ferguson, D. C., Craven, P. D., Minow, J. I. and Wright, K. H., Jr. A Theory of Rapid Charging Events on the International Space Station. 2009. AIAA Paper 2009-3523.
31. Ferguson, D. C. and Wimberly, S. C. 2013. The best GEO daytime spacecraft charging index. In: 51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, 07-10 January 2013. Grapevine (Dallas/Ft. Worth Region), Texas AIAA 2013-0810.pp. 11–17.
32. Ferguson, D.C., Hilmer, R.V., and Davis, V.A. Best Geosynchronous Earth Orbit Daytime Spacecraft Charging Index, Journal of Spacecraft and Rockets, Vol. 52, No. 2, March–April 2015.
33. G. Franchetti, I. Hofmann and M. Aslaninejad. Collective Emittance Exchange with Linear Space Charge Forces and Linear Coupling. Physical Review Letters, Vol. 94, 2005, 194801, pp. 1–4.
34. G. Franchetti, I. Hofmann and G. Turchetti. Micromap Approach to Space Charge in a Synchrotron. AIP Proceedings of the Workshop on Space Charge Physics in High Intensity Hadron Rings, Shelter Island, eds. A.U. Luccio and W. T. Weng, New York, Vol. 448, No.1, 1998, pp. 233–244.
35. Frederickson, A. R. "Radiation-Induced Electrical Current and Voltage in Dielectric Structures," AFCRL-TR-74-0582 (1974).
36. Frederickson, A. R. "Charge deposition, photoconduction, and replacement current in irradiated multilayer structures," IEEE Trans. Nuc. Sci., Vol. 22, Dec. 1975.
37. Frederickson, A. R., Cotts, D. B., Wall, J. A. and Bouquet, F. L. "Spacecraft Dielectric Material Properties and Spacecraft Charging". Washington, DC, AIAA, 1986.
38. Fujita, F., Ruike, M. and Baba, M., "Treeing breakdown voltage and TSC of alumina filled epoxy resin", IEEE Intern. Sympos. Electr. Insul., San Francisco, Vol. 2, pp. 738–741, 1996.
39. Fung, S. F., 1996, "Recent Developments in the NASA Trapped Radiation Models", in Radiation Belts: Modelsand Standards, ed. By J. F. Lemaire, D. Heynderickx, D. N. Baker. Geopys. Monogr. Ser., Vol. 97 (AGU, Washington), pp. 79–91.
40. H. Gandotra, S. Mahajan, S. Jandival and S. Gupta: *Int. J. Sci. Tech. Adv.,* 2018, vol. 4(1), pp. 147–150.
41. Geis, M. W., Efremov, N. N., Krohn, K. E., Twichell, J. C., Lyszczarz, T. M., Kalish, R., Greer, J. A. and Tabat, M. D. Theory and Experimental Results of a New Diamond Surface-Emission Cathode. The Lincoln Lab Journal, 1997, 10(1), 1–18.

42. Ghosh, K., Guo, J. and Lopez-Pamies, O. "Homogenization of Tim-Dependent Dielectric Composites Containing Space Charges, with Applications to Polymer Nanoparticulate Composites", International Journal of Non-Linera Mechanics, 2019.
43. Ginet, G. P., O'Brien, T. P., Huston, S. L., Johnston, W. R., Guild, T. B., Friedel, R., Lindstrom, C. D., Roth, C. J., Whelan, P., Quinn, R. A., Madden, D., Morley, S. and Su, Y.-J., 2013, "AE9, AP9 and SPM: New Models for Specifying the Trapped Energetic Particle and SpacePlasma Environment", Space Sci. Rev., Vol. 179, pp. 579–615.
44. Glatkowski, P., 2003. Society for Advancement of Materials and Process Engineering Proceedings, Vol. 48, p. 2146.
45. Grard, R. J. L., "Spacecraft Potential Control and Plasma Diagnostic Using Electron Field Emission Probes", *Space Sci. Instrum*., 1975, Vol. 1, pp. 363–376.
46. Halbleib, J. A., Kensek, R. P., Mehlhorn, T. A., Valdez, G. D., Sdeltzer, S. M. and Berger, M. J. 1992, ITS 3.0: the integrated TIGER series of coupled electron-photon Monte-Carlo transport codes. SAND91-1634, Sandia Labs, USA.
47. Hanna, R., Paulmier, R. T., Molinie, P., Belhaj, M., Dirassen, B., Payan, D. and Balcon, N. Radiation Induced Conductivity in Space Dielectric Materials. J. Appl. Phys., 2014, Vol. 115, 033713.
48. Hastings, D. and Garrett, H. B., Spacecraft-Environment Interactions, Cambridge University Press, Cambridge, England, UK, 1997.
49. I. Hofmann. Stability of Anisotropic Beams with Space Charge. Physical Review E, Vol. 57, No. 4, 1998, pp. 4713–4724.
50. I. Hofmann and O. Boine-Frankenheim. Resonant Emittance Transfer Driven by Space Charge. Physical Review Letters, Vol. 87, No. 3, 2001, 034802, pp. 1–4.
51. I. Hofmann and G. Franchetti. Nonlinear phenomena in space–charge dominated beams. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 561, Issue 2, 1 June 2006, Pages 137–142.
52. Huang, C., Zhang, Q. M., Yu Li, J., Rabeony, M. 2005. Colossal dielectric and electromechanical responses in self-assembled polymeric nanocomposites. Applied Physics Letters, Vol. 87, 182901.
53. A. A. Ivanov, S. M. Korshunov, N. S. Beryuleva and S. P. Tarasov. Nonlinear interaction of space-charge waves with trapped electrons. Soviet Physics: Journal of Experimental and Theoretical Physics, Vol. 38, No. 5, 1974, pp. 927–930.
54. Iwata, M., Sumida, T., Igawa, H., Fujiwara, Y., Okumura, T., Khan, M. A. R., Toyoda, K., Cho, M., Hatta, S., Sato, T. and Fujita, T. Development of Electron-Emitting Film for Surface Charging Mitigation: Observation, Endurance and Simulations. Proc. of 47th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. 5–8 January, 2009, Orlando, FL, USA, AIAA Paper 2009-560.
55. Jones, W. R.; Jansen, M. *J. Lubrication for Space Applications*, NASA/CR-2005-213424, pp. 1–38.
56. I. Kapchinsky, Theory of Resonance Linear Accelerators, New York, Harwood Academic Publishers, 1985.
57. I. M. Kapchinsky and V. V. Vladimirsky, Proceedings of the International Conference on High Energy Accelerators, CERN, Geneva, 1959, p. 274.
58. Khalil, M. S., Henk, P. O. and Henriksen, M., "The influence of titanium dioxide additive on the short-term DC breakdown strength of polyethylene", IEEE Intern. Sympos. Electr. Insul., Montreal, Canada, pp. 268–271, 1990.
59. V. B. King: *Nanotechnology: Research Advance*, Nova Publishers, Inc., New York, NY, 2007.
60. S. M. Korshunov, Soviet Physics: Journal of Experimental and Theoretical Physics, Vol. 35, 1972, p. 917.
61. B. A. Krick, J.J. Ewin and E.J. McCumiskey: *Tribology Trans*., 2014, vol. 57, pp. 1058–65
62. E. Kutelia, L. Rukhadze, N. Maisuradze, B. Eristavi, S. Bakhtiyarov: *Georgian Eng. News*, 2010, vol. 79(1), pp. 68–73.
63. E. Kutelia, L. Rukhadze, N. Jalabadze, T. Dzigrashvili, O. Tsurtsumia and D. Gventsadze: Nucleation and Growth of Carbon Nanoforms on the Surface of Metallic Plate-Substrates and the *Adv. Mater. Lett.*, 2018a, vol. 9(12), pp. 867–71.

64. E. R. Kutelia, G. Darsavelidze, T. Dzigrashvili, T. Kukava, D. Gventsadze, L. Rukhadze, L. Nadaraia, L. Gventsadze, O. O. Tsurtsumia, N. Jalabadze, I. Kurashvili, S. I. Bakhtiyarov, 2019, "Effect of Fe Cluster-Doped CNT Filler on Internal Friction and Shear Modulus of PTFE", Metallurgical and Materials Transactions A (in press).
65. E. R. Kutelia, G. Darsavelidze, T. Dzigrashvili, D. Gventsadze, O. Tsurtsumia, L. Gventsadze, T. Kukava, L. Rukhadze, L. Nadaraia, I. Kurashvili and S. Bakhtiyarov, 2020, "The Inelastic/Elastic and Tribological Properties of PTFE-Based Nanocomposites Filled with Co Cluster-Doped CNTs", Bulletin of the Georgian National Academy of Sciences: Physical Chemistry, Vol. 14, No. 1, pp. 52-58.
66. E. Kutelia, D. Gventsadze, O. Tsurtsumia, L. Rukhadze, N. Jalabadze, T. Kukava and T. Dzigrashvili: *Adv. Mater. Lett.*, 2018b, vol. 9(5), pp. 320-25.
67. Laghari, J. R. and Hammoud, A. N., 1990, "A Brief Survey of Radiation Effects on Polymer Dielectrics", IEEE Transactions on Nuclear Science, Vol. 37(2), pp. 1076–1083.
68. Lai, S. T., An Overview of Electron and Ion Beam Effects in Charging and Discharging of Spacecraft. IEEE Transactions on Nuclear Sciences, 1989, 36(6), 2027–2032.
69. Lai, S. T., 1991a. "Theory and Observation of Triple-Root Jump in Spacecraft Charging", J.Geophys. Res., Vol. 96, No. A11, pp. 19269–19282.
70. Lai, S. T., 1991b. "Spacecraft Charging Thresholds in Single and Double Maxwellian Space Environments", IEEE Trans. Nucl. Sci., Vol. 19, pp. 1629–1634.
71. Lai, S. T. and Murad, E., 1995. "Spacecraft Discharging Using Vapor of Polar Molecules", in Proc. AIAA Plasmadynamics Lasers Conf., Aug. AIAA-95-1941, p. 11.
72. Lai, S. T. and Wang, J., "Spacecraft Interactions with Beam Emissions", in Proc. IEEE Int. Conf. Plasma Science, San Diego, CA, May 1997.
73. Lai, S. T. and Murad, E., "Mitigation of Spacecraft Charging by Means of Ionized Water", US Patent No. 6463672 B1, Oct. 15, 2002a.
74. Lai, S. T. and Murad, E., "Mitigation of Spacecraft Charging by Means of Polar Molecules", US Patent No. 6500275 B1, Dec. 31, 2002b.
75. Lai, S. T., A Critical Overview on Spacecraft Charging Mitigation Methods. *IEEE Transactions on Plasma Sciences*, 2003, 31(6), 1118–1124.
76. Lai, S. T., Overview of Surface and Deep Dielectric Charging on Spacecraft. Spacecraft Charging. Ed. Lai, S. T., Progress in Astronautics and Aeronautics, Vol. 237, AIAA Publications, 2011, 1–17.
77. Lai, S. T., Cahoy, K., Lohmeyer, W., Carlton, A., Aniceto, R., and Minow, J. 2018, Chapter 16: Deep Dielectric Charging and Spacecraft Anomalies. Extreme Events in Geospace, 419–432.
78. Latham, R. V., "The origin of prebreakdown electron emission from vacuum-insulated high voltage electrodes", 1982. Vacuum, Vol. 32(3), pp. 137–140.
79. J. Li and I.Q. Zhang: *J. Mech. Sci.*, 2009, vol. 223, pp. 2501–07.
80. Liu, L., Dong, Y., Bao, G., Ni, W.-T. and Shaul, D. N. A. 2010, Simulation of ASTROD I test mass charging due to solar energetic particles and interplanetary electrons. Advances in Space Research, Volume 45, Issue 1, Pages 200–207.
81. Long, H.; Wei, Y.; Yu, T. J.; Wang, Z.; Jia C. Y.; Yang, Z. J.; Zhang, G. Y.; Fan, S. S. Modulating Lateral Strain in GaN-Based Epitaxial Layers by Patterning Sapphire Substrates with Aligned Carbon Nanotube Films. Nano Res. 2012, 5, 646–653.
82. Lopez-Pamies, O., Goudarzi, T., Meddeb, A. B. and Ounaies, Z., "Extreme enhancement and reduction of the dielectric response of polymer nanoparticulate composites via interphasial charges", Applied Physics Letters, Vol. 104, 242904, 2014.
83. L.-H. Lyu, Nonlinear Space Plasma Physics. Chapter 1: Introduction to Nonlinear Space Plasma Physics, 2005.
84. Ma, D., Siegel, R.W., Hong, J.-I., Schadler L. S., Martensson, E. and Önneby, C., "Influence of nanoparticle surfaces on the electrical breakdown strength of nanoparticle-filled low-density polyethylene", J. Mater. Res., Vol. 19, No. 3, pp. 857–872, 2003.
85. Ma, D., Akpalu, Y. A., Li, Y., Siegel, R.W. and Schadler, L.S. "Effect of titania nanoparticles on the morphology of low density polyethylene", J. Polymer Sci. Part B, Polymer Phys., Vol. 43, pp. 463–533, 2005.

86. Manners, I., "Ring-Opening Polymerization of Metallocenophanes: A New Route to Transition Metal-Based Polymers", Adv. Organometal. Chem., 1995, Vol. 37, pp. 131–168.
87. Manners, I., "Ring-Opening Polymerization of Silaferrocenophanes with the Channels of Mesoporous Silica: Poly(ferrocenylsilane)-mcm-41 Presursors to Magnetic Iron Manostructures", Adv. Mater., 1998, Vol. 10, pp. 144–149.
88. Marletta, G; Iacona, F. Heat-induced versus particle-beam-induced chemistry in polyimide. *Nucl. Instrum. Methods Phys. Res. B* 1993, 80, 1045–1049.
89. Marletta, G; Iacona, F. Chemical and Physical Property Modifications Induced by Ion Irradiation in Polymers. In *Materials and Processes for Surface and Interface Engineering. NATO ASI Series (Series E: Applied Sciences);* Pauleau, Y., Ed.; Springer: Dordrecht, The Netherlands, 1995; Volume 290, pp. 9–16.
90. McKenna-Lawlor, S. 2014. An overview of the energetic particle radiation hazard en route to and at Mars. International Academics of Astronautics Publication.
91. McKenna-Lawlor, S., Bhardwaj, A., Ferrari, F., Kuznetsov, N., Lal, A. K., Li, Y., Nagamatsu, A., Nymmik, R., Panasyuk, M., Petrov, V., Peitz, G., Pinsky, L., Shukor, M., Singhvi, A. K., Straube, U., Tomi, L. and Lawrence, T. 2015. Recommendations to mitigate against human health risks incurred due to energetic particle irradiation beyond low earth orbit/BLEO, Acta Astronautica, Volume 109, 2015, Pages 182–193.
92. McKnight, D. 2019. Examination of spacecraft anomalies provides insight into complex space environment, Acta Astronautica, Volume 158, 172–177.
93. E. Metral, M. Giovannozzi, M. Martini, R. Steerenberg, G. Franchetti, I. Hofmann, J. Qiang and R.D. Ryne, Space-Charge Experiments at the CERN Proton Synchrotron. 33rd ICFA Advanced Beam Dynamics Workshop on High Intensity and High Brightness Hadron Beams, Benshe, eds. I. Hofmann, J.-M. Lagniel and R. W. Hasse, Vol. 773, No. 1, 2005, pp. 122–126.
94. Milton, G. W., Appl. Phys. A, 1981, Vol. 26, pp. 125–130.
95. D. Moehl and H. Schoenauer, Landau Damping by Non-Linear Space-Charge Forces and Octupoles. In: Proceedings of the IXth International Conference on High Energy Accelerators, Stanford, 1974, pp. 380–384.
96. B. W. Montague, "Fourth-Order Coupling Resonance Excited by Space Charge Forces in a Synchrotron", CERN, Geneva, Switzerland, Rep. CERN-68-38, 1968.
97. Montanari, G. C., Fabiani, D., Palmieri, F., Kaempfer, D., Thomann, R. and Mulhaupt, R., "Modification of Electrical Properties and Performance of EVA and PP Insulation through Nanosructure by Organophilic Silicates", IEEE Trans. Dielectr. Electr. Insul., Vol. 11, pp. 754–762, 2004.
98. D. F. Moore, Principles and Applications of Tribonics, Pergamon Press, New York, NY, 1975, 488 p.
99. Morton, T. and Minow, J. 2002. Floating Potential Probe Langmuir Probe Data Reduction Results. AIAA Paper 2002-0936.
100. Mullen, E. G., Frederickson, A. R., Murphy, G. P., Ray, K. P., Holeman, E. G., Delorey, D. E., Robson, R. and Farar, M. An Automatic Charge Control System at Geoynchronous Altitude: Flight Results for Spacecraft Design Consideration. IEEE Transactions on Nuclear Science, 1997, 44(6), 2188–2194.
101. Mundari, N. D. A., Srivastava, A. K., Toyoda, K. and Cho, M. 2014. Influence of atomic oxygen exposure on surface resistivity of silicon doped polyimide affecting spacecraft charging. Vacuum, Volume 105, 11–16. Simultaneous investigation of magnetospheric plasma and spacecraft charging. Advances in Space Research, 42, 1307–1312.
102. Murray-Krezan, J., Inbody, W. C., Dao, P. D., Dentamaro, A., Fulcoly, D., and Gregory, S. A., "Algorithms for Automated Characterization of Three-Axis Stabilized GEOS using Non-Resolved Optical Observations," Advanced Maui Optical and Space Technologies Conference Technical Papers (2013).
103. NASA-HDBK-4006, Low Earth Orbit Spacecraft Charging Design Handbook (2007).
104. Nelson, J. K., Fothergill, J. C., Dissado, L. A. and Peasgood, W. "Towards an understanding of nanometric dielectrics", IEEE Conf. Electr. Insul. Dielectr. Phenomena, Mexico, pp. 295–298, 2002.

105. Nelson, J. K. and Fothergill, J. C. 2004. Internal charge behavior of nanocomposites. Nanotechnology, Vol. 15, pp. 586–595.
106. Ni, W.-T., Bao, Y., Dittus, H., et al. 2006, Acta Astronaut., Vol.59, p. 598.
107. NIST (2019). http://physics.nist.gov/.
108. Novikov, L. S., Mileev, V. N., Krupnikov, K. K., Makletsov, A. A., Marjin, B. V., Rjazantseva, M. O., Sinolits, V. V. and Vlasova, N. A. 2008.
109. Novikov, L. S., Makletsov, A. A. and Sinolits, V. V. 2016. Comparison of Coulomb-2, NASCAP-2K, MUSCAT and SPIS codes for geosynchronous spacecraft charging, Advances in Space Research, Volume 57, Issue 2, 671–680.
110. Paillous, A. Radiation Damage to Surface and Structure Materials. In: The Behavior of Systems in the Space Environment, Springer: Dordrecht, the Netherlands, 1993, pp. 383–405.
111. Park, C., Ounaies, Z., Watson, K. A., Crooks, R. E., Smith, Jr. J. G., Lowther, S. E., Cennell, J. W., Siochi, E. J., Harrison, J. S. and St. Clair, T. L. 2002, Chem. Phys. Lett., Vol. 364, p. 303.
112. Park, S. H.; Jin, S. H.; Jun, G. H.; Jeon, S.; Hong, S. H. Enhanced Electrical Properties in Carbon Nanotube/Poly (3-hexylthiophene) Nanocomposites Formed through Non-Covalent Functionalization. Nano Res. 2011, 4, 1129–1135.
113. Parks, D. E. and Katz, I. J. "Theory of Plasma Contactors for Electrodynamic Tethered Satellite Systems," 1987. Journal Spacecraft Rockets, Vol. 24, pp. 245–249.
114. Paulmier, T., Dirassen, B. and Rey, R. 2018, Electrostatic behavior of space used materials in regard of internal charging met on spacecrafts, Journal of Electrostatics, Volume 92, 66–74.
115. Pelton, J. N., Allahdadi, F. A., Bakhtiyarov, S. I. and Green, J. L. (2015), "Handbook of Cosmic Hazards and Planetary Defense", Springer, http://refworks.springer.com/cosmichazards
116. Plis, E.A., Engelhart, D.P., Likar, J., Hoffmann, R.C., Cooper, R., and Ferguson, D. Electrical behavior of carbon-loaded Kapton for spacecraft applications, Journal of Spacecraft and Rockets, 2017, Nov. 28, pp. 1–2.
117. Plis, E. A., Engelhart, D. P., Cooper, R., Johnston, W. R., Ferguson, D., and Hoffman, R. Review of Radiation-Induced Effects in Polyimide. *J. Appl. Sci.* 2019, 9, pp. 1–24.
118. Purvis, C. K., Garrett, H. B., Whittlesey, A. C. and Stevens, N. J., "Design Guidelines for Assessing and Controlling Spacecraft Charging Effects", 1984, NASA, NASA TM 2361.
119. Qu, C.; Hu, J.; Liu, X; Li, Z.; Ding, Y. Morphology and Mechanical Properties of Polyimide Films: The Effects of UV Irradiation on Microscale Surface. *Materials*. 2017, 10, 1329 doi:https://doi.org/10.3390/ma10111329.
120. Reinisch, R. F., Gloria, H. R. and Androes, G. M., Photoelimination Reactions of Macromolecules, Photochemistry of Macromolecules, Plenum Press, New York, 1970, pp. 185–217.
121. Richardson, O. W., "On the Negative Radiation from Hot Platinum", Proc. Camb. Phil. Soc., Vol. 11, pp. 286-295, 1901.
122. Roy, M., Nelson, J. K., MacCrone, R. K., Schadler, L. S., Reed, C. W., Keefe, R. and Zenger, W., "Polymer Nanocomposite Dielectrics – The Role of the Interface", IEEE Transactions on Dielectrics and Electrical Insulation, 2005, Vol. 12, No. 4, pp. 629–643.
123. L.N. Rukhadze, E.R. Kutelia, N.I. Maisuradze, B.G. Eristavi, S.I. Bakhtiyarov: *Int. J. Manufac. Technol.*, 2010, vol. 4(2), pp. 75–80.
124. Sawyer, D. M. and Vette, J. I., 1976, "AP-8 Trapped Proton model Environment for Solar Maximum and Minimum", NSSDC/WDC-A-R&S76-06, National Space Science Data Center, Greenbelt, MD.
125. Schneider, T. et al. 2002. Minimum Arc Threshold Voltage Experiments on Extravehicular Mobility Unit Samples. AIAA Paper 2002-1040.
126. Schueler, R., Petermann, J., Schulte, K. and Wentzel, H. P. 1997, Agglomeration and electrical percolation behavior of carbon black dispersed in epoxy resin. J. of Applied Polymer Science, Vol. 63, pp. 1741–1746.
127. Sillars, R. W. 1936, "The properties of a dielectric containing semi-conducting particles of various shapes". Journal of the Institute of Electric Engineers, Vol. 12, pp. 378–394.

128. Sim, A. 2013, Unified Model of Charge Transport in Insulating Polymeric Materials. Logan Utah: Utah State University.
129. Smith, Jr. J. G., Watson, K. A., Thompson, C. M. and Connell, J. W. 2002a. Society for Advancement of Materials and Process Engineering Technical Conference Series, Vol. 34, p. 365.
130. Smith, Jr. J. G., Connell, J. W., Lillehei, P., Watson, K. A. and Thompson, C. M. 2002b. Materials Research Society Spring 2002 Session T. On-line Proceedings. 733E: T3.5.
131. Smith, Jr. J. G., Connell, J. W. Delozier, D. M., Lillehei, P. T., Watson, K. A., Lin, Y., Zhou, B. and Sun, Y.-P. 2004a. Space durable polymer/carbon nanotube films for electrostatic charge mitigation. Polymer, Vol. 45, p. 825–836.
132. Smith, Jr. J. G., Delozier, D. M., Connell, J. W. and Watson, K. A. 2004b, Carbon nanotube-conductive additive-space durable polymer nanocomposite films for electrostatic charge dissipation, Polymer, Volume 45, Pages 6133–6142.
133. Song, Z., Ong, C. and Gong, H. 1997, Secondary and Backscattered Electron Yields of Polymer Surface under Electron Beam Irradiation. Applied Surface Science, Vol. 119 (1), pp. 169–175.
134. Tahara, H., Morishita, A. and Onishi, T. 2004. Ground-based experiments of a spacecraft plasma contactor using a hollow cathode, Vacuum, Volume 73, Issues 3–4, Pages 469–474.
135. Timmermans, M. Y.; Estrada, D.; Nasibulin, A. G.; Wood, J. D.; Behnam, A.; Sun, D. M.; Ohno, Y.; Lyding, J. W.; Hassanien, A.; Pop, Eric., et al. 2012. An Effect of Carbon Nanotube Network Morphology on Thin Film Transistor Performance. Nano Res., 5, 307–319.
136. X. Ting, Z. Zhenghua, X. Zhenxing, Y. Jianwei and J. Minghua: *Adv. Mater. Res.*, 2013, vols. 631–632, pp. 172–75.
137. Toomer, R. and Lewis, T. J., "Charge trapping in corona-charge polyethylene films", Journal of Physics D: Applied Physics, Volume 13(7), pp. 1343–1356.
138. Torkar, K., Fazakerley, A. and Steiger, W. Active Spacecraft Potential Control: Results from the Double Star Project. IEEE Transactions on Plasma Sciences. 2006, 34(5), 2046–2052.
139. Varga, L. and Horvath, E. 1997, Spacecraft 3-dimensional charge deposition modeling. In: Proceedings of the Fourth European Conference on Radiation and Its Effects on Components (RADECS '97) Cannes, IEEE 97th, 82–94.
140. Vette, J. I., 1991, "The AE-8 Trapped Electron Model Environment, NSSDC/WDC-A-R&S 91-24, NASA Goddard Space Flight Center, Greenbelt, MD.
141. Wagner, K. W. 1914, Archiv für Elektrotechnik, Vol. 2, p. 371.
142. Wang, J. and Lay, S. T., "Numerical Simulations on the Effects of Positive Ion Beam Emissions from Negatively Charged Spacecraft", in Proc. AIAA Aerospace Science Meeting, 1996, AIAA-96-0147, p. 12.
143. Wang, J. Y., Wei, L., Peng, F. and Li, Y. 2012. Facile Preparation of Carbon Nanotubes and Graphene Sheets by a Catalyst-Free Refluxing Approach. Nano Res., 5, 640–645.
144. Watson, K. A., Smith, Jr. J. G. and Connell, J. W. 2001. Society for Advancement of Materials and Process Engineering Technical Conference Series, Vol. 33, p. 1551.
145. Watson, K. A., Smith, Jr. J. G. and Connell, J. W. 2003. Society for Advancement of Materials and Process Engineering Technical Conference Series, Vol. 48, p. 1145.
146. Watson, K. A., Ghose, S., Delozier, D. M., Smith, Jr. J. G. and Connell, J. W. 2005, Transparent, flexible, conductive carbon nanotube coatings for electrostatic charge mitigation. Polymer, Vol. 46, pp. 2076–2085.
147. Williams, J. D. and Wilbur, P. J. 1990. Journal Spacecraft Rockets, Vol. 27, pp. 634–641.
148. Wu, M. H., Liu, K. H., Wang, W.L., Sui, Y., Bai, X. D. and Wang, E. G. Ultralong Aligned Single Walled Carbon Nanotubes on Flexible Fluorphlogopite Mica for Strain Sensors. Nano Res. 2012, 5, 443–449.
149. Xu, M.; Futaba, D. N.; Yamada, T.; Yumura, M.; Hata, K. 2010. Carbon Nanotubes with Temperature Invariant Viscoelasticity from –196° to 1000°C. Science, Vol. 330, 1364–1368.
150. J. Ye, D.L. Burris and T. Xie: *Lubricants,* 2016, vol. 4(4), pp. 1–15.
151. J. Ye, A. C. Moore, D.L. Burris: *Tribology Letters*, 2015, vol. 59, pp. 2–11.
152. Ziegler, J. F., 1999. "The Topping of Energetic Light Ions in Elemental Matter", J.pl. Phys. 85, 1249–1272.

Chapter 5
Inertial Morphing as a Novel Concept in Attitude Control and Design of Variable Agility Acrobatic Autonomous Spacecraft

Pavel M. Trivailo and Hirohisa Kojima

Abstract This book chapter presents a systematic overview of the novel concept of "inertial morphing (IM)", first introduced by the authors in 2017 and further expanded in their following publications. It involves deliberate changes of the inertial properties of the system for control of the attitude of the spacecraft.

The "inertial morphing" control concept is essentially based on the realisation that the spinning spacecraft can be seen and utilised as gyroscope itself, instead of utilisation of complex, heavy and energy-consuming gyroscopic devices on-board. Utilisation of the concept, therefore, enables reduction of the weight and dimensions of the conventional systems.

It has been discovered and demonstrated via versatile numerical simulations that IM can be used to enable spacecraft with wide range of attitude control capabilities (e.g. 90° and 180° inversions, de-tumbling and controlled agility acrobatic manoeuvrings). Moreover, it has been also discovered that control of very complex manoeuvres can be achieved with a few only controlled inertial morphing actions (two and three morphings correspondingly for 180° and 90° inversions).

The general control methods presented in this chapter are based on the geometric interpretation of the arbitrary 3D rotational motion of the spacecraft, using angular momentum sphere and kinetic energy ellipsoid in the non-dimensional coordinates. The key control strategies involve combination of installing the angular momentum vector into transition polhodes and installing into transition separatrices.

Reduction in weight and dimensions, simplicity of the implementation of the inertial morphing and simplicity of the attitude control, requiring two or three discrete control actions, make this technology attractive for a variety of applications, especially involving autonomous spacecraft.

One of the remarkable features of the IM control is the ability to access a range of solutions between agile (fast) and prolonged (slow) types and select the most

F. E. Udwadia (✉)
Departments of Aerospace and Mechanical Engineering, Civil and Environmental Engineering, Mathematics, Information and Operations Management, University of Southern California, Los Angeles, CA, USA

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_5

appropriate speed of the undertaking attitude manoeuvre. This added variable agility may be useful, for example, to perform for autonomous spacecraft surveillance, landing or manoeuvring. In particular, the IM may foster effective protection of the spacecraft from hostile environments (asteroids, radiation, etc.), as the spacecraft would be able to quickly expose the most protective surfaces to the sources of danger, hence prolonging survivability of the system. In the other cases of capturing the tumbling spacecraft, the prolonged mode can be selected, allowing more time for the capture and handling.

For the practical implementation of the IM concept, this book chapter also presents a range of conceptual mechanical designs. As Euler's equation for the rotational motion of the rigid bodies paved the way for the development of the theory of gyroscopes and design of various gyroscopic systems, the paradigm of "inertial morphing" may prompt development of new generation of the acrobatic spacecraft with significantly reduced weight and dimensions, reduced cost and enhanced operational capabilities. It may be also possible to design new classes of gyroscopes, possessing an added-on sense of time, which is in contrast to the classical gyroscopes that only possess a sense of orientation.

With a wide spectrum of the presented examples, related to the application of a novel design concept of "inertial morphing", it is believed that presented concept, modelling and simulation of the spinning systems and attitude control method of the spinning systems will be useful not only for the specialists but for a very wide audience, including engineers, scientists, students and enthusiasts of science and space technology.

Keywords Rigid-body dynamics · International Space Station · Spacecraft dynamics · Polhode-to-polhode transfer · Attitude dynamics

5.1 Introduction

There are almost 4900 satellites orbiting the Earth [1], and this number will be non-linearly increasing with time. Continuous control of the attitude is a vital function for spacecraft vehicles. Indeed, communication and observation satellites require directional pointing of their antennae and equipment, for using a single thruster for breaking (used initially for boost) may require 180° attitude reorientation of the spacecraft body ("inversion") to apply thruster force against the motion of the spacecraft, etc. Therefore, attitude dynamics, guidance, navigation and control are the modern research disciplines, requiring new and the most innovative solutions for making new challenging space missions possible and stimulating attention of space engineering community.

Continuous development of new technologies (including miniaturisation of electronic hardware, introduction of new materials) allows significant reduction of the mass of satellites [2]. However, this reduction is in conflict with heavy mass and complexity of the modern attitude control systems, employing gyroscopes. An

attractive alternative of controlling spacecraft without employment of the traditional gyroscopic devices has been proposed by the authors. It involves deliberate changes of the inertial properties of the system, called "inertial morphings (IM)", used for control of the attitude of the spacecraft. "Inertial morphing" control concept is essentially based on realisation that the spinning spacecraft can be seen and utilised as gyroscope itself, instead of utilisation of complex, heavy and energy-consuming gyroscopic devices on-board. It has been recently discovered and demonstrated via versatile numerical simulations that IM can be used to enable spacecraft with wide range of attitude control capabilities (e.g. 90° and 180° inversions, de-tumbling and controlled agility acrobatic manoeuvrings). Moreover, it has been also discovered that control of very complex manoeuvres can be achieved with a few only control inertial morphing actions (two and three morphings correspondingly for 180° and 90° inversions). This book chapter aims to present a systematic overview of the concept of the "inertial morphing", firstly introduced by the authors in 2017 [3] and further expanded in [4–10].

The novel concept of IM enables design and construction of the inertially morphed spacecraft, possessing acrobatic capabilities, and may allow design of new class of gyroscopic systems with a "sense" of time.

5.2 Historical Background

5.2.1 Discovery of the "Garriott's-Dzhanibekov's Effect" in Space

Development of the "inertial morphing" concept was prompted by the flipping motion of the rigid bodies, observed and demonstrated in space.

During his fifth space flight, on June 25, 1985, Vladimir Aleksandrovich Dzhanibekov discovered a spectacular phenomenon: a spinning wing nut in stable flight suddenly, without apparent reasons, changed its orientation by 180° and continued its flight backwards, simultaneously changing its direction of rotation to the opposite. The pattern of the observed as unprovoked 3D flipping motion of the rigid body, which is initially provided with only a one-axis spin, repeated itself in a periodic sequence. This phenomenon was initially widely referred to as Dzhanibekov's effect [11–12]. Vladimir Dzhanibekov himself explained his discovery in various lectures, TV programs and interviews (see Fig. 5.1a).

Performing detailed literature search, we were able to find even earlier demonstrations in space of the flipping motion of the spinning rigid body, dated by 1973. Indeed, interestingly, an experiment with box-shaped space instrument by famous US scientist-astronaut Owen Kay Garriott on-board Skylab 3 in 1973, *12 years before* Dzhanibekov's experiments, demonstrated the flipping motion of the rigid body [13], initiated on purpose by providing it manually with initial energetic spin about the intermediate axis. This immediately resulted in the periodic flipping

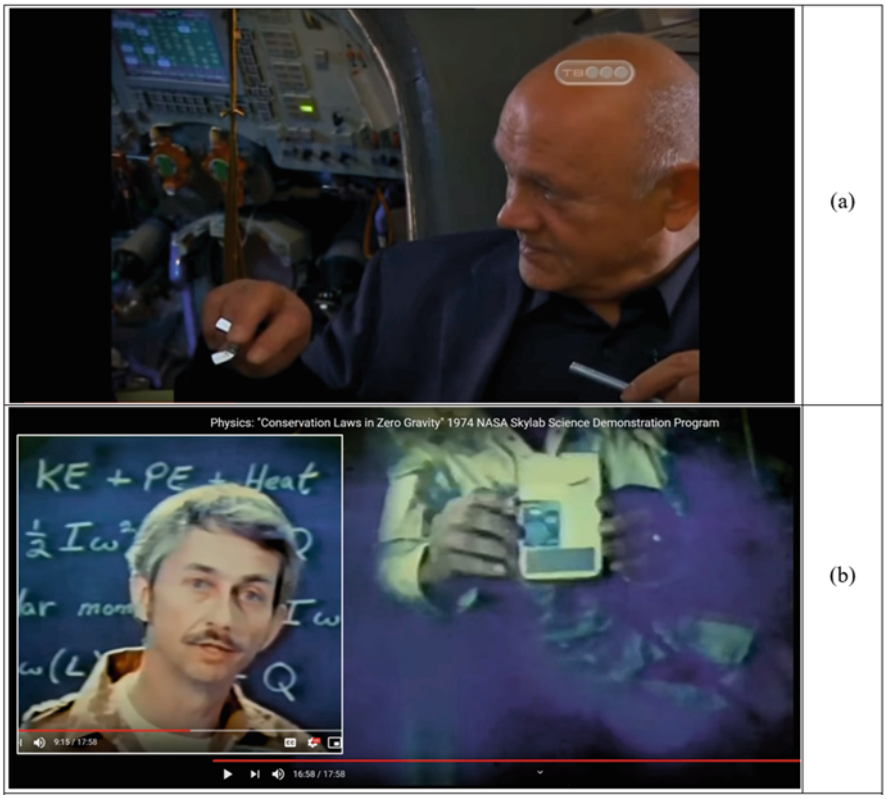


Fig. 5.1 (**a**) Vladimir A. Dzhanibekov, interview at the "Secret Signs" TV program, explaining flipping of the wing nut, observed in 1985 [11–12]; (**b**) Owen K. Garriott, demonstration on-board of "Skylab 3" of the flipping instrument, spun about its intermediate axis, observed in 1973 [13]

motion of the instrument in weightless environment (see Fig. 5.1b). This earlier reference suggests that the use of the "Garriott's-Dzhanibekov's effect" in the future would be more precise name to the observed flipping phenomenon. In both cases, spin to the rigid body was applied in zero gravity by providing a torque impulse about the intermediate axis of inertia of the body, which instantly results in the peculiar rotational motion of the boxed object about this axis with clearly observed periodic flipping about this axis.

Similar experiments have subsequently been run on the International Space Station (ISS), including 30th, 34th and 38th NASA missions. One of the well-known spectacular demonstration involved unscrewed from the base T-handle [14]. These experiments in space clearly demonstrated that a spinning object always rotates in the same direction relative to the observation camera (fixed to the inertial coordinates frame): that means that in the reference frame of the rotating handle, the direction of rotation flips changes each time its orientation flips.

Garriott's-Dzhanibekov's effect has prompted development of theories suggesting that the Earth, similar to the wing nut, performs periodic flips estimated to be at the order of 12,000 years. Evidence in support of this theory includes changes in the Earth's magnetic field [15]. These theories are still debated in the scientific programs [16].

Surprisingly, the Garriott's-Dzhanibekov effect or the "tennis racket theorem" as it is sometimes referred to [17] can be explained by Euler's equations, published in their canonical form in 1758 [18]. During the mid-nineteenth century, Louis Poinsot, a French geometer, developed a geometric interpretation of the physics of rotating bodies that provided a much-welcomed visual counterpart to Euler's algebraic equations [19]. Interestingly, Euler's equations paved the theoretical ground to many scientific manifestations, including Coriolis forces, predicted by Euler but interpreted to the world many years later by French scientist Gaspard-Gustave de Coriolis in 1835. So, heritage by Euler often required time for his ideas to be adopted by scientists. In case of the Euler's equations, it took more than 250 years for scientists to relate the beautiful phenomenon of the flipping motion to these equations. The phenomenon had been conceptually predicted in 1971 by Beachley [20]; however, his work was unnoticed for a very long time and has been left unnoticed, and an in-depth explanation of the phenomenon has only been very recently presented in journal publication [21]. During the last 5 years, interest in the phenomenon has been exponentially increasing. One of the recent interesting references on the topic is authored by Cleve Moler, the founder of Mathworks Company (developing world-famous MATLAB and SIMULINK computer simulation environments), who has been also fascinated with this phenomenon and has a dedicated publication [22].

Derek A. Muller, an Australian-born Canadian science communicator, filmmaker, television personality and inventor, founder of the popular "Veritasium" channel on YouTube, on September 20, 2019, presented a special program [16] dedicated to "Dzhanibekov effect" or "tennis racket theorem", which during the 6 following months attracted almost 7.5 million views, showing an unprecedented interest not only from the scientists and engineers but from the broad community in the discussed topic and its applications.

5.2.2 Demonstrations of the "Garriott's-Dzhanibekov's Effect" on-Board of the ISS

Due to its simplicity and spectacular nature, the Garriott's-Dzhanibekov's effect has become one of the most popular educational and scientific experiments on-board of the International Space Station. It has been reproduced with various rigid-body objects and even liquids. Various videos on these experiments, available in the media and on YouTube, are excellent educational resources.

An amazing visual demonstration in space of the "Garriott's-Dzhanibekov's effect" or "tennis racket theorem" was performed in 2008 by Richard Allen Garriott

Fig. 5.2 Demonstrations of the "Garriott's-Dzhanibekov's effect" on-board of the ISS: (**a**) Richard Garriott; (**b**) Dan Burbank and Anton Shkaplerov, 30th expedition to ISS, 2011; (**c**) Kevin Ford (NASA), 34th expedition to ISS, 2013' (**d**) Koichi Wakata (JAXA), 38th expedition, 2014; (**e**–**f**) "Dancing T-handle" on board of the ISS

de Cayeux [23], a pioneer in commercial space travel and a son of the US scientist-astronaut Owen Kay Garriott. Using nothing more complex than a deck of cards, Richard Allen Garriott demonstrated both stable rotation and tumbling rotation and explained why you can easily spin a rectangular box around two axes but it quickly wobbles out of control if you try to spin it along its intermediate axis (see Fig. 5.2a).

Influence of the shape of the rigid bodies, thus mass distribution in various rigid bodies, including cylinders, cubes and right rectangular prisms, was demonstrated on-board of the ISS by Dan Burbank and Anton Shkaplerov (see Fig. 5.2b), members of the 30th expedition [24].

American astronaut Kevin Ford (NASA) (34th expedition) [25] (see Fig. 5.2c) and Japanese astronaut Koichi Wakata (JAXA) (38th expedition) [26] (see Fig. 5.2d) experimented on-board of the ISS with nothing more complex that pliers. They used

this adjustable geometry tool as an object, capable of intriguing spinning, flipping and tumbling in zero gravity.

One of the most fascinating movies is a continuous short-period flipping of the T-handle on-board of the ISS, fairly called as "dancing T-handle" [14] (see Fig. 5.2e–f). This is a wonderful demonstration of the "Garriott's-Dzhanibekov's effect", which very convincingly illustrates instability of rotation of the rigid body with distinct principal moments of inertia, if the main spin is provided about its principal axis, associated with intermediate moment of inertia.

All these and other demonstrations can be explained with famous Euler's equation.

5.2.3 *Leonard Euler and His Famous Equations for the Rigid-Body Dynamics*

Leonhard Euler (April 15, 1707–Sept. 18, 1783) was a famous Swiss physicist and mathematician (the most eminent of the eighteenth century and one of the greatest in history), who made key contributions to various fields of mathematics and mechanics, leaving long-lasting heritage of more than 500 books and papers (His portrait is presented in Fig. 5.3a.). It has been computed that his publications during his working life averaged about 800 pages a year.

His "Euler's identity" is considered an example of mathematical beauty:

$$e^{i\pi} + 1 = 0 \tag{5.1}$$

called "the most remarkable formula in mathematics" by Richard P. Feynman [27], for its single uses of the notions of addition, multiplication, exponentiation and equality and the single uses of the important constants 0, 1, e, i and π.

In 1988, readers of the *Mathematical Intelligencer* voted it "The Most Beautiful Mathematical Formula Ever". In total, Euler was responsible for three of the top 5 formulae in that poll [28].

His interests are amazingly versatile. Even when dealing with music, Euler's approach is mainly mathematical. His writings on music are not particularly numerous (a few 100 pages, in his total production of about 30,000 pages), but they reflect an early preoccupation and one that did not leave him throughout his life.

Among numerous Euler's works, he developed rigid-body dynamics; very influential publication has a very special place in history. It presented Euler's equations for the dynamics of a rigid body, widely used in modern engineering and science.

Being always fascinated with Euler's scientific work and heritage, the authors were delighted to find in the Euler's archive his original work. It is with greatest pleasure and a profound sense of tribute to Great Euler that we are reproducing in

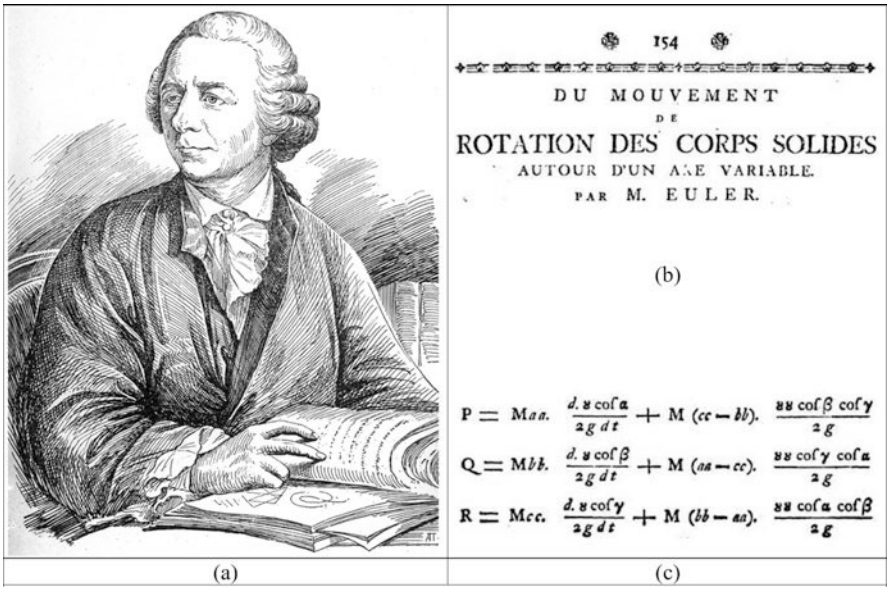


Fig. 5.3 Leonard Euler and his equations for the rigid-body dynamics: (**a**) Leonard Euler's portrait from the University of Tartu collection [29]; (**b**) the title of the historic Leonard Euler's work [18], dated 1758; (**c**) Euler's equations as they appeared in the original L. Euler's work [30]

Fig. 5.3a Euler's portrait from the University of Tartu collection [29] and in Fig. 5.3b the title of the Euler's publication, available from the Euler's archive, and in Fig. 5.3c we show the famous Euler's equations, exactly as they appeared in Euler's original work [30].

In modern language, the Euler's equations in Fig. 5.3c can be written as follows:

$$\begin{cases} \sum N_x = I_{xx}\,\dot{\omega}_x + \left(I_{zz} - I_{yy}\right)\,\omega_y\omega_z \\ \sum N_y = I_{yy}\,\dot{\omega}_y + \left(I_{xx} - I_{zz}\right)\,\omega_z\omega_x \\ \sum N_z = I_{zz}\,\dot{\omega}_z + \left(I_{yy} - I_{xx}\right)\,\omega_x\omega_y \end{cases} \tag{5.2}$$

where x, y, z are the principal axes of inertia fixed to the body; the components of angular velocity in this system are $\omega = (\omega_x, \omega_y, \omega_z)$; the torque is $N = (N_x, N_y, N_z)$; and the diagonal elements of the inertia tensor are I_{xx}, I_{yy} and I_{zz}.

Equations (5.2) are known as "Euler's equations" for a rigid body. They are referred to as equations in principal inertia axes, with the angular velocity components in terms of the angles α, β, γ, which are the angles subtended by the rotation axes with the principal ones fixed in the body. It could be said that these are the Euler angles, although actually they are usually defined by applying the rotation operator to the axes fixed on the body, so that each angle is related to the angular velocities of rotation known as precession, nutation and spin.

5.3 Numerical Modelling and Simulation of the "Garriott's-Dzhanibekov's Effect"

5.3.1 Equations of Motion

Euler's Eq. (5.2), in the general case, can be applied for moments summed about any point P, where P is a point on the rigid body that is attached to a fixed pivot in the inertial reference system. However, in this case the inertia properties should be calculated relative to the point P.

In this work, we will apply the Euler's equations for moments summed about the centre of mass G of the rigid body, free from any external torques ($N_x = N_y = N_z = 0$), and in the further notations, we will imply that I_{xx}, I_{yy} and I_{zz} are principal moments of inertia of the body with respect to the G (which, for brevity, are often denoted as I_x, I_y and I_z):

$$\begin{cases} I_{xx}\ \dot{\omega}_x - \left(I_{yy} - I_{zz}\right)\ \omega_y\ \omega_z = 0 \\ I_{yy}\ \dot{\omega}_y - \left(I_{zz} - I_{xx}\right)\ \omega_z\ \omega_x = 0 \\ I_{zz}\ \dot{\omega}_z - \left(I_{xx} - I_{yy}\right)\ \omega_x \omega_y = 0 \end{cases} \tag{5.3}$$

The matrix form of the above is:

$$\begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \end{Bmatrix} = \begin{Bmatrix} \left(I_{yy} - I_{zz}\right)\ \omega_y\ \omega_z \\ \left(I_{zz} - I_{xx}\right)\ \omega_z\ \omega_x \\ \left(I_{xx} - I_{yy}\right)\ \omega_x\ \omega_y \end{Bmatrix} \tag{5.4}$$

In order to be able to describe instantaneous orientation of a rigid body with respect to a fixed coordinate system, we will use the angles ψ, θ and ϕ, the 3-1-3 Euler angles [31]:

$$\begin{cases} \omega_x &= \dot{\psi}\ \sin\theta\ \sin\phi + \dot{\theta}\ \cos\phi \\ \omega_y &= \dot{\psi}\ \sin\theta\ \cos\phi - \dot{\theta}\ \sin\phi \\ \omega_z &= \dot{\psi}\ \cos\theta + \dot{\phi} \end{cases} \tag{5.5}$$

which can also be written in the matrix form:

$$\begin{bmatrix} \sin\theta & \sin\phi & \cos\phi & 0 \\ \sin\theta & \cos\phi & -\sin\phi & 0 \\ \cos\theta & & 0 & 1 \end{bmatrix} \begin{Bmatrix} \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{Bmatrix} = \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \tag{5.6}$$

For solving the rigid-body dynamics problems, using numerical methods, we combine matrix Eqs. (5.4) and (5.6) into a single equation:

$$\begin{bmatrix} I_{xx} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{yy} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{zz} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sin\theta & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & \sin\theta & \cos\phi & -\sin\phi & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & 1 \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \\ \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{Bmatrix} = \begin{Bmatrix} \left(I_{yy} - I_{zz}\right) \omega_y \omega_z \\ \left(I_{zz} - I_{xx}\right) \omega_z \omega_x \\ \left(I_{xx} - I_{yy}\right) \omega_x \omega_y \\ \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \tag{5.7}$$

This matrix equation can be solved directly, or task for more robust solution can be reformulated in terms of quaternions.

5.3.2 Programming Considerations

Ordinary differential equations can be efficiently solved using Runge-Kutta methods.

MATLAB® has a set of specialised procedures, including ode45, ode23, ode113, ode15s, ode23s, ode23t, ode23tb, ode15i, to deal with various tasks, for example, described by the ordinary differential equation in the classical form:

$$\left\{\dot{\mathbf{x}}\right\} = \{f\,(t, \mathbf{x})\} \tag{5.8}$$

There is also a very useful option enabling solution of the problems, involving the so-called "mass" matrix M:

$$[M\,(t, \mathbf{x})]\left\{\dot{\mathbf{x}}\right\} = \{f\,(t, \mathbf{x})\} \tag{5.9}$$

This option, accessible via the odeset, in some cases can improve efficiency and can also handle cases when the mass matrix is singular (non-invertible). As it can be seen, our matrix Eq. (5.7) corresponds to the format given with Eq. (5.9); therefore, we use MATLAB® ode procedure in conjunction with the "mass matrix" option to simulate dynamic behaviour of the simulated spacecraft models.

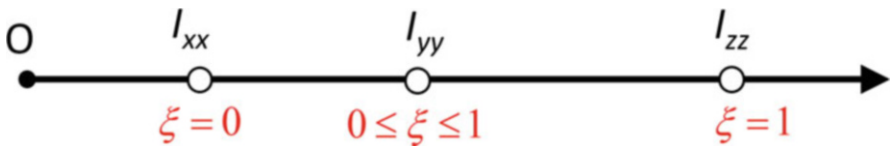


Fig. 5.4 Introduction of the non-dimensional parameters, describing relative values of the principal moments of inertia of the system

5.3.3 *Non-dimensional Formulation of the Equations*

For the main derivations in this chapter, it will be typically assumed that the system has *three distinct* principal moments of inertia, which are *arranged* in the following order: $I_{xx} < I_{yy} < I_{zz}$. For more generic formulations, two non-dimensional parameters, η and ξ, both restricted in their values within the range between 0 and 1 can be introduced:

$$\eta = \frac{I_{xx}}{I_{zz}}; \qquad \xi = \frac{I_{yy} - I_{xx}}{I_{zz} - I_{xx}}; \qquad (0 < \eta < 1; \quad 0 < \xi < 1) \tag{5.10}$$

Parameter ξ in this case would have a similar meaning of the non-dimensional coordinate "counterpart" from the finite element method, defining the current position within the finite element. In the context of this study, ξ is specifying the non-dimensional relative position coordinate of the intermediate value of the moment of inertia between the minimum value of the moment of inertia I_{xx} and the maximum value of the moment of inertia I_{zz} (see Fig. 5.4). In other words, it can be said that ξ is the non-dimensional parameter in the Hermite functions, enabling calculation of I_{yy} using I_{xx} and I_{zz}, using the following relationship:

$$I_{yy} = I_{xx}\,(1 - \xi) + I_{zz}\,\xi \tag{5.11}$$

The zero value of ξ would now correspond to I_{xx}, and unit value of ξ would correspond to I_{zz} and any intermediate value of ξ, expressed via $0 < \xi < 1$, would correspond to I_{yy}. With these notations, we can also derive several relationships, enabling useful conversions in the future:

$$I_{yy} = I_{xx}\left(1 - \xi + \frac{1}{\eta}\right); \qquad I_{zz} = \frac{I_{xx}}{\eta}. \tag{5.12}$$

As illustration, we take $I_{xx} = 2$; $I_{yy} = 3$; $I_{zz} = 4$ [all in kg $\times$ m^2] and can see from Eqs. (5.10, 5.11 and 5.12) that these system's parameters would correspond to $\xi = 0.5$ and $\eta = 0.5$.

Furthermore, in many cases, additional advantages could be gained if the Euler equations can be also rewritten in the non-dimensional quantities [32]:

$$\begin{aligned}
&\frac{d}{d\bar{t}}\overline{H}_{xx} + \left(1 - \frac{I_{yy}}{I_{zz}}\right)\overline{H}_{yy}\overline{H}_{zz} = 0,\\
&\frac{d}{d\bar{t}}\overline{H}_{yy} + \left(\frac{I_{yy}}{I_{zz}} - \frac{I_{yy}}{I_{xx}}\right)\overline{H}_{zz}\overline{H}_{xx} = 0,\\
&\frac{d}{d\bar{t}}\overline{H}_{zz} + \left(\frac{I_{yy}}{I_{xx}} - 1\right)\overline{H}_{xx}\overline{H}_{yy} = 0,
\end{aligned} \tag{5.13}$$

where non-dimensional time is calculated as:

$$\bar{t} = \left(\sqrt{2K_0/I_{yy}}\right)\, t \tag{5.14}$$

where K_0 is the elliptic integral of the first kind.

5.3.4 *Numerical Simulation of the "Garriott's-Dzhanibekov's Effect": Illustration Case*

We reproduce simulation results from [3] for the case study, in which the following parameters were employed: $I_{xx} = 0.3$, $I_{yy} = 0.35$, $I_{zz} = 0.4$ (all in kg $\times$ m^2), corresponding to $\xi = 0.5$ and $\eta = 0.75$, with the initial conditions $\omega_{x0} = 0.1$, $\omega_{y0} = 15$, $\omega_{z0} = 0.1$ (all in rad/s). The main results are shown in Fig. 5.5 for completeness of the presentation.

Figure 5.5a shows that in this illustration case, the initial dominant angular velocity about the system's y-body axis is subject to the periodical change of its initial value to the opposite in sign value, symbolising the flipping motion with 180° change in the orientation of the rigid body. From the simulation results, it can be seen that the period of the flips is equal to $T = 12.3$ s.

Figure 5.5a confirms that during the "flipping" motion, the angular momentum in the system is conserved.

At last, Fig. 5.5b shows that while ψ is monotonically increasing, the ϕ pattern is quite different: there are evident "plateau" segments corresponding to small changes in ϕ around 0°, 180°, 360°, etc. However, the most important observation in the context of this work is the presence of the multiple zero crossings for various components of the angular velocity, in particular, for ω_x, ω_y and ω_z in the presented test case.

5.4 Calculation of the Period of the Flipping Motion

We assume that I_{yy} is intermediate value of the principal moment of inertia. Then the period of the observed unstable motion can be estimated, using Eq. (37.12) in page 154 from the L.D. Landau's reference [33]:

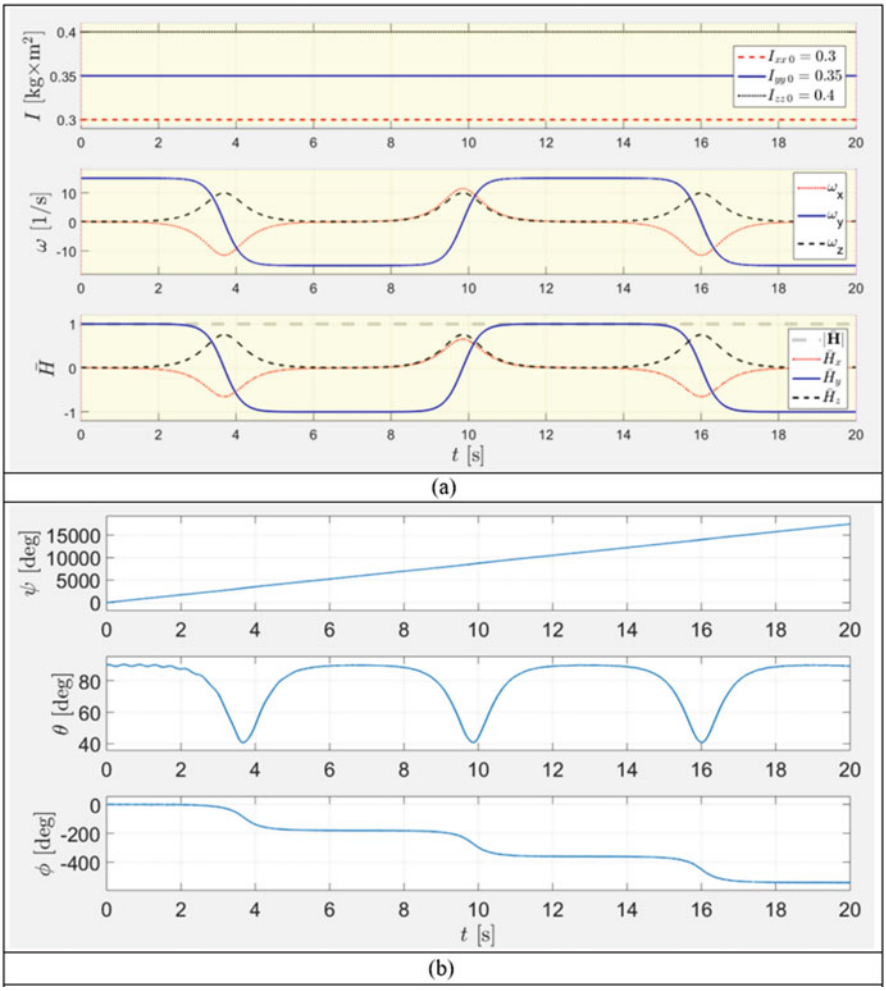


Fig. 5.5 Time histories of the key for the Garriott's-Dzhanibekov's effect flipping motion: (**a**) moments of inertia, angular velocity components and non-dimensional angular momentum components; (**b**) Euler angles

$$\text{If } H^2 > 2K_0\, I_{yy}, \tag{5.15}$$

then

$$T = 4K\sqrt{\frac{I_{xx}I_{yy}I_{zz}}{\left(I_{zz}-I_{yy}\right)\left(H^2-2K_0I_{xx}\right)}} \tag{5.16}$$

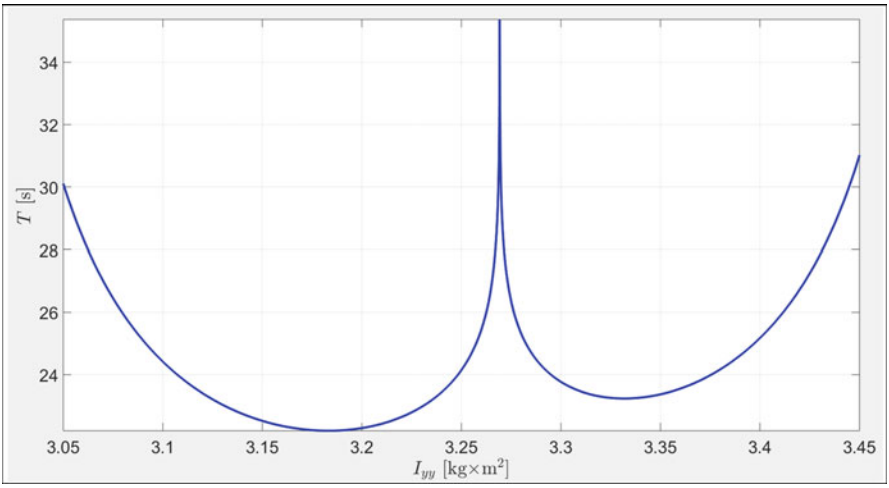


Fig. 5.6 Period T of the unstable flipping motion ("Garriott's-Dzhanibekov's effect" case) as a function of intermediate moment of inertia I_{yy} (for the following example: $I_{xx} = 3$; $I_{zz} = 3.5$ [kg × m^2] ; $\omega_{x,i} = 0.1$, $\omega_{y,i} = 15$, $\omega_{z,i} = 0.1$ [rad/s])

$$\text{If } H^2 < 2K_0\, I_{yy}, \tag{5.17}$$

then

$$T = 4K\sqrt{\frac{I_{xx} I_{yy} I_{zz}}{\left(I_{xx} - I_{yy}\right)\left(H^2 - 2K_0 I_{zz}\right)}}, \tag{5.18}$$

where K is complete elliptic integral of the first kind:

$$K = \int_0^1 \frac{ds}{\sqrt{\left(1 - s^2\right)\left(1 - k^2 s^2\right)}} = \int_0^{\pi/2} \frac{du}{\sqrt{1 - k^2 \sin^2 u}} \tag{5.19}$$

As an illustrative example, let us assume the following parameters of the system: $I_{xx} = 3$, $I_{zz} = 3.5$ (all in kg × m^2), corresponding to $\eta = 0.8571$ with the initial conditions $\omega_{x0} = 0.1$, $\omega_{y0} = 15$, $\omega_{z0} = 0.1$ (all in rad/s). For this case we will use Eqs. (5.15, 5.16, 5.17, 5.18 and 5.19) and will illustrate the influence of the intermediate moment of inertia I_{yy} of the system on the period T of the unstable flipping motion. Resulting plot is presented in Fig. 5.6.

The shape of the plot in Fig. 5.6 is clearly asymmetrical, but enabling variation of the period T within the wide range. Most significantly, Fig. 5.6 shows that the period T of flipping motion is bounded with particular minimum value, and there are two values of I_{yy}, providing with local minima values of the flipping periods T within the range of I_{yy} between the minimum value of the moment of inertia I_{xx} and maximum value of the moment of inertia I_{zz}.

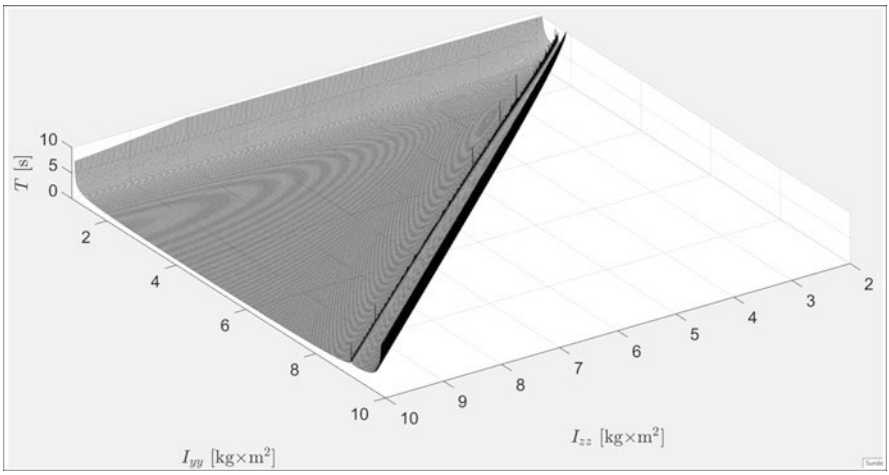


Fig. 5.7 Period T of the unstable flipping motion ("Garriott's-Dzhanibekov's effect" case) of the rigid body, as a function of its moments of inertia I_{yy} and I_{zz}

Setting I_{xx} to a fixed value, and running variation for I_{zz} and I_{yy}, we can also calculate more generic plot, showing influence of these two other principal moments of inertia on the period of the unstable motion. The resultant plot is shown in Fig. 5.7, giving higher resolution for smaller values of T by changing the T-axis limits. This is a very interesting plot, which shows more generic nature of the asymmetry, observed in Fig. 5.6. The plot has also very recognisable "ridge" over the combination of moments of inertia I_{yy} and I_{zz} which results in very high periods.

As a second illustrative example, let us assume the following parameters of the system, in which $I_{xx} = 3$, $I_{zz} = 3.5$ (all in kg×m^2), corresponding to $\eta = 0.8571$, with the initial conditions $\omega_{x0} = 0.1$, $\omega_{y0} = 15$, $\omega_{z0} = 0.1$ (all in rad/s). For this case we will use Eqs. (5.15, 5.16, 5.17, 5.18 and 5.19) and will illustrate the influence of the intermediate moment of inertia I_{yy} of the system on the period of the unstable flipping motion. The resulting plot, presented in Fig. 5.7, is clearly being asymmetrical and could be easily regarded by many as counter-intuitive, as there may be a wrongly perceived assumption of the "symmetrical" influence of I_{yy} on period T.

This plot prompts that variation in the intermediate value of the moment of inertia between I_{xx} and I_{zz} (i.e. changing the ξ value) can allow changes of the period T of the flipping motion within wide range. However, there is a minimum value of the period, which could not be reduced further. For the example shown, the lower threshold of the period is slightly higher than 22.2 s. Also, there is a specific value of the I_{yy} which leads to the infinitely large value of the T. For the example shown, this corresponding value of I_{yy} is approximately 3.27 kg × m^2.

Allowing variation of the I_{yy} and I_{zz} values (i.e. ξ and η non-dimensional parameters), we can also calculate more generic plot, showing influence of these principal moments of inertia on the period T of the unstable motion. The resultant plot is

shown in Fig. 5.7. This is an interesting plot, which shows more generic nature of the asymmetry, observed in Fig. 5.6. However, most significant observation in Fig. 5.7 is that for each of the I_{zz} values, there is a value of I_{yy} which leads to the infinitely large period of the flipping motion. We named this area as "high periods ridge".

The plot in Fig. 5.7 is similar in shape to the plot in Fig. 5.6a, and it similarly prompts that when I_{yy} is approaching any of the other moments of inertia, I_{xx} or I_{zz}, then the period of the flipping is asymptotically approaching infinite values.

Similarly to the examples above, in the second illustrative example, we initially assume initial conditions $\omega_{x0} = 0.01$, $\omega_{y0} = 1.5$, $\omega_{z0} = 0.01$ (all in rad/s) for the system with $I_{xx} = 2$, $I_{zz} = 4$ (all in kg $\times$ m^2), which corresponds to $\eta = 0.5$ (see Eq. 5.10) and plot the flipping period as a function of the intermediate moment of inertia I_{yy}, varying its value in-between the minimum value of the moment of inertia I_{xx} and maximum value of the moment of inertia I_{zz}. The resultant plot is shown in Fig. 5.8a with continuous red line. It allows determination of the flipping period. This value is equal to 47.16 s. Let us now, in addition to the above, consider a similar "variable I_{yy}" experiment, changing only the maximum moment of inertia value from $I_{zz} = 4$ to $I_{zz} = 5$, which would correspond to $\eta = 0.4$. The resultant plot for the period is shown with dotted blue line in Fig. 5.8a. Comparison of the two curves allows to suggest another avenue for manipulation with the period of the flipping motion by changing the ratio between I_{xx} and I_{zz}, i.e. η value. Figure 5.8b also shows a ridge with high values of the flipping periods. It can be observed that for the higher values of I_{zz}, this ridge has more offset towards I_{zz}, than towards I_{xx}.

5.4.1 *Influence of the Value of the Angular Velocity ω_y of the Predominant Spin on the Period T of the Flipping Motion*

In this subsection, we consider systems with non-zero initial angular momentum H_0. In case of the system with predominant spin about its intermediate axis, the major contributors to H_0 are the ω_y and I_{yy}.

Using Eqs. (5.15, 5.16, 5.17, 5.18 and 5.19), we can represent T as a 3D surface plot, explaining influence on the $T(\omega_y, I_{yy})$ function of its two argument. The resultant plot is shown in Fig. 5.9. It clearly reveals the tendency of the periods to become very large, when I_{yy} is approaching to I_{xx} or I_{zz}. However, the plot also reveals the ridge of high value of periods, being asymmetrically in-between I_{xx} and I_{zz}. As surface gradient is very high in vicinity of the ridge, it should be avoided for practical implementations, because in this area T would be very sensitive to small changes in I_{yy}, which would make control of the system period impractical.

As an example, in Fig. 5.9 we also intersect the T surface with two illustrative level values of the period: $T = 30$s and $T = 50$s. The intersection lines show that for

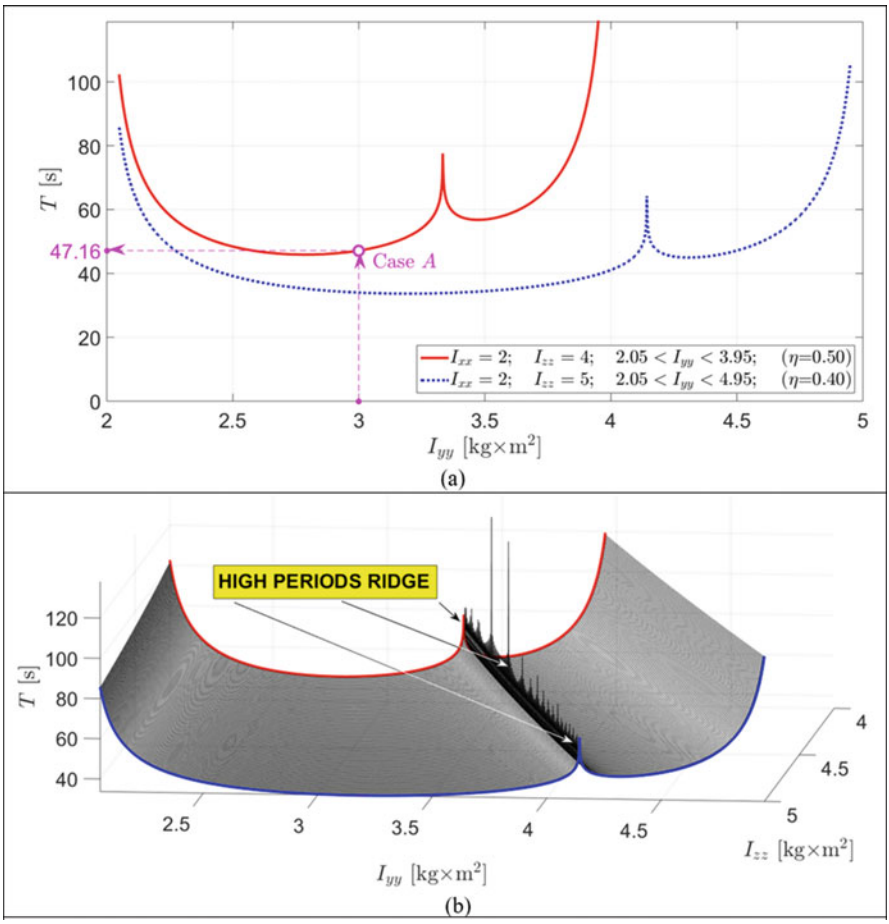


Fig. 5.8 Period of the unstable flipping motion ("Garriott's-Dzhanibekov's effect" case) of the rigid body, as a function of its moments of inertia I_{yy} and I_{zz} for two variation experiments: (**a**) variation of I_{yy} only in two cases $I_{xx} = 2;\ I_{zz} = 4$; and $I_{xx} = 2;\ I_{zz} = 5$ [kg × m^2] ; $\omega_{x,i} = 0.01$, $\omega_{y,i} = 1.5,\ \omega_{z,i} = 0.01$ [rad/s]; (**b**) variation of both, I_{yy} and I_{zz} in the case $I_{xx} = 2$ [kg × m^2] ; $\omega_{x,i} = 0.01, \omega_{y,i} = 1.5,\ \omega_{z,i} = 0.01$ [rad/s]

the desired value of the flipping period T, there are multiple matching combinations of ω_y and I_{yy}; however, if the goal of the selection is to minimise the spin rate, then, there may be two local minimum specific values of I_{yy}. Two contour lines for $T = 30$s and $T = 50$s are shown separately in Fig. 5.10. It shows, that, if, for example, the aim of the design process is to keep ω_y low, for the $T = 50$, there are two solutions for I_{yy}, approximately equal to $I_{yy} = 3.335$ and $I_{yy} = 3.18$.

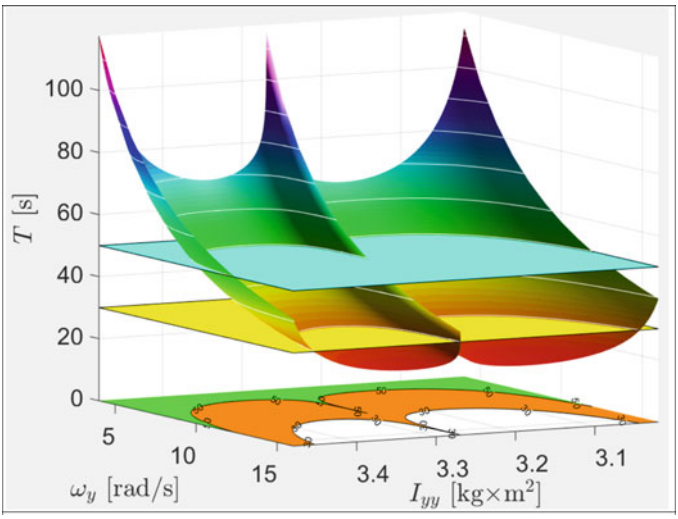


Fig. 5.9 Variation of the period $T(\omega_y, I_{yy})$ of the flipping motion with the changes in the predominant spinning angular velocity ω_y and value of the intermediate moment of inertia I_{yy} for the system with $I_{xx} = 3$, $I_{zz} = 3.5$ kg $\times$ m^2 (i.e. $\eta = 0.8571$)

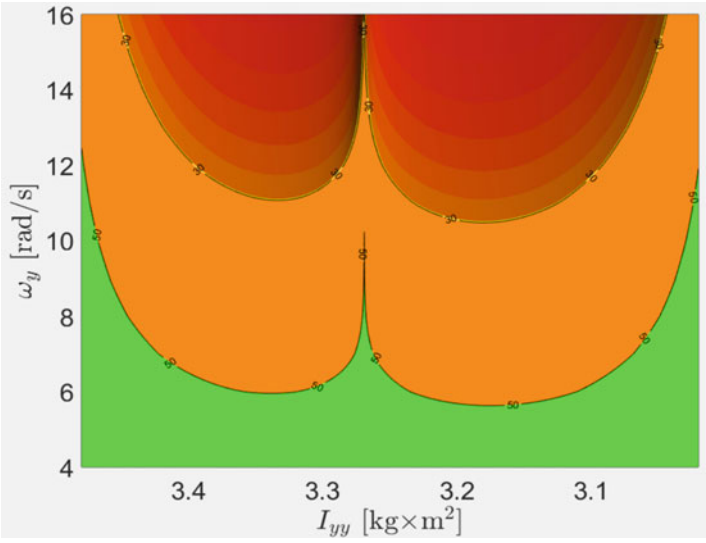


Fig. 5.10 Two labelled contour lines for the ω_y (T, I_{yy}) surface in Fig. 5.9, corresponding to the values of the flipping periods equal to $T = 30$ s and $T = 50$ s

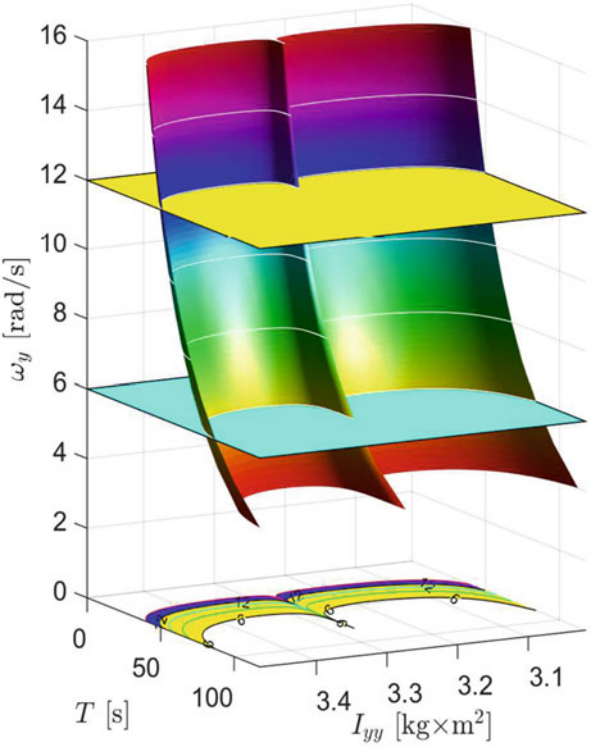


Fig. 5.11 Variation of the period ω_y (T, I_{yy}) of the flipping motion with the changes in the predominant spinning angular velocity T and value of the intermediate moment of inertia I_{yy} for the system with $I_{xx} = 3$, $I_{zz} = 3.5$ kg $\times$ m^2 (i.e. $\eta = 0.8571$)

5.4.2 *Influence of the Value of the Period T of the Flipping Motion on the Angular Velocity ωy of the Predominant Spin*

Results in Fig. 5.9 are presented for the $\omega_y = \omega_y$ (T, I_{yy}), being a function of two arguments, T and I_{yy}. This function is shown in Fig. 5.11 as a 3D surface plot. For the illustration purposes, we assume interest in two special values of angular velocity of the predominant rotation: $\omega_y = 6$ and $\omega_y = 12$ rad/s. The intersection lines of these level panes and the ω_y surface, together with other contour curves, are given in Fig. 5.12.

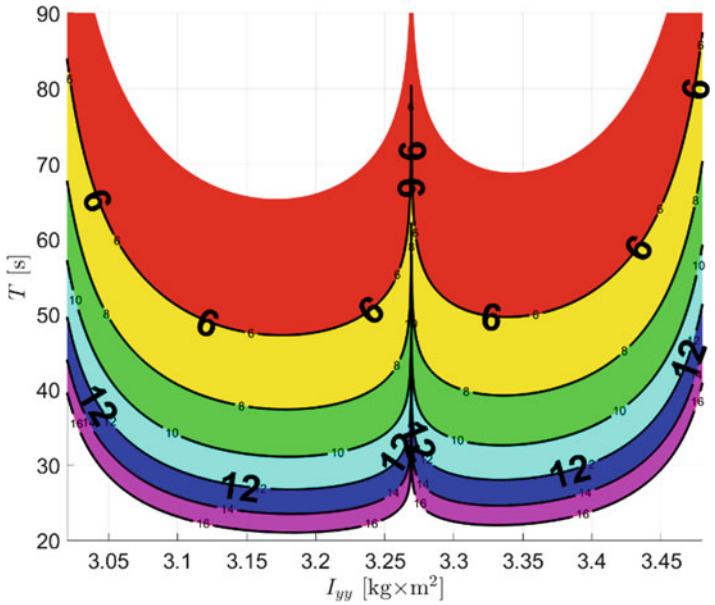


Fig. 5.12 Contour lines for the $T(\omega_y, I_{yy})$ surface in Fig. 5.11 with in-between bands individually coloured. Red band corresponds to the $\omega_y = 4$–6 (rad/s) spin rate range, yellow band to the $\omega_y = 6$–8 range, green band to the $\omega_y = 8$–10, cyan band to the $\omega_y = 10$–12, blue band to the $\omega_y = 12$–14 range and purple band to the $\omega_y = 14$–16 range

5.5 Geometric Interpretation of the 3D Rotational Dynamics of Rigid Objects

5.5.1 General Comments

In this section, we will present geometric interpretation of the 3D rotational dynamics of the rigid body, employing *angular momentum spheres*, *kinetic energy ellipsoids* and *hodographs* – lines produced by the tip of the vector of the non-dimensional angular momentum, changing with time. The current angular momentum is then represented by the vector from the origin to a point on the hodograph.

5.5.2 Angular Momentum Sphere

Let us consider arbitrary free motion of the rigid body. In the context of the rigid-body dynamics, the angular momentum vector is defined as the product of matrix of the moments of inertia and the angular velocity vector. In the following presentation,

for the modelling purposes, we will use the Cartesian coordinate system with its origin at the centre of the mass of the system and its axes x, y and z being the principal axes of inertia of the modelled rigid body. With these, the angular momentum vector can be represented with its components H_x, H_y and H_z:

$$\overrightarrow{\mathbf{H}}_G = H_x \overrightarrow{\mathbf{e}}_x + H_y \overrightarrow{\mathbf{e}}_y + H_z \overrightarrow{\mathbf{e}}_z \tag{5.20}$$

where $\overrightarrow{\mathbf{e}}_x$, $\overrightarrow{\mathbf{e}}_y$ and $\overrightarrow{\mathbf{e}}_z$ are unit vectors along x, y and z orthogonal directions.

For exploring 3D rotational dynamics, including "Garriott's-Dzhanibekov's effect", we will utilise the fundamental law of conservation of angular momentum, implying that the angular momentum can be exchanged between objects in a closed system, but total angular momentum before and after an exchange remains constant (is conserved). Therefore, at any moment of the simulation, the length of the angular momentum vector must remain constant. We can express the squared length of the vector $\overrightarrow{\mathbf{H}}_G$ and equate it to the initial value, which would be known at the beginning of the simulation:

$$[H_x(t)]^2 + \left[H_y(t)\right]^2 + [H_z(t)]^2 = \left|\overrightarrow{\mathbf{H}}_G(0)\right|^2 \tag{5.21}$$

This can be rewritten as follows:

$$\left[\frac{H_x(t)}{H(0)}\right]^2 + \left[\frac{H_y(t)}{H(0)}\right]^2 + \left[\frac{H_z(t)}{H(0)}\right]^2 = 1 \tag{5.22}$$

Now, let us rewrite previous equation in terms of the non-dimensional quantities, $\overline{H}_x$, $\overline{H}_y$and $\overline{H}_z$

$$\left[\overline{H}_x(t)\right]^2 + \left[\overline{H}_y(t)\right]^2 + \left[\overline{H}_z(t)\right]^2 = 1 \tag{5.23}$$

where

$$\begin{aligned}
\overline{H}_x(t) &= \frac{H_x(t)}{H_0} = \frac{I_{xx}\omega_x}{\sqrt{(I_{xx}\omega_x)^2+(I_{yy}\omega_y)^2+(I_{zz}\omega_z)^2}} \\
\overline{H}_y(t) &= \frac{H_y(t)}{H_0} = \frac{I_{yy}\omega_y}{\sqrt{(I_{xx}\omega_x)^2+(I_{yy}\omega_y)^2+(I_{zz}\omega_z)^2}} \\
\overline{H}_z(t) &= \frac{H_z(t)}{H_0} = \frac{I_{zz}\omega_z}{\sqrt{(I_{xx}\omega_x)^2+(I_{yy}\omega_y)^2+(I_{zz}\omega_z)^2}}
\end{aligned} \tag{5.24}$$

Equation (5.23) represents a unit sphere, called angular momentum sphere (AMS).

The graphical interpretation of Eq. (5.23) is a sphere with unit radius in the non-dimensionalised angular momentum coordinates and is shown in Fig. 5.13a.

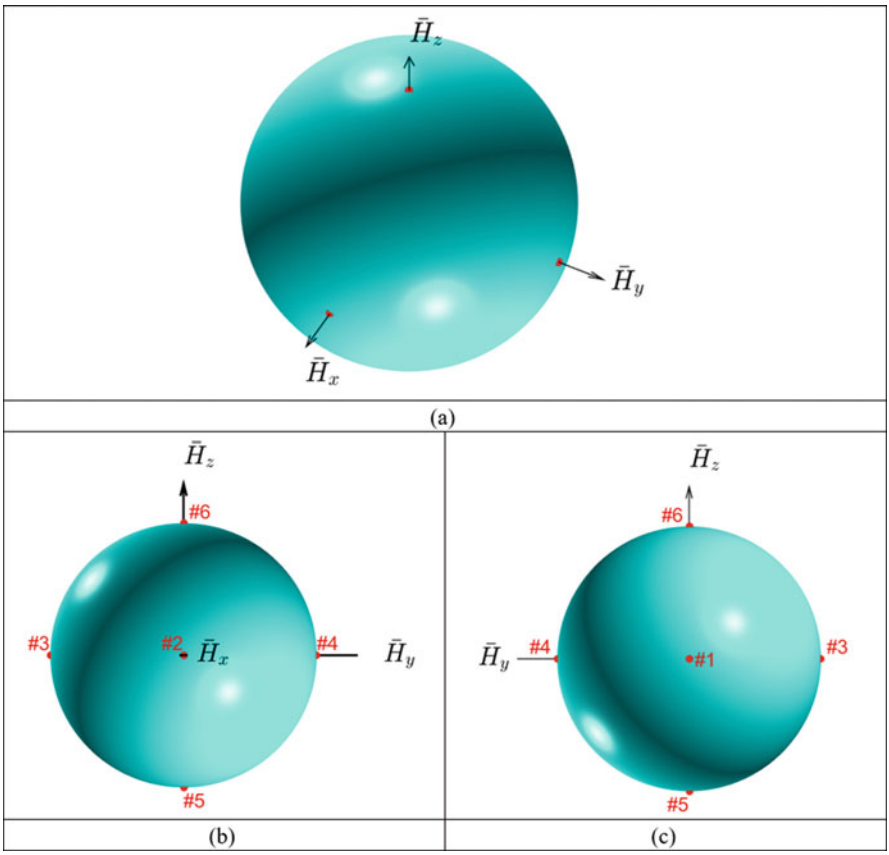


Fig. 5.13 Angular momentum sphere (AMS): (**a**) 3D view of the unit AMS in the non-dimensionalised angular momentum coordinates; (**b**) $+H_x$ view with visible "parking" points #2, #3, #4, #5 and #6 shown; (**b**) $-H_x$ view with visible "parking" points #1, #3, #4, #5 and #6 shown

Amazingly, it would correspond to all rigid bodies and all possible torque-free motions of the rigid bodies.

5.5.3 Utilisation Angular Momentum Sphere and Its Feasible Godographs for the Non-dimensional Angular Momentum Vector as Strategic Basis for the Methods of Attitude Control of the Rotating Systems

In case of a typical Garriott's-Dzhanibekov's effect "flipping" motion, spinning of the system could be seen as a periodic travel of the *hodograph* (line, drawn by the

tip of the non-dimensional angular momentum vector $\overline{\mathbf{H}}$ on the unit AMS) along the closed path on the AMS. The angular momentum godograph is also called polhode and will be discussed in the next section. The motion of the hodograph in the vicinity of the pole (potential "parking" point) is rather slow, and the switch between two points of the intersection of the AMS with the y-axis – negative y point $[0, -1, 0]$ and positive y point $[0, 1, 0]$ – is occurring rather rapidly.

For the demo case, these points are numbered as #3 and #4 in Fig. 5.13b and c. For the selected y rotation example, the simple inversion method [5, 6] can be used, and the spacecraft could be stabilised around these two opposite points, which we will call "potential parking" points or just "parking" points. However, there are two more pairs of the "parking" points on the axes x and z. All six "parking" points are shown in Fig. 5.13b and c.

We have a special interest in the poles, as we see them as parking points to be utilised when the arbitrary 3D motion is to be stabilised and reduced to the regular spin about the body axes. Often, transformation of motion would involve a transfer of the tip of the non-dimensional angular momentum vector from one parking point to another. This could only be achieved if there is a feasible godograph for the vector $\overline{\mathbf{H}}$. And, as it will be seen in the next section, there are only special types of trajectories, separatrices, which are crossing the poles, and these trajectories are for the special type of the rotational motion of the rigid body – "Garriott's-Dzhanibekov's effect" flipping motion. Therefore, we place special emphasis on the establishment of the connection between the two transfer points for the vector $\overline{\mathbf{H}}$, which can be constructed from segments of polhodes and/or separatrices.

In this work, we are proposing to use IM for controlled transfer of the system into unstable "flipping" motion and use transition from one parking point to another parking point for switching trajectory of the hodograph of the vector $\overrightarrow{\mathbf{H}}_G(t)$ to other parking points, i.e. not necessarily being opposite to the established flipping.

If this is achieved, the spacecraft could be stabilised around desired/targeted parking point, as per the morphing procedure in [3, 4]. Stabilisation of the angular momentum vector in any of the six parking points would mean that the system would perform most predominant rotation, associated with only one of the selected body axis, x, y or z, passing through the targeted parking point and almost no rotation about two other orthogonal body axes. This is the reason for aiming to get to any of these points and then to stop flipping motion, as per [3–10].

5.5.4 *Polhodes on the Angular Momentum Sphere*

Let us solve Euler's equations in the matrix form (5.7) and calculate time responses of the same rigid body due to different initial conditions. As contrast illustrations, we consider three cases A, B and C with the following initial conditions, applied to the rigid body with $I_{xx} = 2$, $I_{xx} = 3$, $I_{xx} = 4$ (all in kg $\times$ m^2):

- Case-A: $\omega_{x0} = 0.01$, $\omega_{y0} = 1.5$, $\omega_{z0} = 0.01$ (all in rad/s)

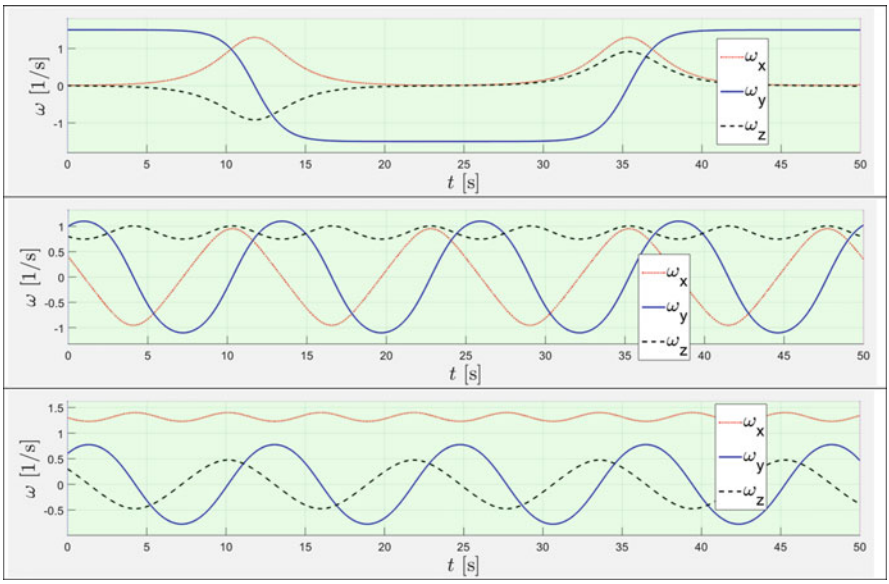


Fig. 5.14 Time histories for angular velocity components $\omega_x,\ \omega_y,\ \omega_z$ for three contrast cases of initial conditions: (***A***) $\omega_{x,i} = 0.01, \omega_{y,i} = 1.5,\ \omega_{z,i} = 0.01$; (***B***) $\omega_{x,i} = 0.4, \omega_{y,i} = 1,\ \omega_{z,i} = 0.8$; (***C***) $\omega_{x,i} = 1.3, \omega_{y,i} = 0.6,\ \omega_{z,i} = 0.3$ (here and further all angular velocities are given in rad/s)

- Case-*B*: $\omega_{x0} = 0.4, \omega_{y0} = 1, \omega_{z0} = 0.8$ (all in rad/s)
- Case-*C*: $\omega_{x0} = 1.3, \omega_{y0} = 0.6, \omega_{z0} = 0.3$ (all in rad/s)

Time histories for the angular velocity components for these three cases are presented in Fig. 5.14.

From a distinct response in Fig. 5.14a, we can conclude that the case-*A* response corresponds to a classical "Garriott's-Dzhanibekov's effect" flipping motion, whereas two other responses correspond to tumbling motion.

Angular momentum sphere can be used in the cases *A*, *B* and *C* not only to mark their respective initial conditions but also to show the corresponding resultant motion as lines on the AMS, called *polhodes*. Polhodes are trajectories of the tips of the non-dimensional angular momentum vectors, $\overline{\mathbf{H}}_A$, $\overline{\mathbf{H}}_B$ and $\overline{\mathbf{H}}_C$ (these three polhodes are marked with the red colour in Fig. 5.15), where we also show for comprehensive visualisation three quiver plots for the $\overline{\mathbf{H}}_A$, $\overline{\mathbf{H}}_B$ and $\overline{\mathbf{H}}_C$ vectors, superimposed over the AMS and polhodes.

We should emphasise that for plotting polhodes on the AMS we are using non-dimensional angular momentum coordinates $\overline{H}_x,\ \overline{H}_y,\ \overline{H}_z$, defined as per Eq. (5.24).

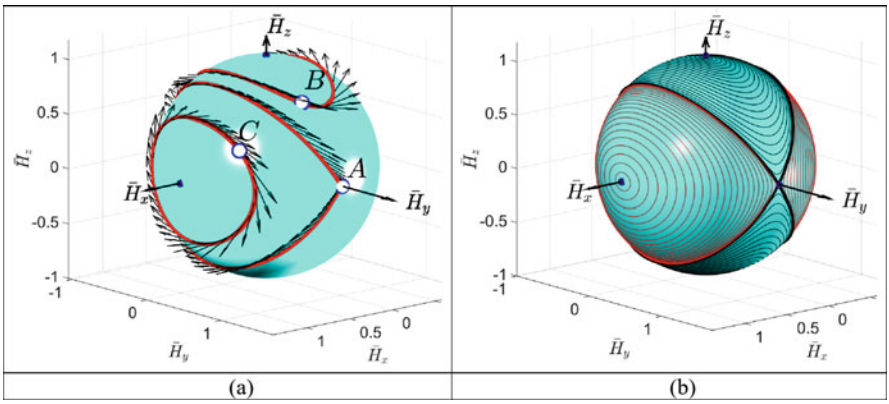


Fig. 5.15 Polhodes: (**a**) for cases *A*, *B* and *C* in Fig. 5.14; (**b**) examples of broad coverage of initial conditions and associated responses

5.5.5 *Kinetic Energy Ellipsoid*

Let us express the kinetic energy of the rotating body in terms of the angular momentum components:

$$\frac{1}{2}I_{xx}\omega_x^2 + \frac{1}{2}I_{yy}\omega_y^2 + \frac{1}{2}I_{zz}\omega_z^2 = \left[\frac{H_x(t)}{\sqrt{2I_{xx}}}\right]^2 + \left[\frac{H_y(t)}{\sqrt{2I_{yy}}}\right]^2 + \left[\frac{H_z(t)}{\sqrt{2I_{zz}}}\right]^2 = K(t) \tag{5.25}$$

The classical demonstrations of the "Garriott's-Dzhanibekov's effect" are typically considered rigid bodies with fixed values of the inertial properties.

However, in our study, it is essential to consider more general case, allowing for the moments of inertia to change with time. The imperative importance of this feature will be explained later.

$$\left[\frac{H_x(t)}{\sqrt{2I_{xx}(t)}}\right]^2 + \left[\frac{H_y(t)}{\sqrt{2I_{yy}(t)}}\right]^2 + \left[\frac{H_z(t)}{\sqrt{2I_{zz}(t)}}\right]^2 = K(t) \tag{5.26}$$

We, then, divide both sides of this equation by constant $H^2(0)$ and rearrange result in terms of non-dimensional quantities $\overline{H}_x$, $\overline{H}_y$ and $\overline{H}_z$:

$$\left[\frac{H_x(t)}{H(0)}\frac{1}{\sqrt{2I_{xx}(t)}}\right]^2 + \left[\frac{H_y(t)}{H(0)}\frac{1}{\sqrt{2I_{yy}(t)}}\right]^2 + \left[\frac{H_z(t)}{H(0)}\frac{1}{\sqrt{2I_{zz}(t)}}\right]^2 = \frac{K(t)}{[H(0)]^2} \tag{5.27}$$

$$\left[\frac{\overline{H}_x}{\sqrt{2I_{xx}(t)}}\right]^2+\left[\frac{\overline{H}_y}{\sqrt{2I_{yy}(t)}}\right]^2+\left[\frac{\overline{H}_z}{\sqrt{2I_{zz}(t)}}\right]^2=\frac{K(t)}{[H(0)]^2} \tag{5.28}$$

Equation (5.26), finally, can be written in its useful final form as follows:

$$\left[\frac{\overline{H}_x(t)}{\frac{\sqrt{2K(t)I_{xx}(t)}}{H(0)}}\right]^2+\left[\frac{\overline{H}_y(t)}{\frac{\sqrt{2K(t)I_{yy}(t)}}{H(0)}}\right]^2+\left[\frac{\overline{H}_z(t)}{\frac{\sqrt{2K(t)I_{zz}(t)}}{H(0)}}\right]^2=1 \tag{5.29}$$

or

$$\left[\frac{\overline{H}_x(t)}{a_x(t)}\right]^2+\left[\frac{\overline{H}_y(t)}{a_y(t)}\right]^2+\left[\frac{\overline{H}_z(t)}{a_z(t)}\right]^2=1 \tag{5.30}$$

Equation (5.30) corresponds to the ellipsoid in the $\overline{H}_x$, $\overline{H}_y$ and $\overline{H}_z$ axis, with the following values of the semi-major axes:

$$\begin{aligned} a_x(t) &= \frac{\sqrt{2\,K(t)\,I_{xx}(t)}}{H(0)} \\ a_y(t) &= \frac{\sqrt{2\,K(t)\,I_{yy}(t)}}{H(0)} \\ a_z(t) &= \frac{\sqrt{2\,K(t)\,I_{zz}(t)}}{H(0)} \end{aligned} \tag{5.31}$$

5.5.6 *Polhodes on the Kinetic Energy Ellipsoids*

Kinetic energy ellipsoids (KEEs) can be used in the study cases *A*, *B* and *C* not only to mark their respective initial conditions but also to show the corresponding resultant motion as lines on the KEEs, called *polhodes*. Polhodes are trajectories of the tips of the non-dimensional angular momentum vectors, $\overline{\mathbf{H}}_A$, $\overline{\mathbf{H}}_B$ and $\overline{\mathbf{H}}_C$ (these three polhodes are marked with the blue colour in Fig. 5.16b, e, h), where we also show for comprehensive visualisation three quiver plots for the $\overline{\mathbf{H}}_A$, $\overline{\mathbf{H}}_B$ and $\overline{\mathbf{H}}_C$ vectors, superimposed over the KEEs and polhodes.

We should emphasise that for plotting polhodes on the KEEs we are using non-dimensional angular momentum coordinates $\overline{H}_x$, $\overline{H}_y$, $\overline{H}_z$, defined as per Eq. (5.24).

Therefore, in Fig. 5.16, in addition to the angular momentum spheres with specific polhodes for the cases *A*, *B* and *C* (Fig. 5.16, left column), we also plotted corresponding kinetic energy ellipsoids (Fig. 5.16, middle column). Then, combining the surfaces in these two columns, we can see that specific polhodes are, in fact, lines of intersection between the corresponding AMSs and KEEs (Fig. 5.16 right column).

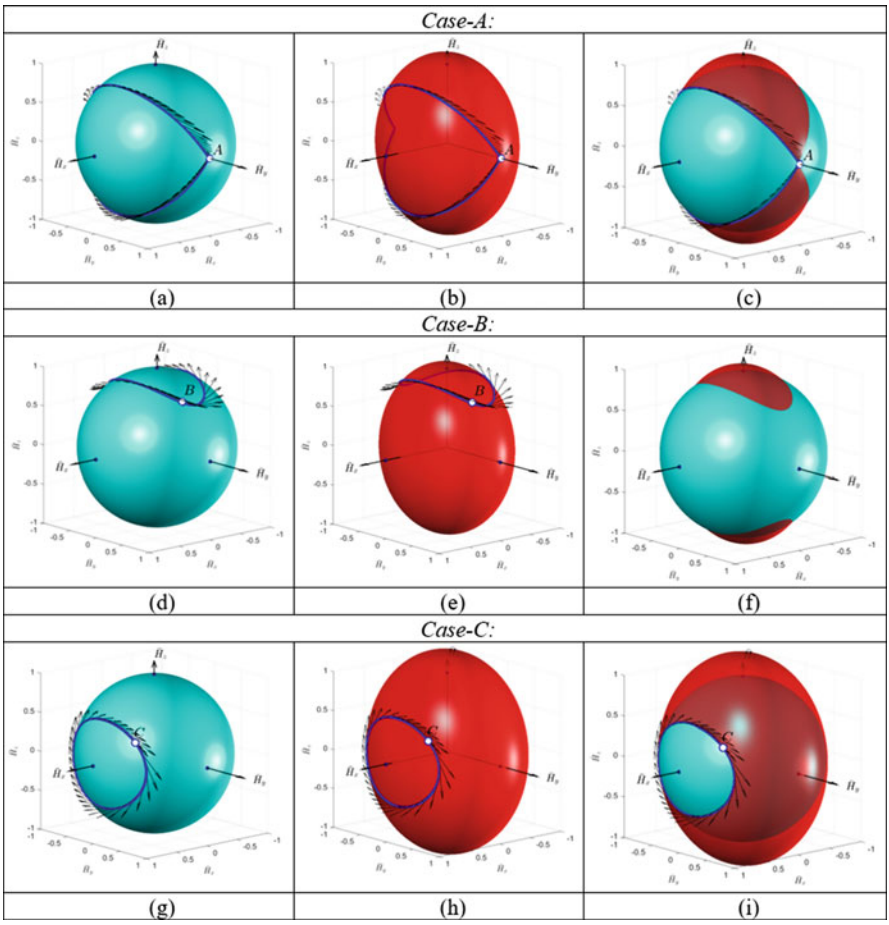


Fig. 5.16 (**a**), (**d**), (**g**) Angular momentum unit spheres (left column); (**b**), (**e**), (**h**) kinetic energy ellipsoids (middle column) for cases *A*, *B*, *C*; (**c**), (**f**), (**i**) superimposed AMSs and KEEs

5.5.7 *Polhodes for Systems with Equal Moments of Inertia*

Let us consider a special case, when two out of three moments of inertia have the same values. For illustration, we simulate the case with $I_{xx} = I_{zz} = 5$, $I_{yy} = 3$ (all in kg $\times$ m^2) and $\omega_{x,i} = 0.2$; $\omega_{y,i} = 1$; $\omega_{z,i} = 0.6$ (all in rad/s). A snapshot from the 3D animation is shown in Fig. 5.17a, time histories of the main spacecraft dynamics parameters are shown in Fig. 5.17b and the co-centred KEE and AMS surfaces are shown in Fig. 5.17c. Figure 5.17a shows that the *y*-body axis is circling, indicating that the motion of the spacecraft is "coning". This is confirmed with the godograph of the *y*-body axis, drawing in 3D a yellow circle, while other body axes, *x* and *z*, are drawing more complex curves (see Fig. 5.17a).

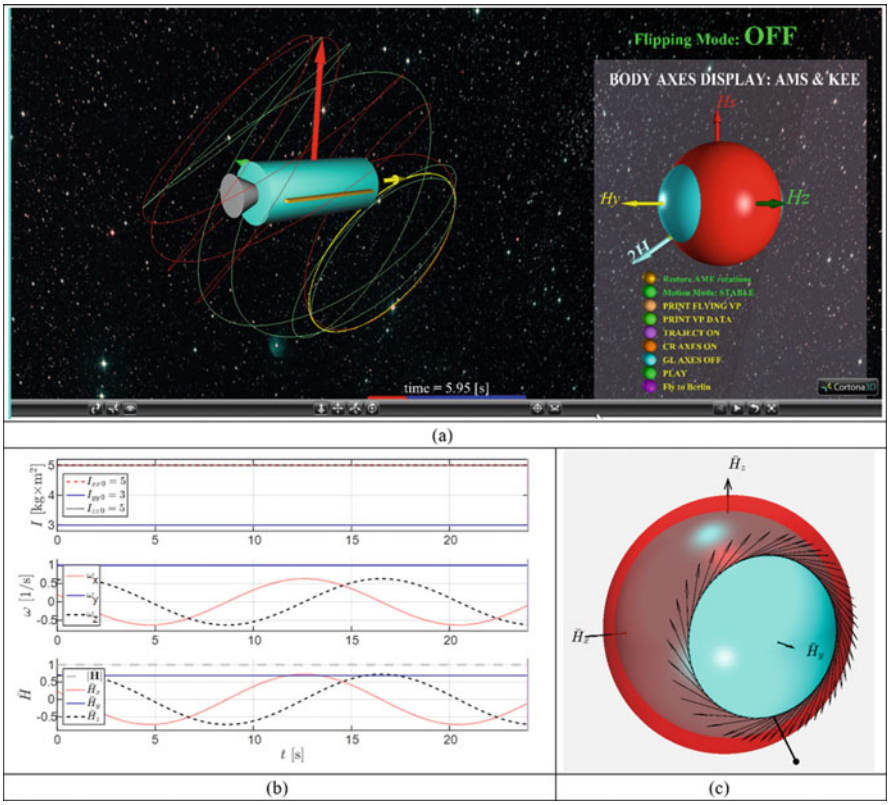


Fig. 5.17 3D rotational motion of the spacecraft with two equal moments of inertia: (**a**) 3D view; (**b**) time histories of the moment of inertia, angular velocity and non-dimensional angular momentum components; (3) collocated AMS and KEE

Figure 5.17c shows that KEE is bulging above the AMS in the xz plane and AMS is intersecting with KEE along the circle, and during the animation process, the tip of the non-dimensional vector $\overline{\mathbf{H}}$ is sliding along this intersection line, called polhode. Velocities of the tip of the vector $\overline{\mathbf{H}}$ are shown with black arrows in Fig. 5.17c. Because vector $\overline{\mathbf{H}}$ has a unit length, it is fully embraced by the AMS and KEE. In order to visualise its current orientation in the body axes, its magnified version with increased length and added small black sphere at the end is shown in Fig. 5.17c. Simulation results show that the values of the semi-major axes of the KEE are equal to $a_x = a_z = 1.1471$ and $a_y = 0.8885$.

For another interesting special case, when all moments of inertia are given with the same number, for example, with $I_{xx} = I_{yy} = I_{zz} = 5$ (all in kg $\times$ m^2) and $\omega_{x,i} = 0.2$; $\omega_{y,i} = 1$; $\omega_{z,i} = 0.6$ (all in rad/s), similar results are presented in Fig. 5.18.

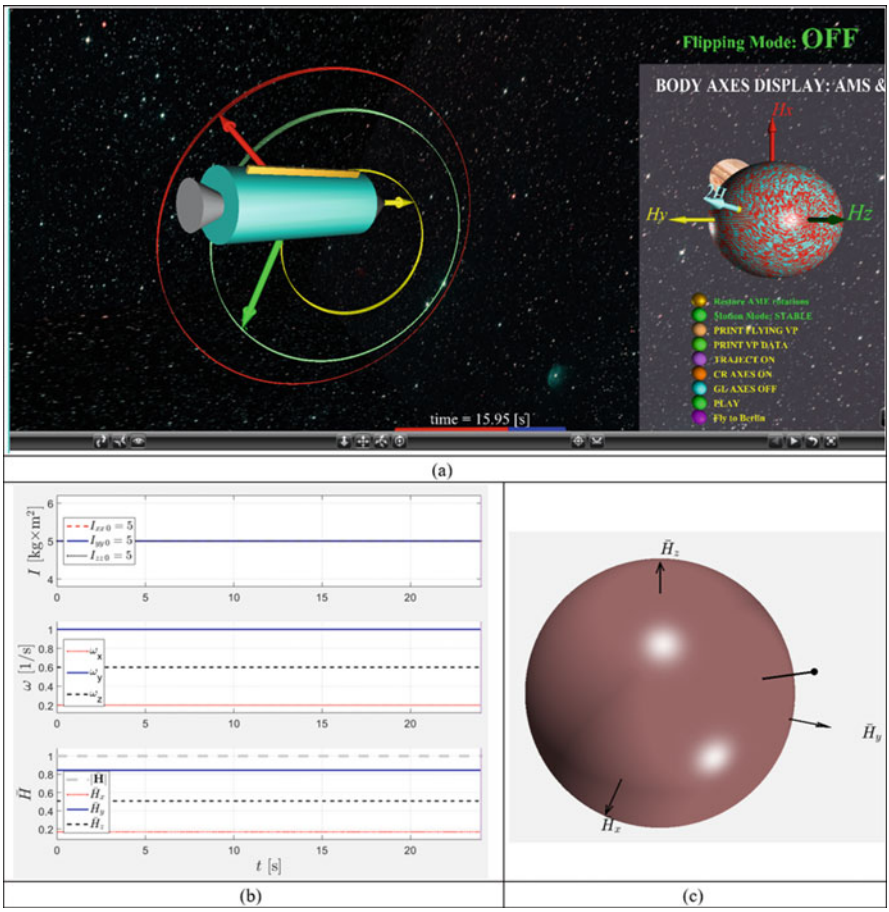


Fig. 5.18 3D rotational motion of the spacecraft with two equal moments of inertia: (**a**) 3D view; (**b**) time histories of the moment of inertia, angular velocity and non-dimensional angular momentum components; (3) collocated AMS and KEE

In this case, both surfaces, AMS and KEE, are given by unit spheres, displayed in Fig. 5.18a on the simulation control panel on the right side with mosaic surface. It is interesting to observe that all body axes, x, y and z, are drawing perfect circles with their tips, as shown in Fig. 5.18a. In this case, components of the angular velocities are not changing their values with time (see Fig. 5.18b), and the angular momentum vector is not changing its orientation in the body axes (see Fig. 5.18).

5.6 Geometric Interpretation of the "Garriott's-Dzhanibekov's Effect", Using Angular Momentum Sphere and Kinetic Energy Ellipsoid

5.6.1 Collocated Angular Momentum Sphere and Kinetic Energy Ellipsoid for the Garriott's-Dzhanibekov's Flipping Motion Example

Figure 5.19 shows collocated $\overline{H}$ sphere (shown in light blue colour) and K ellipsoid (shown in red colour) for the case, considered in Sect. 5.3.4*:* $I_{xx} = 0.3$; $I_{yy} = 0.35$; $I_{zz} = 0.4$ (all in kg $\times$ m^2); $\omega_{x,i} = 0.1$, $\omega_{y,i} = 15$, $\omega_{z0} = 0.1$ (all in rad/s). It is clearly seen from Fig. 5.19a that the value of the z semi-major axis of the K ellipsoid is larger than 1, and therefore, around the "north" and "south" poles, it is extending (extruding/bulging/swelling) outside the $\overline{H}$ unit sphere. The values of the semi-major axes of the KEE can be calculated, using Eq. (5.28), giving the following values: $a_x = 0.9258$; $a_y = 1$; $a_z = 1.0690$. As solution points of the attitude dynamics problem are simultaneously restricted by $\overline{H}$ sphere and K ellipsoid, they must belong to the intersection of $\overline{H}$ and K surfaces.

In order to indicate the direction of motion of the $\overline{H}$ hodograph, in Fig. 5.19c we mark some of the discrete instants for the tip of the $\overline{H}$ vector with their corresponding times. In addition, we are adding the quiver arrow plot, indicating the speed of vector $\overline{H}$ along of the hodograph. Figure 5.19c shows that close to the point, where y-axis intersects with $\overline{H}$ sphere, the density of the points is getting larger, and the length of the quiver arrows is smaller. It indicates that the $\overline{H}$ vector is slowing down, when it is getting more aligned with y-axis, which is an intermediate axis for the system.

The 3D marked plots in Fig. 5.19 conform with the pattern of motion, observed in "Garriott's-Dzhanibekov's effect" experiments, where there is local (in time) stabilisation of the system, where its main rotation is getting closer to the y-axis, and then, the flip is happening rather rapidly (see Fig. 5.19b, c and d).

5.6.2 Conceptual Spacecraft Model, Based on the Flipping Motion

The periodic flipping motion of the system about its intermediate axis prompts design of the spacecraft utilising this peculiar motion. Indeed, if the spacecraft is provided with rotation about its intermediate axis, it would start flipping and would start exposing its instruments to different directions, without any additional involvement of attitude control devices and systems. If, in one particular case, the spacecraft has a camera, scanning instrument or antenna, externally attached to the spacecraft, it could be directed to various directions on its own, due to the flipping

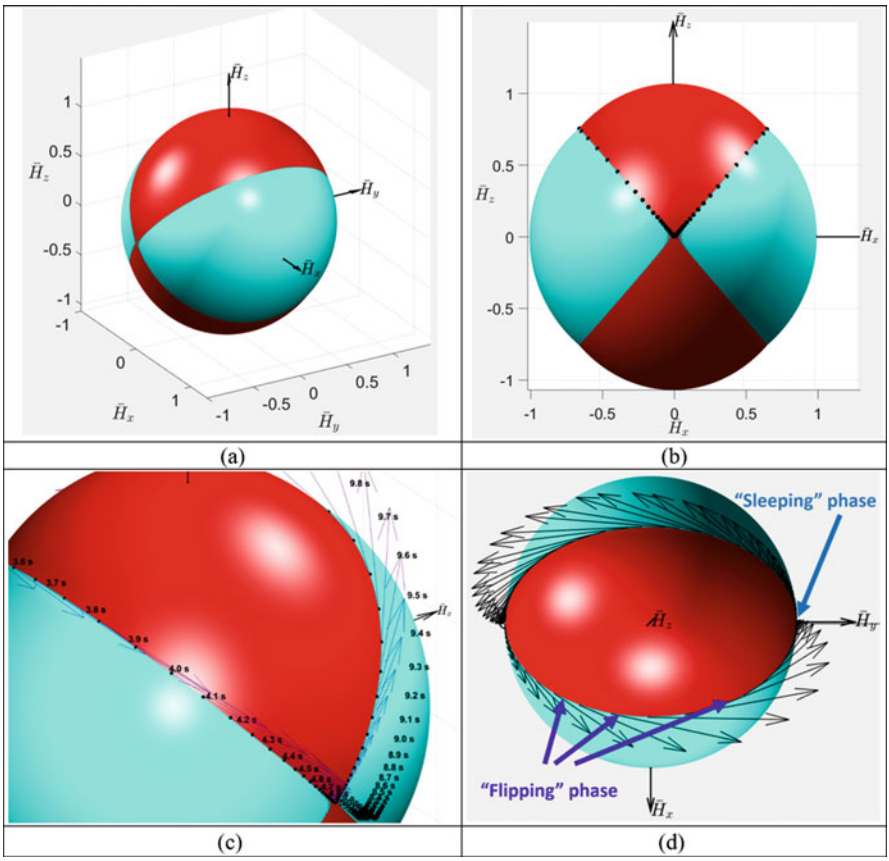


Fig. 5.19 Angular momentum sphere and Kinetic energy ellipsoid for the flipping object (typical "Garriott's-Dzhanibekov's effect" case): (**a**) 3D view of collocated $\overline{H}$ sphere and K ellipsoid; (**b**) 2D "minus" $\overline{H}_y$ view with $\overline{H}$ hodograph marked; (**c**) enlarged $\overline{H}$ hodograph with the selected discrete time points shown; (**d**) "flipping" fast phase and "sleeping" slow phases

motion of the body of the spacecraft. What makes this concept attractive is the following:

- Simplicity of the design and robustness of the flipping motion, which may allow use of the concept for far space autonomous missions, where no real-time interference by the mission specialist would be possible
- Ability to control the period of the flipping motion by proper selection of the principal moments of inertia of the system, which may allow selection of the most efficient agility of the spacecraft

In the presentation of the concept of the flipping spacecraft, we need to make two important comments:

(a) The main disadvantage of the "continuously flipping" spacecraft concept in its basic configuration is its inability to switch flips ON only when they are needed: indeed, it would be impractical for the system to perform flips all the time, including launching and deployment stages. Therefore, it would be imperative to learn on how to switch ON and OFF the flipping motions. The method of initiation of the flipping motion on the spacecraft in stable spin and transferring flips into the stable spin has been discovered in [3] and will be presented in the following section.
(b) Observing the flipping motion of the T-handle [14] or other rigid bodies in the flipping motion demonstrations [23–26], one may get the impression that the instrument, attached to any side of the flipping body, would scan the spacecraft surrounding, pointing in all possible directions. However, our study [8] showed that this impression about full scanning coverage of the instruments on the flipping spacecraft is wrong, and it also showed that attachment of the instrument to different sides of the flipping spacecraft may result in restricted hemispheres for the intended coverage. This aspect is analysed in the next subsection.

5.6.3 *Investigating Orientation of the Sides of the Spacecraft Exposed to the Specific Directions*

As spacecraft may have directional sensing equipment, attached to the sides, let us explore possible exposure of this equipment to the specified directions of interest. For this purpose, let us introduce a semi-transparent *green coloured spherical dome*, embracing the rotating spacecraft (which, in turn, has its rotating body axes system xyz with unit orts $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$). We collocate the centre of the dome (point O) with the centre of the mass of the rotating body. However, most significant, we fix the dome in the global coordinates XYZ, so it is not rotating with the body and its body axes xyz. Then we consecutively plot lines of intersection of the rotating orts $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ with the dome. It must be emphasised that the spheres in Fig. 5.20 are not the bodies of the spacecraft (which may have any arbitrary shape), but the embracing imagined domes.

For the illustration purposes, let us simulate the motion of the spacecraft with the following parameters: $I_{xx} = 2$, $I_{yy} = 4$, $I_{zz} = 3$ (all in kg $\times$ m^2). Let us for $t = 0$ align xyz body axes with XYZ global inertial axes as follows: x is aligned with X, y is aligned with Z and z is aligned with $-Y$. If the spacecraft is installed in orbit with initially provided angular velocity $\omega_{x,i} = 0.01$, $\omega_{y,i} = 0.01$, $\omega_{z,i} = 1$ (all in rad/s), the spacecraft starts "flipping" along axis z, *being an intermediate axis of inertia* (as $I_{xx} < I_{zz} < I_{yy}$).

During this flipping process, we trace all intersections of the orts $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ with the dome and present them as continuous lines with different colours. Results are shown in Fig. 5.20. It should be noted that for each of the computer screen

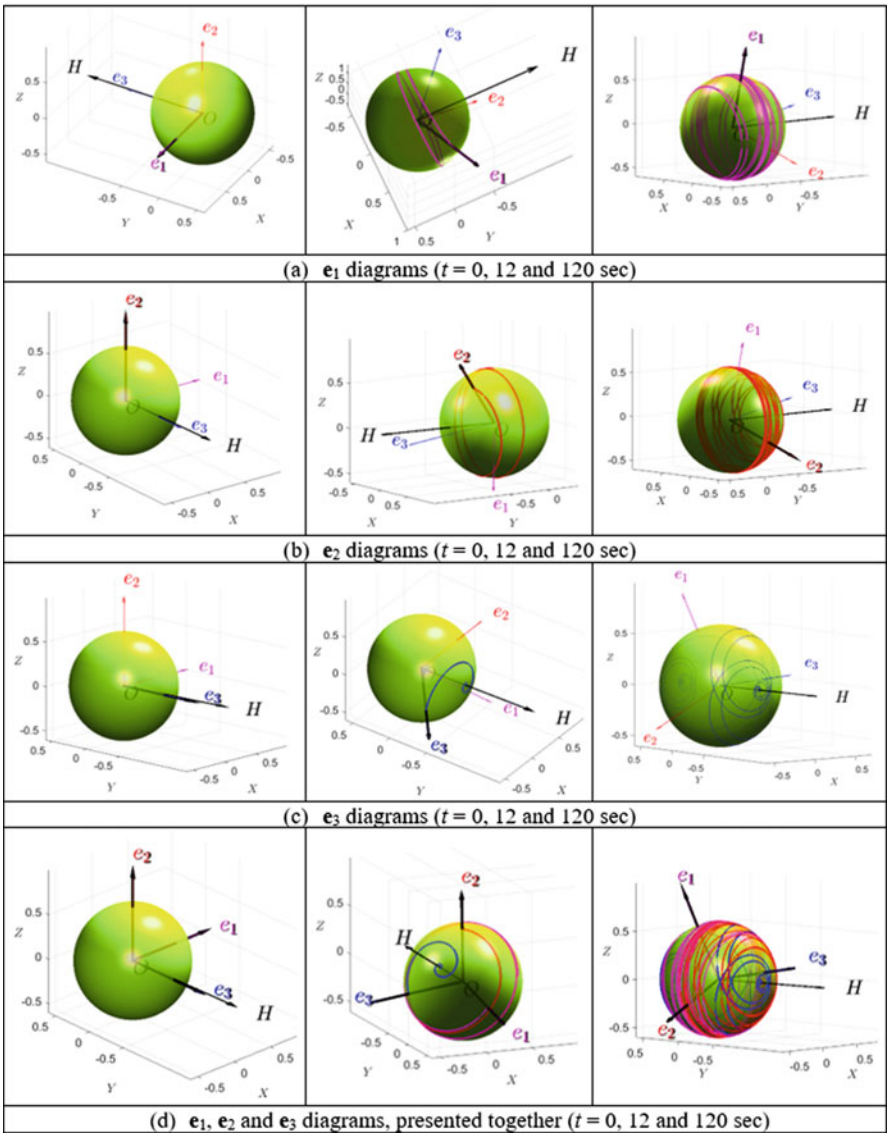


Fig. 5.20 Lines of intersection of the rotating orts $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ with the spherical dome (green), fixed in the global axes system *XYZ*: "ball of wool" lines. (**a**) $\mathbf{e}_1$ diagrams ($t = 0$, 12 and 120 s), (**b**) $\mathbf{e}_2$ diagrams ($t = 0$, 12 and 120 s), (**c**) $\mathbf{e}_3$ diagrams ($t = 0$, 12 and 120 s) and (**d**) $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ diagrams, presented together ($t = 0$, 12 and 120 s)

snapshots in this figure, the individual viewpoint was selected for better observation of the simulation results. Selection of the viewpoints could be clearly understood using the vector of the angular momentum $\boldsymbol{H}$ as a reference, as it is pointing in the same direction in the global coordinates XYZ for all presented snapshots.

Last images for the $\mathbf{e}_2$ in Fig. 5.20b are remarkably interesting and illustrate our new finding! They show that y-body rotating axis, associated in this example with the *maximum* moment of inertia, is "drawing" $\mathbf{e}_2$ intersection lines on the dome only on one hemisphere, bulging towards the angular momentum vector $\boldsymbol{H}$ (we call it H_+ hemisphere), and is never pointing towards the other hemisphere of the dome (shown as H_- hemisphere in Fig. 5.21). This is valid for the direction of y with positive component of the angular velocity along this direction ($\omega_{y,i} > 0$). We have run many other various simulations, confirming that it is a general pattern, so the side, perpendicular to the axis with maximum moment of inertia and associated with positive angular velocity component, is never turned away from the vector $\boldsymbol{H}$ direction.

In Fig. 5.20, initially, vector $\boldsymbol{H}$ is almost aligned with z body axis (which is, in turn, is initially positioned along the $-Y$ global axis); this is because initial values of $\omega_{x,i}$ and $\omega_{z,i}$ (and ultimately $H_{x,i}$ and $H_{z,i}$) are small compared with $\omega_{y,i}$ (and ultimately $H_{y,i}$). Therefore, the 2D plane surface, subdividing H_+ and H_-, is almost parallel to the XZ plane. Discovery of the H_+ and H_- planes is also illustrated on the model of the spacecraft (see Fig. 5.21b), flipping about x body axis and "drawing" godograph on the surrounding dome with its "y" body axis. Figure 5.21b also shows on the right the simulation Virtual Reality control panel, displaying AMS and KEE of the system with enlarged non-dimensional angular momentum vector, sliding along the separatrix.

H_+ and H_- are also shown in Fig. 5.22, where we consider additional contrast case with the following parameters: $I_{xx} = 2$, $I_{yy} = 4$, $I_{zz} = 3$ (all in kg $\times$ m^2) and initial angular velocities $\omega_{x,i} = 0.5$, $\omega_{y,i} = 0.5$, $\omega_{z,i} = 1$ (all in rad/s), which has much more significant initial values of $\omega_{x,i}$ and $\omega_{y,i}$, than in the previous example, hence has large components of $H_{x,i}$ and $H_{z,i}$, as compared with $H_{y,i}$. It results in the subdivision of the dome into two parts (H_+ and H_-) by the inclined 2D plane, shown in white in Fig. 5.22a. Results of the intersection lines of the $\mathbf{e}_2$ ort with the dome are shown in Fig. 5.22a. They somehow resemble "ball of wool" (see Fig. 5.22b), especially with the knitting needles resembling the $\boldsymbol{H}$ and $\mathbf{e}_2$ vectors. However, the simulated resulting "ball of wool" lines are "sitting" on one hemisphere only! This hemisphere is on the side of the plane, perpendicular to $\boldsymbol{H}$ vector (and we will call it H_+ hemisphere). The other side of the hemisphere (H_-) does not have any threads of the "ball of wool".

This discovered new result can be used in the design of various spacecraft missions. For example, in case of the communication mission, if the spacecraft is installed in orbit with predominant rotation about an intermediate axis of inertia and is carrying an antenna, it should be ensured that the initial direction of the angular momentum vector $\boldsymbol{H}$ is consistent with the "source", sensed by antenna,

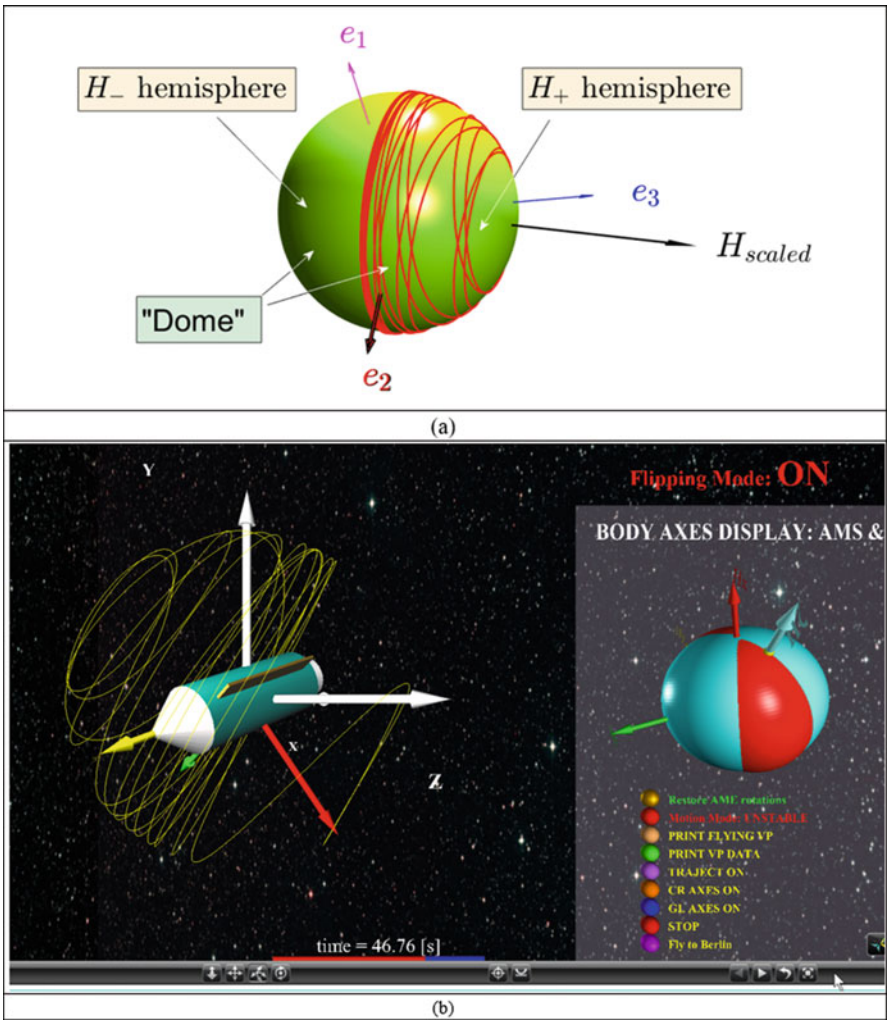


Fig. 5.21 (**a**) H_+ and H_- hemispheres of the "dome" ($I_{xx} = 2$, $I_{yy} = 4$, $I_{zz} = 3$, all in kg $\times$ m^2; $\omega_x = 0.01$, $\omega_y = 0.01$, $\omega_z = 1$, all in rad/s); (**b**) H_+ and H_- hemispheres of the "dome" shown on the flipping acrobatic spacecraft (on the left) concurrently with the AMS/KEE simulation Virtual Reality control panel (on the right)

i.e. with H_+ hemisphere facing the "source"; otherwise, spacecraft communication would be blanked for all instants of the mission. So, it matters which side of the spacecraft, perpendicular to the axis with maximum moment of inertia, is selected: one side would be good for utilising antenna, and the other side would be inoperable/terminal. The exposure "efficiency" of the equipment on the selected sides was explored in reference [34].

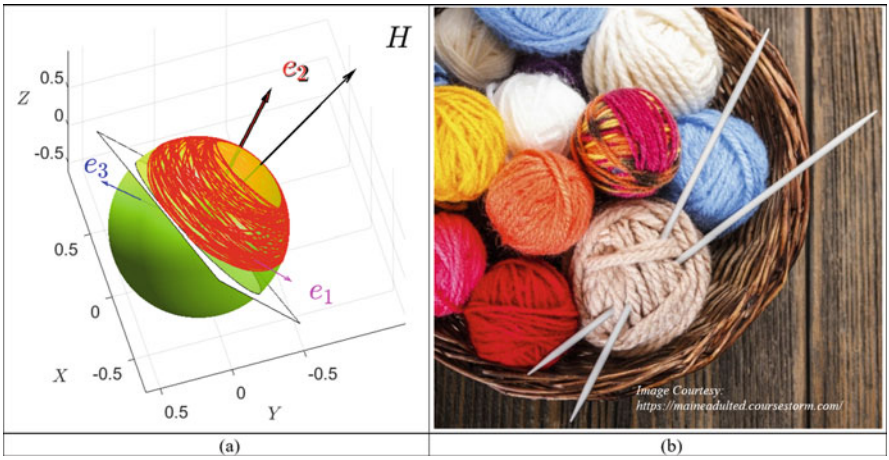


Fig. 5.22 "Ball of wool" lines: (**a**) simulation results for the case $I_{xx} = 2$, $I_{yy} = 4$, $I_{zz} = 3$ (all in kg $\times$ m^2) and initial angular velocities $\omega_{x,i} = 0.5$, $\omega_{y,i} = 0.5$, $\omega_{z,i} = 1$ (all in rad/s); (**b**) original "balls of wool", which prompted the used analogy and terminology

On the same token, in some other cases, when, for example, the spacecraft is subject to directional adhere conditions (heat, radiation, flying debris), it may be advisable to "reinforce" the spacecraft, facing the intended H_- hemisphere, install the spacecraft in orbit with the direction of the initial angular momentum pointing outwards the "danger" and place all sensitive equipment on the side, perpendicular to the axis with maximum moment of inertia and with positive component of the angular velocity along this direction (i.e. "plus" $\mathbf{e}_2$ in the two previously considered illustration cases).

As a summary from this subsection, we present in Fig. 5.23 godographs of all body axes orts $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, for the base case study in Subsect. 5.3.4 ($I_{xx} = 0.3$, $I_{yy} = 0.35$, $I_{zz} = 0.4$ (all in kg $\times$ m^2), corresponding to $\xi = 0.5$ and $\eta = 0.75$, with the initial conditions $\omega_{x0} = 0.1$, $\omega_{y0} = 15$, $\omega_{z0} = 0.1$ (all in rad/s)).

It shows that if the system performs classical "Garriott's-Dzhanibekov's" flipping motion, the godograph of the port with maximum moment of inertia is drawing on the half of the dome, surrounding a flipping object, therefore on this object, where will be a side, which will be only exposed to one half of the semi-space, pointed by the angular momentum vector $\boldsymbol{H}$, and will never be exposed to the second semi-space.

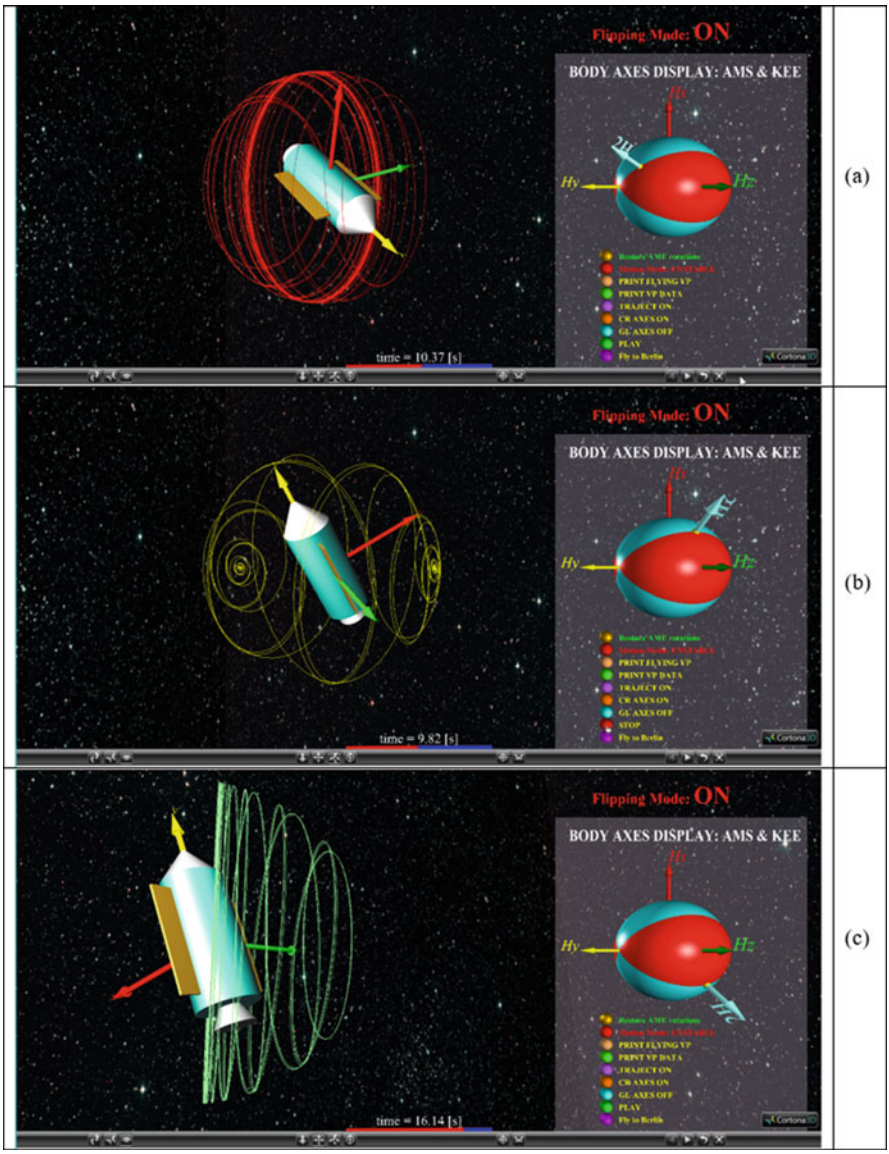


Fig. 5.23 "Ball of wool" lines for body axes orts $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ in base study case in Subsect. 5.3.4: (**a**) red ort $\mathbf{e}_1$ along axis with minimum moment of inertia; (**b**) yellow ort $\mathbf{e}_2$ along intermediate axis; (**c**) green ort $\mathbf{e}_3$ along axis with maximum moment of inertia

5.7 Proposing New Spacecraft Designs/Missions, Utilising Garriott's-Dzhanibekov's Effect and Inertial Morphing

5.7.1 Proposing Method of "Switching ON/OFF" Garriott's-Dzhanibekov's Spacecraft Flipping Motion by Controlled Inertial Morphing

Flipping motion of the rigid body, during which the direction of the angular velocity of the main rotation, let's say, ω_y, is intermittently changing to opposite, is called "Garriott's-Dzhanibekov's effect". It is a consequence of the moment of inertia, associated with the main rotation, being between two other values of the moments of inertia, I_{xx} and I_{zz}, in other words, having an intermediate value among principal moments of inertia.

What if there is a need to stop or suspend for some time the unstable "flipping" motion of the object?

For this purpose, Beachley [20] proposed four types of mass translations. However, this proposition was rather conceptual, as it did not involve equations of motion of the spacecraft with variable inertial properties, did not explain the change in moments of inertia and did not investigate the inversion timing.

In [3], and then in [4–10], we have addressed all these issues in a systematic manner and for the purpose of control of the spacecraft attitude, proposed general method of inertial morphing of the spacecraft, mathematically linking these modifications to the changes in moments of inertia and simulating transitional spacecraft response, using rigorous equations of motion [3]. In particular, we showed that for stabilisation of the spacecraft, there exist two morphing strategies, and after implementation of which, the intermediate moment of inertia becomes the smallest *or* largest among all principal moments of inertia.

In the illustration case, where we selected "y" axis to be the axis of the main rotation, the condition for the unstable "Garriott's-Dzhanibekov's effect"-type motion can be written as:

$$I_{xx} \quad < \quad I_{yy} \quad < \quad I_{zz} \tag{5.32}$$

However, if via special design of the spacecraft, enabling the change of its principal moments of inertia (via mechanical or other means), the targeted value of I_{yy} is in controllable way forcefully "moved" outside the embrace of I_{xx} and I_{zz}, then the condition of instability Eq. (5.32) would no longer be satisfied and the unstable motion would be "switched OFF"!

Conceptually, this proposition can be illustrated with the diagram in Fig. 5.24, which presents *two* sets of solutions. The first conceptual solution set involves reduction of initial value of I_{yy} (which we denote as $I_{yy,i}$) to its new (or final) value $I_{yy,f}$, being smaller than I_{xx} value. And the second solution involves increase of the

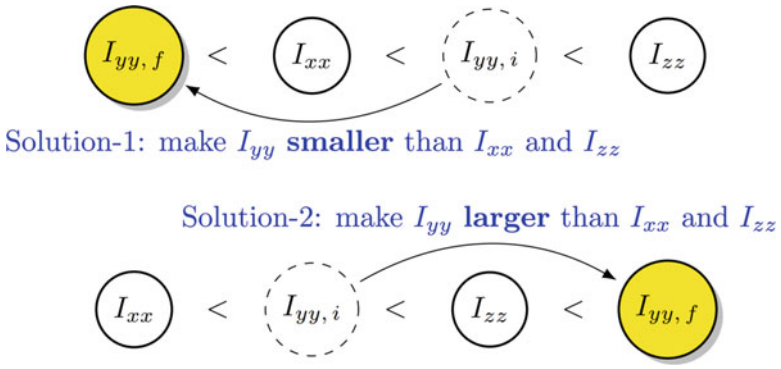


Fig. 5.24 Possible conceptual solutions for stabilizing an unstable spacecraft with its main rotation about the *y*-axis

initial value of I_{yy} (which we denote as $I_{yy,i}$) to its new value $I_{yy,f}$, being larger than I_{zz}.

For the numerical verification of the concept, let us assume the following demonstration values: $I_{xx} = 0.3$, $I_{yy} = 0.35$ and $I_{zz} = 0.4$ (all in kg $\times$ m^2), which are conforming with the general condition in Eq. (5.32) of the flipping unstable motion, which would result if the main rotation about *y*-axis is initiated. And in this case, in order to test the concept of "switching OFF" the flipping motion, we will change $I_{yy,i} = 0.35$ to its new value of $I_{yy,f} = 0.2$ (solution-1) or $I_{yy,f} = 0.5$ (solution-2).

However, in order to proceed with the numerical simulations, we need to expand Euler Eq. (5.3), allowing variations in the moments of inertia of the rigid body.

5.7.2 *Extending Euler's Equations for Rigid-Body Rotations, Allowing Variation of Moments of Inertia*

In order to simulate the cases of the morphing spacecraft with variable moments of inertia, we need to extend classic Euler's Eq. (5.3). We note that the sum of the moments about the centre of mass of a rigid body due to external forces and couples equals to the rate of change of the angular momentum about the centre of mass [35]:

$$\sum \vec{\mathbf{N}} = \left.\frac{d\vec{\mathbf{H}}}{dt}\right|_{Inertial} = \left.\frac{d\vec{\mathbf{H}}}{dt}\right|_{Body} + \vec{\boldsymbol{\omega}} \times \vec{\mathbf{H}} \tag{5.33}$$

Also, the components of the angular momentum vector, $\vec{\mathbf{H}}$, with respect to the body-axis frame can be expressed by the product between the principal moment of inertia matrix I_G and the components of the angular velocity vectors as follows:

$$\vec{\mathbf{H}}(t) = \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \tag{5.34}$$

Therefore, extended Euler's equations can now be written as:

$$\begin{bmatrix} \dot{I}_{xx} & 0 & 0 \\ 0 & \dot{I}_{yy} & 0 \\ 0 & 0 & \dot{I}_{zz} \end{bmatrix} \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} + \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \end{Bmatrix}$$
$$+ \begin{bmatrix} 0 & -\omega_z & -\omega_y \\ -\omega_z & 0 & -\omega_x \\ -\omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix} \tag{5.35}$$

For solving the *morphing* rigid-body dynamics problems, using numerical methods, we combine matrix Eqs. (5.35) and (5.6) into a single equation:

$$\begin{bmatrix} I_{xx} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{yy} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{zz} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sin\theta\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & \sin\theta\ \cos\phi & -\sin\phi & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & 1 \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \\ \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{Bmatrix} = \begin{Bmatrix} (I_{yy} - I_{zz})\ \omega_y\omega_z - \dot{I}_{xx}\ \omega_x \\ (I_{zz} - I_{xx})\ \omega_z\omega_x - \dot{I}_{yy}\ \omega_y \\ (I_{xx} - I_{yy})\ \omega_x\omega_y - \dot{I}_{zz}\ \omega_z \\ \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \tag{5.36}$$

Equation (5.33) is the main equation, used in this work and solved using MATLAB® `ode` MATLAB Runge-Kutta solver, with "mass matrix" option, as per Eq. (5.9). This equation is applicable to the systems which moments of inertia are changing with time.

5.7.3 Six-Mass Conceptual Model of the Spacecraft with Inertial Morphing Capabilities

To demonstrate the feasibility of the controllable behaviour of the spacecraft, let us consider a simple conceptual model of the morphing spacecraft, constructed as an axisymmetric set of three orthogonal dumbbells, each of which has negligible mass of the rod, connecting two equal concentrated masses at its ends. Let us also assume, for conceptual simplicity, that three dumbbells are connected at the middle points of their rods, and the corresponding masses m_x, m_y and m_z are located at the distances r_x, r_y and r_z from the axes of rotation x, y and z, as shown in Fig. 5.25.

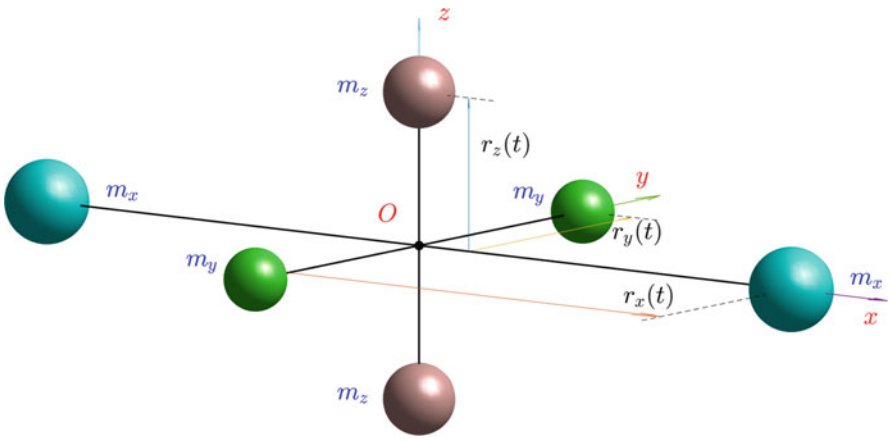


Fig. 5.25 Six-mass conceptual model of the morphing spacecraft

In the illustrated conceptual design, morphing of the spacecraft is achieved via independent synchronized control of the position coordinates $r_x = r_x(t)$, $r_y = r_y(t)$ and $r_z = r_z(t)$ of the masses m_x, m_y and m_z.

The principal moments of inertia of the system can be calculated as follows:

$$\begin{aligned} I_{xx} &= 2m_y r_y^2 + 2m_z r_z^2 \\ I_{yy} &= 2m_z r_z^2 + 2m_x r_x^2 \\ I_{zz} &= 2m_x r_x^2 + 2m_y r_y^2 \end{aligned} \tag{5.37}$$

Then by adding all equations in (5.34), and taking only a half of the left and right hand sides, we can get:

$$\frac{1}{2}\left(I_{xx} + I_{yy} + I_{zz}\right) = 2\left(m_x r_x^2 + m_y r_y^2 + m_z r_z^2\right) \tag{5.38}$$

Then, subtracting from Eq. (5.35) consecutively each of Eq. (5.34), we can obtain:

$$\begin{aligned} r_x(t) &= \sqrt{\frac{I_{yy}(t) + I_{zz}(t) - I_{xx}(t)}{4\,m_x}} \\ r_y(t) &= \sqrt{\frac{I_{zz}(t) + I_{xx}(t) - I_{yy}(t)}{4\,m_y}} \\ r_z(t) &= \sqrt{\frac{I_{xx}(t) + I_{yy}(t) - I_{zz}(t)}{4\,m_z}} \end{aligned} \tag{5.39}$$

Equations (5.39) are very important equations, as they mathematically represent a basis for the concept of "inertial morphing", as formulated for the six-mass model of the morphed spacecraft. They show that at any instant of time t, the exact set of

positions for the masses m_x, m_y and m_z in the model can be uniquely determined, based on the values of the required moments of inertia of the system for this instant of time. Equations (5.39) assume that masses m_x, m_y and m_z have constant values, but they are not imposing these as constraints. In fact, it may be possible in some IM designs to employ masses, which would have variable values, and in these cases, notations m_x, m_y and m_z in Eq. (5.39) should be replaced with $m_x(t)$, $m_y(t)$ and $m_z(t)$. There rather exquisite designs may involve, for example, solidification, evaporation or ablation of the mass materials.

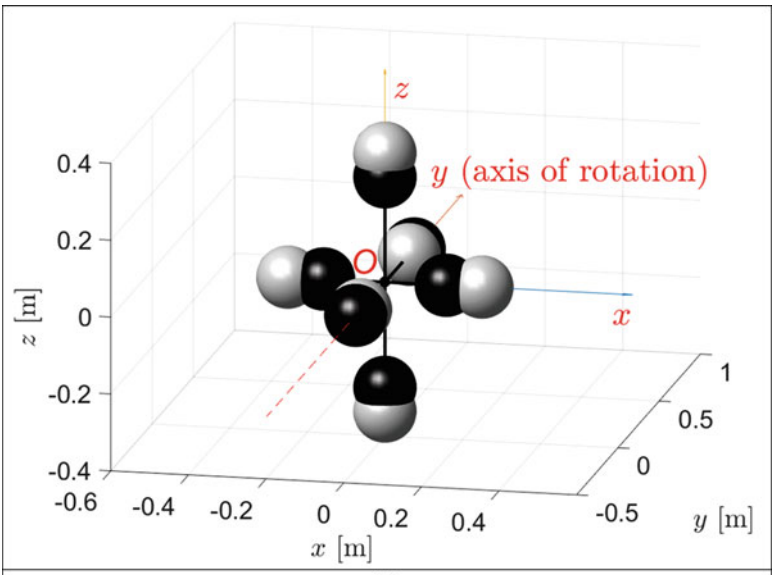
5.7.4 *Conceptual Example of the Morphed Spacecraft, Self-Transferring from Unstable Flipping Motion to Stable No-Flips Spin*

Let us assume, as an example, that the system's parameters are given with the following numbers: $m_x = m_y = m_z = 1$ kg, $I_{xx} = 0.3$ kg $\times$ m^2, $I_{yy} = 0.35$ 0.3 kg $\times$ m^2, $I_{zz} = 0.40$ kg $\times$ m^2, the same as in Sect. 5.3.4 example. Then, for the case of the tumbling spacecraft considered in Sect. 5.3.4, we can find the initial radial positions of the spacecraft masses, using Eq. (5.39):

$$r_{x0} = 0.2500 \text{ m}; \qquad r_{y0} = 0.2958 \text{ m}; \qquad r_{z0} = 0.3354 \text{ m}. \tag{5.40}$$

These values for the unit masses would ensure that the inertial properties of the spacecraft are $I_{xx} = 0.3$ kg $\times$ m^2, $I_{yy} = 0.35$ kg $\times$ m^2, $I_{zz} = 0.4$ kg $\times$ m^2 (this can be proven via Eqs. 5.34). Note that in our example here I_{yy} has an *intermediate* value among all principal moments of inertia: $I_{xx} < I_{yy} < I_{zz}$; therefore, if the spacecraft is provided with the initial angular velocity $\omega_{x0} = \omega_{z0} = 0.1$ rad/s and $\omega_{y0} = 15$ rad/s, with the prevailing rotation about y-body axis, then the spacecraft rotation about this axis would be unstable and the classical "Garriott's-Dzhanibekov's effect" periodic flipping would be observed.

It will be shown in Sect. 9.1 that, if during the "flipping" motion, at the instant, when the angular velocities ω_{x0} and ω_{z0} are close to zeros, the moment of inertia $I_{yy,i} = 0.35$ kg $\times$ m^2 is rapidly changed to its new value, less than I_{xx}, for example, $I_{yy,f} = 0.2$ kg $\times$ m^2, then the nature of the followed motion of the system would change from unstable "flipping" to stable. This would happen, because with the deliberate assignment of the described new values of the moments of inertia, predominant rotation of the system would not be longer along the intermediate axis but would be instantly changed in favour of the rotation about the same body axis but now having its new status of the axis with minimum moment of inertia, and the resulting rotation would become stable, without flips. We call it "switching OFF" the flipping motion manoeuvre. Note: this would require a single instant inertial morphing.



(a)

	r_x	r_y	r_z	I_{xx}	I_{yy}	I_{zz}
Initial	0.2500	0.2958	0.3354	0.30	0.35	0.40
Final	0.1581	0.3536	0.2739	0.30	0.20	0.40

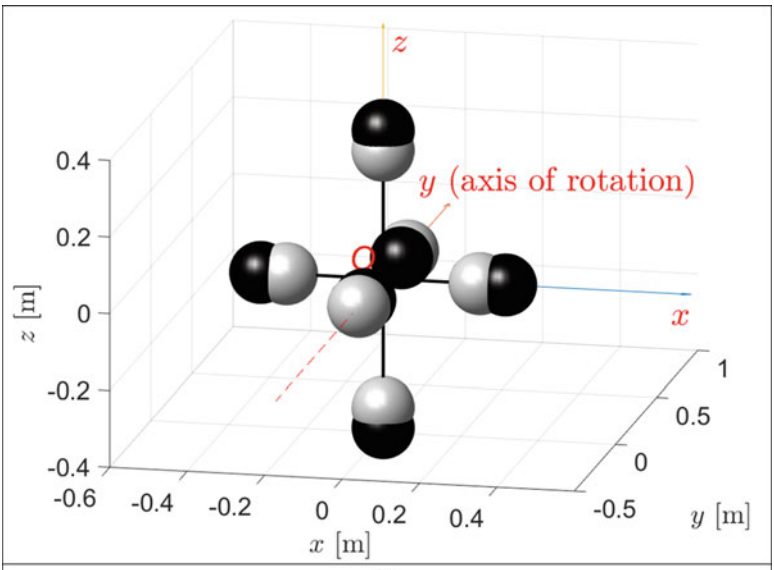
Fig. 5.26 Solution-1 for the conceptual six-mass design of the morphing spacecraft, self-transferring from unstable "Garriott's-Dzhanibekov's effect"-type flipping motion to stable spinning no-flips motion (and vice versa): (**a**) graphical representation, where white spheres correspond to unstable flipping about y – axis of predominant initial rotation and black spheres to stable no-flips spin; (**b**) table representation with radii for all masses specified for initial and final stages of the inertial morphing

The new values of the position radii, corresponding to this particular solution (being one out many solutions in the "solution-1" set), can be calculated using Eq. (5.39) and are shown in the graphical form in Fig. 5.26a and in the table form in Fig. 5.26b:

$$r_{xf} = 0.1581 \text{ m}; \quad r_{yf} = 0.3536 \text{ m}; \quad r_{zf} = 0.2739 \text{ m}. \tag{5.41}$$

The spacecraft masses at these final radius positions are shown in Fig. 5.26a with dark colour.

The flipping motion can be also stopped, using a solution from the "solution-2" set. One of the examples is shown in conceptual Fig. 5.24. For the purpose of the illustration of the concept, let us consider rapid increase of the I_{yy} from its initial value of $I_{yy,i} = 0.35 \text{ kg} \times \text{m}^2$ to its new (final) value of $I_{yy,f}$, being larger than I_{zz}, for



(a)

	r_x	r_y	r_z	I_{xx}	I_{yy}	I_{zz}
Initial	0.2500	0.2958	0.3354	0.30	0.35	0.40
Final	0.3162	0.2236	0.3873	0.30	0.50	0.40

(b)

Fig. 5.27 Solution-2 for the conceptual six-mass design of the morphing spacecraft, self-transferring from unstable "Garriott's-Dzhanibekov's effect"-type flipping motion to stable spinning no-flips motion (and vice versa): (**a**) graphical representation, where white spheres correspond to unstable flipping about y – axis of predominant initial rotation and black spheres to stable no-flips spin; (**b**) table representation with radii for all masses specified for initial and final stages of the inertial morphing

example, $I_{yy,f} = 0.5$ kg $\times$ m^2. The new values of the position radii, corresponding to the "solution-2" in Fig. 5.27b, can be calculated using Eqs. (5.39):

$$r_{xf} = 0.3162 \text{ m}; \quad r_{yf} = 0.2236 \text{ m}; \quad r_{zf} = 0.3873 \text{ m}. \tag{5.42}$$

The spacecraft masses at these radius positions are shown in Fig. 5.27b with dark colour.

The morphing of the spacecraft from the initially unstable configuration [as per Eq. (5.40)], associated with the "flipping" motion, to its final stable configuration [as per Eqs. (5.41) or (5.42) and solution-1 or 2 in Fig. 5.24] is shown in Figs. 5.26a and 5.27a, where masses for the initial unstable (flips) configuration are shown in white, whereas the masses for the final (no-flips) stable configuration are shown in black colour.

Summary for both particular solutions is presented in Figs. 5.26b and 5.27b. It would be important to note that, in the presented cases, it was not obligatory during the morphing of the system and its transition from the "initial" to "final" states to keep both values of I_{xx} and I_{zz} unchanged. However, it was done for purpose to emphasize the role of the I_{yy} in the process of stabilisation of the system.

5.7.5 Geometric Interpretation of the Cases, Where "Garriott's-Dzhanibekov's Effect" Is Controlled

5.7.5.1 Stopping Flipping Motion, Using One Inertial Morphing: Solution-1

At last, in this section, we consider the case of "switching OFF" the flipping motion and fully stabilising the system.

As proposed and illustrated in [3] on the six-mass dumbbell model, full stabilisation of the spinning rigid body (e.g., spacecraft or rocket) can be achieved with a controllable change of all moments of inertia. However, in order to emphasise the importance of the I_{yy} in stabilisation of the system, I_{xx} and I_{zz} were not involved. Application of the applied changes in I_{yy} (shown in Fig. 5.28a) ensures that the periodic change in ω_y is stopped, as illustrated with Fig. 5.28b.

Figure 5.29a shows that while the angular momentum is conserved in the system, the kinetic energy may be a subject to variations. In the presented example, *increase in* the dominant angular velocity component ω_y triggers significant increase in the kinetic energy by 74%, as demonstrated in Fig. 5.29b.

Figure 5.28d of the associated controlled action shows that the kinetic energy ellipsoid has dramatically swollen and is now fully embracing and hiding (inside of the KEE) the angular momentum sphere, which is in great contrast with solution-2, to be illustrated in Fig. 5.30. Nevertheless, both surfaces are touching each other at two points: pole S and another pole on the opposite side of the y-axis.

Conclusion from this subsection is as follows: stabilisation of the flipping motion (i.e. switching OFF the "Garriott's-Dzhanibekov's effect") manipulating with I_{yy} only may be achieved using two avenues – making I_{yy} the smallest moment of inertia or making it largest moment of inertia [3]. The first avenue has been illustrated in Figs. 5.28 and 5.29.

5.7.5.2 Stopping Flipping Motion, Using One Inertial Morphing: Solution-2

One of the two types of solutions is presented in Fig. 5.30.

Solution-2 has been achieved at the "expense" of the kinetic energy E, which has reduced its initial value after the morphing by 30% (see Fig. 5.31a). The geometric interpretation of this controlled action in Fig. 5.30 shows that the kinetic

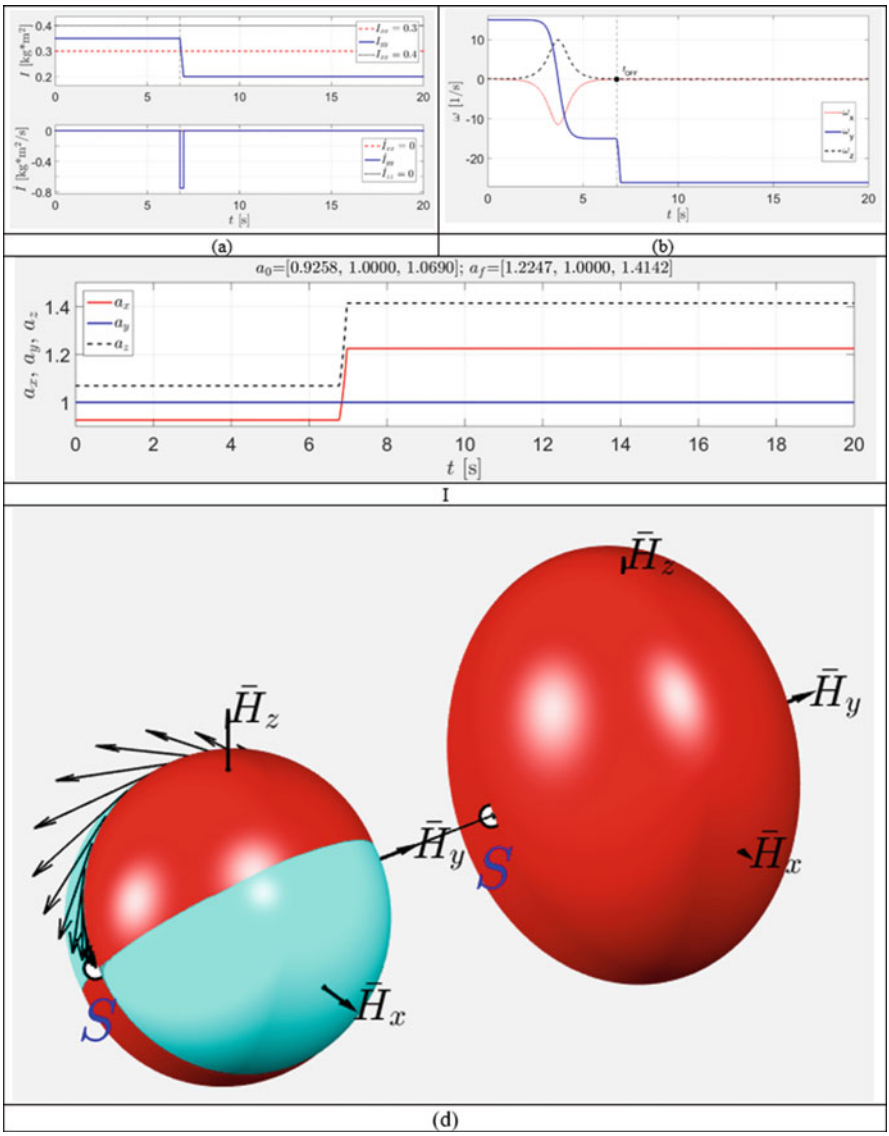


Fig. 5.28 Stopping the flipping motion (solution-1): (**a**) explanation of the inertial morphing applied; (**b**) time history of the resulting angular velocity components ω_x, ω_y, ω_z; (**c**) radii of the KEE; (**d**) geometric interpretation, showing the AMS and KEE before and after application of the inertial morphing

energy ellipsoid has dramatically shrunk and is now fully embraced by the angular momentum sphere, which is in great contrast with solution-1, illustrated in Fig. 5.28. In solution-2, the final KEE is not seen, as is entirely residing inside the shrunk

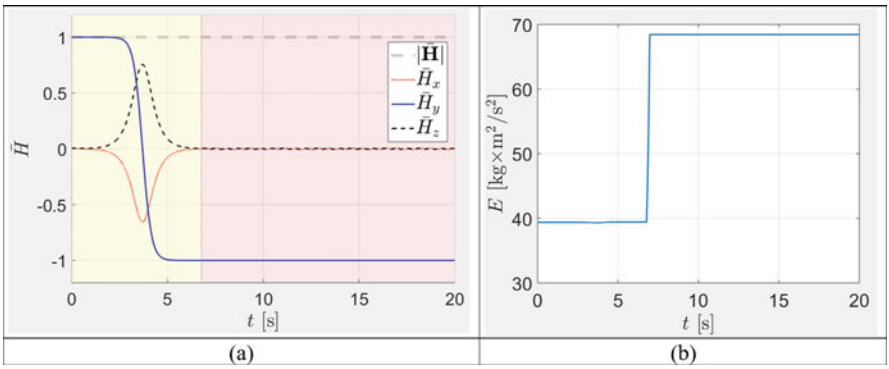


Fig. 5.29 Time histories of the key characteristics of the inertially morphed system: (**a**) non-dimensional angular momentum $\overline{H}$, (**b**) kinetic energy E

AMS. Nevertheless, both surfaces AMS and KEE are touching each other at just two points: pole S and the pole on the opposite side along the y-axis.

5.8 Attitude Dynamics of Spacecraft with Inertial Morphing

5.8.1 Study Case-2: "Switching OFF" Flipping Motion of the Spacecraft After One Flip (Solution-1)

Figure 5.32 shows that at the instant $t = 6.77$ s, the angular velocity ω_y has its highest value and ω_x changes its value from negative to positive. It is believed that this instant, corresponding to the most prominent rotation about the y-body axis, would be the best time to apply morphing to the spacecraft. In our demo case, the moment of inertia I_{yy} is changed from 0.35 to 0.2, as per Fig. 5.32a within relatively short period of time of 0.2 s. Results of the simulation are given with Fig. 5.32b and c. Figure 5.32b shows that the simulated morphing led to the step-type increase of the angular velocity ω_y of the body and did not initiate significant oscillations in ω_x and ω_z. In contrast to Case-1, where ω and H plots had similar shapes, in Case-2 these plots are different.

Figure 5.32b shows that morphing did not change the angular momentum H_y and after the morphing was completed, the value of H_y stayed almost unchanged, evidencing that attempt to stop the "flipping" motion was successful. At last, note that as the stabilised value of $|\phi| = 180°$, the stabilised spacecraft is flying *backwards*, with its initial heading attitude changed by 180°!

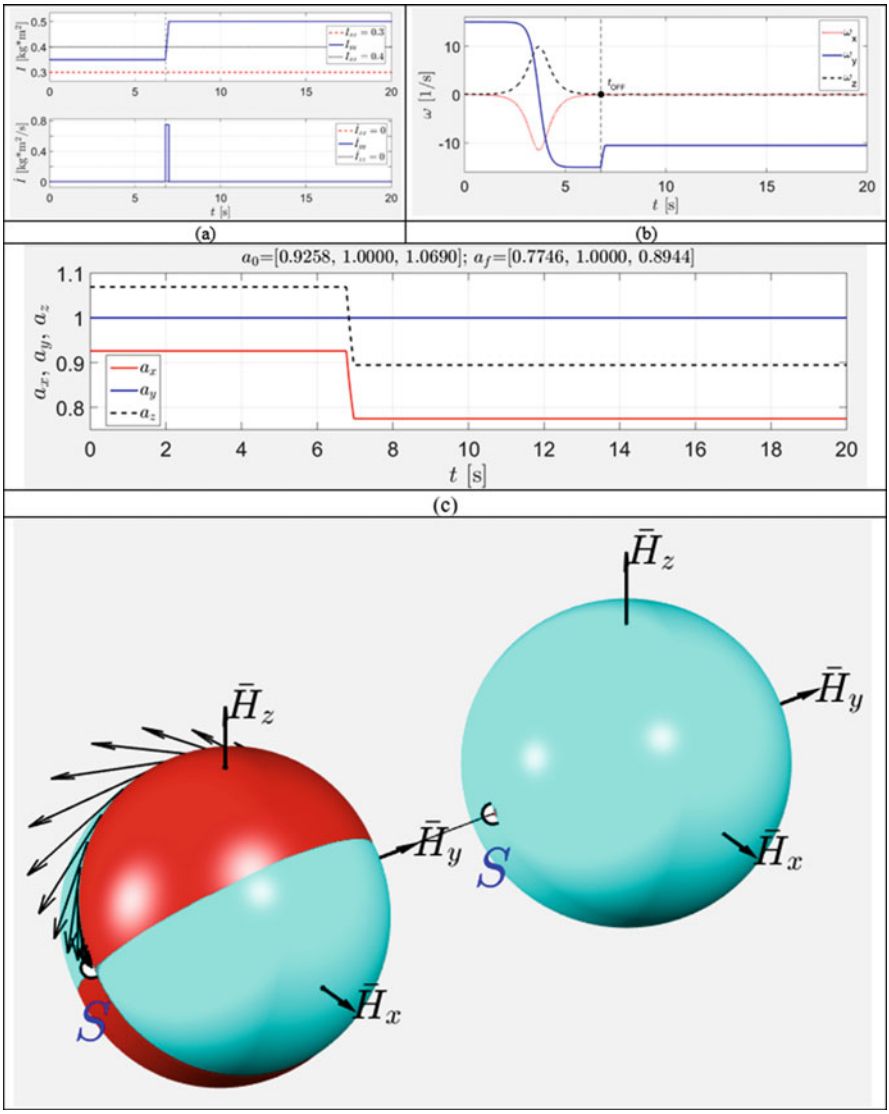


Fig. 5.30 Stopping the flipping motion (solution-2): (**a**) explanation of the inertial morphing applied; (**b**) time history of the resulting angular velocity components ω_x, ω_y, ω_z; (**c**) radii of the KEE; (**d**) geometric interpretation, showing the AMS and KEE before and after application of the inertial morphing

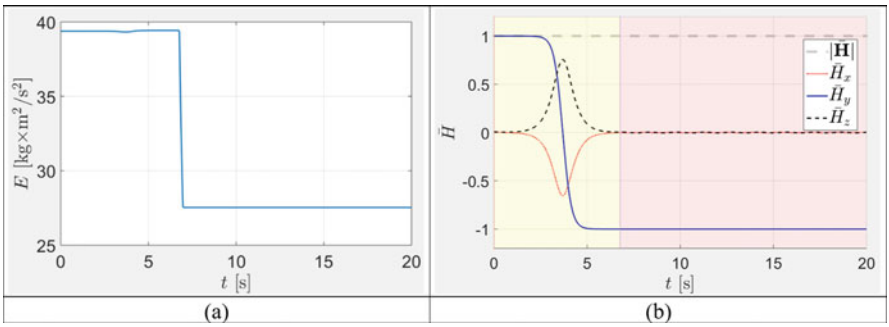


Fig. 5.31 Time histories of the key characteristics of the inertially morphed system: (**a**) kinetic energy E; (**b**) non-dimensional angular momentum $\overline{H}$

5.8.2 *Study Case-3: "Switching OFF" Flipping Motion of the Spacecraft After One Flip (Solution-2)*

It is interesting to observe that stabilisation of the system, illustrated with Fig. 5.33, has been achieved with a controllable change of the moment of inertia I_{yy} (associated with the main rotation of the spacecraft), which initially had its value of $I_{yy,i} = 0.35$, being an intermediate value, surrounded by the smallest $I_{xx} = 0.2$ and largest $I_{zz} = 0.4$ moments of inertia:

$$I_{xx} < I_{yy,i} < I_{zz} \tag{5.43}$$

While keeping values of I_{xx} and I_{zz} unchanged, the value of I_{yy} in the presented experiment was changed from $I_{yy,i} = 0.35$ to the final value of $I_{yy,f} = 0.5$, as per Fig. 5.33a, after which it became the largest principal moment of inertia:

$$I_{xx} < I_{zz} < I_{yy,f} \tag{5.44}$$

Figure 5.33 shows that one of the consequences of the increase of I_{yy} was a reduction from 15 to 10.5 rad/s of the associated angular velocity ω_y of the spacecraft. This simulation result is in perfect agreement with the conservation of the angular momentum of the system, suggesting that ω_y must be reduced by the ratio of $15*(I_{yy,i}/\ I_{yy,f}) = 15*(0.35/0.5) = 10.5$ rad/s.

In contrast to Case-1, where ω and H plots had similar shapes, in the Case-3 these plots are different.

Figure 5.33b shows that morphing did not change the angular momentum H_y and after the morphing was completed, the value of H_y stayed almost unchanged, evidencing that the stopping "flipping" motion has been successful.

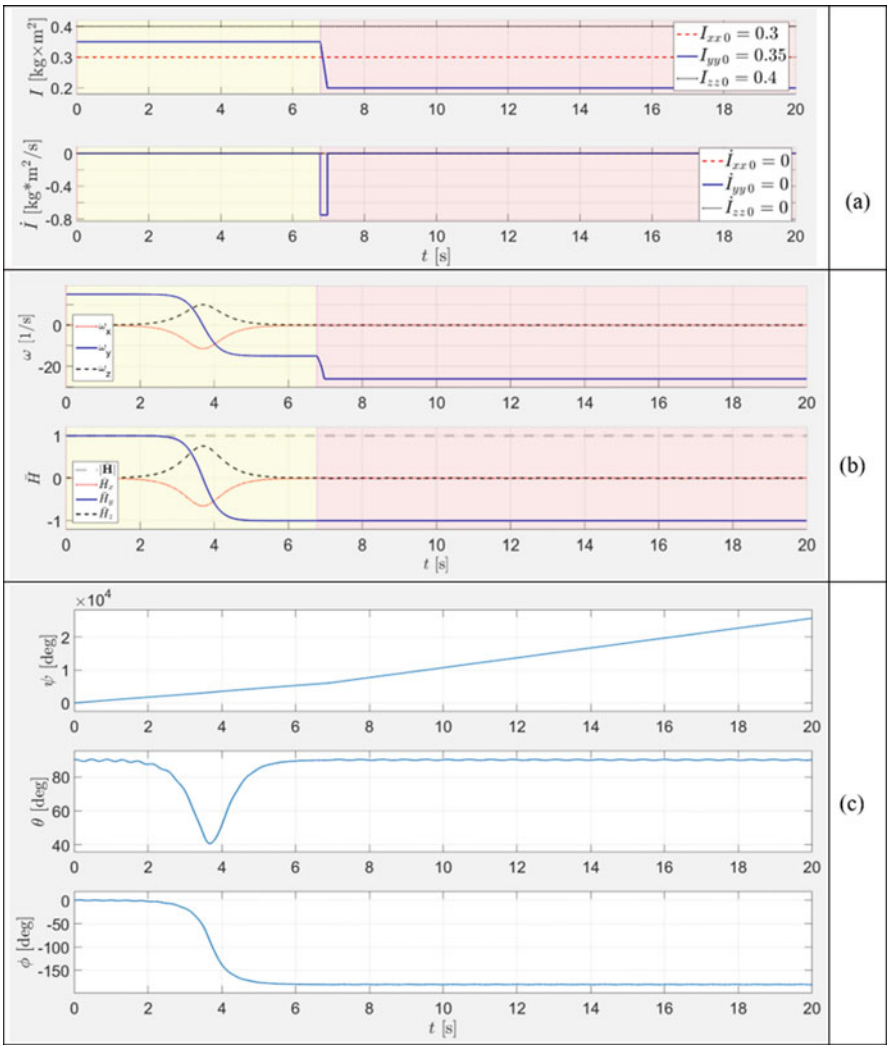


Fig. 5.32 Time histories for Case-2 (switching OFF the flipping motion of the system) of the key parameters: (**a**) controlled moment of inertia I_{yy}; (**b**) angular velocity components; (**c**) angular momentum components; (**d**) Euler angles

5.8.3 *Study Case-4: "Switching OFF" Flipping Motion of the Spacecraft After Two Flips (Solution-1)*

We now demonstrate switching OFF the "flipping" motion of the morphing spacecraft after it performs two flips. The time history of morphing is similar to the one presented in Fig. 5.33, but morphing is starting at $t = 13.54$ s. Results of

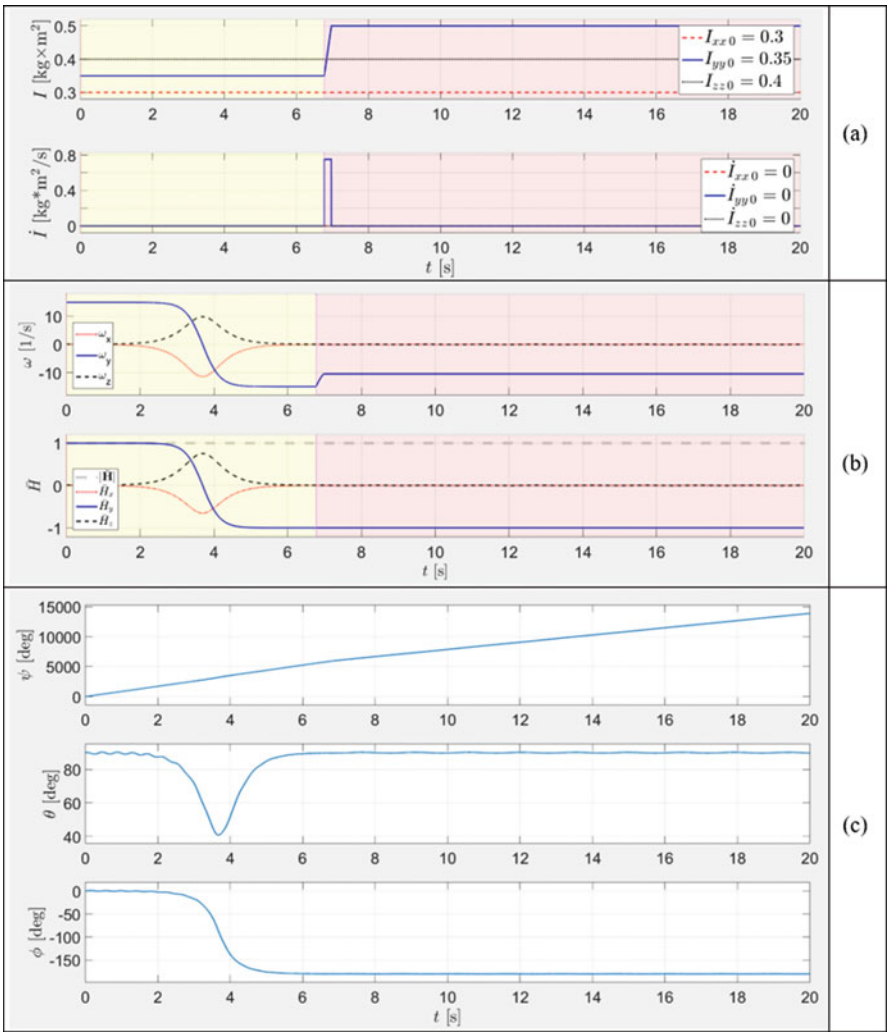


Fig. 5.33 Time histories for Case-3 (switching OFF the flipping motion of the system) of the key parameters: (**a**) controlled moment of inertia I_{yy}, (**b**) angular velocity components, (**c**) angular momentum components and (**d**) Euler angles

this Case-4 are presented in Fig. 5.34. Observed reduction of the angular velocity ω_y is the same, as for the Case-2; however, in Case-2 after the motion is stabilised, the spacecraft continues its flight backwards, whereas in the current case, the stabilised attitude of the spacecraft is the same as at the initial time.

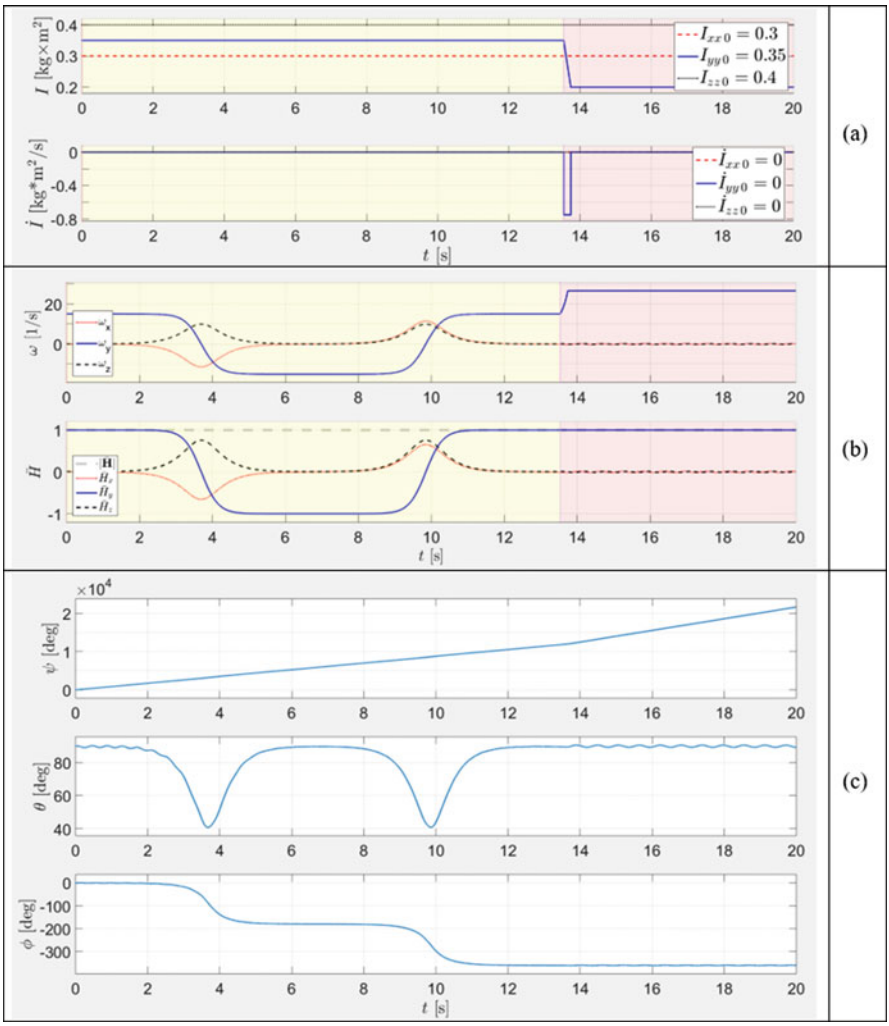


Fig. 5.34 Time histories for Case-4 (switching OFF the flipping motion of the system) of the key parameters: (**a**) controlled moment of inertia I_{yy}, (**b**) angular velocity components, (**c**) angular momentum components and (**d**) Euler angles

5.8.4 Study Case-5: "Switching ON" Spacecraft Flipping Motion

In a similar way as stabilisation, described in Cases 2–4, was achieved, we can initiate the "flipping" motion of the spacecraft. For this, the axis of the major rotation of the system (let's say, y) initially should coincide with the axis of minimal

or maximal moments of inertia, i.e. one of the conditions should be satisfied: $I_{yy} < \min(I_{xx}, I_{zz})$ or $I_{yy} > \max(I_{xx}, I_{zz})$.

In this case, initiated rotation would be stable, without "flipping". To activate the "flipping" motion, morphing of the system should be performed, which should result in I_{yy} becoming an intermediate value between I_{xx} and I_{zz}. In the study Case-5, as illustration, we use the following values: $\omega_x = 0.1$, $\omega_y = 26.25$, $\omega_z = 0.1$ (all – in rad/s), $I_{xx} = 0.3$, $I_{yy,i} = 0.2$, $I_{yy,f} = 0.35$, $I_{zz} = 0.4$ (all – in kg*m^2). The time history of applied morphing is presented in Fig. 5.35a, and the results of the simulation are shown in Fig. 5.35b and c.

5.8.5 *Study Case-6: "Switching ON" Spacecraft Flipping Motion with Following One Flip and "Switching OFF"*

Case-6 represents further development of the Case-5 by switching OFF the "flipping" motion at $t = 9.89$ s, instant of the maximal value of ω_y. The time history of applied morphing is presented in Fig. 5.36a, and the results of the simulation are shown in Fig. 5.36b and c.

5.8.6 *Study Case-7: Control of the Frequency of the Flipping Motion via "Inertial Morphing"*

Case-7 demonstrates the ability of the proposed "inertia morphing" for control of the frequency of the "flipping" motion. In this demo, inertia properties of the "flipping" system are as follows:

- $I_{xx,i} = 0.3$, $I_{yy,i} = 0.395$, $I_{zz,i} = 0.4$ were changed after three 180° flips to
- $I_{xx,f} = 0.3$, $I_{yy,f} = 0.35$, $I_{zz,f} = 0.3 = 0.4$ (all in kg × m^2), as per Fig. 5.37a.

However, the intermediate value of the moment of inertia I_{yy} was still kept within the values of I_{xx} and I_{zz}: $I_{xx} < I_{yy} < I_{zz}$. It is really interesting to observe that it was possible to achieve substantial change of the frequency of the flipping motion. Calculations of the flipping motion periods before and after "morphing" can be calculated, using Eqs. (5.15, 5.16, 5.17, 5.18 and 5.19): $T_i = 21.5$ s and $T_f = 10.8$ s (twice reduction!). This is in perfect correspondence with the results of the simulations, partially shown in Fig. 5.37.

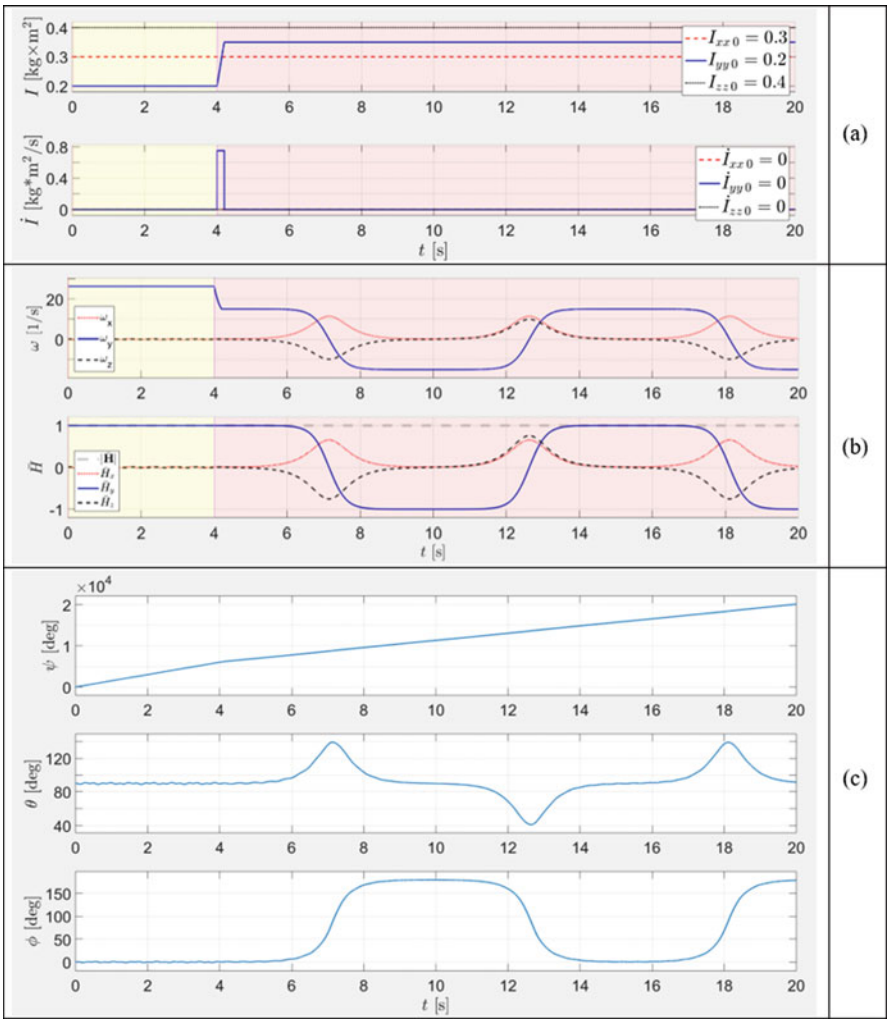


Fig. 5.35 Time histories for Case-5 (switching ON the flipping motion of the system) of the key parameters: (**a**) controlled moment of inertia I_{yy}, (**b**) angular velocity components, (**c**) angular momentum components and (**d**) Euler angles

5.9 Inertial Morphing and the Law of Conservation of Angular Momentum

The law of conservation of angular momentum is one of the fundamental laws in physics, mechanics and quantum mechanics, stating that the total angular momentum of a closed system remains constant. There are numerous fascinating qualitative demonstrations of this law, for example, involving Hoberman sphere

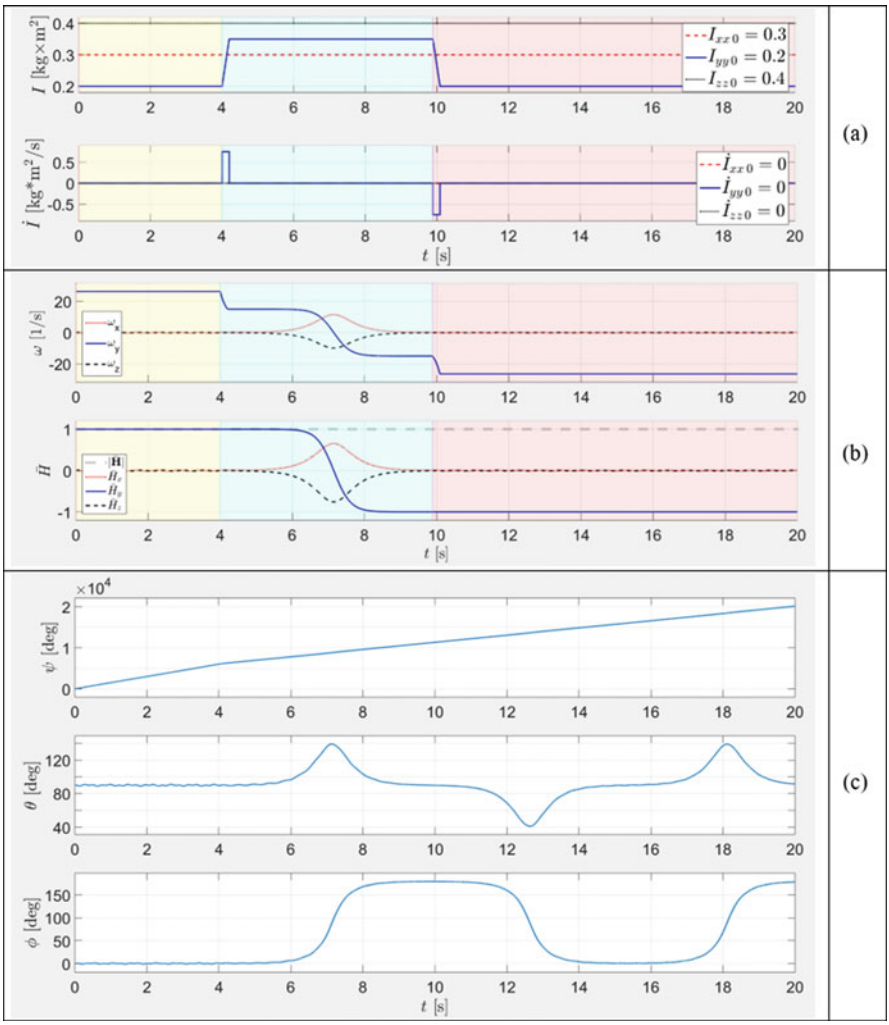


Fig. 5.36 Time histories for Case-6 (switching ON the flipping motion of the system with following one flip and switching OFF) of the key parameters: (**a**) controlled moment of inertia I_{yy}, (**b**) angular velocity components, (**c**) angular momentum components and (**d**) Euler angles

[36], a plastic sphere frame that can be contracted by pulling on a string. Pulling on the string while the sphere is rotating causes the sphere's moment of inertia to decrease and its angular speed to increase, demonstrating conservation of angular momentum.

Figure skaters in a spin (Fig. 5.38a) use conservation of angular momentum: changing their moments of inertia enables them to vary their rate of spin. For the same total angular momentum, they will spin faster by bringing their arms in (lower moment of inertia) and slower by extending their arms (higher moment of inertia).

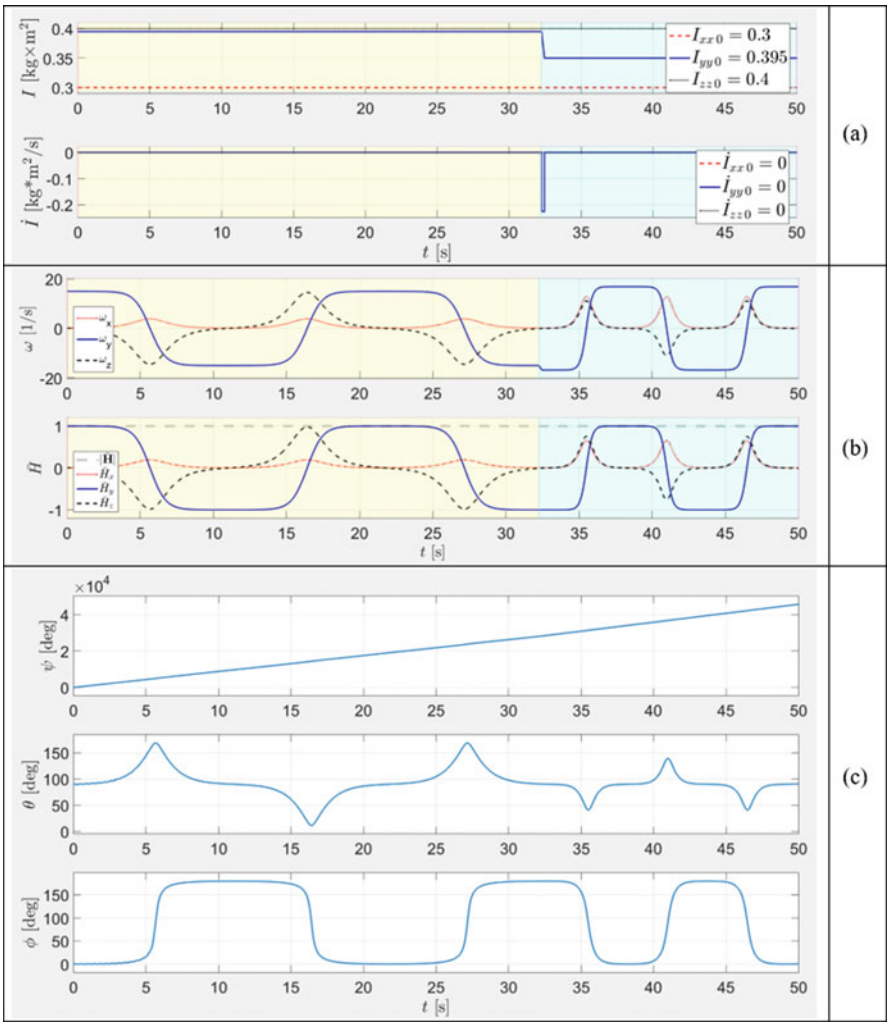


Fig. 5.37 Time histories for Case-7 (control of the frequency) of the key parameters: (**a**) controlled moment of inertia I_{yy}, (**b**) angular velocity components, (**c**) angular momentum components and (**d**) Euler angles

Reduction of the moment of inertia can be achieved by drawing skater's arms and legs. Increase of the moment of inertia can be achieved by moving the skater's arms outwards.

For example, let us assume the initial spin rate of the skater to be equal to 30 rad/s and ability to increase its angular momentum by a factor of 3. Then, using the law of conservation of energy, and ignoring energy losses due to the ice friction and air drag, etc. during the spin, we can estimate a new spin rate of the skater being equal to 10 rad/s.

Fig. 5.38 A figure skater, employing manipulations with the moments of inertia of the human body to control the spin rate: (**a**) fast spin and (**b**) fastest spin by O. Oliver of 342 RMP or 5.7 Hz (19 January 2015, Guinness Book of World Records)

Figure 5.38b shows Olivia Oliver, who achieved the fastest recorded spin on ice skates of 342 RPM (this record was achieved in Warsaw, Poland, on 19 January 2015 and is in the Guinness Book of World Records [37]). This spin speed corresponds to 5.7 Hz or the angular velocity of 35.81 rad/s.

All numerical demonstrations in Figs. 5.29a, 5.32b, 5.33b, 5.34b, 5.35b, 5.36b and 5.37b, involving inertial morphing, showed that the angular momentum was conserved: $|\ \overline{\mathbf{H}}\ |$, the absolute value of the magnitude of the non-dimensional vector of the angular momentum $\overline{\mathbf{H}}$, plotted with grey dashed line, remained its constant value.

Analysis of the numerically simulated changes in the angular velocity of the presented rotating systems was in full agreement with the law of conservation of angular momentum. In particular example, Study Case-3, applying inertial morphing, which resulted in increase of the angular momentum I_{yy} from 0.35 to 0.5 kg $\times$ m^2, has led to the reduction of the angular velocity from 15 to 10.5 rad/s. These changes allow us to express this observation as follows: $I_{yy,i}\omega_{y,i} = I_{yy,f}\omega_{y,f}$.

Figure 5.33 shows that one of the consequences of the increase of I_{yy} was a reduction from 15 to 10.5 rad/s of the associated angular velocity ω_y of the spacecraft. This simulation result is in perfect agreement with the conservation of the angular momentum of the system, suggesting that ω_y must be reduced by the ratio of $15*(I_{yy,i}/\ I_{yy,f}) = 15*(0.35/0.5) = 10.5$ rad/s.

On the surface, there is apparent similarity between the skaters and proposed inertially morphed spacecraft, manipulating with the inertial properties.

However, there is a significant difference: skater is "operating" within the stable modes of motion [38], whereas a morphed spacecraft is actually extracting most advantages from its inertial manipulations in its ability to manipulate with the attitude, as it is able to access unstable motions!

In this respect, we would like to suggest the following analogy, which could be easier understood by a wide audience. Let us consider nothing more complex than heating of water. When this is occurring for the water temperature, being within the range 0–100 °C, let us say 80 °C, then the heating energy, passed on to the water, leads to the *quantitative* change in the water's temperature. However, if the energy is passed on the system, after we reached 100 °C, then we can observe conversion of the water from the liquid state to the gaseous state! This is a change in *quality* of the substance. In a similar way, considering an experiment with water to be cooled, where water has initial temperature, let us say 10 °C, we can see conversion of the matter from liquid to the solid state after we reached 0 °C. This would be also a case with a change in *quality* of the substance.

Similar to the heating/cooling of the water within the range 0–100 °C, manipulation with the moments of inertia, during which their order, let us say, $I_{xx,i} < I_{yy,i} < I_{zz,i}$ is not changing with $I_{xx,f} < I_{yy,f} < I_{zz,f}$, rotational characteristics of the system (like angular velocity components) are changing only in a sense of their quantities, but without the change in the "quality" of the rotational motion. However, similar to the water being evaporated or frozen, with the change of the order of the moments of inertia (due to the applied IM), let us say, from $I_{xx,i} < I_{yy,i} < I_{zz,i}$ to $I_{yy,f} < I_{xx,f} < I_{zz,f}$, (or $I_{yy,f} < I_{zz,f} < I_{xx,f}$ or $I_{xx,f} < I_{zz,f} < I_{yy,f}$ or $I_{zz,f} < I_{xx,f} < I_{yy,f}$), then, in addition to the *quantitative* changes in the rotational parameters of the system, we would be also expecting a *qualitative* change in the motion.

Therefore, it is proposed to distinguish between the inertial morphings, which *do not employ* non-stable modes of motion and is called the *cold inertial morphings* from the inertial morphings, which *employ* non-stable modes of motion and are called the "hot IM". With these definitions, the most useful of the IM for spacecraft, allowing acrobatic capabilities, are "hot inertial morphings".

5.10 Inertial Morphing in Novel Designs of Acrobatic Spacecraft for 180 Degrees Inversions: Method of "Installing into Separatrix" with Pole-Separatrix-Pole Transfer

5.10.1 Applications of Acrobatic Missions

It is believed that new results of this research may have multiple applications. We would like to propose and list a few of them, which may be useful in planning new space missions:

1. *Assistance in establishing formation flight:*

Eject one spacecraft in the forward direction, and then, after one flip, eject another spacecraft in the opposite direction, so that the formation flying is achieved in orbit around the Earth or other planet or celestial body (comet, asteroid).

2. *Thruster direction control*:

To boost (accelerate) or decelerate spacecraft velocity by only one thruster. Of course, deep space mission vehicles change its attitude to boost (escape from the Earth) and to decelerate its velocity when it enters to an orbit around another planet.

Usually other small thrusters (or reaction wheels) are used to change attitude of vehicles. In addition to this, our new proposed method could be also used for changing the attitude of the spacecraft and its following stabilisation in the fixed body frame.

3. *Frequency of "flipping" motion control:*

Inertial morphing can be used to control within wide range the frequency of the "flips" for the "Garriott's-Dzhanibekov's effect" mode. However, we showed that there is a minimum (i.e. low bounding limit) for the period of these oscillations.

5.10.2 Illustrated Description of Application of IM for Thruster Direction Control

Figure 5.39 is dedicated to one (out of many) possible application of the IM. It illustrates inversion of the spacecraft, using two morphing procedures: the first is to activate unstable flipping motion and second is to stop flipping motion. This enables for a single thruster to be used for acceleration of the spacecraft (boosting stage, Fig. 5.39a) and also for its deceleration (braking stage, Fig. 5.39f).

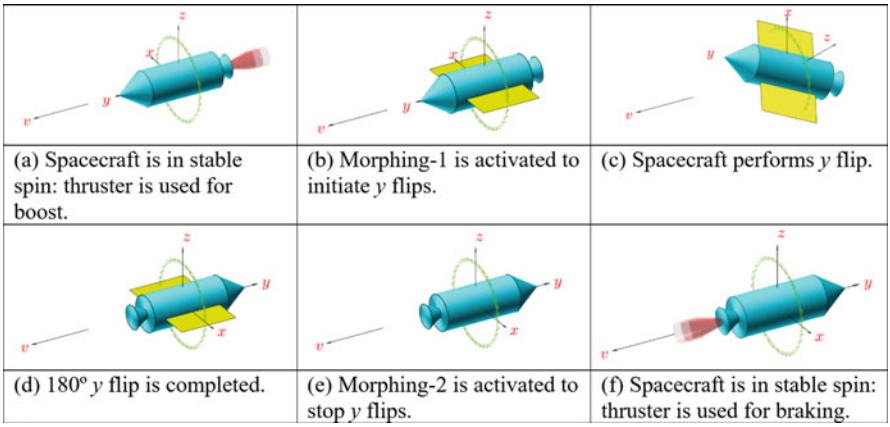


Fig. 5.39 Examples of application of inertial morphing for inversion of spacecraft [9]: (**a**–**f**) 180° inversion, allowing to use the same thruster for the boost and braking stages

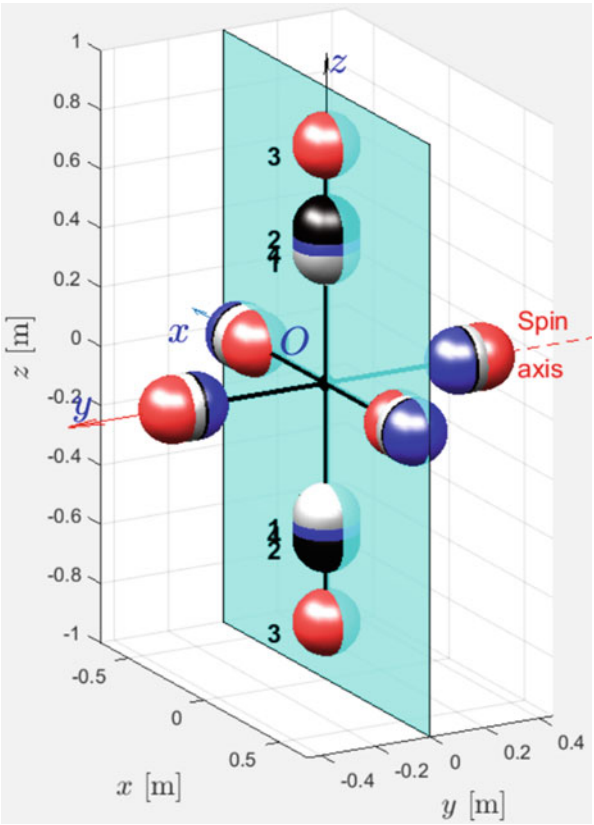


Fig. 5.40 Positions of the masses during spacecraft inversions via inertial morphing for acrobatic manoeuvre in Fig. 5.39(**a**)–(**e**): 180 degrees inversion, illustrated for six-mass spacecraft conceptual model [9]

5.10.3 Fast 180 Degrees Inversion of the Spacecraft

To demonstrate the feasibility of the proposed application, let us consider six-mass model of the spacecraft with corresponding numerical parameters $m_x = 4$ kg, $m_y = 5$ kg; $m_z = 1$ kg. The key requirement to this design would be the ability of the system to reposition six paired masses (e.g. via linear actuators), in accordance to the control considerations. Let us assume that the mission profile would enable instalment of the spacecraft with initial predominant spin about y-axis (with angular velocity $\omega_{y,i} = 8.57$ and other components of ω being small: $\omega_{x,i} = \omega_{z,i} = 0.01$, all in rad/s). If the masses are initially positioned at their locations, shown in Fig. 5.40 with *white* spheres, with position radii equal to $r_{x,i} = 548$, $r_{y,i} = 510$, and $r_{z,i} = 447$ mm correspondingly, then the inertial properties of the spacecraft would be $I_{xx,i} = 3$; $I_{yy,i} = 2.8$; $I_{zz,i} = 5$ (all in kg $\times$ m^2). With these selected parameters, the y spin of the system would be stable, as axis of rotation coincides with the minimal inertia axis.

Let assume that at the instant $t = 10$ s we wish to initiate flipping motion of the system. This can be achieved by applying inertial "Morphing-1", during which

moment of inertia I_{yy} should become an intermediate axis. Aiming for the new moments of inertia to be $I_{xx,1} = 3$; $I_{yy,1} = 3.4$; $I_{zz,1} = 5$ (kg × m^2), we calculate new radii for the masses, using the following relationships, as per Eq. (5.39):

$$\begin{aligned} r_{x1} &= \tfrac{1}{2}\sqrt{\left(I_{yy,1} + I_{zz,1} - I_{xx,1}\right)/m_x} \\ r_{y1} &= \tfrac{1}{2}\sqrt{\left(I_{zz,1} + I_{xx,1} - I_{yy,1}\right)/m_y} \\ r_{z1} &= \tfrac{1}{2}\sqrt{\left(I_{xx,1} + I_{yy,1} - I_{zz,1}\right)/m_y} \end{aligned} \tag{5.45}$$

Resultant values of the new position radii, ensuring transition of the system from stable regular spin to the flipping unstable motion, are as follows: $r_{x1} = 581$, $r_{y1} = 480$ and $r_{z1} = 592$ mm. Therefore, to trigger the spinning motion, it is just necessary to move paired masses from initial positions (shown in Fig. 5.40 with white spheres) to their new positions, shown with *black* spheres. For better perception of the 3D design, an imagined semi-transparent *xz* plane is added to the figure.

For the system, initially satisfying $I_{yy,i} < I_{xx,i} < I_{zz,i}$ condition, rapid assignment at $t = 10$ s of the new moments of inertia, satisfying now $I_{xx,1} < I_{yy,1} < I_{zz,1}$ condition, transfers regular *y* spin motion into unstable spin. Its period can be calculated using Eqs. (5.15, 5.16, 5.17, 5.18 and 5.19) for the corresponding regular spin conditions at $t = 10$ s:

- $I_{xx,1} = 3$, $I_{yy,1} = 4.8$, $I_{zz,1} = 5$, $\omega_{x1} = -0.2917$, $\omega_{y1} = 5.0284$, $\omega_{z1} = -0.6714$.

Calculations give us $T = 17.784$ s; therefore, the time to 180° flip would be $T/2 = 8.8920$ s. At this very moment, rapid "Morphing-2" should be applied to stop flipping phase and stabilise the system. For this morphing, in accordance with [3], any set of new moments of inertia can be selected, strictly satisfying any of the two conditions: $I_{yy,2} < I_{xx,2} < I_{zz,2}$ ("solution I_{yy} min" strategy) or $I_{xx,2} < I_{zz,2} < I_{yy,2}$ ("solution I_{yy} max" strategy). In one case scenario, return to the initial moments of inertia can be implemented. However, as an additional example, we illustrate implementation of the $I_{xx} = 3$, $I_{yy} = 5.2$, $I_{zz} = 5$ ("solution-1" strategy) scenario. To achieve these new inertia characteristics, as per Eq. (5.39), spacecraft control masses should be rapidly moved to their final positions: $r_{x2} = 671$, $r_{y2} = 374$ and $r_{z2} = 894$ mm.

The described morphings (achieved via controlled changes in $r_x(t)$, $r_y(t)$ and $r_z(t)$), time histories for the resulting moments of inertia, angular velocity and angular momentum components are presented in Fig. 5.41. Interestingly, that rapid changes in ω_y and I_{yy} do not lead to similar changes in $\overline{H}_y$. Also, as evidenced by Fig. 5.41c, the total angular momentum $|\overline{\mathbf{H}}|$ is conserved during the flipping acrobatics (Figs. 5.42 and 5.43).

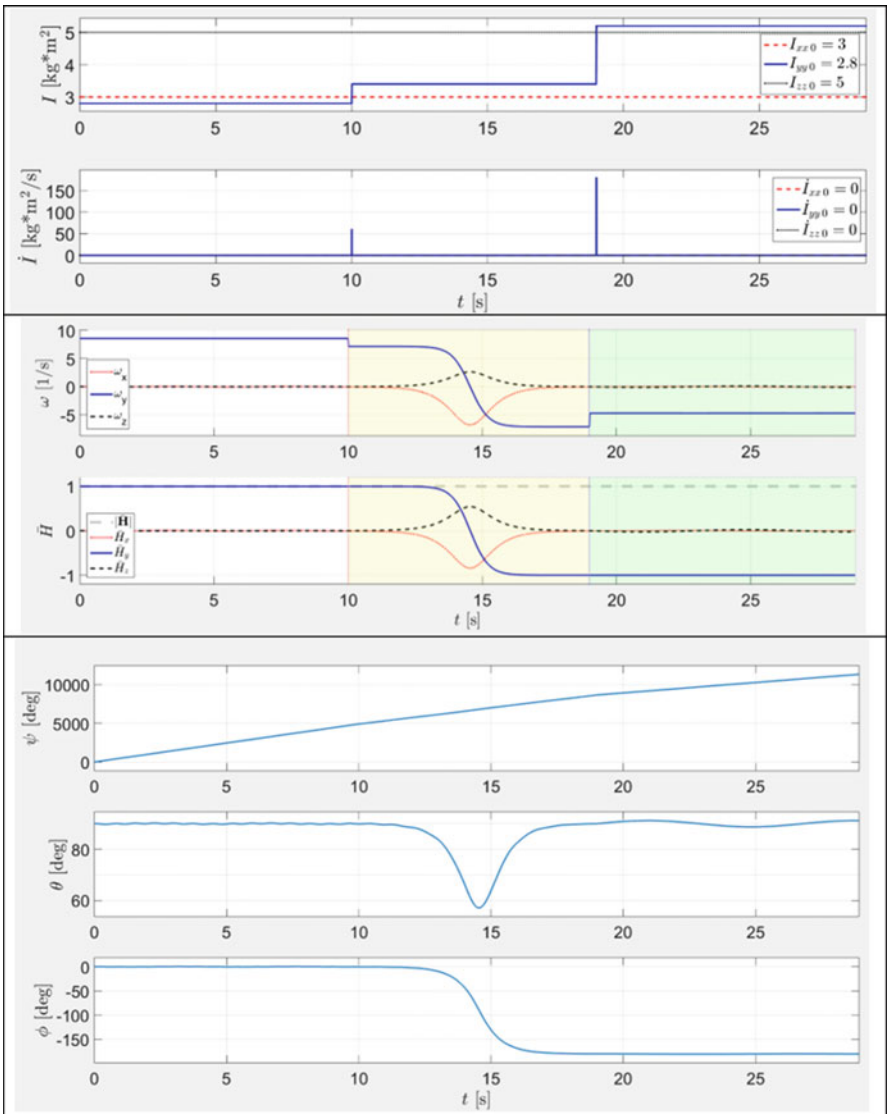


Fig. 5.41 Variation of the system's parameters during fast 180 degrees inversion: (**a**) principal moments of inertia and their derivatives, (**b**) angular velocity and non-dimensional angular momentum components and (**c**) Euler angles

5.10.4 Slow 180 Degrees Inversion of the Spacecraft (Figs. 5.44, 5.45 and 5.46)

The same principle can be used for complete reconfiguration ("repackaging") of the articulated compound spacecraft, consisting, for example, of three segments A_1-B_1, B_2-C_2 and C_3-D_3. If all of these segments are un-docked, they can

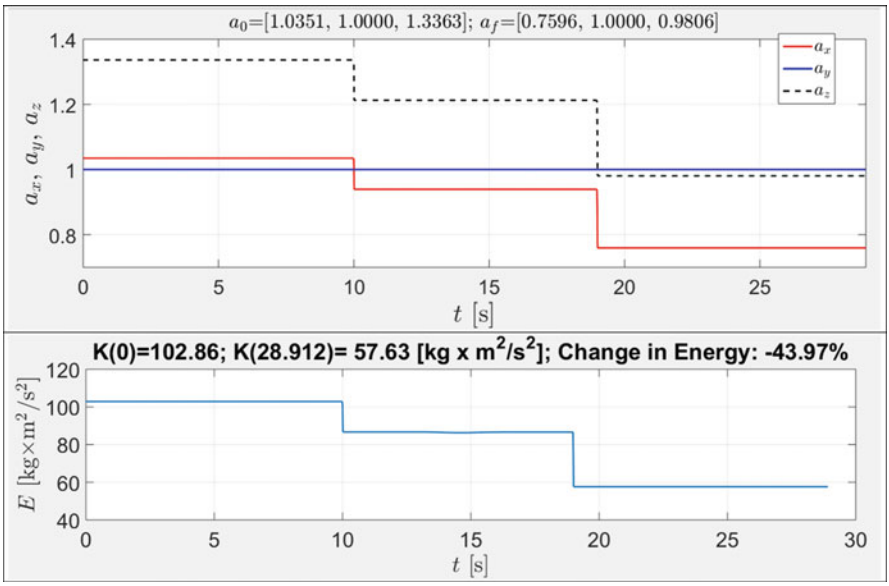


Fig. 5.42 Radii of the KEE and kinetic energy E of the system during 180 degrees inversion

independently perform the same flipping manoeuver, as described above, and then (after stabilisation) docked to a new configuration: B_1-A_1, C_2-B_2 and D_3-C_3. Note that the reconfiguration can be applied to the selected segments only and to the spacecraft with any number of segments. For example, if only central segment B_2-C_2 is inverted, the new configuration would be A_1-B_1, C_2-B_2 and C_3-D_3.

5.11 Inertial Morphing in Novel Designs of Acrobatic Spacecraft for De-tumbling: Method of "Installing into Separatrix" with Polhode-Separatrix-Pole or Polhode-Polhode-Separatrix-Pole Transfer

5.11.1 Application of Inertial Morphing to the Tumbling Spacecraft Model: Observations

In the classical "Garriott's-Dzhanibekov's effect" cases, the rigid body is "given" and its inertial properties are assumed fixed. Moreover, typically, it is assumed that the rigid body/system is provided with most prominent rotation about the y-axis, with two other rotations being very small.

Let us explore more general case where we will also allow other rotations to be more significant. Assume, for example, the following parameters of the system:

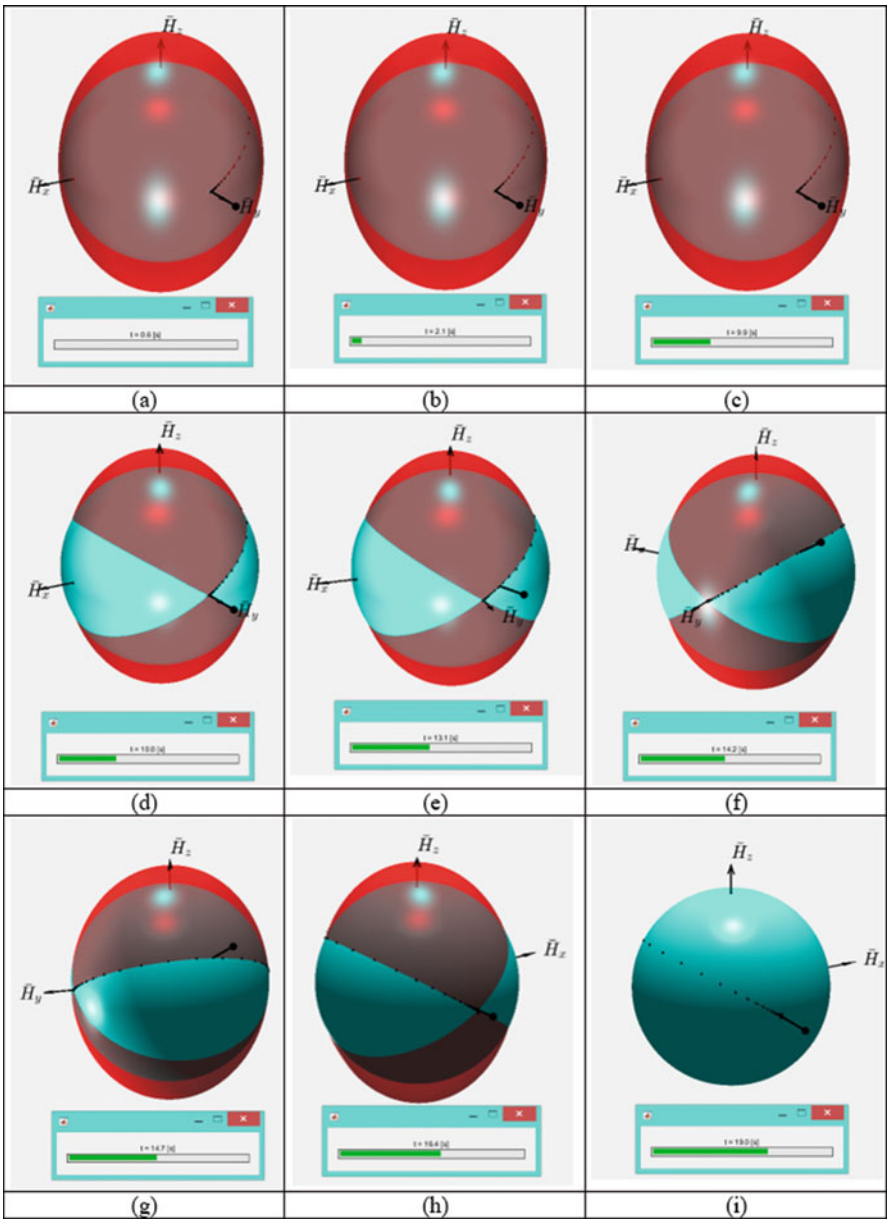


Fig. 5.43 Collocated AMS and KEE for different stages of the 180° inversion: (**a**) $t = 0.6$ s, (**b**) $t = 2.1$ s, (**c**) $t = 9.9$ s, (**d**) $t = 10$s, (**e**) $t = 13.1$ s, (**f**) $t = 14.2$ s, (**g**) $t = 14.7$ s, (**h**) $t = 16.4$ s and (**i**) $t = 19$–28 s

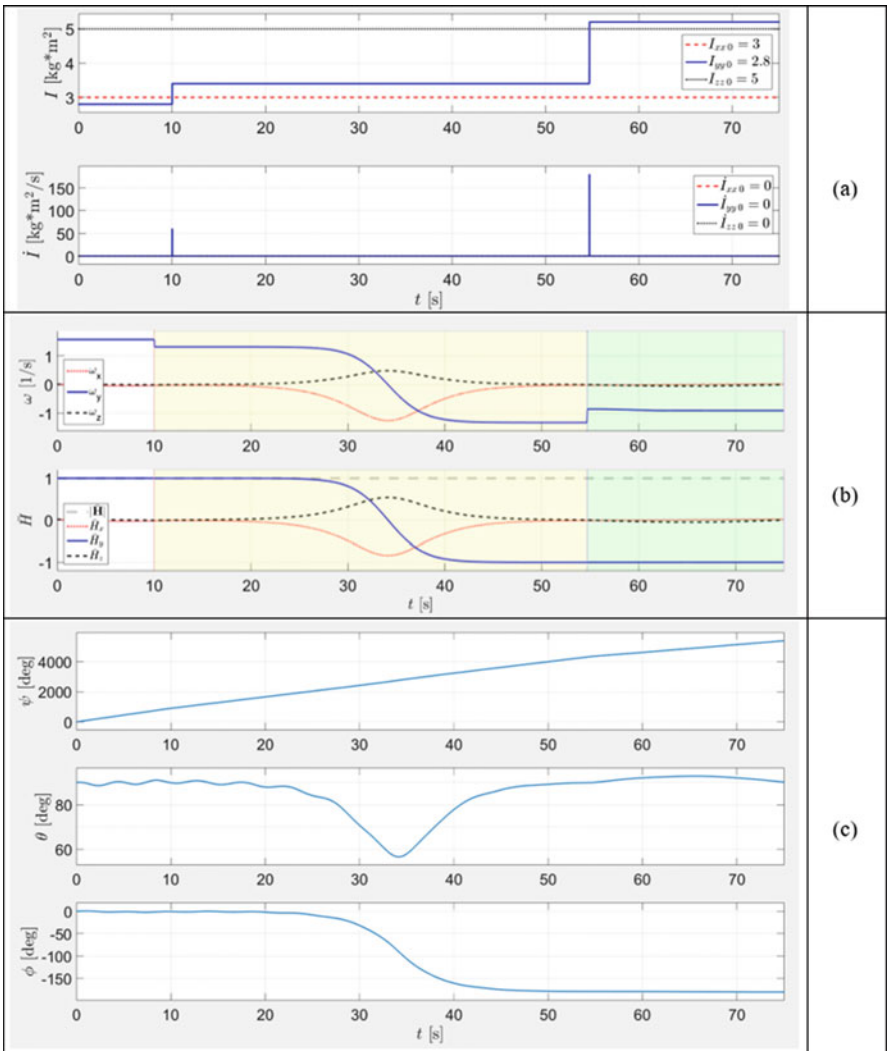


Fig. 5.44 Variation of the system's parameters during slow 180 degrees inversion: (**a**) principal moments of inertia and their derivatives, (**b**) angular velocity and non-dimensional angular momentum components and (**c**) Euler angles

$I_{xx} = 0.3$, $I_{yy} = 0.395$, $I_{zz} = 0.4$ (all in kg $\times$ m^2), with the initial conditions $\omega_{x,i} = 3$, $\omega_{y,i} = 15$, $\omega_{z,i} = 0.1$ (all in rad/s). System can be simulated, using Eq. (5.7). The time history of the angular velocity components is given with Fig. 5.47. Firstly it shows that the initially small x-rotation is quickly evolving and becoming comparable with initially dominant y-rotation. Also, it shows that the pattern of motion is different from the observations in the Garriott's–Dzhanibekov's effect cases: the angular

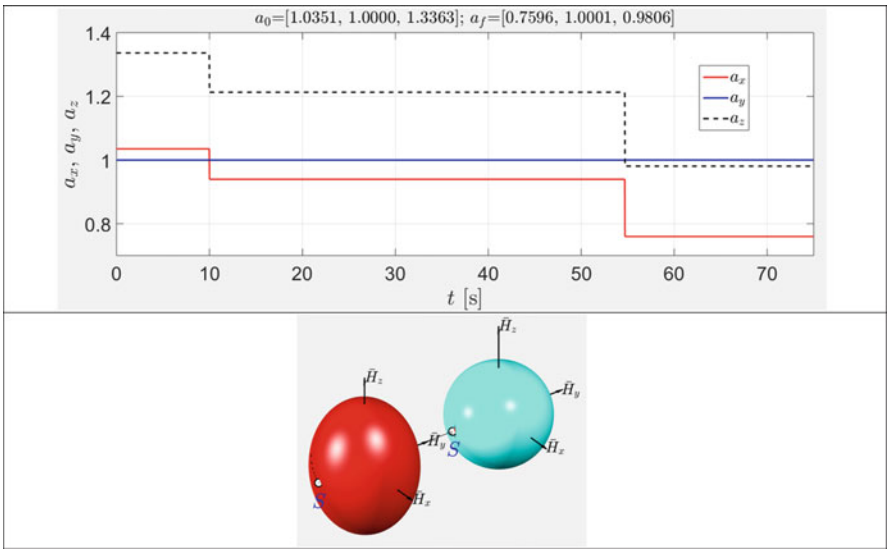


Fig. 5.45 Variation of the system's parameters during slow 180 degrees inversion: (**a**) values of the semi-major axes and (**b**) AMS and KEE for the initial instant (on the left) and after completion of the I (on the right)

velocity curves are free from distinct sharp gradients, which prompts that there are no intermitted fast and slow phases.

Quick analysis of the Euler angles in Fig. 5.47 suggests that the motion of the body can be characterised as tumbling, due to active involvement of several rotations.

The $\overline{H}$ sphere and K ellipsoid can be also constructed for the case. They are shown in Fig. 5.48, which shows two lines of intersection between $\overline{H}$ sphere and K ellipsoid, with only one being the solution hodograph. Figure 5.48 also confirms previous observation: the density of the step points on the hodograph is rather uniform, and the arrows of the quiver plot have almost the same length.

In the following series of experiments, we will be using extended Euler's equations, taking into account variations in the system's moment of inertia.

In the next several experiments, we will be starting simulation of the system with these parameters, $I_{xx} = 0.3$, $I_{yy} = 0.395$, $I_{zz} = 0.4$ (all in kg $\times$ m^2), with the initial conditions $\omega_{x,i} = 3$, $\omega_{y,i} = 15$, $\omega_{z,i} = 0.1$ (all in rad/s), and then will initiate morphing of the system at different times in the experiments.

Firstly, let us apply morphing, as per Fig. 5.49.

Reduction of the y moment of inertia leads to the slight increase in its y angular velocity, as evidenced by Fig. 5.50a with time histories of angular velocity components.

In fact, as the angular momentum is conserved, the ω_y is expected to increase to 15*(0.395/0.350) = 16.9 rad/s. However, this leases to the change in the kinetic

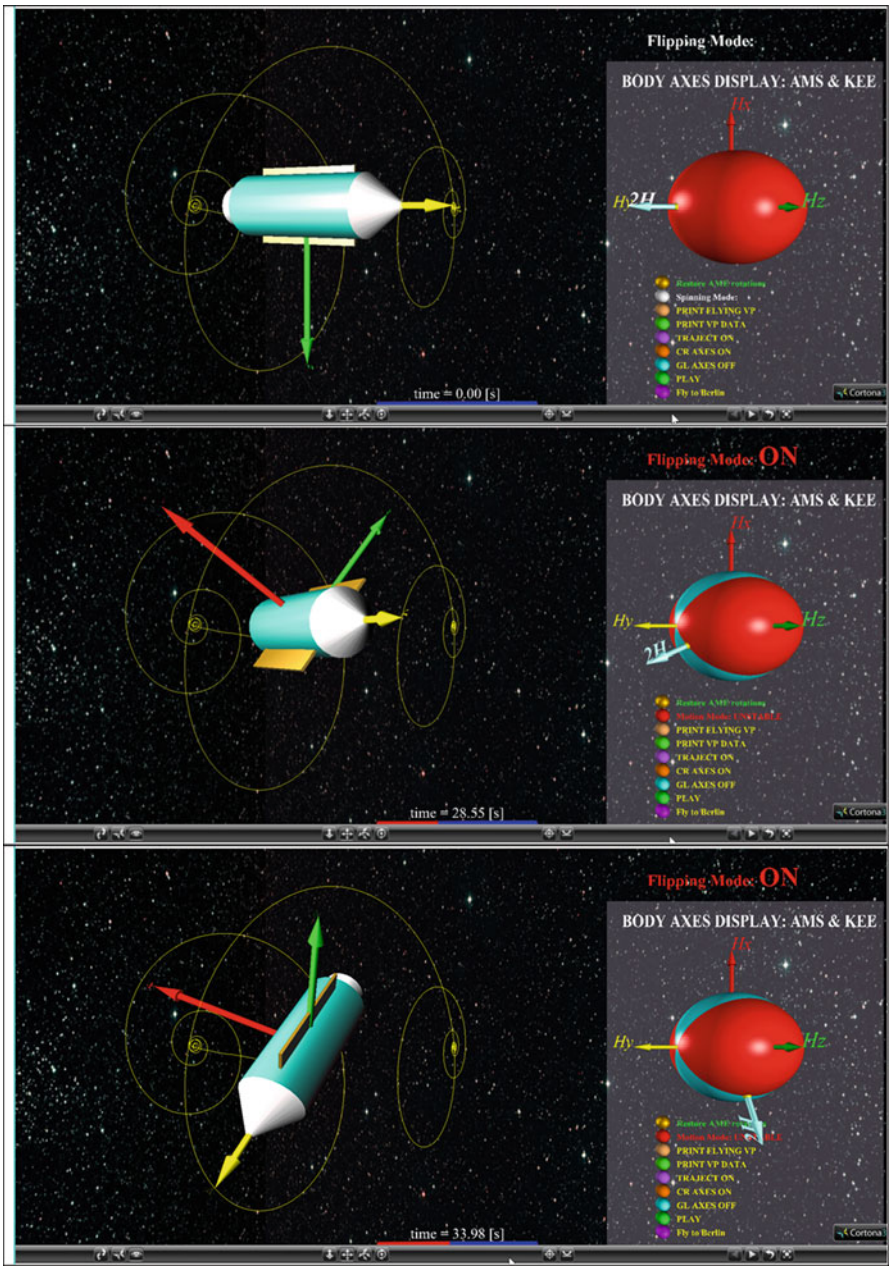


Fig. 5.46 180 degrees inversion of the spacecraft, using IM: flipping acrobatic spacecraft "draws" on the dome with *y* intermediate body axis (on the left) concurrently with the AMS/KEE simulation Virtual Reality control panel (on the right): (**a**) spacecraft initially is in stable spin; (**b–d**) transition stage, when spacecraft flips after unstable mode is activated with first IM; and (**e**) spacecraft is in stable spin, stabilised with the second IM

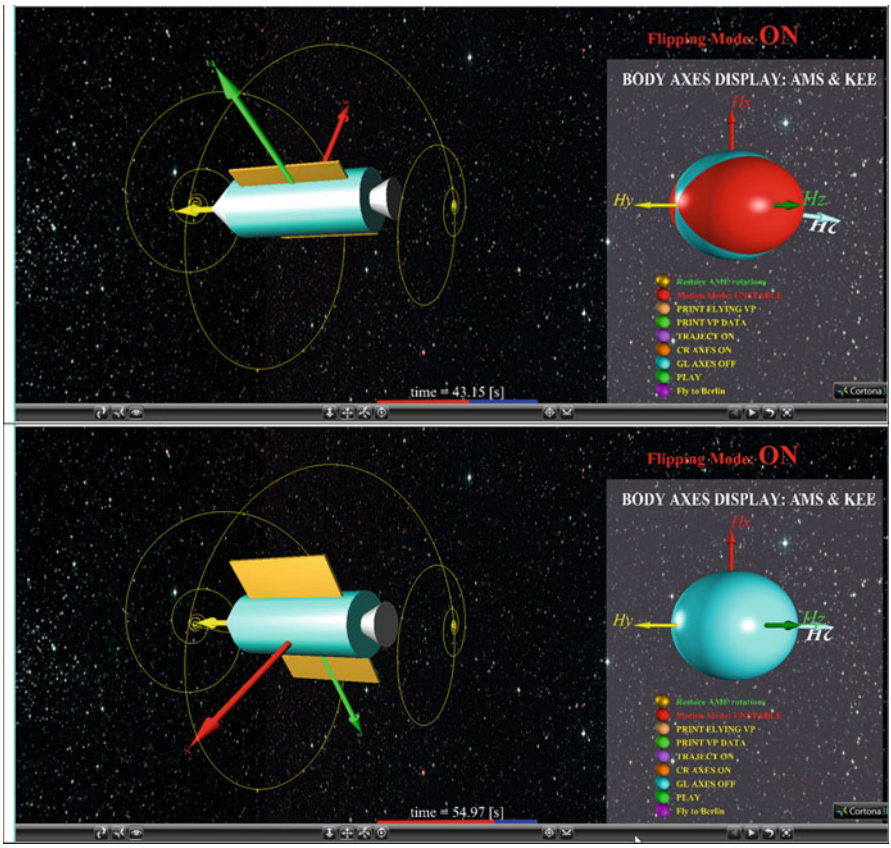


Fig. 5.46 (continued)

energy of the system: change from 45.79 to 51.32 [kg $\times$ m^2/s^2]. As a result of this, the semi-major axes of the energy ellipsoid should be re-calculated, using Eq. (5.12), giving us the following numbers:

- Initial semi-major axis of the KEE: 0.8746, 1.0036, 1.0099
- Final semi-major axis of the KEE: 0.9271, 1.0014, 1.0706.

These numbers suggest that the kinetic energy ellipsoid is now bulging more in the areas of its poles. In view of changed K shape, we now need to show two sets of collocated $\overline{H}$ sphere and K ellipsoid, corresponding to the stages before and after morphing (applied at 3.32 s), as shown in Fig. 5.50.

Note that the angular velocity vector is now escaping from its initial hodograph, leaving it at the point S (corresponding to 3.32 s) and continuing its motion along another trajectory on another $\overline{H}$ sphere from point S.

It is interesting to realise that the hodograph trajectory is very sensitive to the instant of initiation of morphing. To illustrate this, we perform two more

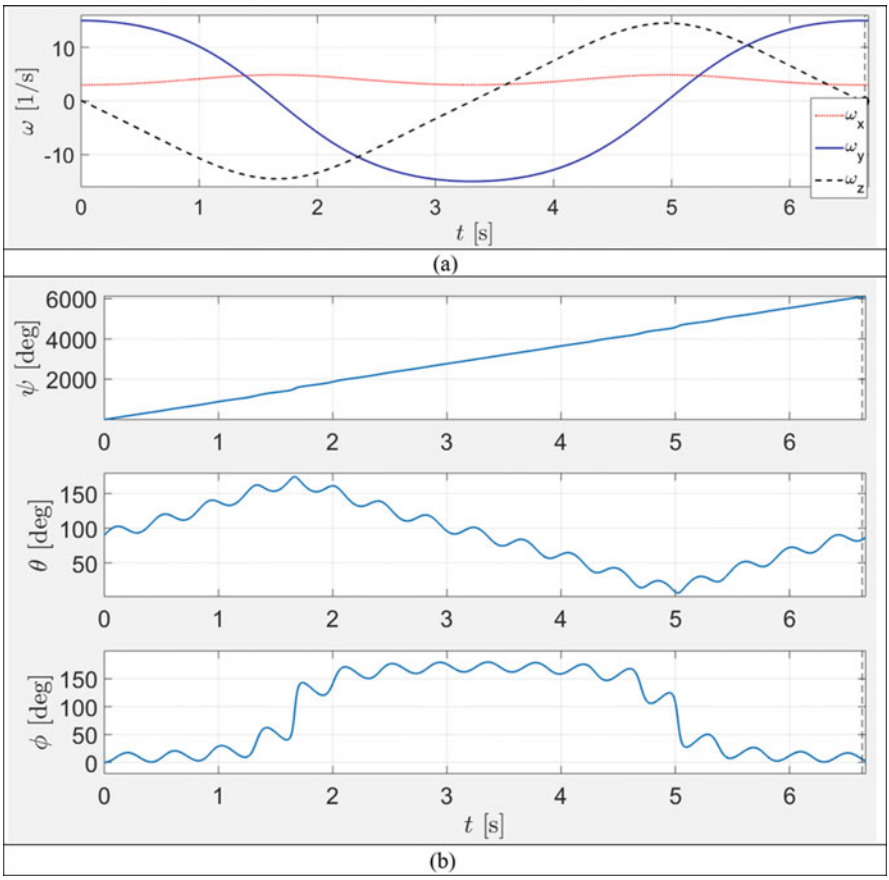


Fig. 5.47 Time histories of the parameters of the study case system with $I_{xx} = 0.3$, $I_{yy} = 0.395$, $I_{zz} = 0.4$ (all in kg $\times$ m^2), with the initial conditions $\omega_{x0} = 3$, $\omega_{y0} = 15$, $\omega_{z0} = 0.1$ (all in rad/s): (**a**) angular velocities and (**b**) Euler angles

experiments, this time, applying morphing at 4 s. Geometric interpretation of this case, presented in Fig. 5.51, shows that the after-morphing hodograph is now circling around the z-axis, which is in contrast to Fig. 5.50, showing coning/circling about x-axis. Figure 5.51 is also suggesting that the period of the circling has been reduced after the morphing.

Two more cases are presented for completeness. They are illustrated with Figs. 5.52 and 5.53. They look similar in terms of the rotational motion; however, their hodographs are at the different poles of the $\overline{H}$ *sphere, on the "North" and on the "South" poles.*

These simple experiments with the application of the same inertial morphing are showing importance of the instant, when this IM is applied, and also they show variety of results to be achieved. In particular, it was possible from common initial

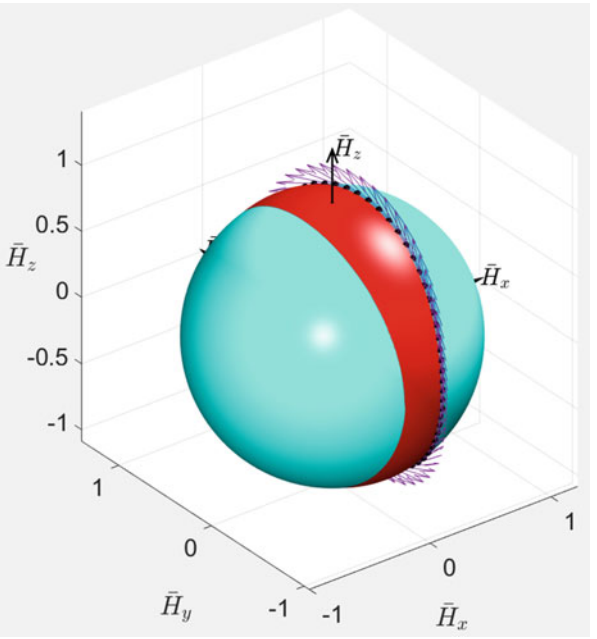


Fig. 5.48 Angular momentum sphere (AMS) and kinetic energy ellipsoid (KEE) for the system with compound rotations

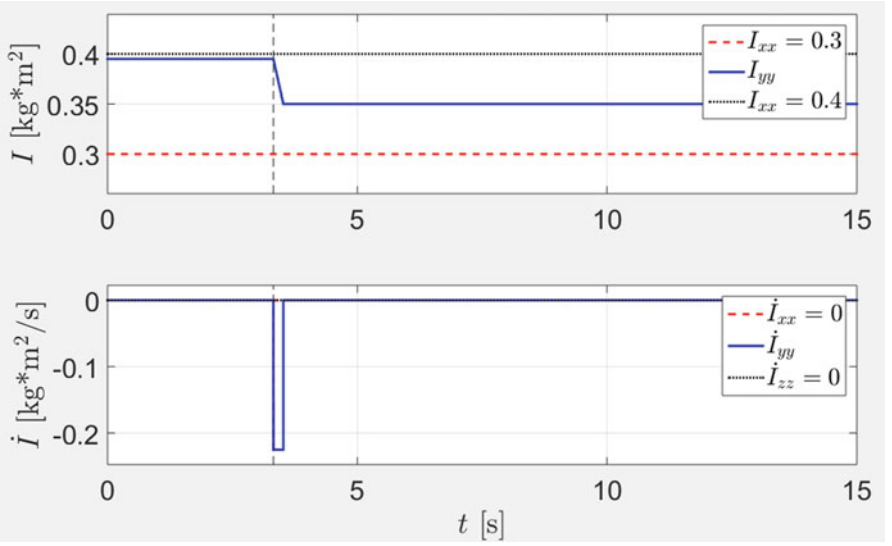


Fig. 5.49 Time history of the controlled manipulation with the moment of inertia I_{yy}

polhode to transfer to new polhodes around various body axes. Essentially, in all of these experiments, we were able to transfer from one polhode to another. In the next

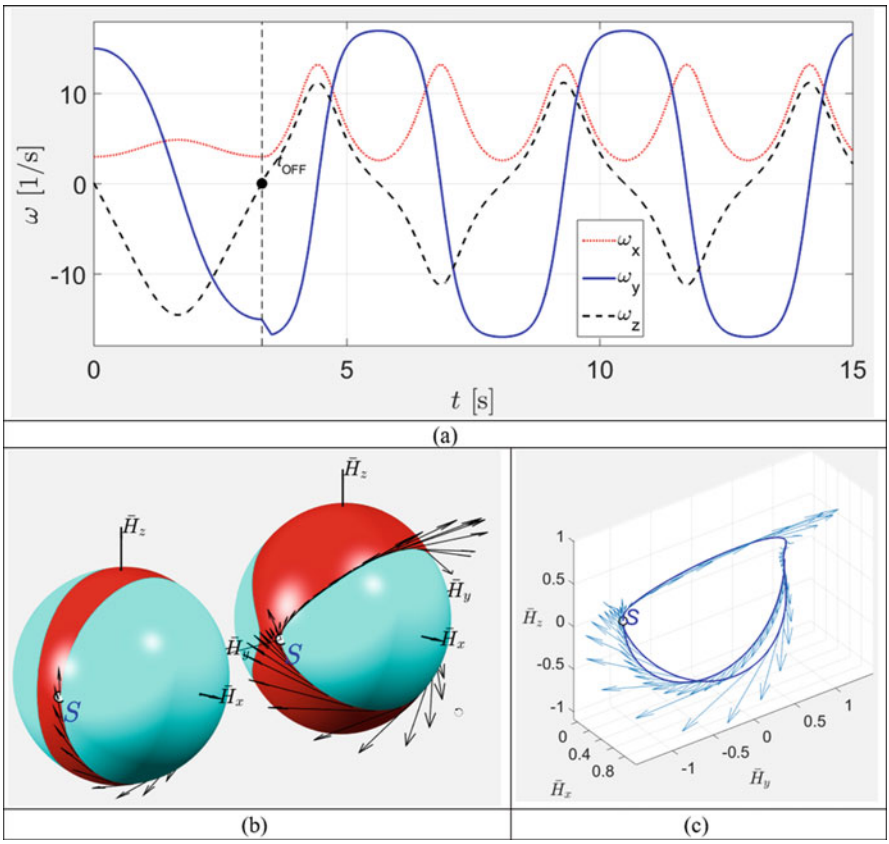


Fig. 5.50 Simulated results for the inertial morphing in Fig. 5.49 applied at $t = 3.32$ s: (**a**) time histories of the angular velocity components; (**b**) AMSs (blue) and KEEs (red) shown before (on the left) and after (on the right) inertial morphing was initiated at $t = 3.32$ s; (**b**) continuous hodograph of the non-dimensional angular momentum vector

section, a rigorous method of selection of the new polhode and timing for the IM will be presented and illustrated with particular examples.

5.11.2 *Formulation of the Conceptual Solution for De-tumbling of the Spacecraft, Using "Installing into Polhode" via "Polhode-to-Polhode" Transfer*

Figure 5.54 shows a conceptual of the "installing into polhode" method with "polhode-to-polhode transfer, with the conjugate transition point shown as white dot. It is important that this can be seen as transition between polhodes, associated

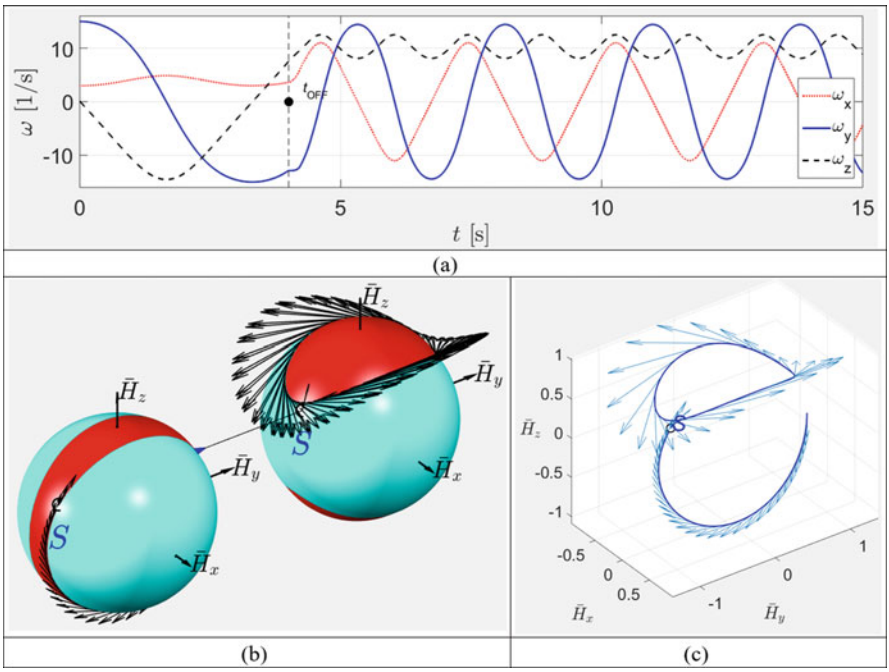


Fig. 5.51 Simulated results for the inertial morphing in Fig. 5.49 applied at $t = 4$ s: (**a**) time histories of the angular velocity components; (**b**) AMSs (blue) and KEEs (red) shown before (on the left) and after (on the right) inertial morphing was initiated at $t = 4$ s; (**c**) continuous hodograph of the non-dimensional angular momentum vector for the case of inertial morphing, initiated at $t = 4$ s

with different poles, being z and y poles in this example. As a particular case, the second polhode can be selected to be a separatrix. And this case is presented in the next section.

5.11.3 *Detailed Example on "Installing into Polhode" via "Polhode-to-Polhode" Transfer (Fig. 5.55)*

Figure 5.56(a) shows non-dimensional angular momentum sphere with two separatrices and sets of representative polhodes for the wide range initial conditions. It also shows, as a blue bold line, a specific polhode (or hodograph of the $\overline{\mathbf{H}}$ vector) for the Phase-1 conditions: $I_{xx} = 2$, $I_{yy} = 3$, $I_{zz} = 4$, $\omega_{x,i} = 0.4$, $\omega_{y,i} = 1$, $\omega_{z,i} = 0.8$.

If the spacecraft possesses inertial morphing capabilities, then the "switch" to any new inertial properties can be simulated and illustrated graphically. Let us assume, for illustration purposes, that the new principal moments of inertia are $I_{xx} = 3.5$, $I_{yy} = 3$, $I_{zz} = 4$. Then, for the Phase-2, its own non-dimensional angular momentum

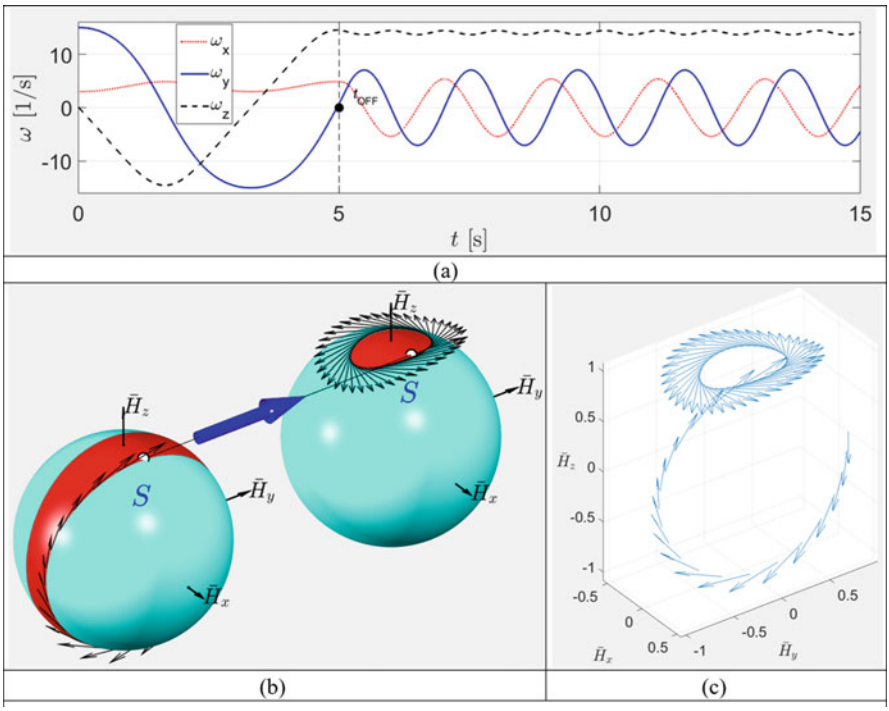


Fig. 5.52 Simulated results for the inertial morphing in Fig. 5.49 applied at $t = 5$ s: (**a**) time histories of the angular velocity components; (**b**) AMSs (blue) and KEEs (red) shown before (on the left) and after (on the right) inertial morphing was initiated at $t = 5$ s; (**c**) continuous hodograph of the non-dimensional angular momentum vector for the case of inertial morphing, initiated at $t = 5$ s

sphere with two separatrices and sets of representative polhodes (for the wide range initial conditions) can be also produced (see Fig. 5.56b). Morphing can be applied at any stage during the execution of Phase-1. For certainty, let us also assume that the morphing is rapidly applied at $t = 21.5$ s instant. Then, the new corresponding angular velocities of the spacecraft could be calculated, using Eqs. (5.36).

5.11.4 Control Method of Installing into Separatrix Using Inertial Morphing: Geometric Interpretation

Using torque-free case Euler's Eq. (5.3) and employing numerical simulator of the rotational motion of the morphing spacecraft [3], let us produce various sets of the feasible polhodes. Figure 5.57 presents three contrast cases, where $I_{xx} < I_{yy} < I_{zz}$ and intermediate value of I_{yy} is (a) close to I_{xx}, (b) almost equally distant from I_{xx} and I_{zz} and (c) I_{yy} close to I_{zz}. These three cases are characterised with the following

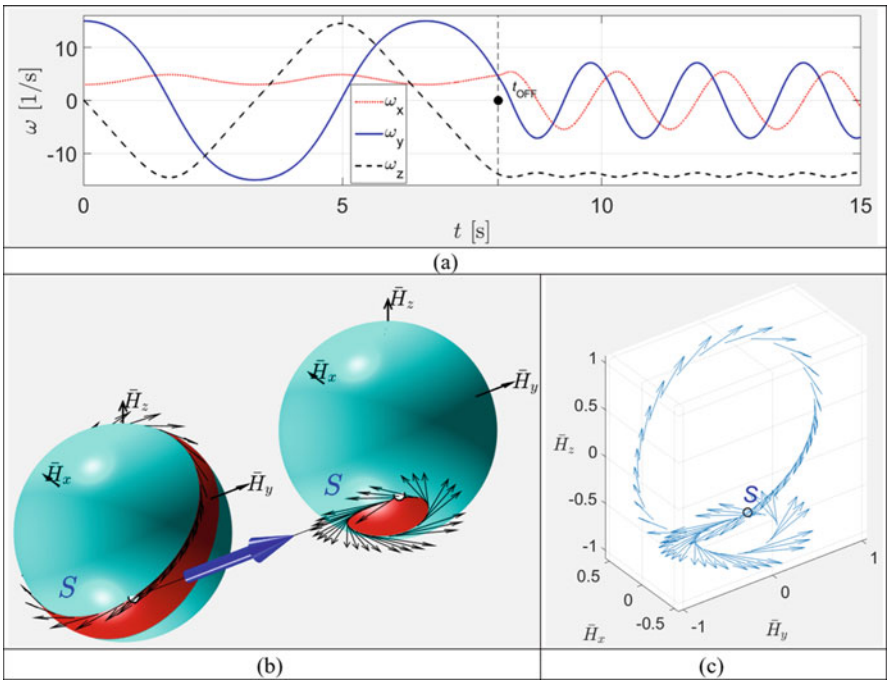


Fig. 5.53 Simulated results for the inertial morphing in Fig. 5.49 applied at $t = 8$ s: (**a**) time histories of the angular velocity components; (**b**) AMSs (blue) and KEEs (red) shown before (on the left) and after (on the right) inertial morphing was initiated at $t = 8$ s; (**c**) continuous hodograph of the non-dimensional angular momentum vector

non-dimensional ratios between the principal moments of inertia I_{xx}: I_{yy}: I_{zz}: (a) 1:1.02:1.26; (b) 1:1.144:1.26; (c) 1:1.22:1.26.

It is quite clear that polhodes can be grouped, based on their "association" with the y (also coinciding with $\overline{H}_x$) and z (also coinciding with $\overline{H}_z$) axes, both being stable axes of rotation. Figure 5.57 distinguishes with bold black lines transition polhodes (called *separatrices*) between two areas of stable periodic rotations about the x and z axes. Note that separatrices pass through the y (coinciding with $\overline{H}_y$) – axis of intermediate inertia.

The separatrices, shown in Fig. 5.57, are critical for the explanation of the "Garriott's-Dzhanibekov's effect" and "tennis racket theorem" [3], during which the object being spun about their intermediate axis y is performing periodic flips, seen as the periodic change of its attitude in the global coordinate system by 180°. As it can be observed from the Fig. 5.57, during the Garriott's-Dzhanibekov's effect classical demonstration, for example, using T-handle [14], the system is usually provided with predominant rotation about its axis, coinciding with the intermediate moment of inertia axis. Using the geometric interpretation, it can be said that at the

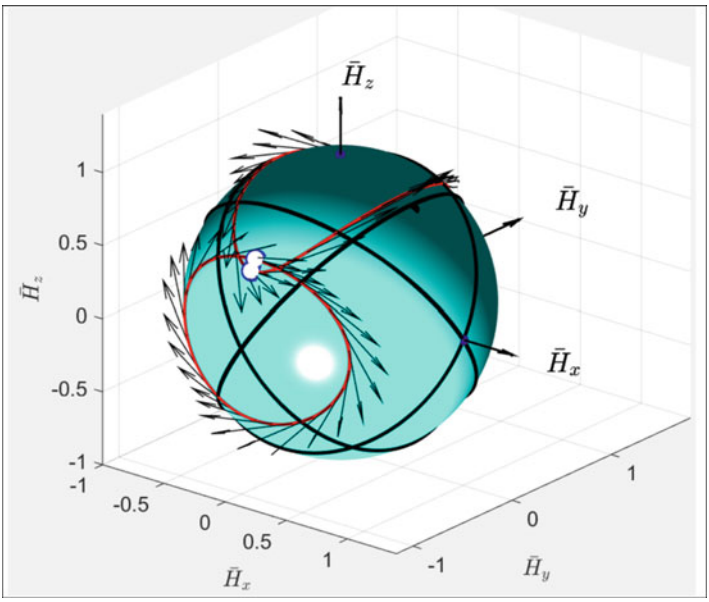


Fig. 5.54 Conceptual demonstration of the "installing into polhode" method with "polhode-to-polhode transfer, with the conjugate transition point shown as white dot

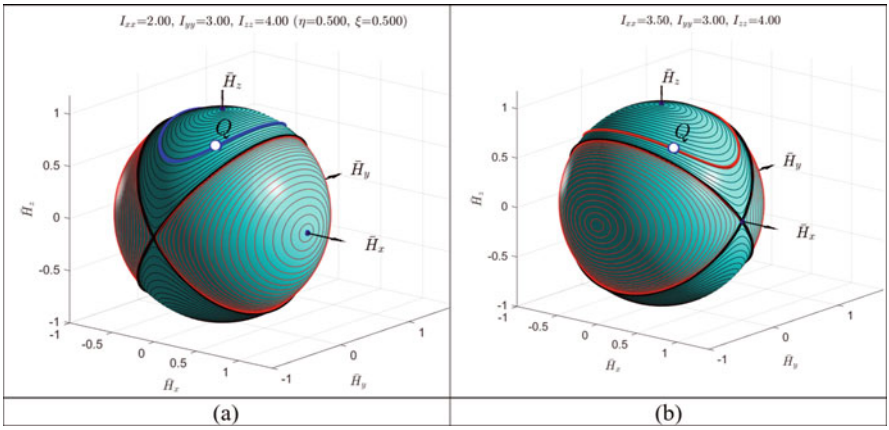


Fig. 5.55 Non-dimensional angular momentum spheres with polhodes and separatrices and truncated specific hodographs for (**a**) Phase-1 (before inertial morphing) conditions – $I_{xx} = 2$, $I_{yy} = 3$, $I_{zz} = 4$, $\omega_{x,i} = 0.4$, $\omega_{y,i} = 1$, $\omega_{z,i} = 0.8$; specific hodograph shown with blue line – and (**b**) Phase-2 (after inertial morphing) conditions: $I_{xx} = 3.5$, $I_{yy} = 3$, $I_{zz} = 4$, $\omega_{x,t_Q} = 0.7133$, $\omega_{y,t_Q} = -0.7318$, $\omega_{z,t_Q} = 0.9016$, $t_Q = 21.5\ s$; hodograph shown with red line

initial instant, the system is "inserted" in one of the saddle points on the y-axis (i.e. point #3 or #4 in Fig. 5.13 notations), where separatrices intersect.

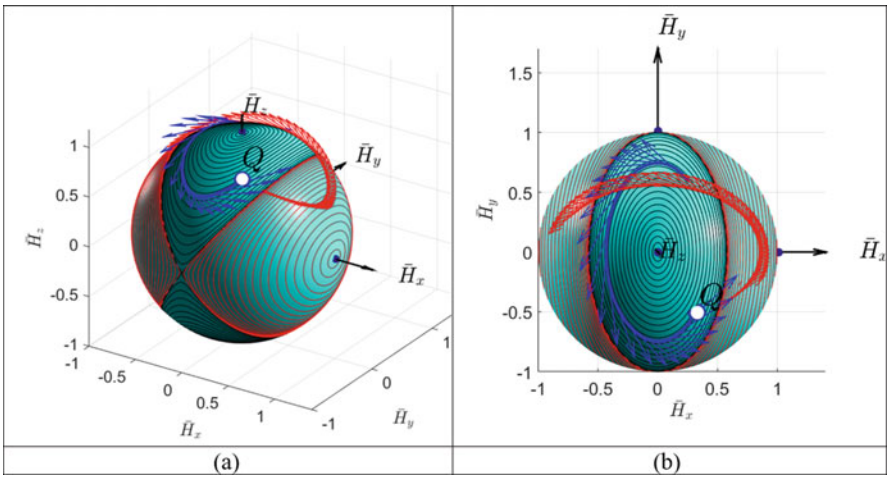


Fig. 5.56 Illustration of the transition between Phase-1 and Phase-2 of the inertial morphing of the system: (**a**) side 3D view; (**b**) z-axis 2D view

Therefore, any small applied disturbance leads to the trigger of the unstable motion along one of the polhodes, standing very close to one of the separatrices, shown in bold black in Fig. 5.57.

For completeness, we should compare the shapes of the polhodes, as seen from various points of view. For this purpose, we take the AMS with plotted polhodes for the case shown in Fig. 5.5c (I_x: I_y: $I_z = 2.5$: 3.05: 3.15). Then we show in Fig. 5.57 (d)–(f) three orthogonal projections of the AMS for case (c).

Polhode projections on the $\left(\overline{H}_y, \overline{H}_z\right)$ plane (shown in Fig. 5.57d) are represented by *ellipses*. In the illustrated case, these ellipses have small eccentricity and resemble circles. However, strictly speaking, they are ellipses.

Polhode projections on the $\left(\overline{H}_x, \overline{H}_z\right)$ plane (shown in Fig. 5.57e) are represented by *hyperbolas*. However, projections of the separatrices are given by straight lines, and we will prove this shortly. These projections of the separatrices along the intermediate axis view are the key for the proposed method.

Polhode projections on the $\left(\overline{H}_x, \overline{H}_y\right)$ plane (shown in Fig. 5.57f) are represented by *ellipses*. For the particular set of selected principal moments of inertia in the demo case, these ellipses have high eccentricity.

Comparison of the polhodes in Fig. 5.57a, b, c, viewed from the intermediate axis direction y shows that the tilt of the separatrices varies within substantial range and is a function of the relative location of the intermediate moment of inertia I_{yy} between the minimum value of the principal moments of inertia (I_{xx}) and the maximum value of the principal moments of inertia (I_{zz}). Next, we establish analytical relations for these parameters.

In this current study we consider various cases of the inertial morphing, during which each of the initially assigned body axes, x, y and z could become an

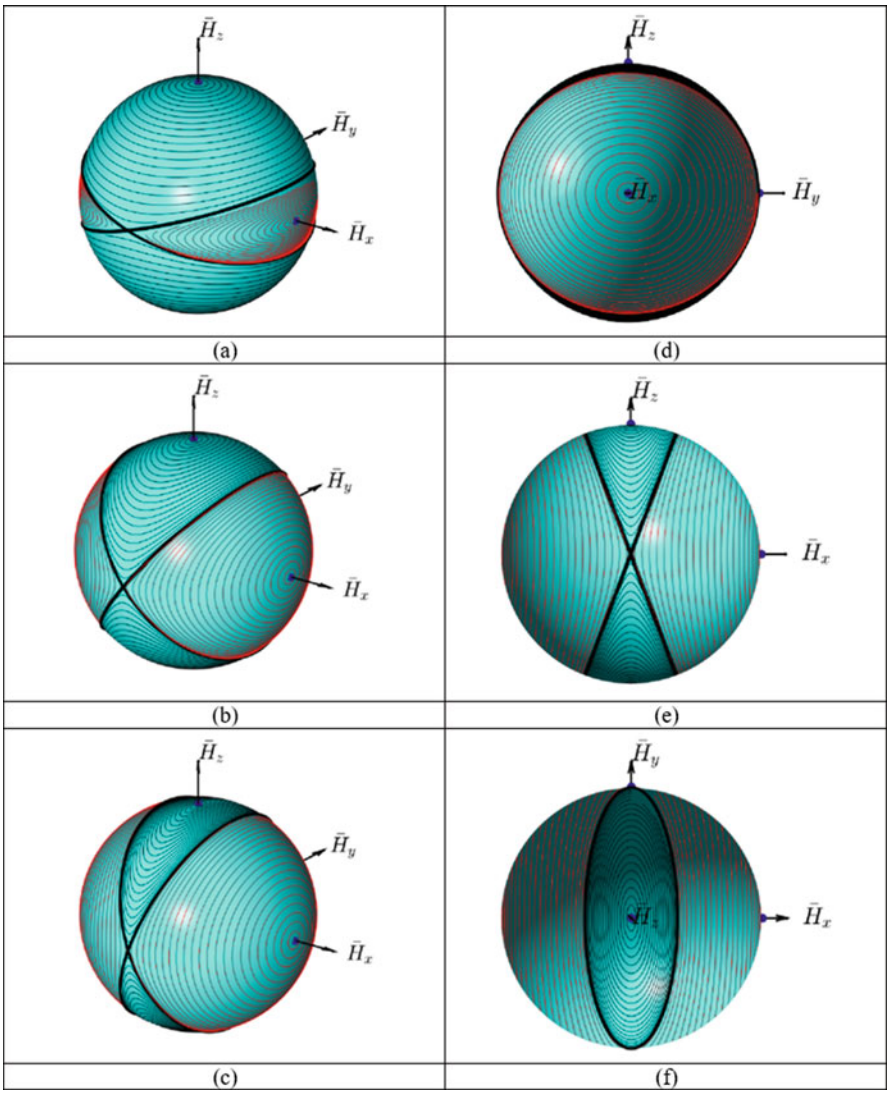


Fig. 5.57 Feasible polhodes for three contrast cases, characterised with the same minimum and maximum values of principal moments of inertia but different intermediate values (hence, ratios I_{xx}: I_{yy}: I_{zz}): (**a**) $I_{xx} = 2.5$; $I_{yy} = 2.55$; $I_{zz} = 3.15$, (**b**) $I_{xx} = 2.5$; $I_{yy} = 2.86$; $I_{zx} = 3.15$, (**c**) $I_{xx} = 2.5$; $I_{yy} = 3.05$; $I_{zz} = 3.15$ (all in kg $\times$ m^2); (**d**) x-view of Case (c); (**e**) minus y-view of Case (c); (**f**) z-view of Case (c)

intermediate axis. However, for derivation of some useful analytical relationships, let us assume first that:

$$I_{xx} < I_{yy} < I_{zz} \tag{5.46}$$

This assumption would also mean that:

$$a_x < a_y < a_z \tag{5.47}$$

If, in view of these assumptions, we divide Eq. (5.23) by a_y^2 and subtract resultant equation from Eq. (5.30), we can eliminate $\overline{H}_y$, which leads us to a *very* important relationship, describing slopes of the projections of the separatrices on the $\left(\overline{H}_x,\ \overline{H}_z\right)$ plane, being normal to the intermediate axis:

$$\frac{\overline{H}_x}{\overline{H}_z} = \sqrt{\frac{a_x^2\left(a_z^2 - a_y^2\right)}{a_z^2\left(a_y^2 - a_x^2\right)}} = \sqrt{\frac{1 - a_{yz}^2}{a_{yx}^2 - 1}}; \qquad \left(\text{only valid if } I_{xx} < I_{yy} < I_{zz}\right) \tag{5.48}$$

where $a_{yz} = a_y/a_z$ and $a_{yx} = a_y/a_x$.

From Eq. (5.31) we can conclude that:

$$\frac{I_{xx}}{a_x^2} = \frac{I_{yy}}{a_y^2} = \frac{I_{zz}}{a_z^2} = const \tag{5.49}$$

In view of this, we can present Eq.(5.48) in an alternative format:

$$\frac{\overline{H}_x}{\overline{H}_z} = \sqrt{\frac{I_x\left(I_z - I_y\right)}{I_z\left(I_y - I_x\right)}} \tag{5.50}$$

Let us call the α-angle from the vertical direction $\overline{H}_z$ to the plane of separatrix a tilt angle. Then:

$$\alpha = \arctan\left(\frac{\overline{H}_x}{\overline{H}_z}\right); \qquad \left(\text{only valid if } I_{xx} < I_{yy} < I_{zz}\right) \tag{5.51}$$

As an example, we produce a particular plot of the values of the α-angle during variation of I_{yy} between given values of $I_{xx} = 2.4$ and $I_{zz} = 3.15$ to be used in the upcoming simulation cases (Fig. 5.58).

However, in this work we are dedicated to the non-dimensional construction of the analytical tools. With this in mind, let us represent results in Eqs. (5.44 and 5.45) in non-dimensional format. For this purpose, let us employ a non-dimensional variable ξ introduced with Eq. (5.10), enabling to express the value of the intermediate moment of inertia via Eq. (5.11), reproduced below:

$$I_{yy} = I_{xx}\left(1 - \xi\right) + I_{zz}\xi \tag{5.52}$$

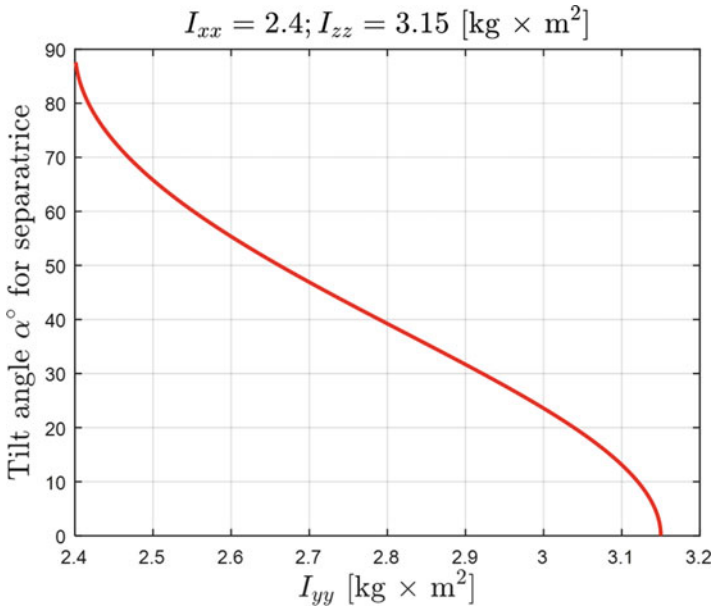


Fig. 5.58 Variation of the α-angle for separatrix (measured from vertical) due to the variation in I_{yy}, based on Eqs. 5.50 and 5.51

Zero value of ξ would now correspond to I_{xx} and unit value of ξ would correspond to I_{zz} and any intermediate value of I_{yy} could be expressed via $0 < \xi < 1$.

With these notations, Eqs. (5.44 and 5.45) can be now expressed as a function of ξ:

$$\alpha = \arctan\left(\sqrt{\left(\frac{I_{xx}}{I_{zz}}\right)\left(\frac{1}{\xi} - 1\right)}\right); \quad \left(\text{only valid if } \; I_{xx} < I_{yy} < I_{zz} \text{ and } 0 < \xi < 1\right) \tag{5.53}$$

Equation (5.47) can be plotted as a 3D surface, corresponding to the α, being a function of two non-dimensional variables: (1) variable ξ, describing relative "location" of I_{yy} between I_{xx} and I_{zz}, and (2) ratio ($\eta = I_{xx}/I_{zz}$), which is subject to the inequalities $0 < I_{xx}/I_{zz} < 1$. This 3D surface plot is shown in Fig. 5.59, together with superimposed plot from Fig. 5.58 and also level lines, corresponding to the angles $\alpha = [10°, 20°, 30°, 40°, 50°, 60°, 70°, 80°]$, i.e. with equal 10° increments.

As mentioned before, this study is not restricted to the cases, stipulated by Eq. (5.40), as in our work intermediate axes may vary during the simulated scenarios. In the cases of the change of the intermediate axes, for example, when intermediate axes change from y to x, (i.e. when after morphing is applied, the new moments

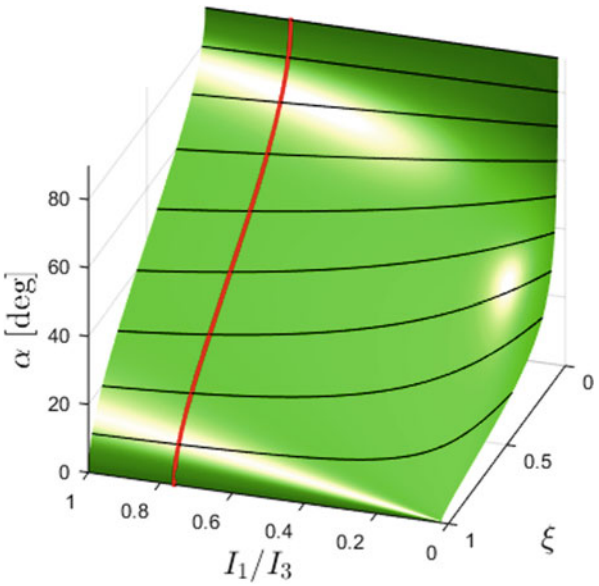


Fig. 5.59 Variation of the α-angle for separatrix (measured from vertical) due to the variations in ξ and ratio (I_{xx}/I_{zz})

of inertia would satisfy $I_{yy} < I_{xx} < I_{zz}$), Eqs. (5.44 and 5.45) must be rewritten as follows:

$$\frac{\overline{H}_y}{\overline{H}_z} = \sqrt{\frac{a_y^2\left(a_z^2 - a_x^2\right)}{a_z^2\left(a_x^2 - a_y^2\right)}} = \sqrt{\frac{1 - a_{xz}^2}{a_{xy}^2 - 1}}; \qquad \left(\text{only valid if } \; I_{yy} < I_{xx} < I_{zz}\right) \tag{5.54}$$

$$\alpha = \arctan\left(\frac{\overline{H}_y}{\overline{H}_z}\right); \qquad \left(\text{only valid if } \; I_{yy} < I_{xx} < I_{zz}\right) \tag{5.55}$$

Similarly, when intermediate axes change from y to z (i.e. when after morphing is applied, the new moments of inertia satisfy $I_{xx} < I_{zz} < I_{yy}$), Eqs. (5.50 and 5.51) must be rewritten as follows:

$$\frac{\overline{H}_x}{\overline{H}_y} = \sqrt{\frac{a_x^2\left(a_y^2 - a_z^2\right)}{a_y^2\left(a_z^2 - a_x^2\right)}} = \sqrt{\frac{1 - a_{zy}^2}{a_{zx}^2 - 1}}; \qquad \left(\text{only valid if } \; I_{xx} < I_{zz} < I_{yy}\right) \tag{5.56}$$

$$\alpha = \arctan\left(\frac{\overline{H}_x}{\overline{H}_y}\right); \qquad \left(\text{only valid if } \; I_{xx} < I_{zz} < I_{yy}\right) \tag{5.57}$$

As this has been mentioned before, if the spinning system is initially provided with a predominant spin about the intermediate axis of inertia, the flipping motion could be initiated. Essentially, this is a very special case of the motion, initiated with particular initial conditions. Geometrically, this would mean that the tip of the angular momentum vector of the system is *initially "inserted"* in one of the saddle points, where separatrices are intersecting. This means that the vector of initial angular momentum $\overrightarrow{\mathbf{H}}_G(0)$ is initially aligned with y-axis! And, as the angular momentum is conserved, the total vector of the angular momentum keeps its initial orientation in the global XYZ coordinate system all the time, allowing AMS to rotate about centre of the mass G in the body-axis system with y-axis periodically coming to the initial alignment with vector $\overrightarrow{\mathbf{H}}_G(t)$. With this, from Fig. 5.57, it is quite obvious that with this particular initial condition, the system has to follow one of the polhodes, located very close to the separatrix.

5.11.5 Control Method of Installing into Separatrix, Using Inertial Morphing: Selection of the IM Parameters and IM Activation Time

Let us consider a task of stabilisation of a spacecraft, which is initially in tumbling motion. Geometric interpretation of the inversion of the spinning body about its intermediate axis also offers an effective solution for this task.

Apart from traditional use of the flipping motion for spacecraft inversion, we are suggesting its another application: it can be used as a "vehicle", pulling the angular momentum vector away from the tumbling polhode and "driving" it along the separatrix to the saddle point, which, after stabilisation at this point, can be called a "parking" point. However, the whole manoeuvre *would require application of the inertial morphing a few (two or more) times*!

First morphing would be required to transfer the system from its initial polhode into separatrix, corresponding to the supplementary case of unstable flipping motion. This is needed to force the $\overline{H}$ vector to get to the saddle point, where two of its components $\overline{H}_x$ and $\overline{H}_z$ are close to zero. Second morphing would be very similar to the one described in the previous section and would be required for stopping the flipping motion of the system.

During flipping process, the system's angular velocity vector "spends" considerable time in the vicinity of the saddle points, where only one component of the angular momentum $\left(\overline{H}_y\right)$ has the most prevailing value, with two other components $\left(\overline{H}_x \text{ and } \overline{H}_z\right)$ having very small values. After leaving the saddle points, the angular momentum vector performs rapid transit to the opposite saddle point, using adjacent

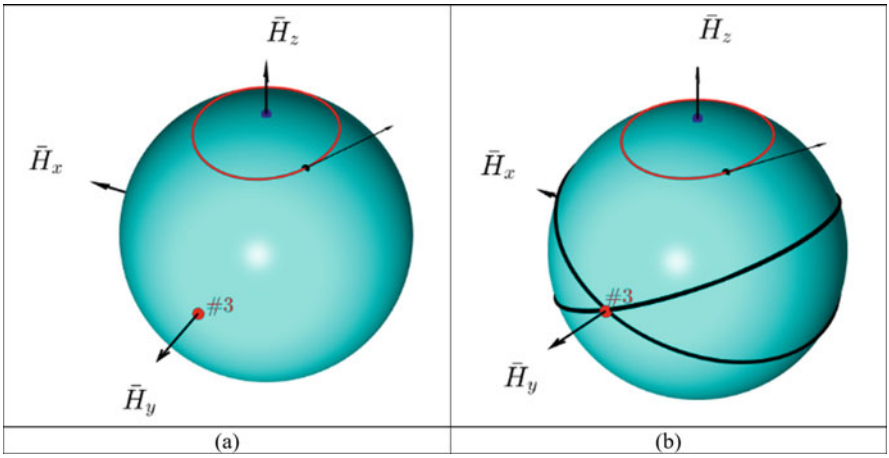


Fig. 5.60 Particular study case: $I_{xx} = 2.4$, $I_{yy} = 2.5$, $I_{zz} = 3.15$ (all in kg $\times$ m^2), $\omega_{x0} = -4.93$, $\omega_{y0} = 2.73$, $\omega_{z0} = 5.75$ (all in rad/s): (**a**) polhode (shown with red line), (**b**) separatrices (shown with black thick lines)

separatrix. During transition stage, all magnitudes of the angular velocity vector have comparable absolute values.

However, in the general case, when initial conditions are not selected in rather unique way, the system performs compound rotation, involving all, x, y and z rotations. One of the particular illustration cases [$I_{xx} = 2.5$, $I_{yy} = 2.4$, $I_{zz} = 3.15$ (all in kg $\times$ m^2), with the initial conditions $\omega_{x,i} = -4.93$, $\omega_{y,i} = 2.73$, $\omega_{z,i} = 5.75$ (all in rad/s)] is presented in Fig. 5.60, where polhode is shown as red line.

For the initiated motion shown, it may be desirable to stabilise spacecraft, reducing compound rotations (with all, ω_x, ω_y, ω_z) involved to a rotation with only y rotation.

This would mean that we need to reduce (shrink) current trajectory, given by the red polhode in Fig. 5.60 to a point, marked as #3. However, this point is only accessible from the polhode, close to the separatrix. Current separatrices, shown in Fig. 5.60b with bold black lines, are not accessible, as the current polhode does not intersect any of the separatrices.

Solving Euler's Eqs.(5.3) for the particular case data, we can determine the maximum projected angle for the point on the polhode, denoted as α_1 in Fig. 5.61. Its value is equal to $|\alpha_1| = 36.1^\circ$.

Using Eqs. (5.60 and 5.61), we can calculate the incline angle of the separatrix, shown in Fig. 5.61 (being a semi-transparent y projection of the AMS in Fig. 5.60b) as α_2. Its value is equal to $|\alpha_2| = 65.8^\circ$.

This prompts a *solution to the task*: in order for the current hodograph of the angular momentum vector trajectory (along polhode) to be switched to the saddle point, or "parking" point (shown as point #3 in Fig. 5.13), we need to aim to get to the separatrix, applying morphing to the system. In other words, we will aim

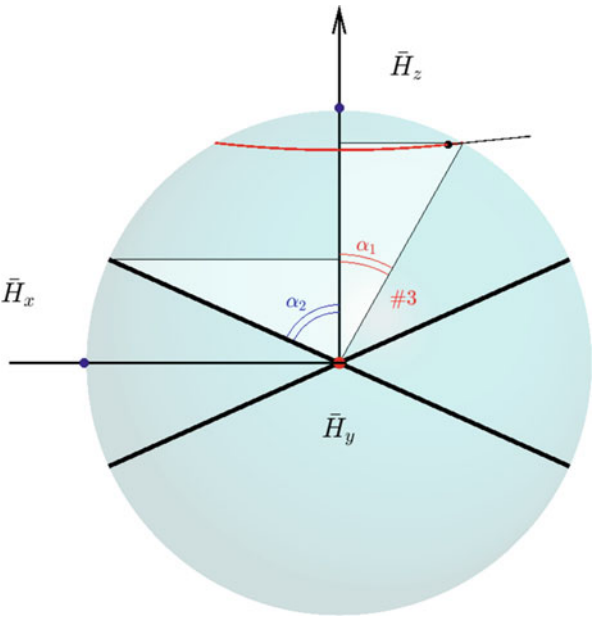


Fig. 5.61 Calculation of the important angles for the illustration study case [$I_{xx} = 2.4$, $I_{yy} = 2.5$, $I_{zz} = 3.15$, $\omega_{x,i} = -4.93$, $\omega_{y,i} = 2.73$, $\omega_{z,i} = 5.75$]: (**a**) α_1, minimum incline angle for the polhode; (**b**) α_2, incline angle for the separatrix plane

to "insert" hodograph of the system into the separatrix. In order to achieve this objective, parameters of the morphing should be carefully selected to ensure that the associated incline angle $|\alpha_{new}|$for the new inertial properties is increased and is exactly equal to $|\alpha_1| = 36.1^{\circ}$.

If we keep the values of I_{xx} and I_{zz} unchanged and if we keep the value of I_{yy} between I_{xx} and I_{zz}, then we can get from Eqs. (5.50 and 5.51) solution $I_{yy} = 2.84$ for this study case. Graphical solution of the task is illustrated in Fig. 5.62.

Moreover, success of implementation of this solution (enabling to get to the hodograph of $\overline{H}$ vector to the parking point #3) is a subject to the right moment of activation of this morphing. In our case, we select the instant, when the angular momentum vector is in the $\left(\overline{H}_x, \overline{H}_z\right)$ plane.

This concludes stage-1 (out of two) for the whole manoeuvre. When hodograph is on the orbit, which is close to the separatrix, "flipping" motion is activated. If we wish to stop flipping, another morphing is needed to convert unstable motion into stable. This can be done via application of the morphing, assigning I_{yy} a new value, which would be the smallest or largest value out of all moments of inertia. Therefore, there could be two classes of solutions to stabilisation of the second stage.

For demonstration purpose, we select solution-2 [5], assigning I_{yy} a value of 3.5, which would make I_{zz} a new intermediate value.

Similar to the stage-1 morphing-1, time of activation of the morphing-2 is also very important. We apply morphing at the instant, when hodograph reaches the desired one (out of two) parking point on the axis y.

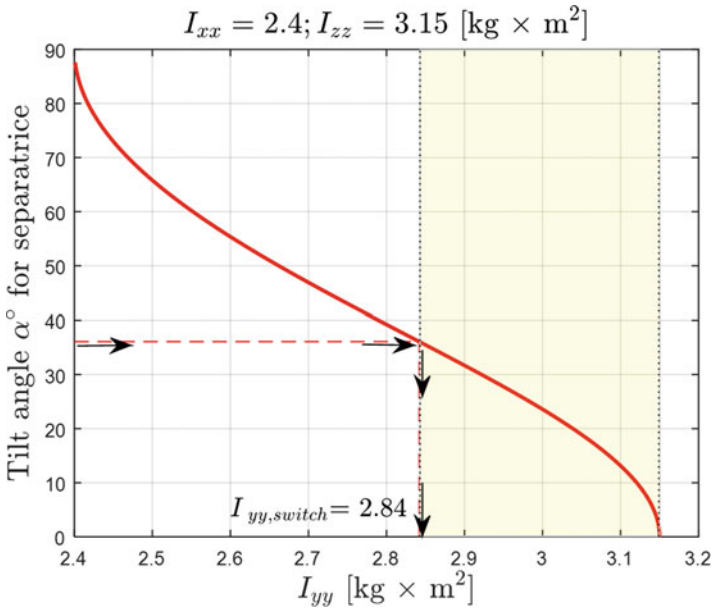


Fig. 5.62 Determination of the *new morphing value of* the principle intermediate moment of inertia I_{yy} for stabilising spin around the y-axis

5.11.6 *Example of the "Flipping"-Assisted Stabilisation (De-tumbling) of the Tumbling Spacecraft, Using Inertial Morphing*

We demonstrate in detail stabilisation of the tumbling spacecraft, using the proposed method. The key to understanding this method is with the graphical interpretation of the attitude dynamics.

Let us assume that the spacecraft with given initial values of the moments of inertia ($I_{xx} = 2.5$, $I_{yy} = 2.4$, $I_{zz} = 3.15$) is originally in arbitrary free rotation, involving all three angular velocities, as shown in Fig. 5.63a.

This motion can be visualised, using intersecting kinetic energy ellipsoid and angular momentum sphere, as shown in Fig. 5.63b. The genuine length $\overline{H}$ vector cannot be used for visualisation as its length is equal to 1, and it would not be seen at any instant, as it would be completely hidden by the embracing angular momentum sphere with unit radius. Therefore, for visualisation of the instantaneous orientation of $\overline{H}$ in Fig. 5.63b, we use a black line with a dot at its end and extruding beyond the surface of the sphere. The godograph of the $\overline{H}$ vector is shown with a black line on the surface of the angular momentum sphere, coming strictly along the intersection between the AMS and KEE.

Let us set a task to control rotations of the system, via the changes of the values of its principal moments of inertia.

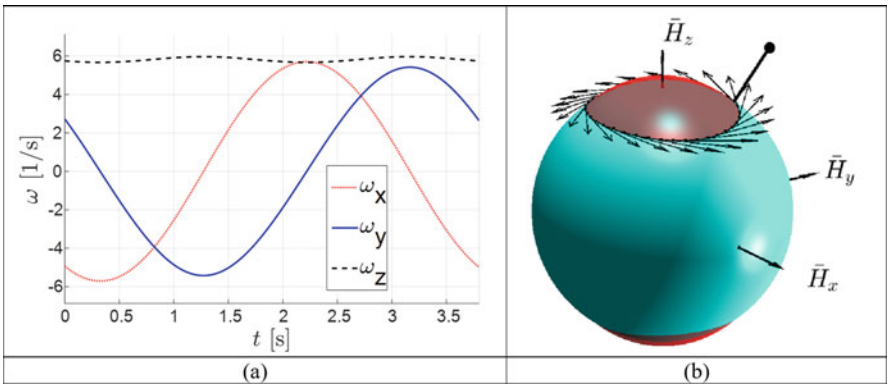


Fig. 5.63 Illustration of the spacecraft tumbling motion: (**a**) time history of ω_x, ω_y, ω_z – components of its angular velocity vector $\vec{\boldsymbol{\omega}}$; (**b**) graphical interpretation of the motion, using KEE and AMS

In each case of using flipping mode for escaping from the closed smooth polhode, we need to apply change to moments of inertia, which could be calculated based on the parameters of the targeted separatrix, using Eqs. (5.50), (5.51) or (5.53).

An example of complete set of morphings, stabilising the system about the y-axis, being initially in the tumbling state as per Fig. 5.64, is presented in Figs. 5.65, 5.66, 5.67 and 5.68. Figure 5.65 explains the sequence and nature of inertial changes. Figure 5.66 gives consecutive snapshots from the simulation process, illustrating changes of the kinetic energy ellipsoid and polhodes – resultant feasible trajectories for the angular momentum vector. Figure 5.68 shows time histories for the angular momentum components and also values of the kinetic energy semi-major axes.

Figure 5.65 explains the sequence and nature of the inertial changes, deliberately applied to the system. Figure 5.66 gives consecutive snapshots from the simulation process, illustrating changes of the kinetic energy ellipsoid and polhodes – resultant feasible trajectories for the angular momentum vector.

It is interesting to observe that at the initial stage of the motion of the system, its $\boldsymbol{e}_2$ body axis ort is "drawing" a pretty spread trajectory on the "dome" (Fig. 5.67a). However, after stabilisation is completed, this trajectory is essentially reduced to the point (Fig. 5.67b). Also, at the last stage of the simulation, trajectories for $\boldsymbol{e}_1$ and $\boldsymbol{e}_3$ are very close to the equatorial plane, which confirms that the stabilised motion is close to the rotation of the body along the direction of the angular momentum vector. The feature of the example is the final direction of the y-body axis system, selected for stabilisation in this example, which is opposite to the direction of **H**. If the goal of stabilisation was to have them both aligned, then third stage should be activated at instant close to 15 s, as evidenced by Fig. 5.68.

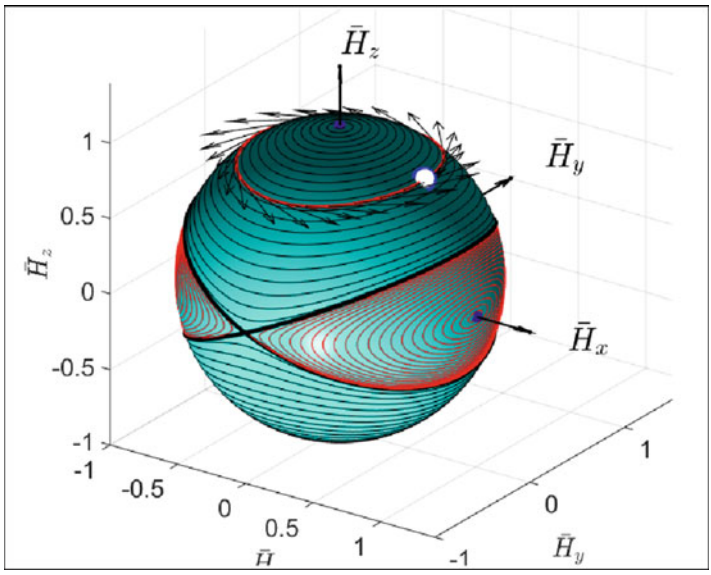


Fig. 5.64 Sets of feasible polhodes for the system with $I_{xx} = 2.5$, $I_{yy} = 2.4$, $I_{zz} = 3.15$

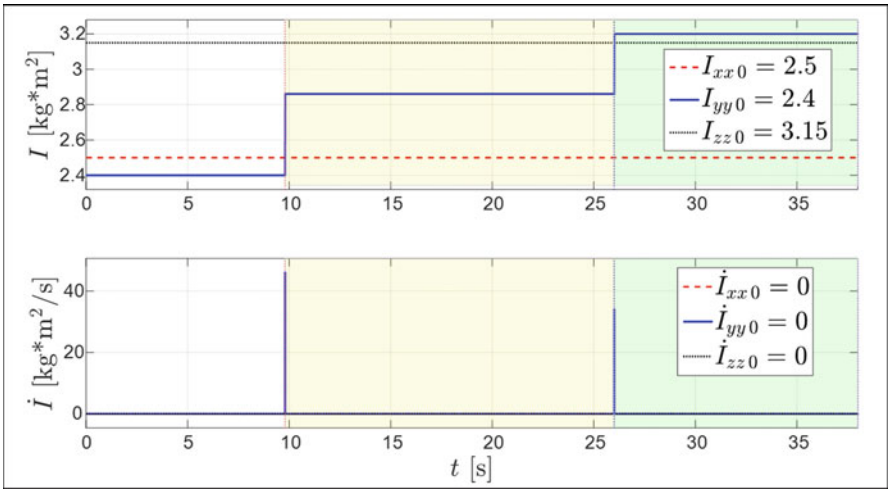


Fig. 5.65 Two-stage stabilisation (de-tumbling) of the tumbling spacecraft via morphing: time history of the I_{xx}, I_{yy}, I_{zz}

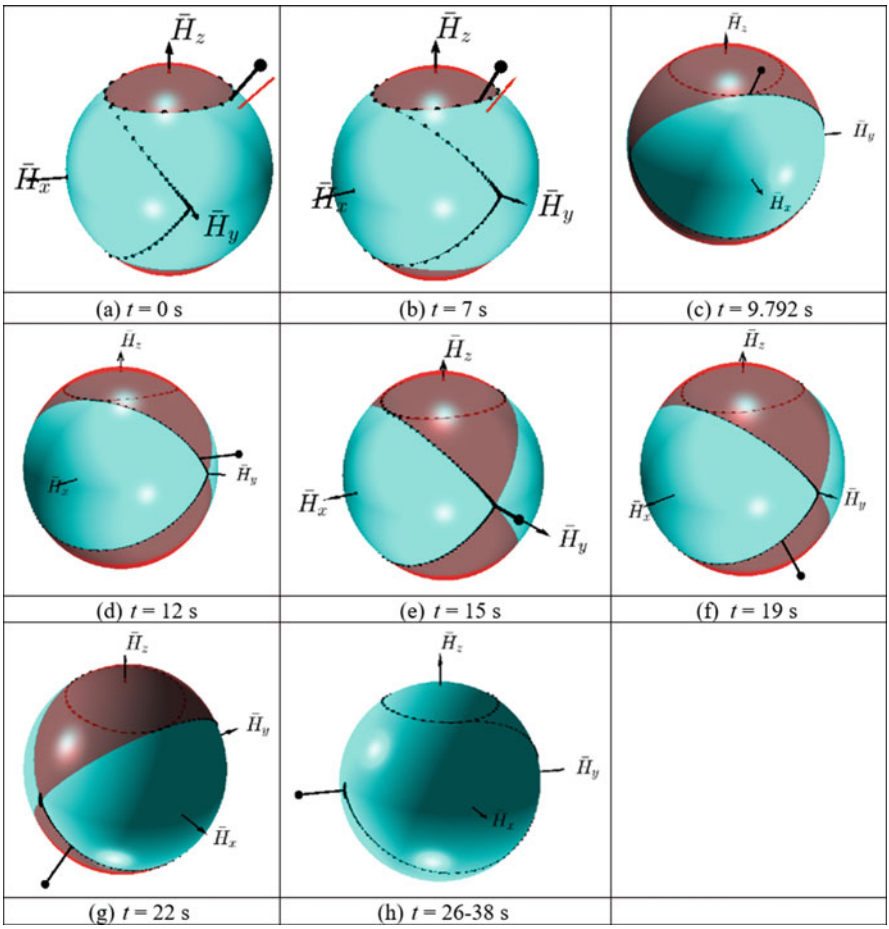


Fig. 5.66 Critical instances of spacecraft stabilisation: (**a**) start of the simulation; (**b**) initially, hodograph is "circling" around z-axis; (**c**) stage-1 ends, and transition to "flipping" is initiated, $t = 9.792$ s; (**d**) approach to the saddle point-1, $t = 12$ s; (**e**) near the saddle point-1 (possible "parking" or stabilisation point), $t = 15$ s; (**f**) passing saddle point-1, $t = 19$ s; (**g**) approach to the saddle point-2, $t = 22$ s; (**h**) stage-2 ends and third stage starts at $t = 26$ s, parking at the stable "saddle point-2 attractor" is activated and de-tumbling (stabilisation) is completed

5.11.7 Reversing Vector of Angular Momentum on the Separatrix, Installing Its Godograph into the Same Separatrix

For any system with $I_{\min}$, I_{int} and $I_{\max}$, regardless of the values of the initial velocities ω_x, ω_y and ω_z, there is a set of two polhodes, which plane has an angle α_{prime} with the $\overline{H}_{\max}$ axis. For example, in case of $I_{xx} < I_{yy} < I_{zz}$, the α-angle

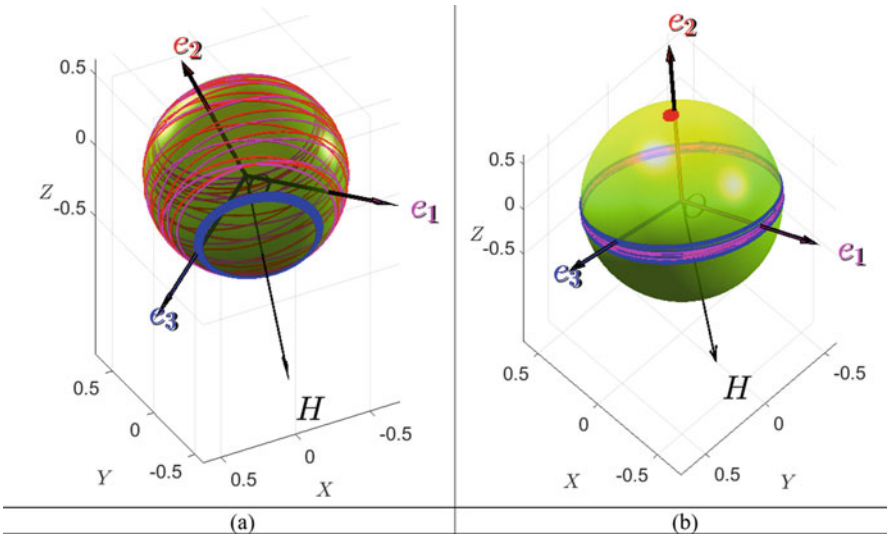


Fig. 5.67 "Balls of wool" for (**a**) the first stage of spacecraft motion with tumbling/coning ($t = 0$ s); (**b**) last stage of stabilisation of the spacecraft ($t = 0.6842$ s) with $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ intersection lines with the "dome"

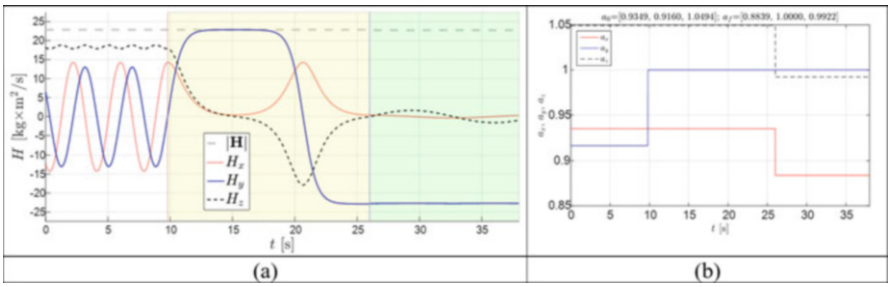


Fig. 5.68 Time history of (**a**) H_{total}, H_x, H_y, H_z; (**b**) a_x, a_y and a_z during two-stage stabilisation (de-tumbling) of the tumbling spacecraft via inertial morphing

can be calculated using Eq. (5.51) or Eq. (5.53). This can be illustrated with the pair of collocated surfaces, AMS and KEE, and for the particular illustration case, $I_{xx} < I_{yy} < I_{zz}$ these two surfaces, which we will call *prime* surfaces, are shown in Fig. 5.69c. However, there is another pair of two surfaces, AMS and KEE, which we will call *dual* surfaces, with $I_{zz} < I_{yy} < I_{xx}$, which have the same intermediate axis and *the same separatrix.*

The "prime" and "dual" surfaces are shown in Fig. 5.69 in each row for the x, y and z flipping motions.

For the dual surfaces in the illustration example, the angle α_{dual} should be measured from the separatrix plane to the axis with maximum moment of inertia, and it is a complimentary angle to α_{prime}:

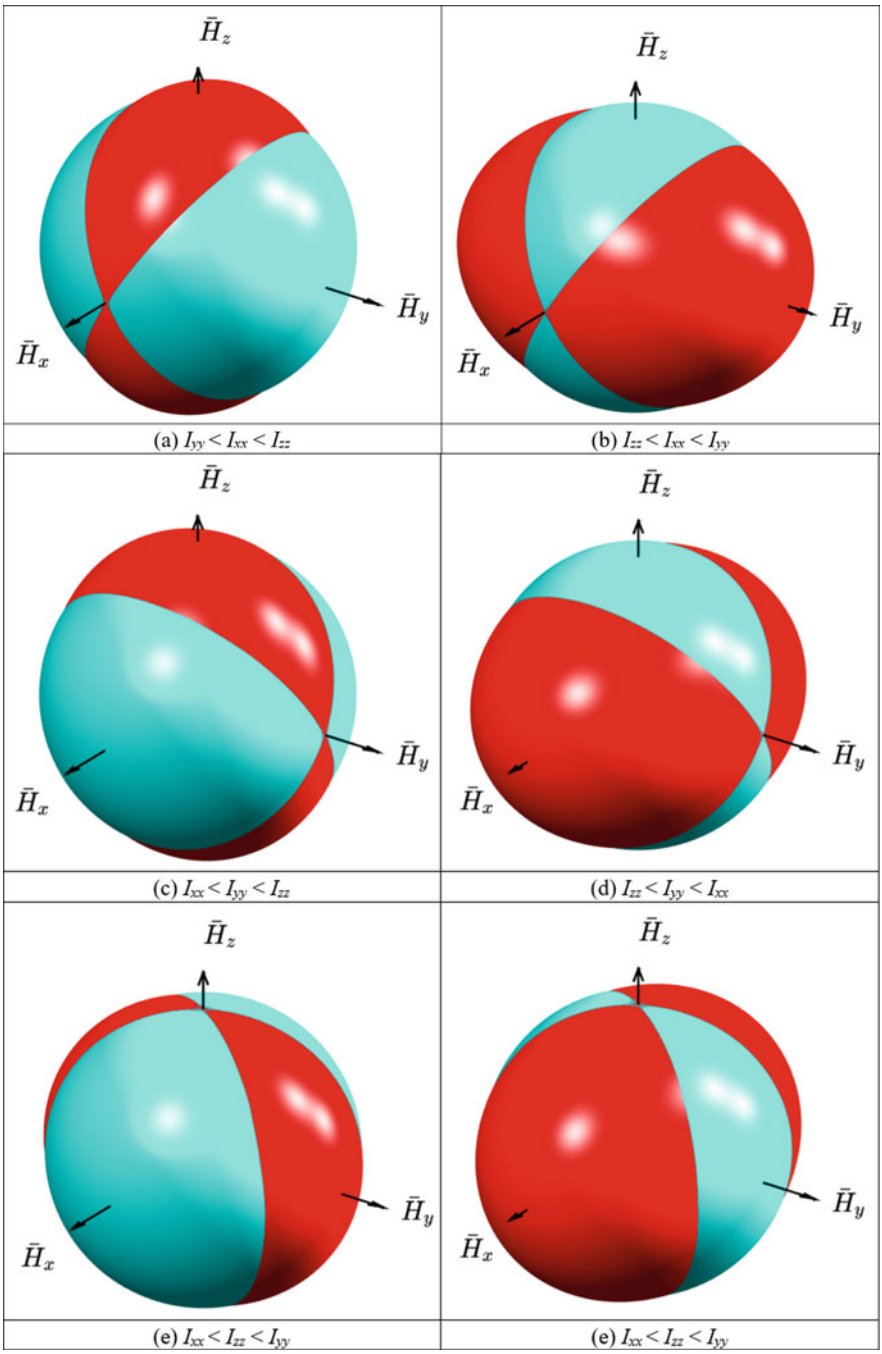


Fig. 5.69 Prime combinations of KEE and AMS for "Garriott's-Dzhanibekov's effect" flipping motion and their dual counterpart combinations: (**a**)–(**b**) for x being an intermediate axis; (**c**)–(**d**) for y being an intermediate axis; (**e**)–(**f**) for z being an intermediate axis

$$\alpha_{\text{dual}} = 90^{\circ} - \alpha_{\text{prime}} \quad \Rightarrow \quad \tan(\alpha_{\text{dual}}) = \tan^{-1}(\alpha_{\text{prime}}) \tag{5.58}$$

The *dual* surfaces for Fig. 5.69c case are shown in Fig. 5.69d, located in the same row. New morphed moments of inertia for the *dual* case can be calculated as follows: we can first set any new values for the minimum and maximum values of the moments of inertia in a dual: $I_{zz,\,\text{dual}} < I_{xx,\,\text{dual}}$. And, then, the intermediate value can be calculated, using adjusted Eq. (5.53), solved for the value of the intermediate moment of inertia and for the current example, and can be written in the explicit form:

$$I_{yy} = \frac{\left[\left(\tan^{-1}\alpha_{\text{prime}}\right)^2 + 1\right] I_{zz}}{\left(\tan^{=1}\alpha_{\text{prime}}\right)^2 + \left(\frac{I_{zz}}{I_{xx}}\right)} = \frac{\left[\left(\tan^{-1}\alpha_{\text{prime}}\right)^2 + 1\right]}{\left(\tan^{-1}\alpha_{\text{prime}}\right)^2 + \eta} I_{\min}, \text{ where } \eta = \frac{I_{\min}}{I_{\max}} = \frac{I_{zz}}{I_{xx}} \tag{5.59}$$

As a particular numerical simulation example, let us consider the case of the system with initial $I_{xx,i} = 0.3$, $I_{yy,i} = 0.35$, $I_{zz,i} = 0.4$ (all in kg $\times$ m^2), and $\omega_{x,i} = 0.1$, $\omega_{y,i} = 15$, $\omega_{z,i} = 0.1$ (all in rad/s). With these conditions, the system starts classical "Garriott's-Dzhanibekov's" flips, during which the godograph of the non-dimensional angular momentum vector is sliding along the line, being close to the separatrix. Angle of the plane of the separatrix with the "z" axis can be determined, using Eqs. (5.50 and 5.51):

$$\alpha = \text{atan}\ \left(\text{sqrt}\left(0.3 * (0.4 - 0.35) / (0.4 * (0.35 - 0.3))\right)\right) * 180/\pi = 40.8934^{\circ}$$

This angle is related to the KEE, bulging over the AMS along the "z" axis. However, the same separatrix would characterise the "dual" KEE, which is bulging over the AMS along the "x" body axis, corresponding to:

$$I_{xx,f} = 0.4 \text{ and } I_{zz,f} = 0.3 \left(\text{all in kg} \times \text{m}^2\right) \text{ and } \alpha_{\text{dual}} = 90^{\circ} - 40.8934^{\circ} = 49.1066^{\circ} \tag{5.60}$$

Equation (5.50) can now be solved for $I_{yy,f}$ or Eq. (5.59) can be used to obtain $I_{yy,f} = 0.336$.

If during the flipping motion at any stage, let say at $t = 9.5$ s if the values of $I_{xx,i} = 0.3$, $I_{yy,i} = 0.35$, $I_{zz,i} = 0.4$, are morphed to the new values $I_{xx,f} = 0.4$, $I_{yy,f} = 0.336$ and $I_{zz,f} = 0.3$, then the angular momentum vector, instead of continuation of its motion along the separatrix, would start moving along the same separatrix, but in the opposite direction, i.e. "backwards" along the previous path of the tip of the vector $\overline{\mathbf{H}}$, as shown in Fig. 5.70!

Results of the animation of the direct and reverse sliding of the godograph of the angular momentum along the same separatrix are shown in Fig. 5.70. On the snapshots, the floating view angle was used, enabling always to the extended in length vector $\overline{\mathbf{H}}$.

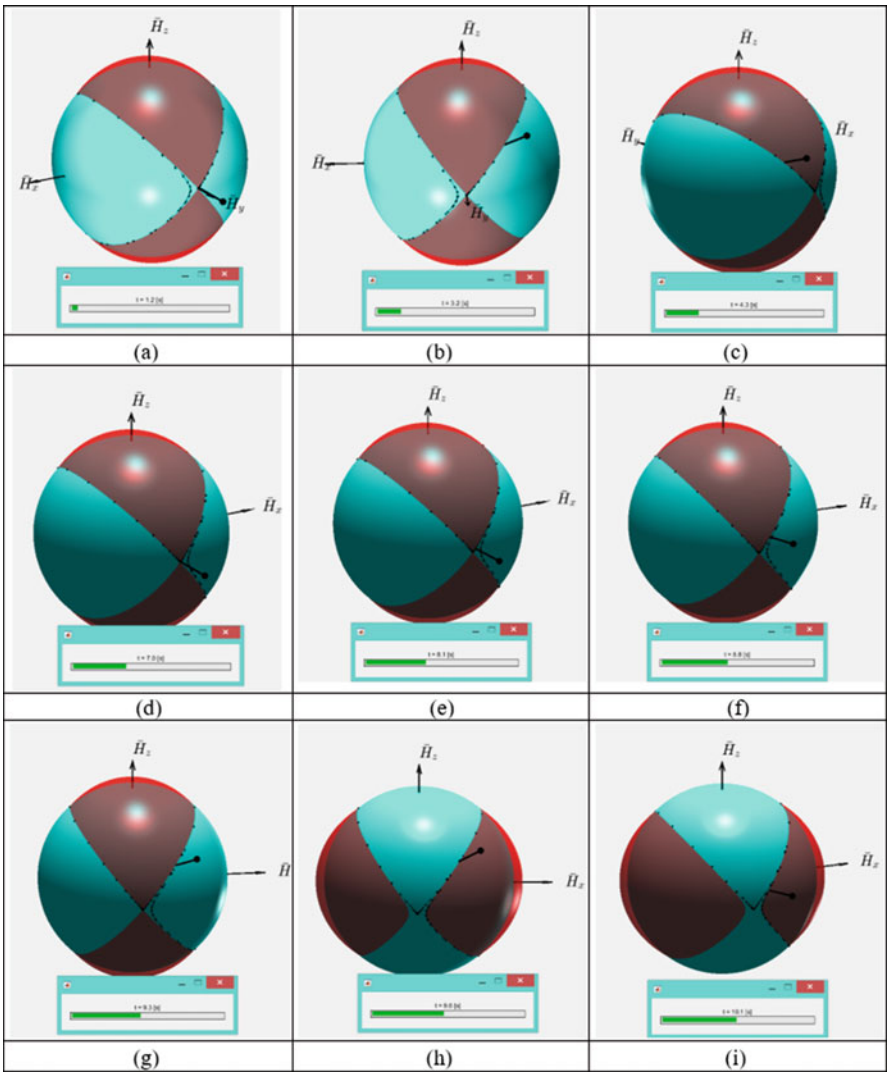


Fig. 5.70 Sliding of the tip of the non-dimensional angular momentum vector along the same separatrix for two different combinations of the moments of inertia: (**a**)–(**g**) direct sliding along "prime" separatrix, (**h**)–(**i**) reverse sliding along coincident "dual" separatrix

In addition to this case, when reversing was activated at the point, being in-between two poles, we also present results for the case, when revering was activated after the non-dimensional angular momentum vector performed full cycle along two semi-separatrices. Results of this simulation case are shown in Fig. 5.71. Figure 5.71b shows slight increase in ω_y immediately after the application of the IM at $t = 13$ s. This is because of the required switch for the I_{yy} value from 0.35 to 0.366,

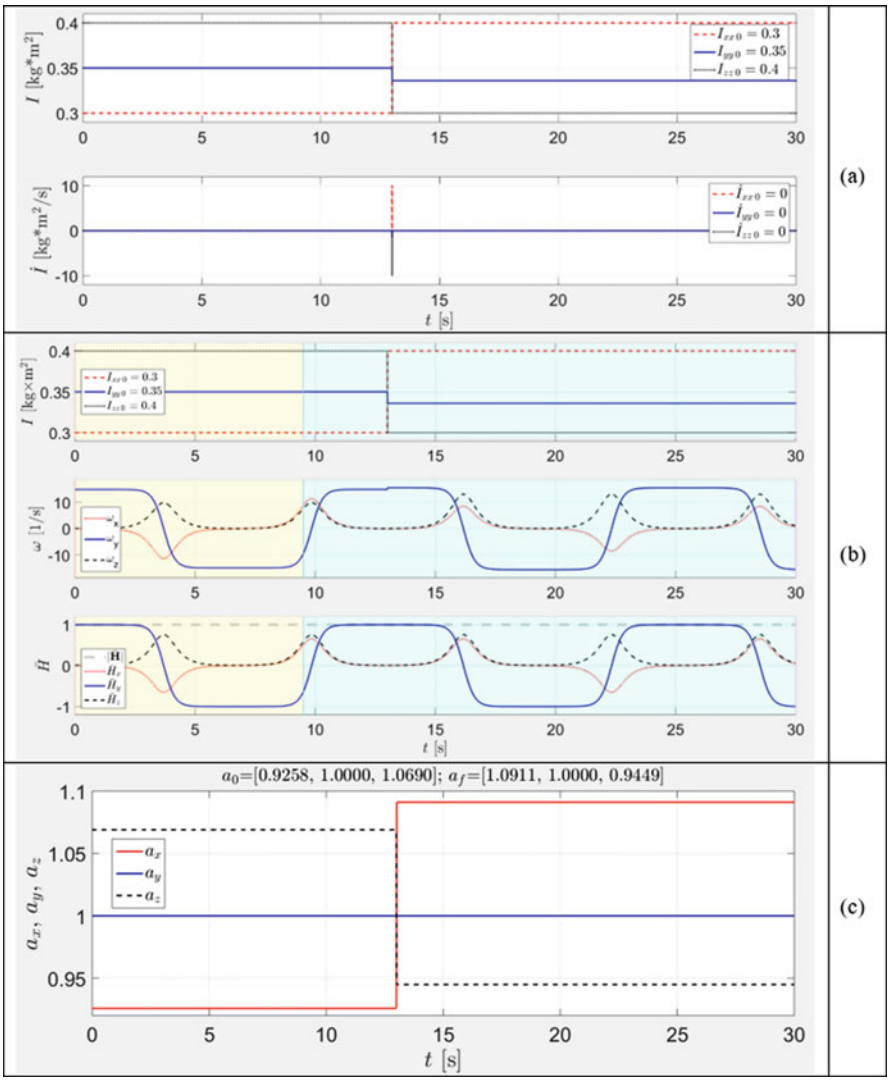


Fig. 5.71 Simulation results for the "Garriott's-Dzhanibekov's effect" flipping motion with godograph sliding along "prime" and "dual" AMSs and KEEs: (**a**) principal moments of inertial and their rates; (**b**) components of the angular velocity non-dimensional angular momentum; (**c**) radii of the KEE

ensuring transition from the "prime" to the "dual" KEE, shown in Fig. 5.72. It is interesting to observe the change of the direction of the motion of the godograph on the "prime" from counterclockwise to opposite after the switch to the "dual" KEE. These directions are shown in Fig. 5.72 with red arrows. Also, the KEEs are shown

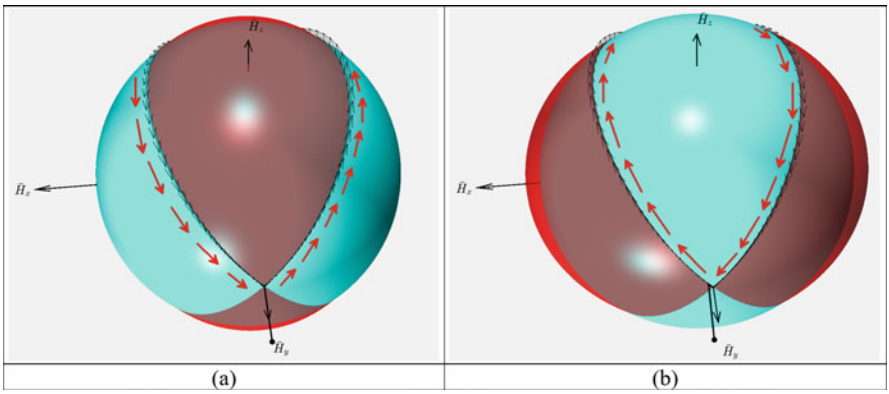


Fig. 5.72 (**a**) CCW is the direction of the godograph on "prime" AMS and KEE; (**b**) CW is the reversed direction of the godograph on "prime" AMS and KEE

as semi-transparent surfaces, enabling to see through and recognise the segments of the AMSs, covered by the KEEs.

5.11.8 *Summary of the Method of Installing of the Godograph of the Non-dimensional Vector of Angular Momentum of Tumbling Spacecraft into Conjugate Separatrix*

In view of the importance of this method, we are presenting its brief summary. The tumbling spacecraft with assumed $I_{xx} < I_{yy} < I_{zz}$ can be represented with the polhode (shown in Fig. 5.73 with red line). For the given I_{xx}, I_{yy} *and* I_{zz} the current separatrix has a dihedral angle with z body axis. *As one of the simplest manoeuvres, leaving y-axis to be an intermediate axis after the applied* morphing, it is proposed to use point #1 or #2 for *installing* godograph of the non-dimensional angular momentum *into a new separatrix*, touching polhode at points #1 and #2, being intersection points between the xz plane and polhode. (Another similar simple manoeuver can involve points #3 and #4 but would require to change intermediate axis from y to x.)

For the given I_{xx} and I_{zz}, we can select I_{yy}, ensuring that the separatrix for the new set of moments of inertia is passing through points #1 and #2. For this selection, we determine α -angle of inclination of the separatrix plane with respect to the z-axis [3]:

$$\alpha = \arctan\left(\sqrt{\frac{I_x\left(I_y - I_z\right)}{I_z\left(I_x - I_y\right)}}\right) = \arctan\left(\sqrt{\eta\left(\frac{1}{\xi} - 1\right)}\right)$$
$$\left(\text{only applicable for the } I_{xx} < I_{yy} < I_{zz} \text{ notations and } 0 < \eta < 1 \text{ and } 0 < \xi < 1\right) \tag{5.61}$$

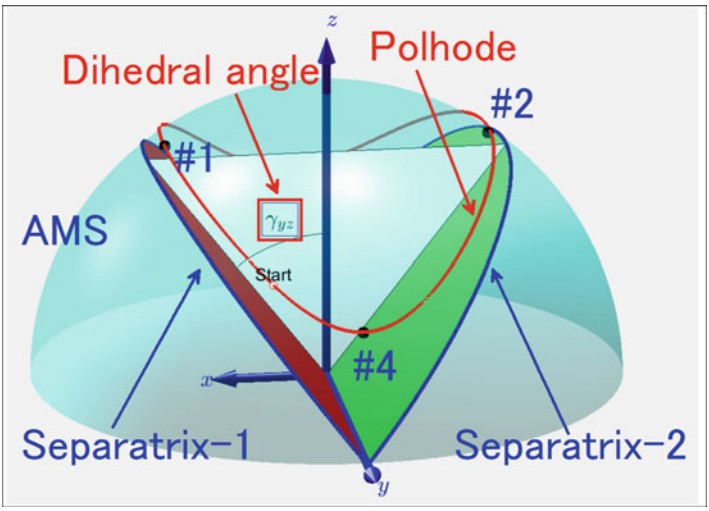


Fig. 5.73 System current separatrices shown together with the current polhode and points, enabling installing into separatrices for possible switch to the new separatrices, changing motion from tumbling to flipping type

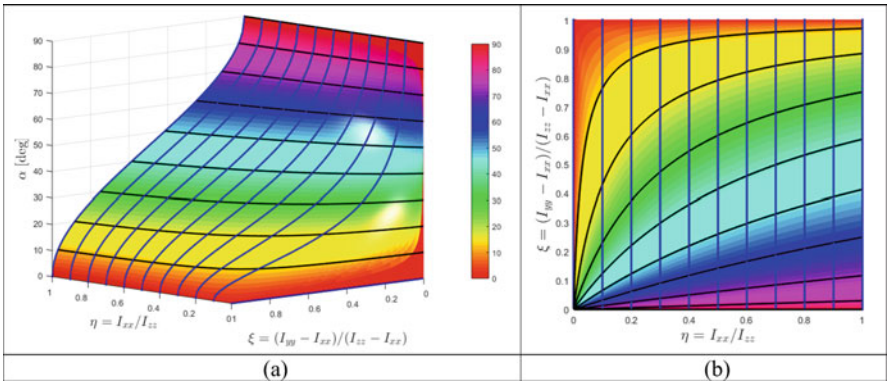


Fig. 5.74 Changes in the α-angle due to the variation in both η and ξ: (**a**) 3D surface plot for $\alpha(\eta, \xi)$ function with colorbar added; (**b**) 2D projection of the $\alpha(\eta, \xi)$ surface with its contour lines: $\eta = 0 : 0.1 : 1; \quad \alpha = 0 : 10 : 90$

Changes in the α-angle due to the variation in both, η and ξ, are shown in Fig. 5.74. Note that for convenience, values of α-angles are presented in degrees.

The method, described in [3], was based on the calculation of the value of the intermediate moment of inertia I_{yy} for the specified α-angle and known values of $I_{xx} = 2.4$ and $I_{zz} = 3.15$. For this formulation, Eq. (5.61) can be rewritten as follows:

$$\xi = \left\{1 + \left[(\tan\alpha)^2/\eta\right]\right\}^{-1} \tag{5.62}$$

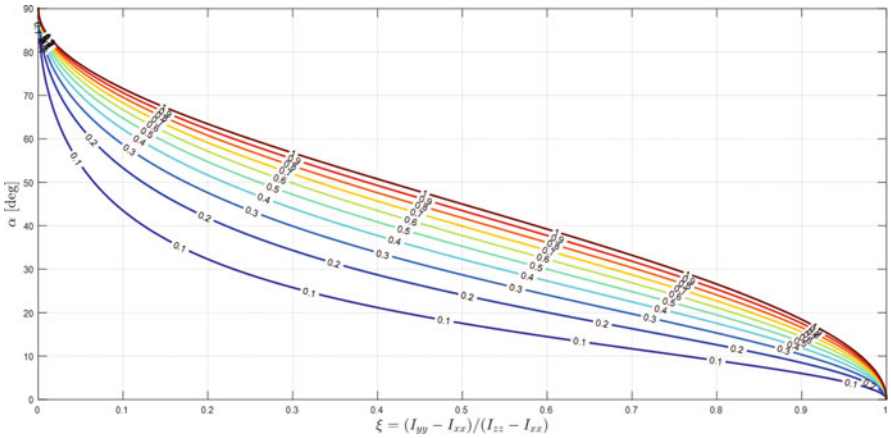


Fig. 5.75 Changes in the α-angle due to the variation in ξ for selected values of $\eta = [1{:}10]/10$

In the particular case considered in reference [4], for $I_{xx} = 2.4$ and $I_{zz} = 3.15$, the corresponding value of η is equal to $\eta = 0.7619$; furthermore, Eq. (5.62) gives $\xi = 0.5907$, which (as per Eq. (5.52)) corresponds to $I_{xx} = 2.8430$.

The generic graphical method, corresponding to this procedure, is illustrated in Fig. 5.75, where α-angle (shown in degrees) is plotted as a function of ξ for various values of $\eta = [.1, .2, .3, .4, .5, .6, .7, .8, .9, 1]$.

Alternatively, if the minimum and maximum values of the moments of inertia are fixed and the α-angle is established, then the intermediate value of the moment of inertia can be calculated, using Eq. (5.50), solved for I_{yy}:

$$I_{\text{int}} = \frac{\left[(\tan\alpha)^2 + 1\right] I_{\min}}{(\tan\alpha)^2 + \left(\frac{I_{\min}}{I_{\max}}\right)} = \frac{\left[(\tan\alpha)^2 + 1\right]}{(\tan\alpha)^2 + \eta} I_{\min}, \text{ where } \eta = \frac{I_{\min}}{I_{\max}} \tag{5.63}$$

5.12 Inertial Morphing in Novel Designs of Acrobatic Spacecraft for 90 Degrees Inversions: Method of "Installing into Separatrix" with Separatrix-to-Separatrix Transfer

Figure 5.76 is dedicated to another possible application of inertial morphing. It illustrates 90° change of the spacecraft spin axis, using three morphing procedures: the first is to activate unstable flipping motion; second is to switch to x separatrix; and the third is to stop tumbling motion with transfer of spacecraft spin from y to a new nominated *body* axis, being x in this illustration example.

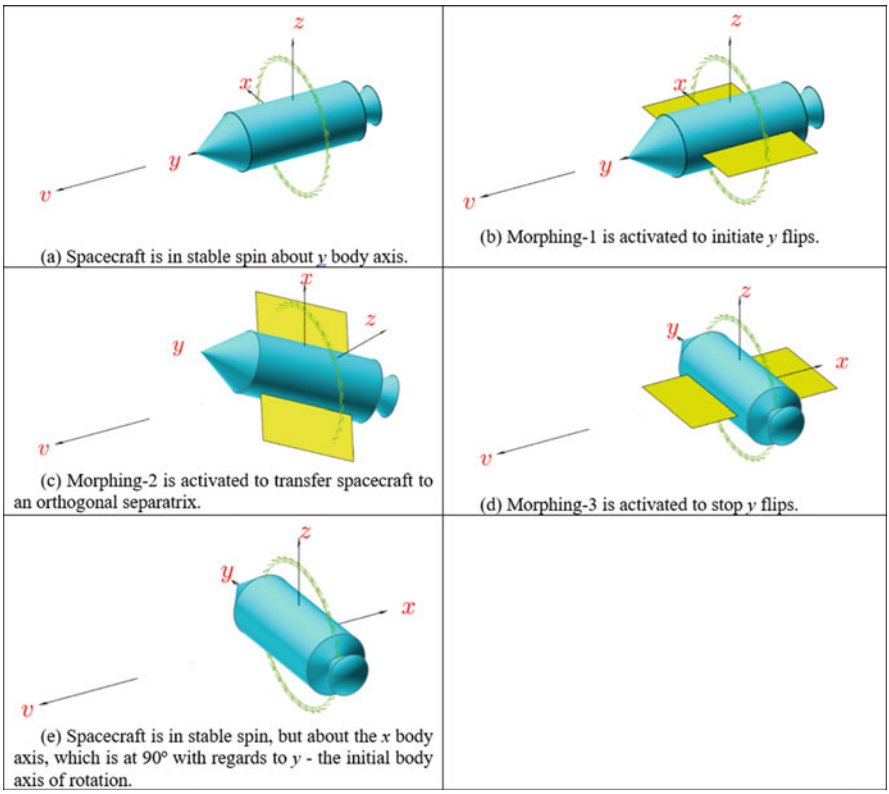


Fig. 5.76 Example of application of inertial morphing for inversion of spacecraft [9]: (**a**–**e**) 90° inversion, allowing change of spacecraft rotation from longitudinal to lateral

System's initial conditions are as follows: $\omega_{x,i} = 0.01$, $\omega_{y,i} = 8.57$, $\omega_{z,i} = 0.01$ (all in rad/s), $I_{xx} = 3$, $I_{yy} = 2.8$, $I_{zz} = 5$ (all in kg $\times$ m^2). It should be noted that in Fig. 5.76 initial and final spin axes are presented in *body*-axis system. And in the global axis system, due to the law of conservation of angular momentum, both initial and final spin orientations are aligned with the same direction of the non-dimensional angular momentum vector $\overline{\mathbf{H}}$.

Spacecraft morphing parameters, corresponding to Fig. 5.76, are shown in Table 5.1 and are also illustrated in Fig. 5.77.

Initially the system is in stable spin (shown as IM0 in table). IM1 was designed to insert the system into y separatrix, similar to shown in Fig. 5.69c. This initiates a y flip, during which IM2 is applied, forcing the system to transfer to x separatrix. So, this acrobatics involves two "installing into separatrices" actions. Important control consideration is as follows: IM2 is applied at the instant when in body axes vector $\overline{\mathbf{H}}$ (to be displayed in body axes, as in Fig. 5.69c) passes *intersection* of x and y separatrices. IM3 is designed to stop x flips in the way shown in Fig. 5.78a. Figure 5.78b shows components of the angular velocity of the spacecraft. It is remarkable

Table 5.1 Spacecraft morphed parameters in the example with 90° inversion

IM	t	r_x	r_y	r_z	I_{xx}	I_{yy}	I_{zz}
Index	s	mm	mm	mm	kg $\times$m^2	kg $\times$m^2	kg $\times$m^2
0	0	548	510	447	3	2.8	5
1	10	565	495	524	3	3.1	5
2	18.25	474	566	806	4.5	3.1	5
3	25.15	403	608	949	5.5	3.1	5
f	35.14	403	608	949	5.5	3.1	5

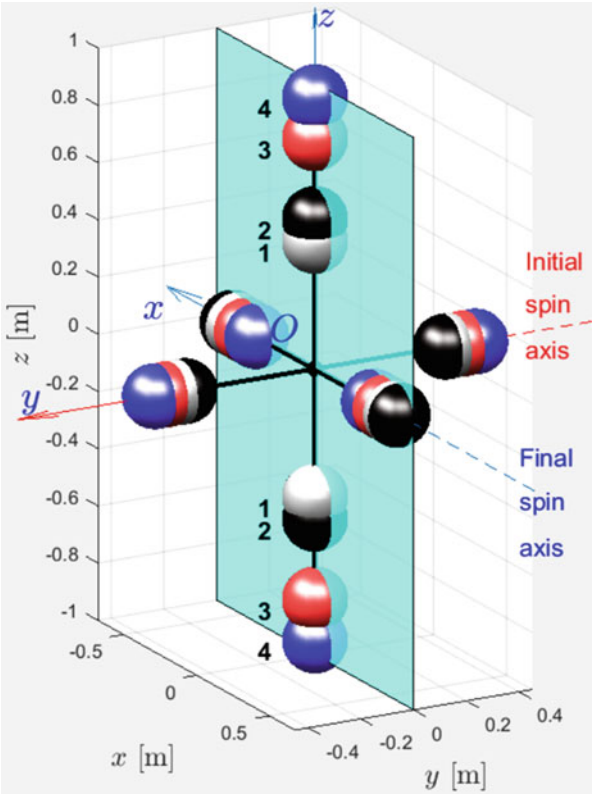


Fig. 5.77 Positions of the masses during spacecraft 90° inversion via inertial morphing for acrobatic manoeuvre in Fig. 5.76 (**a**)–(**e**) [9]: illustrated for six-mass spacecraft model

that, initially, the system had y predominant angular velocity component $\omega_{y,i}$, with two other components being close to zero $\omega_{x,i} = \omega_{z,i} \approx 0$. After three morphings, the system has only one predominant component of the angular velocity, $\omega_{x,i}$, with the values of two other components being close to zero: $\omega_{y,i} \approx 0$, $\omega_{z,i} \approx 0$. At last, Fig. 5.78c shows changes of the radii of the KEE, reflecting the applied changes to the values of the moments of inertia during controlled morphings.

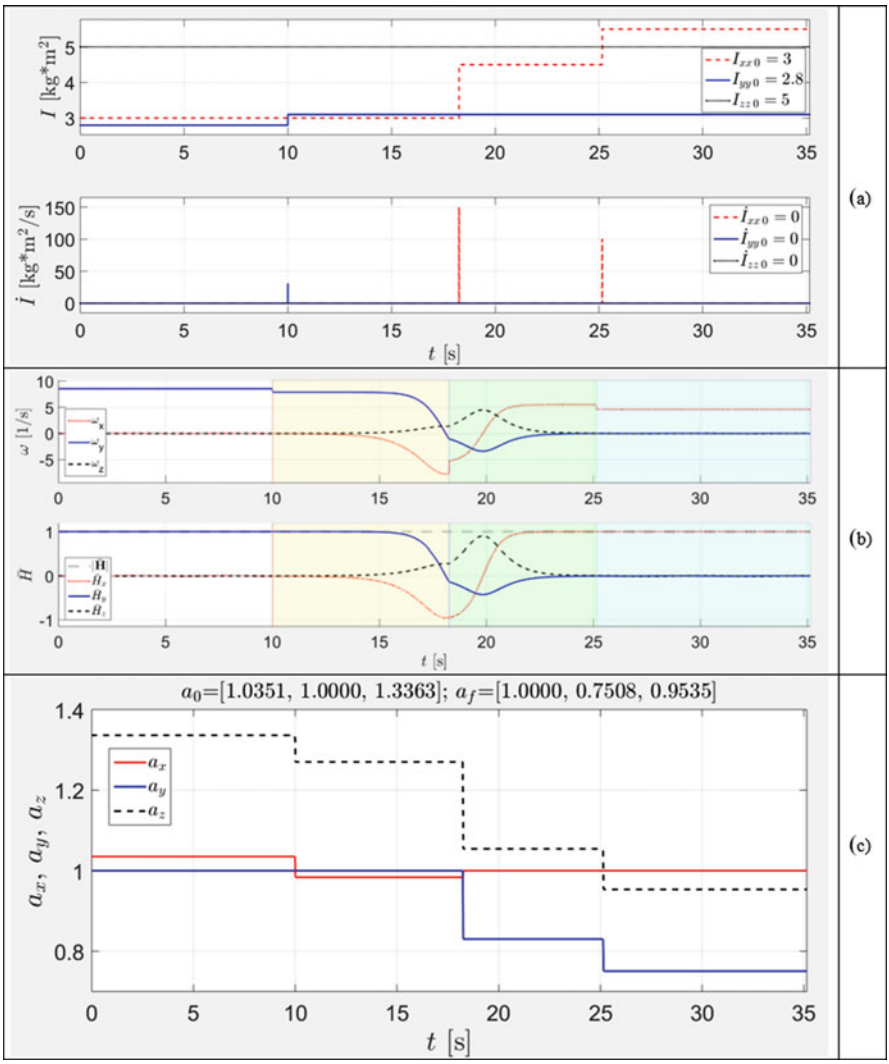


Fig. 5.78 Time histories of the parameters of the system: (**a**) moments of inertia and their time rates, (**b**) components of the angular velocity and non-dimensional angular momentum and (**c**) radii of the KEE

5.13 Demo of Combined Multiphase Inertial Morphing: Consecutive "Parade" of All Three Orthogonal Inversions, Associated with x, y and z Body Axes

In order to demonstrate capability of the proposed method, in Fig. 5.79 we present results for a single simulation case, during which the spinning body is

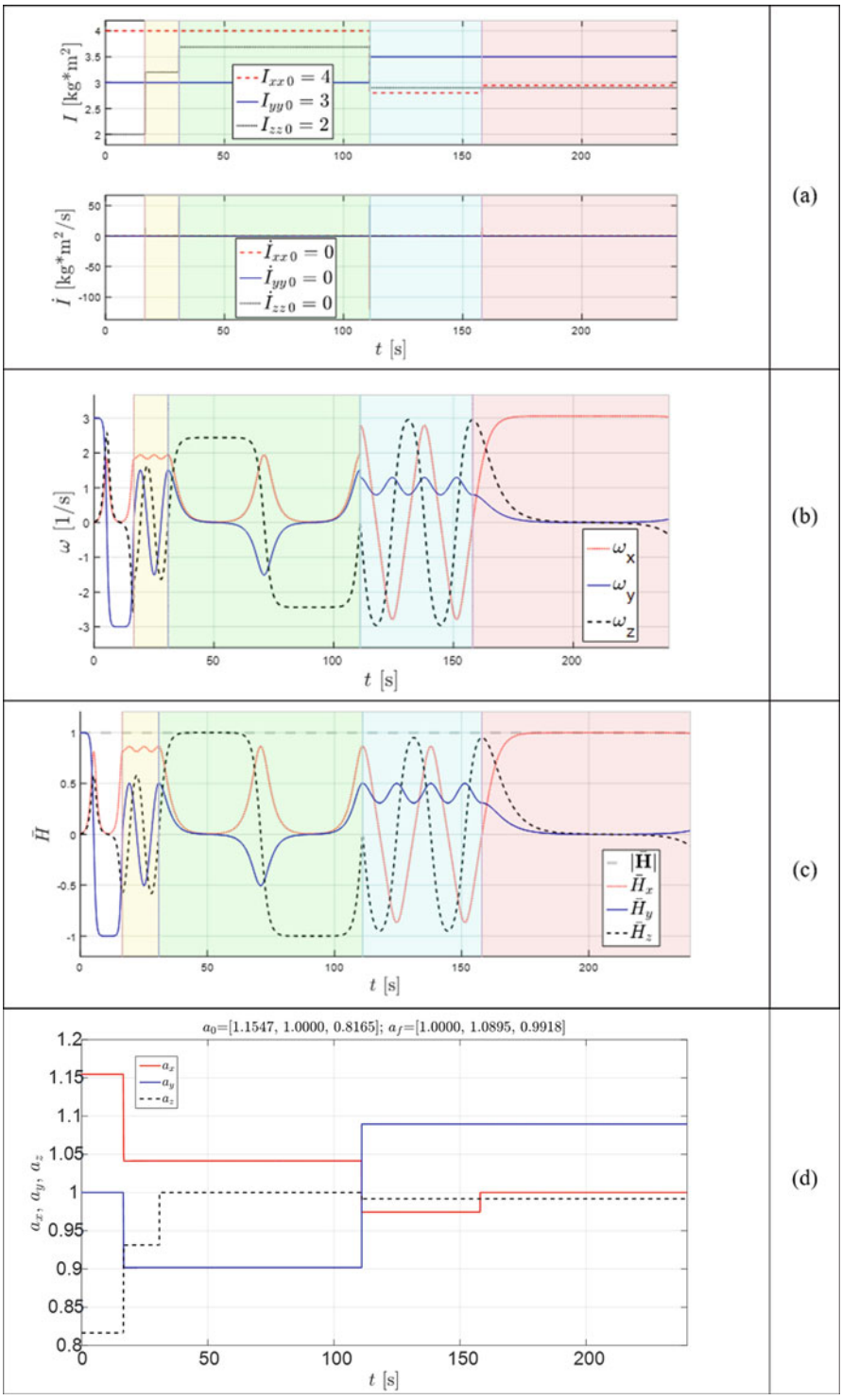


Fig. 5.79 Time history of the (**a**) I_{xx}, I_{yy}, I_{zz}; (**b**) ω_x, ω_y, ω_z; (**c**) H_{total}, H_x, H_y, H_z; (**d**) a_x, a_y and a_z during four-stage "all-axes inversion parade"

reconfigured four times. The carefully selected scenario for the applied inertial morphing (changes in the system, leading to the change of the values of the principal moments of inertia) enables to achieve the following:

(1) Established flipping motion along the y-axis (with possibility for y inversion), distinguished with a white background in Fig. 5.79
(2) Established flipping motion along the z-axis (with possibility for z inversion), distinguished with green background in Fig. 5.79
(3) Established flipping motion along the x-axis (with possibility for x inversion) distinguished with pink background in Fig. 5.79

So, it has been demonstrated that the predominant spin can be consecutively passed on to any of the body axis with multiple possibilities for inversion at any stage of the stabilised motion and then stabilisation of the desirable orientation. In other words, if the object had a cube shape, based on this example, it was possible to perform transition of the spinning motion of the cube, allowing exposure of each of its six faces to the direction of the initial predominant spin. We call this compound demo case "all-axes inversion parade".

In Fig. 5.79 areas with yellow and light blue backgrounds are the transition phases. It can be seen that the angular momentum is conserved. The total simulation time for this case was 240 s, and the instants, when inertial morphings were initiated, were 16.5555 s, 30.8525 s, 111.0420 s and 157.9995 s.

Some of the most critical stages of the "parade" scenario are also illustrated in Fig. 5.80.

It can be observed that the period of inversion stages differs significantly. Periods for the "flipping" stages can be calculated analytically [8], using elliptic integrals. They are becoming very large, especially when the values of two moments of inertia come close.

5.14 Enhancement of the Reorientation and Change of the Spin Axis Using Moment Wheel

For completeness of this chapter, we need to mention another powerful aspect of further enhancement of the spinning spacecraft attitude control capabilities: adding one or a set of moment reaction wheels, which are often used on various space systems [39].

Differential equations of motion of the spacecraft, equipped with wheels, could be presented as follows:

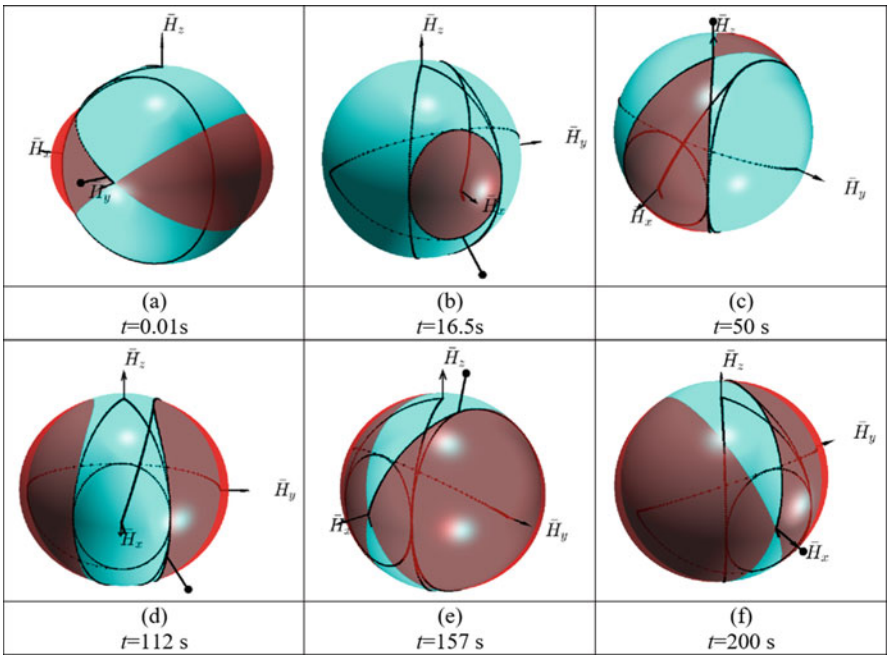


Fig. 5.80 Critical instances of spacecraft "parade" of all three orthogonal inversions: (**a**) y inversion stage (established "flipping"), (**b**) tumble (orbiting x-axis), (**c**) x inversion stage (established "flipping"), (**d**) tumble (orbiting y-axis) and (**e**) x inversion stage (established "flipping") (**f**) at "parking" point: stabilisation opportunity

$$\begin{bmatrix} I_{xx} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{yy} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{zz} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sin\theta\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & \sin\theta\cos\phi & -\sin\phi & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & 1 \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \\ \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{Bmatrix} = \begin{Bmatrix} \left(I_{yy} - I_{zz}\right)\omega_y\omega_z - \dot{I}_{xx}\omega_x \\ \left(I_{zz} - I_{xx}\right)\omega_z\omega_x - \dot{I}_{yy}\omega_y \\ \left(I_{xx} - I_{yy}\right)\omega_x\omega_y - \dot{I}_{zz}\omega_z \\ \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} - \begin{Bmatrix} n_{\omega_1} + \omega_2 l_3 - \omega_3 l_2 \\ n_{\omega_2} + \omega_3 l_1 - \omega_1 l_3 \\ n_{\omega_3} + \omega_1 l_2 - \omega_2 l_1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{5.64}$$

Even simple preliminary cases, involving one wheel and not sophisticated wheel's controls (with one of them presented in Fig. 5.81), enabled us to find sig-

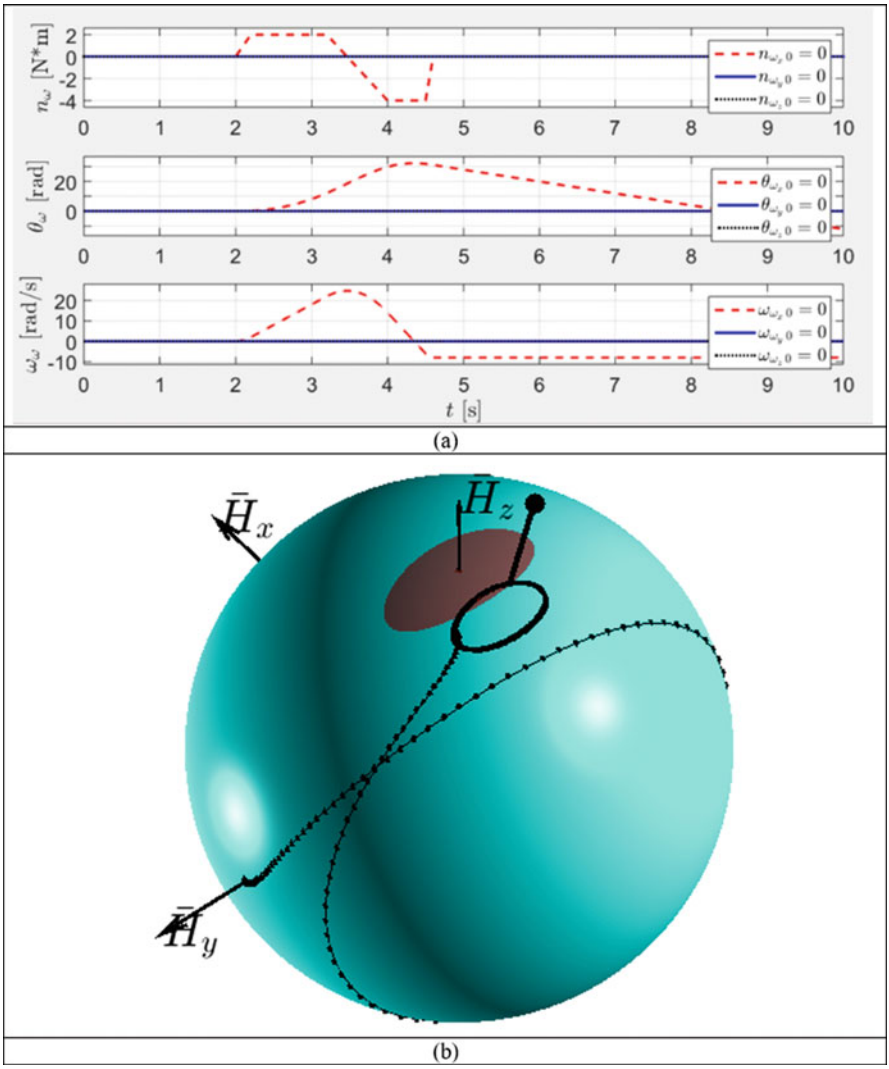


Fig. 5.81 Shift of stabilisation point, achieved with compounding use of the inertial morphing and reaction wheel

nificant influence of this enhancement on performance of the system. In particular, it was possible to significantly influence the period of inversion, make inversions asymmetrical, etc. Authors intend to explore these capabilities in more detail in the future works.

The proposed multistage transfer method can be called a method of "installing into the separatrices". It enables conversion of the predominant rotation about any body axis into rotation about any other body axis.

With the capability of this axis-to-axis transfer, spacecraft essentially could perform three types of inversions, associated with any of three body axes.

In order to demonstrate capabilities of the method, "all-axes inversion parade" was presented, during which the spinning system was transitioned through three consecutive stages with inversion, associated with each of the body axes, x, y and z. This is in contrast with the classical Garriott's-Dzhanibekov's effect demonstration, where only one axis inversion was possible.

This method enables to get precision control of the spacecraft multi-axis inversion without using conventional gyroscopes.

The method is based on the geometric interpretation of the spinning systems. It employs angular momentum sphere and kinetic energy ellipsoid, polhodes and separatrices and their evolution due to the variation of the inertia properties of the system, called inertial morphing. In this chapter, we formalised non-dimensional constructions.

5.15 Animations in Virtual Reality

Computer capabilities of the spinning rigid-body simulator, intended to deal with systems, enabling "inertial morphing", were enhanced with programming of the Virtual Reality block. It enables animated visualisation of the attitude dynamics of the spinning spacecraft.

It allows the operator/designer to observe results of the control actions on the spacecraft, initiating (on the operator's request), for example, transfers of the spacecraft from stable spin to unstable, flipping, mode and vice versa. Virtual Reality not only provides the operator with instant impression on the quality of the executed manoeuvre but also enables observation of the progression of the manoeuvre from various coordinate systems (inertial and body axes).

An example of the simulation of the scenario is shown in Fig. 5.82. During this simulation, the stable spacecraft is transferred into the unstable mode, performs one flip (i.e. changes its axial attitude by 180°) to direct its antenna and then is stabilised by the operator, with flipping completely stopped. Various controls and diagnostics on the progression of the scenario are displayed on the screen, including notification on the current control status, for example, "Garriott's-Dzhanibekov's effect" is "OFF" or "ON".

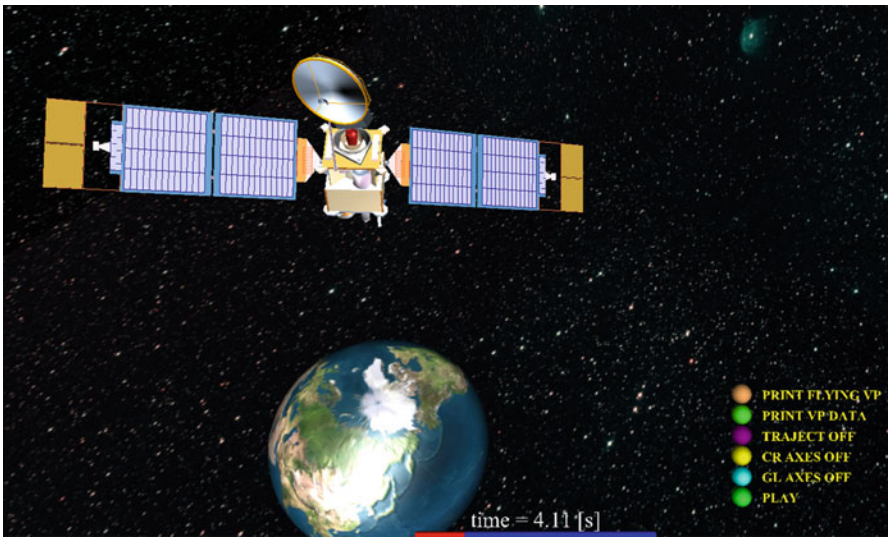


Fig. 5.82 Computer screen snapshot, presenting one of the instants in the animated scenario, involving control of the "Garriott's-Dzhanibekov's effect"

5.16 Examples of the Conceptual Designs of the Inertially Morphed Systems

5.16.1 Example-1 Design, Involving "Six-Masses" Repositioned Along Body Axes

Figure 5.83 shows a simple "six-mass" model of the rotating system, where masses in pairs can be independently repositioned along their respective axes. This reposition is done symmetrically in each pair and can be performed with independent actuators.

Variation of the principal moments of the inertia of the system I_{xx}, I_{yy} and I_{zz} with controlled parameters r_x, r_y and r_z can be described with the following analytical relationships:

$$\begin{aligned} I_{xx} &= 2 \times \left[m_y {r_y}^2 + m_z {r_z}^2 \right] \\ I_{yy} &= 2 \times \left[m_z {r_z}^2 + m_x {r_x}^2 \right] \\ I_{zz} &= 2 \times \left[m_x {r_x}^2 + m_y {r_y}^2 \right] \end{aligned} \tag{5.65}$$

These can be illustrated graphically as 3D surfaces. To be specific, three surfaces, corresponding to I_{xx}, I_{yy} and I_{zz}, are plotted in Fig. 5.84a for the example, given by the following parameters: $m_{1x} = m_{2x} = m_{1y} = m_{2y} = 0.5$ kg, $m_{1z} = m_{2z} = 0.3$ kg – and the radii for the masses m_{1z} and m_{2z} are equal to $r_z = 0.3$ m.

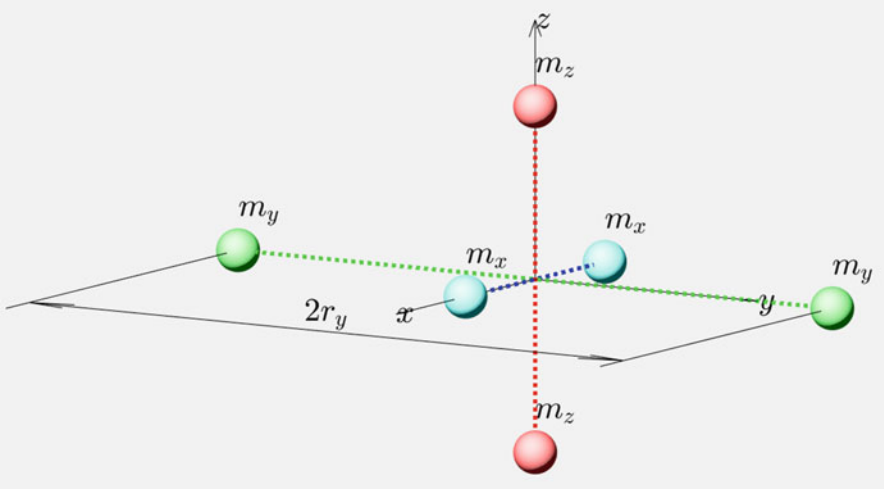


Fig. 5.83 Illustration of the "six-mass" model, with masses symmetrically repositioned in pairs via their translation along the *x*, *y* and *z* axes

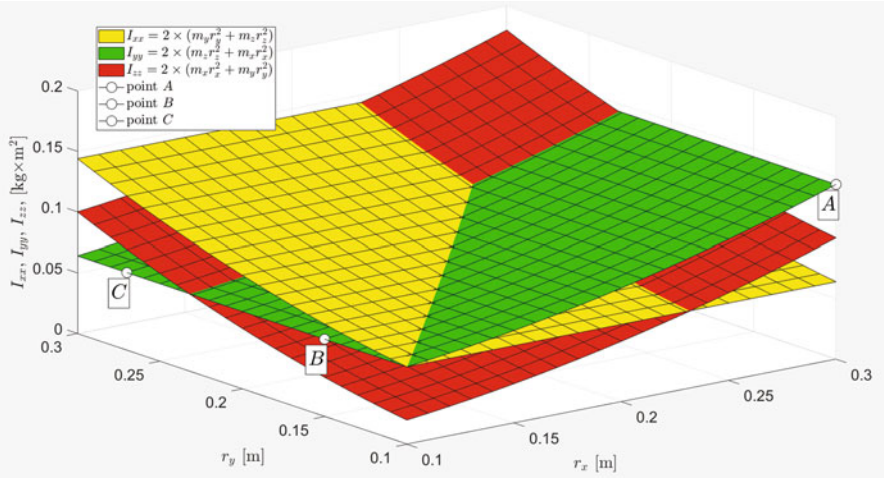


Fig. 5.84 Illustration of the "six-mass" model, with masses symmetrically repositioned in pairs along the *x*, *y* and *z* axes

Figure 5.84 shows that in the illustration example, there could be found numerous combinations of r_x and r_y for which the green surface, corresponding to I_{yy}, could take lowest, intermediate and highest positions. With this, if the system is provided with *y* initial predominant rotation, even for the fixed r_z, proper selection of varied r_x and r_y could switch the rotation about the *y* axis between the stable spin and unstable flipping motion. For example, if the system starts with $r_x = 0.3$ m, $r_y = 0.1$ m,

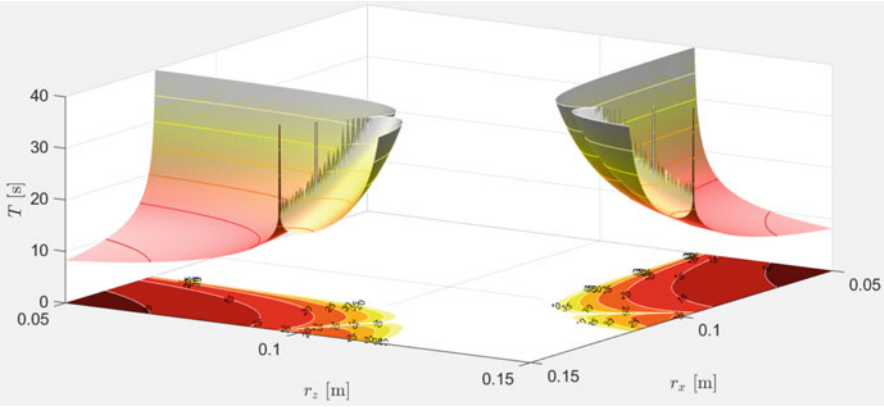


Fig. 5.85 Period of the flipping motion $T = T(r_x, r_z)$ as a function of the x and z positions of the m_x and m_z masses in the six-mass model for fixed $r_y = 0.12$ m

$r_z = 0.3$ m, $\omega_x = 6$ rad/s, $\omega_y = 0.01$ rad/s, $\omega_z = 0.01$ rad/s, we should expect this spin to be stable, because for the given parameters, I_{yy} would be the maximum moment of inertia (see point A in Fig. 5.84). If using translational actuators for the masses m_x and m_y can change position of these masses to the new values, let us say $r_x = 0.1$ m, $r_y = 0.15$ m, and no change to $r_z = 0.3$ m, then for the new radii of the masses, I_{yy} would become an intermediate moment of inertia (shown as point B in Fig. 5.84), and the unstable flipping motion would be triggered. If afterwards, at the time of passing the pole by the system's angular momentum vector, another morphing is applied (setting $r_y = 0.27$ m), and with no changes applied to (leaving them at $r_x = 0.15$ m and $r_y = 0.3$ m), then the moment of inertia I_{yy} (shown as point C in Fig. 5.84) would become the minimal moment of inertia, and the flipping motion of the system would be transferred to the stable spin.

Figure 5.85 complements design process and shows periods of the flipping motions of the "six-mass" design for various combinations of r_x and r_z. This function is shown as a 3D surface for the fixed value of the distance $r_y = 0.12$ m. It confirms that proper selection of r_x, r_y and r_z would not only allow desired manipulations with the principle moments of inertia of the system but would also enable selection of the desired period of the flipping motion. This may be important for the cases, where agile or prolonged manoeuvre would be required.

5.16.2 Example-2 Design: "Scissors" Model for Inertial Morphing

Figure 5.86 shows an example of using the "*XZ* scissors" mechanism for performing changes to the inertial properties of the system. It enables that the system remains

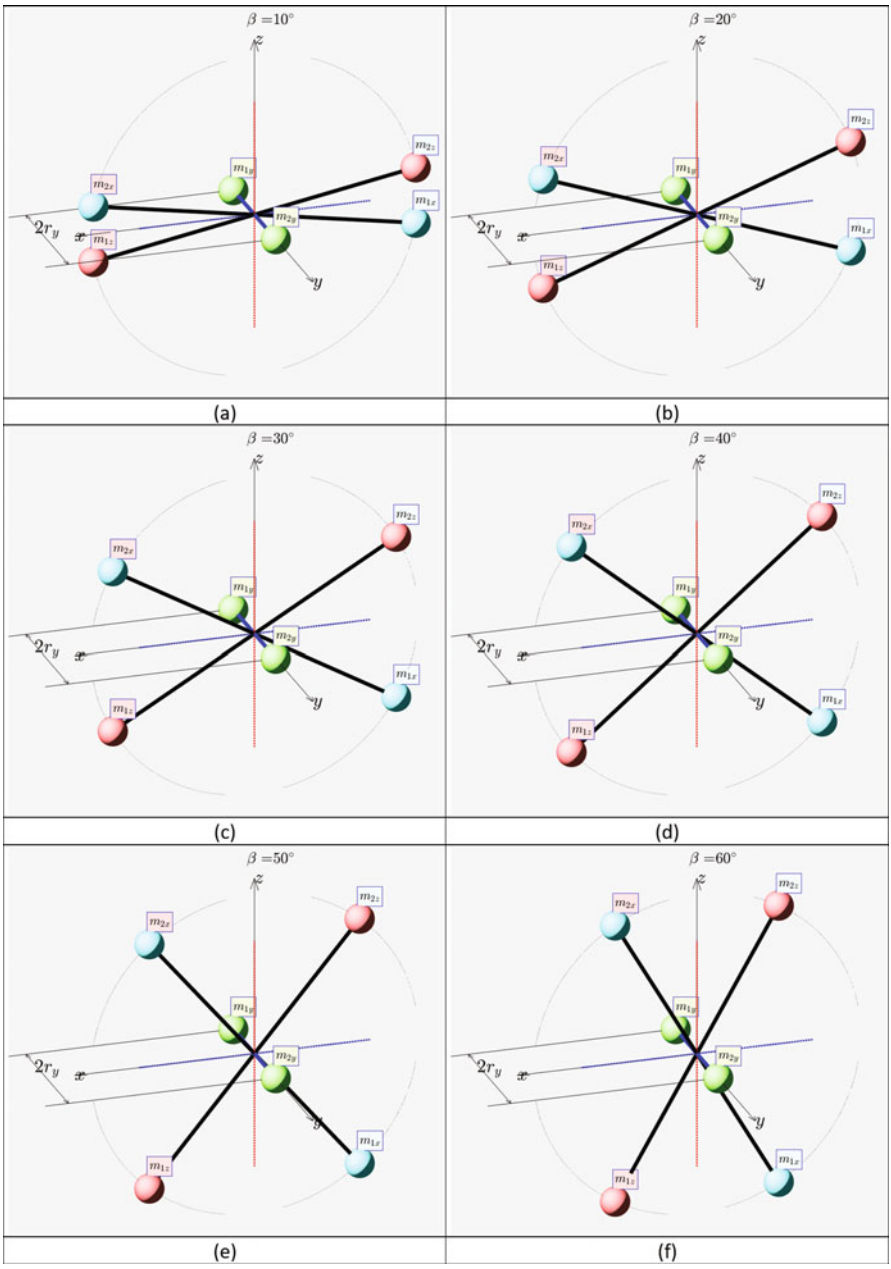


Fig. 5.86 Illustration of the "*XZ* scissors" mechanism, employed to perform controlled inertial morphing of the system: (**a**)–(**h**) correspond to the following angles between the m_x-m_x link and the x axis (also synchronised with the angle between the m_z-m_z link and the x axis): $\beta = 10°$, 20°, 30°, 40°, 50°, 60°, 70°, 80°

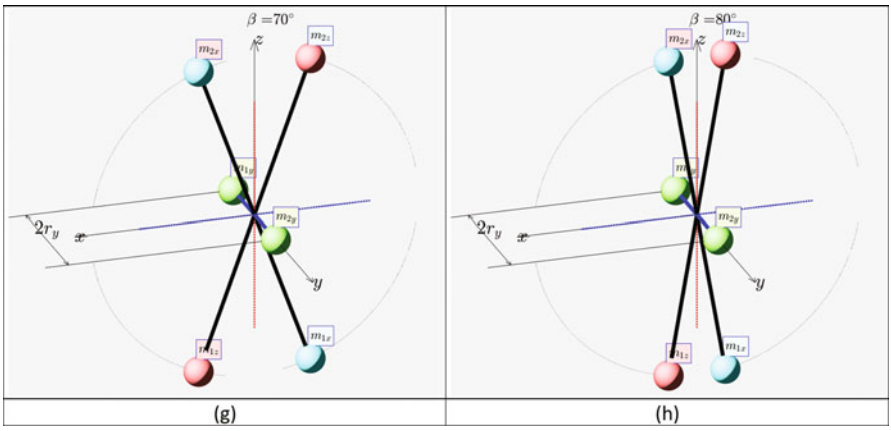


Fig. 5.86 (continued)

dynamically balanced about any of the body axes. Figure 5.86 presents configuration of the same system, but with different angles β between the m_x-m_x links and the "x" axis (being kept by the mechanism the same as the angle between the m_z-m_z links and the x axis), which can be controlled with a single actuator, which will be called β-actuator (not shown in the Fig. 5.86). It is assumed in the presented example, the synchronously symmetric positions r_y of the y masses.

Variation of the principal moments of the inertia of the system I_{xx}, I_{yy} and I_{zz} with both controlled parameters, β and r_y, can be described with the following analytical relationships:

$$\begin{aligned} I_{xx} &= 2 \times \left[m_x (r_x \sin\beta)^2 + m_z (r_z \sin\beta)^2 + m_y r_y^2 \right] \\ I_{yy} &= 2 \times \left(m_x r_x^2 + m_z r_z^2 \right) \\ I_{zz} &= 2 \times \left[m_x (r_x \cos\beta)^2 + m_z (r_z \cos\beta)^2 + m_y r_y^2 \right] \end{aligned} \tag{5.66}$$

These can be illustrated graphically as 3D surfaces. To be specific, three surfaces, corresponding to I_{xx}, I_{yy} and I_{zz}, are plotted in Fig. 5.87(a) for the example, given by the following parameters: $m_{1x} = m_{2x} = m_{1z} = m_{2z} = 0.3$ kg, $m_{1y} = m_{2y} = 0.5$ kg – and the length L of all radii for the masses m_{1x}, m_{2x}, m_{1y}, m_{2y}, m_{1z} and m_{2z} is equal to $r_x = r_y = r_z = L = 0.25$ m.

Figure 5.87b shows the same plot, but with different axes limits and as the top view. It is interesting to observe that if the design with $r_y = 0.25$ m $\beta = 45°$ is initially provided with the predominant rotation about the y axis, the system would remain in the stable spin, and as for the specified parameters, I_{yy} would be the *maximum* moment of inertia, being larger than any other moments of inertia, I_{xx} and

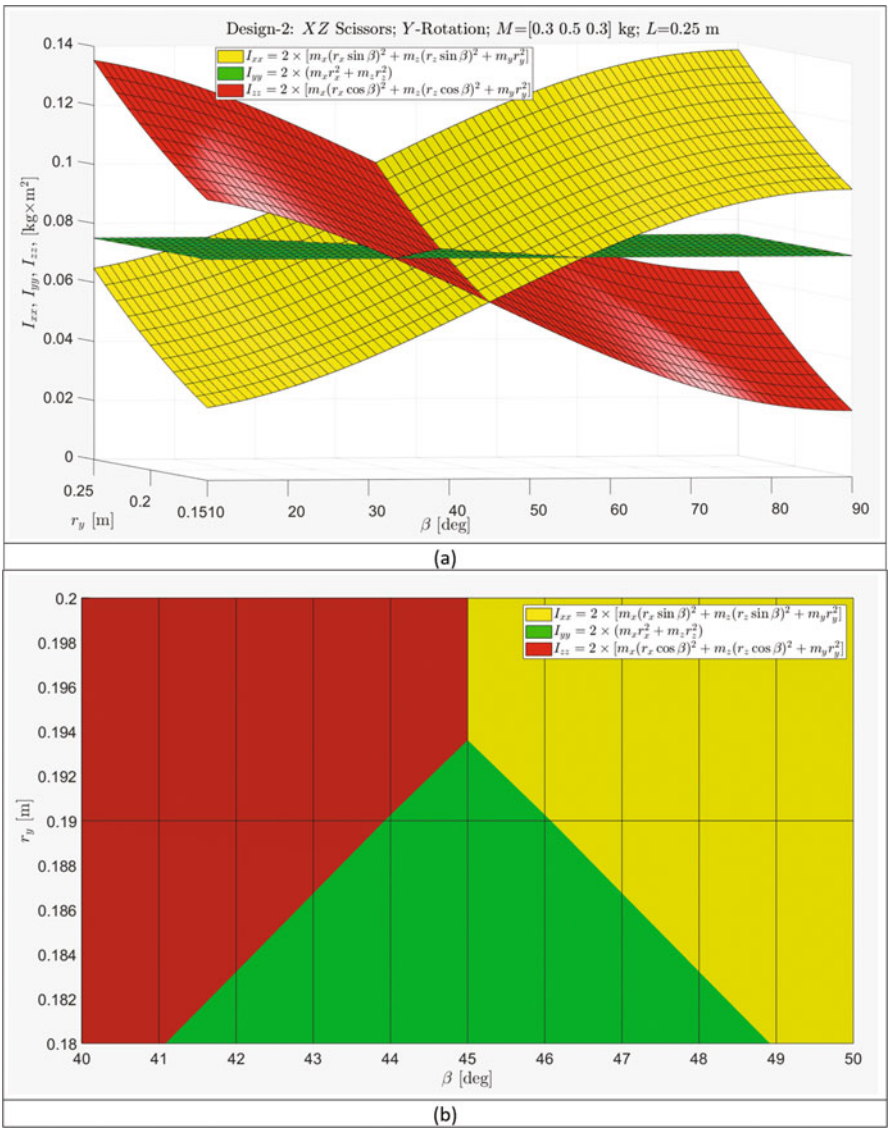


Fig. 5.87 Principal moments of inertia of the system in Fig. 5.86 as functions of β and r_y variables, which can be controlled individually by two actuators: (**a**) 3D representation and (**b**) 2D representation

I_{zz} ($I_{xx} < I_{yy}$ and $I_{zz} < I_{yy}$). However, should the angle θ be increased well above 49° or reduced well below 41° (with the r_y kept unchanged at $r_y = 0.25$ m), then the I_{yy} would become an *intermediate* moment of inertia, and the flipping motion would be triggered with the first, β-actuator. Similarly, the I_{yy} would become an *intermediate*

moment of inertia if the angle β is kept unchanged at $\beta = 45°$ but the r_y value is increased well above $r_y = 0.1935$ m, using the second, β-actuator. Therefore, in the illustration example, transition from stable spin to unstable flipping motion can be independently triggered by any of the two system's actuators or by combined action of two actuators.

Figure 5.88 complements design process and shows periods of the flipping motions of the "scissors" design for various combinations of r_y and β. This function is shown as a 3D surface for the fixed value of the distance to the masses in the xz plane.

5.16.3 *Example-3 Design: Rhombus Model for Inertial Morphing*

Figure 5.89 shows an example of using the "*XZ* rhombus" mechanism for performing changes in the inertial properties of the system. It enables that the system remains dynamically balanced about any of the body axes. Figure 5.89 presents configuration of the same system, but with different angles β between the m_x-m_z links and the x axis, which can be controlled with a single actuator, which will be called β-actuator (not shown in Fig. 5.89). It is assumed in the presented example, the synchronously symmetric positions r_y of the y masses can be varied separately, with another r_y -actuator (also not shown in Fig. 5.89).

Variation of the principal moments of the inertia of the system I_{xx}, I_{yy} and I_{zz} with both controlled parameters, β and r_y, can be described with the following analytical relationships:

$$\begin{aligned} I_{xx} &= 2 \times \left[m_y r_y^2 + m_z (L \sin\beta)^2 \right] \\ I_{yy} &= 2 \times \left[m_x (L \cos\beta)^2 + m_z (L \sin\beta)^2 \right] \\ I_{zz} &= 2 \times \left[m_y r_y^2 + m_x (L \cos\beta)^2 \right] \end{aligned} \tag{5.67}$$

These can be illustrated graphically as 3D surfaces. To be specific, three surfaces, corresponding to I_{xx}, I_{yy} and I_{zz}, are plotted in Fig. 5.90a for the example, given by the following parameters – $m_{1x} = m_{2x} = m_{1y} = m_{2y} = m_{1z} = m_{2z} = 0.3$ kg – and the length L of the m_x-m_z links is equal to 0.25 m.

Figure 5.90b shows the same plot, but with different axes limits and as the top view. It is interesting to observe that if the design with $r_y = 0.15$ m, $\beta = 45°$ is initially provided with the predominant rotation about the y axis, the system would remain in the stable spin, and as for the specified parameters, I_{yy} would be the *maximum* moment of inertia, being larger than any other moments of inertia, I_{xx} and I_{zz} ($I_{xx} < I_{yy}$ and $I_{zz} < I_{yy}$). However, should the angle β be increased well above 53° or reduced well below 37° (with the r_y kept unchanged at $r_y = 0.15$ m), then the I_{yy} would become an *intermediate* moment of inertia, and the flipping motion would be

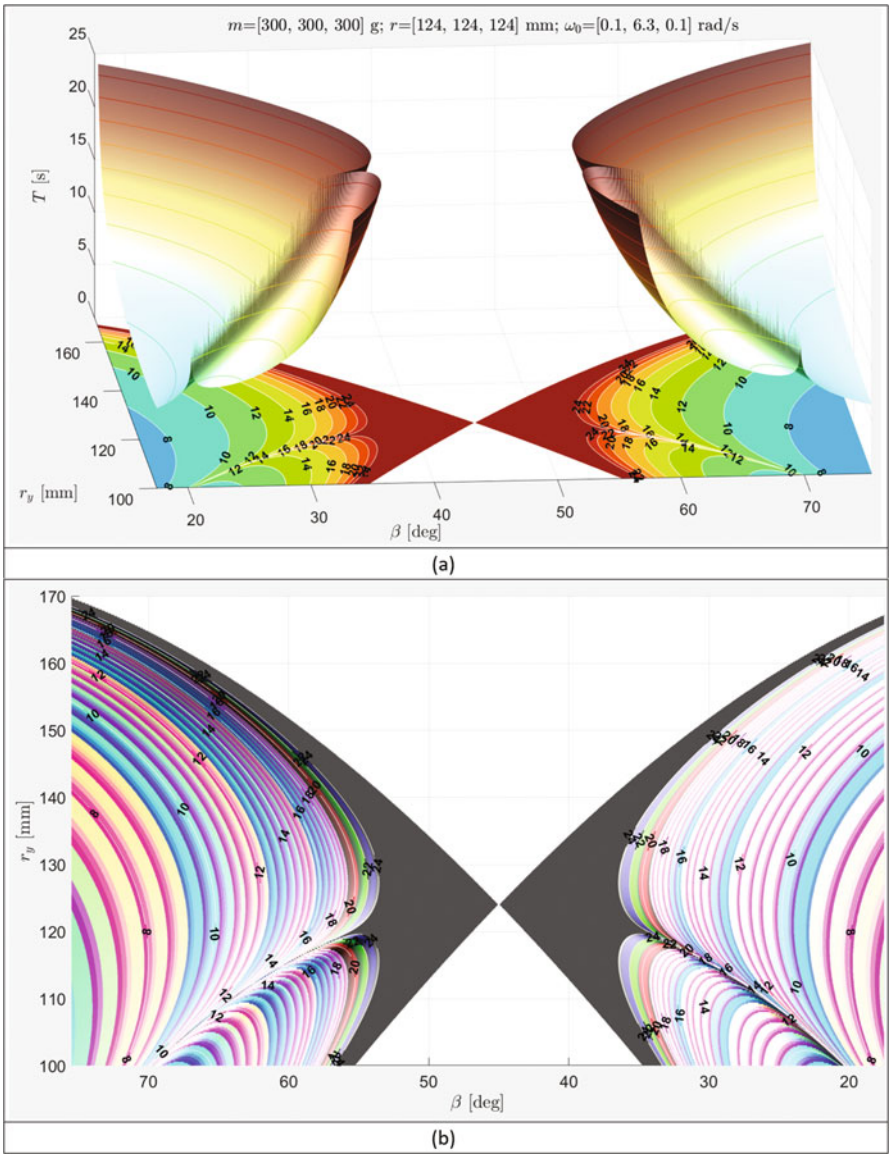


Fig. 5.88 Period of the flipping motion $T = T(r_y, \beta)$ as a function of the r_y and β parameters in the "scissors" model: (**a**) $T = T(r_y, \beta)$ presented as 3D surface, (**b**) $T = T(r_y, \beta)$ presented as 2D plot

triggered with the first, β-actuator. Similarly, the I_{yy} would become an *intermediate* moment of inertia if the angle β is kept unchanged at $\beta = 45°$, but the r_y value is increased well above $r_y = 0.1767$ m, using the second, β-actuator. Therefore, in the

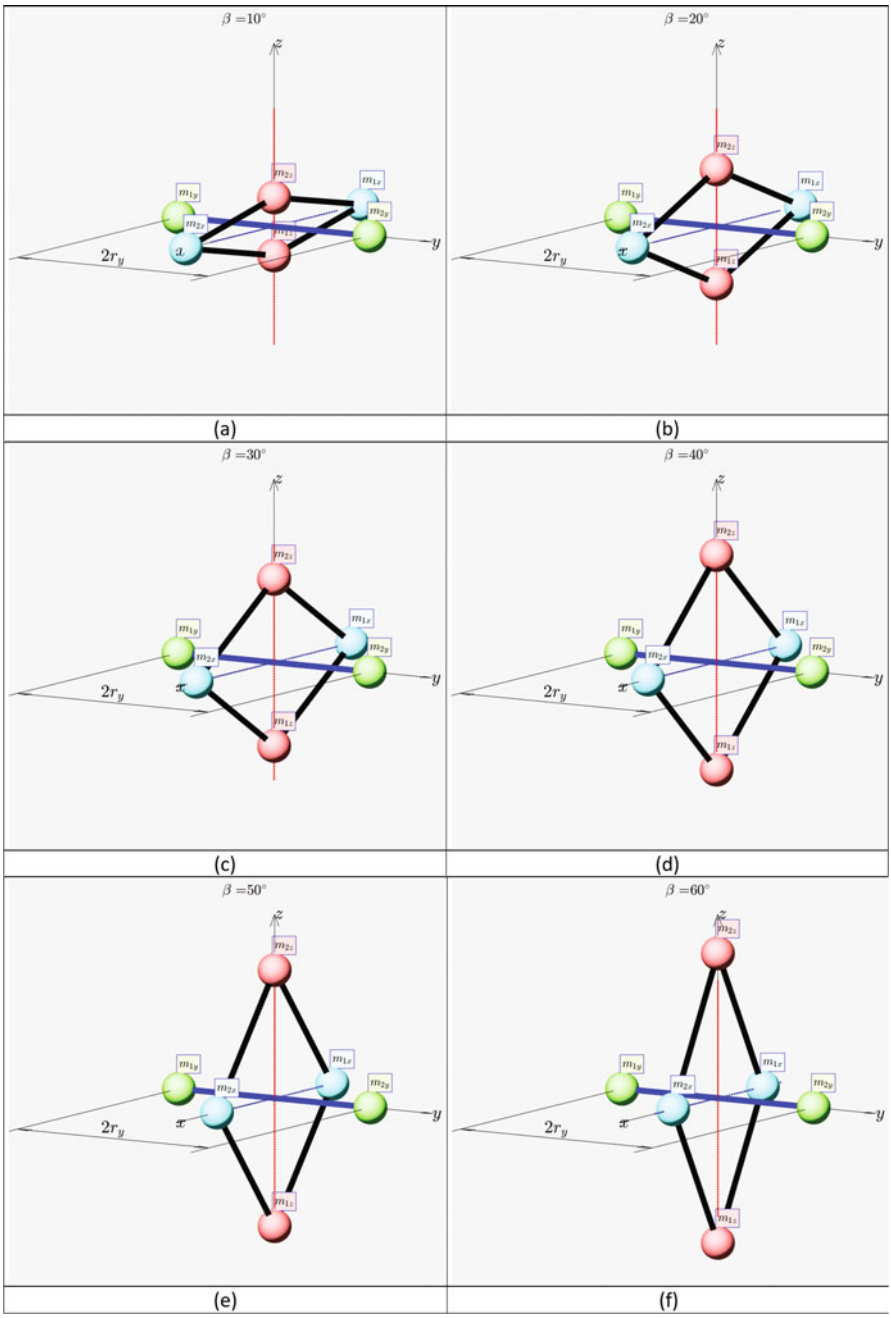


Fig. 5.89 Illustration of the "*XZ* rhombus" mechanism, employed to perform controlled inertial morphing of the system: (**a**)–(**h**) correspond to the following angles between the m_x-m_z links and the "x" axis: $\beta = 10°, 20°, 30°, 40°, 50°, 60°, 70°, 80°$

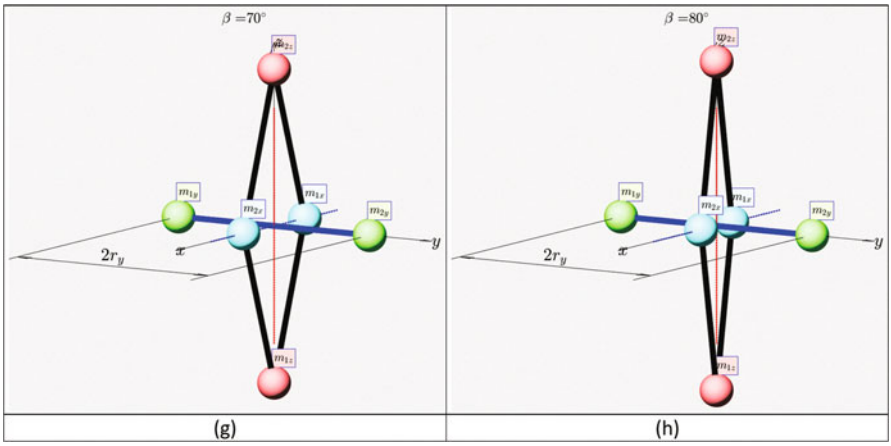


Fig. 5.89 (continued)

illustration example, transition from stable spin to unstable flipping motion can be independently triggered by any of the two system's actuators or by combined action of two actuators.

5.16.4 Example-4 Design: Two Cylinders System

Figure 5.91 shows an example of using the *two cylinders* mechanism for performing changes in the inertial properties of the system. In this conceptual design, two cylinders of length L are synchronously translated in opposite directions to keep symmetry. Initial position of the cylinders is shown with green colour, and an example of the intermediate position with outwards translations is shown with yellow colour. Direction of morphing translations for cylinders is shown with equal size thick black arrows. It enables that the system remains dynamically balanced about any of the body axes. Figure 5.91 presents configuration of the same system, but with different positions of the cylindrical masses along y axis, which can be controlled with a single actuator, which will be called r_y-actuator (not shown in the figure), ensuring the synchronously symmetric repositioning r_y of the y masses.

Variation of the principal moments of the inertia of the system I_{xx}, I_{yy} and I_{zz} with one controlled parameter, r_y, can be described with the following analytical relationships:

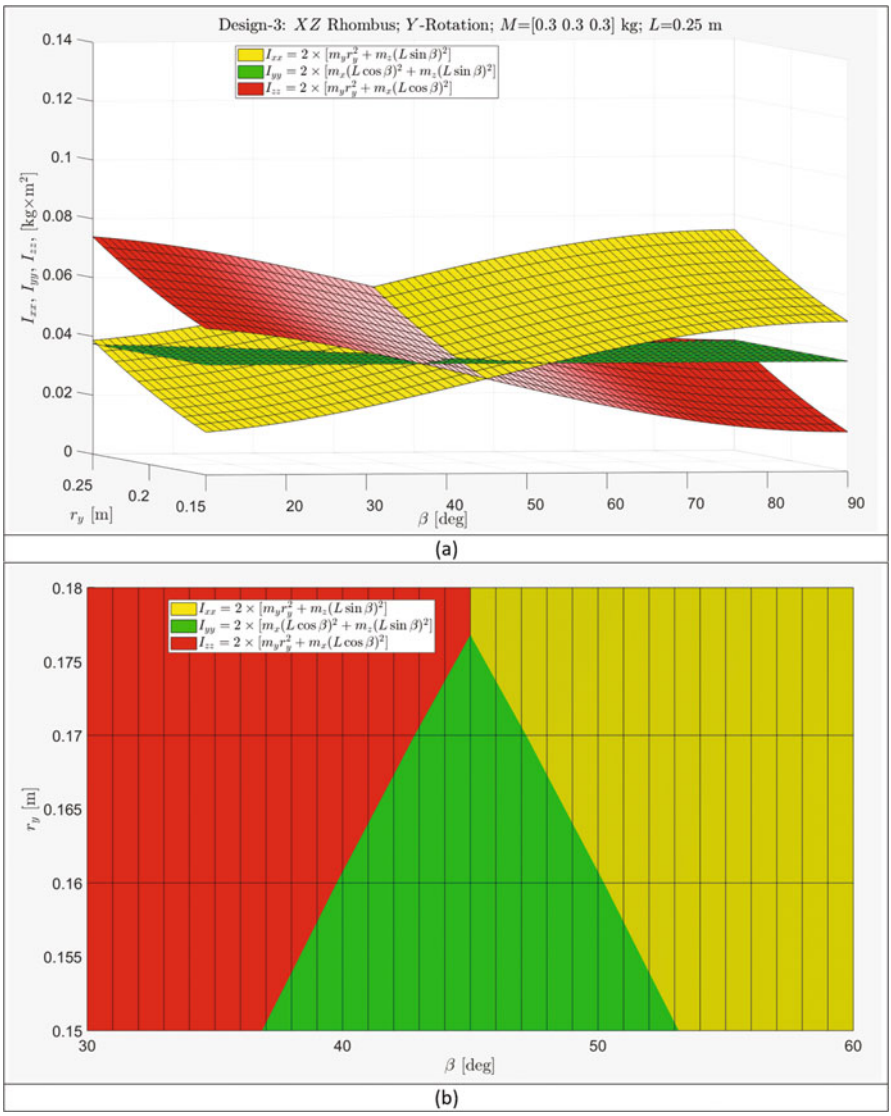


Fig. 5.90 Principal moments of inertia of the system in Fig. 5.89 as functions of β and r_y variables, which can be controlled individually by two actuators: (**a**) 3D representation and (**b**) 2D representation

$$
\begin{aligned}
I_{xx} &= 2 \times \left(\frac{M\ L^2}{12} + M\ {r_y}^2 \right) = I_{yy} + I_{zz} \\
I_{yy} &= 2 \times \left(M\ {r_y}^2 \right) \\
I_{zz} &= 2 \times \left(\frac{M\ L^2}{12} \right)
\end{aligned}
\tag{5.68}
$$

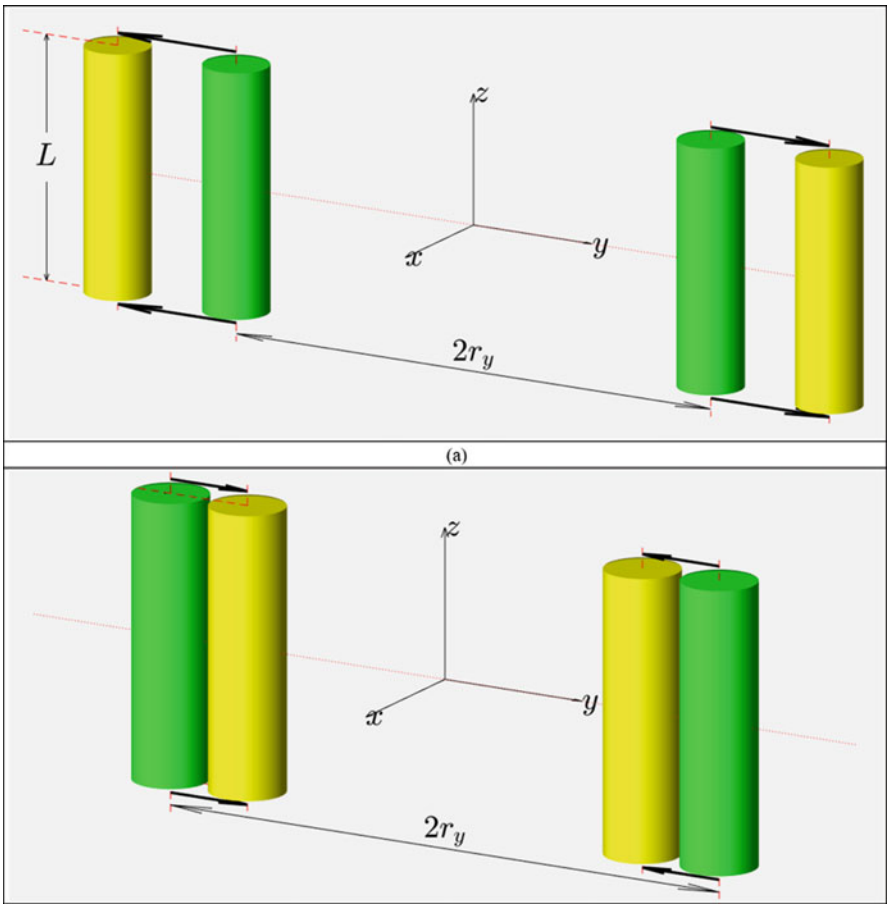


Fig. 5.91 Illustration of the *two cylinders* conceptual mechanism, employed to perform controlled inertial morphing of the system: (**a**) cylinders are moved outwards; (**b**) cylinders are moved inwards

These can be illustrated graphically as 3D surfaces. To be specific, three surfaces, corresponding to I_{xx}, I_{yy} and I_{zz}, are plotted in Fig. 5.92 for the example, given by the following parameters: mass of each cylinders $M = 0.9$ kg – the length L is varied within 0.1–0.3 m range; and the position of the cylinder parameter is varied within 0.1–0.3 m range.

Figure 5.92 shows that the limited options exist for transition to flipping motion, if the initial predominant stable spin of the system is arranged about y or z axes. The feasible combination of L and r_y parameters corresponds to the areas, where I_{yy} and I_{zz} surfaces are changing their bottom-to-top order in the Fig. 5.92.

However, as it can be seen from Eq. (5.62) and from Fig. 5.92, surface I_{xx} is always staying above any of the I_{yy} or I_{zz} surfaces. Therefore, if there is a goal to use the mechanism for transition from stable rotation to unstable flipping motion, this

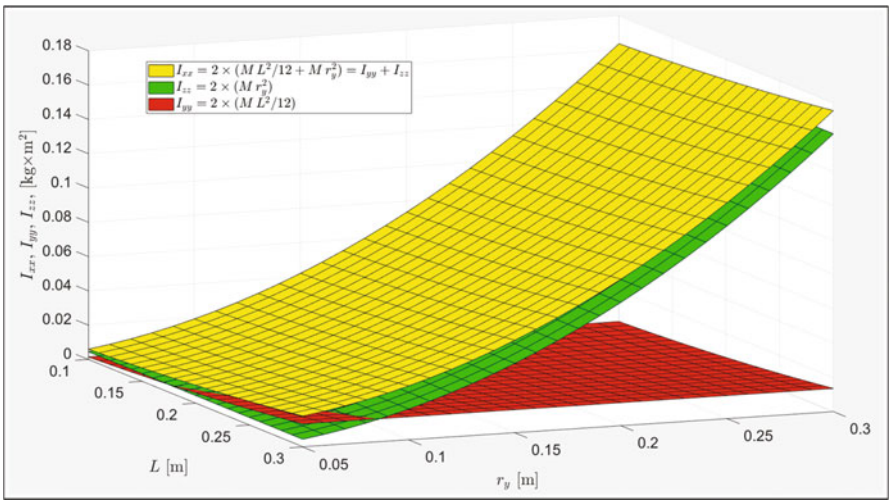


Fig. 5.92 Principal moments of inertia in the "two cylinders" model, as functions of the length of the cylinder L and its distance from the y-axis r_y

mechanism would be inefficient for this intended change, if predominant rotation of the system is selected about x axis.

As Fig. 5.92 shows, at least for large values of a_y (0.1–0.3 m), change of the distances for these cylinders from the axis of rotation is not an effective mean for changing the "status" of the axis of rotation. With this model, the intermediate axis will stay intermediate axis, regardless of the controlled distance. [The green surface is always staying between the red and yellow surfaces.]

5.16.5 Suggestions on Some Practical Implementation of the Inertial Morphing

This section does not aim to present comprehensive collection of the implementation of the methods to control the principal moments of inertia of the spacecraft, which we called "inertial morphing". Nevertheless, for completeness, we wish to present just a few of the methods/concepts, being considered as promising for realization in real spacecraft systems. In these examples, for simplicity of the illustrations, the conceptual model of the spacecraft (Fig. 5.25) will be used. As equations for the moments of inertia are functions of the distances r and masses m, conceptually, there could be two main approaches to the implementation of the inertial morphing: (a) based on variation of r - positions of the masses and (b) based on variation of masses. These approaches are briefly explained below.

(a) The *first approach* to implementation of the inertial morphing is based on the *controlled reposition of the spacecraft masses*, using actuators. Let us consider the following case: $m_x = m_y = m_z = 1$ kg; these masses are initially located at their radii: $r_x = 0.8$ m; $r_y = 1$ m; $r_z = 1.2$ m. Let us assume that the system is equipped with linear actuator (motor and appropriate mechanical system), capable of translational repositioning of the masses m_z via changing the length of r_z from 1.2 to 0.6 m within 1 s. The morphing process is shown in Fig. 5.93a, where initial positions of the masses are shown with white spheres and the final positions with black spheres and where direction of the translations for two m_z masses is shown with two red arrows. Also, for better perception of the 3D design, a semi-transparent yz plane is added to the Fig. 5.93a.

In this example, the positions of the masses on the x and y axes remain unchanged; only r_z is subject to variation (as per Fig. 5.93b). Equation (5.65) permits the calculation of the associated resulting time history of the principal moments of inertia of the system. Figure 5.93c shows that while I_{zz} keeps its value unchanged, during morphing, the I_{xx} value is changing from 4.88 to 2.72 [kg $\times$ m^2] and the I_{yy} is changing from 4.16 to 2.0 [kg $\times$ m^2]. However, most significant in the context of this work is an observation that, in this example, the role of the intermediate moment of inertia (which initially "belongs" to I_{yy}) is "passed" from I_{yy} to I_{zz} (at $t = 0.33$ s) and then is further "passed" to I_{xx} (at $t = 0.67$ s). Consequently, using only one variable r_z in the morphing process, it was possible to arrange for each of the spacecraft axes x, y and z at different stages, to become the *intermediate axis* of rotation.

The method, presented above, can be extended to the actuation of all masses, including m_x and m_y. With this general arrangement, the morphing would permit continuous control of the position of all masses, hence enabling assignment of any arbitrary values to the principal moments of inertia of the system, as per requirements of the morphing scenarios. Of course, these assignments should be compatible with the mechanical/electrical/thermal constrains of the particular designs/implementations of the morphing systems.

A variation of the same method may involve application of the special actuators to reposition large segments of the spacecraft. This idea is illustrated with the controlled change of the angular positions θ_i of the solar panels to manipulate the inertial properties of the spacecraft (see Fig. 5.94).

We envisage that similar implementations of the illustrated principle can be achieved in some other ways. For example, deployment of the masses, to the new destination (in any, inwards, outwards or inclined directions), can be ensured via unconstraining the pre-compressed springs, as per Fig. 5.95. This, however, would permit only a single discrete actuation. Nevertheless, it can be a good choice for the small autonomous system, where the control actions should be minimized and simplified.

Alternatively, for continuous actuation, instead of using solid masses, *heavy liquids* and/or *liquid metals* [41] can be used, which could be controlled via

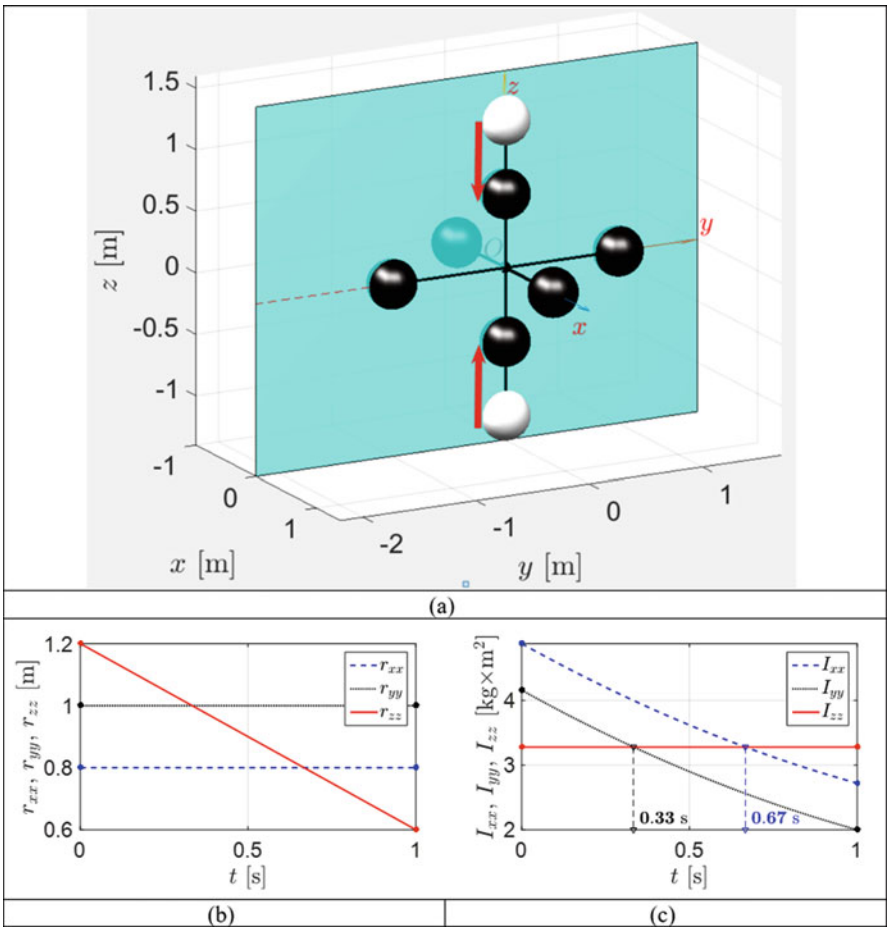


Fig. 5.93 Particular example of inertial morphing via translational reposition of the z dumbbell masses m_z while keeping positions of the x and y masses unchanged ($m_x = m_y = m_z = 1$ kg): (**a**) 3D view of the spacecraft model, (**b**) time history of the position of the masses and (**c**) time history of the resulting principal moments of inertia I_{xx}, I_{yy} and I_{zz}

manipulation with valves and employment of the passive inertial forces and/or controlled magnetic field forces to move these liquid media.

(b) The *second approach* to implementation of the inertial morphing is based on the controlled change of the spacecraft masses and may involve, for example, mass ejection, ablation, evaporation or solidification of the components of the structure, etc.

Geometric reconfigurations of the spacecraft systems (e.g. during deployment of the inflatable components or solar panels and reorientation of the antennae) are a widely used concept and proved to be successful for many space systems

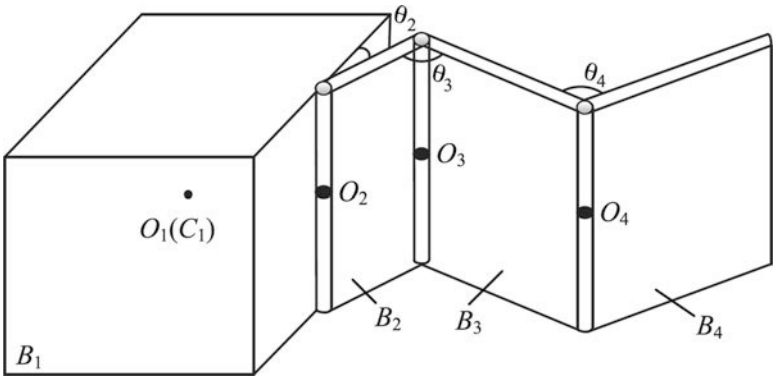


Fig. 5.94 Spacecraft, deploying solar arrays [40]

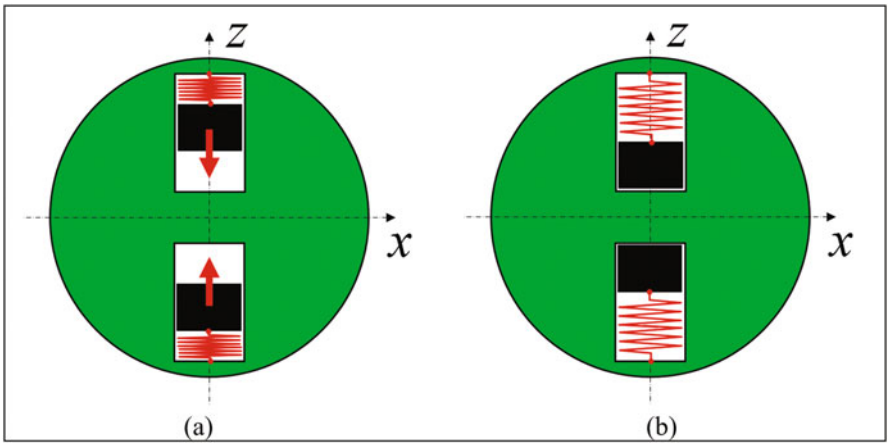


Fig. 5.95 Particular example of inertial morphing via translational reposition of the "z" dumbbell masses m_z (shown with black colour) via release of the pre-constrained compressed springs (shown with red colour): (**a**) initial configuration and (**b**) masses deployed inwards

(e.g. Spartan-207, Hughes/Boeing HS-376, SMART-1 and RAE-B satellites, space probes Rosetta and Down, etc.). However, these are provided to ensure the functionality of the spacecraft, without any relevance to the attitude dynamics objectives [42]. In contrast, concept of spacecraft reconfiguration, explicitly aiming to assist in attitude manoeuvering [3–10], is a very new concept. Indeed, idea of the reconfigurable spacecraft systems, transformable spacecraft, which consist of multiple modules connected with each other by hinges or universal joints, proposed by JAXA [43, 44], is only 3 years old, and further development of the concept is underway [45].

It is believed that this chapter will further contribute to much wider application of the spacecraft reconfigurations with the primary goal to enhance the attitude dynamics capabilities of spacecraft.

5.17 Conclusions

This book chapter presented development of a series of conceptual controlled scenarios, useful for real life of small spacecraft missions. It explored the possible applications of the "inertial morphing" concept, together with the method of "installing into polhodes and separatrices" for the possible applications in efficient spacecraft missions, involving small, low-weight and low-cost autonomous spacecraft with acrobatic capabilities.

We are hopeful that the proposed efficient (few actions) control methods, using inertial morphing, will pave the way for the design of real autonomous spacecraft without using conventional gyroscopic devices or enhance the traditional gyroscopic systems, adding to their operational capabilities and agility or saving energy.

The "inertial morphing" offers a general control platform for spacecraft attitude dynamics manipulations. However, in view of the immediate readiness for the practical application due to simplicity, the 180° and 90° inversions and de-tumbling of the spacecraft should be specially mentioned.

It is believed that the remarkable simplicity of the "OFF" "one control action" method of stopping flipping motion of the systems, rotating about its intermediate axis ("Garriott's-Dzhanibekov's effect"), and the "ON" "one control action" activation of the flipping motion of the systems, being initially set into the regular stable spin about axes with maximum or minimum principal moments of inertia, can be very attractive for the autonomous microsatellites and deep space exploration apparatus. Combination of these results in the "two control actions" method for the 180° inversion of the spinning spacecraft, with additional ability to scale the final rotational speed. This enables to use single thruster for boost and braking. This may be also useful for protection of the most fragile on-board equipment by exposing mostly protected surface towards the environmental hazards (asteroids, radiation, etc.), hence contributing to the prolonged survivability of the spacecraft and adding to the probability of success of the missions with high environmental risks.

Similarly, combination of the "OFF" and "ON" single control actions with the proposed method of "Installing (or, using other word, Inserting) into Separatrix" leads to the "three control actions" method of 90° inversion of the spacecraft.

The third, very important application of the inertial morphing is in combination of the "Installing (or Inserting) into Polhode" or "Installing into Separatrix" ("single control action" each), and "OFF" method allows de-tumbling of the spacecraft from arbitrary spatial tumbling to the regular spin about any of the body axes, selected by the space mission operator. If future spacecraft are designed with the inertial morphing capabilities, de-tumbling manoeuvre, using only "two" or "three control actions", can be useful for facilitating the capture of the spacecraft for servicing or

removal of the spacecraft after its service lifespan. This feature can also be useful for changing axis of rotation to more stable option for the long station keeping.

In summary, the list of the discovered and proposed novel capabilities, offered in spacecraft design by inertial morphing, includes but is not limited to the following:

1. Ability to switch on and off the "Garriott-Dzhanibekov effect", using meticulous inertial morphings
2. Peculiar exposure of the sides of the flipping object to the global axis directions, illustrated with the "ball of wool" method
3. Ability to control flipping periods within the wide range and discovery of the 3D ridge of the parameters with high periods and two valleys with minimum periods, using the proposed "instalment into separatrices" method and "pole-to-pole" transfers
4. Ability to perform 180° inversions with two discrete control actions and ability to perform 90° inversions with three discrete control actions
5. Ability to effectively manipulate with arbitrarily rotational motions of the tumbling systems, using the proposed method of "polhode-to-polhode transfer"
6. Ability to de-tumble spacecraft using combination of "polhode-to-polhode-to-separatrix transfer" or "polhode-to-separatrix" and "instalment into separatrix"
7. Ability to reverse the direction of the slide of the tip of the non-dimensional angular momentum vector, employing "prime" and "dual" separatrices
8. Ability to perform conjugated "parade" of all body axes inversions
9. Ability to enhancement of attitude control of the spacecraft with traditional reaction wheels via added inertial morphing

The novelty of the presented materials, based on [3–9], includes the following:

- Utilisation of deliberately applied changed to the inertial properties of the system for its attitude control and enhancement of the spinning spacecraft attitude dynamics capabilities using inertial morphing
- Generalised geometric interpretation of arbitrary attitude motions of the morphed systems, using non-dimensional angular momentum coordinates, angular momentum sphere and kinetic energy ellipsoid
- Determination of the required IM changes and their phasing
- Development of the methods and mechanisms for inertial morphing
- Development of a general method of de-tumbling of the spacecraft and transition of the rotational motion to the rotation about any nominated body axis
- Reduction of continuous control actions to a limited number of discrete control actions (e.g. two only inertial morphings for 180° inversion and three only morphings for 90° inversion)
- Discovery of the limited exposures of tumbling and flipping bodies to specific orientation
- Establishment of the ranges of the flipping periods
- Suggestion of new spacecraft designs, enabling agile space acrobatic or prolonged attitude capabilities
- Suggestion of new space applications

- Proposition of applications of inertial morphing for the possible future space missions
- Demonstration of the proposed concepts and methods, using versatile advanced non-linear numerical simulations and virtual reality

As Euler's equations paved the way for the development of the theory of gyroscopes and design of various gyroscopic systems, the paradigm of "inertial morphing" may prompt development of new generation of the acrobatic spacecraft with significantly reduced weight and dimensions, reduced cost and enhanced operational capabilities. It may be also possible to design new classes of gyroscopes, possessing an added-on sense of time, which is in contrast to the classical gyroscopes that only possess a sense of orientation.

With a wide spectrum of the presented examples, related to the application of a novel design concept of "inertial morphing", it is believed that presented concept, modelling and simulation of the spinning systems and attitude control method of the spinning systems will be useful not only for the specialists but for a very wide audience, including engineers, scientists, students and enthusiasts of science and space technology.

Appendixes

Nomenclature

ψ, θ, ϕ	Euler angles
$\omega_x, \omega_y, \omega_z$	Components of the angular velocity
$\omega_{x,i}, \omega_{y,i}, \omega_{z,i}$	Initial components of the angular velocity
a_x, a_y, a_z	Values of the semi-major axes of the ellipsoid of the kinetic energy
α	Angle between the plane of the separatrix and axis with maximum moment of inertia
β	Masses position angle in the "scissors" and "rhombus" mechanisms
G	Centre of the mass of the spacecraft
$\vec{\mathbf{H}}(t)$	Angular momentum vector
$H = \mid \vec{\mathbf{H}} \mid$	Value of the system's angular moment
$\overline{H}_x, \overline{H}_y, \overline{H}_z$	Non-dimensionalised components of the angular momentum vector
I_{xx}, I_{yy}, I_{zz}	Principal moments of inertia
k	Parameter in complete elliptic integral k
$K(k)$	Complete elliptic integral of the first kind
K_0	Kinetic energy of the system
l	Angular momentum vector of wheels, expressed in the body-fixed reference frame
m_x, m_y, m_z	Dumbbell masses in six-mass spacecraft model
n_{ω_i}	Control torque applied to the i-th wheel
M	Mass matrix
N_x, N_y, N_z	Torque components
P	Pivot point
P, Q, R	Torque components in original Euler's work
r_x, r_y, r_z	Axial positions of the spacecraft masses in the "six-mass" model
t	Time
T	Period of the flipping motion
x, y, z	Principal body axes of the rigid body
$\mathbf{x}$	Vector of system's states

Acronyms/Abbreviations

ISS	International Space Station
AMS	Angular momentum sphere
KEE	Kinetic energy ellipsoid
IM	Inertial morphing

References

1. Learn Engineering (2019), "How do Satellites work?" ICT #10. – Jul 31, 2019. https://www.youtube.com/watch?v=ror4P1UAv_g (Last accessed on 21 Mar 2021).
2. Selva, D. and Kreci, D. (2012), – A survey and assessment of the capabilities of Cubesats for Earth observation‖, *Acta Astronautica*, 74, pp. 55–68.
3. Trivailo, P.M. and Kojima, H. (2017). "Utilisation of the "Dzhanibekov's Effect" for the Possible Future Space Missions". Paper ISTS-2017-d-047/ISSFD-2017-047. – Proc. of the Joint Conference: 31st ISTS, 26th ISSFD & 8th NSAT, Matsuyama, Japan, 3–9 June 2017. – 10 pp.
4. Trivailo, P.M. and Kojima, H. (2017), "Re-Discovering "Dzhanibekov's Effect" Using Non-Linear Dynamics and Virtual Reality". Paper IAC-17,E1,7,10,x41083. - *Proceedings of the 68th International Astronautical Congress (IAC)*, Adelaide, Australia, 25–29 Sept. 2017. IAC, Vol. 17, pp. 11394–11407 (14 pp).
5. Trivailo, P.M. and Kojima, H. (2018). "Augmented Control of Inversion of the Spinning Spacecraft, using Inertial Morphing", Paper IAC-18,C2,3,5,x45333. – *Proceedings of the 69th International Astronautical Congress*, Bremen, Germany, 01–05 October 2018. 16 pp.
6. Trivailo, P.M., Kojima, H. (2019), "Discovering Method of Control of the "Dzhanibekov's Effect" and Proposing its Applications for the Possible Future Space Missions". – *Transaction of JSASS (The Japan Society for Aeronautical and Space Sciences), Aerospace Technology Japan*, 2019, Vol. 17, No. 1, pp. 72–81. DOI: https://doi.org/10.2322/tastj.17.72.
7. Trivailo, P.M., Kojima, H. (2019), "Enhancement of the Spacecraft Attitude Dynamics Capabilities via Combination of the Inertial Morphing and Reaction Wheels (Keynote Paper and Presentation)". [Peer Reviewed]. - *Transactions of the 18th Australian International Aerospace Congress (AIAC2019), incorporating 27th International Symposium on Space Flight Dynamics (ISSFD)*. Engineers Australia, Royal Aeronautical Society. ISBN: 978-1-925627-21-3. Pp.1093–1122.
8. Trivailo, P. M, and Kojima, H. (2019), "Enhancement of the Spinning Spacecraft Attitude Dynamics Capabilities using Inertial Morphing," Royal Aeronautical Society *The Aeronautical Journal*, v.12, 2019, pp.1–34. DOI: doi:https://doi.org/10.1017/aer.2019.145.
9. Trivailo, P.M., Rittweger, A., Theil, S. (2020), "Utilisation of the Controllable Inertial Morphing for Providing Spacecraft with Acrobatic Attitude Capabilities". – *Proceedings of the IFAC (International Federation of Automatic Control) World Congress-2020 (IFAC2020)*, Berlin, 12-17 July 2020. – 6pp [pp 15105–15110 in Preprints of the 21st IFAC World Congress (Virtual), Berlin, Germany, July 12–17, 2020]. Paper and Video Presentation #2985. Animation in Virtual reality and video presentation: https://youtu.be/wAWjWUuu7z0 (accessed 21 Mar 2021).
10. Trivailo, P.M. and Kojima, H. (2021), De-tumbling of the Spacecraft with Inertial Morphing Capabilities. *IEEE Access*. (Submitted in April 2021), 10 pp.
11. Dzhanibekov, V.A. (2010), Interview at the "Secret Signs" TV Program, https://youtu.be/dL6Pt1O_gSE (published Feb 19, 2010, accessed Mar 21, 2021).
12. Roskosmos Telestudio (2012), "Dzhanibekov's Effect", https://youtu.be/6ozUSgBjeaQ (published on May 15, 2012; accessed Mar 21, 2021).
13. Owen Kay Garriott, O.K. (1974), "Physics: Concedrvation Laws in Zero Gravity" NASA Skylab Science Demonstration Program, https://youtu.be/xdtqVR1CgQg?t=1018 (accessed 19/02/2019).
14. Plasma, B. (2009), "Dancing T-handle in zero-g, HD", https://www.youtube.com/watch?v=1n-HMSCDYtM, (published 03 Mar 2009; accessed 23 Mar 2021; 1,143,240 views).
15. Glatzmaier, G.A.; Coe, R.S. (2015), "Magnetic Polarity Reversals in the Core", Treatise on Geophysics, Elsevier, pp. 279–295, doi:https://doi.org/10.1016/b978-0-444-53802-4.00146-9, ISBN 978-0444538031.
16. Muller, D. (2019), "The Bizarre Behavior of Rotating Bodies, Explained", "*Veritasium*" channel, https://www.youtube.com/watch?v=1VPfZ_XzisU, shown on 20 Sept 2019 (accessed on 23 Mar 2021; 7,491,818 views).

17. Ashbaugh, M. S., Chicone, C. C., and Cushman, R. H. (1991), The twisting tennis racket. *Journal of Dynamics and Differential Equations*, Vol. 3, Issue 1, 1991, pp. 67–85. DOI: https://doi.org/10.1007/BF01049489.
18. Euler, L. (1758), "Du mouvement de rotation des corps solides autour d'un axe variable," M'emoires de l'academie des sciences de Berlin, Vol. 14, Berlin Academy, Berlin, Germany, 1758, pp. 154193.
19. Rimrott F.P.J. (2011), Introductory Attitude Dynamics. Springer-Verlag, 2011, 383pp.
20. Beachley, N.N. (1971), "Inversion of spin-stabilized spacecraft by mass translation – some practical aspects," Journal of Spacecraft, 8, 1971, pp. 1078–1080.
21. Murakami, H., Rios, O. and Impelluso, T.J. (2016), "A theoretical and numerical study of the Dzhanibekov and tennis racket phenomena," *Journal Applied Mechanics*, 83, (Sept. 08, 2016), (11), 111006 (10 pages). Paper No: JAM-16-1017.
22. Moler, C., (2015), "Tumbling Box ODE. (Cleve's Corner: Cleve Moler on Mathematics and Computing)," Posted by Cleve Moler, August 10, 2015. http://blogs.mathworks.com/cleve/2015/08/10/tumbling-box-ode/ (accessed 23 Mar, 2021).
23. Garriott, R. (2009), "Richard Garriott Space Video Blog: Rotational Inertia". https://www.youtube.com/watch?v=fPI-rSwAQNg, (published on 12 Mar 2009; accessed 23 Mar 2021; 100,251 views).
24. Shkaplerov, A., and Burbank, D. (2013), Experiments on board of the International Space Station (Expedition #30, 2011). https://youtu.be/LzVItPwiQyI (published Feb 11, 2013; accessed 23 Mar, 2021; 333,302 views).
25. **Ford**, **K.** (2013), Working with the Spheres Satellites On-board of the ISS (Expedition #34, 2013) https://youtu.be/dsXOxcDSBLQ?t=108 (published Mar 13, 2013; accessed Sept 07, 2017).
26. Wakata, K. (2013), Mysteries of Rotational Movement (On-board of the ISS, Expedition #38, 2014), JAXA. https://youtu.be/QhSN2eua14I?t=309 (published Nov 25, 2013; accessed 23 Mar 2021; 56,642 views).
27. Feynman, Richard (June 1970). "Chapter 22: Algebra". The Feynman Lectures on Physics: Volume I. p. 10.
28. Wells, David (1990). "Are these the most beautiful?". *Mathematical Intelligencer*, 12 (3): 37–41. doi:https://doi.org/10.1007/BF03024015.
29. Portrait of L. Euler. University of Tartu collection, http://dspace.ut.ee/handle/10062/22581?locale-attribute=en, (accessed 23 Mar 2021).
30. The Euler Archive, http://eulerarchive.maa.org/ (accessed 23 Mar 2021).
31. Marsden, J. E. and Ratiu, T. S. (1999), Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems. Springer, 1999. – 693pp. ISBN-13: 978-0387986432, ISBN-10: 038798643X.
32. Ono, T. (2017), A Comprehensive Study of the Dzhanibekov and Tennis Racket Phenomena, UC San Diego Master of Sciences Theses, 2017.
33. Landau, L.D. and Lifshitz, E.M. (1988), *Mechanics*. Vol.1 (In Russian). – 4th ed. Moscow, "Nauka". 1988. – 216 pp.
34. Dorrington, G., Trivailo, P.M. (2019), Utilising the 'Chaotic' Tumbling of CubeSats. [Peer Reviewed]. – *Transactions of the 18th Australian International Aerospace Congress (AIAC2019), incorporating 27th International Symposium on Space Flight Dynamics (ISSFD)*. Engineers Australia, Royal Aeronautical Society. Date of Publication. 24/02/2019. ISBN: 978-1-925627-21-3. Pp.1390–1412.
35. Bedford, A. M. and Fowler, W. (2008), Engineering Mechanics: Dynamics, 5th ed., Pearson, 2008. – 652 pp.
36. Physics Demo (2016), "Angular Momentum Demo: Hoberman Sphere", https://www.youtube.com/watch?v=64t-dVtDwkQ (published August 20, 2016; accessed 31 March 2021; 195,090 views)
37. Oliver, O. (2015), "Fastest ice skating spin", Guinness Book of World Records, 19 January 2015, https://www.guinnessworldrecords.com/world-records/fastest-spin-ice-skating (accessed 31 March 2021).

38. Hanavan E. Jr. (1964), A Mathematical Model of the Human Body. AMRL-TR-64-102. AMRL-TR. Aerospace Medical Research Laboratories (U.S.). 1964 Oct:1–149.
39. Terui, F., Kimura, S., Nagai, Y., Yamamoto, H., Yoshihara, K., Yamamoto, T., Nakasuka, S. (2005), Moon Tracking Attitude Control Experiment of a Bias Momentum Micro Satellite "μ-LabSat", *Trans. Japan Soc. Aero. Space Sci.*, Vol. 48, No. 159, 2005, pp. 28–33.
40. Yao, Q. Ge, X. (2017), Optimal Control of Stretching Process of Flexible Solar Arrays on Spacecraft based on a Hybrid Optimization Strategy, *Theoretical & Applied Mechanics Letters*, Vol. 7, 2017, pp. 258–263.
41. Noack, D. (2019), In-Orbit Verification of a Fluid-Dynamic Attitude Control System. Paper 2019-d-041. – *Proc. of the 32nd ISTS & 9th NSAT (The International Symposium on Space Technology and Science and Nano-Satellite Symposium)*, Fukui, Japan, 15–21 June 2019.
42. Hwang, J.T., Lee, D.Y., Cutler, J.W. and Martins, J.R.R.A. (2014), Large-Scale Multidisciplinary Optimization of a Small Satellite's Design and Operation. – *Journal of Spacecraft and Rockets*, Vol. 51, No. 5, September–October 2014, pp. 1648–1663. DOI: https://doi.org/10.2514/1.A32751.
43. Ohashi, K. Chujo, T. and Kawaguchi, J. (2018), "Motion Planning in Attitude Maneuver Using Non-Holonomic Turns for a Transformable Spacecraft," *The 28th JAXA Workshop on Astrodynamics and Flight Mechanics*, Sagamihara (Japan), 2018. – 6pp.
44. Hernando-Ayuso, J., Baresi, N., and Chujo, T. (2018), "Orbit Design and Insertion for the JAXA Transformable Spacecraft," *The 28th JAXA Workshop on Astrodynamics and Flight Mechanics*, Sagamihara (Japan), Paper B-26, 2018. – 8 pp.
45. Kubo, Y. and Kawaguchi, J. (2020), Analytical Formulation on Non-Holonomic Attitude Motion of Transformer spacecraft," *The 30thJAXA Workshop on Astrodynamics and Flight Mechanics*, Sagamihara (Japan), Paper C-18, 21st of July 2020. – 8 pp.

Chapter 6
A New Strategy for Form Finding and Optimal Design of Space Cable Network Structures

Sichen Yuan and Bingen Yang

Abstract Cable network structures, which are a class of nonlinear flexible structures, have been widely used in infrastructures and spacecrafts. In this work, a new form-finding method, namely, the fixed nodal position method (FNPM), is developed for optimal design of geometric configuration and internal force distribution for cable network structures, to meet the operation requirement of high shape/surface accuracy. Different from conventional methods, which usually adopts a stress-first-and-displacement-later procedure in form finding, the FNPM first assigns nodal coordinates for a cable network structure and then determines the internal force distribution of the structure by a nonlinear optimization process. The highlight of the FNPM is that the prescribed nodal coordinates are unchanged during the form-finding process. This unique feature of fixed nodal positions makes it possible to place the nodes of a cable network structure at desired locations, satisfying complicated structural constrains and yielding high shape/surface accuracy as required. As another advantage, the FNPM in form finding undertakes the assignment of geometric configuration (nodal coordinates) and the determination of internal force distribution separately. This translates into significant savings in computational effort, compared with conventional form-finding methods. The new form-finding method is applied to the optimal design of a large deployable mesh reflector of 865 nodes.

Keywords Force density method · Dynamic relaxation method · Form-Finding method · Optimal design · Space cable network

F. E. Udwadia (✉)
Departments of Aerospace and Mechanical Engineering, Civil and Environmental Engineering, Mathematics, Information and Operations Management, University of Southern California, Los Angeles, CA, USA

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_6

6.1 Introduction

Cable network structures have been of great R&D interest for years. Due to lightweight, foldability, and high stiffness, this type of flexible structures has been widely used in infrastructures and spacecrafts. Examples include the Olympiastadion in Munich with its roof designed as a hanging cable net [1] and space deployable mesh reflectors [2, 3]. In application, a cable network structure is usually required to achieve certain level of geometric or surface accuracy for special functions. For instance, a cable dome needs to retain a desired shape to avoid wrinkling of cover membranes. As another example, a space deployable mesh reflector must maintain a spherical or parabolic shape for signal transmission. In design and operation of a cable network structure, it is essential to maintain a desired shape/surface. Loss of shape/surface accuracy often leads to malfunction or degraded performance of the structure.

One key step in design of a high-precision lightweight cable network structure is *form finding*, which is a process for determination of the geometric configuration and internal force distribution among the members of the structure. In this chapter, the properties of cable network structures are first discussed, and the concepts of form finding, which include the determination of the initial equilibrium configuration and the deformed equilibrium configuration for such a structure in design, are then presented. Afterwards, as the core content of this chapter, a new form-finding method, named the fixed nodal position method (FNPM), is introduced for design of high-precision lightweight space cable network structures. For comparison purposes, two other form-finding methods, the force density method and the dynamic relaxation method, are also described. Also, an FNPM-based optimal structure design procedure is outlined. The advantages of the FNPM in form finding of cable network structures are demonstrated in several demonstrative examples, including a large deployable mesh reflector of 865 nodes.

6.2 Problem Statement

In this section, design of high-precision lightweight space cable network structures for operation under various conditions is considered. The properties of this kind of structures are reviewed, and relevant design requirements are discussed.

The objectives of form finding for a cable network structure in general are to find a geometric configuration described by the nodal positions of the structure and to determine an internal force distribution among the cable members. For clarity of discussion, the following four assumptions about a cable network structure are made:

(A1) Cable members of the structure are connected by pin-joints.
(A2) The cable members can only sustain axial tension forces.

(A3) A level of pre-stress (pretension) is required to stiffen the structure and to avoid slacking of the cable members.
(A4) Boundary nodes of the cable network structure are mounted on a stiff supporting structure. These boundary nodes are treated as fixed nodes.

Form-finding problems may vary significantly, depending on the given information about a cable network structure in consideration. Note that the structure is prestressed according to assumption A3. Form-finding problems for prestressed cable network structures (assumption A3) are of the following two types [4].

(i) Form finding of *initial* equilibrium configuration (FF-IEC): Given a topology of the cable network structure and external forces, determine nodal positions and member internal forces of the structure.
(ii) Form finding of *deformed* equilibrium configuration (FF-DEC): Given a topology of the cable network structure, external forces, and undeformed lengths of cable members, determine nodal positions and member internal forces of the structure.

One major difference between these two types of problems is that the FF-IEC does not require knowledge on undeformed member lengths. In this chapter, FF-IEC problems are first investigated; FF-DEC problems are then addressed.

For FF-IEC of a cable network structure, the equilibrium equation is [5]

$$M\sigma = l \tag{6.1}$$

where M is the equilibrium matrix consisting of direction cosines, σ is a vector of generalized stresses, and l is a vector of generalized loads. Because the cross-section areas and elastic properties of cable members do not have to be specified in FF-IEC [6], it is convenient to replace σ by a vector of member (axial) forces. Thus, the equilibrium equation can be written as

$$MT = -F \tag{6.2}$$

where for a 3-D cable network structure, M is a $3n$ by k matrix, which is obtained from nodal coordinates, with n being the number of nodes and k the number of members; T is a k by 1 vector of member internal forces; and F is a $3n$ by 1 vector that represents the external forces applied to the cable network structure at its nodes.

6.2.1 *Four Types of Structural Assemblies and Extended Maxwell's Rule*

From matrix M in Eq. (6.2), the number m of mechanisms and the state s of self-stress of a cable network structure can be obtained [7]. From Ref. [8], the calculation and physical meaning of m and s are given in Table 6.1.

Table 6.1 Physical meanings of the number m mechanism and state s of self-stress

	Calculation	Physical meanings
m	Left null space of matrix M	Loads which cannot be equilibrated in the initial configuration Nodal displacements that cannot cause member elongations
s	Null space of matrix M	States of self-stress; incompatible member elongations

Table 6.2 Classification of structural assemblies

Assembly type	Value of s and m	Static and kinematic properties
I	$s = 0$ $m = 0$	Statically determinate and kinetically determinate
II	$s = 0$ $m > 0$	Statically determinate and kinetically indeterminate
III	$s > 0$ $m = 0$	Statically indeterminate and kinetically determinate
IV	$s > 0$ $m > 0$	Statically indeterminate and kinetically indeterminate

Table 6.3 Existence and uniqueness of solutions of the equilibrium equation (6.2)

Assembly Type	Value of s and m	Existence and uniqueness of force equilibrium equation
I	$s = 0$ $m = 0$	Eq. (6.2) has a unique solution for any given external force F
II	$s = 0$ $m > 0$	Eq. (6.2) has a unique solution for consistent external force F_c Eq. (6.2) has no solutions for inconsistent external force F_{ic}
III	$s > 0$ $m = 0$	Eq. (6.2) has infinite number of solutions for any given external force F
IV	$s > 0$ $m > 0$	Eq. (6.2) has infinite number of solutions for consistent external force F_c Eq. (6.2) has no solutions for inconsistent external force F_{ic}

According to Ref. [9], there are four types of structural assemblies, depending on the values of m and s, which are listed in Table 6.2. Also, the integers m and s directly determine the existence and uniqueness of the solutions T of Eq. (6.2), as shown in Table 6.3.

For cable network structures of types II and IV, a feasible internal force distribution is not guaranteed to be found for any given external forces, due to the existence of internal mechanisms. The external force vector F is called *consistent* external force vector (denoted by F_c) if at least one feasible internal force distribution exists

that satisfies the equilibrium equation (6.2). Otherwise, F is *inconsistent* external force vector (denoted by F_{ic}). The matrix M can be used to check the consistency of given external forces. If $-F$ lies in the column space of M, F represents consistent external forces. If $-F$ lies in the left null-space of M, F represents inconsistent external forces [9].

Based on the concept of m and s, the extended Maxwell's rule was developed as [10]

$$3n - k - c = m - s \tag{6.3}$$

where c is number of constraints. Equation (6.3) is used as the fundamental rule in design of cable network structures to avoid flexibility and ensure sufficient stiffness. For the preliminary design of a cable net, the requirement is $3n - k - c < 0$. After the form-finding process, the values of s and m should be calculated to verify the satisfaction of $s > 0$, $m = 0$, in order to avoid flexibility and guarantee certain level of pre-tension.

6.2.2 Cable Network, Tensegrity, and Truss Structures

The definitions and classifications of cable network, tensegrity, and truss structures are given as follows. Truss structures, which are called pin-jointed structures, are composed of bar and cable members. Zhang and Ohsaki [11] classified truss structures into two categories: tensile (cable) structures and tensegrity structures. A tensile (cable) structure, which is also named cable network structure, is a truss structure that consists cable members only. Tensegrity structure, introduced by Fuller [12], is a structure based on the use of isolated members in compression (usually bars or struts) inside a net of continuous members in tension (usually cables or tendons), in such a way that the compressed members do not touch each other and the prestressed tensioned members delineate the system spatially. This definition was later changed by Motro [13] as a wider form for referring to prestressed truss structures consist both bars and cables. In such definition, the bar members are allowed to contact with each other.

For a uniform cable member, with constant Young's modulus E and cross-section A, the internal force of the cable is given by

$$T = \begin{cases} EA\frac{L-L_{\text{org}}}{L_{\text{org}}} & L \geq L_{\text{org}} \\ 0 & L < L_{\text{org}} \end{cases} \tag{6.4}$$

where L and L_{org} are the current and the original member lengths. When $L < L_{\text{org}}$, the cable member is in slack stage which is usually undesirable in design of a cable network structure. To avoid the slack stage of cables, in form finding, the cable net

is usually required to be pre-tensioned. In addition, some of the boundary nodes are fixed or under constraints in certain directions.

In a tensegrity structure, cables can only be tensioned, and bars can withstand both compression and tension. Form finding of tensegrity structures has be a topic of research interest in the past [14–16]. A tensegrity structure is required to be self-stressed, which means that its state $s > 0$. During the form-finding procedure, the external loads are considered absent [17]. Generally, none of the nodes of a tensegrity structure is considered fixed. Unlike a cable net which only possesses internal mechanism, a 3-D tensegrity structure has both internal mechanisms and 6 rigid body mechanisms, which are separable by the method introduced in [18]. Due to existence of rigid body mechanisms, the requirements of mechanism and state of self-stress in topology design are $s > 0$ $m \geq 6$. Similar with the preliminary design of a cable net, by extended Maxwell's rule, $3n - k - c < 6$ is firstly checked. m and s are then calculated from the obtained form found by form-finding methods to verify that the tensegrity structure is stiff and stable.

6.2.3 Geometric and Force Constraints

Many cable network structures in engineering applications are designed with geometric constraints [19], which include axial symmetry and some specific lengths requirements for cable members. In general, geometric constraints can be described in the form

$$r_g(p) \geq 0 \tag{6.5}$$

where $r_g(p)$ is a vector of constraint functions, and p is a vector of nodal coordinates given by

$$p = \{x_1 \;\; y_1 \;\; z_1 \;\; \cdots \;\; x_n \;\; y_n \;\; z_n\}^T \tag{6.6}$$

Besides geometric constraints, constraints on internal forces can also be imposed, which is of the general form

$$r_f(T) \geq 0 \tag{6.7}$$

where T is the vector of internal member forces in Eq. (6.2), and r_f (T) is a vector of constraint functions. For instance, in a cable network structure design, the internal forces T of cable members must be tensional to avoid slacks and all the internal forces (of bars and cables) must in equilibrium. This leads to the following force constrains of mixed equality and inequality

$$MT = -F, T > 0 \tag{6.8}$$

Other forms of geometric and force constraints, of course, are possible, depending on applications, and they can be similarly described as in Eq. (6.8).

6.2.4 Desired Internal Force Distribution

In description of force-displacement relation, the stiffness matrix K of a cable network structure can be decomposed into two parts: a geometric stiffness matrix K_g and an elastic stiffness matrix K_e [17]. The K_e corresponds to the current member lengths, nodal positions, and elastic properties of the cable members, and it is independent of the level of pre-stress. The K_g, on the other hand, is related to the nodal positions and the state of pre-stress only. In other words, geometric stiffness vanishes when the cable network structure is not pre-stressed. To obtain a stiff and stable cable network structure, a certain level of pre-stress is required. As mentioned previously, in determination of an initial equilibrium configuration, stress can be represented by the internal forces of the cable members. It is therefore desired to have a specific internal force distribution T_{des} in form finding of a cable network structure.

6.3 The Fixed Nodal Position Method for FF-IEC

In form finding of high-accuracy lightweight space cable network structures, two tasks must be carried out: (task 1) to determine a geometric configuration (nodal positions) to achieve required shape/surface accuracy for a given topology and (task 2) to determine an internal force distribution among cable members for given external forces. A design method, which can successfully undertake these two tasks and, at the same time, warrants required shape/surface accuracy for operation, would be beneficial to R&D of cable network structures in many applications.

To carry on the above-mentioned tasks in form finding, a new method, which is called the fixed nodal position method (FNPM), is developed. The highlight of the FNPM is that it keeps prescribed nodes of a cable network structure unchanged in form finding, while producing a feasible internal force distribution among the structure members. This special feature of fixed nodal positions allows the FNPM to rather freely place the nodes of a cable network structure in design at designated positions, yielding a desired shape or surface with high accuracy. Furthermore, different from conventional form-finding methods, the FNPM carries out the aforementioned tasks 1 and 2 separately; that is, it first assigns desired nodal positions and then determines a viable internal force distribution. This individual execution of the tasks greatly reduces computational effort in terms of computer storage and simulation time, which renders the FNPM more capable of handling complex cable network structures.

The FNPM-based form finding takes three main steps. In the first step, with a given topology, an initial guess of the geometric configuration of the cable network structure in design is made, which complies with the prescribed geometric constraints. In the second step, an internal force distribution among all the cable members is determined by a numerical optimization algorithm. In the third step, the configuration of the cable network structure is optimally and iteratively adjusted such that the internal force distribution satisfies the prescribed force constraints. This three-step form-finding process eventually yields a cable network structure that meets all the design specifications. The steps of the FNPM are detailed in the subsequent subsections.

6.3.1 Initial Guess of Geometric Configuration

Form finding of a cable network structure by the FNPM begins with an initial guess of the geometric configuration (nodal coordinates) of the structure. An initial guess of geometric configuration, of course, depends on different cable network design requirements. An example of initial guess of the nodal coordinates of a large deployable mesh reflector is seen in Ref. [20].

6.3.2 Determination of Internal Force Distribution

In this step of the FNPM, the coordinates of all the nodes are assumed to be unchanged (fixed), which is fundamentally different from those approaches of stress first and displacement later. For different types of structure assemblies shown in Table 6.3, determination of internal force distribution is considered in the following three cases.

Case A. Cable Network Structures of Types I and II with Consistent External Forces
In this case, Eq. (6.2) admits a unique solution of internal force distribution given by

$$T = -M^{+}F_c \tag{6.9}$$

where M^{+} is the *Moore-Penrose inverse* of matrix M. With the unique solution directly obtained. Moreover, if no external forces are applied, $F_c = 0$ and the solution is trivial

$$T = 0 \tag{6.10}$$

Case B. Cable Network Structures of Types II and IV with Inconsistent External Forces

When external forces are inconsistent with the current geometric configuration ($F = F_{ic}$), which leads to an overdetermined problem [6], the solution of Eq. (6.2) does not exist. In this case, adjustment of geometric configuration is needed in a cable network design. Under the circumstances, consider the expression

$$T_{ls} = -M^{+}F_{ic} \tag{6.11}$$

According to Ref. [21], $\|MT_{ls} + F_{ic}\| \leq \|MT + F_{ic}\|$ for any given T, rendering a minimum residue force in the equilibrium equation (6.2). Thus, while it does not satisfy Eq. (6.2), T_{ls} can serve as a useful reference in geometric configuration adjustment, as shall be presented in the next section. For this reason, T_{ls} shall be called the least square internal force vector.

Case C. Cable network Structures of Types III and IV with Consistent External Forces

For a cable network structure of type III or IV with consistent external forces ($F = F_c$), Eq. (6.2) has an infinite number of solutions of internal forces. In this case, due to lack of analytical methods and for best design results, an internal force distribution T can be determined via numerical optimization, which is useful for many engineering problems [22–24]. The numerical optimization process developed for the current form-finding problem is described as follows.

Given a desired internal force distribution T_{des}, one design objective is to minimize the difference between the desired internal force distribution and the internal force distribution T that is obtained from Eq. (6.2). In such an optimization problem, the objective function has several options, such as $f(T) = \|T - T_{\text{des}}\|^2$, $f(T) = \sum_{i=1}^{m} \left|T_i - T_{\text{des},i}\right|$, and $f(T) = (\max(|T_i - T_{\text{des},i}|))$, with the variables being the internal forces contained in vector T. Here, a feasible internal force distribution T must comply with given force constraints. Thus, without loss of generality, the optimization problem in form finding is defined as follows

$$\begin{aligned} &\min f(T) \\ &\text{subject to } r_f(T) \geq 0 \end{aligned} \tag{6.12}$$

This problem can be solved by the active-set method [25] or by some commercial software packages.

Note that for an optimization problem, because its objective function and constraints are usually not unique, the outcome may be significantly different. For instance, an improper selection of objective function and constraints may result in failure of convergence in simulation or unreasonably large computation load. To deal with this issue, the following suggestions are made.

(i) Use a minimum number of variables in the objective function.
(ii) Use analytical form of objective function whenever available.

(iii) If possible, use quadratic objective function and try to avoid nonlinearities in constraints.

These suggestions are aimed to reduce numerical effort and computer storage, which are directly related to the number of variables, the evaluation of gradient Hessian matrices, and the convergence in a numerical optimization process.

A special case of the optimization problem (6.12) is that a quadratic objective function $f(T) = \|T - T_{\text{des}}\|^2$ warrants a global minimizer when the force constraints is not considered. According to the Appendix, the minimizer is given by

$$T = -M^{+} F_c + M_{\text{null}} \alpha \tag{6.13}$$

where M_{null} is the null space of matrix M, and

$$\alpha = M_{\text{null}}^{T} T_{\text{des}} \tag{6.14}$$

With the solution in Eq. (6.13), it remains to verify that T satisfies the force constraints $r_f(T) \geq 0$. If the constraints are satisfied, T given by Eq. (6.13) is the optimal internal force distribution for the form-finding problem and no further computation is necessary.

Even if external forces are consistent, the internal force distribution T given by Eq. (6.13) is not guaranteed to satisfy all the force constraints. Under the circumstances, numerical solution of the optimization problem (6.12) is required. However, a set of feasible internal force distribution may still not be found by the optimization algorithm, which implies the solution of the optimization problem (6.12) does not exist under the given nodal coordinates. This is usually caused by extreme stringent force constraints. This nonexistence of solutions also happens in form finding by other methods. To deal with this issue, a process for adaptation of cable network geometric configuration is developed, as described in the next subsection.

6.3.3 Adaptation of Geometric Configuration

To admit a viable engineering solution in form finding, the geometric configuration (nodal coordinates) of a cable network structure should be adapted in one of the following two cases:

Case 1. The external forces are inconsistent.
Case 2. No internal force distribution can be found to satisfy all the force constraints even with consistent external forces.

To this purpose, a residual force vector is first introduced. In case 1, a residual force vector F_r is defined by

$$F_r = MT_{ls} + F_{ic} \tag{6.15}$$

where T_{ls} is obtained in Eq. (6.11). In case 2, a residual force vector is defined by replacing T_{ls} and F_{ic} by T_{des} and F_c in Eq. (6.15); that is,

$$F_r = MT_{\text{des}} + F_c \tag{6.16}$$

In the geometric configuration adaptation, the nodal coordinates of the cable network structure are updated by

$$p' = p + \gamma u^T \tag{6.17}$$

where p' is a vector of updated nodal coordinates that should comply with the geometric constraints; γ is a step size in nodal coordinate adjustment; and u is a unit vector, representing the adjustment direction. The γ and u are determined in the following optimization task

$$\begin{aligned} &\min \|u - \overline{u}\|^2 \\ &\text{subject to } r_g\left(p'\right) \geq 0 \end{aligned} \tag{6.18}$$

where $\overline{u} = F_r / \|F_r\|$.

The adaptation algorithm (6.17) and the optimization process (6.18) indicate that it is desired to adjust the nodal positions in the directions of residual forces ($u = \overline{u}$). In such a desired case, $u = \overline{u} = F_r / \|F_r\|$ and Eq. (6.17) is reduced to

$$p' = p + \gamma F_r / \|F_r\| \tag{6.19}$$

Because Eq. (6.2) is highly nonlinear regarding the nodal coordinates p', the adjustment direction u and the step size γ should be carefully selected, to avoid the failure of convergence and an unnecessarily large number of iterations in computation. With the updated nodal coordinates p', Eq. (6.2) is updated and the procedures described in Sects. 6.3.2 and 6.3.3 are repeated iteratively until a viable internal force distribution satisfying all the force constraints is obtained.

6.3.4 Procedure of the FNPM

In summary of the previous discussion, a form-finding procedure for cable network structures by the fixed nodal position method is described by a flowchart in Fig. 6.1. The procedure consists of ten main steps, which are explained as follows.

Step 1: Input the data of a form-finding problem, including a shape/surface topology of the cable network structure to be designed, geometric and force constraints, and external forces.

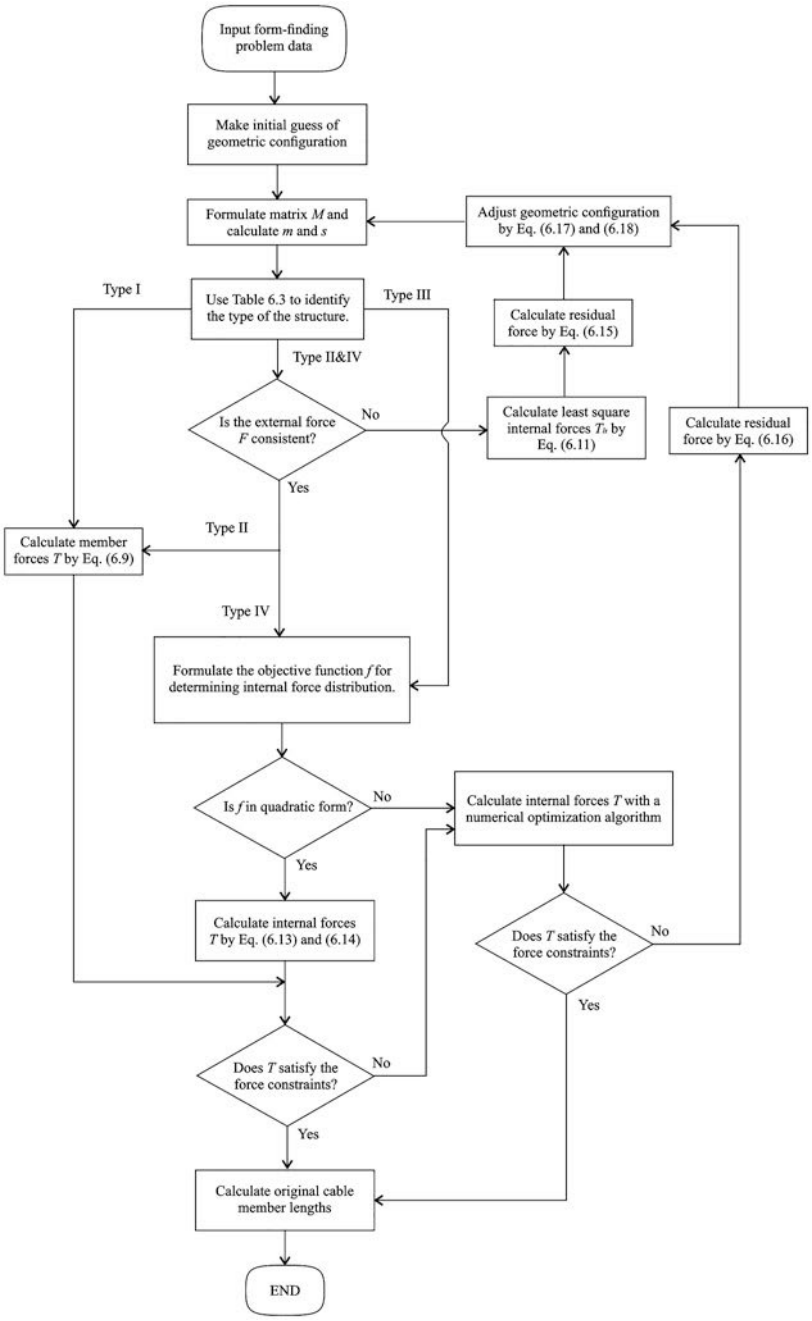


Fig. 6.1 A flowchart of the FNPM

Step 2: Make an initial guess of geometric configuration of the structure that complies with the geometric constraints in Eq. (6.5). The initial guess can be made either intuitively for simple structures or systematically for complicated structures. For example, an initial guess of the geometric configuration for a large deployable mesh reflector can be made by generating the mesh on a plane with straight lines [26] or on a sphere with geodesic curves [20]. The obtained nodal coordinates shall be used in Step 3.

Step 3: With the specified nodal coordinates, formulate the matrix M. Calculate m and s from matrix M. Identify the type of the cable network according to the value of m and s with the information in Table 6.3. For structure type of I, go to step 4. For structure type of III, go to step 5. For structure type of II and IV, check the consistency of the external forces. For structure type of II with consistent external forces, go to step 4. For structure type of IV with consistent external force, go to step 5. For structure type of II or IV with inconsistent external forces, calculate least square internal forces T_{ls} by Eq. (6.11) and go to step 9.

Step 4: Calculate the internal force vector T by Eq. (6.9). And check if the obtained T satisfies the force constraints. If the force constraints are satisfied, go to step 10. Otherwise, go to step 9.

Step 5: Define an objective function f for determining the internal force distribution. If f is in quadratic form as in the optimization problem (6.12), go to step 6. Otherwise, go to step 8.

Step 6: Calculate the internal force vector T by Eqs. (6.13) and (6.14) and go to step 7.

Step 7: Check if the computed T satisfies the force constraints (6.7). If the force constraints are satisfied, the geometric configuration and internal force distribution obtained in the current iteration are the form-finding results and go to step 10. If the force constraints are not satisfied, go to step 8.

Step 8: Form the optimization problem (6.12) and solve it for an internal force distribution by a numerical optimization algorithm. If the obtained T satisfies the force constraints (6.7), the geometric configuration and the internal force distribution in current iteration are the form-finding results, go to step 10. Otherwise, go to step 9 for adaptation of nodal coordinates.

Step 9: First calculate the residual force by Eq. (6.15) for inconsistent external forces or by Eq. (6.16) for consistent external forces. Then update the nodal coordinates by Eq. (6.17). The updated nodal coordinates should comply with the geometric constraints, and for this go back to step 3 for iterative computation.

Step 10: Once the form-finding results (nodal coordinates and internal force distribution that comply with all the geometric and force constraints) are obtained, the deformed lengths and the initial lengths of all the cable members can be computed [27].

The FMPM is different from other methods in that this method keeps the nodal positions (geometric configuration) of a cable network structure fixed or unchanged in determination of a viable internal force distribution (steps 3–9). This process of displacement first and stress later allows the FNPM to place the nodes

of a cable network structure at desired locations with great freedom, which can significantly improve shape/surface accuracy in form finding, as required by design specifications. Furthermore, because the assignment of nodal positions and the determination of internal force distribution are separated, the computational effort in the FNPM-based form finding can be much less than that by a conventional form-finding method.

6.4 Review of the Force Density Method and Dynamic Relaxation Method for FF-IEC

In the past, several form-finding methods have been developed. One commonly used method is the force density method (FDM). The force density method was originally proposed by Linkwitz and Schek [28, 29] in the 1970s and has been popular since then [30–32]. This method applies a simple mathematic transformation to the nonlinear force equilibrium equations of a cable network structure through determination of force densities. Once the force densities are obtained, the force equilibrium equations become a system of linear equations about unknown nodal coordinates, which can then be easily solved.

The dynamic relaxation method (DRM) is another commonly used method in form finding of structures [33–35]. In this method, all the nodes of a cable network structure are viewed as lumped fictitious masses that are connected by fictitious damping, and a force equilibrium configuration is obtained through numerical solution of a system of ordinary differential equations. With fictitious damping, the nodes of the structure will eventually reach the force equilibrium positions even when the initial guess of nodal positions are far from the equilibrium configuration.

In this section, the FDM and DRM are reviewed, for the purpose of comparing them with the proposed FNPM in the numerical investigation as presented in Sect. 6.6.

6.4.1 The Force Density Method

For form finding of cable net, rewrite the force equilibrium equations given in Eq. (6.2) for node i as

$$\begin{aligned} \sum_j \frac{T_{ij}}{L_{ij}} \left(x_i - x_j\right) &= F_x \\ \sum_j \frac{T_{ij}}{L_{ij}} \left(y_i - y_j\right) &= F_y \\ \sum_j \frac{T_{ij}}{L_{ij}} \left(z_i - z_j\right) &= F_z \end{aligned} \tag{6.20}$$

where T_{ij} is the internal force of the member that connects node i and j. L_{ij} is the distance between node i and j. x_i, y_i, z_i are coordinate of node i. F_x, F_y, F_z are external loads in x, y, z directions. Equation (6.20) is a nonlinear system of equations if the nodal coordinates are variables. Force density method linearizes Eq. (6.20) by applying a mathematical transformation. Define force density of the member that connects node i and j as

$$q_{ij} = \frac{T_{ij}}{L_{ij}} \tag{6.21}$$

Equation (6.20) then becomes

$$\begin{aligned} \sum_j q_{ij}\left(x_i - x_j\right) &= F_x \\ \sum_j q_{ij}\left(y_i - y_j\right) &= F_y \\ \sum_j q_{ij}\left(z_i - z_j\right) &= F_z \end{aligned} \tag{6.22}$$

According to the transformation given in [29], Eq. (6.22) can be written in matrix form as

$$\begin{aligned} C^T QCx + C^T QC_f x_f &= F_x \\ C^T QCy + C^T QC_f y_f &= F_y \\ C^T QCz + C^T QC_f z_f &= F_z \end{aligned} \tag{6.23}$$

Q is the force densities diagonal matrix. C and C_f are the two parts of branch node matrix C_S. They are defined according to the classification of fixed and free nodes. For each member j that connects the node i and k, C_S is defined by

$$C_s(j,i) = \begin{cases} +1 & i(j) = 1 \\ -1 & k(j) = 1 \\ 0 & \text{others} \end{cases} \tag{6.24}$$

C and C_f are then defined accordingly

$$\begin{bmatrix} C & C_f \end{bmatrix} = C_s \tag{6.25}$$

Once the force densities are predetermined, the nodal coordinates can be obtained by solving a linear system of equations. From Eq. (6.23), nodal coordinates are calculated as

$$\begin{aligned} x &= \left(C^T QC\right)^{-1} R_x \\ y &= \left(C^T QC\right)^{-1} R_y \\ z &= \left(C^T QC\right)^{-1} R_z \end{aligned} \tag{6.26}$$

where

$$\begin{aligned} R_x &= F_x - C^T Q C_f x_f \\ R_y &= F_y - C^T Q C_f y_f \\ R_z &= F_z - C^T Q C_f z_f \end{aligned} \tag{6.27}$$

While being widely used, the FDM may not work well when a cable network structure has complicated geometric constraints like axial symmetry or has a desired working shape with high accuracy requirement. This is because the FDM-based form finding must obtain force densities for a prestressed structure in design prior to determining a geometric configuration for the structure. This stress-first-and-displacement-later methodology leaves little room for selection of nodal positions to comply with the geometric constraints of the structure. A natural attempt to resolve this issue is to introduce algorithms in a problem of constrained numerical optimization. Nevertheless, by the nature of the FDM, the objective function and constraints of the optimization problem can be quite complicated, which in general do not always guarantee the convergence in numerical optimization, and sometimes even the existence of a solution.

6.4.2 The Dynamic Relaxation Method

In the DRM, the governing equation of motion in x direction of any node i of a structure at time t is given by

$$R_{xi}^t = M_i \dot{v}_{xi}^t + C_i v_{xi}^t \tag{6.28}$$

where R_{xi}^t is residual force at node i in direction x at time t. M_i is fictitious mass at node i. C_i is damping constant at node i. $\dot{v}_{xi}^t$ and v_{xi}^t are the acceleration and velocity of node i in direction x. Note that the forces generated by the stiffness of the structure and external loads are both included in the residual force. By finite difference approximation,

$$R_{xi}^t = \frac{M_i}{\Delta t}\left(v_{xi}^{t+\Delta t/2} - v_{xi}^{t-\Delta t/2}\right) + \frac{C_i}{2}\left(v_{xi}^{t+\Delta t/2} + v_{xi}^{t-\Delta t/2}\right) \tag{6.29}$$

Rewrite Eq. (6.29) as

$$v_{xi}^{t+\Delta t/2} = R_{xi}^t \left(\frac{1}{M_i/\Delta t + C_i/2}\right) + v_{xi}^{t-\Delta t/2}\left(\frac{M_i/\Delta t - C_i/2}{M_i/\Delta t + C_i/2}\right) \tag{6.30}$$

Thus, the average velocity of next time step is able to be predicted by the information in current step. The total x deflection of node i at time $t + \Delta t$ is

$$\delta_{x_i}^{t+\Delta t} = \delta_{x_i}^{t} + \Delta t v_{x_i}^{t+\Delta t/2} \tag{6.31}$$

These calculations are carried out simultaneously for each node of the cable network structure to give the complete displacement at next time step. The force residual is then updated according to the nodal coordinates at current time step by considering the elongation of each member.

$$\Delta R_{xim}^{t+\Delta t} = \frac{T_m^{t+\Delta t}}{L_m^{t+\Delta t}} \left[\left(x_k + \delta_{x_k}^{t+\Delta t} \right) - \left(x_i + \delta_{x_i}^{t+\Delta t} \right) \right] \tag{6.32}$$

where $\Delta R_{xim}^{t+\Delta t}$ is the force residual generated by member m that connects node i and k at time $t + \Delta t$. $T_m^{t+\Delta t}$ is the internal force of member m at time $t + \Delta t$. $L_m^{t+\Delta t}$ is the length of member m at time $t + \Delta t$. For initial configuration determination, $T_m^{t+\Delta t}$ is set to be constant so that the nodal position can be determined from the iterations. At the starting nodal position can be arbitrarily chosen, according to Barnes [33], rapid convergence to the correct geometry is obtained even when the assumed initial geometry is extremely inaccurate.

The dynamic relaxation method was also further developed. For instance, "kinetic damping" was introduced to replace the conventional damping described above [36]. The structure is analyzed by the procedure introduced above with damping coefficient being set to zero. When a local peak in the total kinetic energy of the system is detected, all velocity components are set to zero. The procedure is then repeated until the equilibrium position of the structure is found.

The advantage of the DRM is that it numerically mimics the dynamic behavior of the cable network structure under the assumption that the system will eventually reach its equilibrium position with energy dissipation caused by viscous damping. This method is suitable for form finding of both initial configuration determination and deformed configuration determination. However, the time step and lumped mass need to be carefully selected in order to guarantee the convergence of the algorithm. Also, according to Motro [37], this method will not work effectively in form finding of complicated structures.

Besides the FDM and DRM, other methods for FF-IEC have been developed. Kinematical methods [38] are another type of form-finding methods. In form finding of a truss structure, a kinematical method finds an initial equilibrium configuration by fixing the lengths of bar members and gradually decreasing the lengths of cable members until a minimum of cable member lengths is reached. The procedure of a kinematical method can be carried out via either analytical computation [39] or nonlinear programming [7]. Kinematical methods have been shown to be successful in design of structures with a small number of members. For truss structures with many members, kinematical methods become computationally expensive.

6.5 Methods of FF-DEC

As mentioned in Sect. 6.2, besides form finding of *initial* equilibrium configuration (FF-IEC), form finding of *deformed* equilibrium configuration (FF-DEC) is also pursued. The selection of one of these two types of form-finding problems depends on whether the original (undeformed) member lengths of a cable network structure are known or prespecified. If the original member lengths are to be determined in form finding, FF-IEC is selected. In this case, the corresponding form-finding methods shall have more freedom in determining the configuration and member forces of the structure, and as such, the form-finding solution is not unique. On the other hand, if the original member lengths are prespecified or predetermined, FF-DEC is selected. In this case, with the given member connectivity and external forces, the form-finding solution in terms of the structure shape and member forces is unique. Accordingly, an objective of a form-finding method for FF-DEC is to find the unique solution with least computational effort.

In this section, three methods for FF-DEC problems are introduced.

6.5.1 Singular Value Decomposition Method

Singular value decomposition (SVD) method [5] is utilized in FF-DEC. Force equilibrium equations Eq. (6.2) and compatibility equations Eq. (6.33) are introduced for this method.

$$Bd = \varepsilon \tag{6.33}$$

B is the kinematic matrix. d is the vector of nodal displacements. ε is the vector of member elongation. Matrices M and B have the following relationship.

$$M = B^T \tag{6.34}$$

Apply the technique of singular value decomposition to M,

$$M = UVW^T \tag{6.35}$$

where U is a $3n$ by $3n$ orthogonal matrix; W is a k by k orthogonal matrix. n is number of free nodes. k is number of members V is a $3n$ by k matrix with r positive elements v_{ii} on the leading diagonal; all other elements are zero.

Define the following sub matrices,

$$U_r = \left[u_1, \cdots, u_r \right] \qquad U_{n-r} = \left[u_{r+1}, \cdots, u_n \right] \tag{6.36}$$

$$W_r = \left[w_1, \cdots, w_r \right] \qquad W_{k-r} = \left[w_{r+1}, \cdots, w_k \right] \tag{6.37}$$

According to Table 6.1, for structures with $m > 0$, SVD method starts by checking if the given external force F can be carried by the structure.

$$U_{n-r}^T F = 0 \tag{6.38}$$

Once the external loads pass the check, the internal force in each member is then computed as

$$T = \sum_{i=1}^{r} \frac{u_i^T F}{v_{ii}} w_i + W_{k-r}\alpha \tag{6.39}$$

α contains s free parameters. The value of α is determined by a set of s compatibility equations: the member elongation ε should be orthogonal to the incompatible member elongations.

$$W_{k-r}^T \varepsilon = 0 \tag{6.40}$$

Free variable α is calculated by solving the equation of

$$\alpha = -\left(W_{n_c-r}^T I W_{n_c-r} \right)^{-1} W_{n_c-r}^T \left(I \sum_{i=1}^{r} \frac{u_i^T l}{v_{ii}} w_i + \varepsilon_0 \right) \tag{6.41}$$

where ε_0 is the initial strain and I is the square matrix of member flexibilities. The vector of nodal displacements is calculated as

$$d = \sum_{i=1}^{r} \frac{w_i^T \varepsilon}{v_{ii}} u_i + U_{n-r}\beta \tag{6.42}$$

where β contains m free parameters.

The member forces and nodal displacements are obtained by SVD method. The major computation effort of this method is the singular value decomposition of matrix M. This method is quite efficient comparing with other methods for FF-DEC. However, one important disadvantage of SVD method is that it is only useful for small displacement analysis [8]. Because, in Eqs. (6.39 and 6.42), for simplicity, matrices U, V and W are assumed to be constant. But when nodal positions are changed, these matrices are changed accordingly because matrix M is obtained from nodal positions. Therefore, this method can only deliver accurate results when the nodal displacements are small.

6.5.2 Stiffness Matrix Method

Stiffness matrix method is another commonly used approach for FF-DEC of cable network structures. It was introduced 1970s [40, 41] and was further extended by Murakami [42, 43], Guest [44] and Yuan [22]. Rewrite Eq. (6.2) with nodal coordinates $p = (x, y, z,)$ being variable (the original member lengths are given),

$$M(p)T(p) + F = 0 \tag{6.43}$$

Define $h(p)$ as

$$h(p) = M(p)T(p) + F \tag{6.44}$$

By Taylor series, $h(p)$ can be expanded at any nodal positions p_k as

$$h(p) = h\left(p_k\right) + J\left(p_k\right)\left(p - p_k\right) + o(p - p_k)^2 \tag{6.45}$$

where $J(p_k)$ is the tangent stiffness matrix of the structure, defined by

$$J(p) = \begin{bmatrix} \nabla h_1(p)^T \\ \nabla h_2(p)^T \\ \vdots \\ \nabla h_{3n}(p)^T \end{bmatrix} \tag{6.46}$$

$h_i(p)$ is the ith nonlinear equation of $h(p)$. If p is close to p_k, the high order terms of Eq. (6.45) can be neglected. The approximate solution p^* of Eq. (6.44) can be obtained by solving the linear system of equations

$$J\left(p_k\right) p^* = J\left(p_k\right) p_k - h\left(p_k\right) \tag{6.47}$$

To further improve the accuracy, p^* is used to replace p_k in Eq. (6.45) and repeat the procedure given in Eqs. (6.45, 6.46, and 6.47). The algorithm stops when

$$h\left(p^*\right) < \tau \tag{6.48}$$

where τ is a small value used to determine the accuracy of the algorithm.

Therefore, the nodal coordinates under external force F is obtained. An attractive feature of this method is that it demonstrates quadratic convergence if the current iteration is sufficiently close to the converged solution. On the other hand, when the starting point is remote from a solution, this method can behave erratically [45]. To resolve this issue, external loads are updated by tiny incremental amounts. The solution process starts with the initial forces F^0 and nodal positions p^0, increases

loads gradually at each step, and applies an iteration procedure to eventually obtain acceptable solutions.

Stiffness matrix method is useful for form finding of cable network structures subject to large deformation. But it requires the knowledge of the initial configuration that satisfies

$$M\left(p^0\right)T\left(p^0\right)+F^0=0 \tag{6.49}$$

In addition, stiffness matrix method is based on iteration at incremental forces which is quite computationally expensive to solve for large cable network structures.

6.5.3 Dynamic Relaxation Method

Dynamic relaxation method is also an effective method for FF-DEC. It only requires minor change in the framework introduced in Sect. 6.4.2. Instead of $T_m^{t+\Delta t}$ being set constant in Eq. (6.32), the current internal force in member m is then given by

$$T_m^{t+\Delta t}=T_m^t+\left(\frac{EA}{L^t}\right)e^{t+\Delta t} \tag{6.50}$$

where E is the Young's modulus of the member material. A is the cross-section of each member. Here it is assumed that the cross-section is uniform for all members. $e^{t+\Delta t}$ is the member elongation at time $t+\Delta t$. L_t is the member length at time t.

6.6 Implementation of the Form-Finding Methods

In this section, the form-finding methods introduced in the previous sections are demonstrated in three examples. In the first example, the proposed FNPM is illustrated on a simple 2-D cable net of five nodes and five cable members, in FF-IEC. The FDM and DRM are also applied to this example to compare with the FNPM. In the second example, the FF-DEC solution of a 2-D cable net with known original member lengths is determined by the methods introduced in Sect. 6.5. In the third example, form finding of a large deployable mesh reflector of 865 nodes is performed.

6.6.1 FF-IEC of a 2-D Cable Net

The form finding of a 2-D cable net in Fig. 6.2 is considered in this section, where the cable network has five nodes and five members.

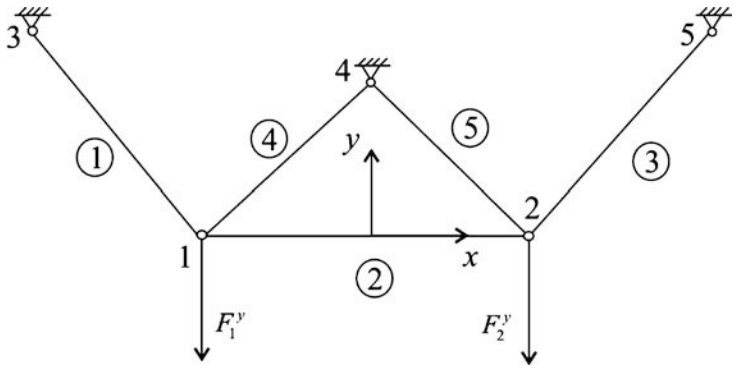


Fig. 6.2 2-D cable net with five nodes and five members

The cable network is mounted at three fixed nodes 3, 4, and 5, with coordinates given by

$$(x_3, y_3) = (-1.5, 1)\ (x_4, y_4) = (0, 0.5)\ (x_5, y_5) = (1.5, 1) \tag{6.51}$$

Because nodes 3, 4, and 5 are fixed ones (boundary nodes), only force equilibrium equations at the movable nodes 1 and 2 are required. The external forces that are applied to nodes 1 and 2 are described by

$$F = \left[F_1^x\ F_1^y\ F_2^x\ F_2^y \right]^T = \left[0\ -2\ 0\ -2 \right]^T \tag{6.52}$$

The structure is required to be symmetric with the y-axis that passes through node 4. Assume that the cable members are linearly elastic, with the same Young's modulus E of 1×10^9 and the same cross-section area A of 1×10^{-7}. Here for simplicity in presentation, non-dimensional values of system parameters have been used. Here, all the structure parameters have non-dimensional units.

The design requirements for the cable network are as follows: (a) the lengths of members 4 and 5 (marked by numbers in circles) are required to be 0.7071; (b) slack is not allowed in all the cable members; and (c) the internal forces of the members fall within the region of 0.7–1.3.

With the previous description, the geometric constraints of the structure are described by

$$\begin{aligned} &x_1 = -x_2,\ y_1 = y_2 \\ &\sqrt{x_1^2 + (y_1 - 1/2)^2} = \sqrt{2}/2 \\ &\sqrt{x_2^2 + (y_2 - 1/2)^2} = \sqrt{2}/2 \end{aligned} \tag{6.53}$$

and the force constraints are given by

$$0.7 < T < 1.3 \tag{6.54}$$

The internal forces (member axial forces) are desired to be uniform; say

$$T_{\text{des}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T \tag{6.55}$$

In this example, the objective function for optimal design of internal force distribution is of quadratic form, $\min\|T_{\text{des}} - T\|^2$. The fixed nodal position method, force density method, and dynamic relaxation method shall be used for the form finding of this given 2-D cable net.

6.6.1.1 The Fixed Nodal Position Method

An initial guess of the geometric configuration of the structure that complies with the geometric constraints (6.53) is made by assigning the coordinates of nodes 1 and 2 as follows:

$$\begin{bmatrix} x_1 & y_1 & x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} -0.5 & 0 & 0.5 & 0 \end{bmatrix}^T \tag{6.56}$$

Matrix M in Eq. (6.2) is obtained from the nodal coordinates:

$$M = \begin{bmatrix} -0.7071 & 1 & 0 & 0.7071 & 0 \\ 0.7071 & 0 & 0 & 0.7071 & 0 \\ 0 & -1 & 0.7071 & 0 & -0.7071 \\ 0 & 0 & 0.7071 & 0 & 0.7071 \end{bmatrix} \tag{6.57}$$

The null space M_{null} and the MP inverse M^+ of M are then calculated as

$$\begin{aligned} M_{\text{null}} &= \begin{bmatrix} -0.4082 & -0.5774 & -0.4082 & 0.4082 & 0.4082 \end{bmatrix}^T \\ M^+ &= \begin{bmatrix} -0.4714 & 0.7071 & -0.2357 & 0 \\ 0.3333 & 0 & -0.3333 & 0 \\ 0.2357 & 0 & 0.4714 & 0.7071 \\ 0.4714 & 0.7071 & 0.2357 & 0 \\ -0.2357 & 0 & -0.4714 & 0.7071 \end{bmatrix} \end{aligned} \tag{6.58}$$

With the obtained matrix M, the values of m and s are calculated as 0 and 1. According to Table 6.3, the assembly of the structure is of type III. The external forces are consistent ($F = F_c$). Since the objective function is in quadratic form, α is calculated by Eq. (6.14) and it is given by

$$\alpha = M_{\text{null}}^T T_{\text{des}} = -0.5774 \tag{6.59}$$

The internal force vector T by Eq. (6.13) is obtained as

$$T = -M^{+}F_c + M_{null}\alpha = \begin{bmatrix} 1.6499 & 0.3333 & 1.6499 & 1.1785 & 1.1785 \end{bmatrix}^T \tag{6.60}$$

It is easy to see that T given in the previous equation does not satisfy the force constraints (6.54). By the procedure in Sect. 6.4, an optimization problem (6.18) is defined as follows:

$$\begin{gathered} \min \left(\|T - T_{\text{des}}\|^2 \right) \\ \text{subject to } MT = -F \text{ and } 0.7 < T < 1.3 \end{gathered} \tag{6.61}$$

A numerical optimization algorithm is used for the optimization problem. However, the internal force distribution that satisfies the force constraints (6.54) still cannot be found.

To deal with this issue, adaptation of the cable network geometric configuration is applied. By Eq. (6.16), the force residual F_r is obtained as

$$F_r = MT_{\text{des}} - (-F) = \begin{bmatrix} 1.0000 & -0.5858 & -1.0000 & -0.5858 \end{bmatrix}^T \tag{6.62}$$

By Eq. (6.17) with a step size γ of 0.6, the nodal positions of the cable network are adapted, which are also required to comply with the geometric constraints (6.53). The unit vector u in Eq. (6.17) is determined via the numerical optimization (6.18) and it is given by

$$u = \begin{bmatrix} 0.6272 & -0.3266 & -0.6272 & -0.3266 \end{bmatrix}^T \tag{6.63}$$

The updated nodal coordinates are then given by

$$p' = p + \gamma u^T = \begin{bmatrix} -0.1228 & -0.1964 & 0.1228 & -0.1964 \end{bmatrix}^T \tag{6.64}$$

The updated geometric configuration and residual forces of the structure are shown in Fig. 6.3.

With the updated nodal coordinates, the form-finding procedure is repeated and it yields the following results:

$$\begin{gathered} \begin{bmatrix} x_1 & y_1 & x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} -0.1228 & -0.1964 & 0.1228 & -0.1964 \end{bmatrix}^T \\ T = \begin{bmatrix} 1.2523 & 0.7376 & 1.2523 & 1.1969 & 1.1969 \end{bmatrix}^T \end{gathered} \tag{6.65}$$

As can be seen, T given in Eq. (6.65) now satisfies the force constraints and it is much closer to T_{des} than the T given in Eq. (6.60). Finally, with the form-finding results given in Eq. (6.65), the deformed lengths of the cable members are obtained as follows:

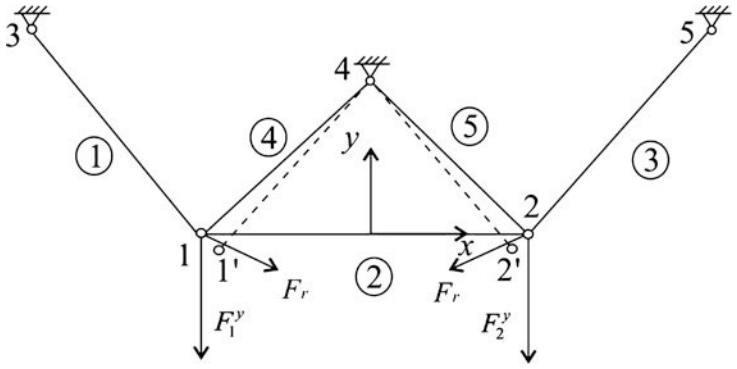


Fig. 6.3 A 2-D cable net with the adapted nodal positions and residual forces

$$L = \left[\, 1.8243 \;\; 0.2456 \;\; 1.8243 \;\; 0.7071 \;\; 0.7071 \,\right]^T \tag{6.66}$$

Also, with the deformed member lengths, the internal force distribution T and the longitudinal rigidity EA of the members, the vector L_{org} of the original members lengths is obtained as follows:

$$L_{org} = \left[\, 1.8018 \;\; 0.2438 \;\; 1.8018 \;\; 0.6988 \;\; 0.6988 \,\right]^T \tag{6.67}$$

6.6.1.2 The Force Density Method

In this example, to obtain consistent form-finding results, the force densities are selected as $q = \left[\, 0.6865 \;\; 3.0033 \;\; 0.6865 \;\; 1.6927 \;\; 1.6927 \,\right]^T$. From Eqs. (6.24 and 6.25), C, C_f and Q are obtained as

$$C = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \; C_f = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \; Q = \begin{bmatrix} 0.6865 & 0 & 0 & 0 & 0 \\ 0 & 3.0033 & 0 & 0 & 0 \\ 0 & 0 & 0.6865 & 0 & 0 \\ 0 & 0 & 0 & 1.6927 & 0 \\ 0 & 0 & 0 & 0 & 1.6927 \end{bmatrix} \tag{6.68}$$

x_f, y_f, F_x and F_y are given in Eqs. (6.51 and 6.52) as

$$x_f = \begin{bmatrix} -1.5 \\ 0 \\ 1.5 \end{bmatrix} \; y_f = \begin{bmatrix} 1 \\ 0.5 \\ 1 \end{bmatrix} \; F_x = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \; F_y = \begin{bmatrix} -2 \\ -2 \end{bmatrix} \tag{6.69}$$

From Eqs. (6.26 and 6.27), coordinate of nodes 1 and 2 are calculated:

$$\begin{cases} \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \left(C^T Q C\right)^{-1} R_x \\ \begin{bmatrix} y_1 & y_2 \end{bmatrix} = \left(C^T Q C\right)^{-1} R_y \end{cases} = \begin{cases} \left(C^T Q C\right)^{-1} \left(F_x - C^T Q C_f x_f\right) \\ \left(C^T Q C\right)^{-1} \left(F_y - C^T Q C_f y_f\right) \end{cases} = \begin{cases} \begin{bmatrix} -0.1228 & 0.1228 \end{bmatrix}^T \\ \begin{bmatrix} -0.1963 & -0.1963 \end{bmatrix}^T \end{cases} \tag{6.70}$$

The internal force of each member is calculated as

$$T = \begin{bmatrix} q_1 L_1 & q_2 L_2 & q_3 L_3 & q_4 L_4 & q_5 L_5 \end{bmatrix}^T = \begin{bmatrix} 1.2523 & 0.7376 & 1.2523 & 1.1969 & 1.1969 \end{bmatrix}^T \tag{6.71}$$

6.6.1.3 The Dynamic Relaxation Method

Dynamic relaxation method is utilized for form finding of the same 2-D cable-net. The objective internal force distribution, which is the $T_m^{t+\Delta t}$, is considered fixed in Eq. (6.32). $T_m^{t+\Delta t}$ is selected as the obtained result given Eqs. (6.65) and (6.71) for comparison. As suggested in [33], for the ith node, Eq. (6.32) is rewritten as

$$v_{x_i}^{t+\Delta t/2} = A_i R_{x_i}^t + B_i v_{x_i}^{t-\Delta t/2} \tag{6.72}$$

where

$$A_i = \frac{\Delta t}{M_i \left(1 + C_i/2\right)} \qquad B_i = \frac{1 - C_i/2}{1 + C_i/2} \tag{6.73}$$

In this example, C and M are selected as

$$C = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}^T \qquad M = \begin{bmatrix} 1 & 1 \end{bmatrix}^T \tag{6.74}$$

For simplicity, Δt is 0.7 in all iterations. in. In order to show the of rate of convergence when the assumed initial geometry is extremely inaccurate, the initial nodal positions $\left(x_{1,2}^0, y_{1,2}^0\right)$ are selected to be far away from the equilibrium position found previously. Specifically,

$$\begin{Bmatrix} x_1^0 & y_1^0 & x_2^0 & y_2^0 \end{Bmatrix}^T = \begin{Bmatrix} -10 & -10 & 10 & 30 \end{Bmatrix}^T \tag{6.75}$$

The updated nodal coordinates for selected iterations are given in Table 6.4.

The iteration stops at the 79th iteration when the residual force satisfies the following condition:

$$\|R_x\| + \left\|R_y\right\| < 1 \times 10^{-6} \tag{6.76}$$

The results of the dynamics relaxation method in this example are consistent with those obtained by the FNPM and the FDM. Note that the rate of convergence

Table 6.4 Nodal coordinates for selected iterations of dynamic relaxation method

Number of iterations	Nodal coordinate
1	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -10.0000 & -10.7840 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} 10.0000 & 29.2160 \end{bmatrix}^T$
5	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -5.4616 & -10.0474 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} 7.0208 & 15.0584 \end{bmatrix}^T$
10	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -0.2014 & -5.7774 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} -0.1699 & -5.1577 \end{bmatrix}^T$
20	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -0.0435 & -3.0352 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} 0.1415 & -3.3833 \end{bmatrix}^T$
30	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -0.0987 & -0.2598 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} 0.1184 & -0.1782 \end{bmatrix}^T$
50	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -0.1230 & -0.1952 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} 0.1231 & -0.1946 \end{bmatrix}^T$
70	$\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T = \begin{bmatrix} -0.1228 & -0.1964 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} 0.1228 & -0.1964 \end{bmatrix}^T$

of the algorithm of dynamic relaxation method can be significantly improved by the modification given Ref. [35].

6.6.2 FF-DEC of a 2-D cable net

The deformed equilibrium configuration of a 2-D cable net, which is almost the same with the structure studied in Sect. 6.1 Fig. 6.2, shall be obtained by the SVD method, the stiffness matrix method, and the dynamic relaxation method, respectively. The 2-D cable net in this example has the original member lengths that are predetermined. The initial nodal positions, external forces, and member material elasticity are different from the structure in the previous example. The initial nodal positions with no external forces and strains are given as

$$\begin{bmatrix} x_1 & y_1 & x_2 & y_2 \end{bmatrix}^T = \begin{bmatrix} -0.1228 & -0.1964 & 0.1228 & -0.1964 \end{bmatrix}^T$$
$$\begin{bmatrix} x_3 & y_3 & x_4 & y_4 & x_5 & y_5 \end{bmatrix}^T = \begin{bmatrix} -1.5 & 1 & 0 & 0.5 & 1.5 & 1 \end{bmatrix} \tag{6.77}$$

Nodes 3, 4, and 5 are fixed. The external forces applied on nodes 1 and 2 are set as $F = \begin{bmatrix} -8 & -8 & 8 & -8 \end{bmatrix}^T$. The initial strain ε_0 is zero. Assume that the members are linearly elastic, with Young's modulus E of 1×10^{10} and cross-section area A of 1×10^{-7}.

6.6.2.1 The Singular Value Decomposition Method

Singular value decomposition is applied to the FF-DEC of the 2D cable net. In singular value decomposition method, only small deformation is considered,

which implies that the geometric stiffness (corresponds to the stiffness due to the reorientation of stressed members) of the structure is negligible. The matrix M that contains the direction cosines is calculated as

$$M = \begin{bmatrix} -0.7549 & 1.0000 & 0 & 0.1737 & 0 \\ 0.6558 & 0 & 0 & 0.9848 & 0 \\ 0 & -1.0000 & 0.7549 & 0 & -0.1737 \\ 0 & 0 & 0.6558 & 0 & 0.9848 \end{bmatrix} \tag{6.78}$$

By singular value decomposition,

$$M = UVW^T \tag{6.79}$$

where

$$U = \begin{bmatrix} 0.6855 & -0.2361 & 0.1733 & 0.6665 \\ -0.1733 & 0.6665 & 0.6855 & 0.2361 \\ -0.6855 & -0.2361 & -0.1733 & 0.6665 \\ -0.1733 & -0.6665 & 0.6855 & -0.2361 \end{bmatrix}$$

$$V = \begin{bmatrix} 1.6377 & 0 & 0 & 0 & 0 \\ 0 & 1.2308 & 0 & 0 & 0 \\ 0 & 0 & 1.1481 & 0 & 0 \\ 0 & 0 & 0 & 0.6966 & 0 \end{bmatrix} \tag{6.80}$$

$$W = \begin{bmatrix} -0.3854 & 0.5000 & 0.2777 & -0.5000 & -0.5238 \\ 0.8372 & -0.0000 & 0.3019 & 0.0000 & -0.4560 \\ -0.3854 & -0.5000 & 0.2777 & 0.5000 & -0.5238 \\ -0.0315 & 0.5000 & 0.6143 & 0.5000 & 0.3488 \\ -0.0315 & -0.5000 & 0.6143 & -0.5000 & 0.3488 \end{bmatrix}$$

By Eq. (6.39) the internal forces under external loads are calculated.

$$T = \sum_{i=1}^{r} \frac{u_i^T F}{v_{ii}} w_i + W_{ne-r}\alpha = \begin{bmatrix} 0.6855 \\ -0.1733 \\ -0.6855 \\ -0.1733 \end{bmatrix}^T \begin{bmatrix} -8 \\ -8 \\ 8 \\ -8 \end{bmatrix} \frac{1}{1.6377} \begin{bmatrix} -0.3854 \\ 0.8372 \\ -0.3854 \\ -0.0315 \\ -0.0315 \end{bmatrix} + \begin{bmatrix} -0.2361 \\ 0.6665 \\ -0.2361 \\ -0.6665 \end{bmatrix}^T \begin{bmatrix} -8 \\ -8 \\ 8 \\ -8 \end{bmatrix} \frac{1}{1.2308}$$
$$\begin{bmatrix} 0.5000 \\ 0.0000 \\ -0.5000 \\ 0.5000 \\ -0.5000 \end{bmatrix} + \begin{bmatrix} 0.1733 \\ 0.6855 \\ -0.1733 \\ 0.6855 \end{bmatrix}^T \begin{bmatrix} -8 \\ -8 \\ 8 \\ -8 \end{bmatrix} \frac{1}{1.1481} \begin{bmatrix} 0.2777 \\ 0.3019 \\ 0.2777 \\ 0.6143 \\ 0.6143 \end{bmatrix} + \begin{bmatrix} 0.6665 \\ 0.2361 \\ 0.6665 \\ -0.2361 \end{bmatrix}^T \begin{bmatrix} -8 \\ -8 \\ 8 \\ -8 \end{bmatrix} \frac{1}{0.6966} \begin{bmatrix} -0.5000 \\ 0.0000 \\ 0.5000 \\ 0.5000 \\ -0.5000 \end{bmatrix} + \alpha$$
$$\begin{bmatrix} -0.5238 \\ -04560 \\ -0.5238 \\ 0.3488 \\ 0.3488 \end{bmatrix} = \begin{bmatrix} 1.3946 \\ 7.8034 \\ 1.3946 \\ 7.1947 \\ 7.1947 \end{bmatrix} + \alpha \begin{bmatrix} -0.5238 \\ -0.4560 \\ -0.5238 \\ 0.3488 \\ 0.3488 \end{bmatrix} \tag{6.81}$$

Free variable α is calculated according to the fact that the member elongation ε must vanish in the subspace of incompatible member elongation. By Eq. (6.41),

$$\alpha = -\left(W_{n_c-r}^T I W_{n_c-r}\right)^{-1} W_{n_c-r}^T \left(I \sum_{i=1}^{r} \frac{u_i^T F}{v_{ii}} w_i + \varepsilon_0 \right) = 0 \tag{6.82}$$

In this 2-D cable-net example, F is a 5 by 5 identical matrix and $\varepsilon_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$. The vector of member elongations is obtained as,

$$\varepsilon = \frac{TL}{EA} = \begin{bmatrix} 0.0025 \\ 0.0019 \\ 0.0025 \\ 0.0051 \\ 0.0051 \end{bmatrix} \tag{6.83}$$

In this example since $m = 0$, U_{n_j-r} and β do not exist. From Eq. (6.42), the nodal displacements are obtained.

$$d = \sum_{i=1}^{r} \frac{w_i^T \varepsilon}{v_{ii}} u_i + U_{n_j-r}\beta = \sum_{i=1}^{r} \frac{w_i^T \varepsilon}{v_{ii}} u_i = \begin{bmatrix} -0.0010 & -0.0050 & 0.0010 & -0.0050 \end{bmatrix}^T \tag{6.84}$$

The nodal coordinates after deformation is then given as

$$\begin{Bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{Bmatrix} = \begin{Bmatrix} x_1^0 \\ y_1^0 \\ x_2^0 \\ y_2^0 \end{Bmatrix} + d = \begin{Bmatrix} -0.1238 \\ -0.2014 \\ 0.1238 \\ -0.2014 \end{Bmatrix} \tag{6.85}$$

6.6.2.2 The Stiffness Matrix Method

For comparison, the FF-DEC for the same structure given subject to same external forces in part (a) shall be done by the stiffness matrix method. Starting from the equilibrium position given in Eq. (6.86), the external forces $F = \begin{bmatrix} -8 & -8 & 8 & -8 \end{bmatrix}^T$ are added up gradually. To ensure fast rate of convergence, the loads are divided into 5 incremental loads. Starting from the initial external loads $F^0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T$ with nodal coordinates $p^0 = \begin{bmatrix} -0.1228 & -0.1964 & 0.1228 & -0.1964 \end{bmatrix}^T$ at the equilibrium state, the first load step is taken as

$$[x_1 y_1 x_2 y_2]^T = [-0.1228 \quad -0.1964 \quad 0.1228 \quad -0.1964]^T \tag{6.86}$$

$$F^1 = F^0 + \frac{1}{5}\left(F - F^0\right) = \begin{bmatrix} -1.6 & -1.6 & 1.6 & -1.6 \end{bmatrix}^T \tag{6.87}$$

From Eq. (6.46),

$$J\left(p^0\right) \equiv J\left(x^0, y^0, z^0\right) = \begin{bmatrix} -4426.71 & 29.54 & 4071.66 & 0 \\ 29.54 & -1607.25 & 0 & 0 \\ 4071.66 & 0 & -4426.71 & -29.54 \\ 0 & 0 & -29.54 & -1607.25 \end{bmatrix} \tag{6.88}$$

The right-hand side of Eq. (6.47) is

$$J\left(p^0\right)p^0 - h\left(p^0\right) = \begin{bmatrix} -4426.71 & 29.54 & 4071.66 & 0 \\ 29.54 & -1607.25 & 0 & 0 \\ 4071.66 & 0 & -4426.71 & -29.54 \\ 0 & 0 & -29.54 & -1607.25 \end{bmatrix} \begin{bmatrix} -0.1228 \\ -0.1964 \\ 0.1228 \\ -0.1964 \end{bmatrix} - \begin{bmatrix} -1.6 \\ -1.6 \\ 1.6 \\ -1.6 \end{bmatrix} = \begin{bmatrix} 1039.40 \\ 313.64 \\ -1039.40 \\ 313.64 \end{bmatrix} \tag{6.89}$$

p^* is obtained by solving Eq. (6.47)

$$p^* = J\left(p^0\right)^{-1}\left(J\left(p^0\right)p^0 - h\left(p^0\right)\right) = \begin{bmatrix} -0.1230 & -0.1974 & 0.1230 & -0.1974 \end{bmatrix}^T \tag{6.90}$$

Table 6.5 Nodal coordinates and number of iterations for each incremental load step of stiffness matrix method

Incremental load $\left[F_{x_1}\ F_{y_1}\ F_{x_2}\ F_{y_2}\right]^T$	Number of iterations	Converged nodal position $\left[x_1\ y_1\ x_2\ y_2\right]^T$
$\left[-1.6\ -1.6\ 1.6\ -1.6\right]^T$	3	$\left[-0.1230\ -0.1974\ 0.1230\ -0.1974\right]^T$
$\left[-3.2\ -3.2\ 3.2\ -3.2\right]^T$	3	$\left[-0.1232\ -0.1984\ 0.1232\ -0.1984\right]^T$
$\left[-4.8\ -4.8\ 4.8\ -4.8\right]^T$	3	$\left[-0.1234\ -0.1994\ 0.1234\ -0.1994\right]^T$
$\left[-6.4\ -6.4\ 6.4\ -6.4\right]^T$	3	$\left[-0.1236\ -0.2004\ 0.1236\ -0.2004\right]^T$
$\left[-8\ -8\ 8\ -8\right]^T$	3	$\left[-0.1238\ -0.2014\ 0.1238\ -0.2014\right]^T$

Then the iterations continue as introduced in Sect. 6.5 until the algorithm converges. Then the forces are increased and the procedure above repeats. Table 6.5 gives the nodal coordinates under each incremental force and the number of iterations for the algorithm to converge.

6.6.2.3 The Dynamic Relaxation Method

The dynamic relaxation method shall also be applied for the FF-DEC of the same 2-D cable net. The initial nodal positions, external forces, and original member lengths are the same with Sects. 6.6.2.1 and 6.6.2.2. The majority of the procedure is the same as it was introduced in Sect. 6.4.2 except member forces in Eq. (6.32) is calculated by Eq. (6.91), not predetermined. Other parameters of the algorithm are the same with the example of FF-IEC, given in Eqs. (6.73) and (6.74). For simplicity, Δt is selected as 0.01.

$$T = \frac{L - L_{\text{org}}}{L_{\text{org}}} EA \tag{6.91}$$

The nodal coordinates for selected iterations are given in Table 6.6.

The iteration stops at the 65th iteration when the residual force satisfies the following condition

$$\|R_x\| + \left\|R_y\right\| < 1 \times 10^{-6} \tag{6.92}$$

Table 6.6 Nodal coordinates for selected iterations of dynamic relaxation method for FF-DEC

Number of iterations	Nodal coordinate
1	$\left[x_1\ y_1\right]^T = \left[-0.1234\ -0.1970\right]\ \left[x_2\ y_2\right]^T = \left[0.1234\ -0.1970\right]^T$
5	$\left[x_1\ y_1\right]^T = \left[-0.1237\ -0.2006\right]\ \left[x_2\ y_2\right]^T = \left[0.1237\ -0.2006\right]^T$
10	$\left[x_1\ y_1\right]^T = \left[-0.1238\ -0.2018\right]\ \left[x_2\ y_2\right]^T = \left[0.1238\ -0.2018\right]^T$
20	$\left[x_1\ y_1\right]^T = \left[-0.1238\ -0.2014\right]\ \left[x_2\ y_2\right]^T = \left[0.1238\ -0.2014\right]^T$
65	$\left[x_1\ y_1\right]^T = \left[-0.1238\ -0.2014\right]\ \left[x_2\ y_2\right]^T = \left[0.1238\ -0.2014\right]^T$

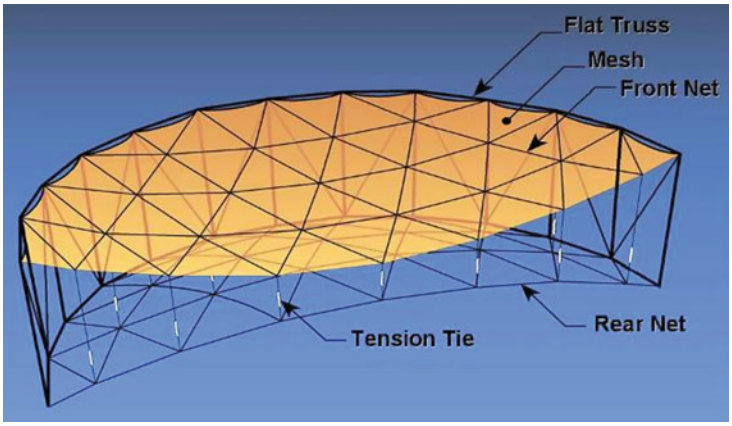


Fig. 6.4 An AstroMesh DMR with deployed working surface

6.6.3 Form Finding of a Large Deployable Mesh Reflector of 865 Nodes

Deployable mesh reflectors (DMR) are high-precision lightweight cable network structures with important space applications [46–49]. In this third example, the proposed FNPM is applied to the structure design of a DMR, as shown in Fig. 6.4. The DMR is a model of AstroMesh DMRs [50]. Upon full deployment, the DMR is supported by a stiff and stable flat truss [27]. The front net (working surface) in the figure, as well as the rear net, is constructed by a mesh of flat triangular facets. The edges of the facets are elastic cable elements interconnected at facet nodes. The nodes of the front and rear nets are also connected by tension ties, the lengths of which can be adjusted. The deployment and operation of such a DMR is described as follows. In deployment of the DMR, folded nets are deployed into highly stretched elastic meshes, with the length of the tension ties properly adjusted such that the facets of the front net eventually form a working surface that is approximate to the desired radio-frequency surface.

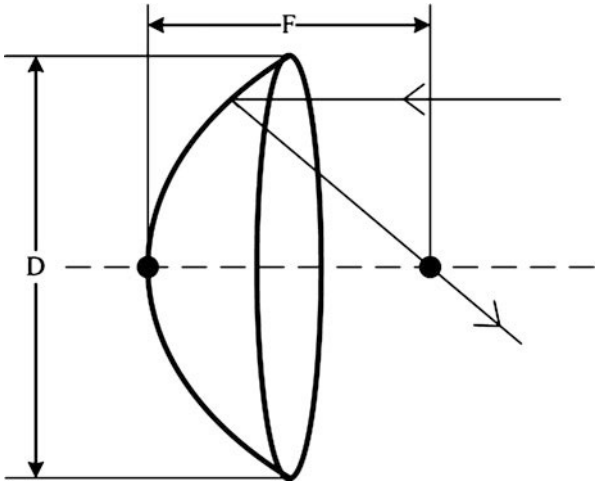


Fig. 6.5 Desired working surface of a deployable mesh reflector

A fully deployed DMR surface can be modeled as a cable network structure, with its members only sustaining axial tensions. Obviously, modeling and design of such a structure must deal with both geometric and material nonlinearities. Let the desired working surface be a paraboloid (see Fig. 6.5), where F and D are the specified focal length F and aperture diameter D in design, respectively. The objective of the DMR design is to obtain a mesh that is as close to the desired working surface as possible. The closeness or surface accuracy is measured by root mean square (RMS) error δ_{rms}, which can be computed by a method proposed by Agrawal, Anderson, and Card [51] and modified by Yuan et al. [52, 53].

In this example, a center-feed and parabola-shaped DMR of 865 nodes is to be designed. The DMRs have a focal length F of 6 m and an aperture diameter D of 12 m. The following uniform internal force distribution is desired

$$T_{\text{des}} = \left[10\ 10 \cdots 10 \right]^T \tag{6.93}$$

Here, the unit of the member internal forces is Newton (N). Due to high accuracy requirement for the reflecting (working) surface of a DMR, the RMS error is required to be

$$\delta_{\text{rms}}(x, y, z) \leq \frac{\lambda}{K} \tag{6.94}$$

where λ is the operating wavelength of the reflector; K is an integer that is usually chosen as $K = 50$, 75, 100, 150 or 200. All the nodes are required to be on a parabolic working surface, which is described by

$$z - H_p = -\frac{1}{4F}\left(x^2 + y^2\right) \tag{6.95}$$

where $H_p = D^2/(16F)$ is the height of desired working surface; and x, y, and z are nodal coordinates.

In operation of a DMR, signal transmission or reflection takes place only on an *effective region*, which is a major portion of the working surface excluding the boundary layer. The boundary layer consists of those boundary nodes that are on the aperture rim and those members that are connected to the boundary nodes. (See Fig. 6.7 for an effective region and a boundary layer of the DMR in consideration.)

To avoid slack and instability and to maintain high surface accuracy, the cable members in the effective region are required to be tensioned in a certain range. On the other hand, the cable members in the boundary layer, which are not in the effective region, can be assigned tensions in a larger range. For the DMR in consideration, the requirements for the internal forces are as follows.

$$7 \leq T_{\text{in}} \leq 14,\ 5 \leq T_b \leq 25 \tag{6.96}$$

where T_{in} is a vector of the internal forces of the members in the effective region; and T_b is a vector of the internal forces of the members in the boundary layer. The unit of member internal forces is Newton (N). The total internal force vector can be written as $T = \left[T_{\text{in}}^T \ T_b^T \right]^T$.

By the FNPM, the initial guess of the geometric configuration of the DMR is made by a systematic mesh generation method, which is namely the pseudo-geodesic method and proposed by Shi et al. [20]. In this method, the mesh geometry is firstly generated on a properly defined spherical reference surface. Then the nodes and members are projected vertically onto the working surface. This method can produce the mesh geometry for a DMR with almost uniform mesh facets (small member length ratio), minimum total member length, competitive surface RMS errors, and large effective regions, for spherical and parabolic working surfaces and with center-feed and offset-feed configurations.

With the geometric configuration generated by the pseudo-geodesic method, the FNPM-based form-finding procedure described in Sect. 6.3 yields a geometric configuration and an internal force distribution for the DMR that satisfy all the geometric and force constraints.

Due to the high complexity of the structure, direct form finding in three dimensions is not available for traditional methods. One of the most commonly used method was an indirect approach developed by Morterolle et al. [54], in which the mesh was generated by force density method in the xy-plane and then projected in the z-direction onto the desired working surface. Desired uniform internal force distribution is then obtained through iterations. In this example, the geometric configurations of the DMRs are also generated by this force density method for comparison.

The final configuration of the 865-node DMR obtained by the FNPM is plotted in Fig 6.6. Also, the reflecting (working) surfaces of the DMR in top view obtained by FNPM and the force density method are shown in Fig. 6.7.

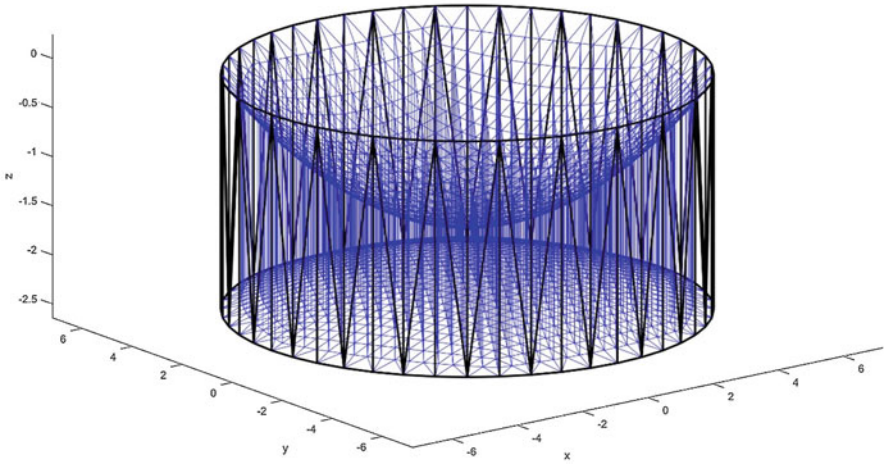


Fig. 6.6 Geometric configuration of the DMR with 865 nodes designed by FNPM

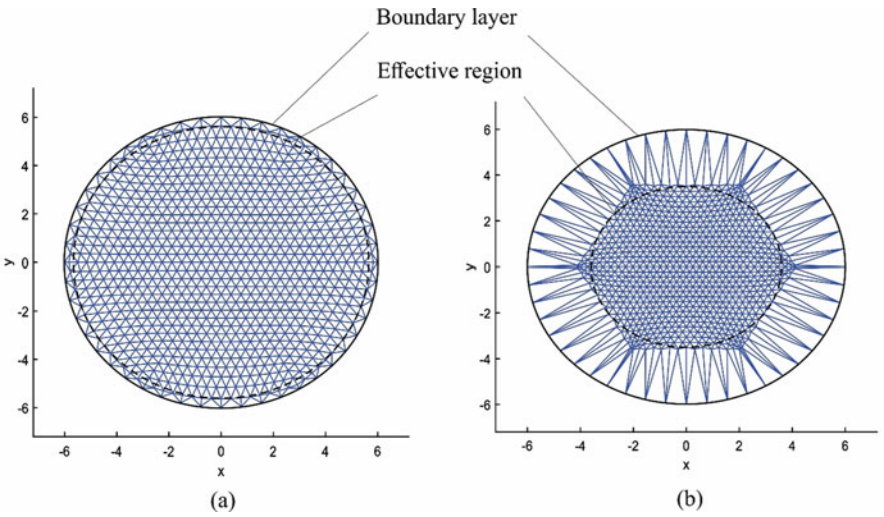


Fig. 6.7 Top view of the DMR with 865 nodes: (**a**) designed by FNPM and (**b**) designed by force density method [54]

The numerical results of form finding are summarized in Table 6.7, where T_{max_in} and T_{min_in}, are the largest and smallest member internal forces in the effective region, shown in Fig. 6.7, respectively; T_{max_b} and T_{min_b} are the largest and smallest member internal forces the boundary layer, respectively; δ_{rms} is the RMS error in millimeter (mm); S_{eff} is the area of effective region in square meters (m^2); and κ is the algorithm accuracy defined in the second example. It is seen from Table 6.7 that the proposed FNPM delivers better form-finding results in both obtaining area of

Table 6.7 Form finding results for the center-feed DMR of 865 nodes

	DMR of 865 nodes designed by force density method [54]	DMR of 865 nodes designed by FNPM
T_{max_in}	9.5908 N	13.9727 N
T_{min_in}	8.1940 N	7.0713 N
T_{max_b}	11.6508 N	20.7099 N
T_{min_b}	11.0450 N	10.3792 N
δ_{rms}	14.2 mm	1.6 mm
S_{eff}	38.4845 m^2	95.0332 m^2
κ	8.4356×10^{-12}	2.4086×10^{-12}

effective region 147% larger and yielding RMS errors 87% smaller than the force density method. The internal force distribution assigned by FNPM is not as uniform as force density method. However, the internal forces of the members vary from 7 to14N in the effective region and from 10 to 21N on the boundary layer, which meets the design requirements.

The convergence process of both methods is shown in Fig. 6.8. The objective function of FNPM and force density method is selected as $\|T - T_{\text{des}}\|/\|T_{\text{des}}\|$. It is observed that both of the two form-finding methods converge in a super linear rate. According to the simulation results, the force density method takes 83 iterations to converge to the value 0.0660. The FNPM only takes 20 iterations to converge at 0.1622. Note that the final objective function values of the two methods are different since the force density method adjusts the nodal coordinates while assigning the internal force distribution and the FNPM treats the nodal coordinates as fixed, such that the desired configuration of the structure is maintained.

6.7 Conclusions

In this chapter, the problem of form finding of cable network structures is presented and several form-finding methods are introduced. The main results contained in the chapter are summarized as follows.

(i) For FF-IEC problems, a new method called the fixed nodal position method (FNPM) is developed. Different from conventional form-finding methods, the FNPM determines the internal force distribution of a cable network structure without changing prescribed nodal coordinates. This unique feature of the new method makes it possible to place the nodes of a cable network structure rather freely at desired locations, satisfying complicated geometric and force constrains and yielding high shape/surface accuracy as required in design.

(ii) In the FNPM-based form finding, the nodal positions of a cable network structure are first assigned (at desired locations), and the internal force distribution is determined via numerical optimization. With this displacement-first-and-

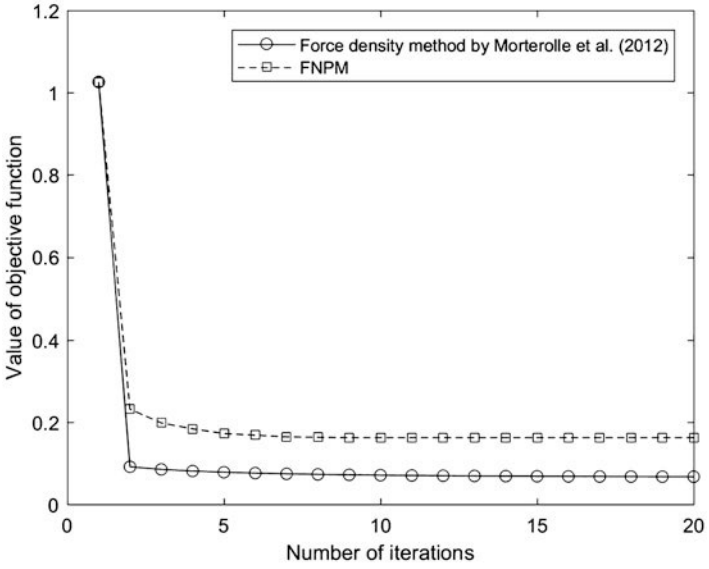


Fig. 6.8 History of convergence for the DMR of 865 nodes

stress-later process, the new method requires much less computational effort than conventional form-finding methods. Moreover, in form finding, the FNPM guarantees a global minimizer if the objective function is of quadratic form.

(iii) For FF-IEC problems, two commonly used methods, namely, the force density method and the dynamic relaxation method, are introduced. For FF-DEC problems, three methods are introduced, and they are the singular value decomposition method, the stiffness matrix method and the dynamic relaxation method.

(iv) The form-finding methods introduced are demonstrated in three numerical examples. In the first two examples, the methods of FF-IEC and FF-DEC are illustrated in design of 2-D cable network structures. In the last example, the FNPM and the force density method are used in design of a large deployable mesh reflector of 865 nodes. Numerical simulations show that the FNPM can deliver results that converge in a super linear rate, have extreme high accuracy, and satisfy all the geometric and force constraints. This indicates that the newly developed method is more suitable than traditional methods in form finding of complicated structures with stringent geometric and force constraints.

Appendix

Global Minimizer and Solution of the Optimization Problem (6.12)

This appendix shows that the solution of internal forces given by Eq. (6.13) and (6.14) is a global minimizer of the optimization problem (6.12).

The general solution of Eq. (6.2) for truss structures of types III and IV with consistent external forces is given by

$$T = -M^{+}F_c + M_{\text{null}}\alpha \tag{A1}$$

where M_{null} is the null space of matrix M and α is an arbitrary s by 1 vector, with s being the states of self-stress. Substituting the solution T into the quadratic objective function in the optimization problem (6.12) yields

$$f = \|T - T_{\text{des}}\|^2 = \alpha^T \left(M_{\text{null}}^T M_{\text{null}}\right)\alpha - 2C^T M_{\text{null}}\alpha \tag{A2}$$

where C is a constant vector given by

$$C = M^{+}F_c + T_{\text{des}} \tag{A3}$$

Consider the following lemma, which is proved in [55].

Lemma
Let f_q be a quadratic function given by

$$f_q = \alpha^T B\alpha + 2g^T\alpha \tag{A4}$$

where α is a vector of variables, and B is any symmetric matrix. The following two statements hold true:

(1) Function f_q attains a minimum if and only if B is positive semi-definite and g is in the range of B. If B is positive semi-definite, then every α satisfying $B\alpha = -g$ is a global minimizer of f_q.
(2) Function q has a unique minimizer if and only if B is positive definite.

Note that function f in Eq. (A2) has the same form as function f_q in (A4), with $g = -C^T M_{\text{null}}$ and $B = M_{\text{null}}^T M_{\text{null}}$, which is positive definite because all the column vectors M_{null} are linearly independent. By the lemma, function f has the unique global minimizer given by

$$\alpha^* = B^{-1}g = \left(M_{\text{null}}^T M_{\text{null}}\right)^{-1}\left(C^T M_{\text{null}}\right) = M_{\text{null}}^T T_{\text{des}} \tag{A5}$$

With α^*, the internal force distribution T^*, which is the minimum deviation from the desired internal force distribution T_{des}, is then obtained as

$$T^* = -M^+ F_c + M_{\text{null}} \alpha^* \tag{A6}$$

References

1. J.H. Argyris, T. Angelopoulos, B. Bichat, A general method for the shape finding of lightweight tension structures, Computer Methods in Applied Mechanics and Engineering, 3 (1974) 135–149.
2. S. Yuan, B. Yang, H. Fang, Self-Standing Truss with Hard-Point-Enhanced Large Deployable Mesh Reflectors, AIAA Journal, 57 (2019) 5014–5026.
3. S. Yuan, B. Yang, H. Fang, Enhancement of Large Deployable Mesh Reflectors by the Self-Standing Truss with Hard-Points, in: AIAA Scitech 2019 Forum, 2019, pp. 0752.
4. S. Yuan, B. Yang, The fixed nodal position method for form finding of high-precision lightweight truss structures, International Journal of Solids and Structures, 161 (2019) 82–95.
5. S. Pellegrino, Structural computations with the singular value decomposition of the equilibrium matrix, International Journal of Solids and Structures, 30 (1993) 3025–3035.
6. R. Haber, J. Abel, Initial equilibrium solution methods for cable reinforced membranes part I—formulations, Computer Methods in Applied Mechanics and Engineering, 30 (1982) 263–284.
7. S. Pellegrino, Mechanics of kinematically indeterminate structures, in, University of Cambridge, 1986.
8. G. Tibert, Deployable tensegrity structures for space applications, Doctoral dissertation, KTH, 2002.
9. S. Pellegrino, Analysis of prestressed mechanisms, International Journal of Solids and Structures, 26 (1990) 1329–1350.
10. C. Calladine, Buckminster Fuller's "tensegrity" structures and Clerk Maxwell's rules for the construction of stiff frames, International Journal of Solids and Structures, 14 (1978) 161–172.
11. J. Zhang, M. Ohsaki, Adaptive force density method for form-finding problem of tensegrity structures, International Journal of Solids and Structures, 43 (2006) 5658–5673.
12. R.B. Fuller, Synergetics: explorations in the geometry of thinking, Estate of R. Buckminster Fuller, 1982.
13. R. Motro, Structural morphology of tensegrity systems, International Journal of Space Structures, 11 (1996) 233–240.
14. R. Motro, H. Nooshin, Forms and forces in tensegrity systems, in: Proceedings of third international conference on space structures, Elsevier, Amsterdam, 1984, pp. 180–185.
15. A. Tibert, S. Pellegrino, Review of form-finding methods for tensegrity structures, International Journal of Space Structures, 26 (2011) 241–256.
16. K. Koohestani, Form-finding of tensegrity structures via genetic algorithm, International Journal of Solids and Structures, 49 (2012) 739–747.
17. D. Veenendaal, P. Block, An overview and comparison of structural form finding methods for general networks, International Journal of Solids and Structures, 49 (2012) 3741–3753.
18. S. Pellegrino, C.R. Calladine, Matrix analysis of statically and kinematically indeterminate frameworks, International Journal of Solids and Structures, 22 (1986) 409–428.
19. M. Masic, R.E. Skelton, P.E. Gill, Algebraic tensegrity form-finding, International Journal of Solids and Structures, 42 (2005) 4833–4858.
20. H. Shi, S. Yuan, B. Yang, New Methodology of Surface Mesh Geometry Design for Deployable Mesh Reflectors, Journal of Spacecraft and Rockets, 55 (2018) 266–281.

21. F.E. Udwadia, R.E. Kalaba, Analytical dynamics: a new approach, Cambridge University Press, 2007.
22. S. Yuan, B. Yang, Design and Optimization of Tension Distribution for Space Deployable Mesh Reflectors, in: 26th AAS/AIAA Space Flight Mechanics Meeting, Univelt, Napa, CA, 2016, pp. 765–776.
23. K.E. Jensen, Numerical Optimization in Microfluidics, in: Complex Fluid-Flows in Microfluidics, Springer, 2018, pp. 95–108.
24. H. Li, Z. Liu, K. Liu, Z. Zhang, Predictive Power of Machine Learning for Optimizing Solar Water Heater Performance: The Potential Application of High-Throughput Screening, International Journal of Photoenergy, 2017 (2017).
25. M. Hintermüller, K. Ito, K. Kunisch, The primal-dual active set strategy as a semismooth Newton method, SIAM Journal on Optimization, 13 (2002) 865–888.
26. A. Tibert, S. Pellegrino, Deployable tensegrity reflectors for small satellites, Journal of Spacecraft and Rockets, 39 (2002) 701–709.
27. S. Yuan, B. Yang, H. Fang, Form-Finding of Large Deployable Mesh Reflectors with Elastic Deformations of Supporting Structures, in: 2018 AIAA Spacecraft Structures Conference, 2018, pp. 1198.
28. I.K. Linkwitz, H.-J. Schek, Einige bemerkungen zur berechnung von vorgespannten seilnetzkonstruktionen, Ingenieur-Archiv, 40 (1971) 145–158.
29. H.-J. Schek, The force density method for form finding and computation of general networks, Computer methods in applied mechanics and engineering, 3 (1974) 115–134.
30. L. Gründig, E. Moncrieff, P. Singer, D. Ströbel, A history of the principal developments and applications of the force density method in Germany 1970–1999, 4th Int. Coll. Computation of Shell & Spatial Structures, (2000).
31. J. Sánchez, M.Á. Serna, P. Morer, A multi-step force–density method and surface-fitting approach for the preliminary shape design of tensile structures, Engineering structures, 29 (2007) 1966–1976.
32. G. Aboul-Nasr, S.A. Mourad, An extended force density method for form finding of constrained cable nets, Case Studies in Structural Engineering, 3 (2015) 19–32.
33. M. Barnes, Form-finding and analysis of tension space structures by dynamic relaxation, in, City University, 1977.
34. A. Day, J. Bunce, The analysis of hanging roofs, Arup Journal, (1969) 30–31.
35. N.B.H. Ali, L. Rhode-Barbarigos, I.F. Smith, Analysis of clustered tensegrity structures using a modified dynamic relaxation algorithm, International Journal of Solids and Structures, 48 (2011) 637–647.
36. M. Barnes, Form finding and analysis of tension structures by dynamic relaxation, International journal of space structures, 14 (1999) 89–104.
37. R. Motro, S. Belkacem, N. Vassart, Form finding numerical methods for tensegrity systems, in: Spatial, Lattice and Tension Structures, ASCE, 1994, pp. 704–713.
38. A. Tibert, S. Pellegrino, Review of form-finding methods for tensegrity structures, International Journal of Space Structures, 26 (2011) 241–255.
39. R. Connelly, M. Terrell, Globally rigid symmetric tensegrities, Structural Topology 1995 núm 21, (1995).
40. J. Argyris, D.W. Scharpf, Large deflection analysis of prestressed networks, Journal of the Structural Division, 98 (1972) 633–654.
41. P. Krishna, Cable-suspended roofs, McGraw-Hill Companies, 1978.
42. H. Murakami, Static and dynamic analyses of tensegrity structures. Part 1. Nonlinear equations of motion, International Journal of Solids and Structures, 38 (2001) 3599–3613.
43. H. Murakami, Static and dynamic analyses of tensegrity structures. Part II. Quasi-static analysis, International Journal of Solids and Structures, 38 (2001) 3615–3629.
44. S. Guest, The stiffness of prestressed frameworks: a unifying approach, International Journal of Solids and Structures, 43 (2006) 842–854.
45. C. Fletcher, Computational techniques for fluid dynamics 2: Specific techniques for different flow categories, Springer Science & Business Media, 2012.

46. R. Nie, B. He, L. Zhang, Y. Fang, Deployment analysis for space cable net structures with varying topologies and parameters, Aerospace Science and Technology, 68 (2017) 1–10.
47. R. Nie, B. He, L. Zhang, Deployment dynamics modeling and analysis for mesh reflector antennas considering the motion feasibility, Nonlinear Dynamics, 91 (2018) 549–564.
48. K. Nakamura, N. Nakamura, S. Ozawa, A. Uematsu, H. Hoshino, T. Kimura, Concept Design of 15m class Light Weight Deployable Antenna Reflector for L-band SAR Application, in: 3rd AIAA Spacecraft Structures Conference, 2016, pp. 0701.
49. Y. Tang, T. Li, X. Ma, Pillow Distortion Analysis for a Space Mesh Reflector Antenna, AIAA Journal, (2017) 3206–3213.
50. M.W. Thomson, Astromesh deployable reflectors for Ku-and Ka-band commercial satellites, in: 20th AIAA International Communication Satellite Systems Conference and Exhibit, 2002, pp. 15.
51. P. Agrawal, M. Anderson, M. Card, Preliminary design of large reflectors with flat facets, IEEE transactions on antennas and propagation, 29 (1981) 688–694.
52. S. Yuan, B. Yang, H. Fang, The Projecting Surface Method for improvement of surface accuracy of large deployable mesh reflectors, Acta Astronautica, 151 (2018) 678–690.
53. S. Yuan, B. Yang, H. Fang, Direct Root-Mean-Square Error for Surface Accuracy Evaluation of Large Deployable Mesh Reflectors, in: AIAA Scitech 2020 Forum, 2020, pp. 0935.
54. S. Morterolle, B. Maurin, J. Quirant, C. Dupuy, Numerical form-finding of geotensoid tension truss for mesh reflector, Acta Astronautica, 76 (2012) 154–163.
55. J. Nocedal, S. Wright, Numerical optimization, Springer Science & Business Media, 2006.

Part II
Applied Design of Engineering Problems

Chapter 7
Application of Genetic Algorithm in Characterisation of Geometry Welds in Spot Weld Process Design

Varshan Beik, Hormoz Marzbani, and Reza N. Jazar

Abstract Vehicle body is the most integral factor shaping the perception of the design and brand image, and arguably it is the biggest contributor to the vehicle dimensional outcome. The weld process design is one of the key influencers to dimensional quality of the body. An optimum process design will ensure the minimum process-driven dimensional variations and hence improved dimensional quality.

Normally in a body production line, there is one geometry setting station for each segment of the process and then a few re-spot stations to complete the welds as per the product design. In the geometry setting station, adequate number of welds and a well-designed weld distribution (optimum process design) is required to ensure the assembly is dimensionally set and fixed. This is not an easy task in high-rate production lines due to the limited available cycle time. This requires the geometry welds to be characterised and prioritised when designing spot weld process. This challenge is more prominent in large and complicated structures such as car body shell. In this research, we aim to characterise the geometry setting welds of body shell as the most complicated structure in vehicle body. This will enable the optimum process design which delivers the minimum dimensional variation.

We first reviewed the assembly sequence planning problems and the application of different methods in solving these problems. We then assessed the application of genetic algorithm in the field and developed three genetic algorithm models including a simple structure, the bodyside assembly, and finally the body shell. The result was significant, and around 99% of non-viable process designs were eliminated. We concluded that the application of genetic algorithm can replace the existing process design methods which mostly rely on individual's experience to design an optimum process.

V. Beik (✉) · H. Marzbani · R. N. Jazar
School of Aerospace, Mechanical, and Manufacturing Engineering, RMIT University, Melbourne, VIC, Australia
e-mail: s3403270@student.rmit.edu.au

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_7

Keywords Genetic algorithm · Vehicle body design · Dimensional quality · Optimum product design · Vehicle spot welds position

7.1 Introduction

Automotive companies face immense competition and a demanding market, so their primary goal is to increase company revenue by producing new products that meet the market demand and are also financially profitable for the corporation.

Krishnan and Ulrich defined product development as "the transformation of a market opportunity and a set of assumptions about product technology into a product ready for sale" [1]. Product development processes are the methods that are used to develop financially profitable products based on the market demand [2].

The traditional serial design approach requires the information from one discipline and/or process to proceed with the subsequent design phase. The implementation of concurrent engineering (CE) and integrated product and process development (IPPD) processes tries to alleviate this problem by allowing process design to be performed in parallel and collaboratively along with the product development. In complex designs with high-quality requirements, geometrical variation control is a great challenge [3]. Geometrical variations are inevitable in the production due to multiple factors. It is essential to optimise the body process design for minimising the process-driven dimensional variation. To better understand it, first we need to have a closer look at the vehicle body structure and process.

The vehicle body is arguably the biggest contributor to the vehicle dimensional outcome and is the most integral factor shaping the perception of the design and brand image. Hundreds of sheet metal parts are joined and assemble the bodyshell (Fig. 7.1). While various welding methods can be used to join these parts, spot welding is the most common technique used to join sheet metals in the automotive industry due to the fast rate of production [4].

The assembly of the body-in-white (BIW) results in the geometrical formation of key body components, and precision in this step is crucial as deviations from the original design would lead to possible systems malfunctioning, added manufacturing cost, and prolonged production time.

The behaviour of an assembly system is never completely predictable as many factors affect the final geometry including the assembly sequence, spot welds position and quantity, as well as the dimensional quality sheet metal parts (Fig. 7.2) [5].

Variation in each of these elements impacts the distortion in the final BIW geometry. These interactions are extremely complex, and optimisation of each of these factors using advanced simulation tools is crucial to predict critical product dimensions. An optimal simulation tool predicts and simulates the propagation pattern of variations of the system early in the process design stage. Despite extensive work that has been done on variation factors and simulation methods, current methods fall short of satisfaction [6]. These traditional methods are not applicable

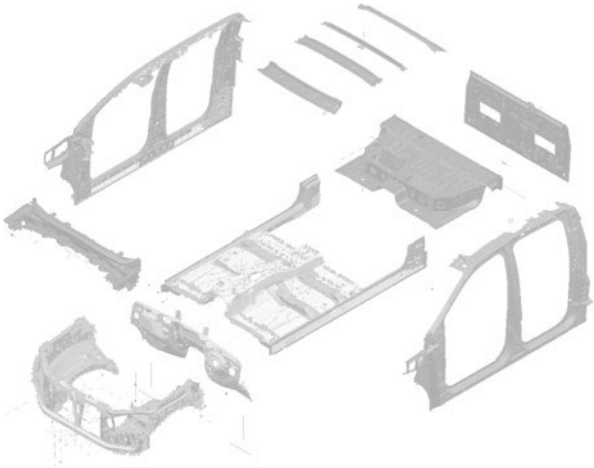


Fig. 7.1 Standard main assemblies shaping a vehicle bodyshell

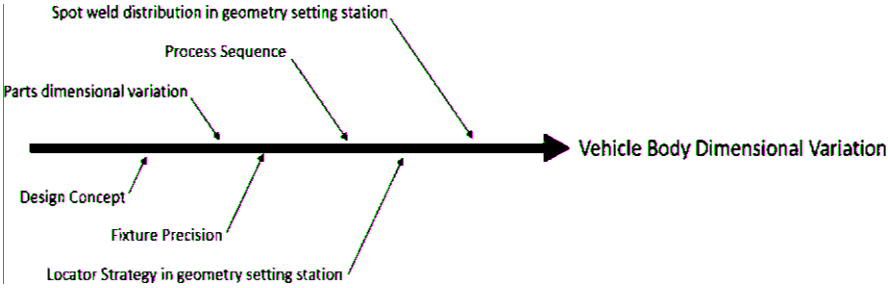


Fig. 7.2 Main contributors to vehicle body dimensional variations

for compliant sheet metal assembly due to possible variations in dimensions during the process. Moreover, existing tools fail to address a multi-levelled production process and are mostly individual machine focused.

7.1.1 Statement of the Problem

From the contributors to the body dimensional variation reviewed in the prior section (Fig. 7.2), spot weld distribution (or called as "process design") in geometry setting station is a hidden factor and a major contributor to the body dimensional variations. This is more significant in high-rate production lines where the station cycle time is very limited. Thus, not many welds can be put in the geometry setting station to dimensionally set the assembly. This process inefficiency will result in dimensional variation of the produced assemblies, which in this research we refer to it as "process driven dimensional variation". So, designing an optimum process or weld distribution can make a huge difference in the dimensional quality of the

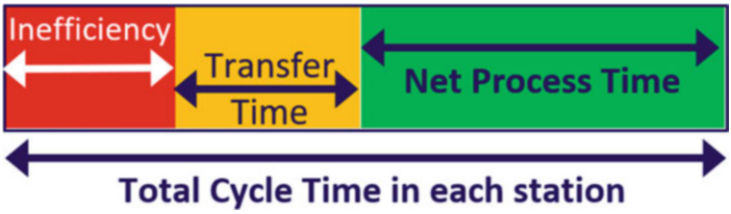


Fig. 7.3 Common elements of cycle time in a body shop welding station

assembly. To clearly understand the problem, let us first review a few definitions in the body shops and then we discuss the problem in more details.

Production Line Rate
Line rate is the number of bodies that can be produced (usually) during an hour. For example, in a production line with a rate of 45 JPH (job per hour), there will be 45 bodies produced each hour.

Gross Cycle Time
Gross cycle time is the total time from the beginning to the end of the process in each station which includes loading parts/assemblies into the fixture, fixture closing, process time (e.g. weld, fastening, sealing, etc.), fixture opening, unloading parts/assemblies, transferring from station to the next station, etc. plus all inefficiency factors (e.g. facility uptime, operator availability, etc.).

Net Cycle Time
Net cycle time is the pure available cycle time excluding inefficiency factors (Fig. 7.3). It equals the gross cycle time multiplied by operator availability (usually 90% and it is there to cover the possible operator's delays or not being available) and multiplied by the facility uptime (usually 85% and it is there to cover the facility break down or urgent maintenance time).

Weld Time
Weld time is the total amount of time required from start to finish a spot weld, including weld gun/robot approaching the weld position, welding, and moving away from the weld position.

One of the main dimensionally critical steps in a body manufacturing line is the framing line in which the main assemblies such as bodysides, roof structure, and underbody are put and joined together in multiple stations to shape the body shell with approximately 500–1000 welds, depending on the size and style of the body (Fig. 7.4).

In a framing line, usually the first geometry setting station is called "framing station" and is the key station in which the geometry is to be set and finalised. In a framing station, the main assemblies are precisely located in the nominal coordinates. This is the only station that all main assemblies are precisely located and as soon as the tooling (pins, locators, etc.) are released and the body moves to

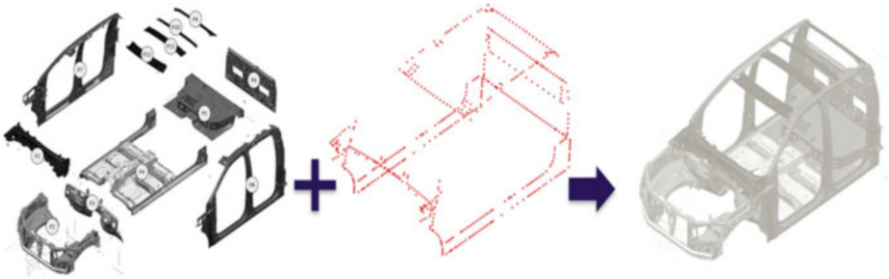


Fig. 7.4 Hundreds of welds connect the main assemblies together to shape the bodyshell

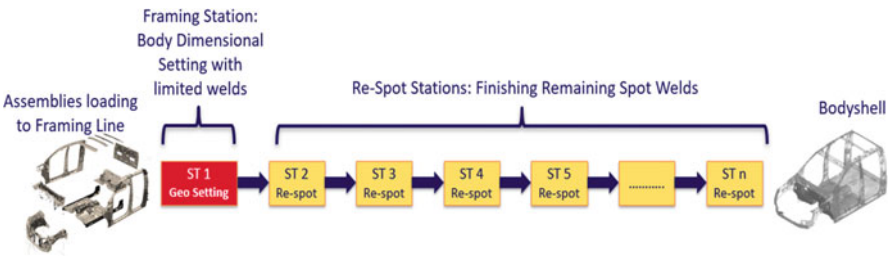


Fig. 7.5 Body framing line

the next station, due to the natural spring back and displacement of the parts, it is not possible to pin or precisely locate the assemblies again in the downstream stations. So, if the body dimension is not fully set and fixed in the framing station, then in the following re-spot stations, with no tooling to locate the assemblies, the welding process will significantly influence the dimension of the body resulting in what we call "process-driven dimensional variations", as it is due to the process inefficiency. Due to the nature of this type of dimensional variations and the challenges in root cause analysis of the defects in upstream processes, it is usually impossible to trace the process-driven dimensional defects. In a well-designed process, the dimension of the body is fully set and fixed in the framing station, so any re-spotting process in the following stations will not impact the dimension of the body (Fig. 7.5).

Now that we reviewed the basics, let us review a simple cycle time and station capacity calculation. The question is how to calculate the station weld capacity in a 45 JPH body production line with a facility uptime of 85% and the availability of 90%. Assume the weld time for each spot weld is 4 s (industry standard).

Production line rate = 45 JPH (meaning 45 bodies have to be produced each hour).

Gross cycle time in each station = 3600/45 = 80 s.

Net cycle time = 80 s $\times$90% $\times$ 85% = 61 s (meaning every 61 second one body has to be completed and leaving the line, and this results in producing 45 bodies in 1 h).

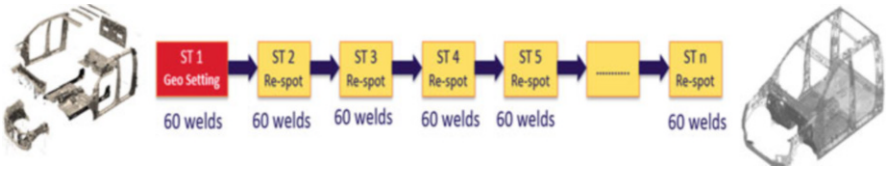


Fig. 7.6 Body Framing diagram of a 45 JPH production line

Total weld time = Net cycle time – Transfer /fixture opening & closing = 61−21 = 40 s.

Assuming weld time for a robot is 4 seconds per weld, then:

$$\text{Weld capacity for each robot} = \frac{\text{Total available weld time}}{\text{Weld time for each spot weld}} = \frac{40}{4} = 10 \text{ welds}$$

This means one robot can put a maximum of ten welds in each cycle. Assuming six robots in a framing station, then the total weld capacity of each station would be:

Station weld capacity = Number of robots in a station x Number of welds per robot = 6 × 10 = 60 welds.

Figure 7.6 illustrates the framing line layout and a simplified weld quantity distribution in each station. In real examples of process design in the industry, the number of welds in each station may not necessarily be the same in all stations as it depends on some other factors such as weld access, gun availability, etc.

The above simplified calculation shows how limited weld capacity in high-rate production lines could be. Now that we understand the weld capacity constrain, we can look at the process design in the framing station.

Problem Definition

In a 45 JPH line of a product which has 850 spot welds to properly join the assemblies in a framing line, a process has to be designed to fully set and dimensionally fix the body with only 60 welds (station weld capacity as per the above calculation). Now the question is that which 60 welds out of the total 850 framing welds can best dimensionally fix the bodyshell. As mentioned before, this is to minimise the dimensional variation on the body. Mathematically there are unlimited answers to this question.

$$\text{Number of possible designs} = \binom{n}{r} \tag{7.1}$$

where “n” is the total number of welds and “r” is the weld capacity of framing station

$$= \binom{850}{60} = \frac{850!}{60!\,(850-60)!} = 8.293419958\text{E}+92.$$

The Industry Gap

Currently the industry mostly relies on individuals' experiences as the main tool to design the body processes. This brings lots of uncertainty and errors, especially in complicated structures such as body framing, and will result in untraceable dimensional variations and vehicle quality issues and consequently additional cost of rectifying the defects. The lack of an engineering solution to design the optimum process that can minimise the dimensional variation in high-rate production lines is apparent.

To ensure our research outcome will benefit the industry, we collaborate with Ford Motor Company to be able to study a real industry problem. This allowed us to work on a real industry challenge, using real product data. The confidentiality deed was signed between Ford and RMIT, and we were given access to some product data. Due to confidentiality and propriety of the data used in this research, throughout this thesis, we only present the findings of our work and will not be able to share any product data that may seem breach of confidentiality.

7.1.2 Research Objectives

In this study, we aim to fully understand one of the most challenging industry gaps in the vehicle body process design. We will apply engineering optimisation methods and CAE to deliver a scientific approach for designing an optimum process to minimise the vehicle body dimensional variation.

The main objectives of this study are:

1. Understanding the vehicle body framing process design and the industry gap
2. Application of genetic algorithm in characterisation of vehicle body geometry setting welds
3. Application of finite element method in assessing the process-driven dimensional variation in the vehicle body
4. Developing a set of design guidelines to enable an optimum process design

7.1.3 Research Questions

After understanding the objectives of the study, it is now time to focus on the research questions. There are many questions that we need to find answers through this research, but the key questions that are critical to successfully achieve the research objectives are as below:

1. How can genetic algorithm help to improve the vehicle body process design?
2. In a vehicle body framing process design, which welds act as geometry setting welds?

3. How can FEM be used to assess the impact of process design on the body dimensional variation?
4. What set of rules could be generated to improve vehicle body process design and minimise the process driven dimensional variations.

   – Developing a set of guidelines to minimise the process-driven dimensional variation based on genetic algorithm assessment and finite element analysis.

Figure 7.7 illustrates a simplified view of the proposed method for optimisation of the body welding process design. The optimal design would have the minimum process-driven dimensional variation.

7.2 Background

7.2.1 Assembly Sequence Planning (ASP)

Assembly planning (AP) is the science of optimising the assembly process of some parts in order to achieve the most feasible way of manufacturing the product [7]. The assembly planning problem has two main aspects: (1) assembly sequence optimisation and (2) optimal allocation of each part on the final product [8].

Assembly sequence planning (ASP) is a critical element of AP. Assembly sequence directly impacts the overall productivity and cost of the assembly process through defining degrees of freedom (DOF), need for fixtures, and the possibility of concurrent testing. To address the ASP problem, it needs to be defined, solved, and optimised. Optimisation refers to the process of choosing the best available solution when more than one result is generated from solving the ASP problem.

7.2.1.1 Assembly Plans

Assembly plans can be of various types [9]. These include coherent/non-coherent, linear/non-linear, sequential/non sequential, monotone/non-monotone, or a combination of the mentioned types [10]. A coherent (C) assembly plan is the one in which every part has direct contact with a previous part (apart from the first part). In a linear (L) assembly plan, all parts are added one by one. A sequential (S) plan is characterised by a step-wise approach to addition of elements. A monotone (M) assembly is defined where each part is placed into its final position regardless of the rest of the assembly. As a result, in a monotone assembly with “m” parts, “m-1” moves are needed to complete the process [10].

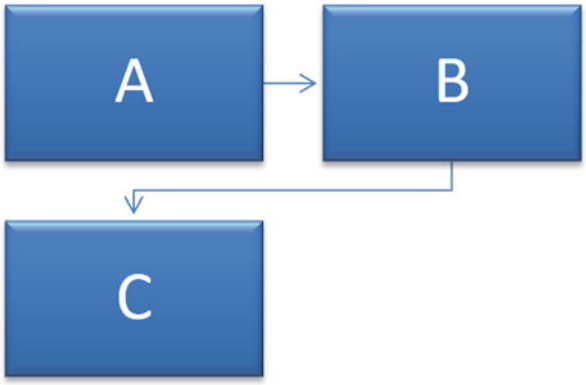


Fig. 7.7 Assembly precedence

7.2.1.2 The Assembly Sequence Planning Problem

The ASP problem is a large scale, NP-hard (nondeterministic polynomial time) combinatorial problem where the number of possible solutions expansively increases with the increase of system components [11]. If a product has "n" components, the number of possible assembly sequences would be n!, and when the components increases to n + 1, the total possible assembly sequences increases to (n + 1)!. As this includes all potential assembly sequences, it is essential to define all feasible sequences, ensuring all possibilities have been considered, and as a result the best option is chosen and the assembly process is optimised [12].

To achieve feasibility, the ASP is constrained by various predefined factors. Two types of constraints apply when solving ASP problems.

1- Absolute constraints which ensure feasibility of the solution.
2- Optimisation constraints which ensure optimal quality of the chosen solution.

Precedence and geometrical constraints are the two most common absolute constraints in ASP [13].

The precedence constraint addresses the relationship between operations and defines the predecessors and successors (Fig. 7.7). The precedence constraint defines that as A precedes B and A precedes C, so B and C will not take place unless A is finished. It is important to note that this constraint does not completely define the sequence of operation [13].

Geometrical constraints address contact between the components and aims to avoid collision. For an assembly to be feasible, the presence of at least one collision-free way for assembly of the components is essential and its absence will result in violation of the geometrical constraint [13].

In summary solving an ASP problem is the process of finding an assembly sequence that meets both precedence and geometrical constrains. Hence, to define the ASP problem, let us imagine a mechanical system (M) with "n" parts:

(1) $M = \{p1, p2, \ldots., pn\}$.

This system can be assembled through "m" assembly operations ($m \geq n$):

(2) $O = \{o1, o2, \ldots., om\}$.

As explained previously, to achieve feasibility the ASP is constrained by geometrical, and precedence constrains.

Geometrical constraints data: where a geometrical connection (G) is present between two operations it is defined as:

(3) $(oi, oj) \in G \iff (oj, oi) \in G$

Precedence constraints data: where an operation precedes another, it is defined as:

(4) $(oi, oj) \in P$

And if the oj does not precede "oi", it is defined as:

(5) $(oj, oi) \notin P$

Quality characteristics for the mechanical system (A) are defined as:

(6) $QCk, k \in N$.

In order to achieve optimisation, the function that demonstrates the performance of the assembly sequence is defined as:

(7) $F = \sum Fi, Fi = 1, \ldots., m$.

When considering the performance of each assembly based on the quality characteristics (QCk):

(8) $Fi = f(QCk, s1, ..., si\text{-}1), i\text{-}1, ..m, k \in N$;

So, the problem will be to define an assembly sequence.
$S = \{s1, \ldots, si, \ldots, sn\}$.
$si \in A$, where:

- Geometrical constraints are met:
  $\forall si \in S, i = 1, ..., n, \exists sj \in S, j \in N$ and $j < i$, so that $(si, sj) \in G$
- Precedence constraints are met:
  $\forall si, sj \in S, i, j = 1, \ldots, n, (si, sj) \in P$
- Optimisation takes place: maximum value for F is achieved.

7.2.1.3 Solving the Assembly Sequence Planning Problem

After absolute constraints are defined, it is time to solve the ASP problem. There are many methods to generate assembly sequences. The method used is mostly affected by the assembly environment. Solving the ASP problem is the process of generating all feasible assembly sequences. Various methods have been developed over the years to address this (Fig. 7.8).

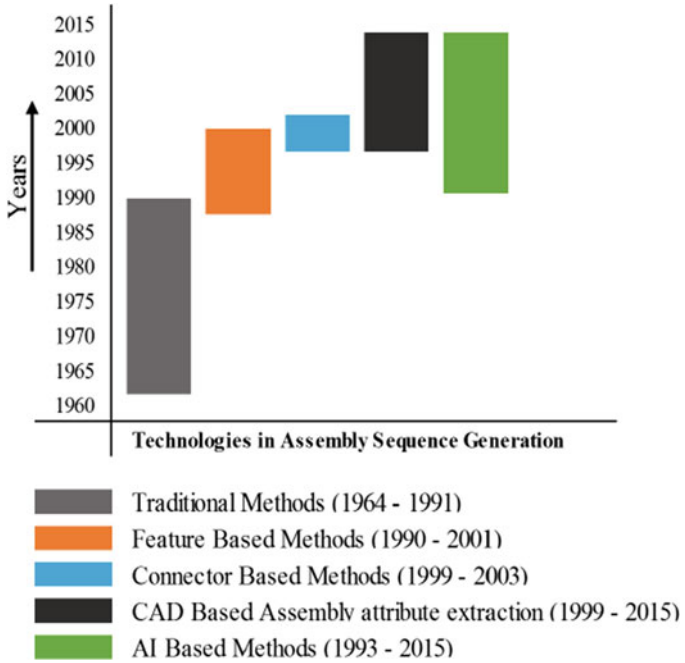


Fig. 7.8 Evolution of assembly sequence generation methods [14]

Traditional methods which evolved in the 1960s were the mainstay of practice until the early 1990s. These methods rely on a highly experienced operator to develop instructions and information that assess feasibility. These methods are extremely time-consuming and require a very large storage space. Feature-based methods use information on the features of the assembly parts to retrieve precedence. These methods are highly reliant on appropriate modelling methodology and skilled user to extract results. Connector-based methods which were developed between 1999 and 2003 are based on the assumption that connectors are primary assembly parts. While these methods guarantee the generation of at least one feasible assembly sequence, they do not cover mechanical feasibility. CAD application greatly changed the interface of assembly planning by offering significant flexibility in retrieval of assembly part attributes. These methods require an operator that has CAD interfacing and geometric reasoning skills. Artificial intelligent techniques used CAD applications to generate assembly sequences. The results are highly dependent on the quality of CAD models. These methods can only generate close to optimal results that are stochastic in nature.

Assembly sequence generating methods can be classified in four trends [14]:

1. Three-step approach
2. Subassembly division
3. Expert systems
4. Case-based reasoning approach

Three-Step Approach

The traditional three-step approach is a general approach and is widely applied to all assembly lines. This approach is based on applying the absolute constraints and finding the feasible sequences that meet these criteria through three steps including: (1) defining precedence constraints, (2) producing all possible assembly sequences, and (3) selection of sequences which meet pre-defined criteria.

Step 1: Defining Precedence Constraints
As mentioned, precedence constraints are considered as absolute constraints which their violation results in infeasibility of the sequence. To define these constraints, three methods can be used: (a) exhaustive method, (b) simplified method, and (c) disassembly method.

(a) *Exhaustive Method*

The exhaustive method uses the information on connection of assembly parts to generate a liaison graph with the aim to map all predecessor and successor relationships (Bourjault's method) [15]. This method is based on the operator answering two questions about each connection in the assembly (Yes/No): (1) Is the connection (L_i) possible to be established in the presence and (2) absence of connection ($L_{j, \ldots,} Lk$).

Lui computerised the Bourjault's method [16]. A software was developed that asks the operator the questions and conducts geometric analysis based on generated liaison graphs and matrices [16]. The fundamental concepts of liaison graphs and tables are explained later in this section. This method is only applicable to assemblies with limited parts and is also operator dependent and hence prone to error.

(b) *Simplified Method*

De Fazio and Whitney aimed to further improve and simplify Bourjault's method by reducing the number of questions asked [17]. The simplified approach uses two questions to define precedence. Which connection must/cannot precede L_i [17]. While this method is much less time-consuming, it is subject to underestimation and omitting some feasible sequences as it is reliant on the user to draw a liaisons graph and answer the questions based on that.

(c) *Disassembly Method (Cut-Set Method)*

The disassembly approach is the most widely used approach in literature. The ability to disassemble a component confirms the precedence relationship. It is much easier to define precedence relationships starting from the final product as all relationships have already been established and can easily be mapped by disconnecting the links. This limits the branching states that should be considered when assessing possible configurations from the end product rather than the initial stage [18]. Homem de Mello and Sanderson [19] applied this approach and aimed to further computerise it. Using a software, they developed an algorithm and an

assembly graph generated by an expert operator, decompositions of the assembly graph were enumerated, and feasible ones were then chosen by the user. This method was semi-automated and as explained involved some user interaction. Baldwin et al. further developed these methods [15, 17, 19] by introducing a computer program that used disassembly analysis to produce all feasible sequences. The program also assisted the user to choose the optimum sequence using predefined criteria [20].

Step 2: Producing all Possible Sequences

An assembly model of a product comprises of various "standard parts" and "connectors" and demonstrates information on the interactions between the parts. The model can be generated once the absolute constraints are defined and characterised. Various techniques can be used to generate all possible assembly sequences. The used methodology is largely based on method used to define the absolute constraints. Such methods to generate assembly sequences include graph and matrix manipulation, case-based search, grouping parts, and knowledge-based approach.

Graphs are commonly used in ASP studies as they guarantee to capture all connections and sequence possibilities in an assembly. Various types of graphs (simple, directed, undirected, etc.) have been widely used in assembly planning and to map connections in assemblies [21]. To apply this method using a graph-based algorithm, the ASP problem is modelled with the aim to develop feasible solutions. A simple directed graph consists of a group of nodes (vertices) and edges (connections between objects with directions). Vertices represent parts of an assembly and edges are the connection (liaison) between these parts. These graphs are widely used in assembly planning due to ease of extraction of data [15, 21]. Liaison graph/matrices demonstrate the liaisons between parts and maps the relationships between parts of an assembly. A graph of liaisons (GL) for an assembly has the following characteristics [15, 17] (Fig. 7.9):

(1) It is simple, meaning that two nodes are only connected with one edge.
(2) There are no loops.
(3) Is connected, all nodes have at least one connection.
(4) Is undirected, meaning if a part (n_1) can be assembled to another (n_2), the reverse is also true.

Information on a graph needs to be converted into matrix for computer processing. Using a table of liaisons (similar to an adjacency matrix), the relationships between the vertices in a finite graph can be clearly demonstrated (1 related, 0 no relation). The links and interferences in a matrix can be manipulated to allocate precedence constraints and sequences [22]. The matrix can be analysed using linear algebra and an assembly sequence can be derived from that. Moreover, a 3-dimensional CAD model of a product can be used to extract corresponding liaison matrices using computer aided programs based on physical and spatial features of assembly parts [23, 24]. Dini and Santochi [25] used three types of matrices including interference, connection, and contact matrices to supply information to an

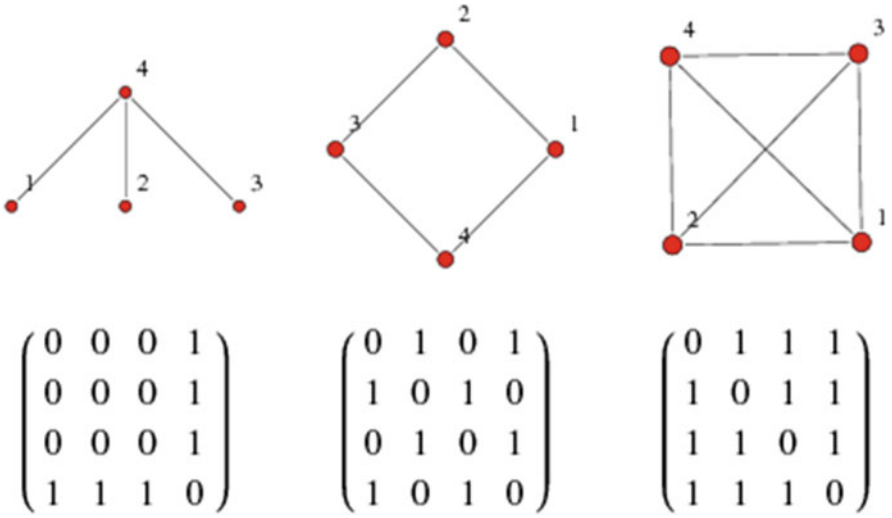


Fig. 7.9 Liaisons and adjacency matrix [21]

automated assembly sequence planner system which did not require any user input during ASP.

The grouping parts method aims to generate sequences from groups of parts rather than individual parts to minimise the exhaustive process of searching through all subassemblies. In this technique, parts are grouped into subassemblies based on specific characteristics. This reduces the search process as unwarranted links are omitted early in the process [21].

Step 3: Selection of Sequences

In order to optimise the assembly sequence, it is critical to choose the most suitable sequence. Optimisation techniques are used to select sequences which meet selection criteria. Several methods have been developed to solve ASP problems.

Subassembly Division

This heuristic method of generating feasible sequences is based on breaking the assembly down into functional subunits. The basis of this method is fastener analysis. Parts are divided into groups based on the method they are connected to each other (such as bolts, pressing fits, etc.). An engineer charts the parts, type of fasteners, and groups which they are divided to and finally manually determines precedence of the parts [26]. This method is very operator dependent and can be prone to error and underestimation.

Expert and Knowledge-Based Systems

Knowledge-based methods are also dependent on expert user involvement to provide extra information on assembly structure, resource, and precedence constraints to generate assembly models. These methods aim to generate enhanced and detailed versions of liaison graphs using this information [27, 28]. Computer-assisted software have been used with the aim to extract information from 3D CAD models with the aim to save time and ease of application.

Case-Based Reasoning

These systems use information of the previously solved similar ASP problems to solve new problems. The system uses a case library data base in order to solve the problem presented to it. The basis of this system is the notion that not all ASP problems need to be solved as solving assembly sequence planning problems is costly [29].

7.2.1.4 ASP Optimisation

After solving the ASP problem, optimisation takes place. This involves selecting the most optimal assembly sequence [21]. Multi-objective optimisation algorithms are highly useful techniques in ASP optimisation. Genetic algorithm is one of the most common methods used for optimisation. Other algorithms used to optimise ASP include simulated annealing (SA) [30], ant colony algorithm (ACA) [31], particle swarm optimisation (PSO) techniques [32], neural networks (NN) [33], immune algorithm (IA) [34, 35], evolutionary algorithm (EA) [35], and hybrid algorithms(HA) [36]. Input data provided to these algorithms can include geometrical, liaison, precedence, subassembly, tool, and interference data based on the type used.

Genetic Algorithm in ASP Optimisation

Genetic algorithm (GA) is the most common and powerful numerical optimisation method used in ASP. Bonneville [37] was one of the first researchers to use GA for ASP optimisation rather than the traditional cut-set method. A group of feasible assembly plans defined by an expert operator were encoded into chromosomes and used as the parent population. The authors used the classical GA approach with two operators of crossover and mutation to generate a set of feasible assembly plans and choose the most suitable one for a sample ten-part product [37].

Dini et al. [22] used classical genetic algorithm to solve the ASP problem with reduced computational time in a 7-part product. The sequences were generation based on fitness functions that took geometrical and precedence constraints into account including feasibility of the sequence, object orientations, gripper changes,

and presence of similar and consecutive assembly operations. The assembly operations were also divided into subassemblies to optimise assembly generation.

Hong and Cho used classical GA to minimise assembly energy in a 13-piece robotic automotive assembly line [38]. This method was heavily reliant on a series of assumptions that had been defined based on extensive knowledge on the assembly system. Hence, this method was heavily dependent on an expert operator. Concurrent assembly planning with genetic algorithm was proposed by Senin et al. [39]. Feasible assemblies were generated through implementing a disassembly approach. This approach is more complex when applied in assemblies with larger number of components. Lazzerini and Marcelloni applied the classical genetic algorithm to generate assembly sequences in a robotic pump assembly line with 11 parts and also a machine assembly line with 8 parts considering assembly directions/orientations and use of grippers and feasibility functions [40].

The ordering genetic algorithm (OGA) has also been used to generate the assembly sequences [41]. This algorithm has three basic concepts: trace, identify the stable subset, and comparison. Firstly, all generated assembly sequences are transformed into feasible ones through application of precedence constraints (trace). This guarantees that invalid sequences will be omitted. Secondly, stable subsets are identified, and lastly, the OGA results are compared against a decision aided method with multiple criteria. While this method is suitable for products with larger number of parts, it is also prone to generating unstable assembly subsets.

Chen and Liu used rule-based adaptive GA to solve the assembly problem (fitness function of sequence with minimum reorientations) in a 19-piece product assembly [42]. In this approach, the genetic operator probabilities are generated using the set of rules which aim to optimise the algorithm. Cut-paste was used in addition to the classical mutation and cross over operators.

Smith and Liu proposed multi-level GA (low level and high level) for optimisation of a 19-component product assembly [43]. The genetic operator probabilities that are used to generate sequences through low level GA (level 1) are synchronously updated by high level GA (level 2) which results in a considerable reduction in the search time in comparison with classical GA approaches.

Aiming to solve the assembly problem in a 9-part block-assembly and a 19-part practical controller assembly, Guan et al. proposed a gene-group-based evolution approach. The sequences were generated using several genetic operators and integrated into an interference matrix to determine geometrical feasibility. Fitness function to generate next generations was assembly cost computed by assembly direction and tool/gripper changes [44].

Using a variety of genetic parameters including cut-paste, reproduction, and break-and-join, an enhanced GA was proposed for automated generation of assembly sequences in a 10-part assembly system [45].

Del Valle et al. used AND/OR graphs and classical GA to generate all feasible assemblies based on the disassembly method [46].

Marian et al. used a bi-phasic approach to solve the ASP problem. The authors used GA to generate feasible sequences in the first phase and then optimisation took

place in phase two by choosing the most suitable ones. Stability and precedence constraints were considered in this method [47, 48].

Bai et al. proposed an effective approach to integrate sequence generation and evaluation [49]. As a semi-automated method, this technique requires initial human and computer interaction to develop a CAD model for the product and then proceeds to an automatic generation of sequences. This method considered geometrical and topological constraints as well as time and cost factors.

Huang et al. introduced the feature mark (FM) concept to solve the ASP problem using classical genetic algorithm. FM was used to characterise the assembly features and constraints. The authors assessed the assembly capability function by assessing the degree of similarity between FM for the assembly parts [50].

Hui et al. applied genetic algorithm and used disassembly feasibility information [51]. Feasibility functions in this study included assembly time and number of tool/gripper changes.

Choi et al. introduced a multi-criteria ASP method with the application of GA to solve the ASP problem to minimise assembly time and number of required part reorientations. A precedence matrix was used to demonstrate the precedence relationships and ensure feasibility of generated sequences [52].

Tseng et al. used genetic algorithm to optimise the assembly sequence in a multi-plant collaborative manufacturing system. The fitness functions that were applied included assembly constraints and cost objectives [53]. Pan et al. introduced assembly ring containing function (ARCF) to assess the fitness of the assembly. The authors used GA together with assembly graphs to define precedence [54].

Table 7.1 lists the summary of the reviewed previous researches and the feasibility function considered in each study.

7.2.2 Automotive Body Welding Optimisation

In the automotive manufacturing industry, thousands of welds are used to join numerous sheet metal parts. Welding is known to be "perhaps the most non-linear problem encountered in structural mechanics" [55]. Optimising the welding process significantly improves the quality of the final product [56]. Various methods have been used by researchers to plan and optimise various aspects of the welding processes including weld number, location, sequence, and quality.

Earlier studies mostly applied numerical and statistical methods to optimise the welding process. Hun and Kang studied optimisation of the quality of welding and durability of the electrodes in automotive body assembly through numerical analysis. The authors derived a 3D finite element code for electrothermal analysis and calculated the nugget shape and size based on that [57]. Li et al.applied a two-stage, sliding-level experiment to analyse the effect of various process conditions on spot weld quality. The abnormal process conditions which were considered included electrode and sheet metal misalignment, separation of the sheet metals (poor fit-up),

Table 7.1 Feasibility function of studies on application of GA in solving ASP problem

	Application					
Studies	Directional alterations	Tool changes	Assembly Time	Stability	Energy	Reliability
Bonneville et al. [37]	X					
Hong and Cho [38]	X			X		
Dini et al. [22]	X	X				
Lazzerini et al. [40]	X	X	X			X
Smith et al. [43]	X	X				
Chen and Liu [42]	X	X				
De Lit et al. [41]	X					
Smith and Smith [45]	X	X		X		
Guan et al. [44]	X	X		X		
Marian et al. [47, 48]	X	X				
Del Valle et at. [46]	X					
Bai et al. [49]			X			
Huang et al. [50]					X	
Hui et al. [51]		X	X			
Choi et al. [52]	X		X			
Tseng et al. [53]	X	X		X		

and electrode size. It was concluded that abnormal conditions significantly affect weld quality.

Use of high current and large electrodes was proposed as a measure to minimise weld quality variations [58]. Tang et al. studied the impact of application of intentional electrode force variations during the solidification phase on weld quality. Dynamic electrical resistance, generated heat, and welding strength were investigated through mathematical and experimental analysis [59]. It was recommended that the weld quality improves by using forging force. Lian et al. introduced a combined data mining and knowledge-based method to identify auto body dimensional variation patterns through a systematic diagnostic approach [60]. The method was used to assess the dimensional variation in the automotive rear end panel. It was concluded that the sever dimensional variation in the case study was due to insufficient fixing and optimisation was carried out by adding welding spots in the rear end panel and underbody. Le Meur et al. estimated thermal contact resistance and partition coefficient of generated heat flux at electrode-sheet interface during spot welding through experimental analysis [61]. They demonstrated that these two parameters are functions of current intensity and electrode force.

Welding sequence optimisation strongly impacts distortion and residual stress in the welded structure; therefore, it has been of interest to various researchers [62]. When exploring different welding sequences for optimal geometry of the automotive body, generation of an experimental sample using all possible sequences is not feasible due to cost, time, and complexity. Hence, common practice is to replace real experiments with computational simulation tools such as FEM. FEA is a strong, accurate, and detailed tool; however, generation of FE models requires extensive engineering expertise and are often very complex. As a result, it is not possible to create FE models for every single welding sequence. Conventional practice is that an FE model is made for a selected sequence based on engineer's experience. Evidently, this is operator dependent and does not necessarily guarantee optimal results. This highlights the importance of welding sequence optimisation (WSO) and application of tools to identify feasible welding sequences and choosing the best to achieve minimal distortion and geometrical variation.

Studies aiming to optimise welding sequences date back to the 1940s [63]. The use of artificial intelligence-based methods started in the 1990s. Fukuda et al. introduced a neural network theory-based approach as an effective method to optimise welding time and sequence. The considered optimisation parameters included welding torch moving distance, shrinkage, and welding symmetry. The study was carried out on a very simple symmetrical structure [64].

Park and Tsai studied the effect of welding sequence on panel distortion using FE analysis and experimental validation [65]. Distortion parameters studied included plate bending/buckling and global bending, residual stress, and shrinkage. Mochizuki et al. used a trial and error approach through generation of all possible sequences to evaluate the effect of welding sequence on residual stresses.

Evolutionary algorithms including GA have been widely used to solve combinatorial problems including the travelling salesman problem (TSP). TSP problem addresses the question "Given a list of cities and the distances between each pair of

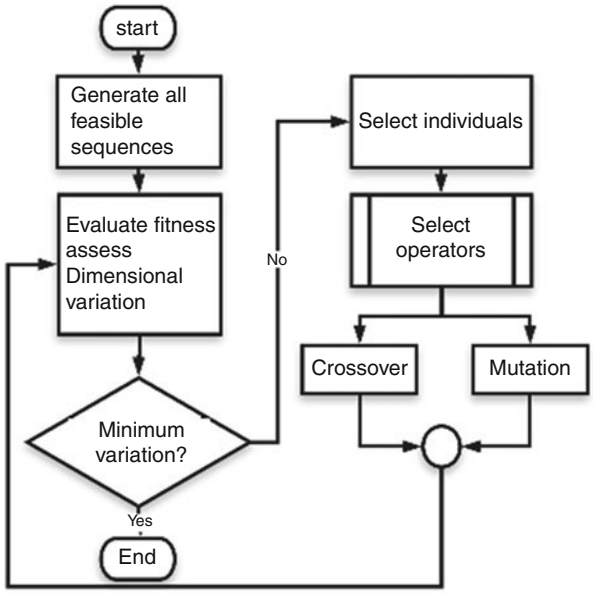


Fig. 7.10 Application of GA in WSO for geometrical improvement [68]

cities, what is the shortest possible route that visits each city and returns to the origin city?" [66]. As discussed previously, ASP problems are combinatorial problems [11] with WSO largely modelled as a TSP [67].

GA has been extensively used for spot welding sequence optimisation aiming at geometrical variation minimisation and has been shown to be an effective method. Details on how GA operates have been discussed extensively in the next chapter. When solving the WSO problem for geometrical variation minimisation, the algorithm typically follows the pattern demonstrated in Fig. 7.10.

Genetic algorithms have been shown to have limitations when used to optimise sequence problems [68]. A prominent drawback which has been identified and largely addressed is "feasibility of offspring" when applying the conventional crossover and mutation operations. Offspring may be not feasible after crossover/mutation is applied to sequencing problems as there is a possibility that repetitive numbers are suggested to permutations (Fig. 7.11). Multiple modifications to the GA have been proposed to overcome this issue. Random key encoding and genetic algorithm for standard cell placement (GASP) [68] are the most widely used modification methods [69].

Huang et al. introduced GASP to address the WSO problem in multiple sample cases including metal plate with 10 bolts, 16 bolts, and vehicle A-pillar subassembly with 14 welds [68]. GASP is a modified GA that uses three types of permutations and one relocation as opposed to the conventional crossover and mutation operations. The fitness function applied in this study was movement distance of

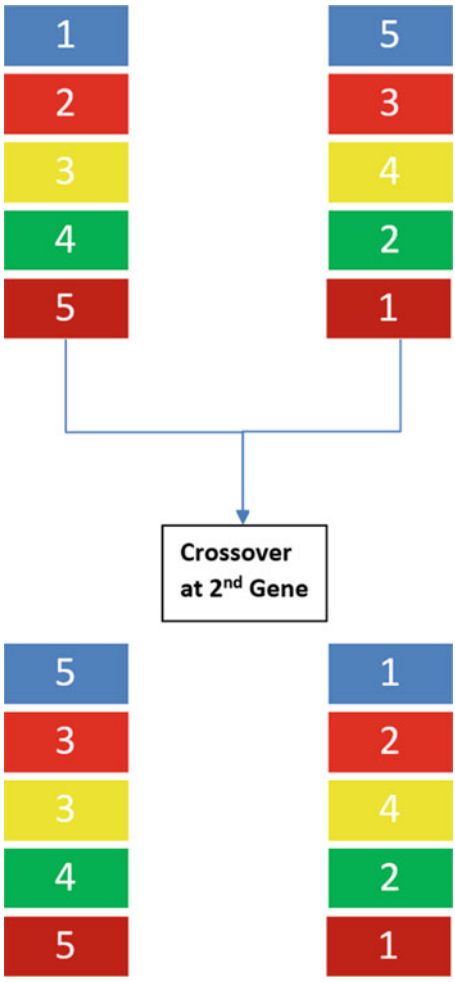


Fig. 7.11 Offspring feasibility problem

the robot arm for the metal plates and degree of deformation of predefined points in the A-pillar assembly case. Finite element models of the A-pillar subassembly were generated and the degree of deformation in all points was added up to assess the fitness function in four different welding sequences (one GASP generated, and three conventional welding sequences defined by author). The method was shown to be efficient for sequence optimisation. The authors did not validate the findings from the FEA. The studied structured were simple and employed a linear weld distribution which is not the case with welding operations in structurally complex assemblies such as bodyshell and BIW.

Damsbo et al. used hybrid genetic algorithm to optimise welding sequence [70]. The authors used welding time and trajectory distance as fitness function while applying precedence constraints as absolute constraints. Rather than using the

conventional binary crossover operation, the authors used the genetic edge recombination method [71] and also modified mutation method applying to mutation subtypes (minor and major). This method was found to be a simple and effective optimisation tool. No experimental validation was carried out in this study.

Kadivar et al. applied GA to optimise welding sequences [62]. Objective function in this study was distortion (radial displacement) and residual stress which was predicted through linking the GA to a thermomechanical FE model. The authors used a transformation method to address the offspring feasibility issue. The results of the GA analysis were validated using experimental data from previous published work.

Xie and Hsieh aimed to minimise deformation together with assembly time through optimisation of welding sequence of automotive body using GA. A multi-objective fitness function was defined through a computer-based program referred to as elastic assembly variation simulation (EAVS), which was used to assess fitness [72].

Kim et al. [73] used a heuristic algorithms to solve the WSO problem. They used "the nearest neighbour" concept to optimise welding sequence with the aim to optimise the movement pattern and time of the robot and minimise distortion. The nearest neighbour concept is an algorithm used to address the TSP which is based on choosing the nearest point to the start point. While this algorithm generates a rapid result, it does not necessarily guarantee the most optimal one. The authors also used the "tabu search" algorithm which implements memory structures based on already-visited options and/or user-defined parameters.

Liao used GA-based algorithm to define optimal welding pattern (positions and number of welds) for dash and front-end panels. Fitness function was calculated through both deterministic (dependent on previous states) and stochastic modes (independent of previous states). Fitness function was calculated based on the degree of deformation [74]. FEA was used to evaluate the fitness of the chosen design. The studied structures were simple and employed a linear weld distribution. The analytic results were not validated.

Chapple et al. [75] studied geometry welds and optimal welding sequence in MINI front subframe tower assembly. Applying a disassembly approach, the authors used HyperStudy, an automated design optimisation software to generate all possible variations of the fully welded model with welds being activated/deactivated. Two optimisation algorithms including adaptive response surface method (ARSM) and GA were used with the fitness function of minimal distortion [75]. A simplified FEA model was used to evaluate the fitness function. The used GA model is not defined in detail in the study, but it was mentioned that the issue of offspring feasibility and repeated sequences was not addressed. As a result, the GA runs were excessive and were not completed. Like other studies, this study also picked a simple automotive part with linear distribution of welds. Results of this study were not validated by experimental analysis.

Sadeghi Tabar et al. evaluated three different evolutionary algorithms to solve the WSO problem for geometrical optimisation. The authors applied random key encoding to overcome the limitations of GA in sequence optimisation. FEA was

used to assess assembly deformations and degrees of variations to assess fitness [76]. It was recommended that other evolutionary algorithms besides GA can also be used to solve the WSO problem. The assessed algorithms in this study were the ant colony optimisation (ACO) and particle swarm optimisation (PSO) algorithms. Experimental validation was not part of this study.

Surrogate models have also been used to solve the WSO problem. These are mathematical data-driven models that mimic more complex models but are simpler and less expensive. Voutchkov et al. developed a "black-box" surrogate algorithm to optimise welding sequence aiming at minimising displacement [77]. The algorithm proposed was developed using weld position and direction data. The "priority list" concept was developed to train the surrogate model. The principle of the priority list concept was that displacement caused by each welding event is dependent on previous welding events. The authors calculated total displacement as the sum of displacements from every individual weld. The cooling stage was ignored in this study. Experimental validation was not done in this study, but results were compared with previous experimental data published by the author group [78].

Asadi et al. also used a surrogate model to optimise welding sequence to achieve minimum distortion in a pipe. They used a modified version of the priority list concept by applying physics of the weld data to train their surrogate model [79, 80]. Bonnaud used the surrogate model developed by Voutchkov et al. together with FEA to minimise residual deformation by finding the optimal welding sequence [81]. Neither of the mentioned studies included experimental validation.

Carlson et al. developed an algorithm based on a systematic search approach to optimise geometrical variations and welding time on A-pillar and part of the front floor consisting of seven welding points [82]. In this study, measurements were done virtually, and an experimental model was not made.

The geometry welds are the welds that fix the geometry and control the dimension of the assembly. The rest of the welds are referred to as re-spot welds with the sole purpose of strengthening the vehicle's body [83]. Common practice in assembly of the BIW is that geometry welds are applied first and then the re-spot ones are welded. Arguably, WSO should be directed exclusively at geometry welds for dimensional optimisation which highlights the importance of identification of geometry welds.

Review of the literature demonstrates that GA together with FEA for assessment of the results is the most widely used and well-validated method in WSO (Table 7.2).

Most of the mentioned studies directed at ASP were carried out based on DfA criteria rather than quality criteria. Dimensional and geometrical quality-based studies have been limited and do not include WSO [84–86]. Only two studies have addressed identification of geometry welds, and both have been limited to simple structures consisting of one row of welds. Most of the studies reviewed have limited or no experimental validation of the results obtained through computational analysis. Also, no studies have addressed real-life industrial-level manufacturing problems. Moreover, no previous study has addressed WSO in the bodyshell as this is the most complex structure with hundreds of welds and multiple welding

Table 7.2 WSO studies and implemented fitness functions

Studies	Method	Torch moving distance	Welding time	Symmetry	Shrinkage	Distortion	Displacement	Residual stress
Fukuda et al. [64]	NN	x	x	x	X			
Huang et al. [68]	GA/FEA	x					x	
Park and Tsai [65]	FE/Exp				X	x		x
Damsbo et al. [70]	GA	x	x					
Mochizuki et al.	Exp					x		x
Kadivar et al. [62]	GA/FEA					x		x
Xie and Hsieh [72]	GA/FEA	x	x	x				
Kim et al. [73]	Nearest neighbour	x	x			x		
Liao [74]	GA/FEA					x		
Voutchkov et al. [77]	Surrogate/FEA						x	
Chapple et al. [75]	GA/FEA					x		
Carlson et al. [82]	SS	x	x				x	
Asadi et al. [79, 80]	Surrogate/ FEA					x		
Bonnaud [81]	GA/FEA			x			x	
Sadeghi et al. [76]	GA/PSO/ACO/FEA					x	x	

stations. The gaps in the present literature demonstrate the necessity of the current study to address the WSO problem in a real-life automotive manufacturing setting for a complex structure, i.e., the bodyshell.

7.3 Optimisation of the Body Processes Using Genetic Algorithm

7.3.1 Genetic Algorithm and Its Applications

The genetic algorithm (GA) is an effective and strong optimisation method that can handle nonlinear design spaces with the aim to maximise or minimise a predetermined function. GA which is based on the theory of evolution and has a search-based algorithm aims to find a global optimum using a set of "predetermined criteria" or an "overall evaluation criterion" (OEC). GA originates from the concept of evolution in reproductive biology, and as a result, most of the terminology used in GA are derived from biology (Fig. 7.12). While like biology, GA is "random" in selection, the operator has the option to set the randomisation degree and the level of control. GA operates using five distinct components: fitness function, set of chromosomes, selection, crossover, and mutation [87].

The fitness function is the function which the genetic algorithm aims to optimise. This is the most important part of the algorithm as it is the only component which the operator can define how the changes to the chromosomes will take place, what conditions and constraints are put in place, and eventually this will determine

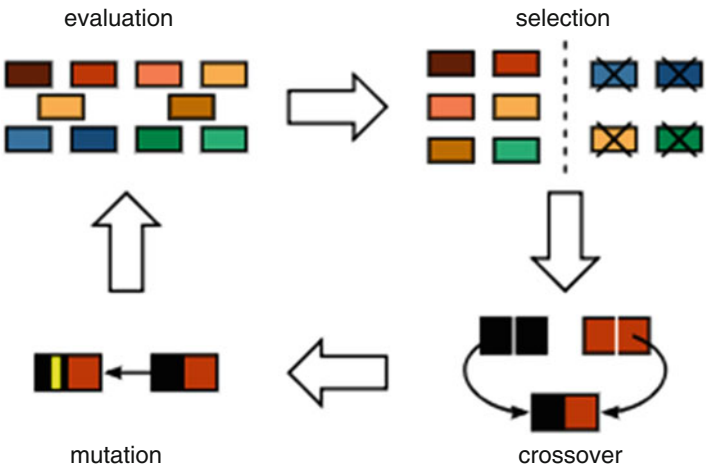


Fig. 7.12 Genetic Algorithm [87]

whether a fit solution is found or not. It is essential that special consideration and time are put in defining this function to improve the algorithm's performance.

After defining the fitness function, the algorithm encodes a subset of solutions into a numerical data structure named a "chromosome" to randomly apply genetic alteration concepts. Classically, only discrete values can be used in GA as it applies a binary string system (0,1). Hence, continuous variables must be discretised and converted into binary strings using tolerance limits that specify the number of divisions the variable will have. To represent the whole design picture, the individual binary strings are pieced end to end to form a long string (chromosome). The number of design combinations (2^n) is based on the length of an individual string (n). Typically, 50–100 combinations are generated by randomly filling the string locations by 0 or 1. The generated strings are then changed into decimal numbers and equivalents and selection follows [88–90].

The algorithm starts with defining the first generation which consists of an assortment of randomly selected chromosomes. The efficacy of each of these chromosomes in solving the problem is assessed based on the defined fitness function.

Selection refers to the process of choosing the fittest subset for breeding next generations based on an operator pre-defined probability distribution. Selection of "parent" designs can be done in various ways including the fitness proportion selection methods such as the roulette wheel method [91] and stochastic universal sampling (SUS), the tournament selection method, rank selection, and random selection[150]. Figure 7.13 illustrates the basics and process flow of solving a genetic algorithm problem.

Fitness Proportion Selection Methods

These methods are the most used parent selection strategies. Parent selection is directly proportional to every individual's fitness. In this selection strategy, fitter individual's likelihood of mating and passing on of genetic material is higher. While these methods are very useful and widely used, they cannot be used when the fitness value is negative. The roulette wheel method [91] and stochastic universal sampling (SUS) are samples of fitness proportion selection methods.

Roulette Wheel Selection

The roulette wheel selects individuals (P^k) randomly based on a probability distribution proportional to their level of fitness by the fitness function (Eq. 7.2).

$$\mathrm{S}\left(P^k\right) = \bigcup_{i=1}^{Npop} s\left(P^k\right) \tag{7.2}$$

where $s\left(P^k\right) = p_j^k$ with probability $\Pr\left[s\left(P^k\right) = p_j^k\right] = \frac{f_j^k}{\sum_{n=1}^{Npop} f_n^{k.}}$

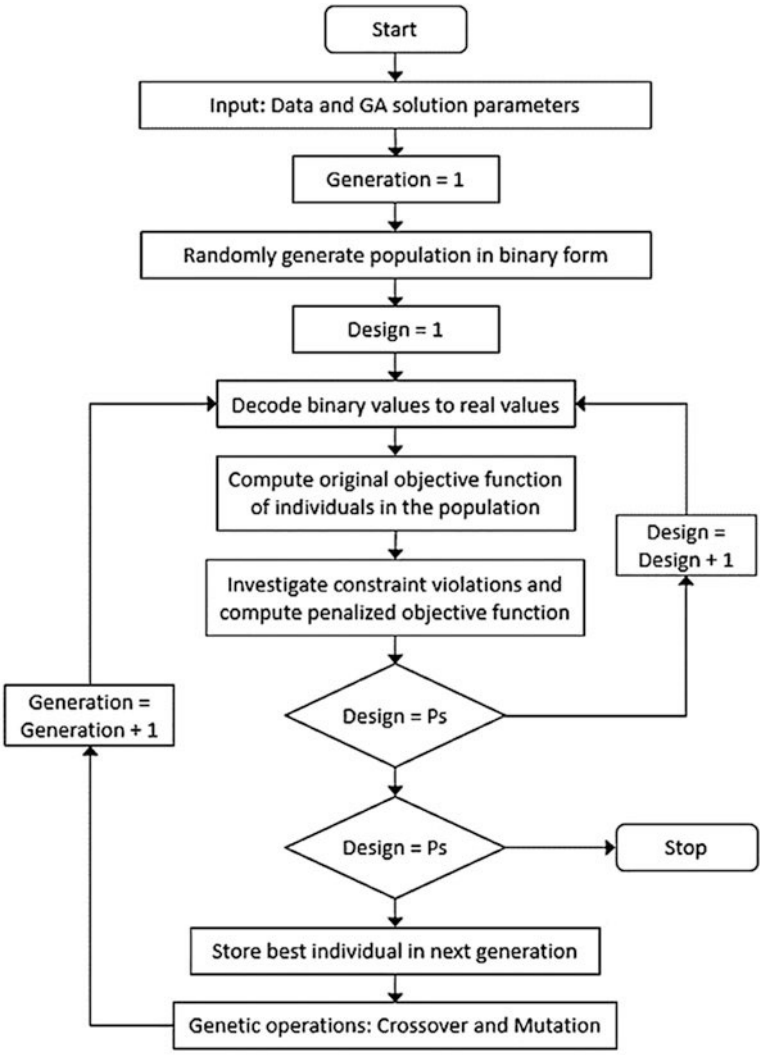


Fig. 7.13 Process flow of solving a GA problem [68]

In other words, the fitter the individual the higher chance of landing in front of the fixed point after the wheel is spun (Fig. 7.14). Multiple spins are required for parents to be selected as there is only one fixed point.

Stochastic Universal Sampling (SUS)
This is very similar to the roulette wheel method in concept. The major difference is that in this method, we have multiple fixed points rather than just one. This allows the selections of multiple parents in one spin and ensures that the individuals with higher fitness values are at least chosen once (Fig. 7.15).

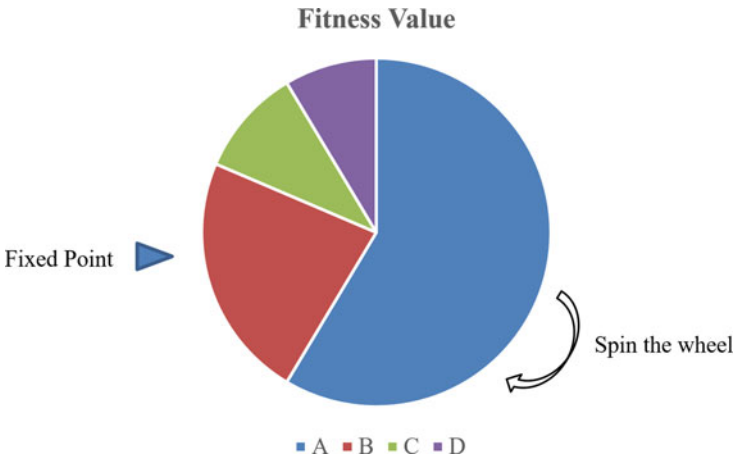


Fig. 7.14 Roulette wheel selection

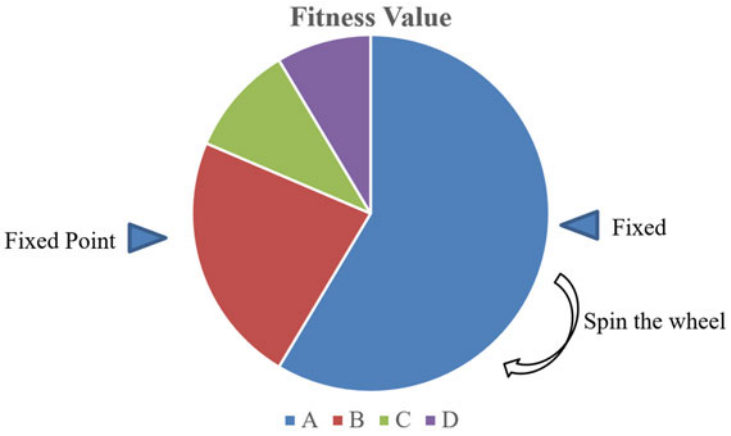


Fig. 7.15 Stochastic Universal sampling

Tournament Selection

In the tournament selection method, a tournament takes place. The designs are randomly compared and the fitter wins. All individuals in the population compete through multiple tournaments and the winners create the parent population [88, 92].

Rank Selection

Rank selection is usually used when the differences in fitness values for various individuals are very small. This results in a rather equal division of the pie and similar probability of selection for each individual (Fig. 7.16).

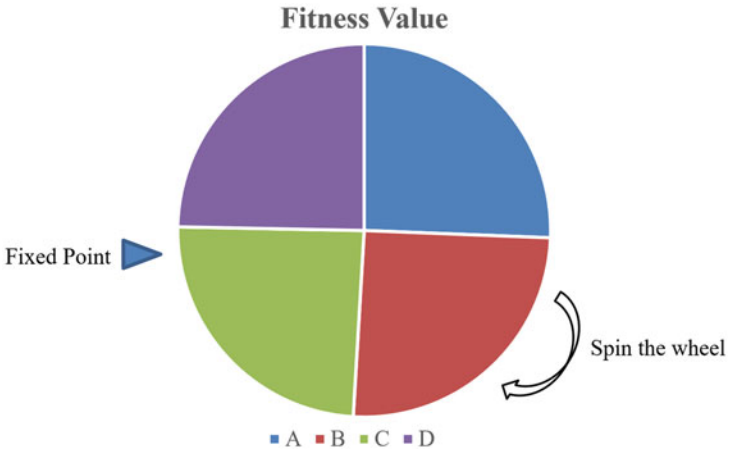


Fig. 7.16 A population with little difference in fitness values

Table 7.3 Rank selection

Chromosome	Rank	Fitness
A	1	8.2
B	2	8.1
D	3	7.9
C	4	7.8

For the fittest individual to be selected, rank selection aims to rank every individual based on their fitness value. Selection will then take place based on rank rather than the absolute fitness value (Table 7.3).

Random Selection

As evident by its name, parents are randomly chosen in this method. As fitness is not a selection factor in this method, the random selection strategy is usually avoided.

After the selection of parents takes places, they are then mated. If a specific fitting probability is met and the bit is chosen, then crossover takes place. Selection of the bit, its location, and size are all done randomly. These chosen subsets are then "mated" to generate a superior combination.

Crossover

The crossover operator is the operation used to combine pairs of individuals previously selected for breeding. Various crossover operators can be used including one point, two-point and uniform. In one-point crossover, the tails of the parents are swapped at a random point and new off-springs are generated (Fig. 7.17) [146].

Two-point or multi-point crossover is an expanded type of the one point where swapped segments alternate (Fig. 7.18). In the uniform method, the chromosome is not divided into segments and each gene is treated separately.

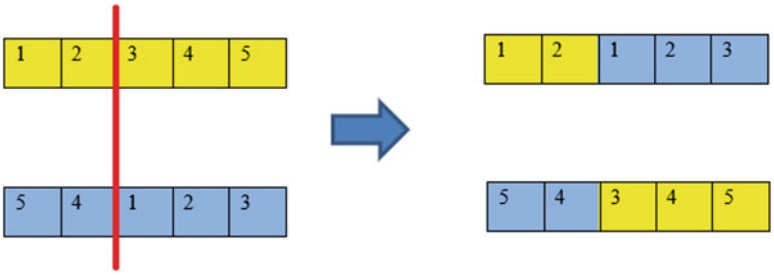


Fig. 7.17 One-point crossover

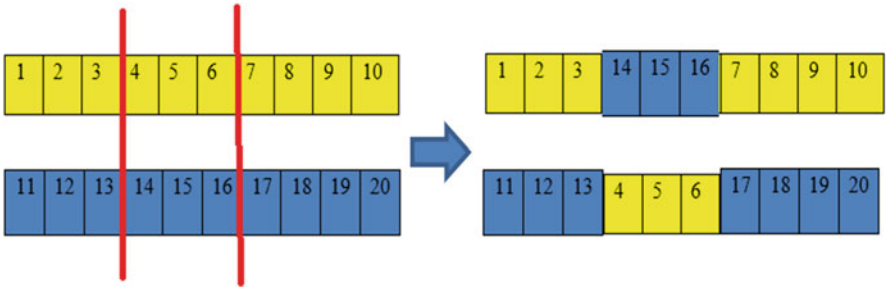


Fig. 7.18 Two-point crossover

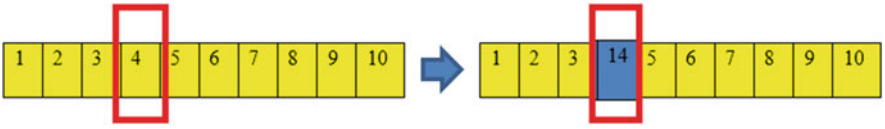


Fig. 7.19 Bit flip mutation

Mutation

Mutation is a function which potentially modifies individual "bits" on the chromosome (swapping a "0" into a "1" and vice versa) with the aim to generate and maintain genetic diversity. Mutation takes place randomly based on the mutation probability which is usually defined as very low such as 0.01.

Mutation is a critical part of the process which explores the search space and is essential in GA convergence [146]. Some of the most used mutation operators include: Bit flip mutation, random resetting, swap mutation, scramble mutation, and inversion mutation.

Bit Flip Mutation

In this mutation, random bits are chosen and flipped as the example illustrated in Fig. 7.19.

Random Resetting

A random permissible value is assigned to a gene which is also randomly selected.

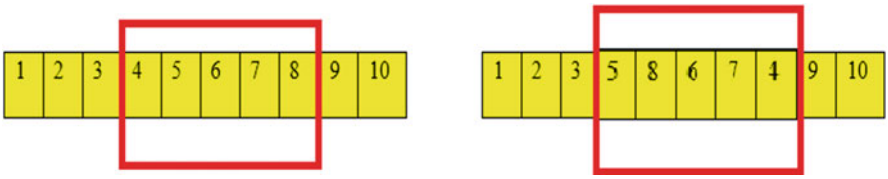


Fig. 7.20 Scramble mutation

Swap Mutation
Two genes are randomly chosen and swapped on a chromosome.

Scramble Mutation
A group of genes on a chromosome are chosen and the values are randomly shuffled (Fig. 7.20).

Inversion Mutation
In this method, the entire string of genes is inverted.

The process of selection, crossover, and mutation is usually continued until the initial population is completely replaced by the second generation and the number of offspring equals the initial population size. Now the cycle repeats in the second generation. Through this ongoing process, the fittest chromosome of each generation (best-so-far) is identified and recorded. The cycle repeats until the maximal fitness value of the best-so-far chromosome is achieved and no further improvement is seen in several successive generations. At this stage, the algorithm has converged to a solution. The process of iteration of the cycles until the algorithm has converged to a solution is referred to as a "run". The aim of each GA run is to generate at least one highly feasible chromosome. This could be the fittest of the best-so-far chromosomes or the fittest chromosome generated in the last generation based on the way the algorithm is programmed.

7.3.2 GA Constraint Handling Methods

Structural optimisation studies are considered constrained problems. In constrained optimisation problems, a function f(x) is to be minimised or maximised subject to variables in the presence of constraints of those variables [93]. This means that some results in the solution space are not feasible. The areas in the solution space that contain feasible results are referred to as "feasible regions" (Fig. 7.21). The usual crossover and mutation operators cannot guarantee a feasible solution in such a scenario.

GA performs best in unconstrained problems, and hence, various constraint handling methods have been introduced in the literature [94–103]. One of the most used techniques applied to address constrained GA problems is the penalising

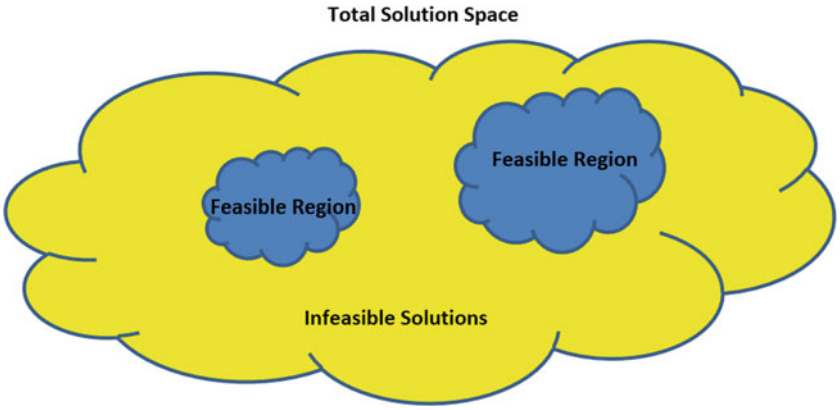


Fig. 7.21 Constrained optimisation problem

strategy. Through this strategy, the constrained problem is turned into a series of unconstrained problems. To achieve this, a "penalty function" is added to the "objective function". The penalty function comprises of a "penalty parameter" multiplied by a degree of violation of the constraints (Eq. 7.3).

$$\begin{gathered} \min f(x) \\ \text{Subject to } c_i\ (x) \leq 0\ \forall_i \in I \\ \min \Phi_k(x) = f\ (x) + \sigma_k\ \ \textstyle\sum_{i \in I} g\ (c_i(x)) \end{gathered} \tag{7.3}$$

Where

$$g\ (c_i(x)) = \max\ (0, c_i(x))\ 2$$

The method by which the penalty function is designed is critical and arguably the most important factor affecting the results [94]. Various strategies for designing penalty functions have been proposed and studied in the literature. An efficient penalty function design is dependent on the problem being addressed, and hence a general design guideline cannot be applicable.

A comprehensive categorisation of penalty functions for constrained problem was proposed by Yeniay in 2005 [97]. Penalty functions were divided in to four groups: methods based on penalty functions, based on search of feasible solutions, based on preserving feasibility of solutions, and hybrid methods [97]. A summary is listed below:

- *Methods based on penalty functions:*

  a. Dynamic penalties
  b. Segregated GA
  c. Death penalty

d. Co-evolutionary penalties
e. Adaptive penalties
f. Static penalties
g. Annealing penalties

- *Methods based on a search of feasible solutions:*

a. Behavioural memory
b. Superiority of feasible points
c. Repairing infeasible individuals

- *Methods based on preserving feasibility of solutions:*

a. Homomorphous mapping
b. Searching the boundary of feasible region
c. GENOCOP system

- *Hybrid methods*

7.3.3 Body Process Modelling for Genetic Algorithm Analysis

To develop the GA model of body process, let us further review the body process elements in detail. As explained in Chaps. 1 and 2, in a body production line, the assemblies are located in a fixture via pins, locators (also called rest points), and clamps. Each of these locating elements has a key coordination in the model with respect to the vehicle coordination system. For a pin, this coordination is the centre position of the pin, which is usually in two directions, as the third direction would be free. In case of a locator, this coordinate is the midpoint of the locator (rest) surface (Fig. 7.22).

For spot welds, this critical coordinate which is also identified in the body CAD data is the position of the centre of the spot weld. Usually, to ease the working with 3D model CAD, we use spheres as representative of spot welds. The centre of the sphere is centre of the spot weld as per the product design (Fig. 7.23).

Parts and assemblies must have some overlaps (loaded on top of each other) to create the weld flanges. So, it is normal to have common pins or locators between some parts (Fig. 7.24).

As we cannot input the CAD models into our genetic algorithm models, we have to use characters, numbers, and coordinates to develop a database that can later be coded for GA model of the body process. To do that, we have extracted three key components of the CAD model into tables as explained below:

1- *Parts/assemblies table:*

A unique number is assigned to each part/assembly to enable referencing it in the model. These numbers will be used throughout the GA optimisation. The outcome of this step would generate a list of unique part numbers, so called as parts table (Fig. 7.25).

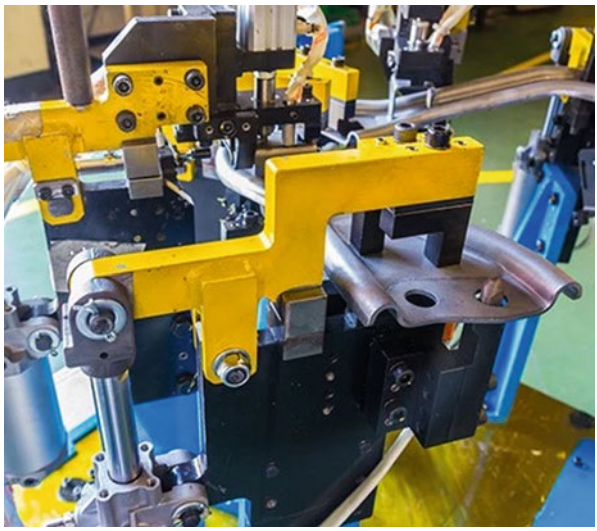


Fig. 7.22 Welding fixtures locating the parts using pins and locators

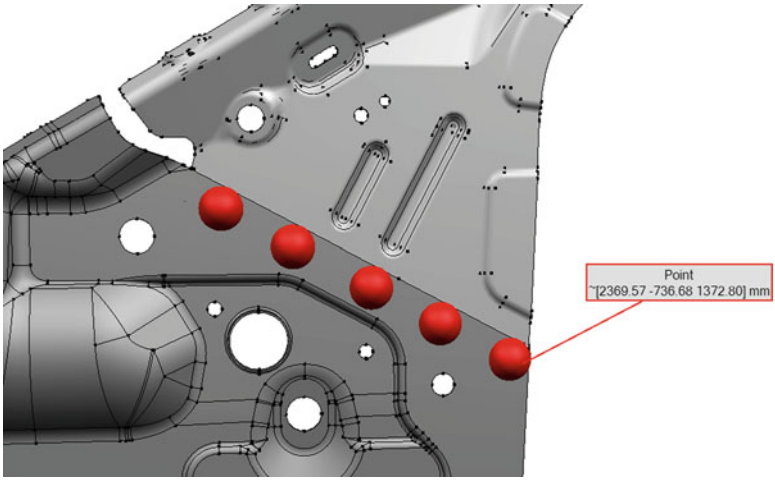


Fig. 7.23 Spot weld coordinate in a car body

2- *Locating elements table:*

Using a 3D software (in this study we used TeamCenter VisMockup application), we extracted the coordinates of all locating elements including pins and locators. By this, we developed the locating table which includes the CAD ID for each locating element, the coordination of the centre point of it, and a code assigned to each (Table 7.4).

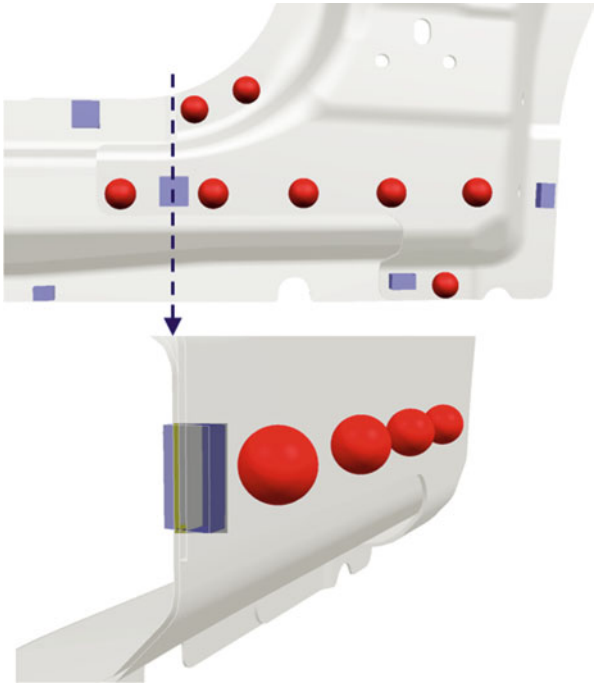
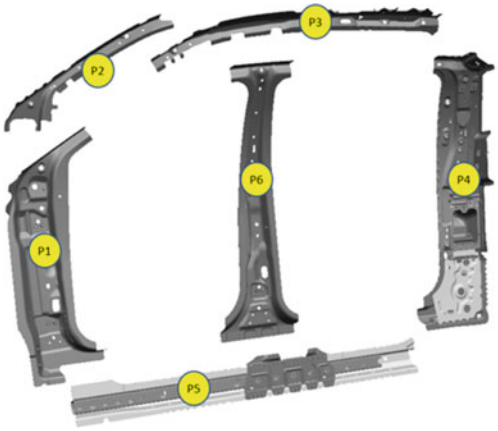


Fig. 7.24 Parts with flange overlap and common locators



Assembly	Code
Front Hinge Pillar	P1
A-Pillar	P2
Roof rail	P3
Rear Pillar	P4
Rocker assembly	P5
B-Pillar	P6

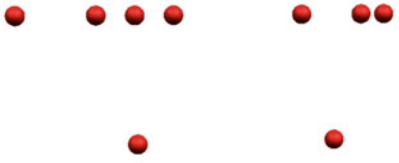
Fig. 7.25 Sample parts/assemblies coding

3- *Welds:*

Using a 3D software (in this study we used Teamcenter VisMockup application), we extracted the coordinates of all welds and listed them in the weld table which included all weld numbers and the 3D coordinates (Fig. 7.26).

Table 7.4 Sample locator table

Locator ID in CAD	Coordinate			Code
	X	Y	Z	
1-26513D20-A-01	2710.00	645.80	1766.00	L1
10-2102556-A-01	2415.00	705.97	1558.00	L2
11-2624320-A-01	3433.00	751.50	740.00	L3
12-2610130-A-01	2270.00	748.55	564.00	L4
12-2624320-A-01	3110.00	751.50	737.00	L5
13-2624320-A-01	3213.00	605.71	1858.00	L6
14-2624320-A-01	3466.00	600.81	1867.00	L7
15-2102556-A-01	2242.00	713.00	1523.37	L8
19-2102556-A-01	2453.00	819.40	638.00	L9
2-2102524-A-01	2710.00	659.42	1766.00	L10



Locator ID in CAD	Coordinate			Code
	X	Y	Z	
WPWDKB3B76000_01_R_00	2,243.00	817.60	633.00	W1
WPWDKB3B76000_02_R_00	2,303.00	817.60	633.00	W2
WPWDKB3B76000_03_R_00	2,363.00	817.60	633.00	W3
WPWDKB3B76000_04_R_00	2,423.00	817.60	633.00	W4
WPWDKB3B76000_05_R_00	2,483.00	817.60	633.00	W5
WPWDKB3B76006_01_R_00	3,389.00	817.60	633.00	W6
WPWDKB3B76006_02_R_00	3,479.00	817.60	633.00	W7
WPWDKB3B76006_03_R_00	3,297.00	817.60	633.00	W8
WPWDKB3B76006_04_R_00	3,205.00	817.60	633.00	W9

Fig. 7.26 Sample weld table extracted from CAD data

By completing the above three steps, we have developed the initial database required for coding into a GA model. Now let us see the platform we can use to develop our GA model.

7.3.3.1 Genetic Algorithm Platforms

There are many platforms and spaces such as Python and MATLAB we could use to develop our coding for genetic algorithm model. There are also several GA optimisation packages available to solve the optimisation problems. The advantage of using a package rather than coding is benefiting from the pre-written codes that enable us to focus more on the actual problem.

One of the widely used packages is modeFRONTIER developed by ESTECO (Fig. 7.27) which is a comprehensive solution for process optimisation in the engineering design process. There was no real preference in which package to use, and the main reason we used modeFRONTIER for our research was due to the package availability at RMIT university software store.

One of the challenges in this study was that the relationship tables between parts, welds, and locators could not be directly fed into modeFRONTIER due to its structure. This was reviewed and discussed with ESTECO application engineers, and it was concluded that coding is needed to convert the tables into the modeFRONTIER acceptable format.

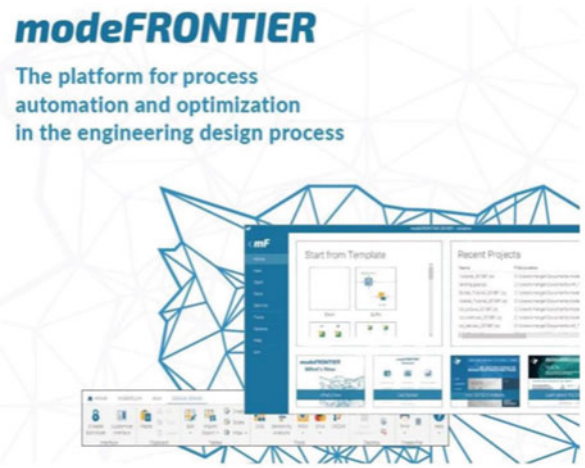


Fig. 7.27 modeFRONTIER optimisation package

7.3.3.2 GA Model Development

GA model development of the body framing process including approximately 850 welds, 12 major assemblies, and 94 locators could be quite challenging. To simplify the GA model development process, we started with a simple assembly (not the real product) to initially develop and test the process. In the next step, one of the main assemblies of a Ford product was developed, and finally after the process and algorithm were tested, the actual model for a body framing process of a real size Ford product was developed. Below, we review these three steps in more details.

Simple Assembly Model Development

This simple assembly includes 5 parts, 9 locators, and 14 spot welds (Fig. 7.28). The purpose of developing this sample model is to test the method and to improve the accuracy of the process.

As you can see in Fig. 7.28, parts A1 to A5 are located using six locators (represented by yellow cylinders in the figure). The product design includes 14 welds (represented by black circles in the figure) that connect the five parts together to make the assembly. The aim is to utilise GA to find the minimum number of welds that can dimensionally set the assembly.

To develop the model, we first need to create the parts, locators, and weld data tables, as explained in Sect. 7.3.3. Using the CAD data, these tables are extracted as below (Table 7.5).

The next step is to define the relationship tables. Each weld is connecting two or more parts together. So, we must define a "weld-part" relationship table. Looking at the assembly (Fig. 7.28), we could extract the below weld-part Table 7.6.

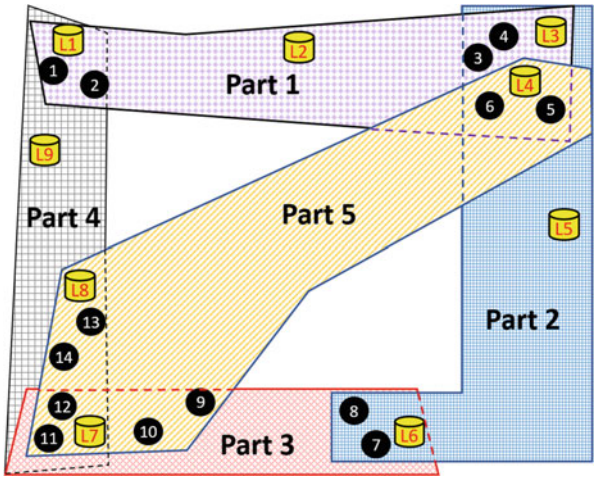


Fig. 7.28 Sample assembly developed to test the process

Also, there is a relationship between locators and panels, meaning which locator is coordinating which panels. As mentioned before, some locators are in the parts overlapping areas so they can be in relation to more than one part. Table 7.7 summarises the locator-part relationship for this assembly.

The next step is to develop a combined relationship table that will help analyse the problem and develop the input needed for modeFRONTIER application. In the combined table, we have the welds as the key elements and all other components of the system including parts and locators listed in relation to each weld. Assessing the weld-parts and locators-parts tables and using VBA coding, we extracted the combined relationship table for this assembly (Table 7.8).

Through VBA coding, we identified all available weld stack-ups. The results were cross-checked with the assembly to ensure the coding is correctly extracting the weld stack-ups. This may seem an easy step just by assessing the 3D model for such a simple assembly, though for a complicated structure such as bodyshell, this can take hours and may contain errors as a result of manual checks. So, we coded every step where feasible. The weld stack-up table for this assembly is summarised in Table 7.9. This table shows how many different stack-up of parts exist in this assembly, regardless of the number of welds for each joint.

To help understand this step, Fig. 7.29 is an illustration of Table 7.9, showing all different stack-ups in the structure, each including some welds.

The next step is to identify all welds in each joint. We have coded extracting the welds included in each joint from the previously developed tables. For example, joint 1 includes a stack-up of parts 1 and 4. Table 7.10 lists the joints and all welds included in each. Note again that in this simple structure, it will be possible to manually extract the welds for each joint from 3D model, though in complicated structure, it will not be an easy task to do so. Hence, we coded this step via VBA.

As a key factor in working with dimensions, we now need to measure the distances between different elements of the structure. This includes the distance

Table 7.5 Parts, locators, and weld data table of the sample structure

Part	Code
Part 1	P1
Part 2	P2
Part 3	P3
Part 4	P4
Part 5	P5

	Coordinate		
Locator	X	Y	Z
L1	3162.86	310.56	102.13
L2	1911.24	302.98	98.33
L3	205.37	321.45	90.28
L4	253.44	285.16	99.50
L5	181.05	23.86	100.01
L6	1620.73	−305.71	103.52
L7	2980.12	−301.12	97.69
L8	3050.72	−93.41	98.11
L9	3230.57	140.32	101.49

	Coordinate		
Weld	X	Y	Z
W1	3170.21	298.35	101.53
W2	3056.48	251.84	103.12
W3	590.12	231.13	98.36
W4	385.23	318.73	99.12
W5	211.37	194.32	98.63
W6	438.65	210.87	99.41
W7	1710.34	−315.23	102.54
W8	1873.22	−291.56	102.31
W9	2489.43	−283.47	97.33
W10	2701.53	−304.55	98.12
W11	3201.71	−315.63	99.63
W12	3159.65	−279.41	101.23
W13	3055.31	−118.94	99.56
W14	3138.24	−203.76	98.41

between each locator to all relevant welds, as well as the maximum distance between welds in each joint. To extract the first list, we have to list all locators and find all welds related (associated) to that locator. For example, as you can see in Fig. 7.28, locator "L1" is locating part 1 (P1) and part 4 (P4). So, all welds on these two parts are assumed related to L1. Using the above tables, we coded this step to list all relevant welds to each locator (Table 7.11).

Then, simply using the coordinates of locators and welds, we calculated the distance between each locator to any of its related welds (Table 7.12).

Table 7.6 Weld-parts relationship table

Weld	Connecting parts		
W1	P1	P4	–
W2	P1	P4	–
W3	P1	P2	–
W4	P1	P2	–
W5	P1	P2	P5
W6	P1	P2	P5
W7	P2	P3	–
W8	P2	P3	–
W9	P3	P5	–
W10	P3	P5	–
W11	P3	P4	P5
W12	P3	P4	P5
W13	P4	P5	–
W14	P4	P5	

Table 7.7 Locator-parts relationship table

Locator	Located parts		
L1	P1	P4	–
L2	P1	–	–
L3	P1	P2	–
L4	P1	P2	P5
L5	P2	–	–
L6	P2	P3	–
L7	P3	P4	P5
L8	P4	P5	–
L9	P4	–	–

We now assess all weld distances in each joint and list the two welds with maximum distance in that joint. This is coded and the results are shown in Table 7.13.

The above extracted tables would characterise each component in the structure and dimension the relation between the components. These data tables can now be used as input to codes for GA model. The GA model was still missing a critical element which was the fitness function. After multiple tries on defining a fitness function, we defined it as the root of sum square of the distance of each weld with respect to the locator with the same stack-up.

After coding the tables to generating a readable format of all the above tables and the fitness function, we developed the GA model in modeFRONTIER package. A schematic of the structure and the GA model is shown in Fig. 7.30.

The model can continue running until the fittest subset of welds are generated. In this simple structure, the fittest subset was achieved as (W1, W5, W7, W12), meaning these welds are the minimum welds required to set the structure, known as geometry welds (Fig. 7.31).

Table 7.8 Combined relationship table

Weld	Connecting parts			Related locators							
W1	P1	P4	–	L1	L2	L3	L4	L7	L8	L9	–
W2	P1	P4	–	L1	L2	L3	L4	L7	L8	L9	–
W3	P1	P2	–	L3	L4	L5	L6	L2	L1	–	–
W4	P1	P2	–	L3	L4	L5	L6	L2	L1	–	–
W5	P1	P2	P5	L7	L8	L3	L4	L5	L6	L2	L1
W6	P1	P2	P5	L7	L8	L3	L4	L5	L6	L2	L1
W7	P2	P3	–	L6	L7	L5	L3	L4	–	–	–
W8	P2	P3	–	L6	L7	L5	L3	L4	–	–	–
W9	P3	P5	–	L4	L6	L7	L8	–	–	–	–
W10	P3	P5	–	L4	L6	L7	L8	–	–	–	–
W11	P3	P4	P5	L7	L8	L9	L1	L6	L4	–	–
W12	P3	P4	P5	L7	L8	L9	L1	L6	L4	–	–
W13	P4	P5	–	L8	L4	L7	L9	L1	–	–	–
W14	P4	P5		L8	L4	L7	L9	L1	–	–	–

Table 7.9 Weld joints (stack-ups)

Joint	Weld stack-ups		
Joint 1	P1	P4	
Joint 2	P1	P2	
Joint 3	P1	P2	P5
Joint 4	P2	P3	
Joint 5	P3	P5	
Joint 6	P3	P4	P5
Joint 7	P4	P5	

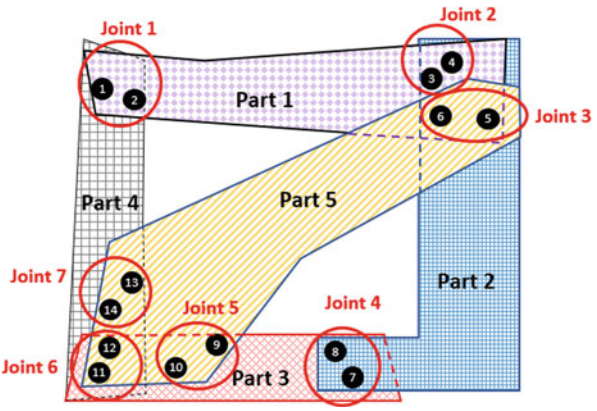


Fig. 7.29 Different joints (stack-ups) exist in the structure

Solving this simple structure enabled us to test the process. The next step would be modelling a real industry structure. We have chosen a bodyside assembly to model and in the next section.

Table 7.10 Joints and welds in each joint

Joint	Weld stack-ups			Weld 1	Weld 2	Weld 3	Weld 4	Weld 5
Joint 1	P1	P4		W1	W2	–	–	–
Joint 2	P1	P2		W3	W4	–	–	–
Joint 3	P1	P2	P5	W5	W6	–	–	–
Joint 4	P2	P3		W7	W8	–	–	–
Joint 5	P3	P5		W9	W10	–	–	–
Joint 6	P3	P4	P5	W11	W12	–	–	–
Joint 7	P4	P5		W13	W14	–	–	–

GA Model Development of Bodyside Assembly

Before modelling the final structure, which is a bodyshell, we needed to test the method we developed in Sect. 7.3.3.2.1 with a real industry assembly. The bodyshell structure will have millions of different solutions and in case of any issues, debugging the method and the process would not be easily possible. So, this step was critical to ensure the method was adequately developed and if any further considerations were required before running the huge bodyshell model.

Bodyside assembly was chosen due to its semi-complicated structure. Bodysides are one of the most critical components of a vehicle body. Any dimensional variation in bodysides may result in aesthetic (e.g. margin and flush between the side doors and body) or functional (e.g. side doors closing effort) issues in the vehicle. The process was assumed a CKD2 level assembly, in which the sub-assemblies (assumed pre-made at supplier site) are being welded together to shape the bodyside (Fig. 7.32).

This bodyside assembly included six sub-assemblies, 62 locators, and 70 spot welds (Fig. 7.33).

As you can see in Fig. 7.32, the structure consists of 6 sub-assemblies including:

1- Front hinge pillar
2- A-pillar
3- Roof rail
4- Rear pillar
5- Rocker assembly
6- B-pillar

These sub-assemblies are located using 62 locators. The product design includes 70 welds which connect the sub-assemblies together to shape the bodyside assembly. The aim is to find the optimum process design (minimum weld combination) that can dimensionally set and fix the assembly.

Similar to the first example, to develop the GA model, we first needed to create the parts, locators, and weld data tables, as explained in Sect. 7.3.3.2.1. Using the CAD data, these tables were extracted as below (Table 7.14).

Table 7.11 List of all relevant welds to each locator

Locator	Related welds														
L1	W1	W2	W13	W14	W12	W11	W3	W6	W4	W5					
L2	W8	W7	W9	W10	W2	W13	W1	W3	W14	W12	W11	W6	W4	W5	
L3	W5	W4	W6	W3	W7	W8	W2	W1							
L4	W5	W4	W6	W3											
L5	W5	W6	W4	W3	W7	W8	W9	W10	W13	W2	W14	W2	W111		
L6	W7	W8	W9	W10	W3	W6	W4	W5	W12	W11					
L7	W12	W14	W13	W11	W10	W9									
L8	W13	W14	W12	W11	W2	W10	W1	W9	W6	W5					
L9	W1	W2	W13	W14	W12	W11	W10	W9	W8	W7	W3	W6	W4	W5	

Table 7.12 List of locators and all relevant weld and distance

Locator	Related welds																											
	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)	Weld	Distance (mm)
L1	W1	14.26	W2	121.51	W13	442.77	W14	514.92	W12	589.98	W11	627.40	W3	2573.97	W6	2726.08	W4	2777.64	W5	2953.78										
L2	W8	598.77	W7	650.08	W9	825.55	W10	996.82	W2	1146.39	W13	1219.39	W1	1278.98	W3	1323.08	W14	1327.52	W12	1377.98	W11	1451.08	W6	1479.47	W4	1526.09	W5	1703.34
L3	W5	127.55	W4	180.10	W6	298.32	W3	395.29	W7	1634.15	W8	1776.98	W2	2858.99	W1	2964.98														
L4	W5	100.11	W4	176.00	W6	190.55	W3	340.99																						
L5	W5	175.14	W6	518.33	W4	538.66	W3	658.59	W7	1566.45	W8	1721.32	W9	2320.75	W10	2561.79	W13	2877.81	W2	2884.46	W14	2965.94	W12	2994.00	W1	3001.74	W11	3189.68
L6	W7	90.12	W8	252.89	W9	849.01	W10	1080.81	W3	1162.06	W6	1290.03	W4	1384.34	W5	1495.44	W12	1539.15	W11	1581.02										
L7	W12	180.87	W14	185.69	W13	197.10	W11	222.07	W10	279.61	W9	498.01																		
L8	W13	25.98	W14	140.84	W12	215.57	W11	266.67	W2	345.33	W10	408.06	W1	409.59	W9	502.60	W6	2620.73	W5	2653.89										
L9	W1	100.17	W2	206.75	W13	312.95	W14	356.27	W12	425.68	W11	436.87	W10	691.23	W9	853.76	W8	1438.40	W7	1587.02	W3	2662.01	W6	2792.81	W4	2850.98	W5	3189.68

Table 7.13 Welds with maximum distance in each joint

Joint	Weld 1	Weld 2	Distance(mm)
Joint 1	W1	W2	122.88
Joint 2	W3	W4	222.83
Joint 3	W5	W6	227.88
Joint 4	W7	W8	164.59
Joint 5	W9	W10	213.15
Joint 6	W11	W12	55.53
Joint 7	W13	W14	118.63

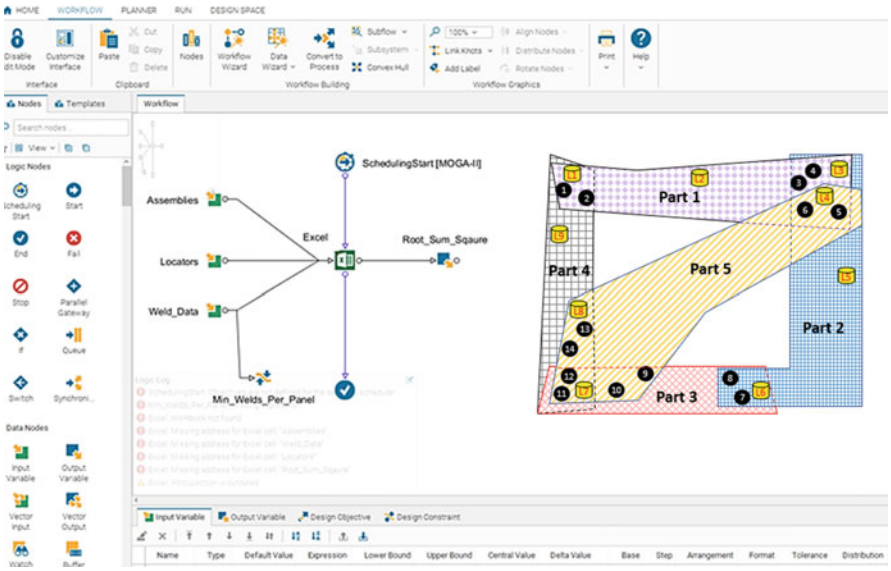


Fig. 7.30 Schematic of GA model developed in modeFRONTIER

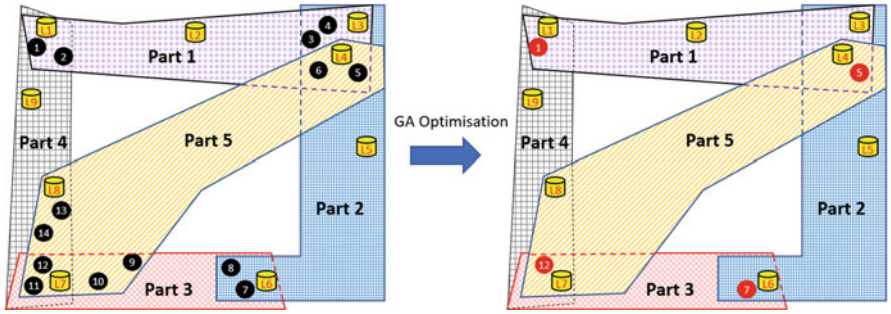


Fig. 7.31 GA optimisation to identify a minimum number of welds to best set the geometry

The next step was to define the weld relationship tables. Each weld is connecting 2 or more parts together. So, we had to define a "weld-part" relationship table. Note

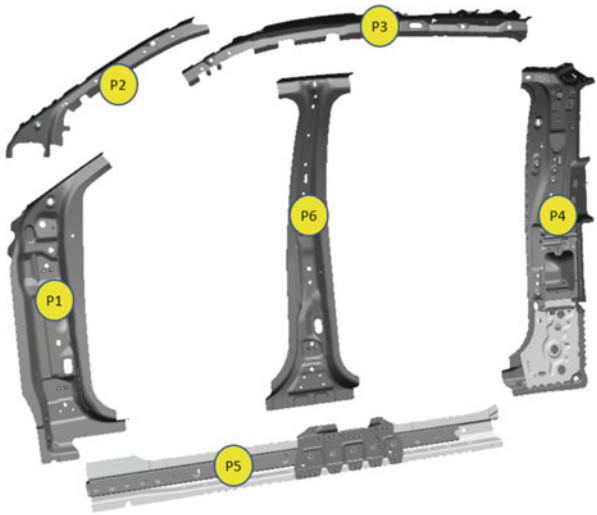


Fig. 7.32 Bodyside structure consisting of six sub-assemblies

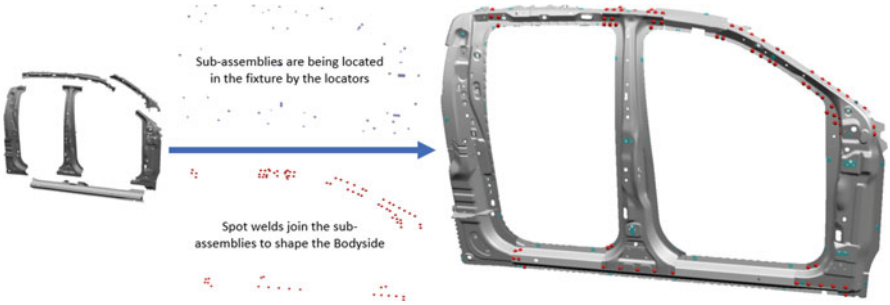


Fig. 7.33 Bodyside assembly with 70 spot welds

Table 7.14 Bodyside assembly part-locator-weld data

Part	Code	Locator	Welds
Front Hinge Pillar	P1	L1-L62	W1-W70
A-Pillar	P2		
Roof rail	P3		
Rear Pillar	P4		
Rocker assembly	P5		
B-Pillar	P6		

this data can be extracted from most of the 3D packages. Specially in complicated structures, it takes hours to go through the 3D model and manually identify the weld stack-ups would include errors. This is usually in the spot weld specification in CAD, so should be easy to get a report from CAD with weld numbers and connecting panel data (Table 7.15).

Like the other assembly, we then needed to develop the relationship between locators and panels, meaning which locator was coordinating which panels. As

Table 7.15 Bodyside weld-parts relationship table

Weld	Connecting parts		Weld	Connecting parts		Weld	Connecting parts	
W1	P5	P1	W25	P3	P6	W49	P1	P2
W2	P5	P1	W26	P3	P6	W50	P1	P2
W3	P5	P1	W27	P3	P6	W51	P1	P2
W4	P5	P1	W28	P3	P6	W52	P1	P2
W5	P5	P1	W29	P3	P6	W53	P1	P2
W6	P5	P6	W30	P3	P6	W54	P1	P2
W7	P5	P6	W31	P3	P6	W55	P1	P2
W8	P5	P6	W32	P3	P6	W56	P1	P2
W9	P5	P6	W33	P3	P6	W57	P1	P2
W10	P5	P6	W34	P3	P6	W58	P1	P2
W11	P5	P6	W35	P2	P3	W59	P1	P2
W12	P5	P6	W36	P2	P3	W60	P1	P2
W13	P1	P5	W37	P2	P3	W61	P1	P2
W14	P1	P5	W38	P2	P3	W62	P1	P2
W15	P1	P5	W39	P2	P3	W63	P1	P2
W16	P4	P5	W40	P2	P3	W64	P1	P2
W17	P4	P5	W41	P2	P3	W65	P1	P2
W18	P4	P5	W42	P2	P3	W66	P1	P2
W19	P4	P5	W43	P2	P3	W67	P3	P4
W20	P4	P5	W44	P2	P3	W68	P3	P4
W21	P3	P6	W45	P2	P3	W69	P3	P4
W22	P3	P6	W46	P2	P3	W70	P3	P4
W23	P3	P6	W47	P1	P2			
W24	P3	P6	W48	P1	P2			

mentioned before, some locators are in the parts overlapping areas so they can be in relation with more than one part. Table 7.16 summarises the locator-part relationship for this assembly.

The combined relationship table would have welds as the key elements and all other components of the system including parts and locators listed in relation to each weld. By assessing the weld-parts and locators-parts tables, we developed the combined relationship table for the bodyside assembly. Through VBA coding, we identified all available weld stack-ups. The weld stack-up table for this assembly is summarised in Table 7.17.

Figure 7.34 is an illustration of Table 7.17, showing all different stack-ups in the structure, each including some welds.

The VBA code to extract the welds included in each joint was developed. For example, Joint 1 includes a stack-up of part 5 and part 1. Table 7.18 lists the joints and all welds included in each one. Unlike the first structure that was reviewed, this larger structure that consists of more welds requires more navigation time to explore the 3D model to identify each joint weld. So, a code was developed to ease the creation of the joints and welds table.

Table 7.16 Bodyside locator-part relationship table

Locator	Locator type	Parts		Locator	Locator type	Parts		Locator	Locator type	Parts	
L1	Hole (Pin)	P1	P5	L22	Slot (Pin)	P4		L43	Surface (Locator)	P4	P5
L2	Hole (Pin)	P2		L23	Slot (Pin)	P5		L44	Surface (Locator)	P5	
L3	Hole (Pin)	P3		L24	Slot (Pin)	P2		L45	Surface (Locator)	P4	
L4	Hole (Pin)	P4		L25	Slot (Pin)	P6		L46	Surface (Locator)	P5	
L5	Hole (Pin)	P6		L26	Slot (Pin)	P4		L47	Surface (Locator)	P4	
L6	Surface (Locator)	P1		L27	Surface (Locator)	P2	P3	L48	Surface (Locator)	P4	
L7	Surface (Locator)	P1		L28	Surface (Locator)	P2	P3	L49	Surface (Locator)	P4	
L8	Surface (Locator)	P1	P2	L29	Surface (Locator)	P2	P3	L50	Surface (Locator)	P5	P6
L9	Surface (Locator)	P3	P6	L30	Surface (Locator)	P2		L51	Surface (Locator)	P5	P6
L10	Surface (Locator)	P3	P6	L31	Surface (Locator)	P2		L52	Surface (Locator)	P5	P6
L11	Surface (Locator)	P3	P6	L32	Surface (Locator)	P2		L53	Surface (Locator)	P5	P6
L12	Surface (Locator)	P5	P1	L33	Surface (Locator)	P3	P4	L54	Surface (Locator)	P5	
L13	Surface (Locator)	P5	P1	L34	Surface (Locator)	P3	P4	L55	Surface (Locator)	P5	
L14	Surface (Locator)	P5	P1	L35	Surface (Locator)	P3	P4	L56	Surface (Locator)	P5	
L15	Surface (Locator)	P1	P2	L36	Surface (Locator)	P4		L57	Surface (Locator)	P5	
L16	Surface (Locator)	P1		L37	Surface (Locator)	P4		L58	Surface (Locator)	P6	
L17	Surface (Locator)	P1		L38	Surface (Locator)	P4		L59	Surface (Locator)	P6	
L18	Surface (Locator)	P1		L39	Surface (Locator)	P3		L60	Surface (Locator)	P6	
L19	Surface (Locator)	P1		L40	Surface (Locator)	P3		L61	Surface (Locator)	P6	
L20	Slot (Pin)	P1		L41	Surface (Locator)	P3		L62	Surface (Locator)	P6	
L21	Slot (Pin)	P3		L42	Surface (Locator)	P4	P5				

Table 7.17 Bodyside weld joints (stack-ups)

Joint	Weld stack-ups	
1	P5	P1
2	P5	P6
3	P4	P5
4	P3	P6
5	P2	P3
6	P1	P2
7	P3	P4

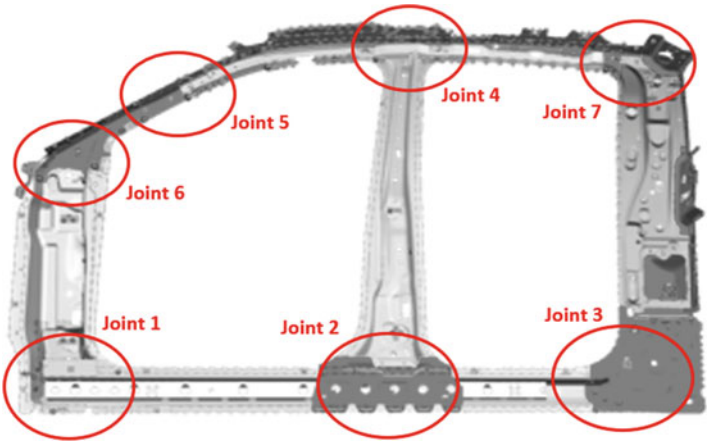


Fig. 7.34 Seven joints (stack-ups) exist in this bodyside structure

The next step was to calculate the key dimensions in the structure. This included the distance between each locator to all relevant welds, as well as the maximum distance between welds in each joint. To extract the first list, all locators needed to be listed and all welds associated to that locator need to be identified. Using the above tables, this step was coded to list all relevant welds to each locator (Table 7.19).

Using the coordinates of these locators and welds, the distance between each locator to any of its related welds was calculated (Table 7.20).

We then assessed all weld distances in each joint and listed the two welds with maximum distance in that joint. This was coded and the results are shown in Table 7.21.

The above developed tables from bodyside structure CAD would characterise each component in the structure and could dimension the relation between the components. These tables could then be used as input data for the GA model coding. Similar to the first structure, we defined the root of sum square of the distance of each weld to the locator with the same stack-up as the fitness function.

The bodyside GA model was developed in modeFRONTIER using the above tables. A schematic of the structure and the GA model is shown in Fig. 7.35.

Table 7.18 Bodyside joints and all welds in each joint

Joint	Stack-ups		Welds in the joint															
1	P5	P1	W1	W2	W3	W4	W5	W13	W14	W15								
2	P5	P6	W6	W7	W8	W9	W10	W11	W12									
3	P4	P5	W16	W17	W18	W19	W20											
4	P3	P6	W21	W22	W23	W24	W25	W26	W27	W28								
5	P2	P3	W35	W36	W37	W38	W39	W40	W41	W42	W29	W30	W31	W32	W33	W34		
6	P1	P2	W47	W48	W49	W50	W51	W52	W53	W54	W43	W44	W45	W46				
7	P3	P4	W67	W68	W69	W70	W55	W56	W57	W58	W59	W60	W61	W62	W63	W64	W65	W66

Table 7.19 List of bodyside key locators with associated welds to each

Locator	Welds in the joint																			
L1	W14	W13	W1	W2	W3	W4	W5	W15												
L8	W54	W65	W66	W62	W64	W55	W52	W53	W63	W49	W56	W50	W51	W57	W58	W59	W60	W61	W47	W48
L9	W26	W24	W25	W23	W27	W34	W28	W29	W30	W33	W22	W31	W21	W32						
L10	W34	W25	W26	W33	W23	W24	W22	W21	W29	W28	W27	W30	W31	W32						
L11	W21	W32	W22	W33	W31	W30	W29	W34	W23	W25	W28	W26	W24	W27						
L12	W1	W15	W2	W3	W14	W4	W13	W5												
L13	W15	W1	W2	W3	W4	W14	W13	W5												
L14	W5	W13	W4	W14	W3	W2	W1	W15												
L15	W54	W55	W65	W56	W66	W64	W57	W62	W58	W63	W53	W59	W52	W51	W50	W60	W49	W61	W48	W47
L27	W45	W44	W40	W35	W39	W36	W38	W37	W41	W43	W46	W42								
L28	W45	W44	W40	W35	W39	W36	W38	W37	W41	W46	W43	W42								
L29	W45	W44	W40	W35	W39	W38	W36	W37	W41	W43	W46	W42								
L33	W67	W70	W69	W68																
L34	W67	W69	W70	W68																
L35	W67	W69	W70	W68																
L42	W16	W19	W17	W18	W20															
L43	W16	W19	W17	W18	W20															
L50	W12	W11	W7	W6	W8	W9	W10													
L51	W10	W9	W8	W11	W12	W6	W7													
L52	W10	W9	W8	W11	W12	W6	W7													
L53	W10	W9	W8	W6	W11	W12	W7													

Table 7.20 List of bodyside key locators and all relevant weld and distance

Locator		L1	L8	L9	L10	L11	L12	L13	L14	L15	L27	L28	L29	L33	L34	L35	L42	L43	L50	L51	L52	L53
Related welds	Weld	W14	W54	W26	W34	W21	W1	W15	W5	W54	W45	W45	W45	W67	W67	W67	W16	W16	W12	W10	W10	W10
	Di (mm)	798.01	2047.45	34.09	1110.19	90.28	111.93	26.27	638.99	774.47	479.85	409.05	584.92	1043.99	1404.54	1156.40	100.13	662.65	952.60	1226.53	1192.12	978.34
	Weld	W13	W65	W24	W25	W32	W15	W1	W13	W55	W44	W44	W44	W70	W69	W69	W19	W19	W11	W9	W9	W9
	Dis. (mm)	818.82	2061.34	44.94	1110.41	117.49	117.73	101.28	677.99	790.30	495.95	414.07	609.56	1076.74	1409.78	1158.78	121.67	663.97	957.70	1268.21	1233.47	1054.64
	Weld	W1	W66	W25	W26	W22	W2	W2	W4	W65	W40	W40	W40	W69	W70	W70	W17	W17	W7	W8	W8	W8
	Dis. (mm)	865.38	2085.44	64.08	1111.43	128.20	164.83	103.04	697.96	803.66	509.34	439.12	613.13	1084.57	1450.26	1199.43	124.85	683.35	1013.39	1313.33	278.45	1130.27
	Weld	W2	W62	W23	W33	W33	W3	W3	W14	W56	W35	W35	W35	W68	W68	W68	W18	W18	W6	W11	W11	W6
	Dis. (mm)	865.38	2090.09	78.77	1115.55	136.36	221.38	134.83	714.95	847.25	524.58	443.79	636.76	1108.91	1464.57	1211.45	143.82	691.58	1070.73	1320.92	1287.22	1208.17
	Weld	W3	W64	W27	W23	W31	W14	W4	W3	W66	W39	W39	W39				W20	W20	W8	W12	W12	W11
	Dis. (mm)	869.53	2098.62	84.02	1137.93	147.72	255.80	181.49	757.10	852.21	557.85	488.58	659.54				161.19	697.73	1133.75	1344.31	1310.64	1217.91
	Weld	W4	W55	W34	W24	W30	W4	W14	W2	W64	W36	W36	W38						W9	W6	W6	W12
	Dis. (mm)	877.77	2101.92	94.08	1139.51	179.70	279.37	190.03	816.35	881.17	601.64	523.38	704.76						1200.52	1363.17	1328.28	1242.90
	Weld	W5	W52	W28	W22	W29	W13	W13	W1	W57	W38	W38	W36						W10	W7	W7	W7
	Dis. (mm)	889.99	2108.35	105.74	1146.37	310.12	288.83	207.71	875.71	882.14	604.79	536.43	710.31						1273.93	1416.02	1381.23	1286.18
	Weld	W15	W53	W29	W21	W34	W5	W15	W62	W37	W37	W37										
	Dis. (mm)	935.45	2116.93	135.22	1151.75	316.46	338.07	234.30	883.85	930.88	661.68	585.57	767.49									
	Weld		W63	W30	W29	W23				W58	W41	W41	W41									
	Dis. (mm)		2121.57	261.07	1178.15	339.65				935.76	717.87	643.51	821.36									
	Weld		W49	W33	W28	W25				W63	W43	W46	W43									
	Dis. (mm)		2121.63	274.07	1178.93	346.41				938.04	756.35	688.26	849.55									
	Weld		W56	W22	W27	W28				W53	W46	W43	W46									
	Dis. (mm)		2130.31	287.83	1180.90	347.39				966.88	760.96	692.24	862.29									
	Weld		W50	W31	W30	W26				W59	W42	W42	W42									
	Dis. (mm)		2133.41	295.44	1184.37	376.37				983.70	784.44	721.40	876.16									
	Weld		W51	W21	W31	W24				W52												
	Dis. (mm)		2142.10	327.04	1187.79	378.36				988.04												
	Weld		W57	W32	W32	W27				W51												
	Dis. (mm)		2148.22	331.22	1192.56	384.60				998.48												
	Weld		W58							W50												
	Dis. (mm)		2176.46							1016.35												
	Weld		W59							W60												
	Dis. (mm)		2202.34							1031.68												
	Weld		W60							W49												
	Dis. (mm)		2228.82							1038.13												
	Weld		W61							W61												
	Dis. (mm)		2255.92							1078.44												
	Weld		W47							W48												
	Dis. (mm)		2262.17							1151.84												
	Weld		W48							W47												
	Dis. (mm)		2268.82							1170.23												

Table 7.21 Bodyside welds with maximum distance in each joint

Joint	Weld 1	Weld 2	Distance (mm)
Joint 1	W5	W15	258.89
Joint 2	W10	W12	379.61
Joint 3	W17	W20	117.18
Joint 4	W21	W27	296.18
Joint 5	W35	W34	809.49
Joint 6	W47	W43	848.12
Joint 7	W67	W68	101.67

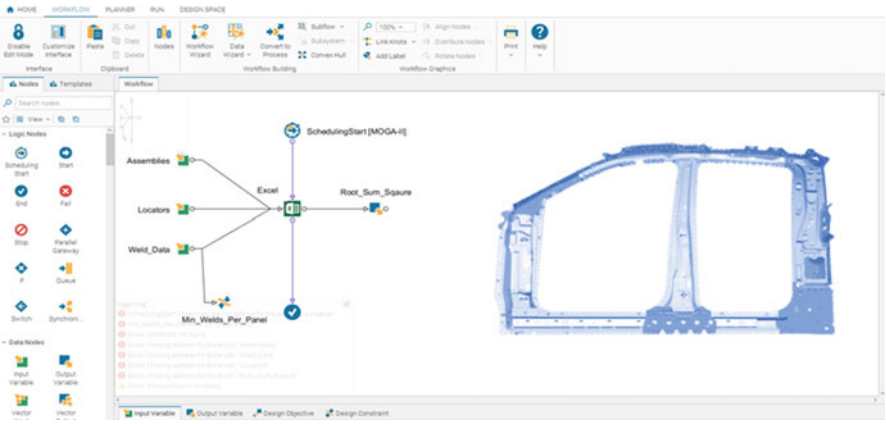


Fig. 7.35 Schematic of bodyside GA model developed in modeFRONTIER

Table 7.22 Optimum bodyside weld design resulted from GA model

GA Outcome			
W14	W15	W11	W36
W13	W5	W10	W38
W54	W55	W9	W47
W65	W45	W7	W43
W26	W44	W17	W46
W24	W67	W20	W48
W34	W70	W27	W49
W25	W69	W35	W50
W21	W16	W29	W68
W32	W19	W40	
W1	W12	W39	

The model could continue running until the fittest subset of welds were generated. In this GA model, the fittest subset of welds was resulted as below (Table 7.22).

In practice, it means these are geometry setting welds (42 welds out of the total of 70 welds) can set the geometry of the bodyside assembly with minimum dimensional variation (Fig. 7.36).

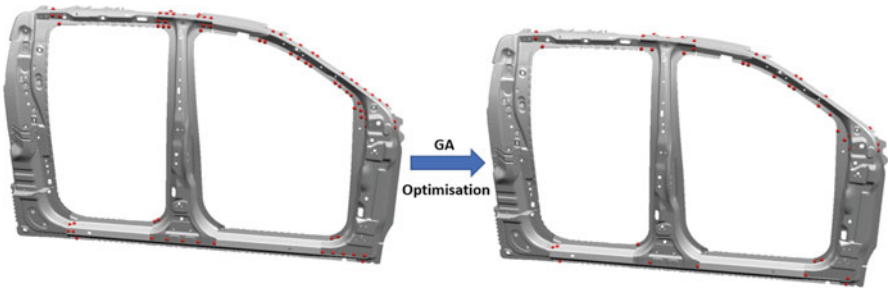


Fig. 7.36 GA optimisation of bodyside welds

Note the process design resulted from the GA model in this structure does not mean this is the minimum number of welds required to set the geometry. To design the optimum process and to identify the minimum required welds, we could continue running the GA model though it would be more reliable to limit the number of designs through GA and then use another method such as finite element analysis (we will do it in the next chapter) to analyse the optimum process design.

GA Model Development for Bodyshell

As explained in Chap. 1, the bodyshell structure problem would have millions of different solutions (process design), and in case of any issues, debugging the method seems unlikely. In the previous two sections, we showed how the initial GA model was developed on a simple structure and then we tested it on a real industry assembly. We now can use the same method to develop a GA model for the bodyshell, which is the main structure investigated in this study.

Bodyshell is the product of body framing line. The framing process is the most dimensionally critical process in body shops as any dimensional variation in this process can directly or indirectly result in the vehicle quality issues and possible systems malfunctioning.

The bodyshell studied in this research consists of 12 assemblies (Fig. 7.37).

The 12 assemblies are located via 94 locators in the framing station and welded together by 850 welds (Fig. 7.38).

As illustrated in Fig. 7.37, the structure consists of 12 assemblies including:

1- Engine compartment
2- Upper dash assembly
3- Lower dash panel
4- Front floor assembly
5- Rear floor assembly
6- Cab back panel
7- Right-hand bodyside
8- Left-hand bodyside

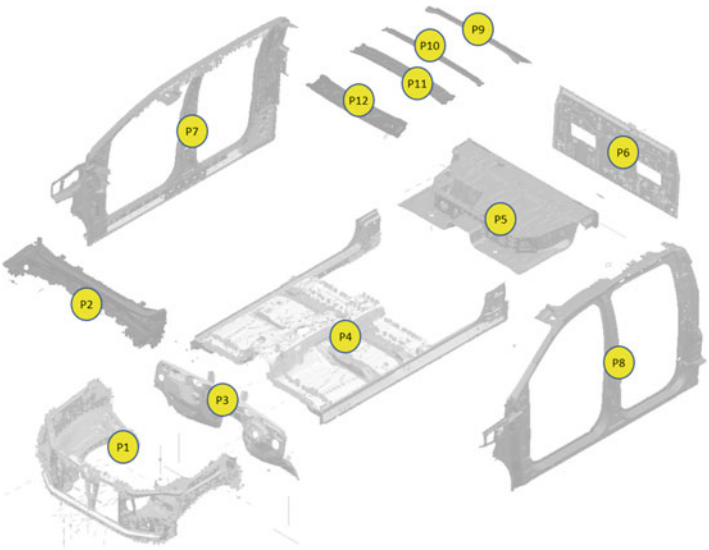


Fig. 7.37 The studied Bodyshell consisting of 12 assemblies

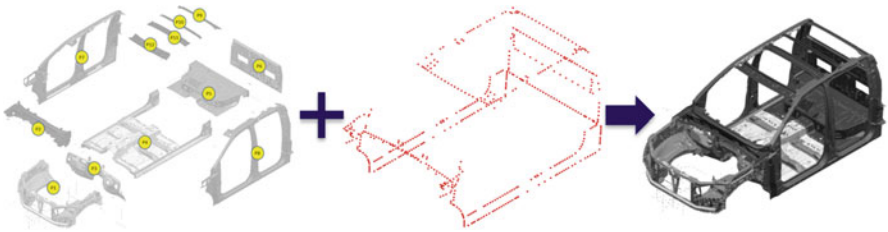


Fig. 7.38 This bodyshell is the product of 12 assemblies joined together via 850 welds

9- Rear header
10- Roof bow#1
11- Roof bow#2
12- Front header

These assemblies are located in the framing station using 94 locators. The product design requires 850 welds to robustly join all these assemblies together to shape the bodyshell. A process design of around 60 welds (station capacity as explained in Chap. 1) out of 850 has unlimited solutions and not all can be tested or verified. As discussed before, due to the limited available time in the framing station in high-rate production lines, an optimum process design is required to minimise the process-driven dimensional variations. Using the findings from the previous two examples we completed in GA modelling, in this case we aim to limit the number of solutions by eliminating non-viable designs and keeping the more fit solutions.

Table 7.23 Bodyshell parts-locators-welds data table

Part	Code	Locator
Engine compartment	P1	L1-L94
Upper dash assembly	P2	
Lower dash panel	P3	
Front floor assembly	P4	
Rear floor assembly	P5	Welds
Cab back panel	P6	W1-W850
Right hand Bodyside	P7	
Left hand Bodyside	P8	
Rear header	P9	
Roof bow#1	P10	
Roof bow#2	P11	
Front header	P12	

The parts/assemblies, locators, and weld data tables can be developed with the same methods explained in Sects. 7.3.3.1 and 7.3.3.2. Unlike the simple structures, any data analysis in this complicated structure shall be done using coding to minimise the errors and to save time. The part and locator-weld data for bodyshell are developed as per Table 7.23.

The weld relationship tables could be a bit more complicated due to the product design and joints in the bodyshell. Each weld is connecting two or more assemblies together. Note as each assembly consists of some parts, the actual connecting panel data extracted from CAD may not be the same for all spot welds in the same joint connecting two assemblies together. This requires an additional assumption of regrouping the welds from panel-to-panel connection to assembly-to-assembly connection. Also, each assembly is assumed that has already been dimensionally fixed in upstream processes. The weld-part relationship table for the bodyshell is extracted from CAD (Tables 7.24, 7.25, 7.26, 7.27, and 7.28).

Locator-part relationship table can now be developed. Usually, in 3D models of locators, they are associated with single panels, so getting a part-locator report from 3D model will have some single panel numbers instead of our 12 assemblies as listed before. In this case, we need to cross-check the data and where needed replace the panel number with the parent assembly. Table 7.29 summarises the locator-part relationship for this assembly.

The combined weld-locator-part relationship table lists the welds as the key elements and all locators and assemblies in relation to each weld. By assessing the weld-part and locator-part tables and using VBA coding, the combined relationship table for the bodyshell were developed.

All available weld stack-ups can be identified manually or through coding, though we used coding due to the complicated structure of bodyshell. The results were randomly cross-checked with the 3D model of bodyshell to ensure the coding is correctly extracting the weld stack-ups. The joint and weld stack-up relationship for the bodyshell is summarised in Table 7.30.

Table 7.24 Bodyshell weld-part relationship table – part 1/5

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies	
W1	P1	P3	P8	W61	P4	P8		W121	P2	P7
W2	P1	P3	P7	W62	P4	P7		W122	P2	P8
W3	P1	P3	P8	W63	P2	P3		W123	P2	P7
W4	P1	P3	P7	W64	P2	P3		W124	P2	P8
W5	P1	P3	P8	W65	P2	P3		W125	P2	P7
W6	P1	P3	P7	W66	P2	P3		W126	P2	P3
W7	P3	P8		W67	P2	P3		W127	P2	P3
W8	P3	P7		W68	P2	P3		W128	P2	P3
W9	P3	P8		W69	P2	P3		W129	P2	P3
W10	P3	P7		W70	P2	P3		W130	P2	P3
W11	P3	P8		W71	P2	P3		W131	P2	P3
W12	P3	P7		W72	P2	P3		W132	P2	P3
W13	P3	P8		W73	P2	P3		W133	P2	P3
W14	P3	P7		W74	P2	P3		W134	P4	P8
W15	P3	P8		W75	P2	P3		W135	P4	P7
W16	P3	P7		W76	P2	P3		W136	P4	P8
W17	P3	P8		W77	P2	P3		W137	P4	P7
W18	P3	P7		W78	P1	P7		W138	P4	P8
W19	P3	P8		W79	P1	P7		W139	P4	P7
W20	P3	P7		W80	P1	P8		W140	P4	P8
W21	P3	P8		W81	P1	P8		W141	P4	P7
W22	P3	P7		W82	P2	P8		W142	P4	P8
W23	P1	P8		W83	P2	P7		W143	P4	P7
W24	P1	P7		W84	P2	P3		W144	P4	P8
W25	P1	P8		W85	P2	P3		W145	P4	P7
W26	P1	P7		W86	P2	P3		W146	P4	P8
W27	P1	P8		W87	P2	P3		W147	P4	P7
W28	P1	P7		W88	P2	P3		W148	P4	P8
W29	P1	P8		W89	P2	P3		W149	P4	P7
W30	P1	P7		W90	P2	P3		W150	P4	P8
W31	P1	P8		W91	P2	P3		W151	P4	P7
W32	P1	P7		W92	P2	P3		W152	P4	P8
W33	P1	P8		W93	P2	P3		W153	P4	P7
W34	P1	P7		W94	P2	P3		W154	P4	P8
W35	P1	P8		W95	P2	P3		W155	P4	P7
W36	P1	P7		W96	P2	P3		W156	P4	P8
W37	P1	P3	P8	W97	P2	P3		W157	P4	P7
W38	P1	P3	P7	W98	P2	P3		W158	P4	P8
W39	P2	P8		W99	P2	P3		W159	P4	P7
W40	P2	P7		W100	P2	P3		W160	P4	P8
W41	P2	P8		W101	P2	P3		W161	P4	P7

(continued)

Table 7.24 (continued)

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies	
W42	P2	P7		W102	P2	P3		W162	P4	P8
W43	P2	P8		W103	P2	P3		W163	P4	P7
W44	P2	P7		W104	P2	P3		W164	P4	P8
W45	P2	P8		W105	P2	P3		W165	P4	P7
W46	P2	P7		W106	P2	P3		W166	P4	P8
W47	P2	P8		W107	P2	P3		W167	P4	P7
W48	P2	P7		W108	P2	P3		W168	P4	P8
W49	P2	P8		W109	P2	P3		W169	P4	P7
W50	P2	P7		W110	P2	P3		W170	P4	P8
W51	P2	P8		W111	P2	P3		W171	P4	P7
W52	P2	P7		W112	P2	P3		W172	P4	P8
W53	P2	P8		W113	P2	P3		W173	P4	P7
W54	P2	P7		W114	P2	P3	P8	W174	P4	P8
W55	P4	P8		W115	P2	P3	P7	W175	P4	P7
W56	P4	P7		W116	P2	P3	P8	W176	P4	P8
W57	P4	P8		W117	P2	P3	P7	W177	P4	P7
W58	P4	P7		W118	P2	P8		W178	P4	P8
W59	P4	P8		W119	P2	P7		W179	P4	P7
W60	P4	P7		W120	P2	P8		W180	P4	P8

This table shows how many different stack-up of assemblies exist in the bodyshell, regardless of the number of welds for each joint. It is critical to check the results to ensure in the final process design each joint has a minimum of one weld. Leaving any joints without any welds will cause significant dimensional variation of the structure.

Figure 7.39 illustrates the grouping as per Table 7.30. These are all joints connecting the assemblies together.

Table 7.31 lists the joints and all welds included in each joint. Unlike the two structures previously reviewed, in this larger structure with more welds, it will take a very long time to navigate through 3D model to identify the weld in each joint. So, we benefited from the codes we developed to help us create the joints and weld tables (Table 7.31).

To calculate the distances in the next step, we first needed to develop a list of locators and all welds associated with each locator. Note that not all locators may have relevant welds that are dimensionally critical to the process (Tables 7.32, 7.33, 7.34, 7.35, and 7.36).

Table 7.25 Bodyshell weld-part relationship table – part 2/5

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W181	P4	P7		W241	P4	P7		W301	P4	P7	
W182	P4	P8		W242	P4	P8		W302	P4	P8	
W183	P4	P7		W243	P4	P7		W303	P4	P7	
W184	P4	P8		W244	P4	P8		W304	P4	P8	
W185	P4	P7		W245	P4	P7		W304	P4	P7	
W186	P4	P8		W246	P4	P8		W306	P5	P8	
W187	P4	P7		W247	P4	P7		W307	P5	P7	
W188	P4	P8		W248	P4	P8		W308	P5	P8	
W189	P4	P7		W249	P4	P7		W309	P5	P7	
W190	P4	P8		W250	P4	P8		W310	P5	P8	
W191	P4	P7		W251	P4	P7		W311	P5	P7	
W192	P4	P8		W252	P4	P8		W312	P5	P8	
W193	P4	P7		W253	P4	P7		W313	P5	P7	
W194	P4	P8		W254	P4	P8		W314	P5	P8	
W195	P4	P7		W255	P4	P7		W315	P5	P7	
W196	P4	P8		W256	P4	P8		W316	P4	P8	
W197	P4	P7		W257	P4	P7		W317	P4	P7	
W198	P4	P8		W258	P4	P8		W318	P4	P8	
W199	P4	P7		W259	P4	P7		W319	P4	P7	
W200	P4	P8		W260	P4	P8		W320	P4	P8	
W201	P4	P7		W261	P4	P7		W321	P4	P7	
W202	P4	P8		W262	P4	P8		W322	P4	P8	
W203	P4	P7		W263	P4	P7		W323	P4	P7	
W204	P4	P8		W264	P4	P8		W324	P4	P8	
W205	P4	P7		W265	P4	P7		W325	P4	P7	
W206	P4	P8		W266	P4	P8		W326	P2	P8	
W207	P4	P7		W267	P4	P7		W327	P2	P7	
W208	P4	P8		W268	P4	P8		W328	P6	P8	
W209	P4	P7		W269	P4	P7		W329	P6	P7	
W210	P4	P8		W270	P4	P8		W330	P6	P8	
W211	P4	P7		W271	P4	P7		W331	P6	P7	
W212	P4	P8		W272	P4	P8		W332	P6	P8	
W213	P4	P7		W273	P4	P7		W333	P6	P7	
W214	P4	P8		W274	P4	P8		W334	P6	P8	
W215	P4	P7		W275	P4	P7		W335	P6	P7	
W216	P4	P8		W276	P4	P8		W336	P6	P8	
W217	P4	P7		W277	P4	P7		W337	P6	P7	
W218	P4	P8		W278	P4	P8		W338	P6	P8	
W219	P4	P7		W279	P4	P7		W339	P6	P7	
W220	P4	P8		W280	P4	P8		W340	P6	P8	
W221	P4	P7		W281	P4	P7		W341	P6	P7	
W222	P4	P8		W282	P4	P8		W342	P6	P8	

(continued)

Table 7.25 (continued)

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W223	P4	P7		W283	P4	P7		W343	P6	P7	
W224	P4	P8		W284	P4	P8		W344	P6	P8	
W225	P4	P7		W285	P4	P7		W345	P6	P7	
W226	P4	P8		W286	P4	P8		W346	P6	P8	
W227	P4	P7		W287	P4	P7		W347	P6	P7	
W228	P4	P8		W288	P4	P8		W348	P11	P8	
W229	P4	P7		W289	P4	P7		W349	P11	P7	
W230	P4	P8		W290	P4	P8		W350	P11	P8	
W231	P4	P7		W291	P4	P7		W351	P11	P7	
W232	P4	P8		W292	P4	P8		W352	P11	P8	
W233	P4	P7		W293	P4	P7		W353	P11	P7	
W234	P4	P8		W294	P4	P8		W354	P11	P8	
W235	P4	P7		W295	P4	P7		W355	P11	P7	
W236	P4	P8		W296	P4	P8		W356	P10	P8	
W237	P4	P7		W297	P4	P7		W357	P10	P7	
W238	P4	P8		W298	P4	P8		W358	P10	P8	
W239	P4	P7		W299	P4	P7		W359	P10	P7	
W240	P4	P8		W300	P4	P8		W360	P12	P8	

Table 7.26 Bodyshell weld-part relationship table – part 3/5

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W361	P12	P7		W421	P4	P7		W481	P2	P7	
W362	P12	P8		W422	P4	P8		W482	P2	P8	
W363	P12	P7		W423	P4	P7		W483	P2	P7	
W364	P12	P8		W424	P4	P8		W484	P2	P8	
W365	P12	P7		W425	P4	P7		W485	P2	P7	
W366	P12	P8		W426	P4	P8		W486	P2	P3	
W367	P12	P7		W427	P4	P7		W487	P2	P3	
W368	P12	P8		W428	P4	P8		W488	P2	P3	
W369	P12	P7		W429	P4	P7		W489	P2	P3	
W370	P12	P8		W430	P4	P8		W490	P2	P3	
W371	P12	P7		W431	P4	P7		W491	P2	P3	
W372	P12	P8		W432	P4	P8		W492	P2	P3	
W373	P12	P7		W433	P4	P7		W493	P2	P3	
W374	P8	P9		W434	P4	P8		W494	P4	P8	
W375	P7	P9		W435	P2	P3		W495	P4	P7	
W376	P8	P9		W436	P2	P3		W496	P4	P8	
W377	P7	P9		W437	P2	P3		W497	P4	P7	
W378	P8	P9		W438	P2	P3		W498	P4	P8	
W379	P7	P9		W439	P2	P3		W499	P4	P7	
W380	P8	P9		W440	P2	P3		W500	P4	P8	

(continued)

Table 7.26 (continued)

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W381	P7	P9		W441	P2	P3		W501	P4	P7	
W382	P8	P9		W442	P2	P3		W502	P4	P8	
W383	P7	P9		W443	P2	P3		W503	P4	P7	
W384	P8	P9		W444	P2	P3		W504	P4	P8	
W385	P7	P9		W445	P2	P3	P8	W505	P4	P7	
W386	P1	P2		W446	P2	P3	P7	W506	P4	P8	
W387	P1	P2		W447	P4	P8		W507	P4	P7	
W388	P1	P2		W448	P4	P7		W508	P4	P8	
W389	P1	P2		W449	P4	P8		W509	P4	P7	
W390	P1	P2		W450	P4	P7		W510	P4	P8	
W391	P1	P2		W451	P2	P8		W511	P4	P7	
W392	P1	P2		W452	P2	P7		W512	P4	P8	
W393	P1	P2		W453	P6	P8		W513	P4	P7	
W394	P1	P2		W454	P6	P7		W514	P4	P8	
W395	P1	P2		W455	P6	P8		W515	P4	P7	
W396	P1	P2		W456	P6	P7		W516	P4	P8	
W397	P1	P2		W457	P6	P8		W517	P4	P7	
W398	P1	P2		W458	P6	P7		W518	P4	P8	
W399	P1	P2		W459	P6	P8		W519	P4	P7	
W400	P1	P2		W460	P6	P7		W520	P4	P8	
W401	P1	P2		W461	P6	P8		W521	P4	P7	
W402	P1	P2		W462	P6	P7		W522	P4	P8	
W403	P1	P2		W463	P6	P8		W523	P4	P7	
W404	P1	P3	P7	W464	P6	P7		W524	P4	P8	
W405	P1	P3	P8	W465	P6	P8		W525	P4	P7	
W406	P1	P3	P7	W466	P6	P7		W526	P4	P8	
W407	P3	P8		W467	P6	P8		W527	P4	P7	
W408	P3	P7		W468	P6	P7		W528	P4	P8	
W409	P3	P8		W469	P6	P8		W529	P4	P7	
W410	P3	P7		W470	P6	P7		W530	P4	P8	
W411	P3	P8		W471	P6	P8		W531	P4	P7	
W412	P3	P7		W472	P6	P7		W532	P4	P8	
W413	P3	P8		W473	P8	P9		W533	P4	P7	
W414	P3	P7		W474	P7	P9		W534	P4	P8	
W415	P3	P8		W475	P8	P9		W535	P4	P7	
W416	P3	P7		W476	P7	P9		W536	P4	P8	
W417	P3	P8		W477	P8	P9		W537	P4	P7	
W418	P3	P7		W478	P7	P9		W538	P4	P8	
W419	P4	P7		W479	P8	P9		W539	P4	P7	
W420	P4	P8		W480	P7	P9		W540	P4	P8	

Table 7.27 Bodyshell weld-part relationship table – part 4/5

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W541	P6	P7		W601	P4	P8		W661	P2	P3	
W542	P6	P8		W602	P4	P7		W662	P2	P3	
W543	P6	P7		W603	P4	P8		W663	P2	P3	
W544	P6	P8		W604	P4	P7		W664	P2	P3	
W545	P6	P7		W605	P4	P8		W665	P2	P3	
W546	P6	P8		W606	P4	P7		W666	P2	P3	
W547	P6	P7		W607	P4	P8		W667	P2	P3	
W548	P6	P8		W608	P4	P7		W668	P2	P3	
W549	P6	P7		W609	P4	P8		W669	P2	P3	
W550	P6	P8		W610	P4	P7		W670	P2	P3	
W551	P6	P7		W611	P4	P8		W671	P2	P3	
W552	P6	P8		W612	P4	P7		W672	P2	P3	
W553	P6	P7		W613	P4	P8		W673	P2	P3	
W554	P6	P8		W614	P4	P7		W674	P2	P3	
W555	P6	P7		W615	P4	P8		W675	P2	P3	
W556	P6	P8		W616	P4	P7		W676	P2	P3	
W557	P6	P7		W617	P4	P8		W677	P2	P3	
W558	P11	P8		W618	P4	P7		W678	P2	P3	
W559	P11	P7		W619	P4	P8		W679	P2	P3	P8
W560	P11	P8		W620	P4	P7		W680	P2	P3	
W561	P11	P7		W621	P4	P8		W681	P2	P3	
W562	P11	P8		W622	P4	P7		W682	P2	P3	P7
W563	P11	P7		W623	P4	P8		W683	P2	P8	
W564	P11	P8		W624	P4	P7		W684	P2	P7	
W565	P11	P7		W625	P4	P8		W685	P2	P8	
W566	P10	P8		W626	P4	P7		W686	P4	P7	
W567	P10	P7		W627	P4	P8		W687	P5	P8	
W568	P10	P8		W628	P12	P7		W688	P5	P7	
W569	P10	P7		W629	P12	P8		W689	P5	P8	
W570	P12	P8		W630	P12	P7		W690	P5	P7	
W571	P4	P7		W631	P12	P8		W691	P5	P8	
W572	P4	P8		W632	P12	P7		W692	P5	P7	
W573	P4	P7		W633	P12	P8		W693	P5	P8	
W574	P4	P8		W634	P12	P7		W694	P5	P7	
W575	P4	P7		W635	P12	P8		W695	P5	P8	
W576	P4	P8		W636	P12	P7		W696	P5	P7	
W577	P4	P7		W637	P12	P8		W697	P4	P8	
W578	P4	P8		W638	P12	P7		W698	P4	P7	
W579	P4	P7		W639	P12	P8		W699	P4	P8	
W580	P4	P8		W640	P12	P7		W700	P4	P7	
W581	P4	P7		W641	P8	P9		W701	P4	P8	
W582	P4	P8		W642	P7	P9		W702	P4	P7	

(continued)

Table 7.27 (continued)

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W583	P4	P7		W643	P8	P9		W703	P4	P8	
W584	P4	P8		W644	P7	P9		W704	P4	P7	
W585	P2	P3		W645	P8	P9		W705	P4	P8	
W586	P2	P3		W646	P7	P9		W706	P4	P7	
W587	P2	P3		W647	P8	P9		W707	P2	P8	
W588	P2	P3		W648	P7	P9		W708	P2	P7	
W589	P2	P3		W649	P8	P9		W709	P6	P8	
W590	P2	P3		W650	P7	P9		W710	P6	P7	
W591	P2	P3		W651	P8	P9		W711	P6	P8	
W592	P2	P3		W652	P7	P9		W712	P6	P7	
W593	P2	P3		W653	P1	P2		W713	P6	P8	
W594	P2	P3		W654	P1	P2		W714	P6	P7	
W595	P3	P7		W655	P1	P2		W715	P4	P7	
W596	P3	P8		W656	P1	P2		W716	P5	P8	
W597	P3	P7		W657	P1	P2		W717	P5	P7	
W598	P3	P8		W658	P1	P2		W718	P5	P8	
W599	P3	P7		W659	P1	P2		W719	P5	P7	
W600	P3	P8		W660	P1	P2		W720	P5	P8	

Table 7.28 Bodyshell weld-part relationship table – part 5/5

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W721	P4	P7		W781	P1	P2		W841	P4	P8	
W722	P4	P8		W782	P1	P2		W842	P4	P7	
W723	P4	P7		W783	P1	P2		W843	P2	P8	
W724	P4	P8		W784	P1	P2		W844	P2	P7	
W725	P4	P7		W785	P1	P2	P7	W855	P6	P8	
W726	P4	P8		W786	P1	P2		W846	P6	P7	
W727	P4	P7		W787	P1	P2	P7	W847	P6	P8	
W728	P4	P8		W788	P1	P2		W848	P6	P7	
W729	P4	P7		W789	P1	P2		W849	P6	P8	
W730	P4	P8		W790	P1	P2		W850	P6	P7	
W731	P4	P7		W791	P1	P2					
W732	P4	P8		W792	P1	P2					
W733	P4	P7		W793	P1	P2					
W734	P4	P8		W794	P1	P2					
W735	P4	P7		W795	P1	P2					
W736	P4	P8		W796	P1	P3					
W737	P4	P7		W797	P1	P3					
W738	P4	P8		W798	P1	P3					
W739	P4	P7		W799	P3	P8					
W740	P4	P8		W800	P3	P7					

(continued)

Table 7.28 (continued)

Weld	Connecting assemblies			Weld	Connecting assemblies			Weld	Connecting assemblies		
W741	P4	P7		W801	P3	P8					
W742	P4	P8		W802	P3	P7					
W743	P4	P7		W803	P3	P8					
W744	P4	P8		W804	P3	P7					
W745	P4	P7		W805	P3	P8					
W746	P4	P8		W806	P3	P7					
W747	P4	P7		W807	P3	P8					
W748	P4	P8		W808	P3	P7					
W749	P4	P7		W809	P3	P8					
W750	P4	P8		W810	P3	P7					
W751	P4	P7		W811	P4	P7					
W752	P4	P8		W812	P4	P8					
W753	P4	P7		W813	P7	P9					
W754	P4	P8		W814	P8	P9					
W755	P4	P7		W815	P7	P9					
W756	P4	P8		W816	P8	P9					
W757	P4	P7		W817	P7	P9					
W758	P4	P8		W818	P8	P9					
W759	P4	P7		W819	P7	P9					
W760	P4	P8		W820	P8	P9					
W761	P4	P7		W821	P7	P9					
W762	P4	P8		W822	P1	P2					
W763	P4	P7		W823	P1	P2					
W764	P4	P8		W824	P1	P2					
W765	P4	P7		W825	P1	P2					
W766	P4	P8		W826	P1	P2					
W767	P4	P7		W827	P1	P2					
W768	P4	P8		W828	P1	P2					
W769	P4	P7		W829	P1	P2					
W770	P4	P8		W830	P1	P2					
W771	P4	P7		W831	P1	P2					
W772	P4	P8		W832	P1	P2					
W773	P4	P7		W833	P1	P2					
W774	P4	P8		W834	P1	P2					
W775	P4	P7		W835	P1	P2					
W776	P4	P8		W836	P1	P2					
W777	P4	P7		W837	P1	P2					
W778	P11	P8		W838	P1	P2					
W779	P11	P7		W839	P1	P2					
W780	P11	P8		W840	P2	P3	P7				

Table 7.29 Bodyshell locator-part relationship table

Locator	Assemblies			Locator	Assemblies			Locator	Assemblies		
L1	P2	P1	P3	L33	P9	P7		L65	P8	P9	
L2	P12	P7		L34	P1	P7		L66	P8		
L3	P11	P7		L35	P7	P9		L67	P12	P8	
L4	P7	P2		L36	P6	P7		L68	P6	P8	
L5	P6	P7		L37	P2	P1		L69	P10	P8	
L6	P7	P7		L38	P3	P7		L70	P8	P12	
L7	P10	P7		L39	P7	P4		L71	P8	P11	
L8	P7			L40	P7	P3		L72	P8	P9	
L9	P11	P7		L41	P7	P2		L73	P4	P5	P8
L10	P7	P6		L42	P7	P4		L74	P9	P8	
L11	P7	P11		L43	P7			L75	P10	P8	
L12	P4	P7		L44	P4	P7		L76	P12	P8	
L13	P7	P4		L45	P2	P7	P3	L77	P8		
L14	P2	P1		L46	P7	P3		L78	P8	P10	
L15	P7			L47	P4	P7		L79	P8	P12	
L16	P7	P4		L48	P2	P1	P3	L80	P9	P8	
L17	P7	P4		L49	P12	P8		L81	P1	P8	
L18	P7	P9		L50	P11	P8		L82	P8	P9	
L19	P7			L51	P8	P2		L83	P6	P8	
L20	P12	P7		L52	P6	P8		L84	P2	P1	
L21	P6	P7		L53	P9	P8		L85	P3	P8	
L22	P10	P7		L54	P10	P8		L86	P8	P4	
L23	P7	P12		L55	P8			L87	P8	P3	
L24	P7	P11		L56	P11	P8		L88	P8	P2	
L25	P7	P9		L57	P8	P6		L89	P8	P4	
L26	P4	P5	P7	L58	P11	P8		L90	P8		
L27	P9	P7		L59	P4	P8		L91	P4	P8	
L28	P10	P7		L60	P8	P4		L92	P2	P8	P3
L29	P12	P7		L61	P2	P1		L93	P83		
L30	P7			L62	P8			L94	P4	P8	
L31	P7	P10		L63	P8	P4					
L32	P7	P12		L64	P8	P4					

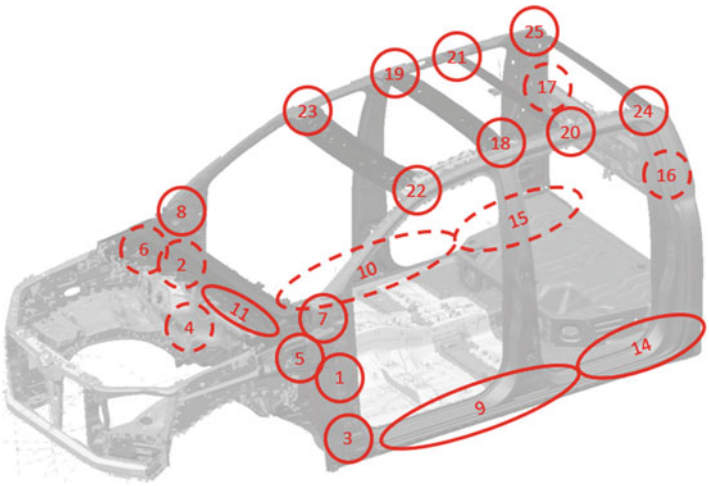


Fig. 7.39 Bodyshell weld joints grouping

Table 7.30 Bodyshell weld joints

Joint	Assemblies		
1	P1	P3	P8
2	P1	P3	P7
3	P3	P8	
4	P3	P7	
5	P1	P8	
6	P1	P7	
7	P2	P8	
8	P2	P7	
9	P4	P8	
10	P4	P7	
11	P2	P3	
12	P2	P3	P8
13	P2	P3	P7
14	P5	P8	
15	P5	P7	
16	P6	P8	
17	P6	P7	
18	P11	P8	
19	P11	P7	
20	P10	P8	
21	P10	P7	
22	P12	P8	
23	P12	P7	
24	P8	P9	
25	P7	P9	
26	P1	P2	

Table 7.31 Bodyshell joints and all welds in each joint

Joints																									
#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24	#25	#26
W1	W2	W7	W8	W23	W24	W39	W40	W55	W56	W63	W114	W115	W306	W307	W328	W329	W348	W349	W356	W357	W360	W361	W374	W375	W386
W3	W4	W9	W10	W25	W26	W41	W42	W57	W58	W64	W116	W117	W308	W309	W330	W331	W350	W351	W358	W359	W362	W363	W376	W377	W387
W5	W6	W11	W12	W27	W28	W43	W44	W59	W60	W65	W517	W518	W310	W311	W332	W333	W352	W353	W759	W760	W364	W365	W378	W379	W388
W37	W38	W13	W14	W29	W30	W45	W46	W61	W62	W66	W519	W520	W312	W313	W334	W335	W354	W355	W761	W762	W366	W367	W380	W381	W389
W404	W405	W15	W16	W31	W32	W47	W48	W134	W135	W67	W458	W459	W314	W315	W336	W337	W751	W752	W848	W850	W368	W369	W382	W383	W390
W406	W407	W17	W18	W33	W34	W49	W50	W136	W137	W68	W460	W461	W709	W710	W338	W339	W753	W754			W370	W371	W384	W385	W391
W408	W409	W19	W20	W35	W36	W51	W52	W138	W139	W69	W462	W463	W711	W712	W340	W341	W755	W756			W372	W373	W777	W778	W392
W440	W441	W21	W22	W80	W78	W53	W54	W140	W141	W70	W464	W465	W713	W714	W342	W343	W757	W758			W763	W764	W779	W780	W393
		W410	W411	W81	W79	W82	W83	W142	W143	W71	W537	W538	W715	W716	W344	W345					W765	W766	W781	W782	W394
		W412	W413	W426	W427	W118	W119	W144	W145	W72	W539	W540	W717	W718	W346	W347					W767	W768	W783	W784	W395
		W414	W415	W428	W429	W120	W121	W146	W147	W73	W541	W542	W466	W844	W731	W732					W769	W770	W785	W786	W396
		W416	W417	W430	W431	W122	W123	W148	W149	W74	W543	W544	W467	W841	W733	W734					W771	W772	W787	W788	W397
		W418	W419	W432	W433	W124	W125	W150	W151	W75	W545	W546	W468		W735	W736					W773	W774			W398
		W420	W421	W434	W435	W326	W327	W152	W153	W76	W547	W548	W469		W737	W738					W775	W776			W399
		W422	W423	W436	W437	W442	W443	W154	W155	W77	W549	W550	W470		W739	W740									W400
		W424	W425	W438	W439	W444	W445	W156	W157	W84	W551	W552	W471		W741	W742									W401
				W483	W481	W446	W447	W158	W159	W85	W553	W554	W472		W743	W744									W402
				W484	W482	W448	W449	W160	W161	W86	W555	W556	W473		W745	W746									W403
						W450	W451	W162	W163	W87	W557	W558	W474		W747	W748									W789
						W452	W453	W164	W165	W88	W559	W560	W475		W749	W750									W790
						W454	W455	W166	W167	W89	W561	W562	W476												W791
						W456	W457	W168	W169	W90	W563	W564	W477												W792
						W485	W486	W170	W171	W91	W565	W566	W478												W793
						W521	W522	W172	W173	W92	W567	W568	W479												W794
						W523	W524	W174	W175	W93	W569	W570	W480												W795
						W525	W526	W176	W177	W94	W571	W572	W487												W796
						W527	W528	W178	W179	W95	W573	W574	W488												W797
						W729	W730	W180	W181	W96	W575	W576	W489												W798
								W182	W183	W97	W577	W578	W490												W799
								W184	W185	W98	W579	W50	W491												W800
								W186	W187	W99	W581	W582	W492												W801
								W188	W189	W100	W583	W584	W493												W802
								W190	W191	W101	W585	W586	W494												W803
								W192	W193	W102	W587	W588	W495												W804
								W194	W195	W103	W589	W590	W496												W805
								W196	W197	W104	W591	W592	W497												W806
								W198	W199	W105	W593	W594	W498												
								W200	W201	W106	W595	W596	W499												
								W202	W203	W107	W597	W598	W500												

(continued)

Table 7.31 (continued)

Joints																								
#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24	#25	#26
								W204	W205	W108	W599	W600	W501												
								W206	W207	W109	W601	W602	W502												
								W208	W209	W110	W603	W604	W503												
								W210	W211	W111	W605	W606	W504												
								W212	W213	W112	W607	W608	W505												
								W214	W215	W113	W609	W609	W506												
								W216	W217	W126	W611	W612	W507												
								W218	W219	W127	W613	W614	W508												
								W220	W221	W128	W615	W616	W509												
								W222	W223	W129	W617	W618	W510												
								W224	W225	W130	W619	W620	W511												
								W226	W227	W131	W621	W622	W512												
								W228	W229	W132	W623	W624	W513												
								W230	W231	W133	W625	W626	W514												
								W232	W233		W627	W628	W515												
								W234	W235		W629	W630	W516												
								W236	W237		W631	W632	W529												
								W238	W239		W633	W634	W530												
								W240	W241		W635	W636	W531												
								W242	W243		W637	W638	W532												
								W244	W245		W639	W640	W533												
								W246	W247		W641	W642	W534												
								W248	W249		W643	W644	W535												
								W250	W251		W645	W646	W536												
								W252	W253		W647	W648													
								W254	W255		W649	W650													
								W256	W257		W651	W652													
								W258	W259		W653	W654													
								W260	W261		W655	W656													
								W262	W263		W657	W658													
								W264	W265		W659	W660													
								W266	W267		W661	W662													
								W268	W269		W663	W664													
								W270	W271		W665	W666													
								W272	W273		W667	W668													
								W274	W275		W669	W670													
								W276	W277		W671	W672													
								W278	W279		W673	W674													
								W280	W281		W675	W676													
								W282	W283		W677	W678													
								W284	W285		W679	W680													

(continued)

Table 7.31 (continued)

Joints																								
#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24	#25	#26
								W286	W287		W681	W682													
								W288	W289		W683	W684													
								W290	W291		W685	W686													
								W292	W293		W687	W688													
								W294	W295		W689	W690													
								W296	W297		W691	W692													
								W298	W299		W693	W694													
								W300	W301		W695	W696													
								W302	W303		W697	W698													
								W304	W305		W699	W700													
								W316	W317		W701	W702													
								W318	W319		W703	W704													
								W320	W321		W705	W706													
								W322	W323		W707	W708													
								W324	W325		W719	W720													
											W721	W722													
											W723	W724													
											W725	W726													
											W727	W728													

Table 7.32 List of body framing key locators with associated welds to each – part 1/5

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12		L13		L14	L16		L17	L18	L20
W107	W371	W351	W327	W329	W383	W357	W349	W333	W351	W211	W207	W325	W303	W387	W137	W299	W135	W325	W385	W371
W109	W373	W349	W83	W331	W379	W359	W351	W329	W349	W289	W233	W323	W305	W401	W303	W153	W305	W323	W379	W369
W105	W367	W355	W119	W333	W377		W353	W339	W355	W213	W209	W321	W241	W403	W305	W62	W303	W145	W377	W373
W401	W369	W353	W40	W335	W385		W355	W331	W353	W287	W275	W319	W135	W395	W209	W60	W293	W321	W375	W363
W387	W363		W52	W337	W381			W341		W215	W137	W279	W271	W389	W135	W58	W255	W143	W383	W367
W103	W361		W121	W339	W375			W343		W285	W303	W277	W293	W399	W207	W301	W137	W319	W381	W365
W63	W365		W123	W341				W337		W189	W305	W317	W269	W397	W205	W56	W271	W139		W361
W115			W50	W343				W335		W235	W241	W235	W255	W391	W275	W169	W257	W317		
W117			W125	W347				W347		W191	W135	W281	W291	W393	W241	W173	W241	W141		
W395			W44	W345				W345		W277	W271	W145	W243	W386	W233	W163	W259	W62		
W101			W115							W283	W269	W283	W257	W402	W187	W175	W269	W60		
W389			W117							W217	W293	W143	W259	W400	W273	W171	W261	W58		
W403			W48							W279	W255	W62	W245	W394	W293	W165	W291	W56		
W399			W46							W193		W285		W398	W231		W243			
W397			W54							W281		W139		W396	W271		W275			
W64			W42							W195		W60		W390	W203		W263			
W85										W219		W287		W388	W229		W245			
W391										W197		W191		W392	W201		W239			
W393										W221		W141			W227		W247			
W87										W145		W58			W255		W209			
W2										W199		W289			W225		W207			
W89										W143		W56			W269		W249			
W4										W325		W193			W199		W265			
W65										W223		W211			W223		W251			
W91										W139		W195			W257		W205			
W6										W323		W213			W197		W267			
W93										W201		W215			W221		W233			
W38										W225		W197			W291		W237			
W95										W321		W189			W259		W273			
W66										W203		W199			W195		W231			
W112										W319		W217			W219		W147			
W126										W141		W201			W243		W187			
W67										W227		W219			W261		W253			
W127										W187		W221			W217		W229			
W131										W317		W203			W245		W203			

(continued)

Table 7.32 (continued)

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L16	L17	L18	L20
W130										W229	W223		W193	W149			
W68										W56	W187		W263	W227			
W97										W58	W225		W189	W159			
W69										W60	W227		W247	W157			
W111										W231	W205		W191	W201			
W70										W62	W229		W239	W225			
W99										W205	W207		W215	W155			
W71										W273	W231		W249	W181			
W98										W301	W209		W213	W179			
W110										W165	W273		W265	W223			
W72										W171	W233		W251	W199			
W96										W175	W275		W235	W177			
W73										W163	W137		W237	W151			
W113										W183	W247		W267	W183			
W74										W157	W263		W211	W221			
W132										W155	W249		W277	W197			
W133										W181	W239		W253	W219			
W75										W161	W251		W279	W195			
W128										W179	W265		W289	W161			
W129										W177	W237		W147	W217			
W94										W151	W267		W287	W295			
W92										W167	W253		W149	W193			
W90										W185	W147		W159	W167			
W76										W153	W149		W285	W189			
W88										W295	W159		W157	W297			
W86										W169	W157		W155	W191			
W84										W297	W183		W283	W185			
W77										W299	W155		W181	W215			
W100										W173	W181		W325	W299			
W102										W239	W179		W183	W213			
W104										W237	W177		W179	W153			
W106										W265	W151		W281	W301			
W108										W253	W161		W177	W169			
W386										W267	W167		W151	W235			
W400										W149	W295		W323	W211			
W402										W147	W185		W145	W173			
W394										W159	W297		W321	W277			
W398										W243	W153		W161	W163			
W396										W257	W299		W143	W175			
W114										W245	W169		W319	W171			
W116										W259	W301		W139	W289			
W390										W247	W165		W167	W165			
W388										W261	W171		W295	W279			
W1										W249	W175		W317	W287			
W3										W263	W163		W297	W285			
W392										W251	W173		W185	W283			
W5										W291	W261		W141	W281			
W37																	

Table 7.33 List of body framing key locators with associated welds to each – part 2/5

L22	L23	L24	L25	L26	L27	L28	L29	L31	L32	L33	L34	L35	L36	L37	L38	L39	L40	L41
W359	W363	W355	W375	W309	W383	W357	W367	W359	W363	W381	W28	W375	W335	W387	W12	W177	W18	W121
W357	W369	W351	W381	W307	W379	W359	W373	W357	W365	W377	W30	W381	W331	W401	W115	W179	W16	W123
	W365	W353	W377	W311	W377		W361		W369	W375	W26	W377	W329	W395	W2	W151	W20	W40
	W371	W349	W379	W315	W385		W371		W371	W379	W34	W379	W341	W403	W4	W181	W22	W119
	W361		W385	W313	W381		W365		W361	W383	W24	W385	W339	W399	W117	W155	W14	W83
	W373		W383	W173	W375		W363		W373	W385	W32	W383	W333	W389	W6	W295	W10	W125
	W367			W301			W369		W367		W36		W337	W397	W38	W157	W8	W115
				W163							W2		W343	W391	W8	W297	W38	W327
				W299							W78		W347	W393	W10	W159	W6	W52
				W297							W4		W345	W386	W14	W299	W4	W117
				W175							W79			W402	W16	W301	W2	W50
				W295							W6			W400	W18	W147	W12	W44
				W171							W38			W398	W20	W267	W115	W48
				W165							W369			W394	W22	W265	W117	W46
				W179							W363			W396		W173		W54
				W177							W371			W390		W163		W42
				W151							W365			W388		W161		
				W181							W373			W392		W183		
				W169							W361					W167		
				W155							W367					W239		
				W153							W353					W175		
				W185							W349					W185		
				W157							W355					W171		
				W167							W351					W149		
				W159							W359					W153		
				W161							W357					W263		
				W183												W169		
				W147												W253		
				W267												W261		
				W149												W165		
				W265												W237		
				W253												W259		
				W239												W251		
				W237												W257		
				W263												W249		

(continued)

Table 7.33 (continued)

L22	L23	L24	L25	L26	L27	L28	L29	L31	L32	L33	L34	L35	L36	L37	L38	L39	L40	L41
				W251												W247		
				W261												W255		
				W249												W245		
				W259												W243		
				W247												W291		
				W257												W293		
				W245												W269		
				W243												W271		
				W255												W135		
				W291												W305		
				W293												W241		
				W269												W303		
				W271												W137		
				W135												W275		
				W305												W209		
				W241												W207		
				W303												W233		
				W137												W273		
				W275												W205		
				W209												W231		
				W233												W229		
				W207												W187		
				W273												W227		
				W205												W203		
				W231												W225		
				W229												W201		
				W187												W223		
				W227												W199		
				W203												W221		
				W225												W197		
				W201												W219		
				W223												W217		
				W221												W195		
				W199												W189		
				W219												W193		
				W197												W215		
				W217												W191		
				W195												W213		
				W189												W211		

(continued)

Table 7.33 (continued)

L22	L23	L24	L25	L26	L27	L28	L29	L31	L32	L33	L34	L35	L36	L37	L38	L39	L40	L41
				W193												W235		
				W215												W277		
				W191												W289		
				W213												W279		
				W211												W287		
				W235												W285		
				W289												W283		
				W277												W281		
				W279												W325		
				W287												W323		
				W285												W145		
				W283												W143		
				W281												W321		
				W325												W139		
				W145												W319		
				W323												W317		
				W143												W141		
				W321												W62		
				W139												W60		
				W319												W58		
				W317												W56		
				W141														
				W62														
				W60														
				W58														
				W56														

Table 7.34 List of body framing key locators with associated welds to each – part 3/5

L42	L44	L45	L46	L47	L48	L49	L50	L51	L52	L53	L54	L56	L57	L58	L59	L60
W305	W171	W50	W10	W291	W106	W370	W350	W326	W328	W382	W356	W348	W332	W350	W210	W324
W303	W175	W109	W14	W243	W108	W372	W348	W82	W330	W378	W358	W350	W328	W348	W288	W322
W135	W165	W44	W8	W245	W104	W366	W354	W118	W332	W376		W352	W338	W354	W212	W320
W137	W169	W121	W16	W269	W400	W368	W352	W39	W334	W384		W354	W330	W352	W286	W318
W293	W163	W52	W18	W247	W386	W362		W51	W336	W380			W340		W214	W278
W241	W153	W107	W38	W249	W102	W360		W120	W338	W374			W342		W284	W276
W271	W173	W123	W20	W255	W114	W364		W122	W340				W336		W188	W316
W255	W185	W105	W22	W257	W116			W49	W342				W334		W234	W234
W275	W167	W119	W6	W259	W394			W124	W346				W346		W190	W280
W209	W161	W125	W4	W293	W100			W43	W344				W344		W276	W144
W269	W183	W48	W2	W251	W388			W114							W282	W282
W207	W155	W117	W12	W271	W402			W116							W216	W142
W257	W181	W115	W115	W261	W398			W47							W278	W61
W259	W157	W83	W117	W263	W77			W45							W192	W284
W205	W301	W40		W237	W396			W53							W280	W138
W291	W299	W103		W239	W84			W41							W194	W59
W233	W179	W46		W265	W390										W218	W286
W261	W159	W63		W253	W392										W196	W190
W243	W297	W12		W241	W86										W220	W140
W273	W295	W101		W135	W1										W144	W57
W231	W177	W54		W267	W88										W198	W288
W245	W149	W327		W149	W3										W142	W55
W187	W147	W64		W305	W76										W324	W192
W263	W151	W42		W147	W90										W222	W210
W247	W253	W2		W303	W5										W138	W194
W229	W267	W85		W183	W92										W322	W212
W239	W265	W87		W137	W37										W200	W214
W203	W237	W4		W159	W94										W224	W196
W249	W239	W6		W157	W129										W320	W188
W227	W251	W89		W275	W128										W202	W198
W265	W263	W65		W161	W75										W318	W216
W201	W249	W91		W155	W133										W140	W200
W251	W261	W38		W181	W74										W226	W218
W225	W247	W93		W167	W132										W186	W220

(continued)

Table 7.34 List of body framing key locators with associated welds to each – part 3/5

L42	L44	L45	L46	L47	L48	L49	L50	L51	L52	L53	L54	L56	L57	L58	L59	L60
W267	W259	W8		W179	W113										W316	W202
W237	W245	W95		W185	W73										W228	W222
W199	W257	W66		W177	W96										W55	W186
W223	W243	W112		W151	W72										W57	W224
W221	W291	W10		W153	W110										W59	W226
W197	W255	W67		W233	W98										W230	W204
W147	W269	W126		W209	W71										W61	W228
W253	W293	W127		W273	W70										W204	W206
W219	W271	W14		W169	W99										W272	W230
W195	W135	W131		W207	W111										W206	W208
W149	W241	W130		W231	W69										W232	W272
W217	W305	W68		W295	W97										W208	W232
W159	W303	W16		W205	W68										W274	W274
W193	W137	W97		W229	W130										W136	W136
W189	W275	W18		W297	W131										W302	W302
W157	W233	W69		W165	W67										W304	W304
W155	W209	W20		W171	W126										W240	W240
W191	W273	W111		W175	W127										W134	W134
W215	W207	W22		W163	W112										W270	W270
W181	W231	W70		W299	W66										W268	W292
W179	W205	W99		W227	W95										W292	W268
W183	W229	W71		W173	W93										W254	W254
W177	W227	W98		W187	W65										W290	W290
W151	W187	W110		W301	W91										W242	W242
W213	W225	W72		W225	W89										W256	W256
W235	W203	W96		W203	W87										W244	W258
W211	W223	W73		W223	W85										W258	W244
W161	W201	W113		W201	W64										W246	W260
W277	W221	W74		W221	W101										W260	W246
W167	W199	W132		W199	W63										W248	W262
W289	W219	W133		W219	W103										W262	W248
W295	W197	W75		W217	W105										W250	W238
W279	W217	W128		W197	W107										W238	W250
W297	W195	W129		W189	W109										W236	W264
W185	W189	W94		W195	W387										W264	W236

(continued)

Table 7.34 List of body framing key locators with associated welds to each – part 3/5

L42	L44	L45	L46	L47	L48	L49	L50	L51	L52	L53	L54	L56	L57	L58	L59	L60
W287	W193	W92		W215	W401										W252	W266
W299	W215	W90		W193	W403										W266	W252
W153	W213	W76		W213	W395										W148	W146
W285	W191	W88		W191	W399										W146	W148
W301	W211	W86		W211	W397										W158	W158
W283	W235	W84		W235	W115										W182	W156
W169	W289	W77		W289	W117										W156	W182
W325	W277	W100		W277	W391										W154	W154
W281	W287	W102		W287	W389										W180	W180
W173	W279	W104		W279	W2										W160	W178
W163	W285	W106		W285	W4										W178	W176
W175	W283	W108		W283	W393										W176	W150
W171	W281	W114		W281	W6										W150	W160
W323	W325	W116		W325	W38										W166	W166
W165	W145			W145											W184	W294
W145	W323			W323											W152	W184
W321	W143			W143											W294	W296
W143	W321			W139											W168	W152
W319	W139			W321											W296	W298
W139	W319			W319											W298	W168
W317	W141			W141											W300	W300
W141	W317			W317											W164	W164
W62	W58			W56											W170	W170
W60	W60			W58											W174	W174
W58	W56			W60											W162	W162
W56	W62			W62											W172	W172

Table 7.35 List of body framing key locators with associated welds to each – part 4/5

L61	L63		L64		L65	L67	L68	L69	L70	L71	L72	L73		L74	L75	L76	L78	L79	L80
W386	W136	W316	W134	W286	W384	W370	W336	W358	W362	W354	W374	W308	W210	W382	W356	W366	W358	W362	W380
W400	W302	W296	W304	W284	W378	W368	W332	W356	W368	W350	W380	W306	W234	W378	W358	W372	W356	W364	W376
W402	W304	W184	W302	W282	W376	W372	W342		W364	W352	W376	W310	W288	W376		W360		W368	W374
W394	W208	W140	W292	W280	W374	W362	W328		W370	W348	W378	W314	W276	W384		W370		W370	W378
W388	W134	W298	W254	W324	W382	W366	W338		W360		W384	W312	W278	W380		W364		W360	W382
W398	W206	W152	W136	W322	W380	W364	W330		W372		W382	W172	W286	W374		W362		W372	W384
W396	W204	W61	W270	W144		W360	W340		W366			W300	W284			W368		W366	
W390	W274	W59	W256	W320			W334					W162	W282						
W392	W240	W57	W240	W142			W346					W298	W280						
W387	W232	W300	W258	W318			W344					W296	W324						
W403	W186	W55	W268	W138								W174	W144						
W401	W272	W168	W260	W316								W294	W322						
W395	W292	W172	W290	W140								W170	W142						
W399	W230	W162	W242	W61								W164	W320						
W397	W270	W174	W274	W59								W178	W138						
W391	W202	W170	W262	W57								W176	W318						
W389	W228	W164	W244	W55								W150	W316						
W393	W200		W238									W180	W140						
	W226		W246									W168	W61						
	W254		W208									W154	W59						
	W224		W206									W152	W57						
	W268		W248									W184	W55						
	W198		W264									W156							
	W222		W250									W166							
	W256		W204									W158							
	W196		W266									W160							
	W220		W232									W182							
	W290		W236									W146							
	W258		W272									W266							
	W194		W230									W148							
	W218		W146									W264							
	W242		W186									W252							
	W260		W252									W238							
	W216		W228									W236							
	W244		W202									W262							
	W192		W148									W250							

(continued)

Table 7.35 (continued)

L61	L63		L64		L65	L67	L68	L69	L70	L71	L72	L73		L74	L75	L76	L78	L79	L80
	W262		W226									W260							
	W188		W158									W248							
	W246		W156									W258							
	W190		W200									W246							
	W238		W224									W256							
	W214		W154									W244							
	W248		W180									W242							
	W212		W178									W254							
	W264		W222									W290							
	W250		W198									W292							
	W234		W176									W268							
	W236		W150									W270							
	W266		W182									W134							
	W210		W220									W304							
	W276		W196									W240							
	W252		W218									W302							
	W278		W194									W136							
	W288		W160									W274							
	W146		W216									W208							
	W286		W294									W232							
	W148		W192									W206							
	W158		W166									W272							
	W284		W188									W204							
	W156		W296									W230							
	W154		W190									W228							
	W282		W184									W186							
	W180		W214									W226							
	W324		W298									W202							
	W182		W212									W224							
	W178		W152									W200							
	W280		W300									W222							
	W176		W168									W220							
	W150		W234									W198							
	W322		W210									W218							
	W144		W172									W196							
	W320		W276									W216							
	W160		W162									W194							
	W142		W174									W188							
	W318		W170									W192							
	W138		W288									W214							
	W166		W164									W190							
	W294		W278									W212							

Table 7.36 List of body framing key locators with associated welds to each – part 5/5

L81	L82	L83	L84	L85	L86		L87	L88	L89		L91		L92		L93	L94	
W27	W374	W334	W386	W11	W176	W286	W17	W120	W304	W172	W170	W278	W49	W105	W9	W290	W278
W29	W380	W330	W400	W114	W178	W284	W15	W122	W302	W162	W174	W284	W108	W107	W13	W242	W284
W25	W376	W328	W394	W1	W150	W282	W19	W39	W134	W174	W164	W282	W43	W109	W7	W244	W282
W33	W378	W340	W402	W3	W180	W280	W21	W118	W136	W170	W168	W280	W120	W115	W15	W268	W280
W23	W384	W338	W398	W116	W154	W324	W13	W82	W292	W322	W162	W324	W51	W117	W17	W246	W324
W31	W382	W332	W388	W5	W294	W322	W9	W124	W240	W164	W152	W144	W106		W37	W248	W144
W35		W336	W396	W37	W156	W144	W7	W114	W270	W144	W172	W322	W122		W19	W254	W322
W80		W342	W390	W7	W296	W142	W37	W326	W254	W320	W184	W142	W104		W21	W256	W142
W1		W346	W392	W9	W158	W320	W5	W51	W274	W142	W166	W320	W118		W5	W258	W138
W81		W344	W387	W13	W298	W138	W3	W116	W208	W318	W160	W138	W124		W3	W292	W320
W3			W403	W15	W300	W318	W1	W49	W268	W138	W182	W318	W47		W1	W250	W318
W5			W401	W17	W146	W316	W11	W43	W206	W316	W154	W140	W116		W11	W270	W140
W37			W399	W19	W266	W140	W114	W47	W256	W140	W180	W316	W114		W114	W260	W316
W368			W395	W21	W264	W61	W116	W45	W258	W61	W156	W57	W82		W116	W262	W55
W362			W397		W172	W59		W53	W204	W59	W300	W59	W39			W236	W57
W370			W391		W162	W57		W41	W290	W57	W298	W55	W102			W238	W59
W364			W389		W160	W55			W232	W55	W178	W61	W45			W264	W61
W372			W393		W182				W260		W158		W11			W252	
W360					W166				W242		W296		W100			W240	
W366					W238				W272		W294		W53			W134	
W352					W174				W230		W176		W326			W266	
W348					W184				W244		W148		W77			W148	
W354					W170				W186		W146		W41			W304	
W350					W148				W262		W150		W1			W146	
W358					W152				W246		W252		W84			W302	
W356					W262				W228		W266		W86			W182	
					W168				W238		W264		W3			W136	
					W252				W202		W236		W5			W158	
					W260				W248		W238		W88			W156	
					W164				W226		W250		W76			W274	
					W236				W264		W262		W90			W160	
					W258				W200		W248		W37			W154	
					W250				W250		W260		W92			W180	
					W256				W224		W246		W7			W166	

(continued)

Table 7.36 (continued)

L81	L82	L83	L84	L85	L86	L87	L88	L89	L91	L92	L93	L94
					W248			W266	W258	W94		W178
					W246			W236	W244	W129		W184
					W254			W198	W256	W128		W176
					W244			W222	W242	W75		W150
					W242			W220	W290	W9		W152
					W290			W196	W254	W74		W232
					W292			W146	W268	W133		W208
					W268			W252	W292	W132		W272
					W270			W218	W270	W13		W168
					W134			W194	W134	W113		W206
					W304			W148	W240	W73		W230
					W240			W216	W304	W15		W294
					W302			W158	W302	W96		W204
					W136			W192	W136	W17		W228
					W274			W188	W274	W72		W296
					W208			W156	W232	W19		W164
					W206			W154	W208	W110		W170
					W232			W190	W272	W21		W174
					W272			W214	W206	W98		W162
					W204			W180	W230	W71		W298
					W230			W178	W204	W70		W226
					W228			W182	W228	W99		W172
					W186			W176	W226	W111		W186
					W226			W150	W186	W69		W300
					W202			W212	W224	W97		W224
					W224			W234	W202	W68		W202
					W200			W210	W222	W130		W222
					W222			W160	W200	W131		W200
					W198			W276	W220	W67		W220
					W220			W166	W198	W127		W198
					W196			W288	W218	W126		W218
					W218			W294	W196	W112		W216
					W216			W278	W216	W66		W196
					W194			W296	W194	W95		W188
					W188			W184	W188	W93		W194
					W192			W286	W192	W65		W214
					W214			W298	W214	W91		W192
					W190			W152	W212	W89		W212
					W212			W284	W190	W87		W190
					W210			W300	W210	W85		W210
					W234			W282	W234	W64		W234
					W276			W168	W288	W101		W288
					W288			W324	W276	W63		W276
					W278			W280	W286	W103		W286

Using the coordinates of these locators and welds, we calculated the distance between each locator to any of its related welds (Tables 7.37, 7.38, 7.39, 7.40, 7.41 and 7.42).

All welds were assessed and the list of two welds with maximum distance in each joint were listed (Table 7.43).

The above developed tables from framing CAD data would characterise each component of the bodyshell system and helped in identifying the relation between the components. These tables were then used as input data for GA model coding. Similar to the other two examples, the root of sum square of the distance of each weld to the locator with the same stack-up as the fitness function.

The bodyshell genetic algorithm model was developed in modeFRONTIER application using the above generated tables and codes. A schematic of the structure and the GA model is shown in Fig. 7.40.

The model can continue running until the fittest subset of welds are generated. In this study, we decided to optimise the weld process design from 850 welds to under 100 welds. The result is shown in Table 7.44.

The GA optimisation eliminated more than 99% of the non-fit solutions by narrowing down the process design from a combination of 60 welds out of 850 welds to a combination of 60 welds out of 97 welds. This optimised group of welds could set the geometry of the bodyshell with minimum process driven dimensional variation (Fig. 7.41).

Note the process design resulted from the GA model in this structure does not necessarily mean this is the minimum number of welds required to set the geometry of the bodyshell. It is prioritising the most effective welds (geometry welds) in minimising the dimensional variations.

The GA assessment is mainly done based on the input data and the fitness function we defined. Finite element analysis can be used to assess the effectiveness of the GA optimisation. Also, as explained in the prior sections, the maximum number of welds (station capacity) in our study was 60 welds. This means the 97 welds resulted of GA optimisation will need to be further assessed and optimised to 60 welds. This is explained in the next chapter.

Table 7.37 Framing key locators and all relevant weld and distance – part 1/6

Locator		L1		L2	L3	L4	L5	L6	L7	L9	L10	L11	L12			L13			L14	L16		
Related welds	Weld	W107	W133	W371	W351	W327	W329	W383	W357	W349	W333	W351	W211	W273	W153	W325	W231	W185	W387	W137	W249	W62
	Dis. (mm)	129.85	860.22	82.43	201.74	133.76	94.63	94.59	73.04	170.24	247.28	170.78	0.92	437.00	1644.91	97.38	653.67	1821.00	138.22	156.18	623.93	1057.80
	Weld	W109	W75	W373	W349	W83	W331	W379	W359	W351	W329	W349	W289	W207	W295	W323	W209	W297	W401	W303	W213	W60
	Dis. (mm)	132.48	910.78	83.22	202.73	146.94	109.45	119.19	74.75	194.97	250.08	198.89	44.35	452.69	1656.71	106.22	659.41	1849.66	181.31	178.51	648.88	1062.58
	Weld	W105	W128	W367	W355	W119	W333	W377		W353	W339	W355	W213	W233	W169	W321	W273	W153	W403	W305	W265	W58
	Dis. (mm)	139.95	919.79	110.27	230.88	148.32	120.74	134.11		199.15	261.13	199.74	59.01	462.00	1683.91	122.00	677.06	1858.39	213.40	198.21	653.95	1067.93
	Weld	W401	W129	W369	W353	W40	W335	W385		W355	W331	W353	W287	W209	W297	W319	W233	W299	W395	W209	W251	W301
	Dis. (mm)	143.24	922.26	133.54	231.86	150.98	158.63	144.05		220.60	263.18	224.42	83.55	476.74	1686.97	142.42	700.56	1882.58	214.52	236.43	655.03	1073.51
	Weld	W387	W94	W363		W52	W337	W381			W341		W215	W275	W299	W279	W275	W169	W389	W135	W235	W56
	Dis. (mm)	148.79	947.41	149.72		216.52	171.22	173.43			266.90		88.41	607.00	1719.20	143.60	838.59	1897.22	231.10	240.22	682.31	1073.84
	Weld	W103	W92	W361		W121	W339	W375			W343		W285	W137	W301	W277	W137	W301	W399	W207	W237	W169
	Dis. (mm)	160.20	991.17	156.61		229.28	172.08	189.62			270.44		121.28	678.19	1747.55	163.65	867.33	1911.51	231.12	256.74	686.50	1087.70
	Weld	W63	W90	W365		W123	W341				W337		W189	W303	W165	W317	W303	W165	W397	W205	W267	W173
	Dis. (mm)	170.39	1034.99	156.92		251.21	188.96				286.10		123.40	745.76	1751.42	165.78	935.09	1962.58	240.27	287.11	690.85	1134.34
	Weld	W115	W76			W50	W343				W335		W235	W305	W171	W235	W305	W171	W391	W275	W211	W163
	Dis. (mm)	171.69	1045.10			256.38	199.27				318.38		144.15	778.58	1761.57	193.98	968.33	1963.14	271.93	294.61	701.95	1135.98
	Weld	W117	W88			W125	W347				W347		W191	W241	W175	W281	W241	W175	W393	W241	W277	W175
	Dis. (mm)	174.32	1078.22			293.68	468.62				534.34		150.60	790.00	1764.95	234.48	1015.56	1963.40	274.18	330.06	716.97	1138.11
	Weld	W395	W86			W44	W345				W345		W277	W135	W163	W145	W135	W163	W386	W233	W253	W171
	Dis. (mm)	180.99	1121.27			304.75	491.30				555.56		154.90	835.81	1768.90	236.07	1026.32	1963.85	1448.96	341.73	741.06	1140.40
	Weld	W101	W84			W115							W283	W271	W173	W283	W271	W173	W402	W187	W279	W165
	Dis. (mm)	188.70	1163.46			317.90							156.89	881.00	1772.53	238.20	1104.29	1964.28	1459.44	346.58	741.40	1149.29
	Weld	W389	W77			W117							W217	W269		W143	W293		W400	W273	W289	
	Dis. (mm)	189.24	1171.61			343.98							158.40	978.00		240.60	1176.14		1459.77	355.25	742.44	
	Weld	W403	W100			W48							W279	W293		W62	W269		W394	W293	W147	
	Dis. (mm)	198.46	1205.79			354.89							166.81	984.72		243.68	1199.18		1462.77	368.63	743.96	
	Weld	W399	W102			W46							W193	W255		W285	W255		W398	W231	W287	
	Dis. (mm)	202.69	1248.11			397.85							168.03	1048.40		246.85	1240.25		1463.72	369.96	778.63	
	Weld	W397	W104			W54							W281	W291		W139	W291		W396	W271	W149	
	Dis. (mm)	204.43	1288.73			431.19							183.39	1055.80		247.67	1275.64		1466.82	380.63	783.76	
	Weld	W64	W106			W42							W195	W243		W60	W243		W390	W203	W159	
	Dis. (mm)	206.31	1329.37			463.52							192.03	1083.87		258.13	1303.25		1491.69	382.27	805.46	
	Weld	W85	W108										W219	W257		W287	W257		W388	W229	W285	
	Dis. (mm)	221.97	1370.05										193.40	1114.22		261.05	1306.48		1495.86	398.09	813.72	
	Weld	W391	W386										W197	W245		W191	W259		W392	W201	W157	
	Dis. (mm)	225.45	1399.59										220.43	1118.87		266.10	1332.32		1516.21	418.59	826.50	

(continued)

Table 7.37 (continued)

Locator		L1		L2	L3	L4	L5	L6	L7	L9	L10	L11	L12			L13			L14	L16		
	Weld	W393	W400										W221	W259		W141	W245			W227	W155	
	Dis. (mm)	233.67	1404.88										228.40	1139.92		270.19	1337.68			422.58	843.11	
	Weld	W87	W402										W145	W247		W58	W261			W255	W283	
	Dis. (mm)	257.98	1407.08										228.72	1153.87		274.10	1367.23			428.13	847.04	
	Weld	W2	W394										W199	W261		W289	W247			W225	W181	
	Dis. (mm)	260.06	1407.79										251.74	1174.66		280.46	1372.15			448.46	857.74	
	Weld	W89	W398										W143	W249		W56	W263			W269	W325	
	Dis. (mm)	296.59	1409.22										253.71	1188.87		291.34	1402.15			448.58	864.85	
	Weld	W4	W396										W325	W263		W193	W249			W199	W183	
	Dis. (mm)	302.55	1411.04										253.82	1209.41		303.06	1406.63			455.39	865.04	
	Weld	W65	W114										W223	W251		W211	W239			W223	W179	
	Dis. (mm)	330.99	1417.63										263.40	1223.87		307.08	1437.08			475.52	866.90	
	Weld	W91	W116										W139	W239		W195	W251			W257	W281	
	Dis. (mm)	336.59	1418.34										279.01	1244.17		340.60	1441.15			490.84	871.76	
	Weld	W6	W390										W323	W237		W213	W265			W197	W177	
	Dis. (mm)	337.66	1432.99										279.48	1258.87		348.19	1475.30			492.56	873.63	
	Weld	W93	W388										W201	W265		W215	W237			W221	W151	
	Dis. (mm)	377.64	1438.47										285.02	1282.23		370.43	1475.69			503.56	878.92	
	Weld	W38	W1										W225	W253		W197	W267			W291	W323	
	Dis. (mm)	412.17	1443.09										298.40	1318.80		378.48	1513.17			510.31	894.38	
	Weld	W95	W3										W321	W267		W189	W253			W259	W145	
	Dis. (mm)	419.37	1451.35										305.92	1320.03		398.11	1534.85			515.56	914.58	
	Weld	W66	W392										W203	W149		W199	W147			W195	W321	
	Dis. (mm)	432.99	1457.65										319.63	1365.00		416.60	1568.29			530.02	923.95	
	Weld	W112	W5										W319	W147		W217	W149			W219	W161	
	Dis. (mm)	455.65	1459.07										332.97	1375.97		426.86	1580.54			532.42	936.65	
	Weld	W126	W37										W141	W159		W201	W159			W243	W143	
	Dis. (mm)	473.73	1478.09										333.01	1445.70		454.90	1634.05			533.56	938.31	
	Weld	W67	W99										W227	W183		W219	W157			W261	W319	
	Dis. (mm)	474.28	654.61										333.40	1452.00		456.49	1657.15			549.13	953.54	
	Weld	W127	W71										W187	W157		W221	W183			W217	W139	
	Dis. (mm)	476.03	696.48										355.21	1471.52		486.83	1666.61			561.97	962.38	
	Weld	W131	W98										W317	W155		W203	W155			W245	W167	
	Dis. (mm)	536.47	708.05										360.48	1493.49		493.35	1675.81			563.12	974.08	
	Weld	W130	W110										W229	W181		W223	W181			W193	W295	
	Dis. (mm)	539.66	739.75										368.40	1514.99		517.76	1692.66			567.71	980.74	
	Weld	W68	W72										W56	W161		W187	W179			W263	W317	
	Dis. (mm)	555.17	754.68										403.39	1527.70		531.91	1703.51			582.87	983.16	
	Weld	W97	W96										W58	W179		W225	W177			W189	W297	
	Dis. (mm)	575.48	787.45										405.80	1531.00		549.18	1711.51			592.13	1011.66	
	Weld	W69	W73										W60	W177		W227	W151			W247	W185	
	Dis. (mm)	604.79	805.86										409.73	1546.61		581.00	1716.88			593.28	1013.52	
	Weld	W111	W113										W231	W151		W205	W161			W191	W141	
	Dis. (mm)	622.84	824.28										412.00	1557.53		598.69	1741.60			605.50	1013.97	
	Weld	W70	W74										W62	W167		W229	W167			W239	W299	
	Dis. (mm)	654.25	855.97										415.12	1566.80		613.18	1780.45			616.75	1044.58	
	Weld	W132											W205	W185		W207	W295			W215	W153	
	Dis. (mm)	857.26											418.17	1607.41		634.78	1818.74			622.79	1049.91	

Table 7.38 Framing key locators and all relevant weld and distance – part 2/6

Locator		L17			L18	L20	L21	L22	L23	L24	L25	L26			L27	L28	L29	L31	L32	L33	L34	L35
Related Welds	Weld	W135	W181	W145	W385	W371	W337	W359	W363	W355	W375	W309	W243	W285	W383	W357	W367	W359	W363	W381	W28	W375
	Dis. (mm)	154.23	646.12	1117.54	196.19	97.67	89.51	88.07	165.12	74.22	57.22	6.06	742.38	1915.49	94.59	88.64	38.78	152.99	157.70	116.68	20.17	198.51
	Weld	W305	W179	W321	W379	W369	W333	W357	W369	W351	W381	W307	W255	W283	W379	W359	W373	W357	W365	W377	W30	W381
	Dis. (mm)	173.34	654.92	1133.53	219.53	117.38	113.20	98.08	173.34	99.95	81.89	21.00	747.33	1950.75	119.19	101.13	61.66	188.19	176.50	129.86	25.36	217.19
	Weld	W303	W223	W143	W377	W373	W343		W365	W353	W377	W311	W291	W281	W377		W361		W369	W375	W26	W377
	Dis. (mm)	191.79	655.67	1141.69	229.24	136.87	134.01		198.45	129.92	108.24	33.60	768.56	1976.91	134.11		97.54		192.33	141.35	70.83	221.67
	Weld	W293	W199	W319	W375	W363	W329		W371	W349	W379	W315	W293	W325	W385		W371		W371	W379	W34	W379
	Dis. (mm)	195.33	659.21	1163.26	238.55	166.72	171.58		220.31	145.87	139.03	45.95	810.75	1992.38	144.05		109.58		223.12	149.93	125.11	231.79
	Weld	W255	W177	W139	W383	W367	W339		W361		W385	W313	W269	W145	W381		W365		W361	W383	W24	W385
	Dis. (mm)	241.45	661.91	1166.16	242.63	177.19	185.55		245.96		180.21	58.67	841.68	2021.81	173.43		124.72		225.17	174.05	140.84	245.71
	Weld	W137	W151	W317	W381	W365	W331		W373		W383	W173	W271	W323	W375		W363		W373	W385	W32	W383
	Dis. (mm)	241.48	667.94	1193.00	259.64	195.27	215.60		249.87		199.71	107.27	934.16	2022.32	189.62		145.07		239.47	205.94	157.55	274.58
	Weld	W271	W183	W141		W361	W341		W367			W301	W135	W143			W369		W367		W36	
	Dis. (mm)	294.61	669.40	1218.51		211.23	243.75		284.40			115.73	959.84	2046.59			163.10		270.00		180.47	
	Weld	W257	W221	W62			W335					W163	W305	W321							W2	
	Dis. (mm)	296.27	687.16	1266.90			276.73					132.40	1017.48	2052.25							210.33	
	Weld	W241	W197	VV60			W347					W299	W241	W139							W78	
	Dis. (mm)	299.05	697.43	1270.90			595.57					134.40	1021.59	2071.67							212.25	
	Weld	W259	W219	W58			W345					W297	W303	W319							W4	
	Dis. (mm)	318.85	718.97	1275.38			619.06					159.44	1050.61	2082.19							221.54	
	Weld	W269	W195	W56								W175	W137	W317							W79	
	Dis. (mm)	319.39	735.76	1280.33								160.27	1118.82	2112.13							230.87	
	Weld	W261	W161									W295	W275	W141							W6	
	Dis. (mm)	350.10	738.26									185.25	1199.02	2125.22							238.25	
	Weld	W291	W217									W171	W209	W62							W38	
	Dis. (mm)	358.07	751.07									185.35	1326.17	2184.01							298.53	
	Weld	W243	W295									W165	W233	VV60							W369	
	Dis. (mm)	374.95	769.50									263.97	1340.66	2186.09							1199.40	
	Weld	W275	W193									W179	W207	W58							W363	
	Dis. (mm)	380.63	774.17									270.42	1351.13	2188.44							1239.70	
	Weld	W263	W167									W177	W273	W56							W371	
	Dis. (mm)	382.00	774.57									271.23	1365.15	2191.08							1244.64	
	Weld	W245	W189									W151	W205								W365	
	Dis. (mm)	397.70	783.42									278.61	1387.47								1287.85	
	Weld	W239	W297									W181	W231								W373	
	Dis. (mm)	414.40	800.36									281.29	1389.66								1299.05	

(continued)

Table 7.38 Framing key locators and all relevant weld and distance – part 2/6

Locator		L17			L18	L20	L21	L22	L23	L24	L25	L26			L27	L28	L29	L31	L32	L33	L34	L35
	Weld	W247	W191									W169	W229								W361	
	Dis. (mm)	422.13	812.56									291.22	1432.45								1340.61	
	Weld	W209	W185									W155	W187								W367	
	Dis. (mm)	422.42	813.11									302.07	1454.89								1351.46	
	Weld	W207	W215									W153	W227								W353	
	Dis. (mm)	446.15	815.99									304.58	1466.83								1693.03	
	Weld	W249	W299									W185	W203								W349	
	Dis. (mm)	447.96	833.23									320.97	1493.94								1697.62	
	Weld	W265	W213									W157	W225								W355	
	Dis. (mm)	450.30	843.50									325.34	1501.24								1759.88	
	Weld	W251	W153									W167	W201								W351	
	Dis. (mm)	474.97	848.72									342.14	1533.00								1764.16	
	Weld	W205	W301									W159	W223								W359	
	Dis. (mm)	480.67	862.12									352.88	1535.68								2027.50	
	Weld	W267	W169									W161	W221								W357	
	Dis. (mm)	486.10	885.76									366.73	1570.14								2065.38	
	Weld	W233	W235									W183	W199									
	Dis. (mm)	486.16	890.31									421.44	1572.06									
	Weld	W237	W211									W147	W219									
	Dis. (mm)	502.96	898.99									424.86	1604.63									
	Weld	W273	W173									W267	W197									
	Dis. (mm)	506.33	924.03									480.02	1611.13									
	Weld	W231	W277									W149	W217									
	Dis. (mm)	526.91	925.32									491.36	1639.14									
	Weld	W147	W163									W265	W195									
	Dis. (mm)	537.52	926.38									516.84	1650.21									
	Weld	W187	W175									W253	W189									
	Dis. (mm)	545.38	929.44									530.47	1673.67									
	Weld	W253	W171									W239	W193									
	Dis. (mm)	552.68	932.70									554.01	1689.30									
	Weld	W229	W289									W237	W215									
	Dis. (mm)	563.68	941.00									582.90	1708.22									
	Weld	W203	W165									W263	W191									
	Dis. (mm)	583.16	945.37									588.12	1728.29									
	Weld	W149	W279									W251	W213									
	Dis. (mm)	592.43	949.97									614.12	1737.25									
	Weld	W227	W287									W261	W211									
	Dis. (mm)	593.86	978.34									622.32	1795.52									
	Weld	W159	W285									W249	W235									
	Dis. (mm)	596.47	1014.40									645.73	1807.08									
	Weld	W157	W283									W259	W289									
	Dis. (mm)	616.51	1048.55									656.60	1839.35									
	Weld	W201	W281									W247	W277									
	Dis. (mm)	621.11	1073.86									677.67	1842.49									
	Weld	W225	W325									W257	W279									
	Dis. (mm)	624.54	1074.12									682.02	1867.41									
	Weld	W155	W323									W245	W287									
	Dis. (mm)	632.25	1103.82									709.90	1878.15									

Table 7.39 Framing key locators and all relevant weld and distance – part 3/6

Locator		L36	L37	L38	L39			L40	L41	L42			L44			L45		L46	L47		
Related Welds	Weld	W335	W387	W12	W177	W271	W145	W18	W121	W305	W219	W145	W171	W271	W323	W50	W14	W10	W291	W169	W323
	Dis. (mm)	46.54	61.81	804.59	15.21	703.46	1766.21	164.78	191.01	156.03	639.98	1033.61	43.93	843.44	1966.45	105.98	527.38	3.89	2.42	626.95	1307.20
	Weld	W331	W401	W115	W179	W135	W143	W16	W123	W303	W195	W321	W175	W135	W143	W109	W131	W14	W243	W207	W143
	Dis. (mm)	85.39	78.97	115.84	34.46	707.25	1790.81	165.15	198.87	158.98	651.82	1048.32	61.83	901.30	1976.72	110.48	557.35	44.23	26.96	642.55	1310.70
	Weld	W329	W395	W2	W151	W305	W321	W20	W40	W135	W149	W143	W165	W241	W321	W44	W130	W8	W245	W231	W139
	Dis. (mm)	127.49	109.11	116.33	50.96	763.64	1790.88	169.29	203.12	168.16	668.94	1057.62	63.38	934.20	1996.42	111.28	559.23	59.75	61.91	645.00	1336.00
	Weld	W341	W403	W4	W181	W241	W139	W22	W119	W137	W217	W319	W169	W305	W139	W121	W68	W16	W269	W295	W321
	Dis. (mm)	162.08	113.00	139.26	69.77	785.76	1815.71	178.16	204.61	186.41	671.32	1077.98	79.51	959.34	2002.00	114.56	572.88	93.95	79.02	667.08	1337.04
	Weld	W339	W399	W117	W155	W303	W319	W14	W83	W293	W159	W139	W163	W303	W319	W52	W16	W18	W247	W205	W319
	Dis. (mm)	196.77	125.94	143.12	104.37	796.18	1820.73	179.28	222.27	262.37	683.35	1081.92	85.96	992.63	2026.40	115.40	576.36	118.74	96.89	678.22	1366.89
	Weld	W333	W389	W6	W295	W137	W317	W10	W125	W241	W193	W317	W153	W137	W141	W107	W97	W38	W249	W229	W141
	Dis. (mm)	202.91	134.09	163.17	117.83	863.54	1850.59	209.96	229.02	293.08	689.97	1107.65	103.19	1060.66	2055.97	115.77	592.88	123.91	131.89	688.60	1390.00
	Weld	W337	W397	W38	W157	W275	W141	W8	W115	W271	W189	W141	W173	W275	W317	W123	W18	W20	W255	W297	W317
	Dis. (mm)	264.52	134.87	213.85	131.97	956.07	1868.91	254.14	251.20	314.78	703.01	1134.02	109.16	1116.84	2056.38	123.52	601.07	143.42	141.69	693.27	1396.74
	Weld	W343	W391	W8	W297	W209	W62	W38	W327	W255	W157	W62	W185	W233	W58	W105	W69	W22	W257	W165	W56
	Dis. (mm)	275.52	171.60	239.63	146.56	1068.22	1923.50	311.06	253.36	316.22	704.10	1181.13	132.68	1261.63	2124.53	134.16	622.67	168.15	148.55	694.44	1459.73
	Weld	W347	W393	W10	W159	W207	W60	W6	W52	W275	W155	W60	W167	W209	W60	W119	W20	W6	W259	W171	W58
	Dis. (mm)	337.03	175.42	275.31	158.73	1093.02	1926.45	360.40	254.64	330.91	720.48	1185.25	168.15	1269.48	2124.54	136.16	625.88	193.69	158.86	707.54	1460.38
	Weld	W345	W386	W14	W299	W233	W58	W4	W117	W209	W191	W58	W161	W273	W56	W125	W111	W4	W293	W175	W60
	Dis. (mm)	359.77	1443.67	312.98	177.98	1093.91	1929.71	380.07	283.02	344.33	728.12	1189.88	204.57	1286.60	2124.81	136.17	640.46	224.47	162.93	712.93	1461.46
	Weld		W402	W16	W301	W273	W56	W2	W50	W269	W215	W56	W183	W207	W62	W48	W22	W2	W251	W163	W62
	Dis. (mm)		1448.17	358.73	206.00	1117.86	1933.30	405.48	288.79	362.20	735.00	1195.02	277.46	1294.38	2124.85	138.04	650.85	263.12	166.88	719.68	1462.96
	Weld		W400	W18	W147	W205		W12	W44	W207	W181		W155	W231		W117	W70	W12	W271	W299	
	Dis. (mm)		1450.62	382.45	219.13	1129.07		428.63	328.81	367.18	735.02		288.48	1311.56		142.53	672.30	298.20	176.01	721.56	
	Weld		W398	W20	W267	W231		W115	W48	W257	W179		W181	W205		W115	W99	W115	W261	W227	
	Dis. (mm)		1450.84	406.63	262.57	1141.85		466.84	374.87	375.54	744.33		292.99	1330.82		143.10	672.34	346.20	177.90	723.60	
	Weld		W394	W22	W265	W229		W117	W46	W259	W183		W157	W229		W83	W71	W117	W263	W173	
	Dis. (mm)		1451.06	431.23	292.38	1183.80		483.04	415.06	399.28	748.20		294.13	1355.11		146.19	714.65	350.74	201.27	726.34	
	Weld		W396		W173	W187			W54	W205	W177		W301	W227		W40	W98		W237	W187	
	Dis. (mm)		1453.28		298.15	1196.00			446.60	400.58	751.70		302.39	1390.08		146.84	725.96		201.88	744.58	
	Weld		W390		W163	W227			W42	W291	W151		W299	W187		W103	W110		W239	W301	
	Dis. (mm)		1476.70		310.53	1217.56			477.44	413.43	757.68		305.01	1398.41		161.03	757.77		227.64	746.74	
	Weld		W388		W161	W203				W233	W213		W179	W225		W46	W72		W265	W225	
	Dis. (mm)		1483.98		317.37	1234.79				419.21	762.07		308.28	1425.04		169.66	773.00		258.86	758.60	
	Weld		W392		W183	W225				W261	W235		W159	W203		W63	W96		W253	W203	
	Dis. (mm)		1501.49		321.80	1251.39				431.76	805.57		308.29	1437.55		174.24	805.59		261.81	783.11	

(continued)

Table 7.39 (continued)

Locator		L36	L37	L38	L39		L40	L41	L42		L44		L45		L46	L47	
	Weld				W167	W201			W243	W211	W297	W223	W12	W73		W241	W223
	Dis. (mm)				322.73	1273.61			433.80	816.81	311.28	1460.01	193.34	824.28		267.01	793.60
	Weld				W239	W223			W273	W161	W295	W201	W101	W113		W135	W201
	Dis. (mm)				323.92	1285.28			437.51	818.48	320.16	1476.69	194.36	842.49		271.31	821.70
	Weld				W175	W199			W231	W277	W177	W221	W54	W74		W267	W221
	Dis. (mm)				325.70	1312.45			456.43	840.40	341.34	1494.97	197.19	874.49		291.71	828.60
	Weld				W185	W221			W245	W167	W149	W199	W327	W132		W149	W199
	Dis. (mm)				334.55	1319.23			460.27	855.35	362.69	1515.84	201.58	877.08		308.01	860.35
	Weld				W171	W197			W187	W289	W147	W219	W64	W133		W305	W219
	Dis. (mm)				340.56	1351.31			463.76	858.35	364.19	1529.94	214.04	880.87		322.89	863.60
	Weld				W149	W219			W263	W295	W151	W197	W42	W75		W147	W217
	Dis. (mm)				348.00	1353.23			464.61	859.38	368.28	1554.98	225.41	929.25		344.08	898.60
	Weld				W153	W217			W247	W279	W253	W217	W2	W128		W303	W197
	Dis. (mm)				348.97	1387.28			487.82	864.95	408.23	1564.91	230.02	939.59		353.31	899.05
	Weld				W263	W195			W229	W297	W267	W195	W85	W129		W183	W189
	Dis. (mm)				353.91	1390.20			490.75	890.16	417.18	1594.14	230.88	942.87		395.01	933.60
	Weld				W169	W189			W239	W185	W265	W189	W87	W94		W137	W195
	Dis. (mm)				366.98	1421.38			497.75	894.30	454.44	1599.88	269.08	965.68		416.20	937.79
	Weld				W253	W193			W203	W287	W237	W193	W4	W92		W159	W215
	Dis. (mm)				369.44	1429.09			500.89	895.33	467.54	1633.29	272.60	1009.41		416.37	968.60
	Weld				W261	W215			W249	W299	W239	W215	W6	W90		W157	W193
	Dis. (mm)				384.75	1455.52			516.27	922.94	492.20	1634.86	307.90	1053.18		445.64	976.56
	Weld				W165	W191			W227	W153	W251	W213	W89	W76		W275	W213
	Dis. (mm)				392.68	1467.91			519.31	930.34	502.24	1664.23	309.27	1063.18		450.00	998.00
	Weld				W237	W213			W265	W285	W263	W191	W65	W88		W161	W191
	Dis. (mm)				403.72	1484.23			534.28	931.08	526.79	1672.34	344.98	1096.40		470.70	1015.27
	Weld				W259	W211			W201	W301	W249	W211	W91	W86		W155	W211
	Dis. (mm)				416.24	1541.88			538.31	951.76	536.97	1723.18	350.56	1139.47		472.67	1057.00
	Weld				W251	W235			W251	W283	W261	W235	W38	W84		W181	W235
	Dis. (mm)				426.32	1546.38			545.48	964.97	561.42	1751.26	383.37	1181.66		502.37	1093.56
	Weld				W257	W277			W225	W169	W247	W289	W93	W77		W167	W289
	Dis. (mm)				439.89	1581.66			548.62	967.75	571.74	1767.49	392.70	1189.64		509.81	1101.34
	Weld				W249	W289			W267	W325	W259	W277	W8	W100		W179	W277
	Dis. (mm)				450.50	1585.28			570.61	989.07	596.09	1786.73	434.19	1223.97		528.59	1128.78
	Weld				W247	W279			W237	W281	W245	W287	W95	W102		W185	W287
	Dis. (mm)				476.03	1606.48			575.34	990.09	606.54	1806.68	435.24	1266.27		550.43	1140.55
	Weld				W255	W287			W199	W173	W257	W279	W66	W104		W177	W279
	Dis. (mm)				501.54	1623.72			575.96	1011.25	621.78	1811.71	449.66	1306.86		561.12	1153.61
	Weld				W245	W285			W223	W163	W243	W285	W112	W106		W151	W285
	Dis. (mm)				502.70	1660.74			578.56	1013.07	641.35	1844.38	472.04	1347.45		585.13	1178.28
	Weld				W243	W283			W221	W175	W291	W283	W10	W108		W153	W283
	Dis. (mm)				530.34	1695.71			609.04	1015.48	669.28	1879.97	482.85	1388.09		587.94	1213.89
	Weld				W291	W281			W197	W171	W255	W281	W67	W114		W233	W281
	Dis. (mm)				553.07	1721.64			613.81	1018.10	687.70	1906.44	491.46	1435.05		595.00	1240.39
	Weld				W293	W325			W147	W323	W269	W325	W126	W116		W209	W325
	Dis. (mm)				562.27	1731.19			622.90	1018.68	746.73	1936.47	494.72	1435.39		618.36	1277.37
	Weld				W269	W323			W253	W165	W293	W145	W127			W273	W145
	Dis. (mm)				618.23	1761.03			627.64	1028.43	751.62	1951.73	495.50			620.00	1285.70

Table 7.40 Framing key locators and all relevant weld and distance – part 4/6

Locator		L48		L49	L50	L51	L52	L53	L54	L56	L57	L58	L59			L60			L61	L63		
Related Welds	Weld	W106	W99	W370	W350	W326	W328	W382	W356	W348	W332	W350	W210	W272	W152	W324	W230	W184	W386	W136	W248	W61
	Dis. (mm)	129.85	708.05	82.43	201.74	133.76	94.63	94.59	73.04	170.24	247.28	170.78	0.92	437.00	1644.91	97.38	653.67	1821.00	138.22	156.18	623.93	1057.80
	Weld	W108	W111	W372	W348	W82	W330	W378	W358	W350	W328	W348	W288	W206	W294	W322	W208	W296	W400	W302	W212	W59
	Dis. (mm)	132.48	739.75	83.22	202.73	146.94	109.45	119.19	74.75	194.97	250.08	198.89	44.35	452.69	1656.71	106.22	659.41	1849.66	181.31	178.51	648.88	1062.58
	Weld	W104	W69	W366	W354	W118	W332	W376		W352	W338	W354	W212	W232	W168	W320	W272	W152	W402	W304	W264	W57
	Dis. (mm)	139.95	750.30	110.27	230.88	148.32	120.74	134.11		199.15	261.13	199.74	59.01	462.00	1683.91	122.00	677.06	1858.39	213.40	198.21	653.95	1067.93
	Weld	W400	W97	W368	W352	W39	W334	W384		W354	W330	W352	W286	W208	W296	W318	W232	W298	W394	W208	W250	W300
	Dis. (mm)	143.24	787.45	133.54	231.86	150.98	158.63	144.05		220.60	263.18	224.42	83.55	476.74	1686.97	142.42	700.56	1882.58	214.52	236.43	655.03	1073.51
	Weld	W386	W68	W362		W51	W336	W380			W340		W214	W274	W298	W278	W274	W168	W388	W134	W234	W55
	Dis. (mm)	148.79	800.01	149.72		216.52	171.22	173.43			266.90		88.41	607.00	1719.20	143.60	838.59	1897.22	231.10	240.22	682.31	1073.84
	Weld	W102	W130	W360		W120	W338	W374			W342		W284	W136	W300	W276	W136	W300	W398	W206	W236	W168
	Dis. (mm)	160.20	821.34	156.61		229.28	172.08	189.62			270.44		121.28	678.19	1747.55	163.65	867.33	1911.51	231.12	256.74	686.50	1087.70
	Weld	W114	W131	W364		W122	W340				W336		W188	W302	W164	W316	W302	W164	W396	W204	W266	W172
	Dis. (mm)	171.69	822.26	156.92		251.21	188.96				286.10		123.40	745.76	1751.42	165.78	935.09	1962.58	240.27	287.11	690.85	1134.34
	Weld	W116	W67			W49	W342				W334		W234	W304	W170	W234	W304	W170	W390	W274	W210	W162
	Dis. (mm)	174.32	881.54			256.38	199.27				318.38		144.15	778.58	1761.57	193.98	968.33	1963.14	271.93	294.61	701.95	1135.98
	Weld	W394	W126			W124	W346				W346		W190	W240	W174	W280	W240	W174	W392	W240	W276	W174
	Dis. (mm)	180.99	885.43			293.68	468.62				534.34		150.60	790.00	1764.95	234.48	1015.56	1963.40	274.18	330.06	716.97	1138.11
	Weld	W100	W127			W43	W344				W344		W276	W134	W162	W144	W134	W162	W387	W232	W252	W170
	Dis. (mm)	188.70	885.59			304.75	491.30				555.56		154.90	835.81	1768.90	236.07	1026.32	1963.85	1448.96	341.73	741.06	1140.40
	Weld	W388	W112			W114							W282	W270	W172	W282	W270	W172	W403	W186	W278	W164
	Dis. (mm)	189.24	909.84			317.90							156.89	881.00	1772.53	238.20	1104.29	1964.28	1459.44	346.58	741.40	1149.29
	Weld	W402	W66			W116							W216	W268		W142	W292		W401	W272	W288	
	Dis. (mm)	198.46	923.86			343.98							158.40	978.00		240.60	1176.14		1459.77	355.25	742.44	
	Weld	W398	W95			W47							W278	W292		W61	W268		W395	W292	W146	
	Dis. (mm)	202.69	947.41			354.89							166.81	984.72		243.68	1199.18		1462.77	368.63	743.96	
	Weld	W77	W93			W45							W192	W254		W284	W254		W399	W230	W286	
	Dis. (mm)	203.88	991.17			397.85							168.03	1048.40		246.85	1240.25		1463.72	369.96	778.63	
	Weld	W396	W65			W53							W280	W290		W138	W290		W397	W270	W148	
	Dis. (mm)	204.43	1031.65			431.19							183.39	1055.80		247.67	1275.64		1466.82	380.63	783.76	
	Weld	W84	W91			W41							W194	W242		W59	W242		W391	W202	W158	
	Dis. (mm)	221.97	1034.99			463.52							192.03	1083.87		258.13	1303.25		1491.69	382.27	805.46	
	Weld	W390	W89										W218	W256		W286	W256		W389	W228	W284	
	Dis. (mm)	225.45	1078.22										193.40	1114.22		261.05	1306.48		1495.86	398.09	813.72	
	Weld	W392	W87										W196	W244		W190	W258		W393	W200	W156	
	Dis. (mm)	233.67	1121.27										220.43	1118.87		266.10	1332.32		1516.21	418.59	826.50	

(continued)

Table 7.40 (continued)

Locator		L48		L49	L50	L51	L52	L53	L54	L56	L57	L58	L59		L60		L61	L63	
	Weld	W86	W85										W220	W258	W140	W244		W226	W154
	Dis. (mm)	257.98	1163.46										228.40	1139.92	270.19	1337.68		422.58	843.11
	Weld	W1	W64										W144	W246	W57	W260		W254	W282
	Dis. (mm)	260.06	1168.83										228.72	1153.87	274.10	1367.23		428.13	847.04
	Weld	W88	W101										W198	W260	W288	W246		W224	W180
	Dis. (mm)	296.59	1205.79										251.74	1174.66	280.46	1372.15		448.46	857.74
	Weld	W3	W63										W142	W248	W55	W262		W268	W324
	Dis. (mm)	302.55	1212.27										253.71	1188.87	291.34	1402.15		448.58	864.85
	Weld	W76	W103										W324	W262	W192	W248		W198	W182
	Dis. (mm)	318.80	1248.11										253.82	1209.41	303.06	1406.63		455.39	865.04
	Weld	W90	W105										W222	W250	W210	W238		W222	W178
	Dis. (mm)	336.59	1288.73										263.40	1223.87	307.08	1437.08		475.52	866.90
	Weld	W5	W107										W138	W238	W194	W250		W256	W280
	Dis. (mm)	337.66	1329.37										279.01	1244.17	340.60	1441.15		490.84	871.76
	Weld	W92	W109										W322	W236	W212	W264		W196	W176
	Dis. (mm)	377.64	1370.05										279.48	1258.87	348.19	1475.30		492.56	873.63
	Weld	W37	W387										W200	W264	W214	W236		W220	W150
	Dis. (mm)	412.17	1399.59										285.02	1282.23	370.43	1475.69		503.56	878.92
	Weld	W94	W401										W224	W252	W196	W266		W290	W322
	Dis. (mm)	419.37	1404.88										298.40	1318.80	378.48	1513.17		510.31	894.38
	Weld	W129	W403										W320	W266	W188	W252		W258	W144
	Dis. (mm)	437.23	1407.08										305.92	1320.03	398.11	1534.85		515.56	914.58
	Weld	W128	W395										W202	W148	W198	W146		W194	W320
	Dis. (mm)	442.26	1407.79										319.63	1365.00	416.60	1568.29		530.02	923.95
	Weld	W75	W399										W318	W146	W216	W148		W218	W160
	Dis. (mm)	445.68	1409.22										332.97	1375.97	426.86	1580.54		532.42	936.65
	Weld	W133	W397										W140	W158	W200	W158		W242	W142
	Dis. (mm)	498.68	1411.04										333.01	1445.70	454.90	1634.05		533.56	938.31
	Weld	W74	W115										W226	W182	W218	W156		W260	W318
	Dis. (mm)	499.44	1417.63										333.40	1452.00	456.49	1657.15		549.13	953.54
	Weld	W132	W117										W186	W156	W220	W182		W216	W138
	Dis. (mm)	503.90	1418.34										355.21	1471.52	486.83	1666.61		561.97	962.38
	Weld	W113	W391										W316	W154	W202	W154		W244	W166
	Dis. (mm)	539.11	1432.99										360.48	1493.49	493.35	1675.81		563.12	974.08
	Weld	W73	W389										W228	W180	W222	W180		W192	W294
	Dis. (mm)	549.35	1438.47										368.40	1514.99	517.76	1692.66		567.71	980.74
	Weld	W96	W2										W55	W160	W186	W178		W262	W316
	Dis. (mm)	575.48	1443.09										403.39	1527.70	531.91	1703.51		582.87	983.16
	Weld	W72	W4										W57	W178	W224	W176		W188	W296
	Dis. (mm)	600.40	1451.35										405.80	1531.00	549.18	1711.51		592.13	1011.66
	Weld	W110	W393										W59	W176	W226	W150		W246	W184
	Dis. (mm)	622.84	1457.65										409.73	1546.61	581.00	1716.88		593.28	1013.52
	Weld	W98	W6										W230	W150	W204	W160		W190	W140
	Dis. (mm)	654.61	1459.07										412.00	1557.53	598.69	1741.60		605.50	1013.97
	Weld	W71	W38										W61	W166	W228	W166		W238	W298
	Dis. (mm)	658.54	1478.09										415.12	1566.80	613.18	1780.45		616.75	1044.58
	Weld	W70											W204	W184	W206	W294		W214	W152
	Dis. (mm)	700.78											418.17	1607.41	634.78	1818.74		622.79	1049.91

Table 7.41 Framing key locators and all relevant weld and distance – part 5/6

Locator		L64			L65	L67	L68	L69	L70	L71	L72	L73			L74	L75	L76	L78	L79	L80	L81	L82
Related Welds	Welds	W134	W180	W144	W384	W370	W336	W358	W362	W354	W374	W308	W242	W284	W382	W356	W366	W358	W362	W380	W27	W374
	Dis. (mm)	154.23	646.12	1117.54	196.19	97.67	89.51	88.07	165.12	74.22	57.22	6.06	742.38	1915.49	94.59	88.64	38.78	152.99	157.70	116.68	20.17	198.51
	Weld	W304	W178	W320	W378	W368	W332	W356	W368	W350	W380	W306	W254	W282	W378	W358	W372	W356	W364	W376	W29	W380
	Dis. (mm)	173.34	654.92	1133.53	219.53	117.38	113.20	98.08	173.24	99.95	81.89	21.00	747.33	1950.75	119.19	101.13	61.66	188.19	176.50	129.86	25.36	217.19
	Weld	W302	W222	W142	W376	W372	W342		W364	W352	W376	W310	W290	W280	W376		W360		W368	W374	W25	W376
	Dis. (mm)	191.79	655.67	1141.69	229.24	136.87	134.01		198.45	129.92	108.24	33.60	768.56	1976.91	134.11		97.54		192.33	141.35	70.83	221.67
	Weld	W292	W198	W318	W374	W362	W328		W370	W348	W378	W314	W292	W324	W384		W370		W370	W378	W33	W378
	Dis. (mm)	195.33	659.21	1163.26	238.55	166.72	171.58		220.31	145.87	139.03	45.95	810.75	1992.38	144.05		109.58		223.12	149.93	125.11	231.79
	Weld	W254	W176	W138	W382	W366	W338		W360		W384	W312	W268	W144	W380		W364		W360	W382	W23	W384
	Dis. (mm)	241.45	661.91	1166.16	242.63	177.19	185.55		245.96		180.21	58.67	841.68	2021.81	173.43		124.72		225.17	174.05	140.84	245.71
	Weld	W136	W150	W316	W380	W364	W330		W372		W382	W172	W270	W322	W374		W362		W372	W384	W31	W382
	Dis. (mm)	241.48	667.94	1193.00	259.64	195.27	215.60		249.87		199.71	107.27	934.16	2022.32	189.62		145.07		239.47	205.94	157.55	274.58
	Weld	W270	W182	W140		W360	W340		W366			W300	W134	W142			W368		W366		W35	
	Dis. (mm)	294.61	669.40	1218.51		211.23	243.75		284.40			115.73	959.84	2046.59			163.10		270.00		180.47	
	Weld	W256	W220	W61			W334					W162	W304	W320							W80	
	Dis. (mm)	296.27	687.16	1266.90			276.73					132.40	1017.48	2052.25							202.58	
	Weld	W240	W196	W59			W346					W298	W240	W138							W1	
	Dis. (mm)	299.05	697.43	1270.90			595.57					134.40	1021.59	2071.67							210.33	
	Weld	W258	W218	W57			W344					W296	W302	W318							W81	
	Dis. (mm)	318.85	718.97	1275.38			619.06					159.44	1050.61	2082.19							220.51	
	Weld	W268	W194	W55								W174	W136	W316							W3	
	Dis. (mm)	319.39	735.76	1280.33								160.27	1118.82	2112.13							221.54	
	Weld	W260	W160									W294	W274	W140							W5	
	Dis. (mm)	350.10	738.26									185.25	1199.02	2125.22							238.25	
	Weld	W290	W216									W170	W208	W61							W37	
	Dis. (mm)	358.07	751.07									185.35	1326.17	2184.01							298.53	
	Weld	W242	W294									W164	W232	W59							W368	
	Dis. (mm)	374.95	769.50									263.97	1340.66	2186.09							1199.40	
	Weld	W274	W192									W178	W206	W57							W362	
	Dis. (mm)	380.63	774.17									270.42	1351.13	2188.44							1239.70	
	Weld	W262	W166									W176	W272	W55							W370	
	Dis. (mm)	382.00	774.57									271.23	1365.15	2191.08							1244.64	
	Weld	W244	W188									W150	W204								W364	
	Dis. (mm)	397.70	783.42									278.61	1387.47								1287.85	
	Weld	W238	W296									W180	W230								W372	
	Dis. (mm)	414.40	800.36									281.29	1389.66								1299.05	

(continued)

Table 7.41 (continued)

	Weld	W246	W190									W168	W228								W360	
	Dis. (mm)	422.13	812.56									291.22	1432.45								1340.61	
	Weld	W208	W184									W154	W186								W366	
	Dis. (mm)	422.42	813.11									302.07	1454.89								1351.46	
	Weld	W206	W214									W152	W226								W352	
	Dis. (mm)	446.15	815.99									304.58	1466.83								1693.03	
	Weld	W248	W298									W184	W202								W348	
	Dis. (mm)	447.96	833.23									320.97	1493.94								1697.62	
	Weld	W264	W212									W156	W224								W354	
	Dis. (mm)	450.30	843.50									325.34	1501.24								1759.88	
	Weld	W250	W152									W166	W200								W350	
	Dis. (mm)	474.97	848.72									342.14	1533.00								1764.16	
	Weld	W204	W300									W158	W222								W358	
	Dis. (mm)	480.67	862.12									352.88	1535.68								2027.50	
	Weld	W266	W168									W160	W220								W356	
	Dis. (mm)	486.10	885.76									366.73	1570.14								2065.38	
	Weld	W232	W234									W182	W198									
	Dis. (mm)	486.16	890.31									421.44	1572.06									
	Weld	W236	W210									W146	W218									
	Dis. (mm)	502.96	898.99									424.86	1604.63									
	Weld	W272	W172									W266	W196									
	Dis. (mm)	506.33	924.03									480.02	1611.13									
	Weld	W230	W276									W148	W216									
	Dis. (mm)	526.91	925.32									491.36	1639.14									
	Weld	W146	W162									W264	W194									
	Dis. (mm)	537.52	926.38									516.84	1650.21									
	Weld	W186	W174									W252	W188									
	Dis. (mm)	545.38	929.44									530.47	1673.67									
	Weld	W252	W170									W238	W192									
	Dis. (mm)	552.68	932.70									554.01	1689.30									
	Weld	W228	W288									W236	W214									
	Dis. (mm)	563.68	941.00									582.90	1708.22									
	Weld	W202	W164									W262	W190									
	Dis. (mm)	583.16	945.37									588.12	1728.29									
	Weld	W148	W278									W250	W212									
	Dis. (mm)	592.43	949.97									614.12	1737.25									
	Weld	W226	W286									W260	W210									
	Dis. (mm)	593.86	978.34									622.32	1795.52									
	Weld	W158	W284									W248	W234									
	Dis. (mm)	596.47	1014.40									645.73	1807.08									
	Weld	W156	W282									W258	W288									
	Dis. (mm)	616.51	1048.55									656.60	1839.35									
	Weld	W200	W280									W246	W276									
	Dis. (mm)	621.11	1073.86									677.67	1842.49									
	Weld	W224	W324									W256	W278									
	Dis. (mm)	624.54	1074.12									682.02	1867.41									
	Weld	W154	W322									W244	W286									
	Dis. (mm)	632.25	1103.82									709.90	1878.15									

(continued)

Table 7.42 Framing key locators and all relevant weld and distance – part 6/6

Locator		L83	L84	L85	L86			L87	L88	L89			L91			L92		L93	L94		
Related Welds	Weld	W334	W386	W11	W176	W270	W144	W17	W120	W304	W218	W144	W170	W270	W322	W49	W13	W9	W290	W168	W322
	Dis. (mm)	46.54	61.81	104.59	15.21	703.46	1766.21	164.78	191.01	156.03	639.98	1033.61	43.93	843.44	1966.45	105.98	527.38	3.89	2.42	626.95	1307.20
	Weld	W330	W400	W114	W178	W134	W142	W15	W122	W302	W194	W320	W174	W134	W142	W108	W113	W13	W242	W206	W142
	Dis. (mm)	85.39	78.97	115.84	34.46	707.25	1790.81	165.15	198.87	158.98	651.82	1048.32	61.83	901.30	1976.72	110.48	556.30	44.23	26.96	642.55	1310.70
	Weld	W328	W394	W1	W150	W304	W320	W19	W39	W134	W148	W142	W164	W240	W320	W43	W73	W7	W244	W230	W138
	Dis. (mm)	127.49	109.11	116.33	50.96	763.64	1790.88	169.29	203.12	168.16	668.94	1057.62	63.38	934.20	1996.42	111.28	567.02	59.75	61.91	645.00	1336.00
	Weld	W340	W402	W3	W180	W240	W138	W21	W118	W136	W216	W318	W168	W304	W138	W120	W15	W15	W268	W294	W320
	Dis. (mm)	162.08	113.00	139.26	69.77	785.76	1815.71	178.16	204.61	186.41	671.32	1077.98	79.51	959.34	2002.00	114.56	576.36	93.95	79.02	667.08	1337.04
	Weld	W338	W398	W116	W154	W302	W318	W13	W82	W292	W158	W138	W162	W302	W318	W51	W96	W17	W246	W204	W318
	Dis. (mm)	196.77	125.94	143.12	104.37	796.18	1820.73	179.28	222.27	262.37	683.35	1081.92	85.96	992.63	2026.40	115.40	592.88	118.74	96.89	678.22	1366.89
	Weld	W332	W388	W5	W294	W136	W316	W9	W124	W240	W192	W316	W152	W136	W140	W106	W17	W37	W248	W228	W140
	Dis. (mm)	202.91	134.09	163.17	117.83	863.54	1850.59	209.96	229.02	293.08	689.97	1107.65	103.19	1060.66	2055.97	115.77	601.07	123.91	131.89	688.60	1390.00
	Weld	W336	W396	W37	W156	W274	W140	W7	W114	W270	W188	W140	W172	W274	W316	W122	W72	W19	W254	W296	W316
	Dis. (mm)	264.52	134.87	213.85	131.97	956.07	1868.91	254.14	251.20	314.78	703.01	1134.02	109.16	1116.84	2056.38	123.52	618.28	143.42	141.69	693.27	1396.74
	Weld	W342	W390	W7	W296	W208	W61	W37	W326	W254	W156	W61	W184	W232	W57	W104	W19	W21	W256	W164	W55
	Dis. (mm)	275.52	171.60	239.63	146.56	1068.22	1923.50	311.06	253.36	316.22	704.10	1181.13	132.68	1261.63	2124.53	134.16	625.88	168.15	148.55	694.44	1459.73
	Weld	W346	W392	W9	W158	W206	W59	W5	W51	W274	W154	W59	W166	W208	W59	W118	W110	W5	W258	W170	W57
	Dis. (mm)	337.03	175.42	275.31	158.73	1,093.02	1926.45	360.40	254.64	330.91	720.48	1185.25	168.15	1269.48	2124.54	136.16	640.46	193.69	158.86	707.54	1460.38
	Weld	W344	W387	W13	W298	W232	W57	W3	W116	W208	W190	W57	W160	W272	W55	W124	W21	W3	W292	W174	W59
	Dis. (mm)	359.77	1443.67	312.98	177.98	1093.91	1929.71	380.07	283.02	344.33	728.12	1189.88	204.57	1286.60	2124.81	136.17	650.85	224.47	162.93	712.93	1461.46
	Weld		W403	W15	W300	W272	W55	W1	W49	W268	W214	W55	W182	W206	W61	W47	W98	W1	W250	W162	W61
	Dis. (mm)		1448.17	358.73	206.00	1117.86	1933.30	405.48	288.79	362.20	735.00	1195.02	277.46	1294.38	2124.85	138.04	672.34	263.12	166.88	719.68	1462.96
	Weld		W401	W17	W146	W204		W11	W43	W206	W180		W154	W230		W116	W71	W11	W270	W298	
	Dis. (mm)		1450.62	382.45	219.13	1129.07		428.63	328.81	367.18	735.02		288.48	1311.56		142.53	676.61	298.20	176.01	721.56	
	Weld		W399	W19	W266	W230		W114	W47	W256	W178		W180	W204		W114	W70	W114	W260	W226	
	Dis. (mm)		1450.84	406.63	262.57	1141.85		466.84	374.87	375.54	744.33		292.99	1330.82		143.10	718.96	346.20	177.90	723.60	
	Weld		W395	W21	W264	W228		W116	W45	W258	W182		W156	W228		W82	W99	W116	W262	W172	
	Dis. (mm)		1451.06	431.23	292.38	1183.80		483.04	415.06	399.28	748.20		294.13	1355.11		146.19	725.96	350.74	201.27	726.34	
	Weld		W397		W172	W186			W53	W204	W176		W300	W226		W39	W111		W236	W186	
	Dis. (mm)		1453.28		298.15	1196.00			446.60	400.58	751.70		302.39	1390.08		146.84	757.77		201.88	744.58	
	Weld		W391		W162	W226			W41	W290	W150		W298	W186		W102	W69		W238	W300	
	Dis. (mm)		1476.70		310.53	1217.56			477.44	413.43	757.68		305.01	1398.41		161.03	768.60		227.64	746.74	
	Weld		W389		W160	W202				W232	W212		W178	W224		W45	W97		W264	W224	
	Dis. (mm)		1483.98		317.37	1234.79				419.21	762.07		308.28	1425.04		169.66	805.59		258.86	758.60	
	Weld		W393		W182	W224				W260	W234		W158	W202		W11	W68		W252	W202	
	Dis. (mm)		1501.49		321.80	1251.39				431.76	805.57		308.29	1437.55		193.34	818.43		261.81	783.11	

(continued)

Table 7.42 Framing key locators and all relevant weld and distance – part 6/6

Locator		L83	L84	L85	L86		L87	L88	L89		L91		L92		L93	L94	
	Weld				W166	W200			W242	W210	W296	W222	W100	W130		W240	W222
	Dis. (mm)				322.73	1273.61			433.80	816.81	311.28	1460.01	194.36	841.11		267.01	793.60
	Weld				W238	W222			W272	W160	W294	W200	W53	W131		W134	W200
	Dis. (mm)				323.92	1285.28			437.51	818.48	320.16	1476.69	197.19	842.89		271.31	821.70
	Weld				W174	W198			W230	W276	W176	W220	W326	W67		W266	W220
	Dis. (mm)				325.70	1312.45			456.43	840.40	341.34	1494.97	201.58	900.04		291.71	828.60
	Weld				W184	W220			W244	W166	W148	W198	W77	W127		W148	W198
	Dis. (mm)				334.55	1319.23			460.27	855.35	362.69	1515.84	211.42	905.37		308.01	860.35
	Weld				W170	W196			W186	W288	W146	W218	W41	W126		W304	W218
	Dis. (mm)				340.56	1351.31			463.76	858.35	364.19	1529.94	225.41	906.04		322.89	863.60
	Weld				W148	W218			W262	W294	W150	W196	W1	W112		W146	W216
	Dis. (mm)				348.00	1353.23			464.61	859.38	368.28	1554.98	230.02	928.11		344.08	898.60
	Weld				W152	W216			W246	W278	W252	W216	W84	W66		W302	W196
	Dis. (mm)				348.97	1387.28			487.82	864.95	408.23	1564.91	230.88	942.32		353.31	899.05
	Weld				W262	W194			W228	W296	W266	W194	W86	W95		W182	W188
	Dis. (mm)				353.91	1390.20			490.75	880.16	417.18	1594.14	269.08	965.68		395.01	933.60
	Weld				W168	W188			W238	W184	W264	W188	W3	W93		W136	W194
	Dis. (mm)				366.98	1421.38			497.75	894.30	454.44	1599.88	272.60	1009.41		416.20	937.79
	Weld				W252	W192			W202	W286	W236	W192	W5	W65		W158	W214
	Dis. (mm)				369.44	1429.09			500.89	895.33	467.54	1633.29	307.90	1049.78		416.37	968.60
	Weld				W260	W214			W248	W298	W238	W214	W88	W91		W156	W192
	Dis. (mm)				384.75	1455.52			516.27	922.94	492.20	1634.86	309.27	1053.18		445.64	976.56
	Weld				W164	W190			W226	W152	W250	W212	W76	W89		W274	W212
	Dis. (mm)				392.68	1467.91			519.31	930.34	502.24	1664.23	332.26	1096.40		450.00	998.00
	Weld				W236	W212			W264	W284	W262	W190	W90	W87		W160	W190
	Dis. (mm)				403.72	1484.23			534.28	931.08	526.79	1672.34	350.56	1139.47		470.70	1015.27
	Weld				W258	W210			W200	W300	W248	W210	W37	W85		W154	W210
	Dis. (mm)				416.24	1541.88			538.31	951.76	536.97	1723.18	383.37	1181.66		472.67	1057.00
	Weld				W250	W234			W250	W282	W260	W234	W92	W64		W180	W234
	Dis. (mm)				426.32	1546.38			545.48	964.97	561.42	1751.26	392.70	1186.85		502.37	1093.56
	Weld				W256	W276			W224	W168	W246	W288	W7	W101		W166	W288
	Dis. (mm)				439.89	1581.66			548.62	967.75	571.74	1767.49	434.19	1223.97		509.81	1101.34
	Weld				W248	W288			W266	W324	W258	W276	W94	W63		W178	W276
	Dis. (mm)				450.50	1585.28			570.61	989.07	596.09	1786.73	435.24	1230.28		528.59	1128.78
	Weld				W246	W278			W236	W280	W244	W286	W129	W103		W184	W286
	Dis. (mm)				476.03	1606.48			575.34	990.09	606.54	1806.68	458.34	1266.27		550.43	1140.55
	Weld				W254	W286			W198	W172	W256	W278	W128	W105		W176	W278
	Dis. (mm)				501.54	1623.72			575.96	I011.25	621.78	1811.71	461.66	1306.86		561.12	1153.61
	Weld				W244	W284			W222	W162	W242	W284	W75	W107		W150	W284
	Dis. (mm)				502.70	1660.74			578.56	I013.07	641.35	1844.38	462.54	1347.45		585.13	1178.28
	Weld				W242	W282			W220	W174	W290	W282	W9	W109		W152	W282
	Dis. (mm)				530.34	1695.71			609.04	I015.48	669.28	1879.97	482.85	1388.09		587.94	1213.89
	Weld				W290	W280			W196	W170	W254	W280	W74	W115		W232	W280
	Dis. (mm)				553.07	1721.64			613.81	I018.10	687.70	1906.44	516.88	1435.05		595.00	1240.39
	Weld				W292	W324			W146	W322	W268	W324	W133	W117		W208	W324
	Dis. (mm)				562.27	1731.19			622.90	I018.68	746.73	1936.47	519.70	1435.39		618.36	1277.37
	Weld				W268	W322			W252	W164	W292	W144	W132			W272	W144
	Dis. (mm)				618.23	1761.03			627.64	I028.43	751.62	1951.73	523.48			620.00	1285.70

Table 7.43 Framing welds with maximum distance in each joint

Joint	Weld 1	Weld 2	Distance (mm)
Joint 1	W1	W37	155.85
Joint 2	W2	W38	155.85
Joint 3	W11	W21	464.58
Joint 4	W12	W22	464.58
Joint 5	W35	W81	302.56
Joint 6	W36	W79	308.59
Joint 7	W41	W326	380.98
Joint 8	W42	W327	380.98
Joint 9	W55	W172	2171.81
Joint 10	W56	W173	2171.81
Joint 11	W108	W109	1377.40
Joint 12	W114	W116	33.48
Joint 13	W115	W117	33.48
Joint 14	W312	W314	104.55
Joint 15	W313	W315	104.55
Joint 16	W336	W344	600.07
Joint 17	W337	W345	600.07
Joint 18	W350	W352	78.41
Joint 19	W351	W353	78.41
Joint 20	W356	W358	40.71
Joint 21	W357	W359	40.71
Joint 22	W366	W368	157.39
Joint 23	W367	W369	157.39
Joint 24	W374	W382	165.53
Joint 25	W375	W383	165.53
Joint 26	W392	W393	1533.44

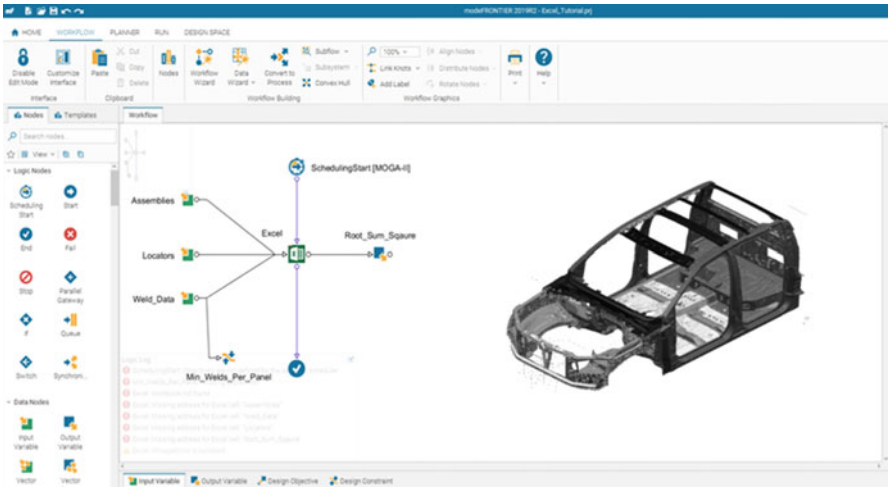


Fig. 7.40 Schematic of bodyshell GA model developed in modeFRONTIER

Table 7.44 Bodyshell optimised weld process design resulted from GA model

Bodyshell GA Outcome				
W5	W109	W173	W281	W351
W6	W114	W158	W300	W356
W13	W115	W159	W301	W357
W14	W116	W180	W312	W358
W27	W117	W181	W313	W359
W28	W118	W182	W314	W362
W39	W119	W183	W315	W363
W40	W134	W198	W316	W364
W41	W135	W199	W317	W365
W42	W136	W240	W324	W376
W61	W137	W241	W325	W377
W62	W140	W252	W326	W378
W63	W141	W253	W327	W379
W70	W150	W260	W330	W388
W77	W151	W261	W331	W389
W82	W164	W274	W332	W394
W83	W165	W275	W333	W395
W98	W168	W278	W348	
W99	W169	W279	W349	
W108	W172	W280	W350	

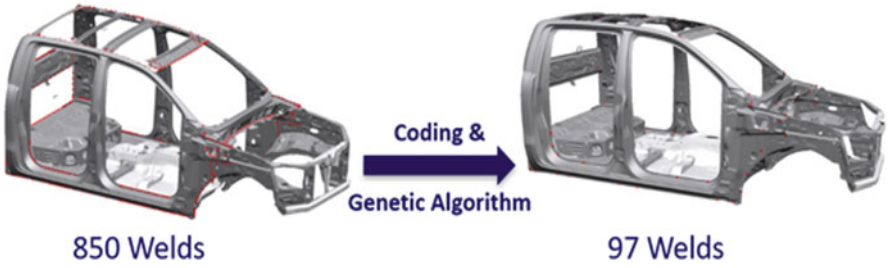


Fig. 7.41 GA optimisation of bodyshell welds eliminated non-fit process designs

References

1. Krishnan V and Ulrich KT. Product Development Decisions: A Review of the Literature. *Management Science* 2001; 47: 1–21. DOI: https://doi.org/10.1287/mnsc.47.1.1.10668.
2. Unger D, Eppinger, Steven. Improving product development process design: a method for managing information flows, risks, and iterations. *Journal of Engineering Design* 2011; 22: 689–699. DOI: https://doi.org/10.1080/09544828.2010.524886.
3. Srinivasan V. Computational Metrology for the Design and Manufacture of Product Geometry: A Classification and Synthesis. *Journal of Computing and Information Science in Engineering* 2006; 7: 3–9. DOI: https://doi.org/10.1115/1.2424246.
4. Kalpakjian S, Schmid SR and Kok C-W. *Manufacturing processes for engineering materials*. 5th ed. Singapore; London: Pearson-Prentice Hall, 2008, p.xxi, 1010 p.

5. Söderberg R, Lindkvist L, Wärmefjord K, et al. Virtual Geometry Assurance Process and Toolbox. *Procedia CIRP* 2016; 43: 3–12. DOI: https://doi.org/10.1016/j.procir.2016.02.043.
6. Hu M, Lin Z, Lai X, et al. Simulation and analysis of assembly processes considering compliant, non-ideal parts and tooling variations. *International Journal of Machine Tools and Manufacture* 2001; 41: 2233–2243. https://doi.org/10.1016/S0890-6955(01)00044-X.
7. Bedeoui A, Benhadj R, Trigui M, et al. Assembly plans generation of complex machines based on the stability concept. *Procedia CIRP* 2018; 70: 66–71. DOI: https://doi.org/10.1016/j.procir.2018.03.030.
8. Nof S, W. Wilbert. *Industrial Assembly*. Chapman & Hall, 1997.
9. Homem de Mello LS and Lee S. *Computer-aided mechanical assembly planning*. Boston: Kluwer Academic Publishers, 1991, p.xi, 446.
10. J.D. W. A constraint-based approach to planning with subassemblies. In: *1990 IEEE International Conference on Systems Engineering* 9–11 Aug. 1990 1990, pp.412–415.
11. Lv HL, C. An assembly sequence planning approach with a discrete particle swarm optimization algorithm. *Int J Adv Manuf Technol* 2010; 50: 761–770.
12. Gottipolu RB and Ghosh K. Representation and selection of assembly sequences in computer-aided assembly process planning. *International Journal of Production Research* 1997; 35: 3447–3466. DOI: https://doi.org/10.1080/002075497194183.
13. Rashid M, Hutabarat W and Tiwari A. *A review on assembly sequence planning and assembly line balancing optimisation using soft computing approaches*. 2012.
14. Delchambre A. Computer-aided Assembly Planning. Dordrecht: Springer Netherlands, 1992, p. 1 online resource (292 pages).
15. Bourjault A. *Contribution a une Approche Methodologique de l'Assemblage Automatise* Universite de Franche Comte 1984.
16. Lui MCM. *Generation and evaluation of mechanical assembly sequences using the liaison-sequence method* 1988.
17. De Fazio TL, Whitney D.E. Simplified Generation of All Mechanical Assembly Sequences. *IEEE Journal of Robotics and Automation RA-3* 1987; 6: 640–658.
18. Sanderson AC. Task Planning and Control Synthesis for Flexible Assembly Systems. *IEEE Transactions on Aerospace and Electronic Systems* 1988; 24: 619-629.
19. Homem de Mello LS and Sanderson AC. A correct and complete algorithm for the generation of mechanical assembly sequences. *IEEE Transactions on Robotics and Automation* 1991; 7: 228-240. DOI: https://doi.org/10.1109/70.75905.
20. Baldwin DF, Abell, T.E., Lui, M.M., De Fazio, T.L., & Whitney, D.E. An integrated computer aid for generating and evaluating assembly sequences for mechanical products. . *IEEE Transactions on Robotics and Automation* 1991; 7: 78–94.
21. Golabi S. Automatic generation of all geometrically feasible assembly sequences using solid modelling. *University of South Australia* 1996.
22. Dini G, F. Failli, et al. Generation of Optimised Assembly Sequences Using Genetic Algorithms. *Annals of the CIRP* 1999; 48: 17–20.
23. Li Y GJ, Shi L, et al. . Assembly modeling for 3d component based on polychromatic sets. . *Adv Mater Res* 2012; 411: 388–392.
24. Bahubalendruni MVAR and Biswal BB. A review on assembly sequence generation and its automation. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 2015; 230: 824–838. DOI: https://doi.org/10.1177/0954406215584633.
25. Dini G, & Santochi, M. . Automated sequencing and subassembly detection in assembly planning. . *Annals of the CIRP* 1992; 41 1–4.
26. Akagi F, Osaki, H., & Kikuchi, S. The method of analysis of assembly work based on the fastener method. *Bulletin of the JSME* 1980; 23: 1670–1675.
27. Huang YFaCSGL. A Framework of Knowledge-based Assembly Planning. . *IEEE International Conference on Robotics and Automation*. Sacramento, California.1991.
28. Woo TC. *Automatic Disassembly and Total Ordering in Three Dimensions*. 1987.

29. Pu P. An Assembly Sequence Generation Algorithm Using Case-Based Search Techniques. *IEEE International Conference on Robotics and Automation*. Nice, France.1992.
30. Milner JM, Graves SC and Whitney DE. Using simulated annealing to select least-cost assembly sequences. In: *Proceedings of the 1994 IEEE International Conference on Robotics and Automation* 8–13 May 1994, pp.2058–2063 vol. 2053.
31. J.F. Wang JHLFZ. A novel ant colony algorithm for assembly sequence planning. *The International Journal of Advanced Manufacturing Technology*; 25: 1137–1143.
32. Zheng Z. PN. A Discrete Particle Swarm Optimization Approach to Compose Heterogeneous Learning Groups. In: *2014 IEEE 14th International Conference on Advanced Learning Technologies* 7–10 July 2014 2014, pp.49–51.
33. Y. CCaP. An integration of neural network and rule-based systems for design and planning of mechanical assemblies. *IEEE Trans Syst Man Cybernet* 1993, p. 1359–1371.
34. P.-B. Cao R-BX. Assembly planning using a novel immune approach. *The International Journal of Advanced Manufacturing Technology* 2007; 31: 770–782.
35. Zeng C, Gu T, Zhong Y, et al. *A Multi-Agent Evolutionary algorIthm for Connector-Based Assembly Sequence Planning*. 2011, p.3689–3693.
36. Shuang B CJaLZ. Microrobot based microassembly sequence planning with hybrid ant colony algorithm. *Int J Adv Manuf Technol* 2008; 38: 1227–1235.
37. Bonneville F, Perrard C and Henrioud JM. A genetic algorithm to generate and evaluate assembly plans. In: *Proceedings 1995 INRIA/IEEE Symposium on Emerging Technologies and Factory Automation ETFA'95* 10–13 Oct. 1995 1995, pp.231–239 vol.232.
38. Hong DS and Cho HS. A genetic-algorithm-based approach to the generation of robotic assembly sequences. *Control Engineering Practice* 1999; 7: 151-159. DOI: https://doi.org/10.1016/S0967-0661(98)00177-4.
39. Senin N, Groppetti, R., Wallace, D. R. Concurrent assembly planning with genetic algorithms. *ROBOTICS AND COMPUTER INTEGRATED MANUFACTURING* 2000; 16: 65–72.
40. Lazzerini BaM, F. *A genetic algorithm for generating optimal assembly plans Artificial Intelligence in Engineering* 2000; 14: 319–329.
41. De Lit P, Latinne P, Rekiek B, et al. Assembly planning with an ordering genetic algorithm. *International Journal of Production Research* 2001; 39: 3623–3640. DOI: https://doi.org/10.1080/00207540110056135.
42. Chen SFaL, Y.J. An adaptive genetic assembly-sequence planner. . *International Journal of Computer Integrated Manufacturing* 2001; 14: 489–500.
43. Smith SSFaL, Y.J. The application of multi-level genetic algorithms in assembly planning. *Journal of Industrial Technology* 2001; 17: 1–4.
44. Guan Q, Liu JH and Zhong YF. A concurrent hierarchical evolution approach to assembly process planning. *International Journal of Production Research* 2002; 40: 3357–3374. DOI: https://doi.org/10.1080/00207540210146152.
45. Smith G.C. S, S. *An enhanced genetic algorithm for automated assembly planning*. 2002, p.355–364.
46. Carmelo Del Valle RMG, Miguel Toro, Eduardo F. Camacho. A Genetic Algorithm for Assembly Sequence Planning. *International Work-Conference on Artificial Neural Networks* 2003: 337–344.
47. Marian RM, Luong LHS and Abhary K. Assembly sequence planning and optimisation using genetic algorithms: Part I. Automatic generation of feasible assembly sequences. *Applied Soft Computing* 2003; 2: 223–253. DOI: https://doi.org/10.1016/S1568-4946(02)00064-9.
48. Marian RM, Luong, L.H. and Abhary, K. A genetic algorithm for the optimisation of assembly sequences. *Computers & Industrial Engineering* 2006; 50: 503–527.
49. Bai Y.W. CZN, Bin H.Z., Hun J. An effective integration approach toward assembly sequence planning and evaluation. *The International Journal of Advanced Manufacturing Technology* 2005; 27: 96–105.
50. Huang J, Du, P.A. and Liao, W.Z. Genetic algorithm for assembly sequences planning based on assembly constraint. . *COMPUTER INTEGRATED MANUFACTURING SYSTEMS* 2007; 13: 756.

51. Hui W, Dong X and Guanghong D. *A genetic algorithm for product disassembly sequence planning*. 2008, p.2720–2726.
52. Choi YK, Lee, D.M. and Cho, Y.B. An approach to multi-criteria assembly sequence planning using genetic algorithms. . *The International Journal of Advanced Manufacturing Technology* 2009; 42: 180–188.
53. Tseng YJ, Chen, J.Y. and Huang, F.Y. A multi-plant assembly sequence planning model with integrated assembly sequence planning and plant assignment using GA. *The International Journal of Advanced Manufacturing Technology* 2010; 48: 333–345.
54. Pan H, Hou WJ and Li TM. Genetic Algorithm for Assembly Sequences Planning Based on Heuristic Assembly Knowledge. *Applied Mechanics and Materials* 2011; 44–47: 3657–3661. DOI: https://doi.org/10.4028/www.scientific.net/AMM.44-47.3657.
55. J. Goldak BP, M. Bibby, J. Moore. Computational weld mechanics. *AGARD Workshop, Structures and Materials 61st Panel Meeting*. 1985.
56. Gene Liao Y. Optimal design of weld pattern in sheet metal assembly based on a genetic algorithm. *The International Journal of Advanced Manufacturing Technology* 2005; 26: 512–516.
57. Huh H and Kang WJ. Electrothermal analysis of electric resistance spot welding processes by a 3-D finite element method. *Journal of Materials Processing Technology* 1997; 63: 672–677. DOI: https://doi.org/10.1016/S0924-0136(96)02705-7.
58. Li W, Cheng S, Hu SJ, et al. Statistical Investigation on Resistance Spot Welding Quality Using a Two-State, Sliding-Level Experiment. *Journal of Manufacturing Science and Engineering* 2000; 123: 513–520. DOI: https://doi.org/10.1115/1.1382595.
59. Tang H, Hou, W., and Hu, S. J. Forging force in resistance spot welding. . In: *Proc Instn Mech Engrs, Part B: J Engineering Manufacture* 2002, pp.957–968.
60. Lian J, Lai XM, Lin ZQ, et al. Application of data mining and process knowledge discovery in sheet metal assembly dimensional variation diagnosis. *Journal of Materials Processing Tech* 2002; 129: 315–320. DOI: https://doi.org/10.1016/S0924-0136(02)00691-X.
61. Le Meur G, Bourouga B and Dupuy T. Measurement of contact parameters at electrode/sheet interface during resistance spot welding process. *Science and Technology of Welding and Joining* 2003; 8: 415–422. DOI: https://doi.org/10.1179/136217103225005589.
62. Kadivar MH, Jafarpur K and Baradaran GH. Optimizing welding sequence with genetic algorithm. *Computational Mechanics* 2000; 26: 514–519. DOI: https://doi.org/10.1007/s004660000195.
63. Meriam JL, DeGarmo, E.P., Jonassen, F. A method for the measurement of residual welding stresses. *Welding Research Supplement* 1946: 340s–343s.
64. Fukuda S and Yoshikawa K. Determination of welding sequence: a neural net approach. *Engineering Analysis with Boundary Elements* 1990; 7: 78–82. DOI: https://doi.org/10.1016/0955-7997(90)90024-4.
65. Park S-C. Distortion mechanisms and control methodology for welding thin-plate panel structures. In: Tsai CL, (ed.). ProQuest Dissertations Publishing, 1998.
66. Wikipedia. Travelling Sales Man Problem, https://en.wikipedia.org/wiki/Travelling_salesman_problem (2018, 2018).
67. Liao Y-F YD-H, Chen C-L. Evolutionary algorithm to traveling salesman problems. . *Computers & Mathematics with Applications* 2012; 64: 788–797.
68. Huang M-W, Hsieh CC and Arora JS. A genetic algorithm for sequencing type problems in engineering design. *International Journal for Numerical Methods in Engineering* 1998; 40: 3105–3115. DOI: https://doi.org/10.1002/(SICI)1097-0207(19970915)40:17<3105::AID-NME200>3.0.CO;2-F.
69. JC. B. Genetic Algorithms and Random Keys for Sequencing and Optimization. *ORSA Journal on Computing* 1994; 6: 154–160.
70. Damsbo M and Ruhoff PT. An evolutionary algorithm for welding task sequence ordering. In: *Artificial Intelligence and Symbolic Computation* (eds Calmet J and Plaza J), Berlin, Heidelberg, 1998// 1998, pp.120–131. Springer Berlin Heidelberg.
71. Whitley D, Starkweather, T., Shaner, D. *The Travelling Salesman and Sequence Scheduling: Quality Solutions using Genetic Edge Recombination.* . Van Nostrand Reinhold, 1991.

72. Xie LS and Hsieh C. Clamping and welding sequence optimisation for minimising cycle time and assembly deformation. *International Journal of Materials and Product Technology* 2002; 17: 389–399. DOI: https://doi.org/10.1504/IJMPT.2002.005465.
73. Kim K-Y, Norman B and O. Nnaji B. *Heuristics for single-pass welding task sequencing*. 2002, p.2769–2788.
74. Liao YG. Optimal design of weld pattern in sheet metal assembly based on a genetic algorithm. *The International Journal of Advanced Manufacturing Technology* 2005; 26: 512–516. DOI: https://doi.org/10.1007/s00170-003-2003-5.
75. Chapple A TZ, Jardine F. Weld Distortion Optimisation using HyperStudy. *The 8th UK Altair Technology Conference*. 2013.
76. Tabar RS, Wärmefjord K and Söderberg R. Evaluating evolutionary algorithms on spot welding sequence optimization with respect to geometrical variation. *Procedia CIRP* 2018; 75: 421–426. DOI: https://doi.org/10.1016/j.procir.2018.04.061.
77. Voutchkov I, Keane AJ, Bhaskar A, et al. Weld sequence optimization: The use of surrogate models for solving sequential combinatorial problems. *Computer Methods in Applied Mechanics and Engineering* 2005; 194: 3535–3551. DOI: https://doi.org/10.1016/j.cma.2005.02.003.
78. Wärmefjord K, Söderberg R and Lindkvist L. Strategies for Optimization of Spot Welding Sequence With Respect to Geometrical Variation in Sheet Metal Assemblies. 2010: 569–577. DOI: https://doi.org/10.1115/IMECE2010-38471.
79. Asadi M and Goldak JA. Combinatorial optimization of weld sequence by using a surrogate model to mitigate a weld distortion. *International Journal of Mechanics and Materials in Design* 2011; 7: 123. DOI: https://doi.org/10.1007/s10999-011-9154-6.
80. Asadi M and Alsoruji G. A Method to Define the Best Weld Sequence Using a Limited Number of Welding Simulation Analysis. 2015: V06BT06A065. DOI: https://doi.org/10.1115/PVP2015-45556.
81. Bonnaud E. Mitigation of weld residual deformations by weld sequence optimization: limitations and enhancements of surrogate models. *Procedia Structural Integrity* 2017; 5: 310–317. DOI: https://doi.org/10.1016/j.prostr.2017.07.176.
82. Carlson JS, Spensieri D, Wärmefjord K, et al. Minimizing Dimensional Variation and Robot Traveling Time in Welding Stations. *Procedia CIRP* 2014; 23: 77–82. DOI: https://doi.org/10.1016/j.procir.2014.03.199.
83. Michael S, Auld D, Klumpp C, et al. A Robotic Platform for Quantitative High-Throughput Screening. *Assay and Drug Development Technologies* 2008; 6: 637–657. DOI: https://doi.org/10.1089/adt.2008.150.
84. Wang H and Ceglarek D. Quality-driven Sequence Planning and Line Configuration Selection for Compliant Structure Assemblies. *CIRP Annals* 2005; 54: 31–35. https://doi.org/10.1016/S0007-8506(07)60043-2.
85. Lai X-M, Xing Y-F, Sun J, et al. Optimisation of assembly sequences for compliant body assemblies. *International Journal of Production Research* 2009; 47: 6129-6143. DOI: https://doi.org/10.1080/00207540802172037.
86. Xing Y, Chen, G., Lai, X., Sun, J., Zhou, J. *Assembly sequence planning of automobile body components based on liaison graph*. 2007, p.157–164.
87. Hinton E, Wood RD and Ghasemi MR. Optimization of trusses using genetic algorithms for discrete and continuous variables. *Engineering Computations* 1999; 16: 272–303. DOI: https://doi.org/10.1108/02644409910266403.
88. Wade B. *Variable Speed Transmission using Planetary Gear System for High Speed Rotorcraft Application*. 2010.
89. Vanderplaats GN. *Numerical optimization techniques for engineering design : with applications*. New York: McGraw-Hill, 1984, p.xvii, 333 p.
90. Gen M and Cheng R. *Genetic algorithms and engineering optimization*. New York: Wiley, 2000, p.xvi, 495 p.
91. Michielssen DSWaE. Genetic Algorithm Optimization Applied to Electromagnetics: A Review. *IEEE Transaction on Antennas and Propagation* 1997; 45: 343-353.

92. Blickle T, and Thiele, L. A Mathematical Analysis of Tournament Selection, Genetic Algorithms. In: *Proceedings of the 6th International Conference* San Francisco, CA, 1995.
93. Sun W and Yuan Y-x. *Optimization theory and methods : nonlinear programming*. New York: Springer, 2006, p.xii, 687 p.
94. Glover F, and Greenberg, H. New Approaches for Heuristic Search: a Bilateral Linkage with Artificial Intelligence. *European Journal of Operational Research* 1989.
95. Cheng FY and Li D. Genetic Algorithm Development for Multiobjective Optimization of Structures. *AIAA Journal* 1998; 36: 1105–1112. DOI: https://doi.org/10.2514/2.488.
96. Gen M and Cheng R. A survey of penalty techniques in genetic algorithms. In: *Proceedings of IEEE International Conference on Evolutionary Computation* 20–22 May 1996 1996, pp.804–809.
97. Yeniay Ö. Penalty Function Methods for Constrained Optimization with Genetic Algorithms. *Mathematical and Computational Applications* 2005; 10. DOI: https://doi.org/10.3390/mca10010045.
98. Srinivas M. PLM. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics* 1994; 24: 656–667. DOI: https://doi.org/10.1109/21.286385.
99. Grefenstette JJ. Optimization of Control Parameters for Genetic Algorithms. *IEEE Transactions on Systems, Man, and Cybernetics* 1986; 16: 122–128. DOI: https://doi.org/10.1109/TSMC.1986.289288.
100. Jong KAD. *An analysis of the behavior of a class of genetic adaptive systems*. University of Michigan, 1975.
101. Goldberg DE and Richardson J. Genetic algorithms with sharing for multimodal function optimization. *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*. Cambridge, Massachusetts, USA: L. Erlbaum Associates Inc., 1987, p. 41–49.
102. Cant-paz E. *A Survey of Parallel Genetic Algorithms*. 1999.
103. Belkadi K, Gourgand M and Benyettou M. Parallel genetic algorithms with migration for the hybrid flow shop scheduling problem. *Journal of Applied Mathematics and Decision Sciences* 2006; 2006: 17. DOI: https://doi.org/10.1155/jamds/2006/65746.

Chapter 8
The Past, Present and Future of Motion Sickness in Land Vehicles

Muhammad Rehan Siddiqi, Hormoz Marzbani, and Reza. N. Jazar

Abstract The formal understanding of motion sickness has refined over the years, peculiarly in context of land vehicles. Since land vehicles are on-going a transitional phase in technology the perception of how motion sickness impacted passengers of land vehicles in the twentieth century to the twenty-first century has kept evolving in leaps and bounds. The problems that were previously faced or that had not surfaced up earlier will indeed transgress as we move up in the order or level of automation in land vehicles. Hence, this chapter elaborates how motion sickness was perceived in the early versions of land vehicles and how it modified as and when technology in cars made them more ride comfortable, and what we predict can be faced in context to motion sickness of passengers in Autonomous Cars.

Keywords Motion sickness · Land vehicles · Vehicle ride comfort · Vomiting · Drowsiness · Nausea · Autonomous vehicles

8.1 Introduction

Motion sickness is a complex syndrome and is best illustrated by the Poison Theory [1] by Michel Triesman in 1977. Defined by her as a defence mechanism of the body in retaliation against the neurotoxins. In simpler words the so-called poison ingested into the body gives a misperception to our nervous system and makes our brain conclude that motion is being perceived but not being visualized or vice versa. This disconnect gives rise to many symptoms of motion sickness such as nausea, empathy, headache and finally concluding with an obnoxious feeling of vomiting.

To better understand the Poison theory a detail emphasis on understanding the organs of balance is required, our bodies constitute of three main organs of balance (Fig. 8.1); (1) vestibular system presents inside the inner part of our ears, (2) vision is what our eyes view on a daily basis and (3) the proprioceptive sensors present in our

M. R. Siddiqi (✉) · H. Marzbani · R. N. Jazar
School of Engineering, RMIT University, Melbourne, VIC, Australia
e-mail: rehan.siddiqi@rmit.edu.au

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_8

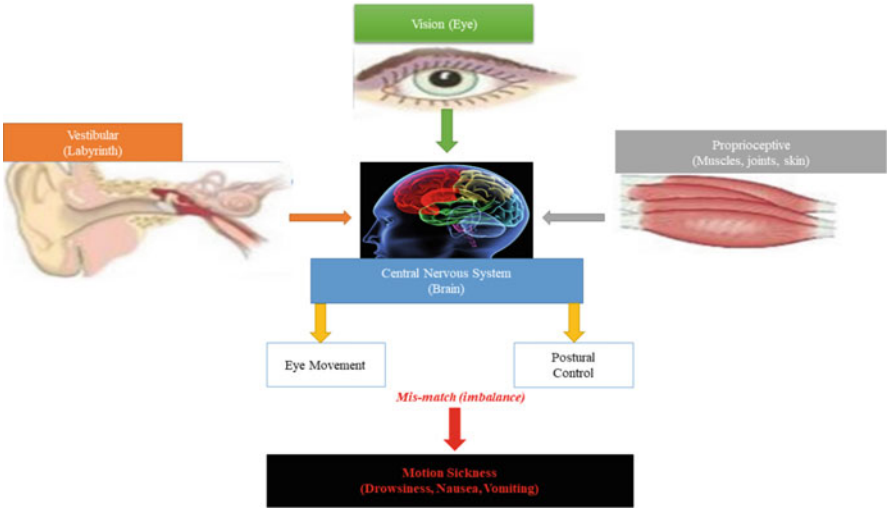


Fig. 8.1 Three organs of balance; vestibular, vision and proprioceptive [2]

muscles, responsible for sensing posture changes in our bodies. These three organs of the body interact with our central nervous system at lightning speeds to give a sense of stability. If any of the organs of balance fails to fulfil its duty the central nervous system develops a mismatch and hence develops a feeling of uneasiness which develops symptoms of motion sickness. The most acknowledged theory in psychology to support this was presented by Reason and Brand as early as in 1970s and was named as the "'motion sickness theory" [3].

The most fitting example of motion sickness in land vehicles has to be reading a book whilst in a car, which is moving at a certain level of speed fluctuations, your ear fluid recognizes this movement signalling to your nervous system about this motion. However, visually your eyes see no movement, while your nerves receive a signal from your eyes, that you are in fact stationary. Hence, a mismatch arises. Making symptoms of motion sickness, such as nausea, vomiting, dizziness, sweating, etc.

Motion sickness is a disconnection of the sensors, which ignites a mismatch in the brain and brain makes you feel sick. Motion sickness can be divided into 3 parts: Type 1—visual present but absence of vestibular and somatosensory inputs. Type 2—Vestibular present but absence of visual and somatosensory inputs. Type 3—Somatosensory input present but absence of visual and vestibular inputs. Although the ultimate manifestation of motion sickness is most definitely vomiting as shown by plenty of studies [3, 4] and [5], but many intermediate symptoms also exist such as Sweating, pallor (dis-colouration), flatulence, burping, salivation, apathy, Nausea, retching. Looking into further detail one can notice greater changes in the human biology while dwelling under motion sickness, as listed below in Table 8.1:

Many studies [6–8] and [9] have described one important facet of motion sickness as "Sopite syndrome", which refers to the profound drowsiness and

Table 8.1 List of bodily changes likely to occur in a person when under the effect of motion sickness

Human feature	Symptom/impact
Cardio	Changes in blood rate + blood pressure
Respiratory	Yawning (alterations in the respiring rate)
Gastrointestinal	Changes in lactic dehydrogenase concentration
Body fluids	Changes in lactic dehydrogenase concentration
Blood	Increase in haemoglobin, pH, plasma proteins
Urine	17-hydroxy corticosteroids, catecholamines
Temperature	Decrease in body temp, coldness
Visual	Ocular imbalance, small and dilated pupils, Nystagmus

persistent fatigue felt by a person as result of prolonged exposures to low intensity motion stimulation, usually yawning has been a potential marker for this syndrome. The syndrome has a lasting effect for up to several hours or even days depending upon the duration of the exposure to low intensity motion. Recent study presented in [10] showed an increase in motion sickness on passengers who were more sleep deprived than passengers who had enough sleep and the study also proved that performance on many tasks deteriorated with increase motion sickness. However, some simple detection tasks like perceptual vigilance seemed unuttered. Earlier research studies by Lackner [11] indicate that, much of the motion sickness is unnoticed and goes unrecognized by people until nausea or vomiting are not elicited.

8.2 What Was Motion Sickness in Land Vehicles of Old?

The earliest evidence of land vehicles causing motions sickness in passengers were mentioned in the works of Robert Barany [11]. In 1921 Barany came up with a term called *railway nystagmus* which in medical terms is known as Optokinetic nystagmus (Fig. 8.2), his investigations notified the back and forth movement of the eyes of railway passengers when looking at the side view from inside the train caused nausea and vomiting. However, it was not until the 1970s when these symptoms were generalized as motion sickness, credit to the motion sickness theory.

Many theories have been developed over the years some of these theories like evolutionary theory [29] and [30], ecological theory of motion sickness [31] (ecological theory of orientation) and sensory conflict theory of motion sickness [32] will be briefly mentioned here. Evolutionary theory essentially states that motion sickness is a response to poisoning, for example, a mismatch between two sensors of the body produces a mismatch or imbalance in a person's body. The person's body takes this difference and produces an involuntary action to exit the body, so this so-called poison needs to be removed from the body, which causes actions like vomiting to take place.

Fig. 8.2 Illustrates Optokinetic nystagmus testing setup

Ecological theory is based on the hypotheses that, postural instability causes motion sickness this theory was developed from a study presented in [33] subjects were blind folded while using a joystick to keep themselves upright on a device mimicking the inverted pendulum behaviour. The subjects were affected with the device's upright and could not make out if they were upright or not. These results indicated that, dynamic balance influences the perceived upright. This ecology holds true when a passenger is sitting in a car or plane, for example, the movement of the car signals to our brain, implies that we are moving when in fact we are stationary, but the relative motion of our car gets perceived as motion. Lastly the sensory conflict theory is the most widely accepted theory of motion sickness. Agreed by many studies that, any elicit motion induces motion sickness. The main theme of this theory relates voluntary commands such as vision, touch, hearing and vestibular activity with musculature movements like arm movements or body movements. Almost all of these movements also affect the soft tissues of body.

Alongside with these theories many remedies have been suggested by studies to lower the risk of having motion sickness one such way is to avoid exposure to provocative situations entirely [34]. Another remedy can be to gradually introduce exposure to such provocative situations which is found to be a very effective solution in reducing the susceptibility of motion sickness [35] and [36]. This fact has been proven in the long series of experiments held in the slow rotation room (Pensacola) where it was shown that, by making head movements in the beginning at very low velocities of rotation and then progressively increasing the head movement rotations

to higher velocities a person can adapt up to rotations of 25 rpm without eliciting any symptoms of motion sickness. Another remedy can be the use of anti-motion sickness drugs. These drugs can help in reducing symptoms even when higher levels of stimulations are experienced [37]. But unfortunately, these drugs are central nervous system depressants and induce drowsiness and known to cause placebo effects where the brain convinces an individual into a fake treatment. However, these remedies (and many other not mentioned) need to reduce their undesired effects such as drowsiness.

In 1970s majority of the standards and investigations revolved around the motion sickness theory. Most work done in that era was in the field of psychology understanding what may aggravate motion sickness in people. Psychologists loosely blamed vibrations felt by land vehicles due to road or railway induced vibrations as the number one cause in aggravating motion sickness in passengers. British Scientists in the late 1980s become the very first people to specify a manual (called the British standard BS 6841:1987 [12]) relating vibrations and shock felt in land vehicles to human health and more importantly to motion sickness. Standards presented in BS 6841:1987 initiated a gap in research that required further investigations to be carried out. This led to a sharp rise in experiments conducted to investigate the impacts of vibrations on posture that eventually led to motion sickness. The investigations were put together by Riccio and Stoffregen, who came up with a new unknown theory which they co-named the "postural instability theory" [13]. The theory proposed loss of posture of a passenger as being a necessary and sufficient condition preceding motion sickness. The theory backed with experimental results made a lot of reasoning sense in some cases of land vehicles but not in all cases. Such as why was it when a person would be driving on a straight highway, subjected to reading a book would still undergo motion sickness?

These and many more instances raised doubts in scientists if postural instability was let alone the only cause of initiating motion sickness in people specifically in land vehicles. In the mid-90s Kennedy and Stanney answered these questions and let alone cleared many doubts about motion sickness. There investigations and findings formulated the problem as indicating, that both motion sickness and postural instability were secondary order effects generating from a single source [14]. These theories and standards addressed, required attention from engineers to further investigate the root cause of motion sickness aggravating in land vehicles, as land vehicles had transitioned into Level 2 automation [15].

The occurrence of motion sickness has a profound effect on human beings during road transport. In an early study regarding motion sickness [38] the authors suggested motion sickness being linked with long unbroken rides. Especially associating these rides on winding routes and not over rough roads. In a laboratory study presented in [39] and [40] for sea showed that, vertical oscillations conducted at low frequencies (below 0.5Hz) caused sickness, however, study presented in [41] discredited this claim with respect to most vehicles on roads and credited the impact of motion sickness in road transport to low-frequency horizontal motions.

Further impactful researches [42] and [43] produced comparative studies between vertical and fore-and-aft accelerations of similar magnitudes and

frequencies. Results indicated that, fore-and-aft accelerations significantly to be more nauseogenic than vertical motion. These results have been further supported by other studies presented in [44] and [45]. However, these studies were not conducted under typical normal driving conditions, leading to larger scale of studies into this field.

Drivers rarely deemed to fall victim to motion sickness as compared to passengers a fact supported by studies of late [46, 47] and [48], the following studies also indicated a few key factors of producing low-frequency fore-and-aft and lateral accelerations such as; the driving behaviour of the driver, route taken by the vehicle, vehicle characteristics and speed of the vehicle.

Griffin and Turner provided with majority of these investigations and elaborated the vehicle parameters that lead passengers in motion sickness. Conducting variety of big scale experimentation on land vehicles during the same era in which the British standards were formalized. However, the key research which investigated over 3000 land vehicle passengers for elaborating the impact of drivers, routes and vehicles on motion sickness surfaced like never before [16]. Classifying the forces emerged during cross country roads as the most aggravating manoeuvres of land vehicles that lead to motion sickness in passengers. Further works of Griffin in the following years led to greater insights into the frequencies and phase relationships of lateral acceleration and roll. It was determined that weighted frequency range from 0.1 Hz to 0.5 Hz and in-phase relationship of roll and lateral accelerations were critical in impacting motion sickness in passengers.

They investigated the impact on passengers' motion sickness in coaches with regards to different routes, different vehicles, different vehicle forces, different varieties of driving style and the seating position of the passenger in the coach. Private buses or coaches were hired to investigate the motion sickness felt by 3256 passengers, which yielded over 110 h of travelling. All journeys included a combination of routes from highways to cross country roads. A total of 5 different vehicles were used with 17 different drivers driving those vehicles. Passenger motion sickness was measured by a means of questionnaires given to each passenger just 5 min before the end of the journey. In total six vehicle forces were measured and recorded (3 translational and 3 rotational), in the x, y and z directions by means of an accelerometer. Over all the readings were summed for over 56 journeys to find a relationship of each key factor with motion sickness.

Results indicated that the highest symptom of motion sickness felt by most of the passenger was feeling hot (28.5%) and the least was vomiting only felt by 1.7% of passengers. The power spectral density functions were calculated for each of the forces measured. At frequencies below 0.5Hz fore-and-aft and lateral accelerations were the highest translational forces and yaw being the dominant one in the rotational forces. These forces were overall higher in cross country roads as compared to motorways during the 56 journeys recorded. Statistics showed very high positive correlation between illness ratings and lateral and fore-and-aft forces, with lateral forces producing the highest correlation.

Different vehicles did not produce any significant values or correlation with respect to motion sickness. However, seating positions showed a lot of variations.

People Sitting in the front had the lowest motion sickness as compared to people seated in the middle and the worst being seated at the rear. The study concluded by stating low-frequency lateral accelerations, felt during cornering (cross country roads) with passengers sitting at the rear end of the vehicle were the most likely to undergo motion sickness in vehicles and suggested to educate drivers to take slower and more relaxed turns during journeys to ensure a lower levels of motion sickness felt by passengers during journeys in buses or coaches.

An experimental study on low-frequency motion cars [49] further investigated the impact lateral acceleration has on motion sickness. The following four studies were analysed for vehicle forces and motion sickness: 1. Repeated measurements with one driver and one car. 2. Variability in measurements between vehicles. 3. Variability in measurements between drivers. 4. Comparison of three types of vehicles.

In the first study a single driver was used for driving a car over a suburban route repeating this test 40 times in total. The power spectral densities of each x, y and z were captured. The power spectral density for lateral and horizontal accelerations resulted in very high values of about 5–12 $(ms^{-2})^2$/Hz over the lower frequency range of less than 0.1 Hz. Vertical acceleration overall produced very low power spectral values of about a maximum of 0.15 $(ms^{-2})^2$/Hz. Even the Motion sickness dosage values (MSDV) were higher for horizontal and lateral accelerations (mean value of 19.3$(ms^{-2})^2$.s and 16.5$(ms^{-2})^2$.s, respectively) as compared to vertical acceleration (a mean value of 1.5$(ms^{-2})^2$.s). The results for each test drive were almost identical ensuring that the same driving style and route do ensure the same levels of acceleration and motion sickness. Also, since the MSDV and acceleration values recorded were reasonably high, indicating roads with a lot of turns and junctions (such is the case for suburban areas in the UK) are most likely to produce higher motion sickness levels in passengers.

In the second study a single driver drove 10 different cars on the same route. Vertical acceleration and MSDV values due to vertical acceleration produced yet again very low values, but with huge variations amongst each car (due to mostly varying suspension mechanisms amongst cars). The mean results produced for x and y MSDV values were 20.7 $ms^{-1.5}$ and 19.3 $ms^{-1.5}$. Again, recorded values for the horizontal and lateral accelerations were very high (maximum of 21$(ms^{-2})^2$/Hz and 5$(ms^{-2})^2$/Hz, respectively). Variations in the results do indicate that different cars produce slightly different values of MSDV and accelerations, but the deviation for horizontal acceleration is much more than the lateral accelerations indicating that, cornering is mainly influenced by the style of driving and braking/accelerating coincides with the weights associated with each car.

In the third study a total of eight different drivers were used to drive the same car on the same route. As the general trend seen in the other 2 previous studies lateral and horizontal accelerations were higher and same goes for MSDV values as well, but the interesting part of the results was that this time around the variations in the accelerations and MSDV values was a lot more evident than in the other 2 studies. So much so that, the standard deviation for x and y MSDV values was 4.8 $ms^{-1.5}$ and 3.3 $ms^{-1.5}$, respectively. The variations produced in the results for the third

study were almost double that, of the first and second studies. Indicating driving behaviour impacts highly the motion sickness levels in a passenger.

In the final study a comparison was made between 2 manually driven cars and a single automatic car. Once again values for MSDV and accelerations for lateral and horizontal were very high and low for vertical acceleration. Also, almost all readings were identical. An indication that, all cars in general produce high enough accelerations to cause motion sickness so it is highly unlikely that, a passenger stays clear of falling motion sick in one car as compared to another. The study [49] has drawn some important conclusions, for example, suburban areas produced high levels of lateral and horizontal accelerations and frequencies lower than 0.1 Hz that, can lead to high motion sickness levels in passengers and that, these values are more dependent on the behaviour of the driver than the characteristic of the vehicles themselves. Such as driving with higher speeds and acceleration around turns.

To further the discussion, roll and lateral acceleration and their impact on motion sickness a study presented in [50] needs to be thoroughly analysed. Most studies prior to this one studied motion sickness when both lateral acceleration and roll were combined and in phase, but this particularly studied the impact of having phase difference between roll oscillation and lateral acceleration. At intervals of 1 week 20 subjects were tested in a cabin with 30-min exposures of combined 0.2 Hz sinusoidal lateral acceleration and roll oscillations. Roll oscillations were differed in phase every week by a phase difference of; (a) 0° (week 1), (b) 14.5° delay (week 2), (c) 29° delay (week 3) and (d) 29° advance (week 4), during the exposure subjects gave their sickness ratings at intervals of 1 min. The results indicated maximum sickness levels were experienced at 0° delay, lower sickness levels were found at 14.5° and 29° delays, but the lowest sickness level were found to be 29° advance. Indicating that, sickness level reduces with either a delay/advance in roll oscillations with respect to lateral acceleration, with lowest sickness levels found to be in advance roll oscillations.

Other studies indicate [51, 52] and [53] roll oscillations on its own is not usually highly provocative of motion sickness specially at lower frequencies. In study presented in [52] found that, ±8° of roll oscillations provoked low levels of sickness, even when these tests were done on different frequencies the result produced low level of sickness. Simply indicating that, if the angular displacement is the same at different frequencies the frequency of roll oscillations has no impact on motion sickness of the passengers, but when roll oscillations are combined with lateral acceleration and when specifically, lateral acceleration is fully compensated with roll oscillations motion sickness increases [50, 54, 55] and [56].

In another study by Griffin [57] he investigates the relationship of roll and pitch magnitudes on motion sickness. 120 subjects were divided into groups of 6 and were exposed to 30 min of 0.2Hz sinusoidal roll or pitch oscillations in a closed cabin during which, there sickness levels out of a maximum of 7 (indicating the most sickness). Results suggested that, over the six symptoms of motion sickness mild nausea was the highest reported symptom. Oscillations at the lowest magnitude (±1.83°) produced the lowest rating in the sickness levels as for the intermediate magnitude (±3.66°) and the highest magnitude (±7.32°) levels of motion sickness

were higher than the lowest magnitude. However, in all the three scenarios the sickness levels had no significant difference and overall the highest level of motion sickness value recorded was almost 2. Indicating that, roll and pitch oscillations both do show a trend of increasing motion sickness with an increase in magnitude, but reach a saturation point since there was no increase in motion sickness levels between intermediate to the highest magnitude roll and pitch oscillations.

The results about roll and pitch oscillations demonstrate that the probability of motion sickness and severity of symptoms increases with their magnitudes, but the increase is either not to strong or that apparent at times [57]. One key factor helping the subjects to lower their motion sickness is that most of these experiments allow the subjects is to gain control of their postures such as moving the head or the upper body towards the backrest or headrest a study presented in [58] has found that a backrest is enough for the passenger to reduce his sickness levels. So most of these subjects presented a natural reaction to these magnitudes to reduce motion sickness or discomfort caused due to postural instability.

Based on these invaluable findings a Human-Comfort Factor Map (HCoM) was proposed based on human preferences and a framework for a safe and comfortable navigation of passenger vehicles implementing the algorithm, which is a Graph-based search path planning algorithm [59]. Regarding the methodology and considering the comfort state space, four major human-comfort factors were integrated into a geometrical map generated by SLAM framework in order to build a comfort map. Additionally, participants were instructed to drive an autonomous robotic wheelchair in two distinct experiments aiming to extracting comfort factors and building the comfort map based on human preferences. As a result, it was revealed that reducing the travel time, or increasing the velocity and acceleration (negotiating the shortest path), does not necessarily increase passenger's comfort.

It was not until the early twenty-first century that lateral acceleration magnitudes for discomfort were defined [17]. The range of 1.8 m/s^2–3.6 m/s^2 was defined as the magnitude of lateral acceleration that produces the most comfortable ride. Anything beyond this level would either produce medium discomfort and anything more than $5m/s^2$ would lead to lethal uncomfortable rides. An experimental study in China utilized these comfort thresholds where roads were classified into 3 groups (6 lanes, 4 lanes and 2 lanes) [60]. Six lane roads produced about 98 percentiles of their lateral accelerations' below 1.8 m/s^2 with a very few values above the mark of 1.8m/s^2 with the maximum reaching almost the 3.5 m/s^2 mark. Four lane roads produced about 91 percentiles of their lateral accelerations' below 2 m/s^2 with quite a few (more than 6 lane roads) more above the 2 m/s^2 mark with the maximum reaching almost the 6 m/s^2 mark. The poorest performance in terms of comfort with respect to lateral acceleration was by 2 lane roads its roads produced about 50 percentiles below the 2 m/s^2 mark with highest lateral accelerations reaching up to 8 m/s^2.

The radius alongside the lateral acceleration was also investigated, in general as the radius increased the lateral acceleration of the roads decreased. Since 2 lane roads had smaller radiuses (majority of the roads less than 500 m radius) there lateral accelerations were higher. As for 4 lane roads there, roads had a more fifty-fifty balance with majority of the roads having a radius between 100–1000m hence

having lateral accelerations mainly between 0–2 m/s^2. Lastly, the 6 lane roads had majority of their road radius above or around the 1000m mark hence they produced the lowest lateral accelerations.

The driving speed was also found to be related inversely with the lateral acceleration meaning the higher the driving speed the lower the lateral acceleration experienced by the vehicle. Overall, the mean value of lateral acceleration for 6 lanes was 0.7853, for 4 lane roads was 1.1246 and for 2 lane roads was 1.9934. These results clearly indicate the problem for motion sickness would increase on highways with smaller number of lanes because since lateral acceleration increases, radiuses are smaller with decreasing levels of comfort, so it is expected that the motion sickness levels will start increasing.

The studies of until the beginning twenty-first century started the movement of understanding the forces of mechanics that are present in the land vehicles that increase motion sickness tendencies in passengers. As robotics and automation in cars was exceeding in leaps and bounds and entering the Level 3 automation extra research was required in understanding ways of preventing the rise of concerns with motion sickness in land vehicles. The next part of this chapter elaborates the present scenario of motion sickness in land vehicles as we dwell into Level 3 automation of cars.

8.3 What Is Motion Sickness in Cars of Today?

At the present moment we are dealing with Level 3 automation in land vehicles where features like cruise control, automatic braking systems, parking sensors, fuel efficient engine injectors and many more such technologies are making driving a less tedious task. But are these tasks reducing the impact of motion sickness in passengers of land vehicles? A question that we look to answer in the following parts of this chapter.

Technology of today presents these experimentation's to be done via simulations rather than asking passengers to sit in a car and report these motion sickness ratings. Dynamic simulators are often used to read human behaviour (like motion sickness), self-motion perception and steering angle effort by simulating the motion perception in humans through the stimulation of the vestibular system, somatosensory systems and vision [61].

Unfortunately, due to the multi-sensory stimulation of the sensors during dynamic simulation the task of optimizing the simulator is highly complex, for example, in studies presented in [62, 63] and [64] the motion produced because 1:1 simulator is overestimated. An idea presented in study [63] allowed some technical techniques to be developed to avoid this overestimation such as scale factor (also called a motion gain) or tilt-coordination (combination of tilt and translation). However, both these variables were found to be highly dependent on the level of simulated acceleration [64] and [65].

Fig. 8.3 SHERPA simulator used for automotive studies [118]

Lateral and longitudinal movements differ a lot meaning, the simulator for a lateral manoeuvring is far more complex than for the longitudinal manoeuvring. Since lateral manoeuvring adds rotation movements like yaw and roll motions. In a study presented in [66] motion components like yaw, roll was considered and were found to be highly dependent on the steering angle of the car while cornering, in that very study the results suggested that motion gains for lateral and rotational acceleration should be within the range of 0.4–0.75. The studies limitation was that, the lateral accelerations of the simulator were kept the same throughout the simulations.

This limitation was overcome by the study presented in [67] the respect roles of lateral acceleration and lateral, roll and yaw motions for cornering behaviour on a dynamic driving simulator were studied. Generally, the motion gain used for dynamic simulators is kept the same for all types of different paths in a simulator. The present study [67] further investigates the cornering behaviour with regards to different motion gain settings (a scaling factor for simulating the multi-sensory organs like vestibular system, somatosensory system and vision).

All the simulation tests were done using the dynamic driving simulator SHERPA [68], 27 volunteers took part in this experiment between the age of 22 and 49 with none of them having significant experience with the driving simulator. The path designed for the simulator was a slalom (two lane track) with cones (to create the sinusoidal pathway) placed on both sides of the road to limit the person from going off the track. The velocity of the car was kept constant to 70km/h with three varying scenarios of lateral accelerations; (1) 1 m/s^2, (2) 2 m/s^2 and (3) 4 m/s^2. With post spacing varying for each scenario (scenario 1—86.39 m, scenario 2—61.09m and scenario 3—43.19 m). Drivers of the simulator were told to perform a slalom course and told to stay with in the path made with the mini cones. A total of 25 different varying motion gains were applied for lateral, roll and yaw gain under each lateral acceleration scenario. So, a grand total of 75 tests were carried out with the dynamic simulator (Fig. 8.3).

All participants started the simulator with a familiarization stage of 10 min (driving through a rural area). Each session was then conducted with 25 trials and after a few hours the next session of 25 trials was conducted and then after a gap of another few hours the third trial was conducted. All participants after completing the trials took part in subjective evaluation of the realism of the dynamic simulator and motion sickness susceptibility questionnaire. During the experiment only 4 subjects felt sick enough to stop the simulator. These all occurrences occurred during the highest lateral motion gains; however, the remaining participants were able to complete the trials with a misery average score of about 1.2. After performing Principle component analysis (PCA) on the subjective evaluation of realism of the dynamic simulator, the result reached no consensus hence the data was centred to identify 2 homogeneous groups of G1 and G2. Group G1 results indicated that, for a realistic simulator the lateral motion gain was proposed to equal to 0.5, roll motion gain equal to 0.5 and yaw motion gain equal to 0. As for Group G2 the values proposed were 0.85 for lateral motion gain, 1 for roll motion gain and 0 for yaw motion gain.

As for the objective analysis the study [67] calculated the steering wheel reversal rate using PCA. A consensus was reached for the slalom levels. Lateral motion gain was the main significant factor for all the slalom levels the participants took. For the trial one conducted at 1 m/s^2 that, the number of reversals decreased with an increase in lateral motion gain, which basically meant that more steering corrections were required at lower lateral motion gains. The analysis also suggested that, roll and yaw gains had no effect on the driving performance of the participants.

The second and the third trials (2 m/s^2 and 4 m/s^2, respectively) results showed that, as lateral acceleration increased the lateral motion gain had to be reduced because the best Steering wheel reversal rates were found to be 0.5 and 0.25, respectively, for the second and third trails. As for the roll and yaw motion gains both indicated a value for 1 would be best in the second and third trials.

Furthermore, objective analysis was done on the results using the path root mean square error. It was found that, once again the main component impacting the driving accuracy of the participants was the lateral motion gain. What is more interesting was that, extremes of lateral motion gain (either very close to 0 or 1) the drivers took wider trajectories to accomplish the path successfully with a lot of difficulty. The values between 0.25 to 0.5 the participants looked in greater control than whenever the lateral motion gain was outside this range. From this study it was concluded that, three motion gains should not be set to the same value when simulating the cornering manoeuvre and should be varied according to the lateral acceleration levels, for example, the major component lateral motion gain should be reduced with an increase in lateral acceleration. As for roll motion gain a value of 1 was perceived as the most realistic driving simulator. However, yaw did only slightly influence the final perceptions of the participants and so it seemed that, yaw was masked behind the two major components of lateral and roll gains.

The usage of such experimentation to generate mathematical models has been widely used in studying postural instability in passengers in order to do simulations instead of actual live experiments. One such mathematical model presented in the

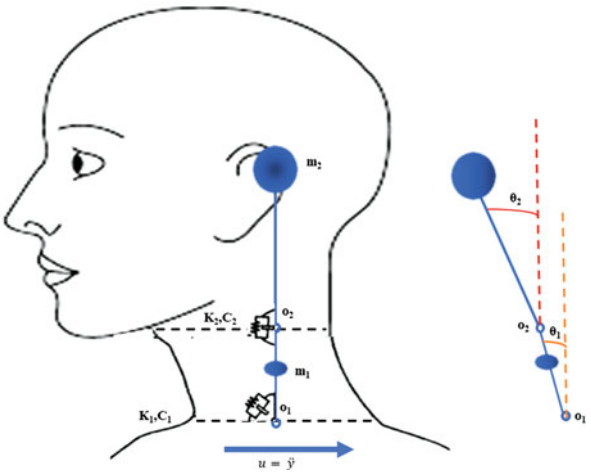


Fig. 8.4 The head neck complex (HNC) modelled as a double inverted pendulum [69]

study [69] developed a head neck complex (HNC) model and tested it with its response to trunk horizontal vibrations. The model is based upon a double inverted pendulum such as shown below (Fig. 8.4):

As seen the HNC model is a combination of 2 parallel connections between a spring and a damper with these parallel connections in series to each other with 2 different masses at the ends of each pendulum labelled as m1 and m2. Tests were carried out on 4 subjects and their results compared with the HNC model when subjects were exposed to the same trunk horizontal vibrations as the HNC model. The comparison of the results was almost identical indicating that, a lot of time, money, health and accuracy can be achieved by just simply testing different vehicular forces like lateral acceleration, roll, pitch, etc. with such accurate models depicting our body. And further experiments applied to these models can help us predict how do vehicle forces (which impact or increase motion sickness) impact postural instability to cause motion sickness in people without any discrepancies.

The very first practical model representing motion sickness based on the theory of subjective vertical concept (a concept first introduced by Bos and Bles in 1998 [18] and [19] based on motion sickness theory) was presented by Wada and Kamaji from 2010–2015 [20]. The model comprises of the inner vestibular system of the human into different control blocks as shown in Fig. 8.5.

The model contains two major organs of the vestibular system that contribute towards motion sickness, namely otolith (OTO) and semi-circular canal (SCC). The model takes lateral acceleration and the angular velocity felt on the head of the passenger to calculate the motion sickness impact felt on the passenger due to subjective vertical conflict. The importance of having such a model is that it, firstly it allows results of motion sickness to be investigated through simulation as the values generated by the model showed high correlation with passengers actually tested on an experiment done on road by the same team who made the model in the first place. Secondly, the model clarifies how vestibular system (present inside the inner ear)

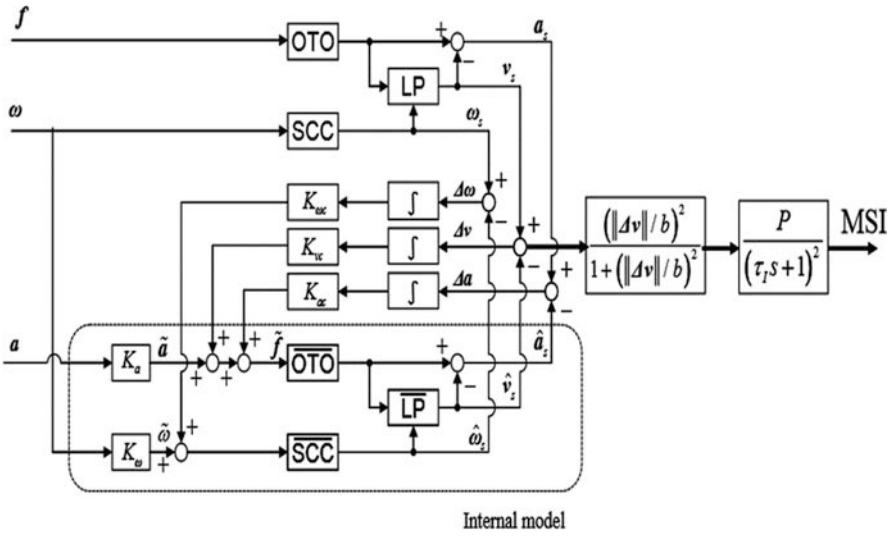


Fig. 8.5 The model of Kamaji based on 6 degrees of freedom subjective vertical conflict theory [20]

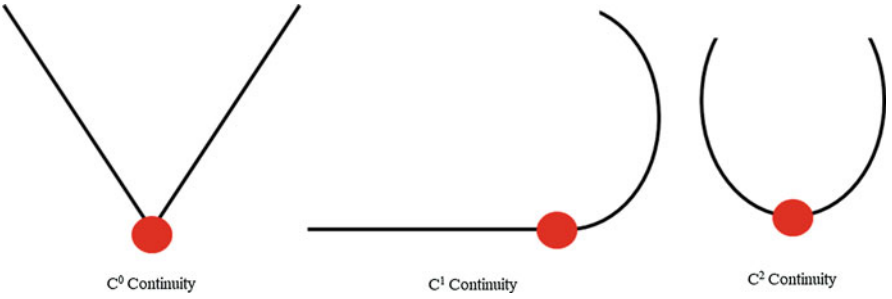


Fig. 8.6 Example of C0, C1 and C2 continuities

develops to build up motion sickness symptoms in humans and more importantly indicates that lateral acceleration and the angular velocity felt by the person and on his head are the main critical factors increasing the incidence of motion sickness in passengers.

Later, in 2015, talks of creating road tracks that produced continuous curvature was taking place to improve ride comfort for land vehicles through handling comfort of vehicles. One such design for betterment of roads was presented by Hormoz using clothoid [70], also known as spiral curve and able to produce curves with C0, C1 and C2 continuity (Fig. 8.6).

Although clothoid produced curves with continuous curvature but due to them not having a closed form and extreme computational complexities the idea of using them for road design for better ride comfort was not quite possible.

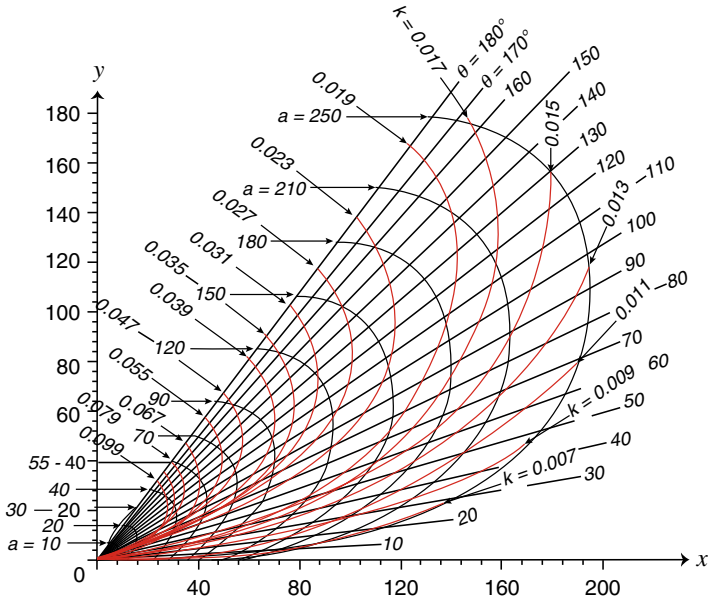


Fig. 8.7 Road design graph for cloithoid presented in [70]

The study presented in [70] simplified the clothoid by defining the Fresnel integrals using a magnifying factor 'a' and developing a design graph of the relationship between the clothoid and parameters of magnification factor a, curvature k, and slope h. Using this graph (shown in Fig. 8.7), the complexity of designing a clothoid for mobile robots was reduced, however, the time response of the system was too long to be practically used in real time for mobile robots or autonomous vehicles.

Now days interpolating curve planners like Bezier and B-spline/NURBS are used [71] some works of these curves are present in navigating an autonomous mobile robot a study presented in [72] used NURBS to define the trajectory for an autonomous mobile robot. Initially the 3D environment of the robot was captured via a DIS radar then these objects were represented in the computer using AUTOCAD and then applying Principle component Analysis (PCA) these objects were defined in planar view (2D). A Path from start to goal was generated using parametric curves then the discontinuities were removed using NURBS to define a smooth, and optimal path for the robot to reach the goal while avoiding obstacles during the course.

Robot motion planning done using NURBS in a dynamic environment is presented in [73]. A collision free path for the robot was designed using NURBS. In this simulation-based solution obstacles were defined as polygons and a mathematical model or path for the robot was designed using vector addition to avoid all obstacles in its course. NURBS and its offsets were utilized to smooth the curve and to ensure that the final path is a 100 percent collision free path. Additionally, the algorithm

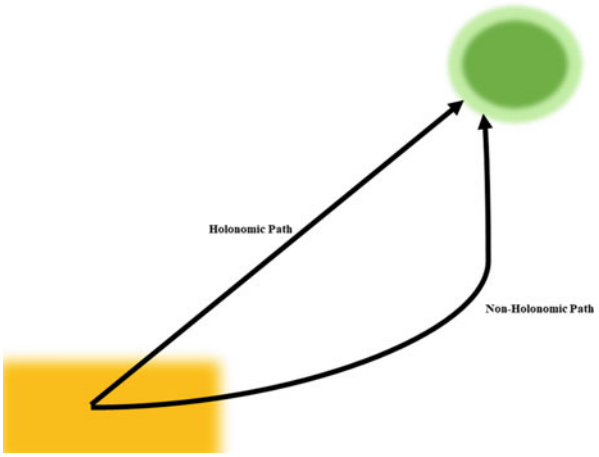


Fig. 8.8 Difference between a holonomic path and non-holonomic path

produced can detect dynamic changes made to the environment as the number obstacles is not defined in the mathematical model.

Furthermore, NURBS have been successfully used for path planning and generating a smooth path from the starting point with a certain orientation to the destination point with different orientation for a humanoid robot in the study presented in [74]. In their objective they strive to make the movement of the humanoid as smooth as possible with lowest number of jerks to make the movement of the humanoid as close as possible to a human movement.

Most of the robots mentioned above are based on holonomic movements, but as we very well know a car does not support holonomic movements (Fig. 8.8), which basically means that it is not able to translate in the direction of y-axis by rotating on its own axis and so its integral to see studies which use parametric curves like Bezier and B-spline/NURBS on cars or at least car-like robots. One such study presented in [75] uses NURBS to optimize and smoothen the paths produced using S-roadmaps. Surfaced roadmaps were used to produce the shortest path for car-like robots and then later using NURBS these curves or paths were optimized and smoothened out to reduce the discontinuities existing in the path.

A study presented in [76] practically shows the benefits of using B-spline over any other curve. In this study the author designs an algorithm using B-spline curves which, produces a continuous and smooth path for car-like robots. Firstly, he starts with a discussion clarifying the need to use parametric equations over standard cubic curves like clothoid, since clothoid lacks an efficient approximation method (increasing the computational complexity of the program) the author relies on these parametric curves to produce a definite approximation of the path. Secondly, he practically justifies his selection of B-spline curve over Bezier curves by plotting the two curves on a zigzag pattern. The result shows that b-spline curve produces a path which is much smoother and continuous than a Bezier curve. Additionally, the B-spline curve followed the path much more closely than the Bezier curve. Another

very important and valid point of having to use B-spline was its local control ability as compared to Bezier curves global control, which is most suitable when designing robot paths in dynamically changing environments.

The study further clarifies the need of using B-spline curves with a weightage of 1 across all its points (a NURBS curve with weight equal to 1 is a B-spline curve) and using a 3rd order cubic B-spline instead of any other order. To keep the algorithm simple, the authors choose equal weights across all the points (or weight equal to a one) and since 2nd order B-spline curves produced a very high curvature value as compared to a 5th order B-spline curve. It was clear that having a high curvature would produce discontinuous paths with uncertainty in its smoothness. Hence it was recommended to use higher orders of B-spline curve to produce better results but, since higher order B-spline curves will increase the complexity of the program and having too low of an order will increase the curvature to that extent that the path will not remain smooth and will produce discontinuity thus an order of 3 was chosen, which was a value between the 2nd and 5th order B-spline curves. However, the authors do suggest that more investigation can be done on behalf of the order of B-spline and their weightages.

Once the selection for designing the curve had been justified the next step was to prepare the algorithm. So, an algorithm was based on a 3rd order B-spline curve but to reduce the maximum curvature to a certain limit mid-point insertion were proposed and the proposed algorithm in the study was finalized as a 3rd order B-spline curve using mid-point insertions to smoothen the curves even further. This fact was proven by the authors by testing the 3rd order b-spline curve independent of the mid-point insertion, for which the curve profile was very high and then testing the 3rd order b-spline with mid-point insertion, for which the curvature profile was half of the curvature profile produced by the 3rd order b-spline only. So, this insertion of mid-points allows the 3 points on the curve to be very easily changed into 5 points without increasing the order of the curve, that helps in saving computational complexity and maximum curvature profile.

Lastly the developed algorithm in the study was tested against; (1) previous curves used in studies [77, 78] and [79], (2) combined with path planning method called the Rapid exploring random tree (RRT) and (3) Tested with actually human steering using a small scaled car-like robot, undergoing various manoeuvres such as seven segment turns [77], narrow passage [78], a random zigzag pattern [79], lane changing, cornering, obstacle avoidance and clutter manoeuvring.

The Results indicated that as compared to all the previous curves used in the studies presented in [77, 78] and [79] the b-spline algorithm proposed in the study [76] produced smoother paths with continuity and did not exceed the maximum curvature limit imposed. As the for the path planning using RRT was successfully implemented with the b-spline algorithm on the narrow passage path producing smooth and continuous path and path did not exceed the maximum curvature imposed by the path. Also, the computation time of the RRT with the proposed algorithm was very relatively very small (80 milliseconds). Lastly the result for testing the proposed algorithm on a car-like robot produced successful results in terms of producing smooth, continuous and most importantly producing almost

identical steering angles to the human steering. Concluding that the proposed algorithm very closely resembles the steering done by a human himself or in simpler words the proposed algorithm is capable of mimicking human steering when undergoing different manoeuvres like lane changing, cornering, obstacle avoidance and clutter manoeuvring.

The same authors this time added further to their work in study [76] by adding field results to their proposed algorithm. This study presented in [80] proposes a novel manoeuvre planning method that results in attenuating disturbances, avoiding abrupt changes in steering angle and longitudinal velocities whilst reducing high frequency disturbances, lateral acceleration and yaw disturbances.

Results of this proposed algorithm were first compared against a previous study presented in [81] the proposed algorithm was used to draw a path on a shape that consisted of successive alternating turns (increasing the complexity of manoeuvring) while the study presented in [81] used on the same path a Bezier curve. The result showed that, the proposed algorithm which used B-spline curves produced much smoother and continuous paths, also when comparing the frequency response of the two curves the Bezier curve's response showed disturbances existing in the higher frequency as for the proposed program had eliminated these high frequency disturbances.

The proposed algorithm was then compared with a study presented in [82], the manoeuvre this time was the famous lane changing manoeuvre. The previous study used a sinusoidal function to execute the lane changing manoeuvre, the results showed almost identical results in terms of acceleration and velocity, but unlike the sinusoidal wave the proposed algorithm has the benefit of being used for re-planning of paths for various manoeuvres including the lane changing manoeuvres for which the proposed algorithm did very well to produce almost identical results.

The third comparison of the proposed algorithm was made against a study presented in [83] the roundabout manoeuvre was used. The previous study [83] used simply a straight line and arc to design the manoeuvre. Once again, the proposed algorithm showed a reduction in lateral acceleration and side slip. Lastly the B-spline planner algorithm was combined with a controller to autonomously drive an experimental vehicle to carry out field testing for numerous manoeuvres like lane changing, obstacle avoidance and turning. The Numerical simulation results and the experimental field test results both showed similar lateral acceleration profiles and peak values, however, the slight differences were due to the road disturbances measured while driving the autonomous experimental vehicle, but overall the proposed method and the field results agree to each other.

A comprehensive review presented in this region in [84] on different motion strategies in context of autonomous vehicles with more emphasis on improving safety, comfort and fuel consumption optimization. It was demonstrated that all of path planning strategies could be classified into four major groups, based on the way of their implementations, named Sampling, Interpolating, Graph-bases search and Numerical Optimization. Finally, a brief description of each technique, their contributions to different research teams in terms of motion planning have been summarized below:

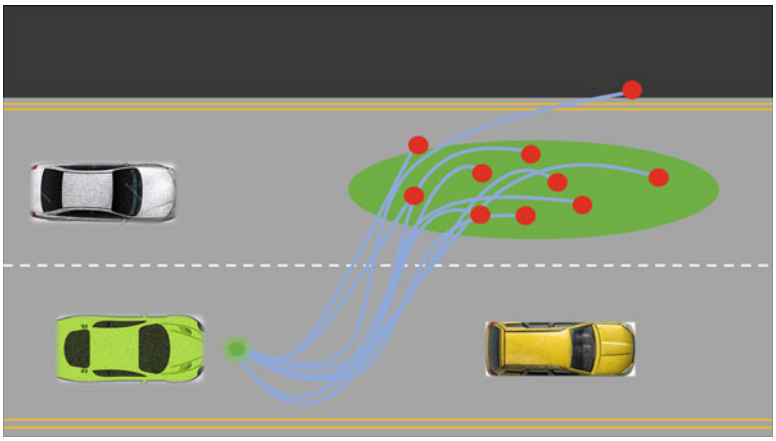


Fig. 8.9 Imaginary illustration of how path planning is done using RRT for a single lane change manoeuvre

- Graph Search Based Planners: Classified into 2 main groups, namely (1) A × Algorithm Family and (2) State Lattices. The A algorithm utilizes a heuristic function for searching making it reduce computational time, but resulting on non-continuous paths. As for state lattices it can handle multiple dimensions such as velocity and acceleration, but usually the solution is quite costly and based on a very low resolution.
- Sampling Based Planners: Belonging to the Rapid exploring Random Tree (RRT) family, are able to quickly find a solution and are suitable for local and global planning. However, since the patterns are randomly generated the paths are not smooth and very jerky leading non-continuous paths.
- Interpolating curves: Four major curves belong to this group, namely (1) Cloithoid, (2) Bezier, (3) Spline Curves and (4) Polynomial curves. We have already mentioned about the major advantages of theses curves in detail earlier and their potential is much more directed towards minimizing motion sickness unlike other groups of motion planning strategies (RRT or Graph-based). Hence, it is safe to say that once a path has been established using RRT or Graph-based approach then applying these interpolating curves for limiting motion sickness will be inevitable (Fig. 8.9).

Thus, it is paramount to learn about a few critical path planning studies that have improved the path planning process. A study presented in [85] provides a good overview of the path planning techniques used up to date. The overview starts by stating some earlier research presented in [86–88] that attempted to finding an optimal path by representing the map into an appropriate free space. So simply these researchers defined a static environment of the map with cluttered objects and then applied Quadtree representation to represent their environment or map. Eventually an A* search algorithm was used to solve the path planning problem. However, the

problem of this method presented in [86, 87] and [88] was that, it did not localize the effect of obstacles in representation which made the algorithm time consuming.

Another research presented in [89] represented obstacles and the robot as polygons in the environment. This time generalized cones with A* search algorithm was used to specify the path planning. The downside of this approach was that the path was not successfully found when the environment was too cluttered. Enhancing on the existing work obstacles and the robot were now represented in the work space as polyhedral objects as presented in study [90]. This method used the visibility graph algorithm together with the A*search algorithm to find the optimal path. The path found was always a collision free path but failed when dynamic obstacles were presented. The A*search algorithm was later modified to better function with less computational time was named as modified A* algorithm, this algorithm was utilized in study presented in [91] where in the work space was defined by a convex polygon moving and rotating in free space while avoiding obstacles. This solution found the optimal path but took ample amounts of time to solve other algorithms like depth first search found a path much faster than using modified A* search.

A totally unique proposal to handle complex environments with cluttered obstacles is presented in [92] the algorithm used was artificial potential algorithm it helped robot in navigation by going in a very smooth way, however, an everlasting problem with artificial field is the problem of local minima especially when the environment is cluttered with a lot of obstacles. Sometimes due to multiple minimums of potential field the robot ended at a point other than the destination point. A study presented in the early 2000s presented in [93] used the genetic algorithm to define a work space which time varying and having an unknown environment. This solution worked well with complex for each earlier traditional algorithms failed. The best part of this solution was that it did not matter how the obstacles are planted in the environment and since it was time varying it produced results in real time.

Presently researches like studies presented in [76] and [80] have started using methods like probabilistic map (PRM) or Rapid exploring random tree (RRT), these concepts due to their efficient and low processing times added with high accuracy have become the norm for being used in robotics and autonomous vehicles. The programming behind these concepts is presently very easy with many tool boxes present in software's like MATLAB. Hence the other methods previously mentioned slowly disappear as these newer more efficient algorithms over take them in the race for path planning tools.

However, it is important to note that these path planning tools do not alone help in resolving the motion sickness problem which is why transition curves are so important because they will guide the vehicle to an ergonomic solution which hopefully when investigated might tell us how it may help in reducing motion sickness in passengers of autonomous vehicles.

Meanwhile parallel to this the first few designs of Level 4 automation (implication of self-driving technology) cars had appeared as seen in Fig. 8.10.

A few researchers brought to attention the implications of level 4 automation of land vehicles in 2016 [21], based on the envisaged scenarios seen in Fig. 8.10.

Fig. 8.10 A look into the future autonomous land vehicle [119, 120]

As land vehicles are predicted to transform into entertainment transportation pods for the infinite socio-economic benefits, it would at the same time severely worsen motion sickness in passengers of land vehicles as opposed to motion sickness felt now at present in level 3 automation of cars. As all people entering such land vehicles would only be passengers and the design of these cars encourages them to participate into activities like watching a movie, playing games, working on the laptop, reading a book, or just simply having a nap. Involvement in such activities would kick start the motion sickness theory together with postural instability theory as well. Meaning that passengers would undergo severe mismatching of organs of balance and symptoms of motion sickness would make it impossible to make the ride a pleasant one for them. In another study [94] passengers were subjectively examined to respond to a rating for driving in a car; (1) with no task at all, (2) with a visual task like using a hand-held device or book. A total 21 male and 10 females completed this study in a mini-van for a route of 15 min exposed to several numerous curves and gentle slopes. The motion sickness severity was highest when participants were subjected to a visual task as compared to no task.

The study presented in [96] presented an opinion that, the following studies previously mentioned utilized unrealistic driving test conditions with a very few participants, so in this study the authors tested the vehicles on two realistic roads which were accumulated of equal number of right and left turns, and 26 male and 26 females participated in this study. The participants taken were made sure to be belonging to a wide range of age. Each participant had to go through 20 min of the test track once with low acceleration and no task, second with low acceleration and a task, third with moderate acceleration and no task and lastly moderate acceleration with task. The task varied from reading a book, to using a hand-held device to visual search on a tablet. The motion sickness rating was conducted from scale of 0–10. Were a ‘0’ corresponds to no motion sickness at all and ‘10’ means the participants wants to stop the trial.

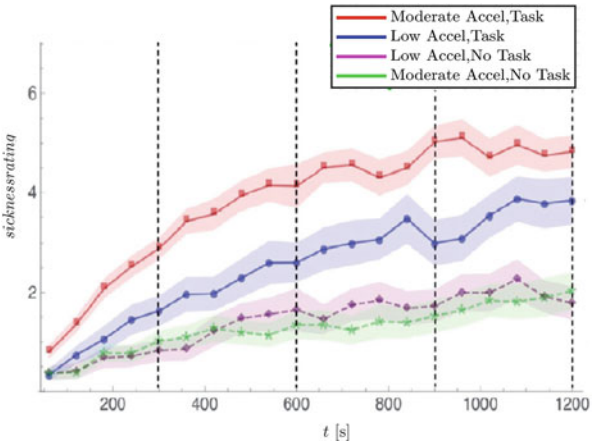


Fig. 8.11 Motion sickness ratings from the study [96]

The results showed that, low acceleration with a task saw a linear (almost a straight line) trend of increase with test track time. Also, this test as compared to others had the highest value rating of motion sickness at the end of 20 min (almost reaching a rating of 4). As for the low acceleration and moderate acceleration without any tasks showed very similar trends throughout the 20 min. They both kept increasing in a linear fashion, but, however, the results saw a lot of oscillating values between both readings and the motion sickness rating trend was like an increasing linear sine wave. But overall these two tests produced the highest motion sickness rating of only 1.5 as compared to a rating of 4 in the first scenario.

Lastly the fourth scenario of having moderate acceleration and a task showed very high motion sickness values at the very beginning (almost reaching up to 3.5 as earlier as in 5 min of the test track), but then the motion sickness went drastically down to as low as to 1 towards the 15 min mark and then showed a sudden peak again between 17–18 min after which it steady to a low value of 2 towards the end of the test at 20 min. All of these results can be seen in Fig. 8.11.

The authors of this paper [96] further analysed by splitting up the motion sickness rating of any participants over the age of 60. Observations clearly show that, all motion sickness rating increased with an increase in the duration of the test. Moderate acceleration with task produced almost a rating of 6, low acceleration with task produced a final rating of 4, moderate and low acceleration both produced a final motion sickness rating of 2. These isolated results (of participants above the age of 60) indicate that, people with ages above 60 are seemingly more susceptible to motion sickness and it gets way worse when these people are involved in a task.

Based on these implicit implications' studies done by Elbanhawi from 2015–2017 [22] and [23] to improve the handling behaviour of level 4 land vehicles using transition curves like Bezier and B-splines proved to be profitable. As steering angles produced were having continuous curvatures and mimicking the handling behaviour that of a real human driver. However, the investigations were more

focused on producing real life like manoeuvrability in level 4 automated land vehicles and not targeted in reducing motion sickness.

Similarly, works presented in [95] used Proportional-Integral-Derivative controller (PID) to enhance the driving comfort of a passenger in an autonomous vehicle a directional closed-loop path tracking control system is integrated into the model which enables the AV model to accurately follow different path profiles. In this regard, firstly, a well-established PID controller is utilized and all of three control parameters are defined by trial and error (KP, KI and KD). In addition, Steering Angle and Path Tracking Errors (which comes from the difference between the direction of the vehicle's trajectory and the slope of the desired path at the closest point to the vehicle) are penalized through a specific equation. Secondly, a multi-objective optimization (with equal weighting factors) is implemented to improve the handling behaviour of the vehicle which ensures enhancement of handling comfort. Then utilising these five objective functions for optimising the controller's gain; (1) Path Tracking Error, (2) Travel Time, (3) Accelerations of the Vehicle Body, (4) Variation Frequency of the Steering Angle and (5) Instability of the Motion. Results indicated that the gain controllers were adjusted successfully to optimize the path mainly by reducing the speed of the vehicle (reduction in acceleration) and path tracking error (the two major contributors causing errors).

The very first strategies to solve the problem of motion sickness were not seen until the year 2020 and 2021 by two important works illustrated by Saruchi [24]. They presented a novel control strategy to utilize the head tilt angles felt by passengers during a slalom track using Radial Basis Function Neural Network to reduce the motion sickness incidence in passengers. Although results have only been validated to include simulation results based on the 6 degree of freedom subjective vertical conflict model, but it was the first of many steps seen to solve the problem of motion sickness in level 4 automation of cars.

Another recent investigation [97] in which a car was modelled as a mass of particle and using a cost function dependant on journey time and MSDV it was shown that issue of motion sickness can be addressed and reduced by more than 50%. However, further investigation is necessary on the weightage factors of the cost function plus as before MSDV was again only exclusive including accelerations in the x and y directions and neglected jerk or other important aspects impacting motion sickness on passengers. However, many researchers around the world are currently implementing strategies to minimize its impact, but very few have practically applied tools to do so. as this has been realized by engineers around the world as the primary issue that might in the end limit Autonomous vehicles to just a simple mere technology that came but could not conquer.

At present it can be deduced that vehicle parameters can create a great impact on reducing or increasing motion sickness specifically the following five major motion sickness thresholds:

– Lateral acceleration: To limit motion sickness its best to keep the lateral accelerations during a turn to 1.8 m/s^2 to 3.6 m/s^2. Going beyond this will aggravate symptoms of motion sickness in passengers.

– Frequency spectrum of lateral acceleration: As per the British Standards the frequency range of 0.1 Hz to 0.5 Hz needs to be suppressed as much as possible, yet no minimum magnitude has yet been defined.
– Phase relationship of roll and lateral acceleration: To minimize motion sickness the roll and lateral acceleration should be at best 29° apart (delay/advance) in phase.
– Steering angles: Gentle and subtle changes in steering angles reduces the resonant body frequencies leading to a better ride comfort in passengers.
– Successive opposing turns: The path taken by the land vehicle if has a lot of successive turns can lead to high levels of mismatch in the organs of balance and hence cause huge amounts of motion sickness in passengers. Additionally, if these turns are successive and opposing in nature it would further worsen the motion sickness in passengers.

Unfortunately, the level 3 technology we have today in cars is not apt to help us reduce these motion sickness thresholds. Whilst we know that the motion sickness incidence expected to befall us is much greater in level 4 automated land vehicles as compared to level 3. Hence, its paramount to discuss how we can tackle this in the future. The next chapter focuses on the not so distant future existence of motion sickness in autonomous land vehicles (AV).

8.4 What Motion Sickness Beholds in Cars of Future?

It is no surprise to say that Diels is responsible for setting the tone for the future research of motion sickness implications on level 4 and level 5 (fully autonomous) land vehicles. And it is well known that these land vehicles will have huge socio-economic benefits to our society. So, it is safe to say that the next couple of years of work will be done to safeguard this technology to bring it into existence. The following chapter sets an outlook for the reader to know what those steps might be to minimize motion sickness in our future land vehicles to optimize their potential in our lives. Figure 8.12 reveals all stages of automation and is a good indication of what level we are now and what is expected in coming few years.

There are two main factors driving motion sickness to exist in future land vehicles, they are:

– Vehicle driven parameters: Since the future of land vehicle will be embedded with self-driving technology. The way land vehicles will be driven, can now be perfected and so all those motion sickness thresholds talked about earlier can be addressed. Some of these solutions will be discussed ahead in this part of the chapter.
– Passenger driven parameters: Future envisaged designs clearly indicate that the passenger will have total freedom to do whatever they want to do in these future land vehicles, since they will be transformed into a mini entertainment transportation pod. This will cause severe mismatching of organs of balance that

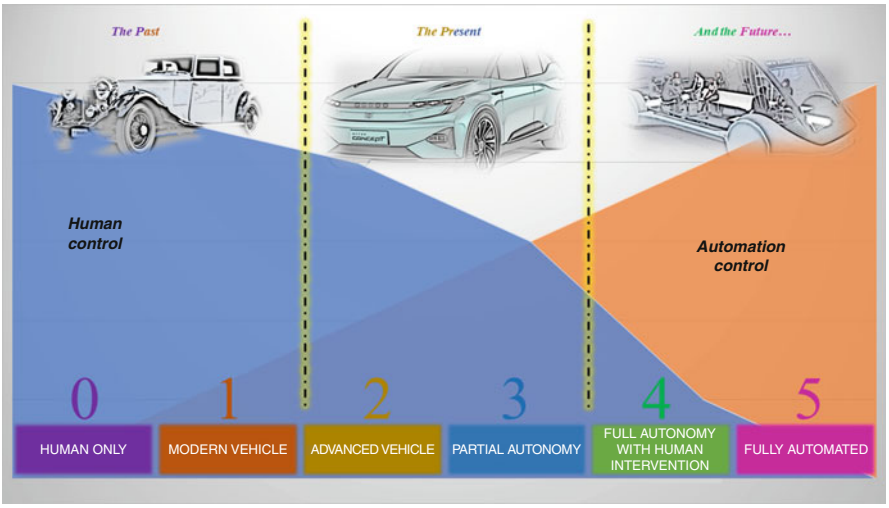


Fig. 8.12 Revealing all levels of automation in land vehicles from the past, present to the future

will induce severe motion sickness symptoms in passengers. Hence, we will be discussing the potential solutions that ergonomist might take to minimize this.

Engineers are developing strategies for making the self-driving technology as safe as possible by implementing ergonomic paths for reducing the motion sickness thresholds to a minimum. Transition curves previously have proven to produce continuous curvature paths and improved handling behaviour and engineers around the world have predicted that a detail investigation into these can help produce paths with which motion sickness can also be minimized. The ideal transition curves for implementing this would be Bezier and B-spline curves as they are the most current and modern techniques used by designers to design curves. Moreover, they are mathematically very easy to compute and super-fast when it comes to processing, additionally they produce realistic curvatures and radii for implementing various land vehicle manoeuvres.

However, the characteristics of these transition curves can be further extended, such as local control and uniformity by using Non-Uniform Rational Basis Splines (NURBS). These characteristics allow paths to be more feasible and present more control over the lateral movement of the object travelling on those curves. These characteristics can profoundly be fruitful in delegating with the thresholds of motion sickness and hence evolve the paths. Making them ergonomic enough to limit motion sickness in passengers.

While ergonomic paths on their own are very useful and may present a viable solution to the problem of motion sickness in future autonomous cars, but to optimize and minimize the problem application of a cost function with an appropriate selection of proper control strategy and controllers might be required. To begin with proper selection of objective functions are required to setup a cost function that

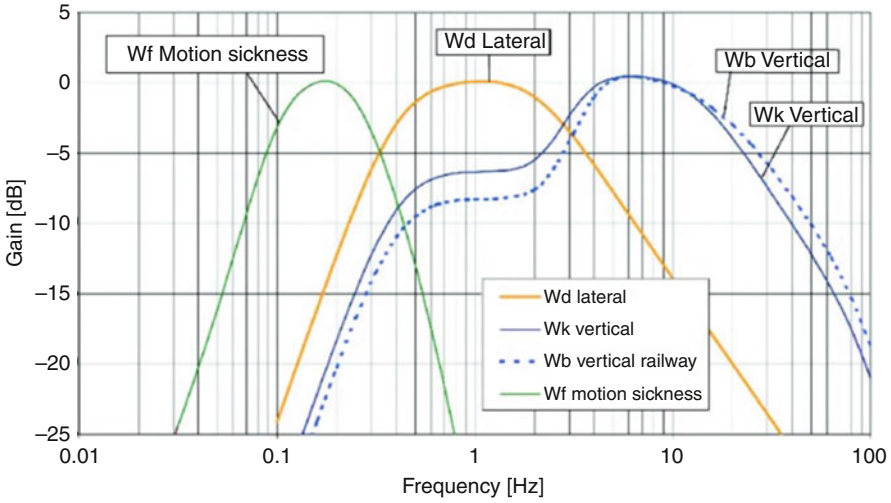


Fig. 8.13 Various weighted frequency filters as per defined by ISO 2631 and British standards, with weighted frequency filter shown in green colour [117]

would minimize motion sickness. Following are a few recommendations that maybe useful to solve this problem:

1. Fast Fourier Transform (FFT) spectrum: These are the Root Mean Square (RMS) values of the FFT spectrum of steering angle variation targeted at reducing high-magnitude low-frequency values, this would reduce the human body resonance frequencies and help in reducing motion sickness in passengers.
2. Motion Sickness Dosage Value (MSDV): These are the weighted frequency values (Fig. 8.13) of a land vehicle's lateral acceleration. This aims to reduce the frequency component of lateral acceleration produced in the range of 0.1 Hz to 0.5 Hz in order to reduce motion sickness in passengers.
3. Motion Sickness Incidence (MSI): Defined using the 6 degree of freedom subjective vertical conflict model. It is defined as the variation of a passenger's head tilt (roll angles) has on their motion sickness symptoms. Minimizing this would lead to increased postural stability and reduction in motion sickness.
4. Lateral Error: The lateral deviation (Fig. 8.14) of the autonomous car from the paths that the vehicle is subjected to follow. This ensures that the vehicle follows the ergonomic paths most efficiently. Following these paths would lead to minimizing the motion sickness in a passenger.
5. Illness rating (IR): A simple numerical value calculated by multiplying an empirically value of 1/50 by the MSDV values of the vehicle to identify the sickness levels of passengers. Reducing this would reduce the rate at which sickness symptoms will aggravate in a passenger.
6. Acceptable Jerk Value (AJV): Jerk is the derivative of lateral acceleration, and as per the recommendations and its strong relationship to subjective evaluation of

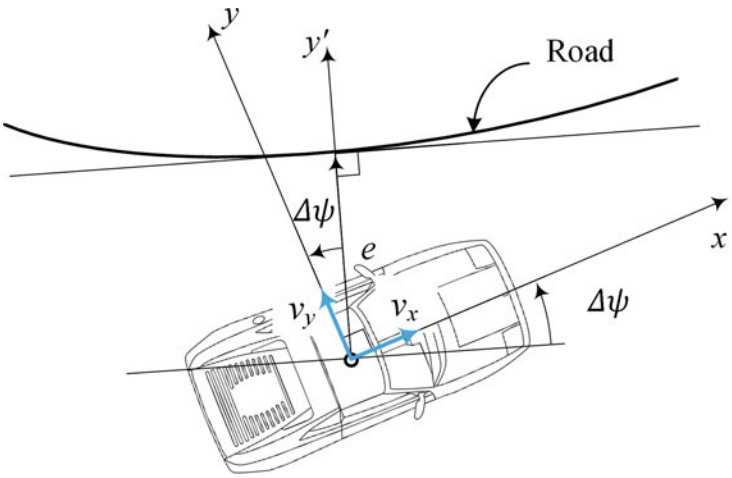


Fig. 8.14 Shows how lateral error is calculated [121]

passengers' sickness levels [25]. AJV is a much better way to calculate motion sickness symptoms then jerk on its own. AJV is calculated by summating the RMS value of jerk with the highest peak square value of jerk. Minimizing AJV will lead to a better shift quality and ride comfort and hence reduce the motion sickness in the passengers.

For optimizing these objectives traditional controllers such as the Proportional-Integral-Derivative (PID) controller can be used together with Particle Swarm Optimization (PSO) to minimize motion sickness. However, newer, robust and much profound controllers such as the fuzzy logic and model predictive controllers would be a better choice. Since they are commonly used in automotive industries, for example, the use of fuzzy logic to define a control strategy for eliminating motion sickness by optimizing the head tilt angles, or optimizing the lateral error to keep the autonomous car on track using model predictive controllers.

Concentration on using Model Predictive controllers (MPC) is increasing mainly due to two valid reasons presented in [99] this study overviews the MPC. Initially the study elaborates using MPC's over PID controller. As MPC can be placed anywhere in the process levels and its capable of handling a wide variety of constraints. Mainly the study [99] emphasizes on understanding how MPC has been deployed in the field of automation for vehicles. Specifically, in the field of; (1) steering and braking control, (2) Yaw and Roll dynamics control and (3) Active and semi-active suspension control.

In the field of steering and braking initially MPC controllers were utilized for active safety in longitudinal dynamics for Anti-locking braking (ABS) and for traction control systems (TCS). Next line of applications was based on Non-linear MPC (NMPC) based on the bicycle model such as mentioned in the study [100] which defined its main objective function to stabilize vehicle and its trajectory

against physical disturbances like wind. Results indicated the strength of this design when the controller performed the double lane change manoeuvre. Studies presented in [101] compared two types of MPC controllers, linear time varying (LTV MPC) and NMPC. Results indicated that LTV MPC produced better stability and were computationally more realisable. An approach utilized in the study [105] for braking and steering used MPC and results of the controller where then compared with [100] and [101]. Controllers were tested on snowy and icy tracks together with a double lane change manoeuvre at very high speeds. Results produce stable and successful trajectory following and introduced obstacle avoidance in constraints.

So far studies presented in [100, 101] and [105] designed MPC based on known trajectories the study presented in [106] applied control to four unknown vehicle trajectories. The study utilized the bicycle model and the vehicle model were formulated using linear hybrid dynamic system in piece wise affine form with respect to front plus rear tire slip angles. Results showed stability and fast track response. Recent studies like [107] and [108] have utilized design controllers that are capable of progressively augmenting driver steering input and more recent studies have presented design of an active system to; (1) prevent collisions, (2) road obstacles and (3) road way departure. Further scope of work for which MPC can be utilized in the field of steering and braking is to deal with extreme cornering while tackling issues of over-steer and under-steer. Another area requiring development is the use of hap-tic feedback to inform driver of the augmented steering.

Usage of MPC for yaw and roll dynamics has been studied in [109]. Using fast NMPC (FNMPC) methodology based on offline approach produced highly damped behaviour but stability for reverse driving. Further a study in 2011 [110] based on the MPC technique integrated a control algorithm for vehicle in active steering and dynamics yaw control. The MPC was designed to solve problems like adjusting weight control, input variables together with tires limitation and force allocation to produce results that enhance body vehicle stability. Study using NMPC on a 12th order car model presented in [111] involved roll dynamics to improve the dynamics vehicle behaviour at high speed on high friction surfaces whilst smoothing path following and steering control. Future work on roll with MPC can be done on other under predicted values like roll rates and roll angles inclusion of these factors will improve correlation and bring in realistic factors like tire deformation and bushing stiffness.

Use of MPC has been quite successfully done on active and passive suspension systems. Study [112] presents the first ever suspension system designed using MPC. The design was based on second degree of freedom (DOF) quarterly model with hard constraints using single and multiple step inputs results illustrated increased ride comfort and road handling. Scenarios such as riding over a rough road using a half car model were implemented using MPC in the study presented in [113] moreover neural networks together with MPC on a non-linear quarter car model were applied in [114] and [115], resulting in better tracking then a constant gain PID controller. Recently adopted designs incorporating roll mode, different profiles and non-linear full car vertical dynamics model in [116] resulted in reduction of acceleration and increase in ride comfort. FMPC was tested on half car non-linear

model to conduct a comparison between; (1) sky hook, (2) linear quadratic and 3) clipped control. Response produced by MPC showed better handling and comfort over all road classes considered. However, a lot of studies have been conducted on suspension systems, but future work can be done on working with electromagnetic suspensions by involving additional uncertainties of state for suspension control (due to delay) plus including sensor measurements, actuator band limit, different MPC control strategies and influence of time delay in loop. Having discussed the countless number of applications that MPC are involved with in land vehicles of tomorrow. The choice and selection of the controller is not limited to these only as many other control tools like sliding mode controller and linear quadratic regulator are often used to optimize autonomous tasks in land vehicles as well.

The land vehicles of tomorrow may well be fitted with devices that could measure the motion sickness in passengers by evaluating the vehicle driven parameters. Hence very recently, Bos has presented the first ever portable motion sickness prediction device for automated vehicles [98]. It utilizes a cost function to predict the motion sickness which is based on; (1) low-frequency motion of the vehicles (measured using MPU 6050 together with an Arduino Nano), (2) head tilt gaze function, where a head up results in a low gain (small contribution to motion sickness) and a head down results in a high gain value and (3) lastly utilizing the temperature module present in Arduino Nano to develop the ambient temperature function. Plus, the device has the capability of showing motion sickness recovery when the vehicle or the passenger is not undergoing any motion. Measurements produced by the device support previous studies, such as, participants having their head down had higher MSDV values as compared to people with their head up. Overall, the study presented a device that successfully predicts the passengers motion sickness, however, due to limited data available these measurements are only catering for a small population of people, plus many factors such as jerk etc... are not included in the cost function of the device making, indicating that the following cost function needs much needed correction for it to match a larger scale of population

The strategies discussed above may help in limiting motion sickness by reducing vehicle driven parameters but not necessarily impact as much of the passenger driven parameters. To limit motion sickness a lot of it depends on how ergonomists can design the vehicle interior to limit motion sickness behaviours/movements of passengers. But to understand better, ergonomists must keep in mind a few less popular human traits that may increase motion sickness in passengers of autonomous cars:

1. Habituation: passengers of land vehicles show lower levels of habituation as compared to sickness at sea or in flight. For instance, people on cruise get habituated as the journey progresses, where they feel more sickness on the first day but on the second day symptoms decrease. Unfortunately, this is not the case with land vehicles.
2. Susceptibility to postural instability: One in ten thousand people are susceptible to motion sickness and such people are also prone to postural instability. This

happens because such people have; (1) high sensitivity to stimulation, (2) slower adaptation rates to changes and (3) slower decay of elicited symptoms. So, to define a single generic strategy for all passengers might not be beneficial for people who are more susceptible to motion sickness than others.

3. Activities prior to sitting in an AV: Motion sickness is found to be higher if a person is sleep deprived or having an empty stomach. Maybe a certain driving setting for people who slept less than optimal hours and a different setting for people sleeping the optimal hours might be needed by the vehicle, by simply asking the passenger a few routine questions before he starts his/her journey.
4. Age & Gender: Females show more susceptibility to motion sickness as compared to males. In terms of age motion sickness peaks are in the age group of 16–20. Similar strategies to point number 3 might work here as well.

It is worth mentioning a few critical studies that have indicated the importance of passenger driven parameters on aggravating motion sickness such as the study presented in [102] considered testing the impact of motion sickness with regards to vision and task performance. The participants of this study went through several test such as (1) blind folding them to block their vision, (2) restriction on the forward and rear-view windows and altering the seat positions. The twenty male participants were assessed through a questionnaire after participating in each test condition. The severity of motion sickness was found to be the most when participants were not allowed to see the front view.

The study in [103] investigated variations in in-built vehicle displays. Twenty male participants were tested with different varieties of displays (no visual screen, compensated pitch motion displays and forward located displays) in the vehicle. As a result, a reduction in motion sickness was observed for the compensated displays.

In a more recent study [104] the role of vision and head position on motion sickness severity was conducted. Small passenger cars were driven on a slalom track at a speed of 30km/h for 30 laps. Half of the participants were told to tilt their heads towards the centrifugal force intentionally with their eyes open and the other half were told to tilt their heads away against the centrifugal force and with their eyes open. Results indicated participants with their eyes open and tilting towards the centrifugal force were less susceptible to motion sickness. These studies have severed as an eye opener to understanding the impact of passenger driven behaviours in the future of land vehicles. These interesting points raised allow us to dwell upon how ergonomists in the future can address these problems.

With self-driving technology embedded into land vehicles of future seat allocation in land vehicles will change. Ergonomists can use this to their advantage and have revolving seats that would force the passengers to manoeuvre their heads towards the bend (more like the head tilt strategy of a driver Fig. 8.15). As experimentational studies have proven that the head roll angles of a driver cause lower motion sickness as their heads are towards the bend of the curve [26]. On the contrary in a traditional land vehicle passengers move their heads away from the bend of the curve, such a head strategy increases motion sickness. Hence, designing revolving seats that would ensure the passengers of future land vehicles would have their heads aligned with the bend could do wonders for postural stability.

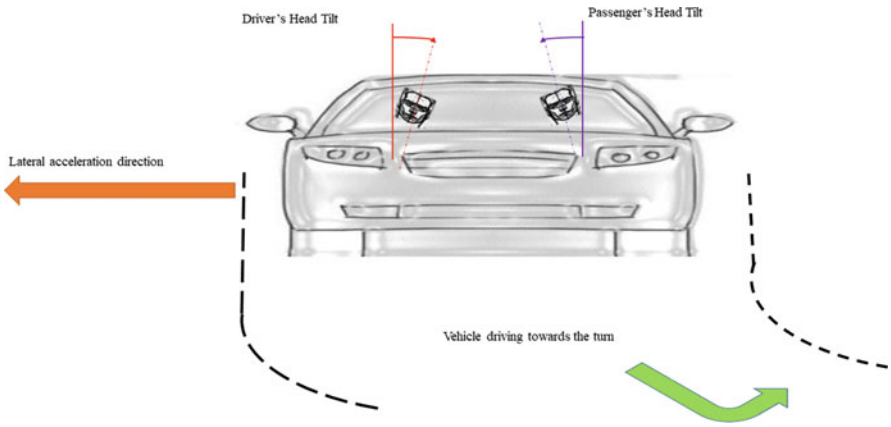


Fig. 8.15 Head tilt strategies of a driver and a passenger of a land vehicle

Where a passenger sits is as important as the seating arrangements inside a car in context to reducing motion sickness. As passenger seating placement changes the frontal viewing angles change, either giving them a greater view or vice versa. Studies indicate that passengers are more prone to motion sickness symptoms if their frontal viewing angle is barred [27]. Designing panoramic windows throughout the interior of a car (360 degrees viewing angle) or using augmented reality to allow passengers a full frontal viewing angle might be one of many design challenges that ergonomist have to implement to reduce motion sickness in future land vehicles. Passengers would like to fully realize the potential of self-driving technology in future land vehicles by utilizing their time either working or emailing via a tablet, laptop or even their mobile phones. Ergonomists must present an interior design that would allow the users screens to holographically appear on the windows of the vehicle. This would ensure, that while the passenger is working on writing an email, his eyes will also have a background view of the outside, ensuring visually induced motion sickness to reduce.

With multiple passengers inside the land vehicle options of movie night, playing games like monopoly, chess or even a game of cards would increase, which in turn increases the likelihood of passengers getting sick. Studies have shown using roll compensated visual displays (for movie night viewing) has proven to reduce motion sickness in passengers [28]. Hence, screens should support this feature and games of these nature should be uploaded in the vehicle from before so they too can be viewed on the roll compensated screens.

However, with all this said and done there is still a chance that the passenger does something unexpected and that is where the real challenge lies for ergonomists and engineers. It is very much possible that the passenger would like to have a meal on the way in his futuristic land vehicle. In such a case a lot of variables come into action, for example, what type of meal are they having as the changes in posture would be very different if the person is having a pizza or just a simple bowl of soup. Such ambiguous behaviours need to be dealt with, by using artificial

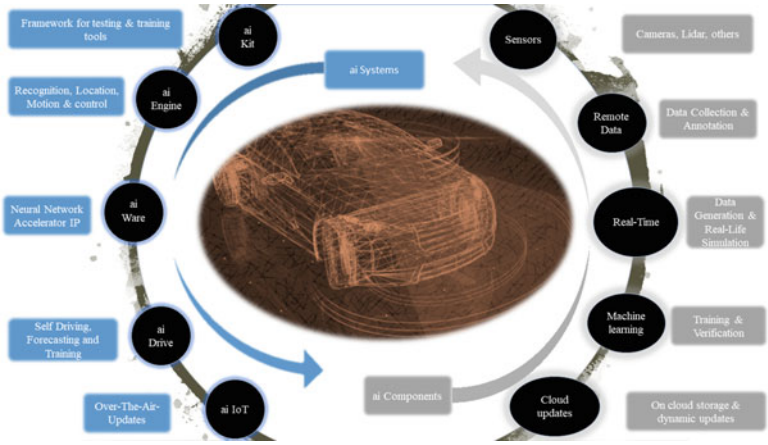


Fig. 8.16 The AI automotive ecosystem of level 5 vehicles

intelligence. Artificial intelligence would help to predict or forecast the type of movements generally the passenger would make when doing such an unexpected behaviour and would modify the car settings (both passenger driven and vehicle driven) to best adapt to these unexpected patterns of behaviour and help in limiting motion sickness. Figure 8.16 shows how the engineers anticipate the Ecosystem of Artificial Intelligence (ai) would be in self-driving Autonomous Vehicles.

It would be excellent to see the impact of integrating these features in our future cars to limit motion sickness as much as possible by both means. However, one thing is rest assured that motion sickness is a real threat to the future of self-driving technology and it needs to be addressed to its very core in order to secure the acceptance of this technology in the future land vehicles and without this being addressed it would be an extreme case of exaggeration to see level 4 or 5 automated cars on the road as early as 2021.

References

1. Treisman, M. (1977). Motion sickness: an evolutionary hypothesis. Science, 197(4302), 493–495.
2. Samuel, A. J., Solomon, J., & Mohan, D. (2015). A critical review on the normal postural control. Physiotherapy and Occupational Therapy Journal, 8(2), 71.
3. Reason, J. T., & Brand, J. J. (1975). Motion sickness. Academic press.
4. Cunningham, Micheal, 2015. Autonomous Vehicles: Human factors issues and future research. Proceedings of the 2015 Australasian Road Safety Conference.
5. Sivak, M., & Schoettle, B. (2015). Motion sickness in self-driving cars. Michigan, USA. University of Michigan, Transportation Research Institute (UMTRI).
6. Graybiel A, Knepton J (1976) Sopite syndrome: a sometimes-sole manifestation of motion sickness. Aviat Space Environ Med 47:873–882.
7. Lawson BD, Mead AM (1998) The sopite syndrome revisited: drowsiness and mood changes during real or apparent motion. Acta Astronaut 43:181–192.

8. Matsangas P, McCauley M (2014b) Sopite syndrome: A revised definition. Aviat Space Environ Med 85:672–673.
9. Kaplan E, ventura J, Pierobon A, Bakshi A, DiZio P, Lackner JR (2014) Influence of sleep deprivation on susceptibility to motion sickness and performance on cognitive and motor tasks (in preparation).
10. Lackner JR, Graybiel A (1984a) Elicitation of motion sickness by head movements in the microgravity phase of parabolic flight maneuvers. Aviat Space Environ Med 55:513–520.
11. Barany, R. (1921). Modern labyrinthology. The Laryngoscope, 31(7), 401–403.
12. British Standards Institution. (1987). Measurement and evaluation of human exposure to whole-body mechanical vibration and repeated shock. BSI.
13. Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. Ecological psychology, 3(3), 195–240.
14. Kennedy, R. S., & Stanney, K. M. (1996). Postural instability induced by virtual reality exposure: Development of a certification protocol. International Journal of Human-Computer Interaction, 8(1), 25–47.
15. Taxonomy, S. A. E. (2014). Definitions for terms related to on-road motor vehicle automated driving systems. Society of Automotive Engineers (SAE): Troy, MI, USA.
16. Turner, M., & Griffin, M. J. (1999). Motion sickness in public road transport: the effect of driver, route and vehicle. Ergonomics, 42(12), 1646–1664.
17. Voigt, A. P., Fenno, D. W., & Borchardt, D. W. (2003). Evaluation of vehicle speeds on freeway-to-freeway connector ramps in Houston (Vol. 4318, No. 1). Texas Transportation Institute, Texas A & M University System.
18. Bles, W., Bos, J. E., De Graaf, B., Groen, E., & Wertheim, A. H. (1998). Motion sickness: only one provocative conflict? Brain research bulletin, 47(5), 481–487.
19. Bos, J. E., & Bles, W. (2002). Theoretical considerations on canal–otolith interaction and an observer model. Biological cybernetics, 86(3), 191–207.
20. Wada, T., & Kamij, N. (2015). A mathematical model of motion sickness in 6DOF motion and its application to vehicle passengers. arXiv preprint arXiv:1504.05261.
21. Diels, C., & Bos, J. E. (2016). Self-driving carsickness. Applied ergonomics, 53, 374–382.
22. Elbanhawi, M., Simic, M., & Jazar, R. (2015). In the passenger seat: investigating ride comfort measures in autonomous cars. IEEE Intelligent transportation systems magazine, 7(3), 4–17.
23. M. Elbanhawi, M. Simic, and R. Jazar, "Improved manoeuvring of autonomous passenger vehicles: Simulations and field results," Journal of Vibration and Control, vol. 23, no. 12,pp. 1954–1983, 2017.
24. Saruchi, S. A., Ariff, M. H. M., Zamzuri, H., Amer, N. H., Wahid, N., Hassan, N., & Kadir, Z. A. (2020). Lateral control strategy based on head movement responses for motion sickness mitigation in autonomous vehicle. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 42(5), 1–14.
25. Huang, Q., & Wang, H. (2004). Fundamental study of jerk: evaluation of shift quality and ride comfort (No. 2004-01-2065). SAE Technical Paper.
26. Wada, T., Konno, H., Fujisawa, S., & Doi, S. I. (2012). Can passengers' active head tilt decrease the severity of carsickness? Effect of head tilt on severity of motion sickness in a lateral acceleration environment. Human factors, 54(2), 226–234.
27. Griffin, M. J., & Newman, M. M. (2004). Visual field effects on motion sickness in cars. Aviation, space, and environmental medicine, 75(9), 739–748.
28. Kato, K., & Kitazaki, S. (2006). A study for understanding carsickness based on the sensory conflict theory (No. 2006-01-0096). SAE Technical Paper.
29. Money KE (1990) Motion sickness and evolution. In: Crampton GH (ed) Motion and space sickness. CRC Press, Boca Raton.
30. Treisman M (1977) Motion sickness: an evolutionary hypothesis. Science 197:493–495.
31. Riccio GE, Stoffregen TA (1991) An ecological theory of motion sickness and postural instability. Ecol Psychol 3:195–240.
32. Oman CM (1982) A heuristic mathematical model for the dynamics of sensory conflict and motion sickness. Acta Otolaryngol Suppl 392:1–44.

33. Smart LJ, Pagulayan RJ, Stoffregen TA (1998) Self-induced motion sickness in unperturbed stance. Brain Res Bull 47:449–457. https://doi.org/10.1016/S0361-9230(98)00103-8.
34. James R. Lackner, 2014, Motion sickness: more than nausea and vomiting, Experimental brain research, 2014, 2493–2510.
35. Graybiel A, wood CD (1969) Rapid vestibular adaptation in a rotating environment by means of controlled head movements. Aerosp Med 40:638–643.
36. Yen-Pik Sang F, Billar J, Gresty MA, Golding JF (2005) Effect of a novel motion desensitization training regime and controlled breathing on habituation to motion sickness. Percept Mot Skills 101:244–256.
37. Cohen B, Dai M, Yakushin SB, Raphan T (2008) Baclofen, motion sickness susceptibility and the neural basis for velocity storage. Prog Brain Res 171:543–553. https://doi.org/10.1016/S0079-6123(08)00677-8.
38. Guignard J.C and McCauley, M.E. 1990, The acceleration stimuli in motion sickness, in G.H.Crampton (ed.), Motion and Space Sickness (Boca Raton: CRC Press), 123–152.
39. Alexander. S. J, Cotzin, M. Hill, C.J.Ricciuti, E.A and Wendt G.R, 1945, Westleyan University studies of motion sickness VI. Predictions of sickness on a vertical accelerator by means of motion sickness history questionnaire, Journal of Psychology, 20, 25–30.
40. Lawther A, and Griffin M.J 1987, Prediction of the incidence of motion sickness from the magnitude, frequency and duration of vertical oscillation, Journal of the Acoustical Society of America, 82, 957–966.
41. Griffin M.J , 1990, Handbook of human vibration (London: Academic press).
42. Golding J.F, Markey H.M and Scott J.R.R, 1995, the effects of motion direction, body axis and posture on motion sickness induced by low frequency linear oscillation, Aviation, Space and Environmental medicine, 66, 1046–1051.
43. Golding J.F and Markey H.M 1996, Effect of frequency of horizontal linear oscillation on motion sickness and the somatogravic illusion, Aviation, Space and Environmental Medicine, 67, 121–126.
44. Probst, Krafczyk, Buchele and Brandt, 1982 Visual Prevention of motion sickness in cars, Archive for psychiatrie und Nervenkranheiten, 231, 409–421.
45. H.Vogel, R.Kohlhaas and Baumgarten, R.J Von 1982, Dependence of motion sickness in automobiles on the direction of linear acceleration, European Journal of Applied physiology, 48, 399–405.
46. A. Rolnick and R.E.Lubow, 1991, why is the driver rarely motion sick? The role of controllability in motion sickness, Ergonomics, 34, 867–879.
47. R.J.Koppa and G.G.Hayes, 1976, Driver inputs during emergency or extreme vehicle manoeuvres, Human Factors, 18, 361–370.
48. M.Turner, 1992, A description of low frequency motion in road coaches, in proceedings of United Kingdom Informal group of human response to vibration (Southampton: University of Southampton), 323–339.
49. M.J Griffin and M.M Newman, 2004, An experimental study of low-frequency motion in cars, Journal of Automobile engineering Vol. 218 Part D.
50. Joseph J.A & Griffin M.J, Motion sickness from combined lateral and roll oscillation: effect of varying phase relationships. Aviat Space Environ Med 2007; 78:944–50.
51. Förstberg J. Effects from lateral and/or roll motion on nausea on test subjects: studies in a moving vehicle simulator. Proceedings of the 34th UK Group Meeting on Human Response to Vibration; 22 – 24 September 1999.Dunton, Essex, UK: Ford Motor Company; 1999: 209–18.
52. Howarth HVC, Griffin MJ. Effect of roll oscillation frequency on motion sickness. Aviation Space Environ Med 2003 ;74: 326–31.
53. Wertheim AH, Bos JE, Bles W. Contributions of roll and pitch to sea sickness. Brain Res Bull 1998 ;47: 517–24.
54. Donohew BE.Effect on motion sickness of the percentage of roll compensation during lateral oscillation. Proceedings of the 38th United Kingdom Conference on Human Response to Vibration; 17–19 September 2003 Alverstoke, Gosport, UK: The Institute of Naval

Medicine;2003: 79–87.
55. Donohew BE. Effect on motion sickness of the percentage of roll-compensation of 0.1 Hz lateral oscillation. Proceedings of the 39th United Kingdom Conference on Human Response to Vibration; 15–17 September 2004. Ludlow, Shropshire, UK: RMS Vibration Test Laboratory; 2004: 47–56.
56. Förstberg J. Effects from lateral and/or roll motion on nausea on test subjects: studies in a moving vehicle simulator. Proceedings of the 34th UK Group Meeting on Human Response to Vibration; 22 – 24 September 1999. Dunton, Essex, UK: Ford Motor Company;1999: 209–18.
57. Joseph JA, Griffin MJ. Motion sickness: effect of the magnitude of roll and pitch oscillation. Aviation Space Environ Med 2008; 79:390–6.
58. Mills KL, Griffin MJ, Effect of seating, vision and direction of horizontal oscillation on motion sickness. Aviation Space Environ Med 2000 ; 71 : 996–1002.
59. Y. Morales, N. Kallakuri, K. Shinozawa, T. Miyashita and N. Hagita, (2008), 'Human-Comfortable Navigation for an Autonomous Robotic Wheelchair', IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, 2737–2743
60. Jin Xu, Kui Yang, Yiming Shao and GongYuan Lu, An experimental study on lateral acceleration of cars in different environments in Sichuan, Southwest China. Hindawi Publishing Corporation Discrete Dynamics in Nature and Society Volume 2015, Article ID 494130, 16 pages.
61. A. Kemeny and F. Panerai, "Evaluating perception in driving simulation experiments.", Trends Cogn Sci, vol. 7, no. 1, January. 2003, pp. 31–37.
62. E. L. Groen and W. Bles, "How to use body tilt for the simulation of linear self motion", J. Vestib. Res. Equilib. Orientat., vol. 14, no. 5, 2004, pp. 375–385.
63. A. Stratulat, V. Roussarie, J.-L. Vercher, and C. Bourdin," Improving the realism in motion-based driving simulators by adapting tilt-translation technique to human perception", 2011, pp. 47–50.
64. A. M. Stratulat, V. Roussarie, J.-L. Vercher, and C. Bourdin, "Perception of longitudinal acceleration on dynamic driving simulator", presented at Proceeding of the Driving Simulation Conference 2012, Paris, France, 2012, pp. 33–40.
65. D. R. Berger, J. Schulte-Pelkum, and H. H. Bülthoff, "Simulating believable forward accelerations on a stewart motion platform", ACM Trans. Appl. Percept., vol. 7, no. 1, January. 2010, pp. 1–27.
66. A. Berthoz et al. "Motion Scaling for High-Performance Driving Simulators", IEEE Trans. Hum.-Mach. Syst., vol. 43, no. 3, May. 2013, pp. 265–276.
67. Savona, Florian & Stratulat, Anca & Diaz, Emmanuelle & Honnet, Vincent & Houze, Gilles & Vars, Philippe & Masfrand, Stéphane & Roussarie, Vincent & Bourdin, Christophe. The Influence of Lateral, Roll and Yaw Motion Gains on Driving Performance on an Advanced Dynamic Simulator. 10.13140/2.1.3720.9600.
68. Chapron, Thomas, and Colinot, Jean-Pierre, "The new PSA Peugeot-Citroën Advanced Driving Simulator Overall design and motion cue algorithm", presented at Proceeding of the Driving Simulation Conference 2007, North America, Iowa City, 2007, pp. 44–52.
69. Fard, Mohammad & Ishihara, Tadashi & Inooka, Hikaru. (2003). Dynamics of the head-neck complex in response to the trunk horizontal vibration: modeling and identification. Journal of biomechanical engineering. 125. 533–9.
70. Marzbani, H., Simic, M., Fard, M., & Jazar, R. N. (2015). Better road design for autonomous vehicles using clothoids. In Intelligent Interactive Multimedia Systems and Services (pp. 265–278). Springer, Cham.
71. D. González, J. Pérez, V. Milanés and F. Nashashibi, (2016), 'A Review of Motion Planning Techniques for Automated Vehicles, IEEE Transactions on Intelligent Transportation Systems, Vol. 17, No. 4, 11351145.
72. Belaidi H, Hentout A, Bouzia B, Bentarzi H, Belaidi A, NURBS trajectory generation and following by an Autonomous Mobile Robot Navigating in 3D environment, The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent

Systems June 4–7, 2014, Hong Kong, China.
73. A. K. Singh, A. Aggarwal, M. Vashisht and R. Siddavatam, "Robot motion planning in a dynamic environment using offset Non-Uniform Rational B-Splines (NURBS)," 2011 IEEE International Conference on Industrial Technology, Auburn, AL, 2011, pp. 312–317.
74. A. J. Schmid and H. Woern, "Path planning for a humanoid using NURBS curves," IEEE International Conference on Automation Science and Engineering, 2005., Edmonton, Alta., 2005, pp. 351–356.
75. E. K. Xidias and N. A. Aspragathos, "Continuous curvature constrained shortest path for a car-like robot using S-Roadmaps," 21st Mediterranean Conference on Control and Automation, Chania, 2013, pp. 13–18.
76. Elbanhawi M, Simic M, Reza NJ, Journal Intelligent Robot Syst (2015) 80 (Suppl 1):S23–S56 DOI 10.1007/s10846-014-0172-0.
77. Huh, U.-Y., Chang, S.-R.: A G2 continuous pathsmoothing algorithm using modified quadratic polynomial interpolation. Int. J. Adv. Robot. Syst. 25(11) (2014). https://doi.org/10.5772/57340.
78. Zhou, F., Song, B., Tian, G.: Bézier curve based smooth path planning for mobile robot. J. Inf. Comput. Sci. 8(12), 2441–2450 (2011).
79. Kwangjin, Y., Jung, D., Sukkarieh, S.: Continuous curvature path-smoothing algorithm using cubic Bezier spiral curves for non-holonomic robots. Adv. Robot. 27(4), 247–258 (2013). https://doi.org/10.1080/01691864.2013.755246.
80. Elbanhawi M, Simic M, Reza NJ, Journal of Vibration and Control 2017, Vol. 23(12) 1954–1983, https://doi.org/10.1177/1077546315605666.
81. Kwangjin Y and Sukkarieh S (2010) An Analytical Continuous-Curvature Path-Smoothing Algorithm. Robotics, IEEE Transactions on 26: 561–568.
82. Zheng Z (2014) Recent developments and research needs in modeling lane changing. Transportation Research Part B: Methodological 60: 16–32.
83. Perez J, Godoy J, Villagra J, et al. (2013) Trajectory generator for autonomous vehicles in urban environments. Robotics and Automation (ICRA), 2013 IEEE International Conference on. pp. 409–414.
84. D. González, J. Pérez, V. Milanés and F. Nashashibi, (2016), 'A Review of Motion Planning Techniques for Automated Vehicles', IEEE Transactions on Intelligent Transportation Systems, Vol. 17, No. 4, 1135–1145.
85. N. Sariff and N. Buniyamin, "An Overview of Autonomous Mobile Robot Path Planning Algorithms," 2006 4th Student Conference on Research and Development, Selangor, 2006, pp. 183–188.
86. B.R Donald, "Motion Planning With Six Degrees of Freedom," Massachusetts Institute Technology Artificial Intelligence Laboratory, Technical report AIM-791, 1984.
87. Rodney A.Brooks, "Solving the find path problem by representing free space as generalized cones" A.I Memo No 674, Massachusetts Institute Technology, May 1982.
88. Subbarao Kambhampati and Larry S. Davis, "Multiresolution path planning for mobile robots," in International Journal of Robotics Research, 5(10):90–98, Spring 1986.
89. R.A Brooks and T.Lorenzo-Perez. "A subdivision algorithm in configuration space for findpath with rotation," in IEEE Transactions on Systems, Man and Cybernetics, Mar./Apr.1985, SMC-15(2), pp 225–233.
90. T.Lozano-Perez and M.A Wesly. "An algorithm for planning collisionfree paths among polyhedral obstacles," in Communication of the ACM,22(10), Oct 1979, pp 560–570.
91. Kuo-Chin Fan and Po-Chang Lui, "Solving the find-path problem in mapped environments using modified A* search algorithm," in IEEE Transactions on Systems, Man, and Cybernetics, vol.24, no.9, September 1994, pp 1390–1396.
92. O.Khatib, "Real time obstacle avoidance for manipulators and mobile robots," in International Journal of Robotics Research, 5(10):90–98, Spring 1986.
93. Christoph Niederberger,ejan Radovie, and Markus Gross, "Generic path planning for real time application," in Proceedings of the Computer Graphics International (CGI 04), 2004, pp 1–8.

94. Isu, N., T. Hasegawa, I. Takeuchi, and A. Morimoto. 2014. "Quantitative Analysis of Time-Course Development of Motion Sickness Caused by in-Vehicle Video Watching." Displays 35(2): 90–97. https://doi.org/10.1016/j.displa.2014.01.003.
95. Mohajer, N., Nahavandi, S., Abdi, H., & Najdovski, Z. (2020). Enhancing Passenger Comfort in Autonomous Vehicles Through Vehicle Handling Analysis and Optimization. IEEE Intelligent Transportation Systems Magazine.
96. Monica L.H.Jones , Victor C. Le, Shiela M. Ebert, Kathleen H. Sienko, Matthew P. Reed and James R. Sayer, Motion sickness in passenger vehicles during test track Operations, 2019, Ergonomics Vol.62. No 10, 1357–1371, doi.org/10.1080/00140139.2019.1632938.
97. Htike, Z., Papaioannou, G., Velenis, E. and Longo, S., 2020, May. Motion Planning of Self-driving Vehicles for Motion Sickness Minimisation. In 2020 European Control Conference (ECC) (pp. 1719–1724). IEEE.
98. Salter, S., Diels, C. and Thake, D., 2019. Motion sickness prediction device for automated vehicles. Int J Mech Produc Eng, 7, pp.68–74.
99. Yakub, Fitri & Mori, Y.. (2013). Model predictive control for car vehicle dynamics system -comparative study. 2013 IEEE 3rd International Conference on Information Science and Technology, ICIST 2013. 172–177. 10.1109/ICIST.2013.6747530.
100. F. Borrelli, P. Falcone, T. Keviczky, J. Asgari, and D. Hrovat, "MPC based approach to active steering for autonomous vehicle systems," Int. J. Vehicle Autonomous Systems, vol. 3, no. 2/3/4, pp. 265–291, 2005.
101. P. Falcone, F. Borrelli, J. Asgari, H. E. Tseng, D. Hrovat, "Predictive active steering control for autonomous vehicle systems," IEEE Trans. on Control System Technology, vol. 15, no. 3, pp. 566–580, 2007.
102. Griffin, M. J., and M. M. Newman. 2004b. "Visual Field Effects on Motion Sickness in Cars." Aviation, Space, and Environmental Medicine 75(9): 739–748.
103. Kato, K., and S. Kitazaki. 2006. "A Study for Understanding Carsickness Based on the Sensory Conflict Theory." SAE Technical Paper. https://doi.org/10.4271/2006-01-0096.
104. Wada, T., and K. Yoshida. 2016. "Effect of Passengers' Active Head Tilt and Opening/Closure of Eyes on Motion Sickness in Lateral Acceleration Environment of Cars." Ergonomics 59(8): 1050–1059. https://doi.org/10.1080/00140139.2015.1109713.
105. Paolo Falcone, Francesco Borrelli, Jahan Asgariy, H. Eric Tsengy, and Davor Hrovat, "A Model Predictive Control Approach for Combined Braking and Steering in Autonomous Vehicles," Mediterranean Control and Automation, 2007.
106. Bernardini. D, Di Cairanoz. S, Bemporad. A, Tsengy. H.E, "Drive-by-wire vehicle stabilization and yaw regulation: a hybrid Model Predictive Control design," IEEE Decision and Control, pp.7621–7626, 2009.
107. Craig E. Beal, "Applications of Model Predictive Control to Vehicle Dynamics for Active Safety and Stability," A Dissertation for the degree of doctor of philosophy, 2011.
108. Gray. A, Gao. Y, Lin. T, J. K. Hedrick, Tseng. E, and Borrelli. F, "Predictive control for agile semi-autonomous ground vehicles using motion primitives," American Control Conference, 2012.
109. M. Canale and L. Fagiano, "Vehicle yaw control using a fast NMPC approach," IEEE Conference on Decision and Control, pp. 5360–5365, 2008.

110. Li. Gang, Z. C. Fu, Z. Hong-Yu and H. Wei, "Vehicle active front steering and yaw moment integrated control," IEEE on transportation, Mechanical, and Electrical engineering, Changchun, China, pp.787–790, 2011.
111. G. Palmieri, P. Falcone, H. E. Tseng and L. Glielmo, "A preliminary study on the effect of roll dynamics in predictive vehicle stability control," IEEE Conference on Decision and Control, Cancun, Mexico, pp. 5354–5359, 2008.
112. R. Mehra, J. Amin, K. Hedrick, C. Osorio, and S. Gopalasamy, "Active suspension using preview information and model predictive control," IEEE Int. Conference on Control Applications, pp. 860–865, 1997.
113. B.K. Cho, G. Ryu, and S. J. Song, "Control strategy of an active suspension for a half car model with preview information," Int. Journal of Automotive Technology, vol. 6, pp. 243–249, 2005.
114. Dahunsi, O. A, Pedro, J. O, and Nyandoro O. T, "Neural Network Based Model Predictive Control of a servo-Hydraulic Vehicle Suspension System," IEEE AFRICON, Nairobi, Kenya, pp.1–6, 2009.
115. Jing xu and Juntao Fei, "Neural Network predictive control of vehicle suspension", IEEE ICISE, Changzhou, China, pp. 1319–1322, 2010..
116. C. Gohrle, A. Wagner, A. Schindler and O. Sawodny, "Active suspension controller using MPC based on full-car model with preview information," American Control Conference, Montreal, Canada, pp.497–502, 2012.
117. Jiang, Y., Chen, B. K., & Thompson, C. (2019). A comparison study of ride comfort indices between Sperling's method and EN 12299. International Journal of Rail Transportation, 7(4), 279–296.
118. Hybrid Simulator for Automotive Studies and Research | LAMIH. Uphf.fr. (2021). Retrieved 8 March 2021, from https://www.uphf.fr/LAMIH/en/SHERPA.
119. Amyx, S. (2021). Future connected cars: Focus on back-seat comfort, versatility. IoT World Today. Retrieved 9 March 2021, from https://www.iotworldtoday.com/2017/06/08/future-connected-cars-focus-back-seat-comfort-versatility/.
120. AG, R. (2021). Rinspeed AG - Re-Think! Re-Invent! Re-Create!. Rinspeed.eu. Retrieved 9 March 2021, from https://www.rinspeed.eu/en/.
121. Jazar, R. N. (2019). Advanced vehicle dynamics. Cham: Springer International Publishing.

Chapter 9
Vehicle Vibration Analysis of the Quarter-Car Model Considering Tire-Road Separation

Dang Quy Nguyen, Sina Milani, Hormoz Marzbani, and Reza N. Jazar

Abstract A real suspension system has inherent nonlinearities that are often ignored when modeling, and a more accurate system than the traditional ones must be able to account for these effects. This chapter examines one such nonlinearity which is the possible separation of the tire from the ground. To accomplish this aim, a nonlinear model will be formulated using nondimensionalization which would incorporate the tire-road discontinuities. The number of input dimensionless parameters is minimized by developing the differential equation system. Once the tire loses contact with the ground, the separation time is a crucial criterion for evaluating ride comfort. Thus, the examination of the time response, frequency response, and fraction of the separation time has been sensitively conducted. The separation dynamics are numerically investigated to determine which values of suspension characteristics can postpone and avoid the tire-road separation phenomenon and eliminate the potentially dangerous vibrations.

Keywords Vehicle vibrations · Quarter-car vibrations · Tire-road separation · Suspension design · Vehicle safety · Ride comfort

9.1 Introduction

In order to study the main vibrating mode of vehicles, a quarter-car model is simple enough to employ and investigate new optimization concepts [10]. Furthermore, Jazar [7] argues that the most useful model of a vehicle suspension to investigate and optimize its body vibrations is a two degrees of freedom (*2DOF*). This model

D. Q. Nguyen (✉)
School of Engineering, RMIT University, Melbourne, VIC, Australia

S. Milani · H. Marzbani · R. N. Jazar
Faculty of Vehicle and Energy Engineering, Le Quy Don Technical University, Hanoi, Vietnam
e-mail: sina.alamdarimilani@rmit.edu.au; hormoz.marzbani@rmit.edu.au; reza.nakahiejazar@rmit.edu.au

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_9

includes the fundamental features of the practical vehicle vibrations, and it is an effective tool to examine the ride behavior and simulate suspension's functionality.

In vehicle suspensions, the tire is often assumed to be in contact with the ground. However, it is inaccurate at high excitation frequencies and road amplitudes; therefore, the possibility of separation phenomenon between the tire and the surface provides a model closer to the real system [6].

The issue of vehicle vibration analysis regarding tire separation is traced back to the 1980s when the researchers suggested that the wheel would not remain in contact with the road at large amplitudes or high frequencies of surface undulations. However, detection and investigation of the separation period by the available differential equations is complicated and requires delicate computer simulations. Thus, Gillespie [4] and Guiggiani [5] state that almost recent mathematical models dismiss the separation condition to avoid complexities for numerical simulation and optimization purposes. Jazar [7] also introduced this shortcoming while mentioning the limits of the linear quarter-car systems. Reza-Kashyzadeh et al. [17] have examined frequency domains of vehicle suspension systems and identified its stability when the tire is in contact with the road. Furthermore, the interaction of the wheel with the surface, having lost contact with the road, has hardly ever been examined [14].

Most of the studies have acknowledged the tire separation phenomenon could happen at high damping ratios and high input amplitudes. Nevertheless, researchers hold distinct attitudes about the tire separation impacts on ride comfort, road holding, and safety for passengers.

Related to the vertical force, Wong [20] has suggested that the condition, where the tire takes off the surface, is important from the attitude of road holding. He stated the normal force decreases to zero when the wheel separates from the ground plane, so the tire would not be creating any forces on the road. Captain, Boghani, and Wormley [2] also applied the criteria that the vertical tire force equals zero when the tire is in no-contact with the road. Wageman [19] has declared that a tire loses contact with the ground at any time, and the tire becomes incapable of generating force. Moreover, Jiao [9] pointed out that the vertical force acting on the road from the tire equals zero until the tire gets in contact with the ground again. Reza-Kashyzadeh et al. [17] and Dobaj [3] also concluded that these forces equal to zero when the wheel jumps from the road because the friction force is dependent on the vertical tire load.

The traveling velocity has also an impact on the tire-road separation. Lindgren [11] suggested that a wheel could leave the surface for a while at high speed. However, as indicated by Barbosa [1], by increasing traveling speed, the likelihood of separation phenomenon is possible. A computer simulation of the suspension vibration, when its tire has jumped over from the road by changing suddenly longitudinal speed, is also investigated in [13]. Khazaie et al. [10] pointed out that when the suspension system is forced by vertical excitations traveling on large undulations with high velocities, its tire can jump from the road, which could lead to breaking the contact with the surface.

Wageman [19] declared that the separation between the tire and the road might occur in several cases, such as the ground having a large bump to launch the vehicle into the air or suspension setups that are prone to lifting the outer wheels during cornering. Jiao [9] stated that when the vertical motion of the unsprung mass is too high, it will lead to the tire being no more in contact with the surface. In addition, Reza-Kashyzadeh et al. [17] and Dobaj [3] suggested that if the sprung mass has a large vertical deflection, the tire may lose contact with the ground. Milliken, Milliken, and Olley [15] also suggested that the wheel can be thrown clear of the road profile at several periods. Once the separation appears, flying times of the wheel have been measured to validate the quality of the vehicle suspension. Moreno-Ramírez and Tomas-Rodríguez [16] claimed that a shorter separation time would present a better design and increase the control over road perturbations. The vehicle vibration with four tires that have lost contact with the road has also been discussed in [12].

As presented by Khazaie et al. [10], the vehicle suspension system has only two states, namely, in-contact and no-contact. These two states are entirely distinct both physically and mathematically; therefore, the governing equation system of oscillations will be discontinuous or piecewise. To determine the separation, they conclude that an unloaded tire would lose contact with the road when the gap from the wheel center to the road surface is greater than the wheel's relaxed radius. Similarly, Svahn, Jerrelind, and Dankowicz [18] also argued that the tire is in contact with the ground if the unsprung mass has a vertical motion smaller than the relaxed radius of the tire. By contrast, the free fall will happen if the unsprung mass has displaced a distance greater than the geometric radial distance. Wageman [19] claimed that loss of contact is possible to occur approximately at the peak dynamic deflection of the tire and the separation phenomenon could appear close to two natural frequencies corresponding to the 2DOF system.

A tire-road separation period has a major impact on losing vertical forces between the tire and the ground, which results in a loss of the traction force. Thus, the system would completely lose control of its steering and braking system. Furthermore, the separation phenomenon causes also a significant decrease in ride comfort, road holding, and vehicle safety. With the considerations explained above, this chapter will identify proper suspension and road parameters to suppress the potentially dangerous oscillation, provide ride comfort for passengers, as well as minimize the no-contact time for improving vehicle handling and stability capacities. The approach of analysis and most results could also be employed on any other multi DOF system.

9.2 Dynamic Equations of Motion

The *2DOF* quarter-car model as shown in Fig. 9.1 is used to study the vertical dynamics of vehicles considering tire-road separation [6]. The masses of the system symbolized by m_s and m_u are sprung and unsprung masses, respectively. The

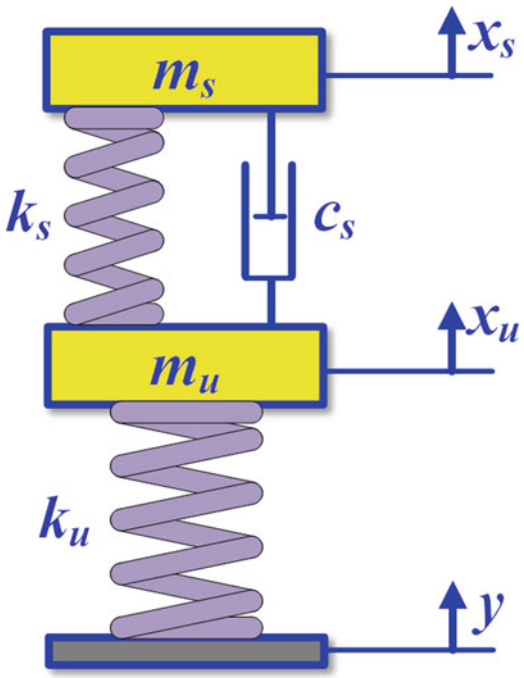


Fig. 9.1 A quarter-car model and free body diagram

stiffnesses of the tire and main spring are k_u, k_s. Neglecting the tire's damping to simplify the system does not affect the accuracy of its vertical response [18], because the value of the tire damping, compared to the primary suspension damping, is too small; therefore, the only damper is assumed to be the main shock absorber of the suspension presented by c_s. The displacements of the masses and the road are denoted by x_s, x_u, and y.

The equations of motion of the sprung and unsprung masses in the contact state are:

$$m_s\ddot{x}_s + c_s\left(\dot{x}_s - \dot{x}_u\right) + k_s\left(x_s - x_u\right) = 0 \tag{9.1}$$

$$m_u\ddot{x}_u + c_s\left(\dot{x}_u - \dot{x}_s\right) + \left(k_u + k_s\right)x_u - k_s x_s = k_u y \tag{9.2}$$

The tire is in contact with the road when the vertical displacement of the tire is smaller than its relaxed radial dimension. However, the no-contact state only occurs if the tire displaces a distance greater than the relaxed radius [18]. Therefore, the separation condition is the relative displacement between displacements of the unsprung mass and road excitation starting greater than the static tire compression:

$$\begin{aligned} x_u - y &\geq x_{ST} \text{ no-contact} \\ x_u - y &< x_{ST} \text{ in-contact} \end{aligned} \tag{9.3}$$

where $x_{ST} = \frac{(m_u+m_s)g}{k_u}$

In the no-contact state (3), the motion of the unsprung mass follows [10]:

$$m_u\ddot{x}_u + c_s\left(\dot{x}_u - \dot{x}_s\right) + k_s\left(x_u - x_s\right) + \left(m_u + m_s\right)g = 0 \tag{9.4}$$

The motion equations are switched from (9.1) and (9.2) to (9.1) and (9.4) for the sprung and unsprung masses, respectively, when the separation condition (9.3) happens. Therefore, the system would be nonlinear and more complicated.

9.3 Time Response

To investigate vertical vibrations in the dimensional field, assuming that the system vibrates harmonically with a function , where y_o is the road amplitude, ω is the input frequency.

Using parametric inputs of a practical quarter-car model in Table 9.1 [10], the time response presents in Fig. 9.2. The separation phenomenon occurs at a high frequency $\omega = 30$ *rad/s* as shown in Fig. 9.2a; the indicator variable I identifies which is the separation state (I = 0.2) or the contact state (I = 0). Figure 9.2c indicates the time response in an in-contact zone while, Fig. 9.3b denotes that a transient separation can appear at the initial period. However, the separation also happens at high road amplitudes as described in Fig. 9.3d.

Once the tire loses its road contact, the displacements of sprung and unsprung masses x_s, x_u unpredict in a linear system. An example of the change of x_s, x_u depicts in Fig. 9.3; the solid curves show the displacements of the real system compared to dotted curves without the separation assumption; they only coincide in each contact period (I = 0). Using such a system improves the modeling accuracy; however, the system becomes more complicated to investigate the dynamic response in a long computational time with the discontinuous equation system. To solve this issue, the number of input parameters needs to minimize by introducing nondimensional parameters.

$$X_u = \frac{x_u}{x_0} \qquad X_s = \frac{x_s}{x_0} \qquad \tau = \omega_s t \tag{9.5}$$

Taking derivative to obtain:

Table 9.1 The data of a quarter-car model

Parameters	Value [unit]	Parameters	Value [unit]
m_s	49.5 kg	k_s	15,760 N/m
m_u	5.5 kg	k_u	197,900 N/m
c_s	3000 ns/m	x_{ST}	0.0027 m

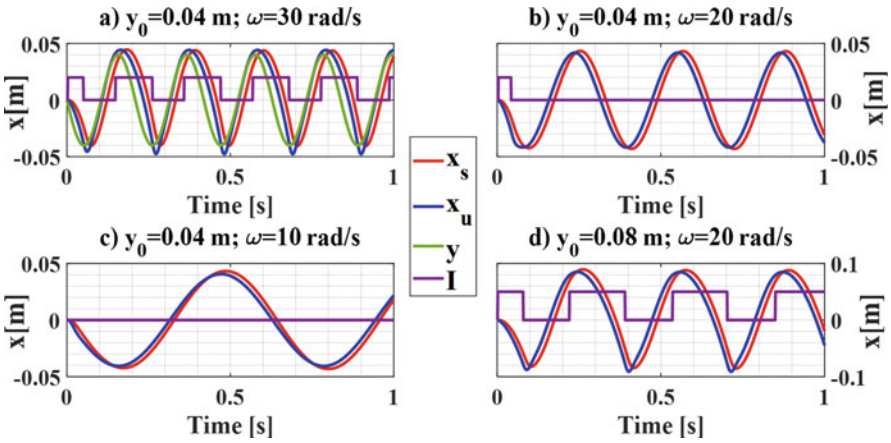


Fig. 9.2 Time response in contact and no-contact zones

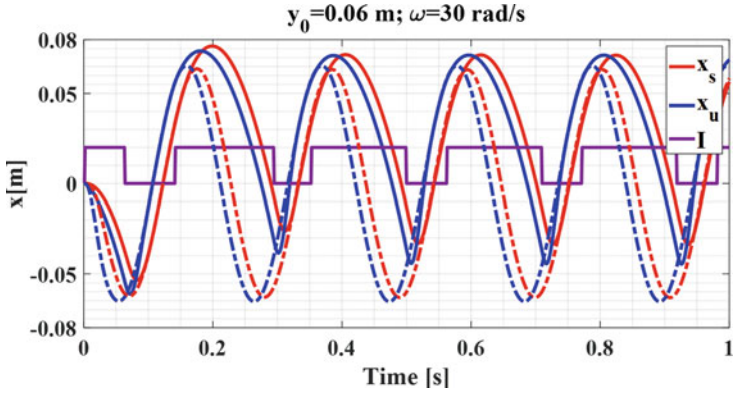


Fig. 9.3 Comparison of displacements between contact and no-contact assumptions using the characteristic parameter $x_0 = x_{ST}$ to make the new variables

$$\dot{x}_u = \frac{dx_u}{dt} = \omega_s x_0 X_u' = \omega_s x_0 \frac{dX_u}{d\tau}; \dot{x}_s = \frac{dx_s}{dt} = \omega_s x_0 X_s' = \omega_s x_0 \frac{dX_s}{d\tau}$$

$$\ddot{x}_u = \frac{dX_u'}{dt} = \omega_s^2 x_0 X_u'' = \omega_s^2 x_0 \frac{d^2 X_u}{d\tau^2}; \ddot{x}_s = \frac{dX_s'}{dt} = \omega_s^2 x_0 X_s'' = \omega_s^2 x_0 \frac{d^2 X_s}{d\tau^2}$$

Introducing the following nondimensional parameters:

$$\alpha = \frac{\omega_u}{\omega_s} \quad r = \frac{\omega}{\omega_s} \quad \xi = \frac{c_s}{3m_s w_s} \quad \varepsilon = \frac{m_s}{m_u} \quad Y = \frac{y}{x_0} \tag{9.6}$$

$$\omega_s = \sqrt{\frac{k_s}{m_s}} \qquad \omega_u = \sqrt{\frac{k_u}{m_u}} \qquad X_{ST} = \frac{x_{ST}}{x_0} = 1 \qquad X_R = \frac{x_R}{x_0} = n \tag{9.7}$$

where α can be called the stiffness ratio because α is a function of the stiffnesses below:

$$\alpha = \sqrt{\frac{\varepsilon k_u}{k_s}} \tag{9.8}$$

The dimensionless parameters for existing design quarter-car models are $\alpha = [110]$, $\varepsilon = [220]$, $\xi = [0\,2]$, and $r = [0\ 200]$ [6]. However, the tire-road separation will be examined with typical parameters, for the sake of making the figures understandable.

Based on these introducing parameters, the motion equation of the sprung mass is rewritten:

$$\begin{aligned} &m_s w_s^2 x_0 X_s'' + c_s w_s x_0 \left(X_s' - X_u'\right) + k_s x_0 \left(X_s - X_u\right) = 0 \\ &X_s'' + \tfrac{c_s}{m_s w_s} \left(X_s X_u'\right) + \tfrac{k_s}{m_s w_s^2} \left(X_s - X_u\right) = 0 \\ &X_s'' + 2\xi \left(X_s' - X_u'\right) + \left(X_s - X_u\right) = 0 \end{aligned} \tag{9.9}$$

The motion equation of the unsprung mass in the contact state is equivalent:

$$\begin{aligned} &m_u w_s^2 x_0 X_u' + c_s w_s x_0 \left(X_u' - X_s'\right) + \left(k_s + k_u\right) x_0 X_u - k_s x_0 X_s = k_u x_0 Y \\ &X_u'' + \tfrac{c_s}{m_u w_s} \left(X_u' - X_s'\right) + \tfrac{k_s + k_u}{m_u w_s^2} X_u - \tfrac{k_s}{m_u w_s^2} X_s = \tfrac{k_u}{m_u w_s^2} Y \\ &X_u'' + 2\omega\xi \left(X_u' - X_s'\right) + \varepsilon \left(X_u - X_s\right) + \alpha^2 \left(X_u - Y\right) = 0 \end{aligned} \tag{9.10}$$

and for the no-contact state:

$$\begin{aligned} &m_u w_s^2 x_0 X_u'' + c_s w_s x_0 \left(X_u' - X_s'\right) + k_s x_0 \left(X_u - X_s\right) + \left(m_u + m_s\right) g = 0 \\ &X_u' + \tfrac{c_s}{m_u w_s} \left(X_u' - X_s'\right) + \tfrac{k_s}{m_u w_s^2} \left(X_u - X_s\right) + \tfrac{m_u + m_s}{m_u w_s^2 x_0} g = 0 \\ &X_u'' + 2\varepsilon\xi \left(X_u' - X_s'\right) + \varepsilon \left(X_u - X_s\right) + \alpha^2 = 0 \end{aligned} \tag{9.11}$$

The condition of switching between the equations is:

$$\begin{aligned} X_u - Y \geq 1 = X_{ST} \quad &\text{no-contact} \\ X_u - Y < 1 = X_{ST} \quad &\text{in-contact} \end{aligned} \tag{9.12}$$

Nondimensionalization transforms three variables x_u, x_s, t and five original suspension parameters m_u, m_s, k_u, k_s, c_s to three new variables X_u, X_s, τ and three parametric inputs α, ε, ξ. Thus, the complexity of the investigation would decrease

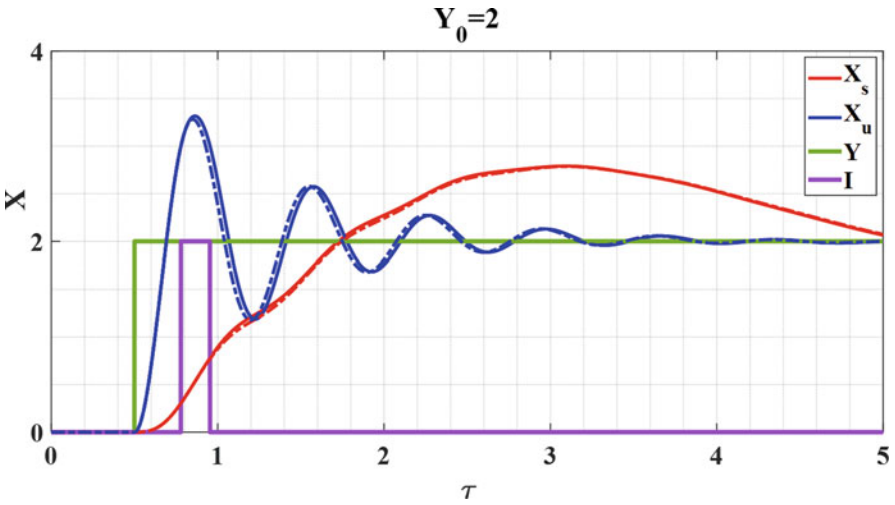


Fig. 9.4 Time response for a step excitation

significantly and could facilitate the scale-up of numerical results to real conditions [8].

In transient response analysis, the road excitation Y is assumed to be a step input, a half sine wave, as well as a bump. The step function used for the separation examination is given below:

$$\begin{aligned} Y &= \frac{y}{x_0} = \frac{y_0}{x_0} = Y_0 \ (\tau \geq \tau_0) \\ Y &= 0 \ (\tau < \tau_0) \end{aligned} \tag{9.13}$$

The transient response for a step input is denoted in Fig. 9.4; the wheel only separates in a while after hitting a step of height $Y_0 = 2$. Comparison and examination of the time response plots denote that the oscillation frequencies of the tire X_u is higher than the oscillation frequencies of the sprung mass X_s. The tire also stops vibrating much faster than the body. The displacements of the sprung and unsprung masses X_s, X_u shown by solid curves deviate slightly from dashed curves with no-separation assumption.

The examination and optimization of the transient response are less important than the frequency domain because the dynamic behaviors change remarkably in responses for the system going over a harmonic input. Therefore, we will observe the time response in a sinusoidal function:

$$Y = \frac{y}{x_0} = \frac{y_0 cos\omega t}{x_0} = Y_0 cosr\tau \tag{9.14}$$

The road amplitude is multiple Y_0 of the static tire compression, X_{ST}, usually one to three times. For a practical amplitude $Y_0 = 1$, the time response of given

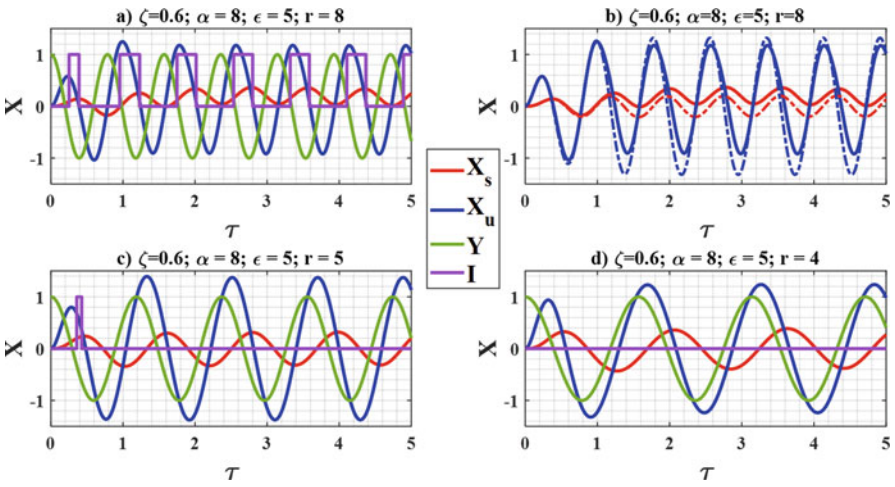


Fig. 9.5 Time response in the nondimensional field

nondimensional parameters is presented in Fig. 9.5. The tire always contacts with the ground at a low frequency ratio, $r = 4$, as depicted in Fig. 9.5d, while it can separate from the surface at a high frequency ratio, $r = 8$. When the separation appears, the displacements X_s, X_u shown by the solid curves in Fig. 9.5a deviate from the dotted curves in Fig. 9.5b with no-separation. The transient separation in Fig. 9.5c will be ignored to calculate the separation time in the next section.

The vertical velocity and acceleration of the sprung and unsprung masses are vital criteria to optimize damping and mainspring forces. The velocity and acceleration responses will be more correct if the separation dynamics are considered, and they are expressed by solid curves in Figs. 9.6 and 9.7 compared to the dotted curves without the separation.

Using the input parameters from Fig. 9.7 and 9.8 reveals phase portraits after a transient period. The phase portraits between vertical acceleration A and vertical velocity V of the sprung and unsprung masses are expressed in Fig. 9.8a, b while Fig. 9.8c, d demonstrate the phase portraits of velocity V and displacement X. The acceleration and velocity are harmonic outputs without separation assumption providing an elliptic relationship, but it will change in a real system considering the separation.

9.4 Frequency Domain

The frequency response program evaluates the time response over a range of input frequencies, road amplitudes, and damping ratios. The maximum from the steady-state result matrix is calculated to obtain the output amplitudes of sprung and

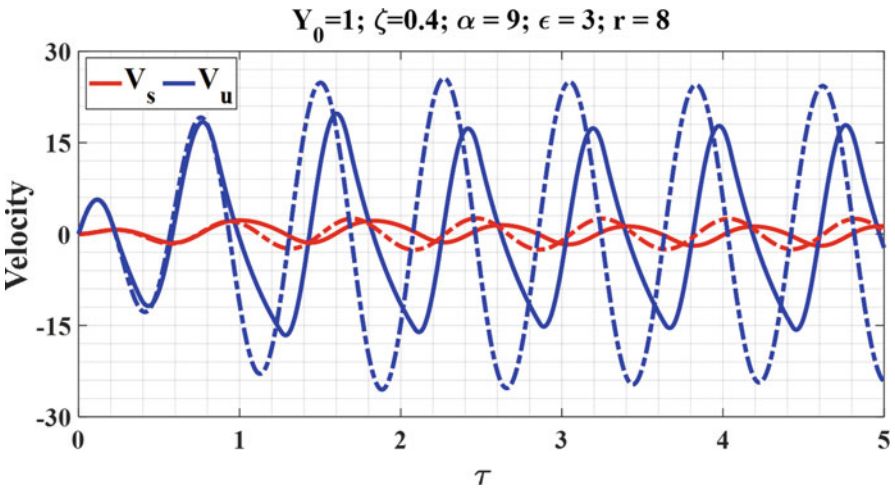


Fig. 9.6 Vertical velocities in a separation zone

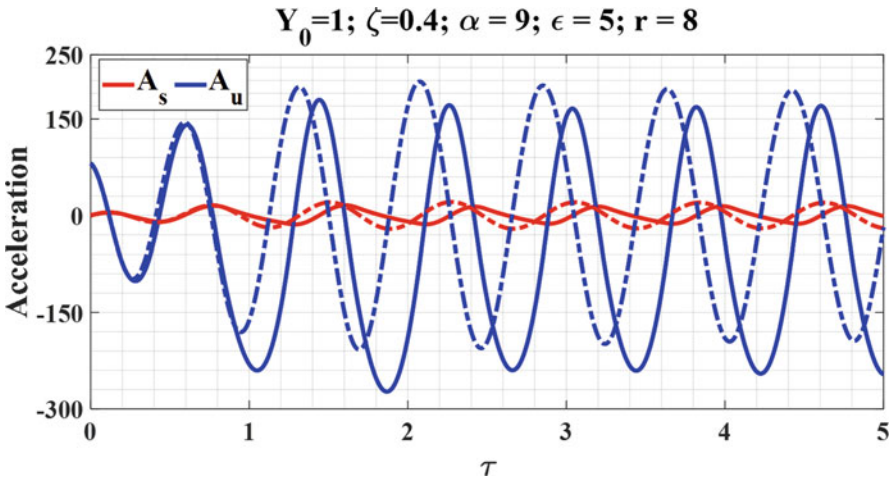


Fig. 9.7 Vertical accelerations in a separation zone

unsprung masses. The frequency response is investigated for the system traveling on harmonic excitation (14). In the in-contact dynamics, the displacements X_s and X_u will be a harmonic form:

$$X_s = A_1 \sin r\tau + B_1 \cos r\tau = |X_s| \sin\ (r\tau + \varphi_s) \tag{9.15}$$

$$X_u = A_2 \sin r\tau + B_2 \cos r\tau = |X_u| \sin\ (r\tau + \varphi_u) \tag{9.16}$$

where the phases of X_s, X_u are:

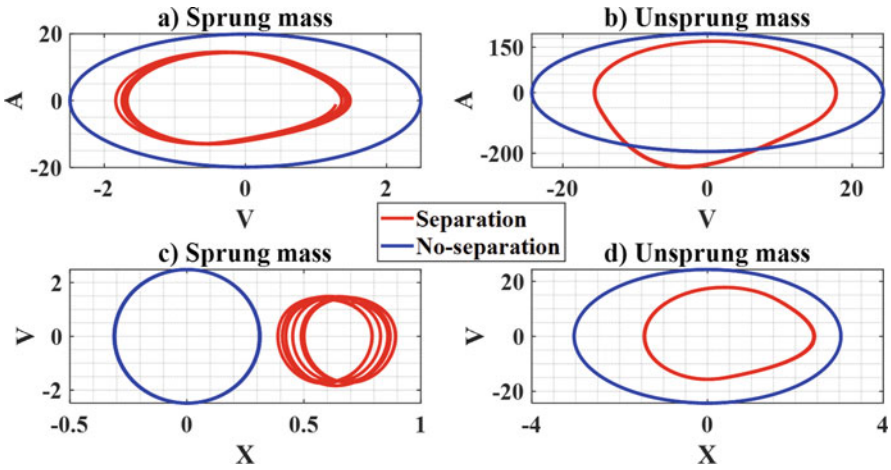


Fig. 9.8 Phase portraits in a no-contact zone

$$\tan\varphi_s = \frac{B_1}{A_1};\ \tan\varphi_u = \frac{B_2}{A_2} \tag{9.17}$$

Substituting X_s, X_u in the in-contact Eqs. (9.9 and 9.10) makes two algebraic equations.

$$\begin{aligned}&\left(A_1 - A_2 - 2\xi r\left(B_1 - B_2\right) - r^2 A_1\right)\sin r\tau \\ &+\left(B_1 - B_2 - 2\xi r\left(A_1 - A_2\right) - r^2 B_1\right)\cos r\tau = 0\end{aligned} \tag{9.18}$$

$$\begin{aligned}&\left(\alpha^2 A_2 - r^2 A_2 - \varepsilon\left(A_1 - A_2\right) + 2\xi\varepsilon r\left(B_1 - B_2\right)\right)\sin r\tau \\ &+\left(\alpha^2 B_2 - r^2 B_2 - \varepsilon\left(B_1 - B_2\right) - 2\xi\varepsilon r\left(A_1 - A_2\right)\right)\cos r\tau - \alpha^2 Y_0\cos r\tau = 0\end{aligned} \tag{9.19}$$

An algebraic balance of the coefficients of $sin\,\omega\,t$, $cos\,\omega\,t$ in these equations will search the unknown coefficients:

$$[A]\begin{bmatrix}A_1\\B_1\\A_2\\B_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\\\alpha^2 Y_0\end{bmatrix} \tag{9.20}$$

where

$$[A] = \begin{bmatrix}1-r^2 & -2re\xi & -1 & 2r\xi \\ 2r\xi & 1-r^2 & -2r\xi & -1 \\ -\varepsilon & 2r\xi\varepsilon & \alpha^2 - r^2 + \varepsilon & -2r\xi\varepsilon \\ -2r\xi\varepsilon & -\varepsilon & 2r\xi\varepsilon & \alpha^2 - r^2 + \varepsilon\end{bmatrix} \tag{9.21}$$

The coefficients are:

$$A_1 = \frac{2r^3\xi\alpha^2\left(\alpha^2 - r^2\right)}{Z_1^2 + Z_2^2}Y_0 \tag{9.22}$$

$$B_1 = \frac{r^2\left(r^2 - \alpha^2\right) - \left(4r^2\xi^2 + 1\right)\left(r^2\left(1 + \varepsilon\right) - \alpha^2\right)}{Z_1^2 + Z_2^2}\alpha^2 Y_0 \tag{9.23}$$

$$A_2 = \frac{2r^5\alpha^2\xi\varepsilon}{Z_1^2 + Z_2^2}Y_0 \tag{9.24}$$

$$B_2 = \frac{r^2\left(r^2\left(\alpha^2 - r^2 + \varepsilon + 2\right) - 2\alpha^2\right) - \left(4r^2\xi^2 + 1\right)\left(r^2\left(1 + \varepsilon\right) - \alpha^2\right)}{Z_1^2 + Z_2^2}\alpha^2 Y_0 \tag{9.25}$$

The modulus of oscillations of the sprung and unsprung masses in the contact state would be:

$$|X_s|^2 = A_1^2 + B_1^2 = \frac{4r^2\xi^2 + 1}{Z_1^2 + Z_2^2}\alpha^4 Y_0^2;\ |X_s| = \sqrt{A_1^2 + B_1^2} \tag{9.26}$$

$$|X_u|^2 = A_2^2 + B_2^2 = \frac{4r^2\xi^2 + r^2\left(r^2 - 2\right) + 1}{Z_1^2 + Z_2^2}\alpha^4 Y_0^2;\ |X_u| = \sqrt{A_2^2 + B_2^2} \tag{9.27}$$

where:

$$Z_1 = r^2\left(r^2 - \alpha^2\right) + \left(\alpha^2 - \left(1 + \varepsilon\right)r^2\right) \tag{9.28}$$

$$Z_2 = 2r\xi\left(\alpha^2 - \left(1 + \varepsilon\right)r^2\right) \tag{9.29}$$

The amplitudes X_s, X_u of an undamped system $\xi = 0$ for resonant situations tends to infinity at natural frequency ratios r; the resonance appears when Z_1 will be zero at the natural frequency ratios.

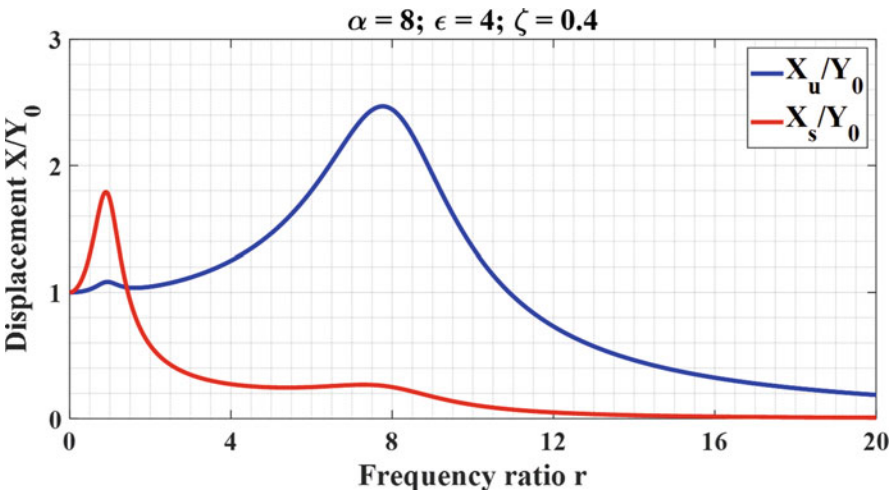


Fig. 9.9 Frequency response with no-separation assumption

$$r_{n1} = \sqrt{\frac{1}{2}\left(1+\varepsilon+\alpha^2\right) - \frac{1}{2}\sqrt{\left(1+\varepsilon+\alpha^2\right)^2 - 4\alpha^2}} \tag{9.30}$$

$$r_{n2} = \sqrt{\frac{1}{2}\left(1+\varepsilon+\alpha^2\right) + \frac{1}{2}\sqrt{\left(1+\varepsilon+\alpha^2\right)^2 - 4\alpha^2}} \tag{9.31}$$

The vertical motions are proportional to amplitude Y_0 proving the linear system in a contact state. Based on the Eqs. (9.26) and (9.27), the original variables x_s/y_0, x_u/y_0 of the in-contact frequency response could plot by utilizing new ones X_s/Y_0 and X_u/Y_0:

$$\frac{X_s}{Y_0} = \frac{x_s}{y_0} \qquad \frac{X_u}{Y_0} = \frac{x_u}{y_0} \tag{9.32}$$

An example of the frequency domain X_s/Y_0, X_u/Y_0 as functions of input frequency ratio r with no-separation assumption is illustrated in Fig. 9.9. There are two maxima corresponding to the 2DOF system; indeed, the fluctuations of the dynamic behaviors match the natural frequency ratios calculated from (9.30) and (9.31). The body bounce X_s is obvious when r closes to $r_{n1} = 0.97$, while the wheel's hoop is obvious when $r$extends to $r_{n2} = 8.25$.

Considering the contact-free, the real response must be recalculated from (9.9) and (9.11). A comparison of the vertical oscillations between contact and in-contact assumptions is demonstrated in Fig. 9.10, and its dashed curves show the oscillations X_s/Y_0, X_u/Y_0 in Fig. 9.9. The wheel's hoop of the real system happens earlier around$r_{n2} = 6.5$. Furthermore, the ride comfort reduces significantly because the

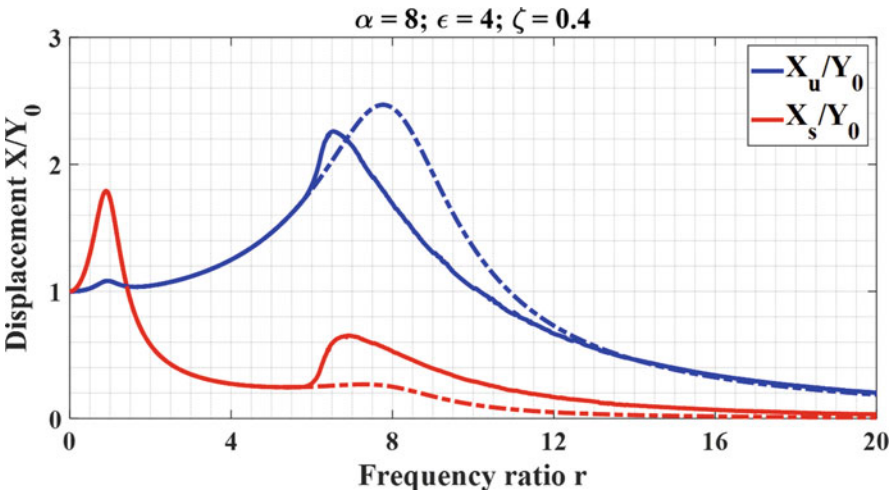


Fig. 9.10 Frequency response of real system considering the separation

body vibration X_s/Y_0 deviates suddenly after a point $r = 5.75$; it is called the separation point, and an identification of the separation position will be revealed in the next section. Therefore, identifying the maximum achievable stability responses is also a vital step forward in terms of vehicle safety.

The dynamic responses of the system are sensitive to the input frequency and damping ratios. Figure 9.11 illustrates the frequency response of unsprung mass in response to different damping ratios. In the linear models, the output amplitudes do pass through the two nodes. Considering the separation phenomenon, however, the response for a critically damped system $\xi = 1.0$ does not pass through the second node. Furthermore, the amplitude of unsprung mass is just proportion to the value of damping ratio after the first node at $r_1 = 0.91$ and before the second node $r_2 = 2.81$.

Figure 9.12 shows that the frequency response of sprung mass for all damping coefficients does pass through two nodes. The amplitude of sprung mass is just proportion to the value of the damping ratio after the first node at $r_1 = 1.21$ and before the second node at $r_2 = 4.34$.

9.5 Separation Boundary

A boundary of losing tire-road contact will be identified in the in-contact frequency response. The boundary involves a set of separation points where the relative displacement of the unsprung mass and the road excitation reaches the tire static deflection. To determine the boundary, a new variable of the relative displacement is introduced:

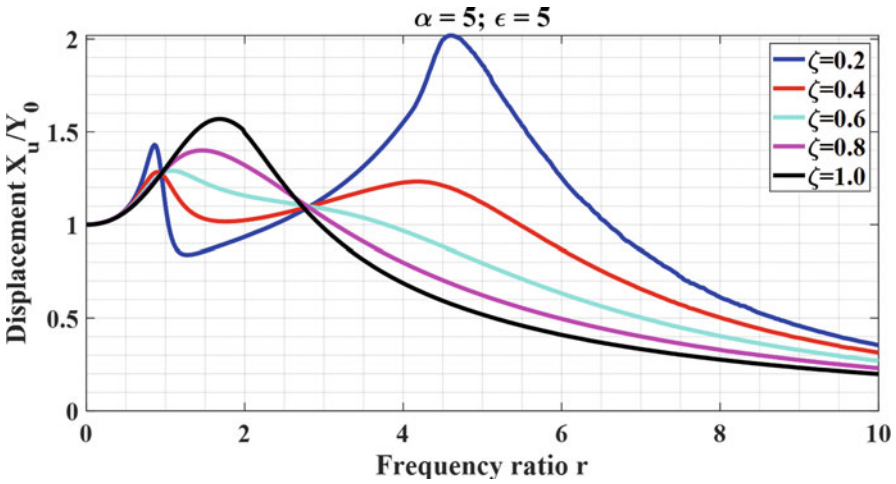


Fig. 9.11 Frequency responses of unsprung mass regarding different damping ratios

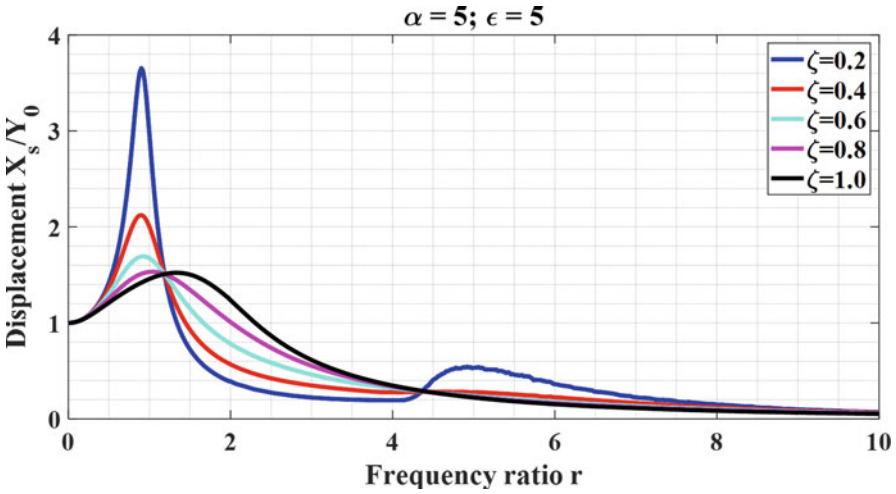


Fig. 9.12 Frequency responses of sprung mass regarding different damping ratios

$$x_t = x_u - y_0 \cos r\ \tau \tag{9.33}$$

The separation condition (9.12) is equivalent to $X_t = 1$ when the road separation occurs, separating the regions of in-contact $X_t < 1$ and no-contact $X_t > 1$.

$$X_t = X_u - Y = A_2 \sin r\ \tau + (B_2 - Y_0) \cos r\tau = |X_t| \sin\ (r\tau + \varphi_t) \tag{9.34}$$

where the phase of relative displacement is identified as $tan\ \varphi_t = (B_2 - Y_0)/A_2$.

The solution provides the quadratic relative amplitude:

$$|X_t|^2 = A_3^2 + B_3^2 = \frac{4\xi^2 r^2(\varepsilon+1)^2 + \left(r^2-1\right)^2 + \varepsilon^2 - 2\varepsilon\left(r^2-1\right)}{Z_1^3 + Z_2^2} r^4 Y_0^2 \tag{9.35}$$

The amplitude of the relative displacement for the in-contact system would be:

$$|X_t| = \sqrt{A_3^1 + B_3^2} \tag{9.36}$$

In the nonlinear system, the relative displacement X_t is plotted by using the motion Eqs. (9.10) and (9.11) while substituting $X_t = X_u - Y$, while X_t of the in-contact frequency response can be plotted by employing directly a new dimensionless variable X_t/Y_0 as a function of the input frequency ratio r shown in (25) and we have:

$$\frac{X_t}{Y_0} = \frac{x_t}{y_0} \tag{9.37}$$

Figure 9.13 describes the real relative frequency response as shown by the solid curve, which will deviate significantly from the dotted curve with no-separation assumption after the position S corresponding to $r = 5.75$. After the relative displacement X_t hits the horizontal line $y = 1$, the dynamic responses of the linear system would be invalid; this interaction indicates the limiting frequency ratio called the separation point. The no-contact zone SZ appears after this point meaning that the system would be in no-contact with the road after passing this point.

The decreasing value of the stiffness ratio could obtain two separation zones as denoted in Fig. 9.14. With $\alpha = 3$, the relative displacement of the real system (solid curve) has a rapid fluctuation around the first natural frequency compared to the response ignoring the separation (dotted curve). Consequently, the dynamic response has a considerable error even at a low frequency ratio if the tire-road separation is dismissed. The frequency response is only correct when the curve of the relative displacement is under the line $y = 1$. There are three intersections – S, P, and Q – indicating contact zone for $r < 0.74$ or $1.21 < r < 2.38$ and no-contact zone for $0.74 < r < 1.21$ or $r > 2.38$.

The separation condition would be $X_t = |X_t| \sin(r\tau + \varphi_t) \geq 1$ while $\sin(r\tau + \varphi_t) \leq 1$; therefore, the separation happens when $|X_t| \geq 1$; and the separation boundary can be determined directly by solving the implicit function: $|X_t| = 1$.

The stiffness ratio is a crucial feature to study and optimize behaviors of the suspension system. Therefore, we need to find separation points with respect to the stiffness and excitation frequency ratios. A set of the separation points is called the separation boundary.

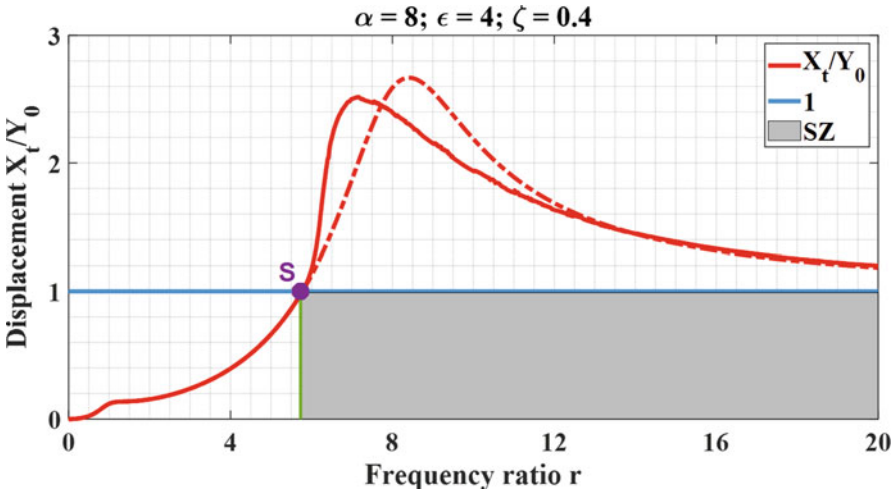


Fig. 9.13 The relative frequency response with one separation point

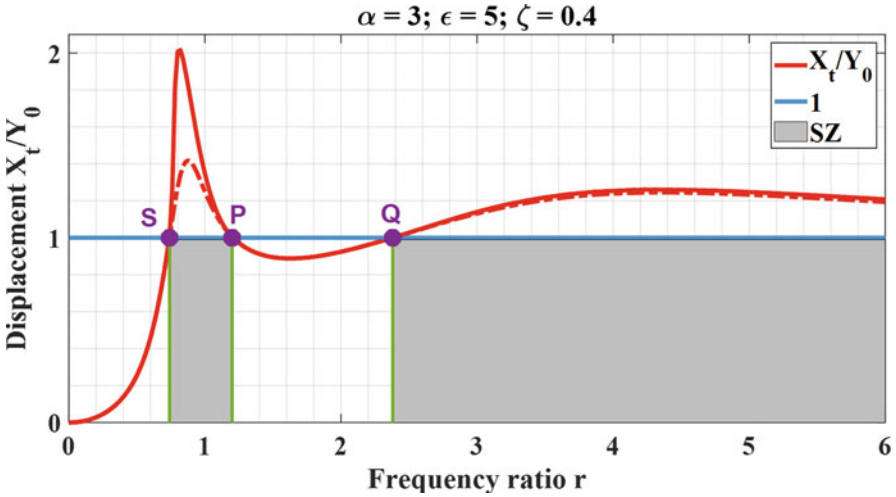


Fig. 9.14 The relative frequency response with three separation points

For a nominal value of $\varepsilon = 4$ and a given damping ratio, say $\xi = 0.4$, the effect of the stiffness ratio α on the no-contact zone is only a continued curve regarding a range of damping ratio as presented in Fig. 9.15. The curve divides the plane (r, α) to two areas, every point P$(r = 2, \alpha = 2)$ under the solid curve at which the wheel keeps contacting with the ground, whereas, at above the condition, the wheel starts separating from the surface. Therefore, the curve is called the separation boundary. The graph is useful to suggest an optimal technique in a step forward to improve the contact capacity of vehicles by reducing the no-contact area.

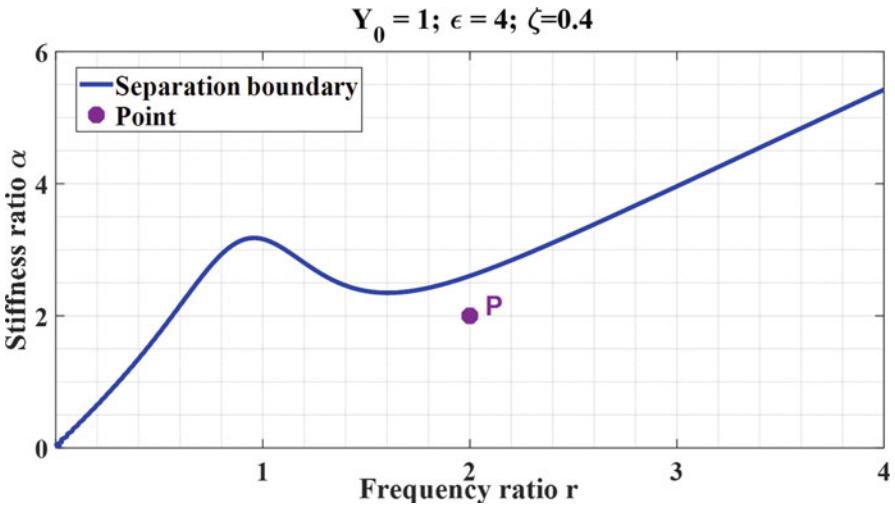


Fig. 9.15 The separation boundary regarding stiffness ratio

With a low damping ratio, $\xi = 0.2$, the separation boundaries regarding the stiffness ratio α will be parabolic and exponential curves as depicted in Fig. 9.16. These curves split the plane (r, α) into contact zone and separation zones including every point which is under the parabolic path or to the right of the exponential branch. Each separation point on the curve can be identified for every value of the stiffness ratio, for instance, $\alpha = 3$; there would be three intersecting points, $r_1 = 0.73$, $r_2 = 1.02$, and $r_3 = 2.86$, indicating two regions: in-contact for $r < 0.73$ or $1.02 < r < 2.86$ and no-contact for $0.73 < r < 1.02$ or $r > 2.86$.

Figure 9.17 demonstrates the impact of the stiffness ratio α on the in-contact zone by changing the damping ratios ξ. For each ξ, the in-contact zone is above the corresponding curve, while the below area of the curve presents a no-contact zone; therefore, the point $P(r = 2, \alpha = 2)$ is in the separation zone with $\xi = 0.4$, but it is in the contact zone with $\xi = 0.3$. It can be inferred from the graphs that a reduction of the damping ratio leads to a reduced separation zone.

Increasing the value of mass ratio ε can reduce considerably the separation zone as compared in Fig. 9.18. For example, keep a constant damping ratio $\xi = 0.3$; the point $P(r = 2, \alpha = 2)$ lies in the contact zone with $\varepsilon = 4$ in Fig. 9.18a, while it is in the in-contact zone in Fig. 9.18d with $\varepsilon = 10$.

Figure 9.19 denotes the separation frequency depending on the road inputs. The relationship between the input amplitude Y_0 and frequency ratio r has the same trend at high damping ratios, and the separation frequency for a critically damped system $\xi = 1.0$ seems to be linear, whereas they deviate from the trend at low damping ratios $\xi = 0.2$, $\xi = 0.4$. The tire separates from the ground at conditions above each separation curve; let us explain by using a checking point $P(r = 2, Y_0 = 2)$; this point is just in the contact zone with $\xi = 0.2$, and it is in the

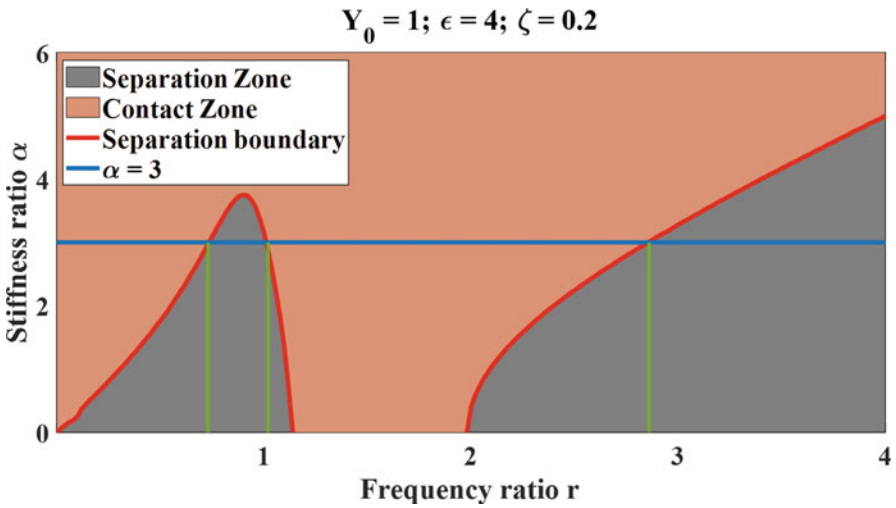


Fig. 9.16 The separation boundary with a low damping ratio

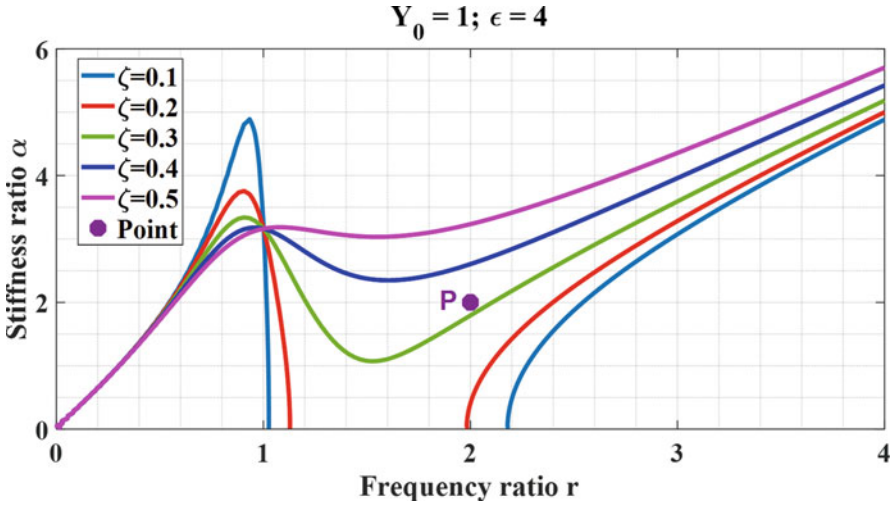


Fig. 9.17 Separation boundary regarding stiffness ratio with different damping ratios

separation zone for all other damping ratios. Furthermore, for $Y_0 = 1$, the horizontal line $Y_0 = 1$ will cut the five separation curves at five points in Figure 9.19. For instance, with a low damping ratio $\xi = 0.2$ the separation point is at $r = 4.23$, while it is at $r = 1.45$ with a high damping ratio $\xi = 1.0$. Consequently, a low damping ratio is beneficial not only for reducing the separation zone but also for postponing the separation at low input amplitudes.

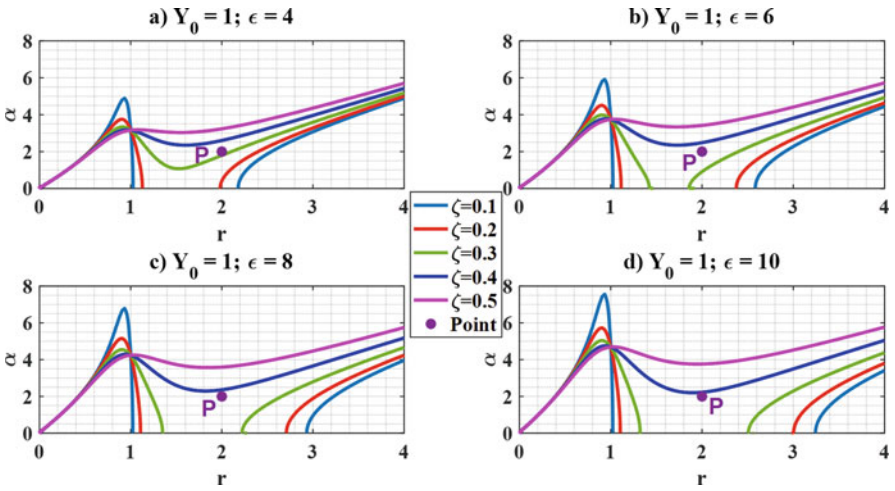


Fig. 9.18 Separation boundary regarding stiffness ratio by increasing mass ratio

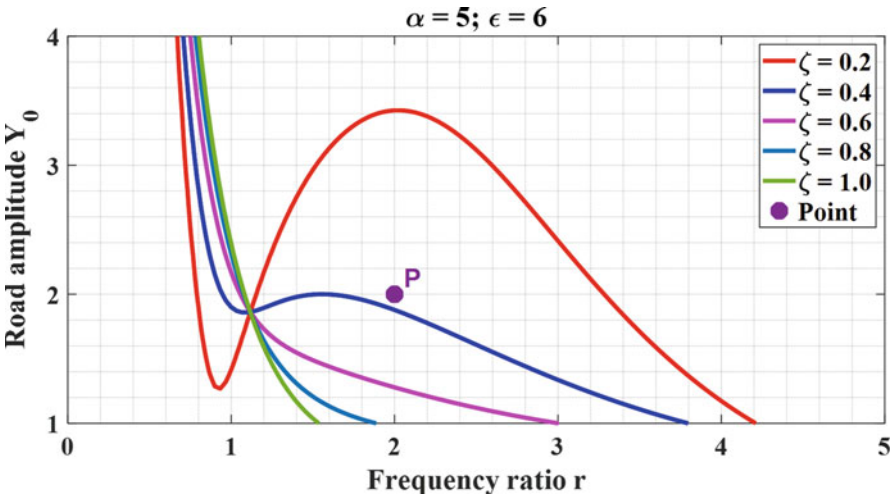


Fig. 9.19 Separation boundary regarding road characteristics

The mass ratio is also another important suspension characteristic; a relationship between the mass and frequency ratios in responding for separation condition has been built. Figure 9.20 shows that the separation boundary for a given damping ratio $\xi = 0.6$ is determined by using the mass ratio ε as a function of the frequency ratio. The contact zone is left of the separation curve, while the separation phenomenon occurs at higher frequency ratios.

For $\xi = 0.4$, Fig. 9.21 reveals a separation boundary regarding the mass ratio for a set of nominated parameters with in-contact and contact zones. The

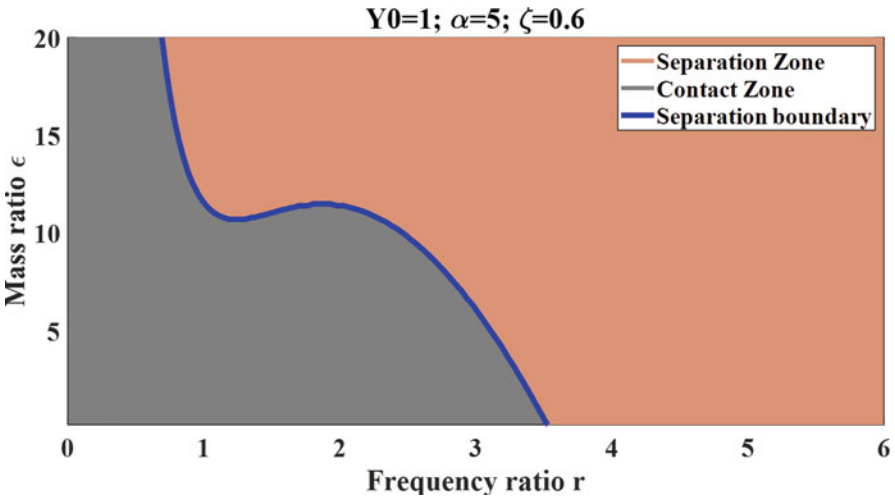


Fig. 9.20 Separation boundary regarding mass ratio with $\xi = 0.6$

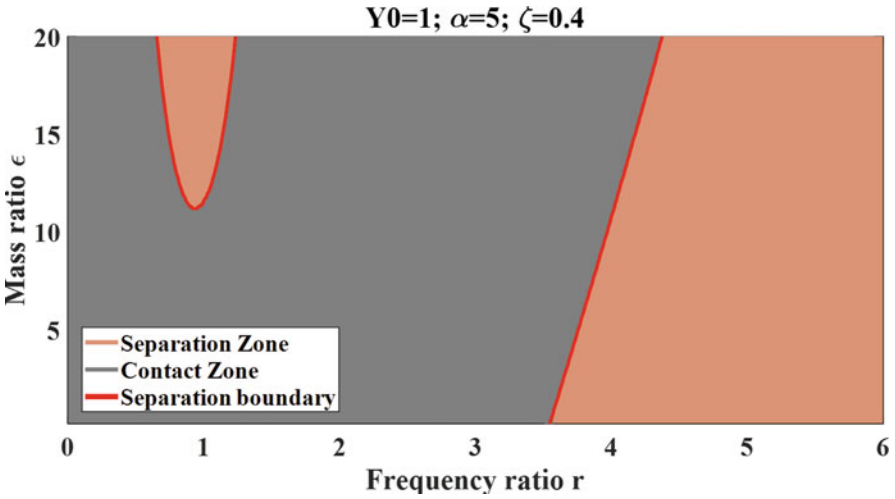


Fig. 9.21 Separation boundary regarding mass ratio with $\xi = 0.4$

separation boundary involves two branches. The tire-road separation appears with high frequency ratios $r \geq 3.54$ or very low frequency ratios closed to the first natural frequency ratio.

Let us keep a constant stiffness ratio $\alpha = 5$; the effect of mass ratio ε on the no-separation and separation zones has been determined for different values of the damping ratio as presented in Fig. 9.22. Theoretically, the mass ratio can equal to 0 if the sprung mass is zero, and it can happen when the model is a wheel. For low damping ratios $\xi = 0.2$ and $\xi = 0.4$, every point is above the parabolic curve, or

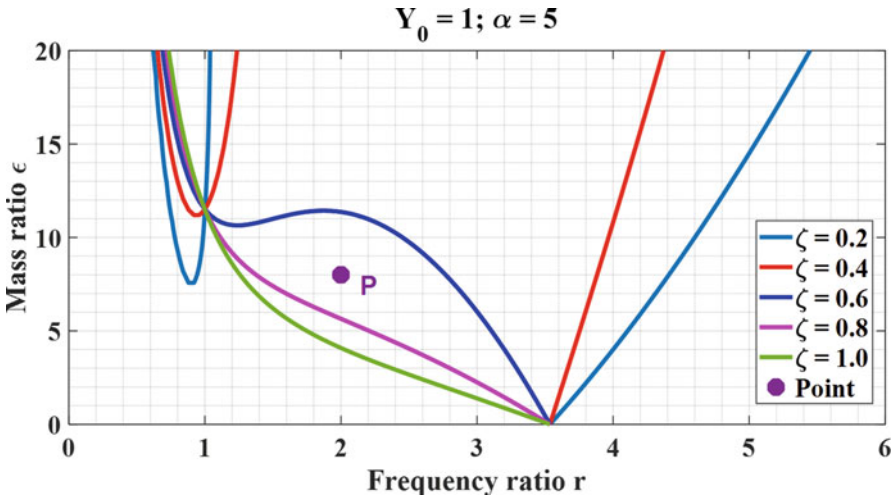


Fig. 9.22 Separation boundary regarding mass ratio with different damping ratios

points are right of the second curve at which there would be tire-road separation. For other damping ratios, the separation zone is identified above the corresponding separation curve. Thus, decreasing the value of the damping ratio ξ can result in a reduced separation area because the point P ($r = 2, \varepsilon = 8$) is just in a no-contact zone with $\xi = 0.8$ and $\xi = 1.0$.

To investigate the effect of stiffness ratio on the separation zone, let us keep the same parametric input in Fig. 9.22, reducing the stiffness ratio, say $\alpha = 3$; the separation boundary associated with $\xi = 0.6$ in Fig. 9.23 moves under the point P ($r = 2, \varepsilon = 8$). Therefore, the separation zone is increased significantly at low stiffness ratios, meaning that the model with a higher α has a wider in-contact area.

For a given damping ratio in Fig. 9.23, say $\xi = 0.4$, the separation zone is above the parabolic branch and right the second path as explained in Fig. 9.24. For a given mass ratio, say $\varepsilon = 6$, the horizontal line $\varepsilon = 6$ extends to an associated curve $\xi = 0.4$ at $r_1 = 0.68$, $r_2 = 1.27$, and $r_3 = 2.44$ indicating the limiting frequency ratio with different zones. Separation areas are $0.68 \leq r \leq 1.27$ or $r \geq 2.44$ and contact areas with $r < 0.68$ or $1.27 < r < 2.44$.

Let us keep a constant damping ratio $\xi = 0.4$, changing the value of stiffness ratio α in a practical range to investigate the effect of mass ratio on the limit of separation frequency ratio as illustrated in Fig. 9.25. It can be noticed that the curves$\alpha = 3$ and $\alpha = 5$ were described in Figs. 9.21 and 9.24, respectively. The separation zone is inside two separation curves for $\alpha = 1, \alpha = 2$, while it is above the parabolic curve or right of the exponential curve for other α. Thus, the contact area will be wider at higher stiffness ratios.

The effect of the damping ratio on the separation boundary is investigated in Fig. 9.26 by keeping a constant mass ratio $\varepsilon = 5$. For low stiffness ratios $\alpha = 1 - 3$, input parameters of the system are left of the first line or under the parabolic curve

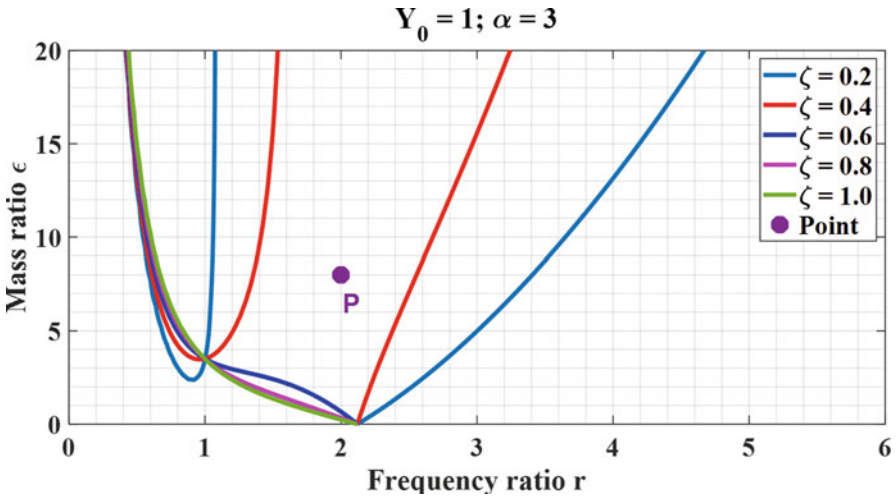


Fig. 9.23 Separation boundary regarding mass ratio with a low stiffness ratio

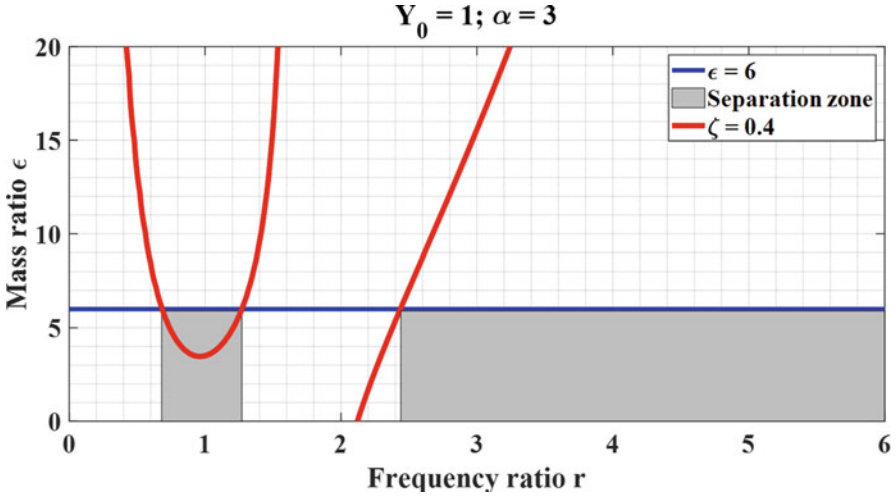


Fig. 9.24 Separation boundary regarding mass ratio with $\alpha = 3$, $\xi = 0.4$

that keeps its tire always in contact to the road with any damping ratio as shown in Fig. 9.27. However, the separation zone is under the parabolic curve or right of the second curve for other stiffness ratios. Therefore, the separation square will significantly reduce when the stiffness ratio has risen.

For a nominal value of $\varepsilon = 5$ and a given stiffness ratio, say $\alpha = 3$, the effect of ξ on the limit of separation frequency ratio is presented in Fig. 9.27. For a given damping ratio, say $\xi = 0.4$, the horizontal line $\xi = 0.4$ hits curve $\varepsilon = 5$ at $r_1 = 0.73$, $r_2 = 1.21$, and $r_2 = 2.41$ indicating the limit of the separation.

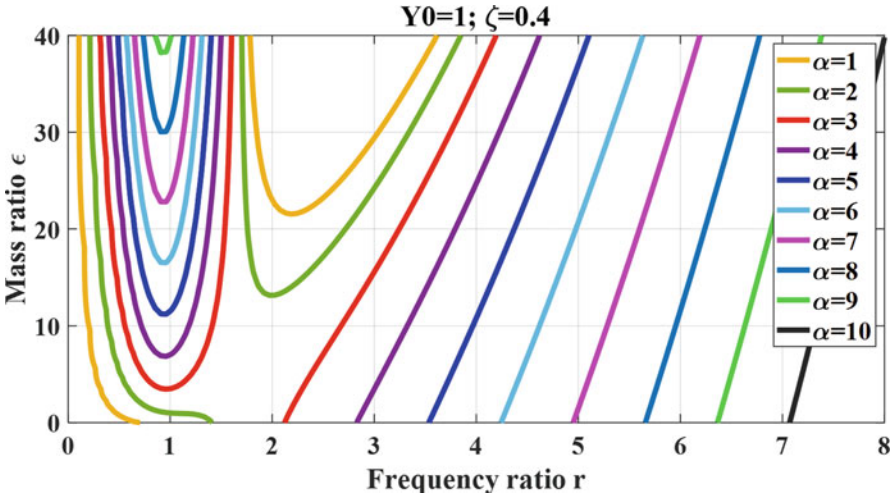


Fig. 9.25 Separation boundary regarding mass ratio with different stiffness ratios

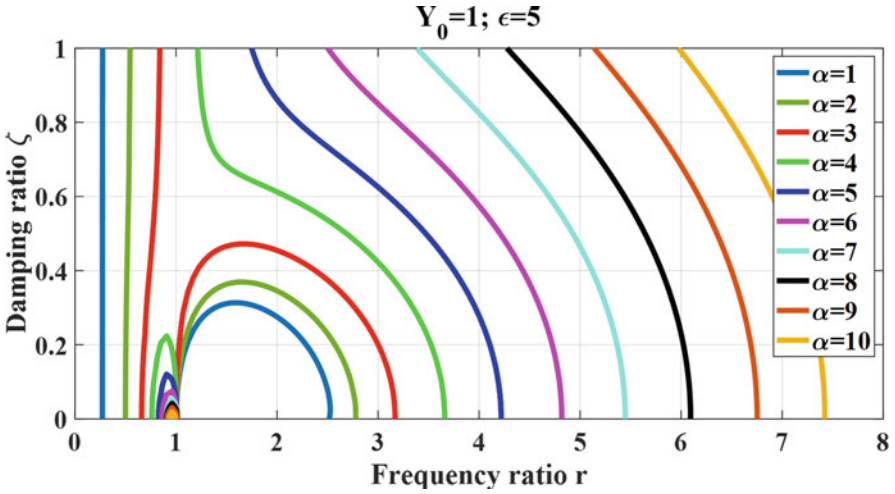


Fig. 9.26 Separation boundary regarding damping ratio with different stiffness ratios

The effect of mass ratio ε on the separation boundary is presented in Fig. 9.28 by keeping a constant stiffness ratio $\alpha = 5$. The separation zone is similar to Fig. 9.26 indicating that the separation zone declines when the stiffness ratio grows up, meaning that the wheel has better contact with the ground at high stiffness ratios. At high mass ratios, for example $\varepsilon = 14$, no-contact regions are left of the first curve, where the slope of the curve is nearly vertical or under closed parabolic curve.

Figure 9.29 reveals the frequency ratio limit for the happening of separation in regard to a high mass ratio $\varepsilon = 10$ extracted from Fig. 9.28. A horizontal line $\xi = 0.2$

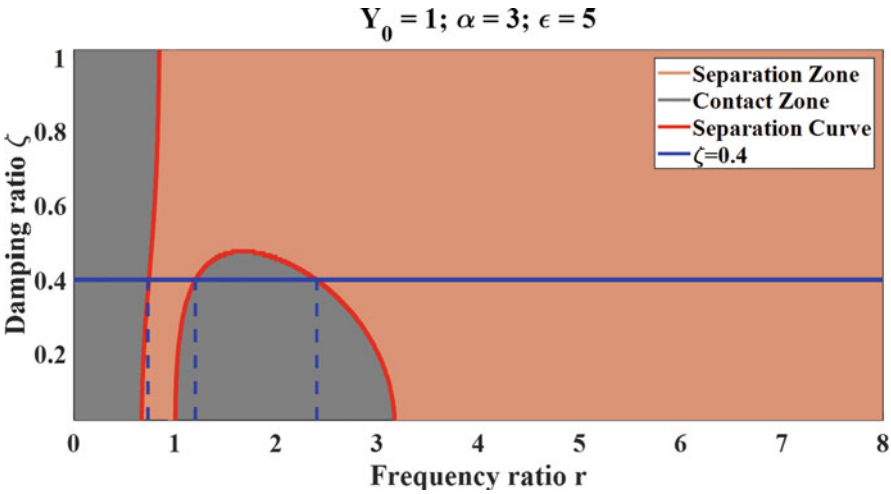


Fig. 9.27 Separation boundary regarding damping ratio for a low stiffness ratio

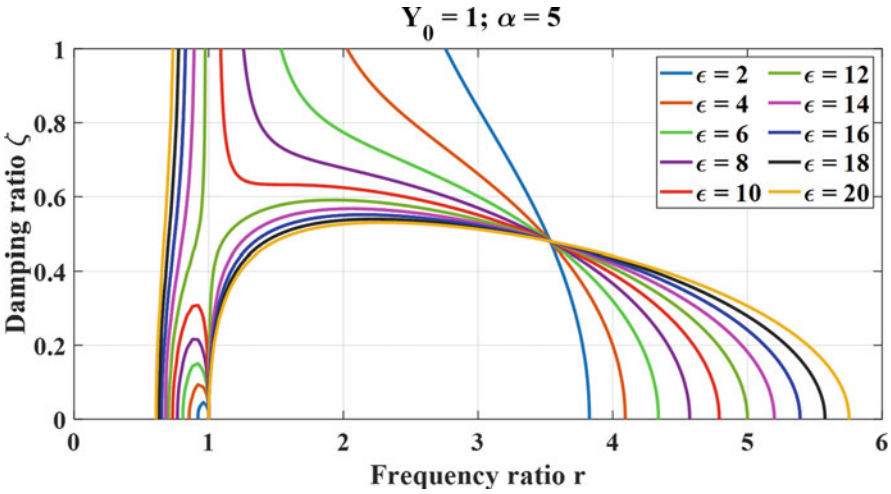


Fig. 9.28 Separation boundary regarding damping ratio for different mass ratio

begins from $r = 0$ andextends to curves $\varepsilon = 10$ at $r_1 = 0.78$, $r_2 = 0.98$, $r_3 = 4.59$ presenting the frequency ratio limit with two separation zones. In this case, the model with higher damping ratios can dismiss the separation at low frequency ratio.

Separation frequency boundaries could be regarded as the limitation of accuracy of linear systems; beyond these borders, the linear system might not yield an accurate approximation of dynamic responses. Thus, these boundary graphs are very useful to determine accurately contact and in-contact areas, the limits of separation, as well as design desirable suspension parameters to avoid the contactless process.

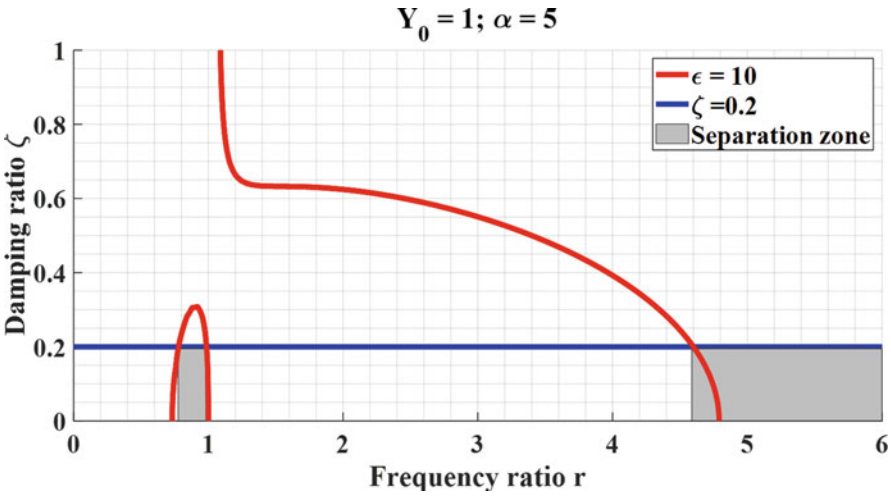


Fig. 9.29 Separation boundary regarding damping ratio for $\varepsilon = 10$

9.6 Separation Duration

In order to analyze deeper the effect of parametric dynamics on a no-contact zone, the tire-road separation duration will be introduced. The separation duration curves provide understandable system behavior to allow designers to select the best values of suspension parameters for postponing or suppression of the tire-road separation phenomenon. Duration of separation D_s is a fraction of the oscillation interval in which the tire is separated from the surface:

$$D_s = \frac{Time\ in\ no-contact\ state}{Investigation\ interval} x\ 100 \tag{9.38}$$

Comparison of the relative displacement to the separation condition in the steady-state response determines the separation duration. Thus, the interval of investigation time is sufficiently long to diminish the transient response within the first half of the total time interval.

A range of the frequency ratio r from 6 to 10 changes the road excitation Y to impact the separation duration D_s. The separation duration is calculated by using the indicator variable I. The separation indicator describes which state that the system is in at each evaluation point. The no-contact and contact states are presented by switching from 1 to 0. The indicator line can compute the fraction of the oscillation duration in which the separation phenomenon occurs by summation of the total of these values over the investigation interval. Figure 9.30a, b, c demonstrate the time response with 27.02, 31.18, and 26.18 percent of separation time, respectively. These separation durations are plotted by solid circles A, B, and C indicating the discrete sequence data as shown in Fig. 9.30d.

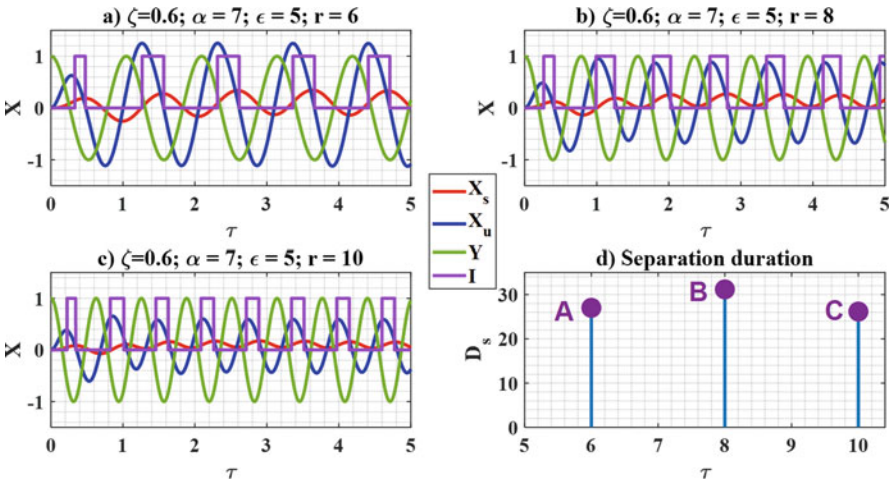


Fig. 9.30 Time response with separation duration

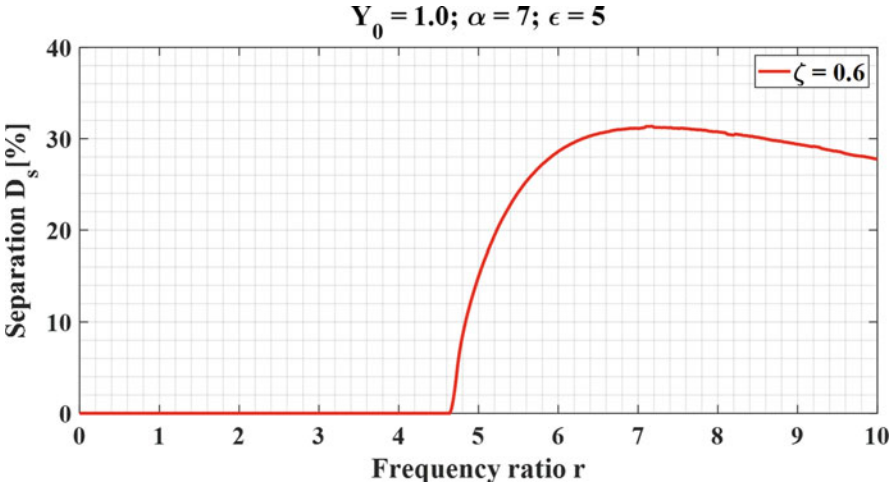


Fig. 9.31 Separation duration for a range of frequency ratio

Firstly, the separation duration is examined following a range of frequency ratio r. Figure 9.31 demonstrates the separation duration for the range of r. The calculation of separation time is only conducted in separated systems, meaning that the loss of contact will occur after a separation point. The tire keeps contacting with the surface until it separates at the limit of frequency ratio $r = 4.63$. The separation time increases rapidly after this separation point and reaches a peak roundly 31.5% at $r = 7$; it has a gradual decrease at higher frequency ratios.

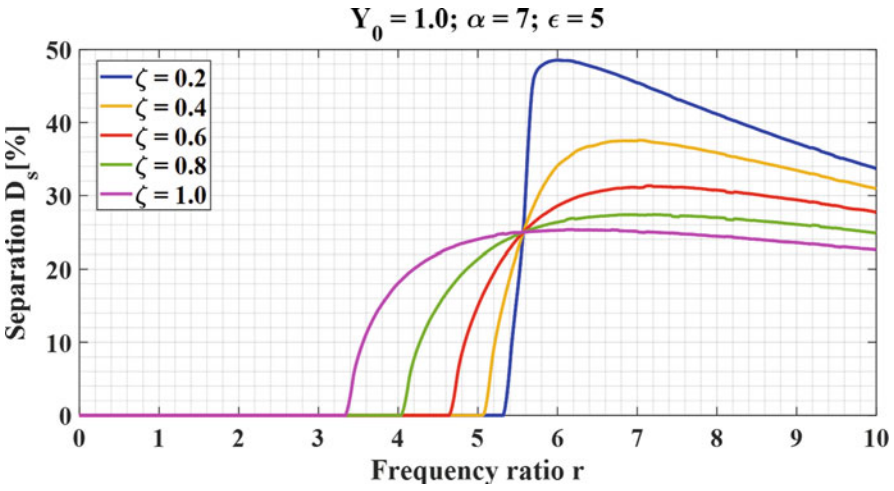


Fig. 9.32 Separation duration for a range of frequency ratio with different damping ratios

For a set of input parameters, changing the value of the damping ratio provides more curves presenting the separation time as shown in Fig. 9.32. The separation happens later at low values of damping ratio. For example, for $\xi = 1.0$ the separation point is at $r = 3.34$, while the system continuous contacts at frequencies beyond this point then separate at $r = 5.48$ for $\xi = 0.2$. Initially, the percentage of separation time goes rapidly up following the value of frequency ratio, and the separation time curves will pass through one node at $r = 4.58$ and have an opposite trend after this intersection. At the node, the system has separated 24.51% of moving time at any damping ratio. Especially, the percentage of separation time for a low damping ratio $\xi = 0.2$ climbs exponentially upward nearly 50% at $r = 6$.

Let us keep the same set of input parameters using Fig. 9.32; increasing the value of the mass ratio, $\varepsilon = 9$, provides a lower separation duration D_s as inferred from Fig. 9.33. There is a similar shape of the separation duration to Fig. 9.32. Although D_s is under 22% at high damping ratios from $\xi = 0.6$ to $\xi = 1.0$, the separation occurs earlier, for instance, the separation point for $\xi = 1.0$ is $r = 2.19$.

To evaluate the stiffness ratio α on the separation duration, a comparison of $\alpha = 3$ to Fig. 9.32 with $\alpha = 7$ will be carried out. Figure 9.34 denotes the fraction of separation time in response to a low stiffness ratio $\alpha = 3$. The system has two nodes at $r_1 = 0.81$ and $r_2 = 3.43$. Between the two nodes in Figure 9.34, the no-contact time is proportional to the value of damping, and it might equal zero at low damping ratios, $\xi = 0.2$ and $\xi = 0.4$. The percentage of no-contact time for almost damping ratio has so rapidly risen to approximately 46% at low excitation frequency ratios closed to the first natural frequency ratio. However, the separation time can decrease at higher frequency ratios.

Secondly, the separation duration following a range of road amplitude Y_0 is investigated. For given parameters, Fig. 9.35 illustrates that the percentage of

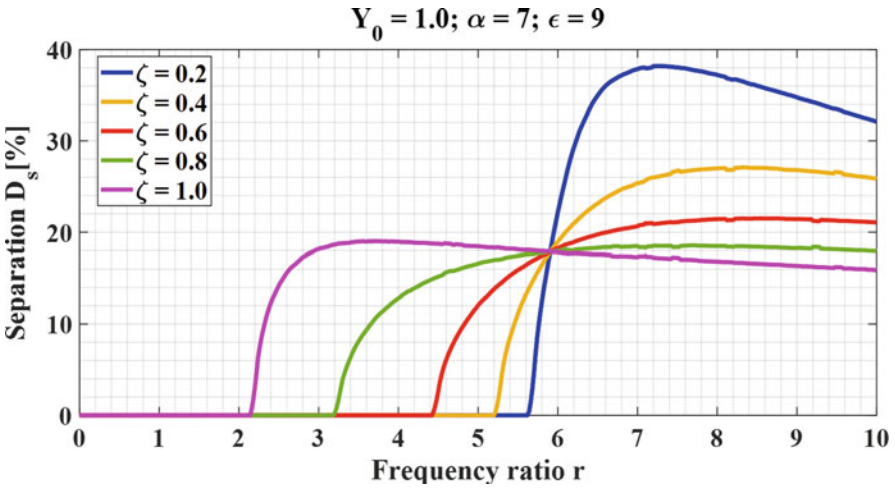


Fig. 9.33 Separation duration for a range of frequency ratio with a high mass ratio

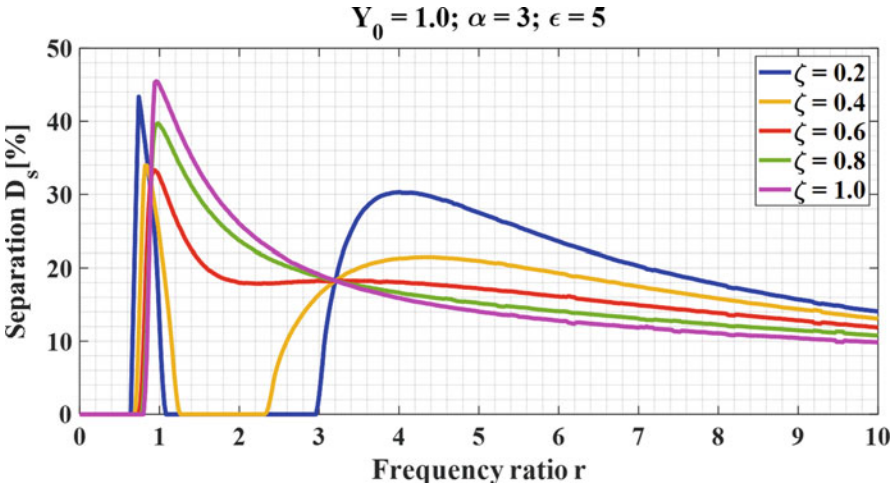


Fig. 9.34 Separation duration for a range of frequency ratio with a low stiffness ratio

separation time will rise following the road amplitude. The total separation time is not proportional to values of damping ratio; the separation time curve with a low damping ratio is higher than others. Furthermore, the curves have a slight distinction with high damping ratios. That means the effect of road amplitude is slightly at high damping ratios.

Thirdly, we analyze the separation duration in response to the designed stiffness ratio with the shifting value of the damping coefficient. Figure 9.36 illustrates the effect of stiffness ratio on separation duration for a given frequency ratio $r = 3$. The separation time is not proportional to the value of damping ratio before passing

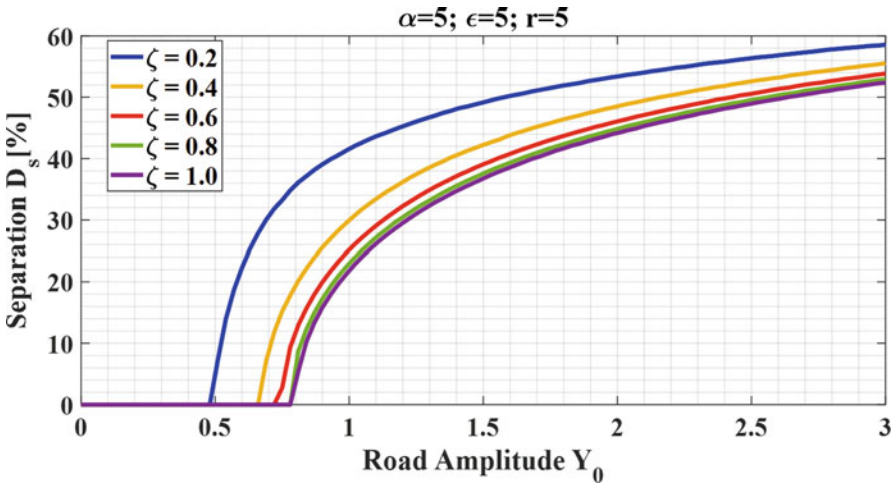


Fig. 9.35 Separation duration depends on road amplitude

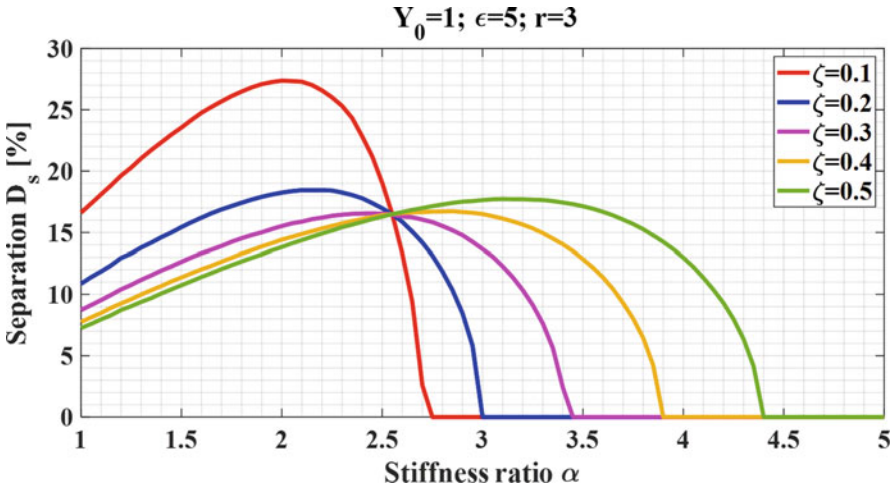


Fig. 9.36 Separation time depends on stiffness ratio

a node at $\alpha = 2.55$. The curves of separation time have a similar tendency with a damping ratio greater than 0.4. To reduce the separation time, the system should have a high stiffness ratio because its tire may recontact with the ground for any damping ratio when the stiffness ratio is greater than 4.4.

In addition, we investigate the fraction of separation time depending on the designed mass ratio. Figure 9.37 shows the effect of mass ratio on the fraction of no-contact time at a low stiffness ratio $\alpha = 3$. The data points to signs of a slowdown in separation duration; it seems to depict a linear relationship with the mass ratio. The system has more in-contact time at high mass ratios. Let us give an example; for

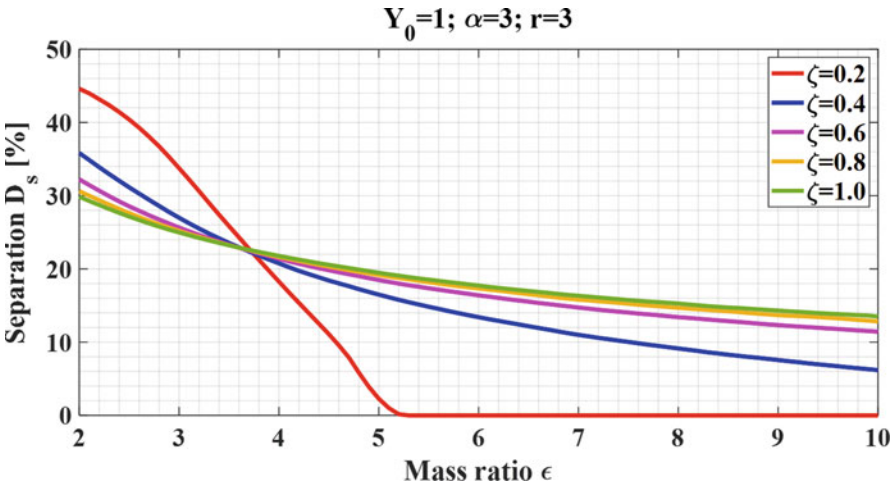


Fig. 9.37 Separation time depends on mass ratio

given inputs $r = 3$ and $\varepsilon = 6$, the separation time is just under 20% for any damping ratio, and the tire contact with the road for a low damping ratio $\xi = 0.2$.

Additionally, we observe the separation time in reply to a range of the suspension damping ratio. Figure 9.38 provides an indication of separation time considering the damping ratio at a high mass ratio $\varepsilon = 9$. The separation time has a proportional growth rate regarding the value of damping and stiffness ratios. Instead of rising in a nonlinear fashion at the beginning at low damping ratios, they are increasing much more quickly with $\alpha = 5$ and $\alpha = 6$and come to over 20% at $\xi = 1.0$. Although at a low stiffness ratio $\alpha = 1$ the system starts early losing its road contact, the total of no-contact time is always under 5%. However, at a high stiffness ratio $\alpha = 6$, the tire-road separation occurs later at a high damping ratio $\xi = 0.68$.

The effect of any two input parameters on the duration of tire separation could be examined by a surface in the three-dimensional coordinate. Figure 9.39 demonstrates the separation duration depending on the stiffness and mass ratios for $\xi = 0.4$, $r = 5$. At high stiffness ratios, the contact state does not happen with any mass ratio, while no-contact time has increased starting from a mass ratio of under 10% at $\alpha = 1$ until reaching a peak of nearly 45% at $\alpha = 6$, $\varepsilon = 2$. Following the chart, the suspension system with high mass and low stiffness ratios has a better contact period than other ones.

The main goal in the optimization of vehicle suspension systems is a reduction of the vertical acceleration absolute because it presents the transmitted force to the passengers. Therefore, the investigation of separation time regarding the acceleration of sprung mass provides an insight understanding of the real system. The nondimensional acceleration of the sprung mass $\ddot{X}_s$ is determined directly from the equations of motion. Figure 9.40 denotes a relationship between the separation time with the acceleration of sprung mass A_s in response to the frequency ratio $r = 3$.

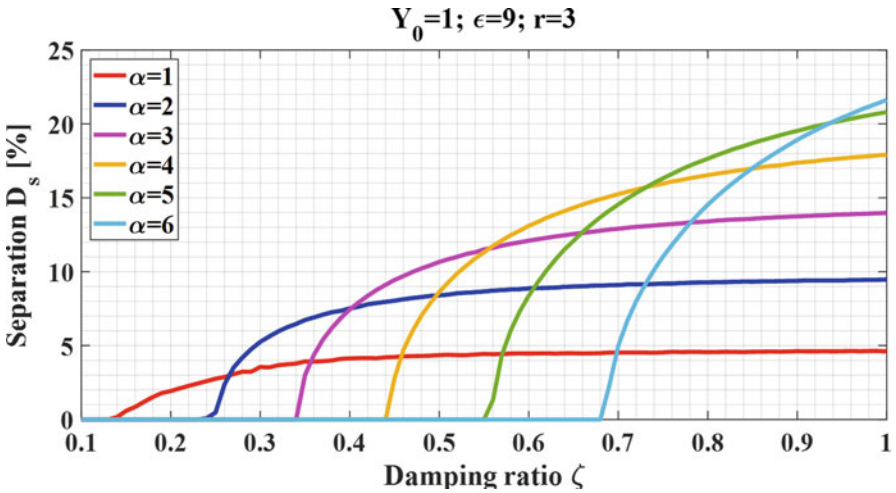


Fig. 9.38 Separation time depends on damping ratio

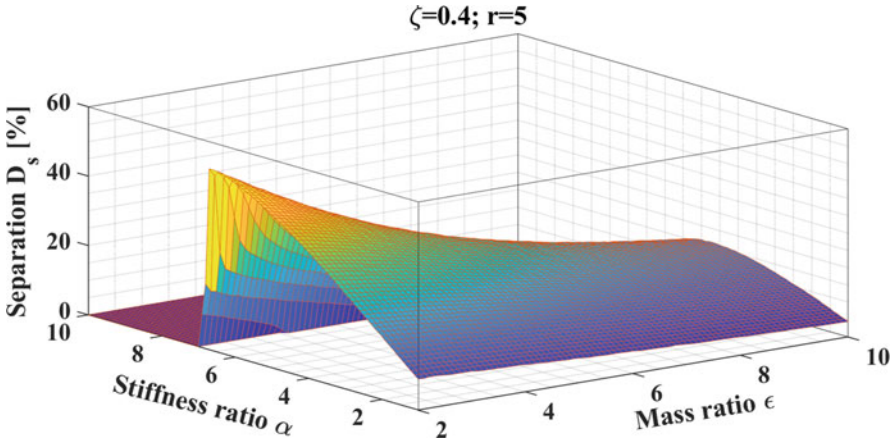


Fig. 9.39 Separation time in three-dimensional space

To minimize both separation time and acceleration, the designer should choose a low damping ratio $\xi = 0.1$. In this case, the tire is always contacted with the road and the acceleration of the sprung mass is minimum at 1.82.

For a high input frequency ratio $r = 5$, the sprung mass acceleration achieved the minimum nearly 2.6 at $\xi = 0.13$ as shown in Fig. 9.41, but the separation time is too high at this point with 47.51%. To gain a minimum value of the separation time, the suspension should have a high damping ratio to $\xi = 1.0$, although the acceleration of sprung mass A_s goes up to over 4.41. Therefore, in this set of input parameters, the system needs to balance the ride comfort and tire-road separation criterion to determine the optimal damping ratio.

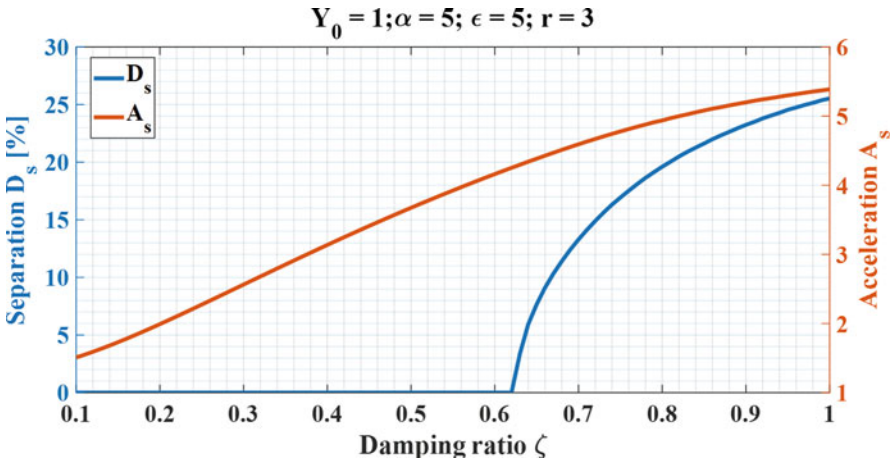


Fig. 9.40 Separation time and body acceleration for a range of damping ratio with $r = 3$

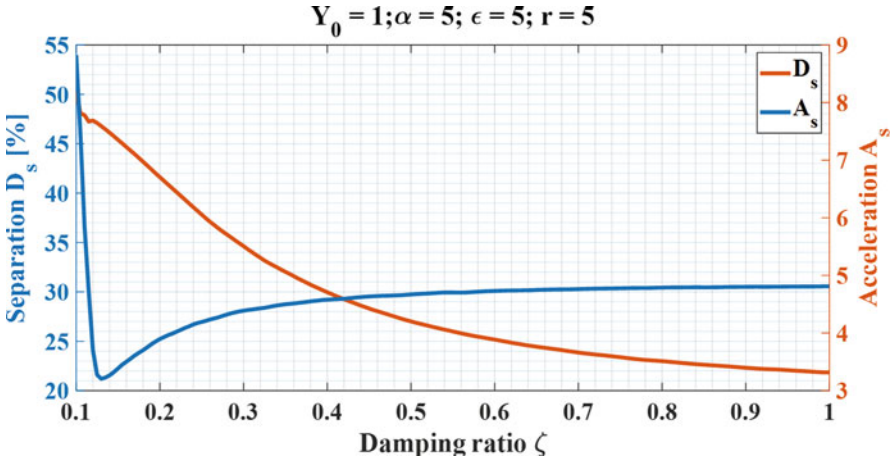


Fig. 9.41 Separation time and body acceleration for a range of damping ratio with $r = 5$

To identify the optimal input parameters for reducing both separation time D_s and body acceleration A_s, the tire-road separation time with respect to body acceleration will be examined in a wider range of data. Figure 9.42 expresses the tire-road separation time with respect to body acceleration graphically for a set of input parameters. The horizontal axis is the value of the body acceleration, and the vertical axis is the value of separation time. There are two sets of curves that make a mesh. The first set, which is almost parallel as shown by dotted curves, has a constant stiffness ratio α. The second set has a constant damping ratio ξ as shown by solid

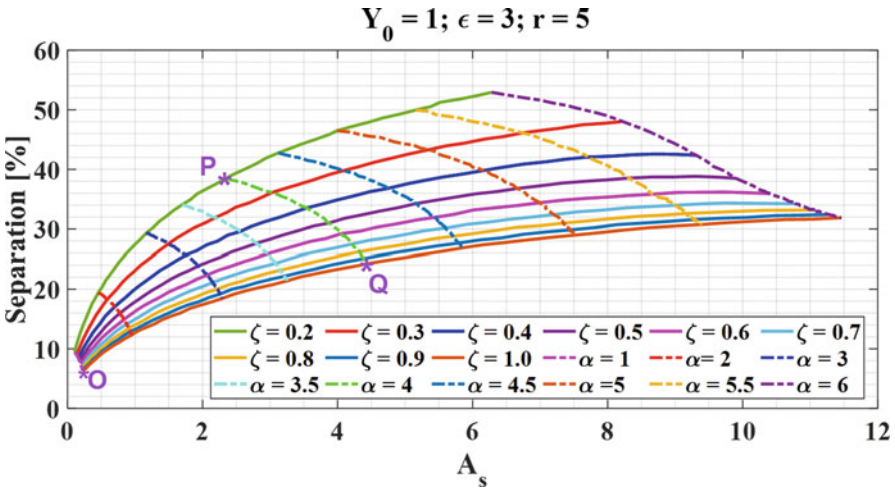


Fig. 9.42 Separation time respects body acceleration

curves. For $\alpha = 4$, the no-contact time D_s is roughly 38.11%, and the acceleration A_s obtains nearly 2.4 at $\xi = 0.2$ presented by intersection P, while D_s is just 24.31% and A_s achieves over 4.2 at $\xi = 1.0$ indicated by intersection Q. In this nominal set of input data, the separation time and the body acceleration have both decreased at low stiffness and high damping ratios. Thus, the optimal value of α, ξ to obtain a minimum of both separation time and body acceleration is at point O.

Figure 9.43 compares the no-contact time with respect to body acceleration A_s by changing the value of the frequency ratio. For $\alpha = 7$ in Fig. 9.43a with $r = 9$, D_s is 35.01%, and A_s achieves roughly 4.5 for $\xi = 0.2$ at point *P,* while they are just 15.12% and approximately 5.1, respectively, for $\xi = 1.0$ at point *Q*. Reducing the frequency ratio could result in a significant decline separation time as illustrated in Figs. 9.43c, d with $r = 5$, $r = 3$ because point P($\xi = 0.2, \alpha = 7$) gives the separation duration of 0%, and the point Q($\xi = 1.0, \alpha = 7$) has the separation duration of 18%, whereas the body acceleration does show a slight increase; the wheel can always be in contact with the road when the stiffness ratio is greater than 7, as depicted in Fig. 9.43d.

These curves presented the tire-road separation time to describe vital ride behaviors that could be completely absent in linear systems. Furthermore, the most important benefit of the investigation of separation time is a basis for reducing the loss of tire-road contact on future vehicle designs.

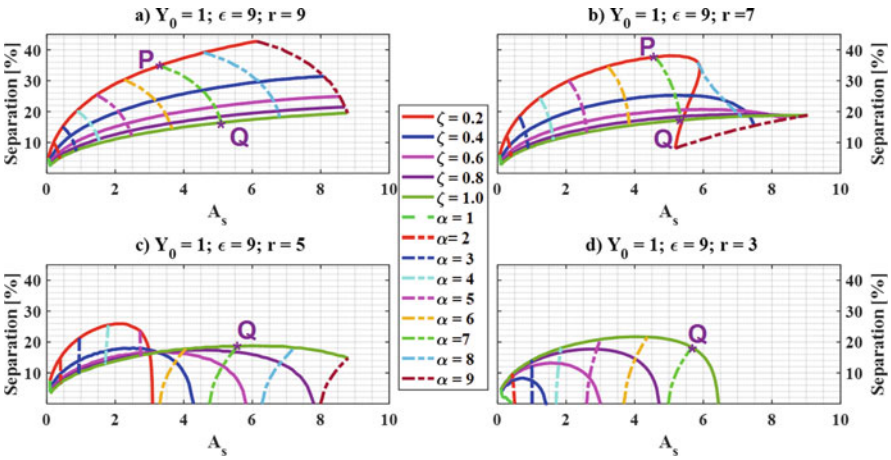


Fig. 9.43 Separation time respects body acceleration with $\varepsilon = 9$

9.7 Conclusion

This chapter presents the vehicle vibration study considering the tire-road separation, the governing differential equations have been transformed into a nondimensional field, and the novel method allows an examination of the separation phenomenon in a wider range of input parameters including damping, mass, stiffness, frequency ratios, and road amplitude. The time response and frequency domain have been demonstrated graphically by using the discontinuous equation system to present the dynamic behaviors of the nonlinear system. The limiting values of acceptable parametric suspension and road excitation for accuracy of linear models are identified by the separation condition based on the relative displacement. Collecting these separation points provides the separation boundary that separated the in-contact and no-contact zones. Sensitivity analysis of separation duration is carried out to allow designers to choose the optimal parameters for reduction or suppression of the separation phenomenon as well as undesirable vibration to broaden new avenues for future research.

Notation

Description	Dimensionless symbol	Dimensional symbol [unit]
Sprung mass		m_s[kg]
Unsprung mass		m_u[kg]
Suspension spring stiffness		k_s[N/m]
Tire stiffness		k_u[N/m]
Main suspension damper		c_s[Ns/m]
Sprung mass natural frequency		ω_u[rad/s]
Unsprung mass natural frequency		ω_s[rad/s]
Excitation frequency		ω [rad/s]
Second natural frequency		ω_{n2}[rad/s]
Stiffness ratio	α	
Mass ratio	ε	
Damping ratio	ξ	
Excitation frequency ratio	r	
Natural frequency ratio	r_{n2}	
Amplitude of road	Y_0	y_0[m]
Road excitation	Y	y [m]
Time	τ	y [s]
Initial time	τ_0	
Displacement of sprung mass	X_s	x_s[m]
Displacement of unsprung mass	X_u	x_u[m]
Relative displacement	X_t	x_t[m]
Static tire deflection	X_{ST}	x_{ST}[m]
Amplitudes	$\|X_s\|$, $\|X_u\|$, $\|X_t\|$	$\|x_s\|$, $\|x_u\|$,$\|x_t\|$ [m]
Phases of displacements	φ_s, φ_u, φ_t	
The body acceleration	A_s	
Unsprung acceleration	A_u	
Factor to nondimensional transfer		x_0[m]
The acceleration of gravity		g [m/s^2]
Number of static tire deflection	n	
Coefficients of harmonic solutions	A_1, A_2, A_3, B_1, B_2, B_3	
Symbols of harmonic solutions	Z_1, Z_2, Z_3, Z_4	
Separation duration	D_s	

References

1. Barbosa, RS 2011, 'Vehicle dynamic safety in measured rough pavement', *Journal of Transportation Engineering*, vol. 137, no. 5, pp. 305–310.
2. Captain, K, Boghani, A & Wormley, D 1979, 'Analytical tire models for dynamic vehicle simulation', *Vehicle System Dynamics*, vol. 8, no. 1, pp. 1–32.

3. Dobaj, K 2015, 'Influence of car wheel suspension parameters on improvement of active safety and ride comfort', *Czasopismo Techniczne*, vol. 2015, no. Mechanika Zeszyt 3-M (21) 2015, pp. 13–22.
4. Gillespie, T.D., 1992. *Fundamentals of vehicle dynamics* (Vol. 400). Warrendale, PA: Society of automotive engineers
5. Guiggiani, M., 2014. The science of vehicle dynamics. *Pisa, Italy: Springer Netherlands*, p.15.
6. Jazar, RN 2013, *Advanced vibrations: A modern approach*, Springer Science & Business Media.
7. Jazar, R.N., 2017. Vehicle vibrations. In *Vehicle Dynamics* (pp. 819–881). Springer.
8. Jazar, R.N., 2020. Dynamic Dimensional Analysis. In *Approximation Methods in Science and Engineering* (pp. 87–189). Springer, New York, NY.
9. Jiao, L 2013, *Vehicle model for tire-ground contact force evaluation*.
10. Khazaie, A., Hussaini, N., Marzbani, H. and Jazar, R.N., 2018. Quarter Car Suspension Model with Provision for Loss of Contact with the Road. In *Nonlinear Approaches in Engineering Applications* (pp. 167–208). Springer.
11. Lindgren, G 1981, 'Jumps and bumps on random roads', *Journal of Sound and Vibration*, vol. 78, no. 3, pp. 383–395.
12. Lukoševičiene, O, Bogdevičius, M, Nagurnas, S & Pečeliūnas, R 2004, 'Comparative research of a motor car motion in the case of the loss of contact with the road surface', *Transport*, vol. 19, no. 1, pp. 20–23.
13. Lukoševičiene, O & Sokolovskij, E 2006, 'Movement of the vehicle being braked when some wheels are incapable of braking or have lost touch with the road surface', *Transport*, vol. 21, no. 1, pp. 8–11.
14. Lukoshevichiene, O & Sokolovskij, E 2001, 'Movement of a vehicle after loosing the touch of the wheels with road surface', *Transport Engineering*, vol., pp. 29–34.
15. Milliken, W.F., Milliken, D.L. and Olley, M., 2002. *Chassis design: principles and analysis* (Vol. 400). Warrendale, PA: Society of Automotive Engineers.
16. Moreno-Ramírez, C. and Tomas-Rodríguez, M., 2014, July. Non linear optimization of a sport motorcycle's suspension interconnection system. In *2014 UKACC International Conference on Control (CONTROL)* (pp. 319–324). IEEE
17. Reza-Kashyzadeh, K., Ostad-Ahmad-Ghorabi, M.J. and Arghavan, A., 2014. Investigating the effect of road roughness on automotive component. *Engineering Failure Analysis*, *41*, pp.96–107
18. Svahn, F, Jerrelind, J & Dankowicz, H 2009, 'Suppression of bumpstop instabilities in a quarter-car model', in *Non-smooth Problems in Vehicle Systems Dynamics*, Springer, pp. 137–147.
19. Wageman, DA 2015, Development of a Non-Linear Multi-body Dynamics Model for Active Suspension Control Design and Evaluation of a Formula Student Racecar, thesis.
20. Wong, JY 2008, *Theory of ground vehicles*, John Wiley & Sons.

Chapter 10
Nonlinear Model Predictive Control Real-Time Optimizers for Adaptive Cruise Control: A Comparative Study

Sadegh Tajeddin, Yasaman Masoudi, and Nasser L. Azad

Abstract Nonlinear Model Predictive Controllers are among the most popular techniques for developing intelligent control systems in automotive applications, including Adaptive Cruise Control Systems. Nonlinear Model Predictive Controllers, however, can be computationally expensive and many studies in the last decade have been dedicated to decrease their computational time. In this regard, by applying the so-called indirect optimization approach of Generalized Minimal Residual method to the Adaptive Cruise Control problem for Toyota Plug-in Hybrid Prius, this chapter contributes to the literature in three ways; first, it provides a comparative study of indirect real-time optimization methods in the context of Nonlinear Model Predictive Adaptive Cruise Control. These methods are implemented by the MPsee tool, which facilitates automatic code generation, implementation, and calibration of Nonlinear Model Predictive Controllers for MATLAB and Simulink users. Second, real-time implementability of proposed real-time optimization approaches is investigated by conducting Hardware-in-the-Loop experiments. Finally, this study argues that utilizing a Perturbed Chord modified Newton method instead of a Standard Newton method improves the computational speed up to 50%.

Keywords Nonlinear model predictive control · Real-time optimization · Indirect optimization methods · Generalized minimal residual method · Automatic code generation · Perturbed chord method

10.1 Introduction

Rapid advancement of affordable and powerful computational devices has made automakers more interested in implementation of smart control systems rather

S. Tajeddin · Y. Masoudi (✉) · N. L. Azad
Systems Design Engineering Department, University of Waterloo, Waterloo, ON, Canada
https://uwaterloo.ca/smart-hybrid-electric-vehicles-systems-lab
e-mail: stajeddin@uwaterloo.ca; ymasoudi@uwaterloo.ca; nlashgarianazad@uwaterloo.ca

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_10

than costly modification of mechanical hardware [1]. These state-of-the-art control systems use the information distributed by Intelligent Transportation Systems (ITSs) to strategically improve driving experience, fuel consumption, and safety [2, 3]. Adaptive Cruise Control (ACC) systems are examples of recent achievements in this area [4–8], where information received from different sensory (e.g., radar) and communication (e.g., V2X) devices is integrated with modern control techniques to optimize the speed profile.

Real-Time Optimization (RTO) methods including Model Predictive Controllers (MPCs) have several applications in the automotive control area [9]. MPCs are proved to be highly effective in utilizing various information in the process of decision making [10–12], and given their advantages, MPCs have been identified as excellent methods for ACC systems [11, 13, 14]. Nonlinear Model Predictive Control (NMPC) [15] is a vastly studied type of MPC, where embedding a nonlinear mathematical model of the system has made great levels of prediction accuracy achievable. NMPC could be computationally expensive compared to its rivals, which has led to a great amount of research to improve its speed [16–19]. In this context, NMPC is formulated as an optimal control problem (OCP), and dealt with different approaches often categorized as dynamic programming (DP) and direct and indirect methods (DMs and IMs) [20, 21]. DP provides the global optimal solution of the optimization problem, yet is not suitable for real-time implementations. DMs include turning problem into a nonlinear programming problem (NLP) using the so-called discretization methods, while IMs use the necessary optimality conditions to construct a boundary value problem (BVP), which deals with numerical solution of differential equations [22, 23].

Literature in the automotive control area has suggested different applications for each of these techniques. While DMs are the common approaches in engineering problems [24] including obstacle avoidance [25], cruise control [26, 27], and fuel consumption minimization [28, 29], IMs provide a higher accuracy which has made them suitable for certain applications [20]. Moreover, the reported hassles of IMs, such as computational effort, instability, and initialization issues could be studied and medicated further [20, 24]. IMs could be easily implemented for online solutions and several examples of them for real-time control can be found; by applying different optimization methods to ACC problem, Saerens et al. [21] demonstrated the potential of IMs in particular cases. Bu et al. [30] applied an indirect MPC for ACC, with online parameter identification to incorporate parameter uncertainties. Petit et al. [24] suggested an indirect approach to solve the energy management problem of an electric vehicle (EV). Through this methodology and using the structure of the system dynamics, they improved the computational expense as well. Moreover, Dib et al. [31] studied eco-driving for EVs using an indirect online implementable method which was benchmarked against DP. Kock et al. [32] implemented a predictive speed controller, an extended version of the conventional cruise controllers, using indirect optimization and obtained a semi-analytic solution. The results were benchmarked against a PI cruise controller. Sciarretta et al. [33] suggested an indirect analytical solution for Eco-driving (ED) in different platforms, where ED is formulated as an OCP with the main objective

of energy minimization with state, control, and terminal constraints. Another real-time example of indirect approaches is the integration of Generalized Minimal Residual method (C/GMRES), showcased by a robotic application [34]. Further, Di Cairano et al. [9] found C/GMRES effective for handling nonlinear OCPs with equality and a few inequality constraints. Perturbed Chord method in [18] has demonstrated promising improvements in NMPC real-time implementation as well. Utilizing such indirect RTO methods, scholars have been able to develop NMPCs for automotive applications[35, 36]. In line with the above-mentioned studies and given the potential of IMs, the authors apply an indirect GMRES-based method in this study.

The proposed RTO methods preprocess OCP and transform it into a set of iterative online computations. From the implementation standpoint, synthesis, and calibration of NMPCs could be facilitated by developing automatic code generators. AutoGenU [37] in Maple based on single-shooting and C/GMRES, ACADO Toolkit [38] based on direct multiple-shooting and Real-Time Iteration (RTI), and SCDE, a design environment in Maple based on Standard Newton iterations [19] are examples of previous advances in this area. Moreover, an extension to the ACADO toolkit for long prediction horizons is presented in [39]. Here the authors use *MPsee* [40], their developed tool for rapid generation of NMPCs with different RTO methods. Similar to AutoGenU, MPsee is a mathematical program that generates NMPC codes based on GMRES solver. However, this tool works in MATLAB environment, where user can specify the RTO technique (e.g., multiple-shooting or single-shooting; Newton/GMRES or C/GMRES) and generate relative Simulink blocks for the controller. The auto-generated codes are compatible with MATLAB coder, and MATLAB can easily generate the C codes associated with the designed controller.

Despite many studies on development of novel RTO methods [34, 41, 42], few have compared such methods in terms of computational speed and performance in automotive control applications [17], specifically at a Hardware-in-the-loop (HIL) scale. The main contribution of this work is conducting such a comparative study in the context of ACC systems. To do so, the authors extend their previous research on developing ACC for the baseline vehicle, Toyota Plug-in Hybrid Prius, where an earlier version of MPsee was applied [43]. The problem which was previously formulated via single-shooting is now formulated in a multiple-shooting fashion as well, and the associated results are compared. Additionally, different combinations of the controller settings are implemented to examine their real-time implementability through HIL experiments. Moreover, the results are not limited to Standard Newton, and in this case, a perturbed method [18] is added and its effects are investigated.

The rest of this chapter is organized as follows: NMPC formulation and the applied RTO methods as well as the MPsee tool for automatic code generation are described in Sect. 10.2. Next, the considered ACC problem is explained in Sect. 10.3. Section 10.4 includes the simulation results along with a thorough comparative discussion on the performance and computational speed of the devised

approaches. Finally, Sect. 10.5 concludes this study and suggests the associated future work.

10.2 NMPC Formulation and Implementation

In this section NMPC formulation, the GMRES-based optimizers utilized to solve the nonlinear root finding problem in real-time, and the methods applied in code generator are explained in detail.

10.2.1 Problem Formulation

The main task of an NMPC controller is to minimize a cost functional, which could be described in the general form of

$$J = \phi(x(T), p(T)) + \int_0^T L(x(\tau), u(\tau), p(\tau))d\tau, \tag{10.1}$$

subject to

$$g(x(\tau), u(\tau), p(\tau)) = 0 \, and \tag{10.2}$$

$$h(x(\tau), u(\tau), p(\tau)) \leq 0. \tag{10.3}$$

Here T represents the prediction horizon length and $0 \leq \tau \leq T$ is the prediction horizon time domain. Also, $\phi(.)$ is the terminal cost associated with the final state in the prediction horizon, $L(.)$ is the trajectory cost, and g and h represent the equality and inequality constraints, respectively.

A so-called control-oriented model is used at the heart of the NMPC controller for predicting states of the system over the prediction horizon. Such a model could be represented using the generic state equations

$$\dot{x}(t) = f(x(t), u(t), p(t)), \tag{10.4}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input vector, and $p(t) \in \mathbb{R}_p^n$ represents the known time-varying parameters in the system.

For the sake of implementation, we discretize the nonlinear optimization problem defined in (10.1) into N time-steps:

$$J = \phi(x_N, p_N) + \sum_{i=0}^{N-1} L(x_i, u_i, p_i)\Delta\tau, \tag{10.5}$$

subject to

$$x_{i+1} = x_i + f(x_i, u_i, p_i)\Delta\tau, \tag{10.6}$$

$$x_0 = x(t), \tag{10.7}$$

$$g(x_i, u_i, p_i) = 0, \ and \tag{10.8}$$

$$h(x_i, u_i, p_i) \leq 0, \tag{10.9}$$

where $\Delta\tau = T/N$ defines the stepping time inside the prediction horizon.

Among different methods suggested for handling inequality constraints [44], here the exterior penalty method is used, as it seems to be the most robust method with tremendously less tuning effort. Using this method, a relaxation of the objective function defined in (10.5) becomes

$$J = \phi(x_N, p_N) + \sum_{i=0}^{N-1} L(x_i, u_i, p_i)\Delta\tau + \sum_{j=0}^{l} \beta_j h_j(x, u)^2(x_i, u_i, p_i), \tag{10.10}$$

where

$$\beta_j = \begin{cases} 0 & if \ \ h_j(x, u) \leq 0 \\ \alpha_j & if \ \ h_j(x, u) > 0. \end{cases} \tag{10.11}$$

Moreover, applying the optimality necessary conditions results in

$$x_{i+1} = x_i + f(x_i, u_i, p_i)\Delta\tau, \tag{10.12a}$$

$$\lambda_i = \lambda_{i+1} + \left(\frac{\partial H}{\partial x}\right)^T (x_i, u_i, \lambda_{i+1}, \mu_i, p_i)\Delta\tau, \tag{10.12b}$$

$$\left(\frac{\partial H}{\partial u}\right)^T (x_i, u_i, \lambda_{i+1}, \mu_i, p_i) = 0, \tag{10.12c}$$

$$g(x_i, u_i, p_i) = 0, \tag{10.12d}$$

$$x_0 = x(t), \ and \tag{10.12e}$$

$$\lambda_N = \left(\frac{\partial \phi}{\partial x}\right)^T (x_N, p_N), \tag{10.12f}$$

where λ and μ denote Lagrange multipliers for dynamics of the system and constraints, respectively, and H represents the Hamiltonian:

$$H(x, u, \lambda, \mu, p) = L(x, u, p) + \lambda^T f(x, u, p) + \mu^T g(x, u, p). \tag{10.13}$$

The Two Point Boundary Value Problem (TPBVP) described in (10.12a–10.12f) can be written as a root finding problem [34], with single or multiple-shooting formulations. In a single-shooting approach, sequences $\{u_i | i = 0, \ldots, N-1\}$ and $\{\mu_i | i = 0, \ldots, N-1\}$ are taken as the optimization variables. Here the nonlinear equation

$$F(U(t), x(t)) = \begin{bmatrix} \left(\frac{\partial H}{\partial u}\right)^T (x_0, u_0, \lambda_1, \mu_0, p_0) \\ g(x_0, u_0, p_0) \\ \vdots \\ \left(\frac{\partial H}{\partial u}\right)^T (x_i, u_i, \lambda_{i+1}, \mu_i, p_i) \\ g(x_i, u_i, p_i) \\ \vdots \\ \left(\frac{\partial H}{\partial u}\right)^T (x_{N-1}, u_{N-1}, \lambda_N, \mu_{N-1}, p_{N-1}) \\ g(x_{N-1}, u_{N-1}, p_{N-1}) \end{bmatrix} = 0 \tag{10.14}$$

is to be solved, where $U(t)$ is in fact the vector of unknowns and defined as

$$U(t) = [u_0^T(t), \mu_0^T(t), u_1^T(t), \mu_1^T(t), \ldots, u_{N-1}^T(t), \mu_{N-1}^T(t)]^T. \tag{10.15}$$

In a multiple-shooting approach, sequences $\{x_i | i = 1, \ldots, N\}$ and $\{\lambda_i | i = 1, \ldots, N\}$ are added to the optimization variables, and the nonlinear equation becomes

$$F(U(t), x(t)) = \begin{bmatrix} \left(\frac{\partial H}{\partial u}\right)^T (x_0, u_0, \lambda_1, \mu_0, p_0) \\ x_1 - x_0 - f(x_0, u_0, p_0)\Delta\tau \\ \lambda_1 - \lambda_2 - \left(\frac{\partial H}{\partial x}\right)^T (x_1, u_1, \lambda_2, \mu_1, p_1)\Delta\tau \\ g(x_0, u_0, p_0) \\ \vdots \\ \left(\frac{\partial H}{\partial u}\right)^T (x_i, u_i, \lambda_{i+1}, \mu_i, p_i) \\ x_{i+1} - x_i - f(x_i, u_i, p_i)\Delta\tau \\ \lambda_{i+1} - \lambda_{i+2} - \left(\frac{\partial H}{\partial x}\right)^T (x_{i+1}, u_{i+1}, \lambda_{i+2}, \mu_{i+1}, p_{i+1})\Delta\tau \\ g(x_i, u_i, p_i) \\ \vdots \\ \left(\frac{\partial H}{\partial u}\right)^T (x_{N-1}, u_{N-1}, \lambda_N, \mu_{N-1}, p_{N-1}) \\ x_N - x_{N-1} - f(x_{N-1}, u_{N-1}, p_{N-1})\Delta\tau \\ \lambda_N - \left(\frac{\partial \phi}{\partial x}\right)^T (x_N, p_N) \\ g(x_{N-1}, u_{N-1}, p_{N-1}) \end{bmatrix}, \tag{10.16}$$

where the vector of unknowns is redefined as

$$U(t) = [u_0^T(t), \mu_0^T(t), x_1^T(t), \lambda_1^T(t), \ldots, u_{N-1}^T(t), \mu_{N-1}^T(t), x_N^T(t), \lambda_N^T(t)]^T. \tag{10.17}$$

10.2.2 Applied Algorithms

Last section demonstrated how an NMPC problem is transformed into a nonlinear root finding problem. Here, the methods applied to handle this problem are described. Consider $U^*(t_k)$ to be the solution to the equation $F(U(t), x(t)) = 0$ for the instant t_k. A Standard Newton iteration could be applied to find the solution to the next sampling time t_{k+1}

$$\hat{U}(t_{k+1}) = \hat{U}(t_k) - \left(\frac{\partial F}{\partial U}\Big|_{t_k}\right)^{-1} F(\hat{U}(t_k), x(t_k)), \tag{10.18}$$

where $\hat{U}(t_k)$ is an approximation of $U^*(t_k)$, which is used as in most applications the exact solution of the previous sampling time $U^*(t_k)$ is not at hand. The linear algebraic equation (10.18) could be solved with the so-called Forward-Difference Generalized Minimal Residual (FDGRMES), where only the product of a Jacobian and a vector is required at each iteration. Here, this method is referred to as *Newton/GMRES*. See [45] for details of Newton-iterative methods such as FDGMRES.

Alternatively, a predictor term could be considered to result in the so-called predictor-corrector path following method formula

$$\hat{U}(t_{k+1}) = \hat{U}(t_k) - \left(\frac{\partial F}{\partial U}\Big|_{t_k}\right)^{-1} \left(\frac{\partial F}{\partial x}\Big|_{t_k}\right) (x(t_{k+1}) - x(t_k)) - \left(\frac{\partial F}{\partial U}\Big|_{t_k}\right)^{-1} F(\hat{U}(t_k), x(t_k)). \tag{10.19}$$

Linear equation (10.19) is solved using FDGMRES algorithm and forms the basis of the C/GMRES method (see [34] for proof of convergence).

In addition to the methods discussed above, [18] proposes a modified Newton method, where the Jacobian obtained at the initial point is used again in the subsequent iterations and computation of the Jacobian at every iteration is avoided. For example, the modified version of Eq. (10.18) becomes

$$\hat{U}(t_{k+1}) = \hat{U}(t_k) - \left(\frac{\partial F}{\partial U}\Big|_{t_0}\right)^{-1} F(\hat{U}(t_k), x(t_k)). \tag{10.20}$$

This method is also referred to as perturbed chord method, originally introduced in [46]. It is proved that under appropriate conditions a modified Newton iteration still

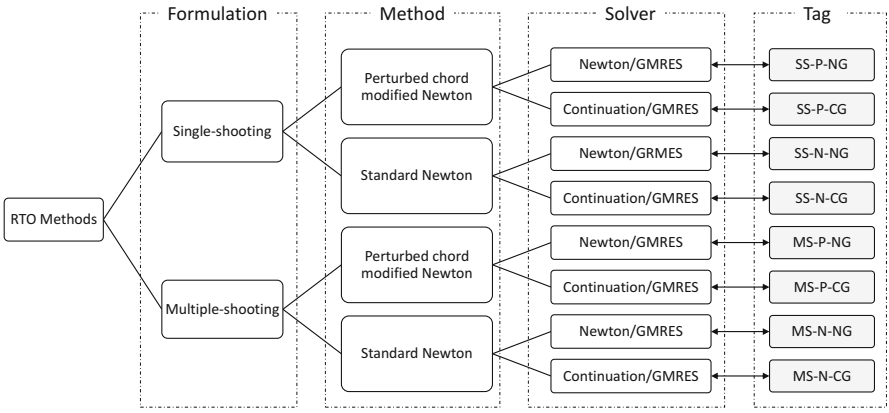


Fig. 10.1 Different GMRES-based RTO methods for NMPC implementation and their name tags

converges to a solution. Different RTO methods applied in this study are categorized in Fig. 10.1.

10.2.3 Implementation of Automatic NMPC Code Generation

In this section our developed Automatic NMPC Code Generation tool, MPsee, is explained. MPsee is a MATLAB-based tool that collects all the necessary information about the OCP, including the objective function, dynamics of the system, and the constraints, along with the solver options selected by the user. Based on this information, the tool generates the controller as MATLAB files, which can in turn be used for Model-in-the-loop (MIL) simulations and HIL experiments.

MPsee works in three consecutive phases: in the first phase, user provides the OCP definition through a Graphical User Interface (GUI). The GUI has predefined fields for the objective function, control-oriented model, equality and inequality constraints, the terminal cost, and so on. Moreover, the user can define the so-called user-defined parameters (UDPs). Additionally, time-varying parameters (TVPs) of the control-oriented model or objective function could be defined in the GUI, and the values of these parameters are received as signals during the simulations. For example, in an optimal tracking problem the reference signal could be defined as a TVP in the problem definition. It should be notified that TVPs can be defined frozen or dynamic, as an array over the prediction horizon. In the second phase, Hamiltonian and its partial derivatives are derived offline using MATLAB symbolic toolbox. This procedure significantly reduces the online computational load by preprocessing the problem offline. The resulting expressions are exported as MATLAB functions and repeatedly called during the next online calculations. In the

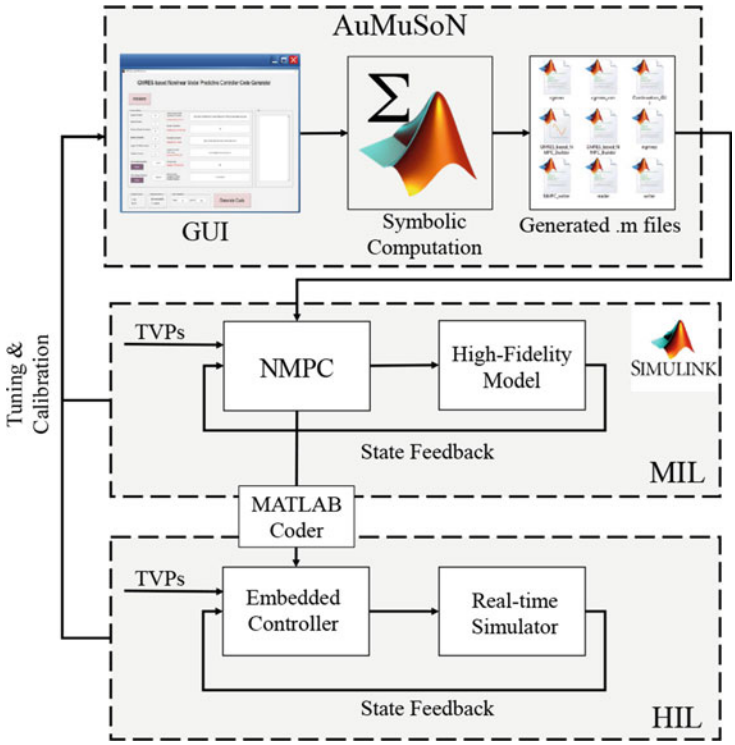


Fig. 10.2 Automatic NMPC code generation tool in simulation loops

third phase, the generated symbolic expressions are rearranged as several MATLAB functions, all of which contribute to forming the nonlinear equation $F(U(t), x(t))$ online. MPsee also generates a SIMULINK file that can be used for MIL simulations with small modifications. The main two blocks in SIMULINK *.mdl file*, namely the controller and the plant are shown in Fig. 10.2. The plant block includes the control-oriented model of the system, which can be then replaced with a high-fidelity model for MIL simulations. The controller block includes the NMPC execution code, which is generated based on the user-defined solver options and during the simulation calls other generated functions if necessary. The files generated by MPsee include five *.m files*, which are problem-dependent MATLAB functions. Using these along with three other problem-independent functions, SIMULINK simulation file can be run. MPsee output files are compatible with MATLAB C code generator, and the C codes necessary for HIL experiment are generated by MATLAB.

10.3 Adaptive Cruise Control Formulation

This study investigates the ACC problem for a Plug-in Hybrid Electric Vehicle (PHEV) as a common application of NMPC technique in automotive field. Figure 10.3 illustrates the system block diagram. The system is designed to both regulate the cruising speed of a PHEV on hilly roads and minimize the energy cost. Here receding horizon predictive control is used to to generate an optimal velocity trajectory using the road elevation profile ahead of the vehicle.

Using the vehicle longitudinal dynamics, the so-called control-oriented model can be written as

$$\begin{bmatrix} \dot{s}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} v(t) \\ u - \frac{1}{m} F_r \end{bmatrix}, \qquad F_r = \frac{1}{2}\rho A_f C_d v^2(t) + mg \sin\theta(s(t)) + \mu mg \cos\theta(s(t)). \tag{10.21}$$

Here, the travelled distance (s) and velocity (v) are the states of the system. $\theta(s(t))$ is the road grade which is given to the controller as a dynamic time-varying parameter inside the prediction horizon. It is assumed that the vehicle receives information about the road grade ahead via V2I, road maps and GPS. ρ denotes air density, A_f and C_d are the frontal area of the vehicle and the drag coefficient, respectively, and μ represents rolling resistance coefficient. $u(t)$ is the input, i.e., the vehicle longitudinal acceleration.

The objective function of the NMPC controller is

$$J = \frac{1}{2}\omega_1(s(T) - s_{ref}(T))^2 + \int_0^T \{\frac{1}{2}\omega_2(v(\tau) - v_{ref})^2 + \frac{1}{2}\omega_3(u(\tau) - \eta_{reg} u_{reg}^{max}(\tau))^2\} d\tau, \tag{10.22}$$

subject to

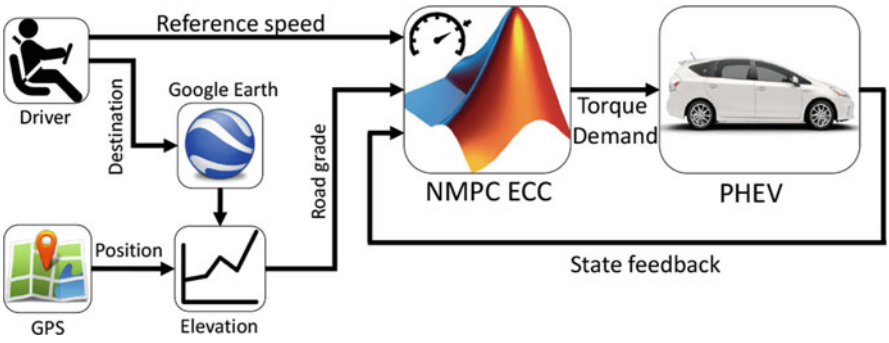


Fig. 10.3 Schematic illustration of the ACC problem

$$\begin{cases} u_{min} \leq u(\tau) \leq u_{max} \\ v_{min} \leq v(\tau) \leq v_{max}, \end{cases} \tag{10.23}$$

where v_{ref} is the cruise reference speed given to the controller by the driver as frozen time-varying parameter. ω_1,ω_2 and ω_3 are the weighting factors of the different cost terms. The first term on the right-hand side (RHS) is the terminal cost where $s_{ref}(T)$ is the target distance travelled by the end of each prediction window. This term helps reducing the trip time. The second RHS term refers to the tracking of the reference velocity as demanded by the driver. The last term on the RHS represents the torque demand minimization, where η_{reg} and $u_{reg}^{max}(\tau)$ represent the efficiency and the maximum deceleration achievable by regenerative braking at each time step. The ACC as described above is developed and tuned by executing MPsee tool.

10.4 Results and Discussion

In this section, performance of the proposed controller is showcased for a high-fidelity model of Toyota Plug-in Hybrid Prius [47]. This model consists of the longitudinal dynamics, tire characteristics, and hybrid powertrain architecture. Additionally, a driver model and a rule-based energy management controller are included in the model. A series of MIL simulation and HIL experiment results are presented to evaluate the controller performance.

10.4.1 MIL Simulations

MIL simulations are performed on an Intel Core i7 PC. A prediction horizon of size $N = 20$ and a time step of $\delta\tau = 1$s are chosen. Constraints on this system are in the form of (10.23) and the speed boundary limits are defined as $v_{min} = v_{ref} - 10$km/h and $v_{max} = v_{ref} + 10$km/h. The objective function's weighting factors are considered as $\omega_1 = 15$, $\omega_2 = 1$ and $\omega_3 = 15$ after proper tuning. Moreover, the exterior penalty weights in (10.11) are defined as $(\alpha_j)_i = 1/(i+1)$ for the jth inequality constraint violated at ith time step of the prediction horizon. This type of parameter definition leads to prioritizing the constraints violated at closer time-steps over those occurring farther. Figure 10.4 shows the simulation results for a 2km track, where the elevation profile of the track includes a rise followed by a flat part and then a fall as illustrated. In this scenario, it is presumed that the driver sets the reference speed to 80km/h for the first half of the track and then increases it to 100km/h for the rest of the trip. This scenario illustrates speed tracking characteristics of the NMPC as well as its potential in saving energy. For the sake of comparison, a PID controller is fine-tuned and tested in the same scenario. The Torque plot in Fig. 10.4 indicates a smoother torque demand profile leading to

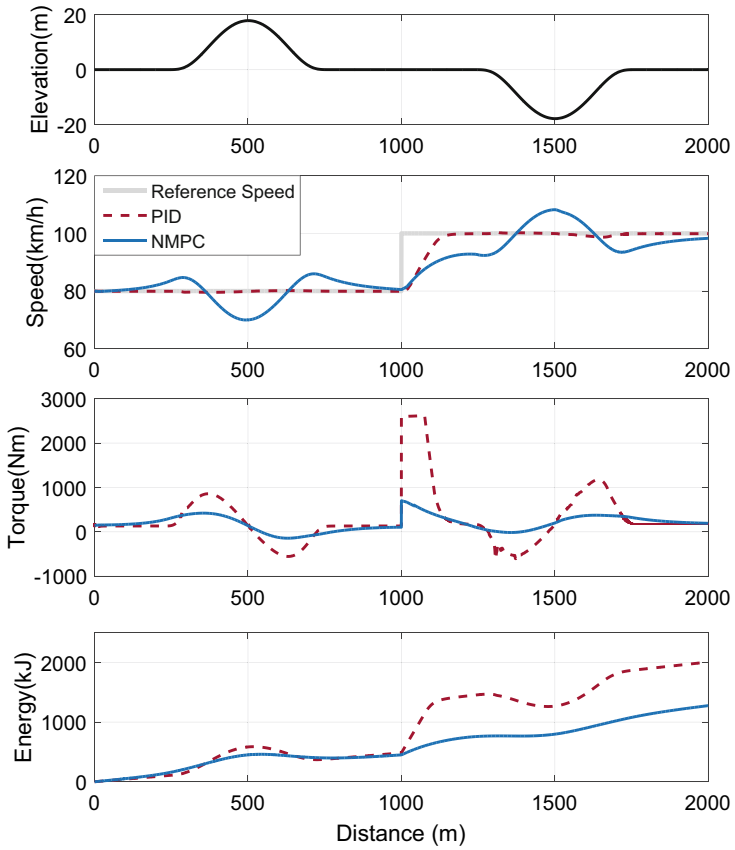


Fig. 10.4 Ecological speed control of a vehicle driving over a hilly road

a better ride comfort due to less jerk and harsh acceleration or deceleration when using the NMPC. The amount of energy saving using the proposed NMPC directly depends on the road elevation profile and the variations in the reference speed set by the driver. As such, while for the first half of the track the NMPC-based ACC saves an insignificant amount of energy compared to the PID, the pit in the second half of the track together with the increase in the reference speed by the driver create a proper condition for the NMPC to optimize the torque demand. Thus, an overall 36% improvement in the energy consumption is observed at the end of this trip, which comes at the cost of deviating from the reference speed within the permissible boundaries, in this case $\pm$10km/h. It is noteworthy that the performance shown in Fig. 10.4 represents the behavior of the NMPC approach regardless of its real-time optimization method.

In real-time implementation context, depending on the systems turnaround time, one can apply the Newton, modified Newton, or predictor-corrector Newton type iteration more than once in each sampling time. These iterations are often referred

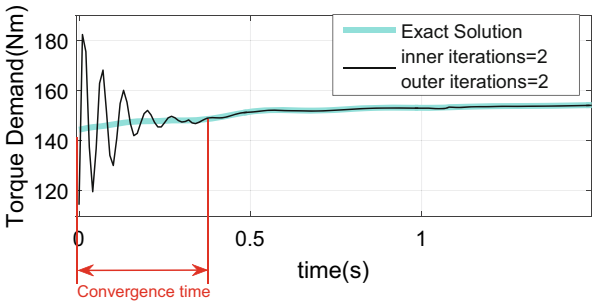


Fig. 10.5 Time to converge to an exact optimal solution for an NMPC with 2 inner iterations and 2 outer iterations

to as the *outer* iterations. Another factor significantly affecting the computational speed is the number of *inner* iterations inside the linear solver. Simulation results indicate that for a sufficient number of optimization iterations performed every sampling time, the optimality of the solution and hence the performance of the NMPCs with different real-time optimizers are the same. Different implementation methods of the real-time optimization, however, cause a relatively large variation of computational speeds, which is investigated further in the next section. Before such a comparative study, a prerequisite analysis is carried out to find the best combination of inner and outer iterations that presents a certain accuracy or convergence. Here the rate of convergence is measured through calculating the time it takes from the start of the control loop for the NMPC control input to converge to a small finite limit of an exact optimal solution as illustrated in Fig. 10.5.

A set of simulations with different maximum number of inner and outer iterations have been carried out for the two main shooting methods, and the measured convergence times are presented in Figs. 10.6 and 10.7. It is inferred from these figures that for both methods increasing the number of inner iterations has a more significant impact on reducing the convergence time. Moreover, a primary study of these simulations on PC reveals that the computational speed largely depends on the total number of inner iterations performed in every sampling time, i.e., product of numbers of inner and outer iterations. Figure 10.8 shows computational time as a function of these two numbers for both methods.

According to (10.14) and (10.16) and the selected prediction horizon size, nonlinear equation dimension is 20 and 60 for single- and multiple-shooting methods, respectively. Even though theoretically the same number of inner iterations as the dimension of the equation is adequate to achieve convergence, in practice an acceptable solution is obtained within a fewer number of inner iterations. Here it is observed that a minimum of 2 and 14 inner iterations are required for the single- and multiple-shooting approaches to be convergent, respectively. Besides, the exterior penalty constraint handling technique necessitates a minimum of 2 outer iterations since the calculated solution must be checked for constraint violation at least one time every sampling time. Considering Figs. 10.6, 10.7, and 10.8,

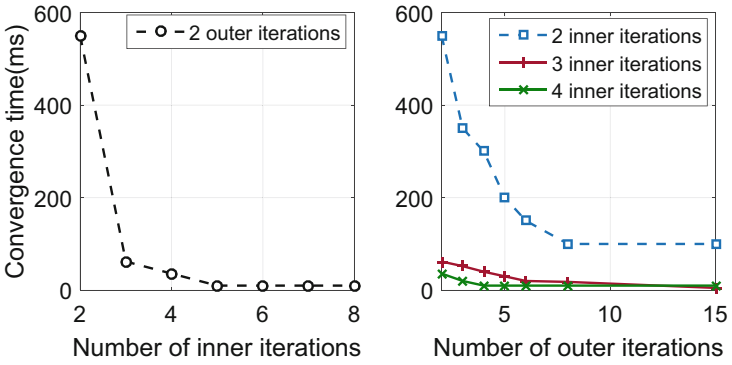


Fig. 10.6 The effect of increasing inner and outer iterations on the NMPC convergence in single-shooting approach

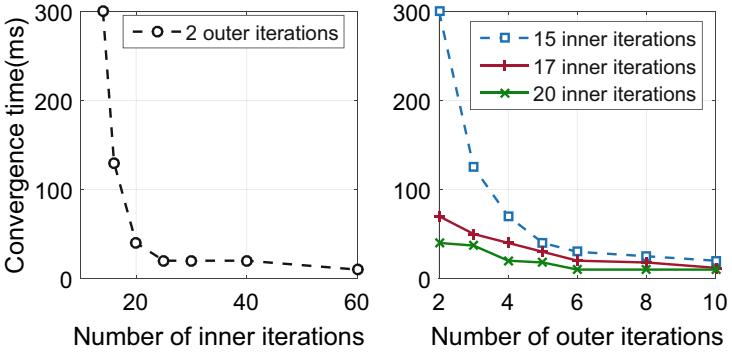


Fig. 10.7 The effect of increasing inner and outer iterations on the NMPC convergence in multiple-shooting approach

it is concluded that selecting a larger number of inner iterations per one outer iteration is a more computationally efficient way to achieve higher convergence rates. As such, 5 and 15 inner iterations are chosen for single- and multiple-shooting methods, respectively, while 2 outer iterations are performed for both approaches every sampling time.

10.4.2 HIL Experiments

Real-time implementability is one of the biggest challenges in control design. Here HIL tests with prototype ECUs are performed to investigate real-time implementability and compare computational speed of several RTO methods in ACC application. At the early stages of vehicle development, physical prototyping

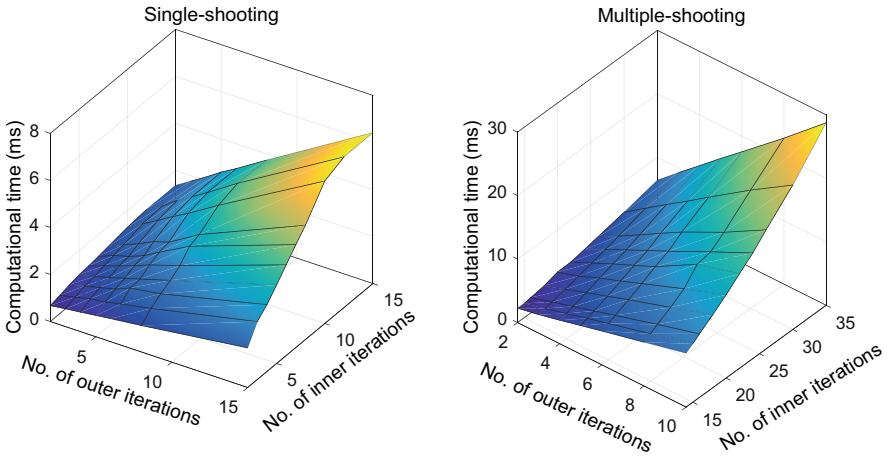


Fig. 10.8 The effect of increasing inner and outer iterations on the computational time of the single- and multiple-shooting approaches

can be costly and time-consuming, making rapid prototyping methods such as HIL testing more popular. HIL experiments are beneficial when merely virtual simulations are not sufficient for evaluating the system performance. HIL methods require less hardware than physical prototyping, and reduce the validation time and cost. Moreover, HIL tests facilitate repeatability in the design process, as various operating conditions can be subsequently simulated in the lab environment. Additionally, HIL techniques reduce the hazards and costs of destructive tests. All of these advantages make HIL testing an ideal calibration method for different phases of control development.

Figure 10.9 schematically demonstrates the HIL simulation platform used in this study. The dSPACE system consists of a MicroAutoBoxII (MAXII) prototype ECU and a DS-1006 real-time simulator communicating via CAN. The specifications for the hardware components are presented in Table 10.1. In this configuration, the NMPC is programmed onto the prototyped ECU and the Autonomie high-fidelity model of the vehicle plus the environment model are run on the real-time simulator. Meanwhile, the system's performance and signals from both platforms are monitored and recorded into the host PC with dSPACE ControlDesk software.

dSPACE includes associated MATLAB libraries, through which embeddable C codes can be generated. Any user-generated Simulink model and MATLAB code can be compiled using Real-Time Workshop code generator with the corresponding compiler of the target device.

As explained earlier, MPsee tool generates MATLAB codes and Simulink blocks that can be compiled into C code through this rapid programming method. Therefore, the task of iterative evaluation and calibration of the controller on a HIL platform is significantly facilitated (see Fig. 10.2). Here, the prototype ECU's turnaround time is used to measure computational speed and assess real-time implementability. Therefore, the MicroAutoBoxII prototype ECU's turnaround time

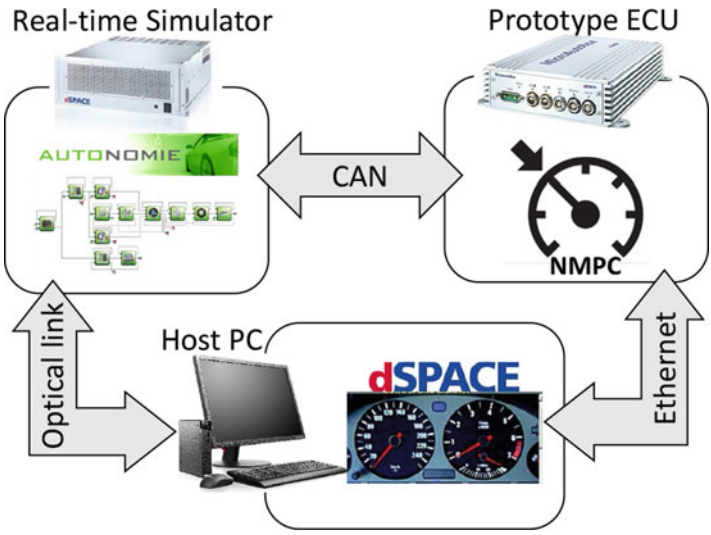


Fig. 10.9 Schematic layout of the HIL simulation platform

Table 10.1 dSPACE hardware specifications

Specification	Real-time simulator	Prototype ECU
Hardware	DS-1006	MicroAutoBoxII
Processor	Quad-Core AMD 2.8 GHz	DS-1401 PowerPC 750GL 900 MHz
Memory	1GB local, 4x128 MB global	16 MB main, 16 MB nonvolatile
I/O board	DS-2202	DS-1511

is reported for every HIL test in Figs. 10.10, 10.11, 10.12, where every color bar represents the range of turnaround times measured during the simulation. Figure 10.10 presents a comparison between different RTO implementation methods in terms of computational speed. Every method is recognized with a two-part label. The first part identifies the type of Newton approach to the nonlinear problem: Standard Newton (SN) or Perturbed Chord (PC). Then, the second part identifies the type of the linear solver used to solve for the Newton step: Newton/GMRES (NG) or Continuation/GMRES (CG). According to Fig. 10.10, methods based on single-shooting are executed significantly faster than the ones based on multiple-shooting. This is mainly due to the fact that using multiple-shooting greatly enlarges the size of the nonlinear equation and requires a higher number of inner iterations. Moreover, it is inferred that implementing a perturbed chord method has a more significant impact on improving the computational speed of the single-shooting approaches; turnaround time is reduced to almost half of its value when switching to perturbed chord method in single-shooting while this reduction is not as much for multiple-shooting. Besides, since in the Standard Newton approaches FDGMRES algorithm is used, the number of $F(U, x)$ evaluations are directly related to the number of

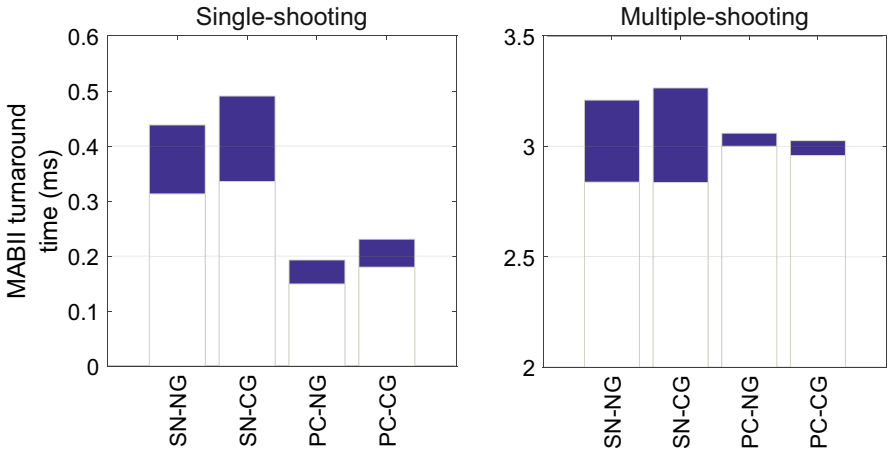


Fig. 10.10 Comparison of the computational speed of different RTO methods in HIL simulation

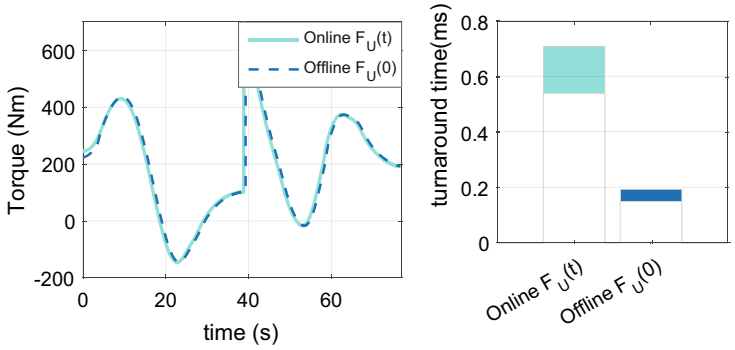


Fig. 10.11 The effect of different Jacobian calculation techniques on the response (left) and the computational speed (right) of Perturbed Chord method with Newton/GMRES linear solver in the single-shooting approach

inner iterations performed every sampling time. This number is variable during a simulation as it depends on accuracy of the initial guess. As such, the turnaround time is widely distributed. Whereas, turnaround time is less floating in the perturbed chord method where GMRES algorithm is used instead of FDGMRES.

In case of using perturbed chord RTO method, there are two approaches for computation of Jacobian. These approaches are compared for the single-shooting formulation with Newton/GMRES linear solver in Fig. 10.11. The first option is to numerically calculate the Jacobian, F_U, online once every sampling t at the initial outer iteration (Online $F_U(t)$). The second option would be precomputing the Jacobian offline at the initial state (Offline $F_U(0)$) and then approximating the Jacobian every sampling time with the offline precomputed matrix. It is observed that replacing an online computed Jacobian with an offline precomputed one at

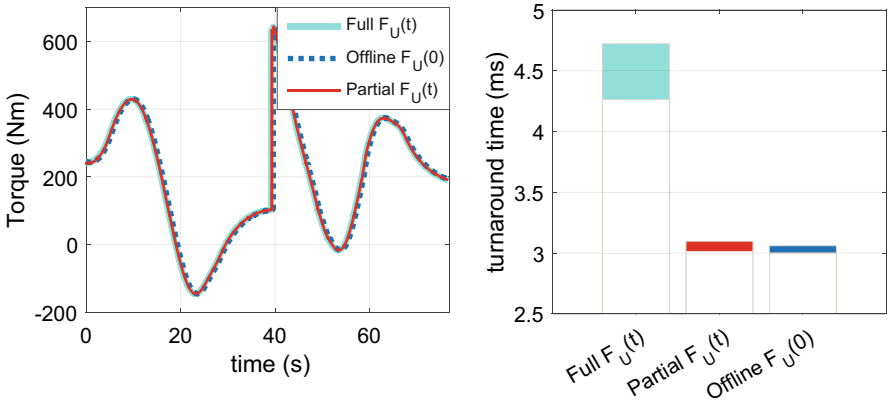


Fig. 10.12 The effect of different Jacobian calculation techniques on the response (left) and computational speed (right) of Perturbed Chord method with Newton/GMRES linear solver in the multiple-shooting approach

the initial state notably reduces the turnaround time while generating almost the same solution. A similar conclusion can be made for the multiple-shooting approach by studying Fig. 10.12. However, sparsity of the Jacobian matrix in this case allows for to considering a third option in computation of the Jacobian. In this technique, a symbolic operation is performed offline to detect the zero elements of the Jacobian and eliminates the need for numerically calculation of those elements online. Consequently, the Jacobian matrix is only partially computed online every sampling time. The results for this method of implementation are represented in Fig. 10.12 as *Partial* $F_U(t)$. It is observed that by using this partial approach a low turnaround time, close to that of the case with offline precomputed Jacobian, is achieved. Furthermore, this method increases robustness as the Jacobian is partially computed online every sampling time rather than being estimated by a precomputed value.

10.5 Conclusion and Future Work

In this work, we proposed a comparative study of NMPC-based RTOs in the context of ACC. The controllers were formulated using a so-called indirect method, and implemented by the MATLAB-based automatic NMPC code generation tool called "MPsee" for the baseline vehicle, namely Toyota Plug-in Hybrid Prius. To evaluate the performance of the proposed RTOs, several MIL simulations on artificial road profiles and driving scenarios as well as a series of HIL experiments with a prototype ECU in real-time were carried out. It was observed that GMRES-based methods with a single-shooting formulation were faster than those working based on a multiple-shooting formulation. Furthermore, utilizing a perturbed chord

modified Newton method instead of the Standard Newton method improved the prototype ECU's turnaround time. This improvement was more significant in case of single-shooting approach than multiple-shooting and turnaround time dropped by approximately 50%.

One could use condensing to reduce the size of the linear equation and execute multiple-shooting methods faster. As such, an area for future work would be investigating the effect of condensing and comparing single- and multiple-shooting approaches in term of the computational cost then. Furthermore, despite running faster in real-time, the perturbed chord modified Newton method requires offline pre-computation of the Jacobian in several working points that may cause losing robustness to uncertain constraint violations and remote working points. Future research is necessary to address these issues. Another desirable area for future studies is investigating the impacts of using MATLAB C code generation on the computational time of controller, as tailored C code generation is expected to improve the computational time.

Acknowledgments The authors would like to thank NSERC and Toyota for supporting this study.

References

1. JN Hooker. Optimal driving for single-vehicle fuel economy. *Transportation Research Part A: General*, 22(3):183–201, 1988.
2. Matthew Barth and Kanok Boriboonsomsin. Energy and emissions impacts of a freeway-based dynamic eco-driving system. *Transportation Research Part D: Transport and Environment*, 14(6):400–410, August 2009.
3. Nicholas J Kohut, J Karl Hedrick, and Francesco Borrelli. Integrating Traffic Data and Model Predictive Control to Improve Fuel Economy. In *12th IFAC Symposium on Control in Transportation Systems*, 2009.
4. Sagar Akhegaonkar, Sebastien Glaser, Lydie Nouveliere, and Frederic Holzmann. Advanced Microsystems for Automotive Applications 2007. In *Advanced Microsystems for Automotive Applications 2014*, pages 15–27. Springer, 2014.
5. Yugong Luo, Tao Chen, Shuwei Zhang, and Keqiang Li. Short Papers Intelligent Hybrid Electric Vehicle ACC With Coordinated Control of Tracking Ability , Fuel Economy , and Ride Comfort. pages 1–6, 2015.
6. Payman Shakouri and Andrzej Ordys. Nonlinear Model Predictive Control approach in design of Adaptive Cruise Control with automated switching to cruise control. *Control Engineering Practice*, 26:160–177, May 2014.
7. Sadegh Tajeddin, Mahyar Vajedi, and Nasser L Azad. A Newton/GMRES approach to predictive ecological adaptive cruise control of a plug-in hybrid electric vehicle in car-following scenarios. *IFAC-PapersOnLine*, 49(21):59–65, 2016.
8. Bijan Sakhdari and Nasser L. Azad. A distributed reference governor approach to ecological cooperative adaptive cruise control. *IEEE Transactions on Intelligent Transportation Systems*, 19:1496–1507, 2018.
9. Stefano Di Cairano and Ilya V. Kolmanovsky. Real-time optimization and model predictive control for aerospace and automotive applications. *2018 Annual American Control Conference (ACC)*, pages 2392–2409, 2018.

10. Kaijiang Yu, Junqi Yang, and Daisuke Yamaguchi. Model predictive control for hybrid vehicle ecological driving using traffic signal and road slope information. *Control Theory and Technology*, 13(1):17–28, 2015.
11. Tao Chen, Yugong Luo, and Keqiang Li. Multi-objective adaptive cruise control based on nonlinear model predictive algorithm. *Proceedings of 2011 IEEE International Conference on Vehicular Electronics and Safety*, pages 274–279, 2011.
12. D.Q. Mayne, James Rawlings, Christopher Rao, and P. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 06 2000.
13. M. Vajedi and N. L. Azad. Ecological adaptive cruise controller for plug-in hybrid electric vehicles using nonlinear model predictive control. *IEEE Transactions on Intelligent Transportation Systems*, 17(1):113–122, 2016.
14. Bijan Sakhdari, Ebrahim Moradi Shahrivar, and Nasser L. Azad. Robust tube-based mpc for automotive adaptive cruise control design. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6, 2017.
15. Lars Grüne and Jürgen Pannek. Nonlinear model predictive control. In *Nonlinear Model Predictive Control*, pages 45–69. Springer, 2017.
16. Yuichi Shimizu, Toshiyuki Ohtsuka, and Moritz Diehl. A real-time algorithm for nonlinear receding horizon control using multiple shooting and continuation/Krylov method. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 19(8):919–936, 2009.
17. Mike Huang, Hayato Nakada, Ken Butts, and Ilya Kolmanovsky. Nonlinear model predictive control of a diesel engine air path: A comparison of constraint handling and computational strategies. *IFAC-PapersOnLine*, 48(23):372–379, 2015.
18. Ken Butts, Asen Dontchev, Mike Huang, and Ilya Kolmanovsky. A perturbed chord (Newton-Kantorovich) method for constrained nonlinear model predictive control. *IFAC-PapersOnLine*, 49(18):253–258, 2016.
19. Kevin Walker, Behzad Samadi, Mike Huang, Jürgen Gerhard, Kenneth Butts, and Ilya Kolmanovsky. Design environment for nonlinear model predictive control. Technical report, SAE Technical Paper, 2016.
20. Benjamin Passenberg. *Theory and algorithms for indirect methods in optimal control of hybrid systems*. PhD thesis, Technische Universität München, 2012.
21. Bart Saerens, Hesham A Rakha, Moritz Diehl, and Eric Van den Bulck. A methodology for assessing eco-cruise control for passenger vehicles. *Transportation research part D: transport and environment*, 19:20–27, 2013.
22. Max Leo Correa Córdova. High performance implementation of mpc schemes for fast systems. 2016.
23. Harald Waschl, Ilya Kolmanovsky, Maarten Steinbuch, and Luigi Del Re. *Optimization and optimal control in automotive systems*, volume 455. Springer, 2014.
24. Nicolas Petit and Antonio Sciarretta. Optimal drive of electric vehicles using an inversion-based trajectory generation approach. In *The 18th IFAC World Congress*, pages 14519–14526, 2011.
25. Janick V Frasch, Andrew Gray, Mario Zanon, Hans Joachim Ferreau, Sebastian Sager, Francesco Borrelli, and Moritz Diehl. An auto-generated nonlinear mpc algorithm for real-time obstacle avoidance of ground vehicles. In *2013 European Control Conference (ECC)*, pages 4136–4141. IEEE, 2013.
26. Michael Dellnitz, Julian Eckstein, Kathrin Flaßkamp, Patrick Friedel, Christian Horenkamp, Ulrich Köhler, Sina Ober-Blöbaum, Sebastian Peitz, and Sebastian Tiemeyer. Development of an intelligent cruise control using optimal control methods. *Procedia Technology*, 15:285–294, 2014.
27. Thomas Stanger and Luigi del Re. A model predictive cooperative adaptive cruise control approach. In *2013 American Control Conference*, pages 1374–1379. IEEE, 2013.
28. Bart Saerens, Jeroen Vandersteen, Tim Persoons, Jan Swevers, Moritz Diehl, and Eric Van den Bulck. Minimization of the fuel consumption of a gasoline engine using dynamic optimization. *Applied energy*, 86(9):1582–1588, 2009.

29. Bart Saerens, Hesham Rakha, Kyoungho Ahn, and Eric Van Den Bulck. Assessment of Alternative Polynomial Fuel Consumption Models for Use in Intelligent Transportation Systems Applications. *Journal of Intelligent Transportation Systems*, 17(4):294–303, October 2013.
30. Fanping Bu, Han-Shue Tan, and Jihua Huang. Design and field testing of a cooperative adaptive cruise control system. In *Proceedings of the 2010 American Control Conference*, pages 4616–4621. IEEE, 2010.
31. Wissam Dib, Alexandre Chasse, Philippe Moulin, Antonio Sciarretta, and Gilles Corde. Optimal energy management for an electric vehicle in eco-driving applications. *Control Engineering Practice*, pages 1–9, February 2014.
32. P Kock, HJ Welfers, B Passenberg, S Gnatzig, O Stursberg, and AO Ordys. Saving energy through predictive control of longitudinal dynamics of heavy trucks. *VDI Berichte*, pages 53–68, 2008.
33. Antonio Sciarretta, Giovanni De Nunzio, and Luis Leon Ojeda. Optimal ecodriving control: Energy-efficient driving of road vehicles as an optimal control problem. *IEEE Control Systems Magazine*, 35(5):71–90, 2015.
34. Toshiyuki Ohtsuka. A continuation/GMRES method for fast computation of nonlinear receding horizon control. *Automatica*, 40(4):563–574, 2004.
35. Md. Abdus Samad Kamal, Masakazu Mukai, Junichi Murata, and Taketoshi Kawabe. Model Predictive Control of Vehicles on Urban Roads for Improved Fuel Economy. *IEEE Transactions on Control Systems Technology*, 21(3):831–841, May 2013.
36. Dongmei Wu, Yang Xia, Masatoshi Ogawa, Harutoshi Ogai, and Taketoshi Kawabe. Automobile optimal driving control using surrounding information based on model predictive control. *SICE Journal of Control, Measurement, and System Integration*, 3(6):435–441, 2010.
37. Toshiyuki Ohtsuka. A tutorial on C/GMRES and automatic code generation for nonlinear model predictive control. In *2015 European Control Conference (ECC)*, pages 73–86. IEEE, 2015.
38. Boris Houska, Hans Joachim Ferreau, and Moritz Diehl. Acado toolkit–an open-source framework for automatic control and dynamic optimization. *Optimal Control Applications and Methods*, 32(3):298–312, 2011.
39. Milan Vukov, Alexander Domahidi, Hans Joachim Ferreau, Manfred Morari, and Moritz Diehl. Auto-generated algorithms for nonlinear model predictive control on long and on short horizons. In *52nd IEEE Conference on Decision and Control*, pages 5113–5118. IEEE, 2013.
40. S. Tajeddin. MPsee, https://www.mathworks.com/matlabcentral/fileexchange/64608-mpse, 2017.
41. Moritz Diehl, H Georg Bock, Johannes P Schlöder, Rolf Findeisen, Zoltan Nagy, and Frank Allgöwer. Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *Journal of Process Control*, 12(4):577–585, 2002.
42. Victor M Zavala and Lorenz T Biegler. The advanced-step NMPC controller: Optimality, stability and robustness. *Automatica*, 45(1):86–93, 2009.
43. Sadegh Tajeddin and Nasser L Azad. Ecological cruise control of a plug-in hybrid electric vehicle: A comparison of different GMRES-based nonlinear model predictive controls. In *2017 American Control Conference (ACC)*, pages 3607–3612. IEEE, 2017.
44. Andreas Fischer. A special newton-type optimization method. *Optimization*, 24(3–4):269–284, 1992.
45. Carl T Kelley. *Iterative methods for linear and nonlinear equations*. SIAM, 1995.
46. Leonid Vital'evich Kantorovich. Functional analysis and applied mathematics. *Uspekhi Matematicheskikh Nauk*, 3(6):89–185, 1948.
47. Amir Taghavipour, Ramin Masoudi, Nasser L. Azad, and John McPhee. High-fidelity modeling of a power-split plug-in hybrid electric powertrain for control performance evaluation. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 55843, page V001T01A008. American Society of Mechanical Engineers, 2013.

Chapter 11
Influence of Lateral Asymmetry on Car's Lateral Dynamics

Dai Q. Vo, Van Tan Vu, Hormoz Marzbani, Mohammad Fard, and Reza N. Jazar

Abstract Research on vehicle dynamics usually works on the assumption that vehicle is bilaterally symmetrical. In practice, however, cars may have laterally asymmetrical characteristics due to different reasons. This investigation looks into how the lateral asymmetry affects the dynamic behavior of a steering vehicle. In order to do so, an unconventional planar model of laterally asymmetric vehicle is built to derive the equations of motion; Matlab/Simulink platform is utilized to simulate the dynamic responses of the vehicle to a ramp steer input. The responses are then put together with those of the baseline for comparison. The results show that, the vehicle with its center of gravity (CoG) deviating toward rotation center produces more lateral grip than the symmetric car; whilst the vehicle with its CoG being on the other side of rotation center provides less lateral grip than the baseline. The most important outcome of this investigation is that, a laterally asymmetrical vehicle turning to the heavier side can produce about 7% higher lateral grip capacity than turning to the lighter side.

Keywords Vehicle dynamics · Bilaterally symmetrical · Lateral asymmetry · Lateral acceleration · Lateral grip · Steering car · Cornering car

D. Q. Vo (✉)
Faculty of Vehicle and Energy Engineering, Le Quy Don Technical University, Hanoi, Vietnam
e-mail: dai.voquoc@lqdtu.edu.vn

V. T. Vu
Faculty of Mechanical Engineering, University of Transport and Communications, Hanoi, Vietnam
e-mail: vvtan@utc.edu.vn

H. Marzbani · M. Fard · R. N. Jazar
School of Engineering, RMIT University, Melbourne, Australia
e-mail: hormoz.marzbani@rmit.edu.au; mohamad.fard@rmit.edu.au; reza.jazar@rmit.edu.au

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_11

11.1 Introduction

In the literature on lateral vehicle dynamics, it has been widely assumed that the car's CoG is on the geometric longitudinal centerline. That assumption simplifies the models and is good enough for most of the cases from studies providing basic understanding of vehicle dynamics [3, 4, 12] to research developing high tech vehicles in recent years [6–8, 10, 11].

However, in reality a car may be laterally asymmetric—the CoG locates at a point away from its geometric longitudinal centerline. The lateral asymmetry of a car can be attributed to different reasons such as manufacturing aspects, how the car seats are occupied, and/or how the cargo is loaded. When the car is not bilaterally symmetrical, the static normal load under each wheel and the lateral transfer of normal load from the inner to the outer wheels (herein, there is no difference in the use of wheel and tyre) when cornering will be different from those of bilaterally symmetrical car. As the lateral force of the tyre is a non-linear function of the normal force, the redistribution of the normal force during the turn will facilitate or lessen lateral grip of the asymmetrical vehicle. In this paper, how the lateral location of the car's CoG affects the dynamic behavior of a steering car is investigated. A dynamic model of the asymmetrical car is built to generate the equations of motion for the case. The dynamic responses of the car to a slowly increasing steer input will be obtained to examine how much the lateral dynamics of an asymmetrical car differ from that of the baseline car.

11.2 Modelling

In order to investigate the effect of the CoG's lateral location on the dynamics of the steering car, a planar dynamic model of car as shown in Fig. 11.11 is built. The x-axis passes through its center of gravity C and is parallel to the geometric longitudinal centerline. This model can capture the main features of a laterally asymmetrical car such as uneven static normal load distribution between the left and the right wheels, the lateral load transfer of the vertical forces when cornering, and the unconventional kinematic relations between the car's CoG and its wheels. The change of the car's yaw inertia due to the asymmetry is insignificant. Therefore, it is neglected in this analysis.

The lateral dynamics of such a car is governed by the three following equations of motion:

$$m\dot{v}_x - m\omega_z v_y = -F_{yfI}\sin\delta_{fI} - F_{yfO}\sin\delta_{fO} + F_{xfI}\cos\delta_{fI} + F_{xfO}\cos\delta_{fO} \tag{11.1}$$

$$m\dot{v}_y + m\omega_z v_x = F_{yfI}\cos\delta_{fI} + F_{yfO}\cos\delta_{fO} + F_{xfI}\sin\delta_{fI} + F_{xfO}\sin\delta_{fO} + F_{yrI} + F_{yrO} \tag{11.2}$$

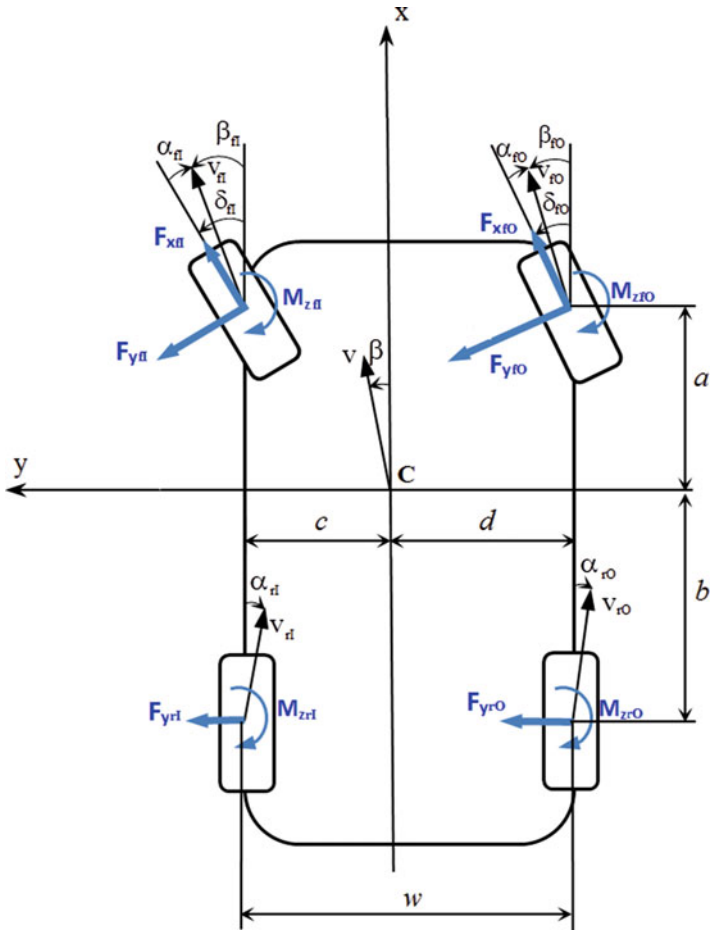


Fig. 11.1 Vehicle dynamic model of laterally asymmetric car

$$\dot{\omega}_z I_z = -(M_{zfI} + M_{zfO} + M_{zrI} + M_{zrO}) - c(F_{xfI}\cos\delta_{fI} - F_{yfI}\sin\delta_{fI}) + d(F_{xfO}\cos\delta_{fO} - F_{yfO}\sin\delta_{fO}) + a(F_{yfI}\cos\delta_{fI} + F_{xfI}\sin\delta_{fI} + F_{yfO}\cos\delta_{fO} + F_{xfO}\sin\delta_{fO}) - b(F_{yrI} + F_{yrO}) \quad (11.3)$$

Herein, F_x, F_y, and M_z, respectively, represent longitudinal, lateral forces, and aligning moment of the tyres. Also, the subscripts f, r, I, and O, respectively, represent front, rear, inner, and outer tyres of the car. F_x is the longitudinal force required to keep the longitudinal velocity of the car constant. The tyre lateral force F_y and aligning moments M_z are described by the Magic Formula [1] to include their non-linear characteristics:

- Lateral force:

$$F_y(F_z, \alpha) = D \sin[C \tan^{-1}(B\Phi)] \tag{11.4}$$

where

$$\Phi = (1 - E)\alpha + \frac{E}{B} \tan^{-1}(B\alpha) \tag{11.5}$$

$$B = \frac{a_3 \sin[a_4 \tan^{-1}(a_5 F_z)]}{CD}; C = 1.30 \tag{11.6}$$

$$D = a_1 {F_z}^2 + a_2 F_z; E = a_6 {F_z}^2 + a_7 F_z + a_8 \tag{11.7}$$

- Self-aligning moment:

$$M_z(\alpha, \gamma) = D \sin[C \tan^{-1}(B\Phi)] \tag{11.8}$$

where

$$\Phi = (1 - E)\alpha + \frac{E}{B} \tan^{-1}(B\alpha) \tag{11.9}$$

$$B = \frac{a_3 {F_z}^2 + a_4 F_z}{CDe^{a_5 F_z}}; C = 2.40 \tag{11.10}$$

$$D = a_1 {F_z}^2 + a_2 F_z; E = a_6 {F_z}^2 + a_7 F_z + a_8 \tag{11.11}$$

here B, C, D, and E are called stiff factor, shape factor, peak factor, and curvature factor, respectively; a_i (i =1...8) are the coefficients obtained through regression techniques to best match the mathematical formula with the experiment data [1, 2, 9].

F_z is the tyre vertical force. It is a function of longitudinal and lateral accelerations. Unlike conventional model of laterally symmetrical car, the vertical forces under tyres of the asymmetrical vehicle are as follows [5]:

$$F_{zf1} = \frac{b}{L}\frac{d}{w}mg - \frac{h_g}{L}\frac{d}{w}ma_x - \frac{b}{L}\frac{h_g}{w}ma_y + \frac{h_g^2}{Lw}m\frac{a_x a_y}{g} \tag{11.12}$$

$$F_{zfO} = \frac{b}{L}\frac{c}{w}mg - \frac{h_g}{L}\frac{c}{w}ma_x + \frac{b}{L}\frac{h_g}{w}ma_y - \frac{h_g^2}{Lw}m\frac{a_x a_y}{g} \tag{11.13}$$

$$F_{zrI} = \frac{a}{L}\frac{d}{w}mg + \frac{h_g}{L}\frac{d}{w}ma_x - \frac{a}{L}\frac{h_g}{w}ma_y - \frac{h_g^2}{Lw}m\frac{a_x a_y}{g} \tag{11.14}$$

$$F_{zrO} = \frac{a}{L}\frac{c}{w}mg + \frac{h_g}{L}\frac{c}{w}ma_x + \frac{a}{L}\frac{h_g}{w}ma_y + \frac{h_g^2}{Lw}m\frac{a_x a_y}{g} \tag{11.15}$$

The side slip α of each tyre is obtained from the asymmetrical kinematic relationship between the center of gravity C and each wheel. The side slip angles for the four wheels are written as:

$$\alpha_{fI} = \delta_{fI} - \tan^{-1}\frac{v_y + ar}{v_x - cr} \tag{11.16}$$

$$\alpha_{fO} = \delta_{fO} - \tan^{-1}\frac{v_y + ar}{v_x + dr} \tag{11.17}$$

$$\alpha_{rI} = -\tan^{-1}\frac{v_y - br}{v_x - cr} \tag{11.18}$$

$$\alpha_{rO} = -\tan^{-1}\frac{v_y - br}{v_x + dr} \tag{11.19}$$

After integrating formula from (11.4) to (11.19) into Eqs. (11.1), (11.2), and (11.3), the equations of motion are solved in Simulink environment for examining the dynamic behavior of the laterally asymmetrical steering car.

11.3 Lateral Dynamics of Asymmetrical Car

In order to examine the effects of lateral asymmetry on the dynamics of a steering car, a baseline car $1994 Ford Taurus Gsedan$ with parameters shown in Table 11.1 is chosen [2, 9]. The responses of the symmetrical and asymmetrical cars to a slowly increasing steering input of 1 deg/s (at ground wheels), at velocity of 70 km/h are obtained. More specifically, the three following configurations are investigated:

- A symmetrical car with $c = d = \frac{w}{2}$
- A car with CoG deviating toward the inner side of the turn, $c = \frac{w}{2} - 0.1(\text{m})$
- A car with CoG deviating toward the outer side of the turn, $c = \frac{w}{2} + 0.1(\text{m})$

The two asymmetrical configurations can occur in practice when only two seats on the same side of the car are occupied.

Table 11.1 Vehicle parameters

Parameter	Value	Parameter	Value
m(kg)	1530	b(m)	1.636
I_z(kgm^2)	4607	h_g(m)	0.5185
a(m)	1.14	w(m)	1.55

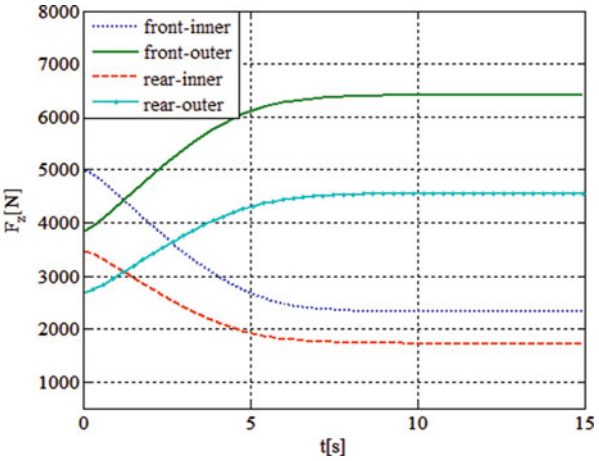


Fig. 11.2 The normal load under each tyre of a car with CoG deviating toward the rotation center

11.3.1 Car's CoG Deviating Toward the Inner

The normal load under each tyre of an asymmetrical car in the case where its CoG is toward the inner side of the turn is illustrated in Fig. 11.2. Because of the lateral asymmetry, the static normal load of each inner tyre (associated with $t = 0$) is greater than that of the opposite outer tyre. When the car is cornering, the lateral acceleration gives a rise to a normal load transfer from the inner to the outer tyres: the vertical force of the inner tyre decreases while that of the opposite tyre increases by the same amount.

Figure 11.3 visualizes how the car with CoG deviating toward the inner differs from symmetrical car in terms of normal loads under front tyres. As can be seen from the graph, there is an increase in normal load of the inner tyre of the asymmetrical car compared to that of the baseline. By contrast, the vertical force of the outer tyre decreases by the same amount compared to that of the symmetrical car. Therefore, the difference between normal loads of inner and outer tyres of the asymmetrical car in this case is less than that of the symmetrical car. This feature is also retained for rear axle.

It has been widely acknowledged that the normal load transfer causes a loss in the total lateral force of each axle. This is because lateral force is a non-linear function of the normal load, as demonstrated in Fig. 11.4. Also, the larger the difference between normal loads of inner and outer tyres ($2\Delta F_z$), the greater the loss of total lateral force (ΔF_y).

Because the inner-outer difference of normal load for the car with CoG deviating toward the inner is less than that of the baseline car, the normal load transfer effect of this asymmetrical car is weaker—the lateral force loss for this asymmetrical

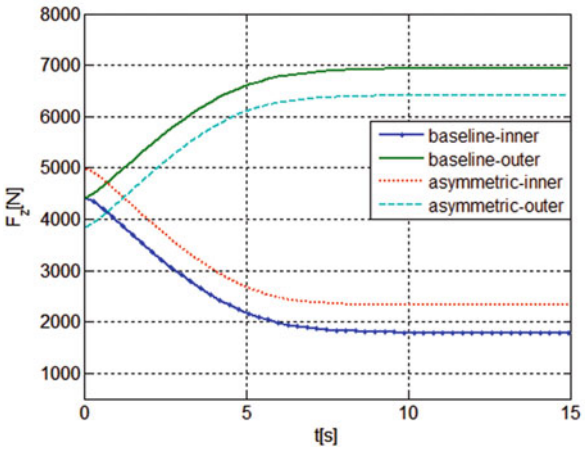


Fig. 11.3 The normal load under front tyres: car with CoG deviating toward the rotation center vs symmetric car

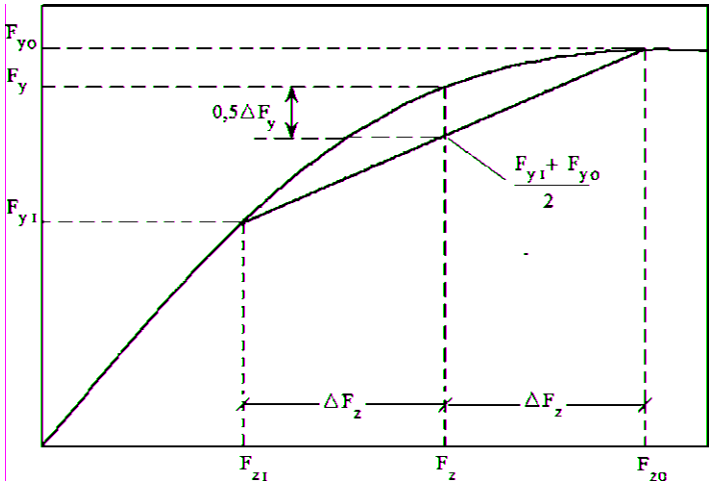


Fig. 11.4 Load transfer effect: normal load transfer causes loss of total lateral force

vehicle is less than that of the baseline. To be more specific, Fig. 11.5 compares the lateral forces of the front tyres for the two cases. It is shown that, there is a rise in the inner tyre lateral force and a drop in the outer tyre lateral force of the asymmetrical car but the rise surpasses the drop. Therefore, the resultant lateral force of the axle for the asymmetrical case is greater than the baseline. This is the reason why the asymmetrical car in this case can produce more lateral force than the baseline vehicle, as shown in Fig. 11.6.

Figure 11.7 demonstrates the variation in the lateral acceleration of a car when its center of gravity is laterally shifted toward the inner. As can be observed, when the car's CoG moves toward the inner side of the turn its lateral grip is improved at any steering angle. Although the improvement of lateral acceleration is only around

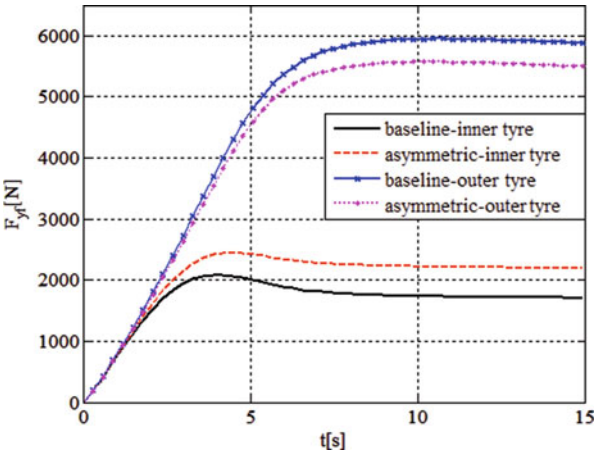


Fig. 11.5 The lateral forces under front tyres: car with CoG deviating toward the rotation center vs symmetric car

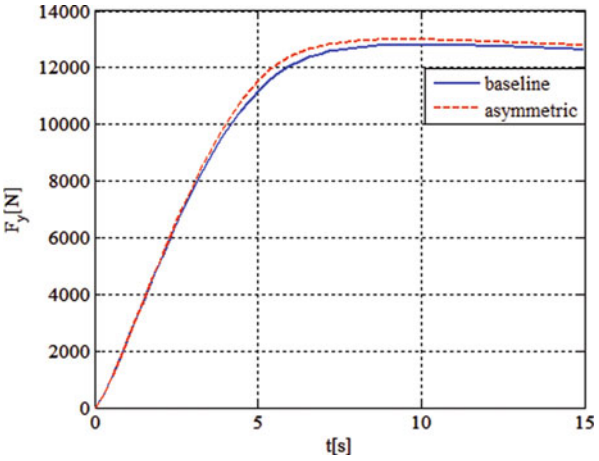


Fig. 11.6 The car resultant lateral force: car with CoG deviating toward the rotation center vs symmetric car

0.2 m/s^2, it is of importance in the region where the lateral force and acceleration are saturated—the car cannot produce more grip to track the turn despite increasing steering angle.

11.3.2 Car's CoG Deviating Toward the Outer

Figure 11.8 illustrates the vertical load under each tyre of an asymmetrical car in the case where its CoG deviates toward the outer side of the rotation center. The static vertical load of each inner tyre (associated with $\delta = 0$) is lower than that of the opposite outer tyre due to the fact that the car's CoG is located on the outer side of rotation center.

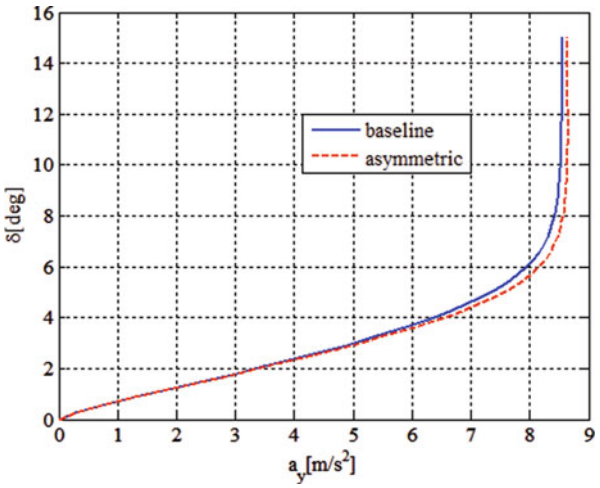


Fig. 11.7 The car lateral acceleration: car with CoG deviating toward the rotation center vs symmetric car

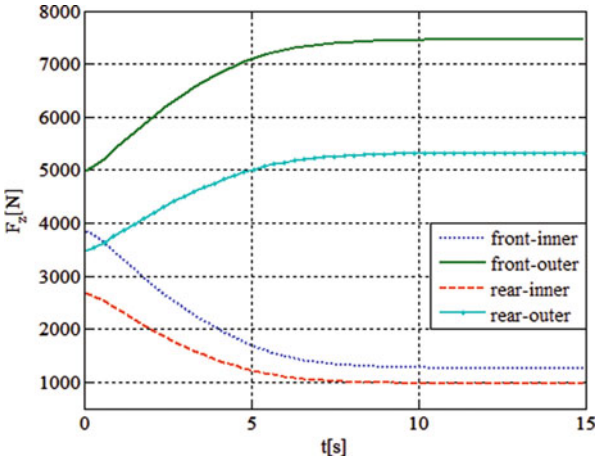


Fig. 11.8 The normal load under each tyre of the car with CoG deviating toward the outer side of the turn

When the car is tracking the turn, the normal load of the inner side transfers to the outer side. In comparison to vertical load distribution of the symmetrical car in Fig. 11.9, it is clear that the inner-outer normal load difference of the asymmetrical vehicle is greater than that in the baseline. Therefore, the lateral force loss in this asymmetrical vehicle is greater than the baseline (the normal load transfer effect of the asymmetrical car is stronger). Figure 11.10 gives a closer look at the lateral forces developed at front tyres for the two cases. It appears that, there is a drop in the inner tyre lateral force and a rise in the outer tyre lateral force of the asymmetrical car but the drop surpasses the rise. This causes the asymmetrical car a lateral force loss compared to the baseline vehicle, as shown in Fig. 11.11.

Figure 11.12 depicts how the car with CoG deviating toward the outer side of the turn diverges from the symmetrical car in terms of lateral grip. As can be observed, there is a decline, around 0.25 m/s^2, in the lateral acceleration of the asymmetrical

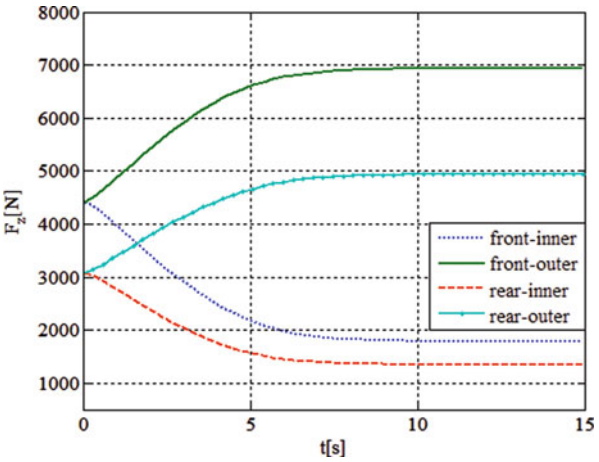


Fig. 11.9 The normal load under each tyre of a laterally symmetrical car

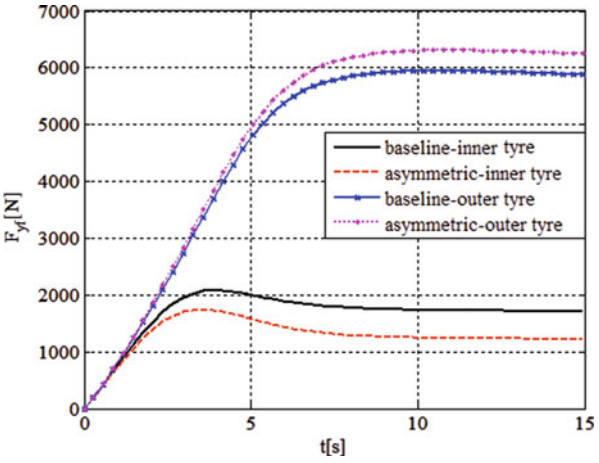


Fig. 11.10 The lateral forces under the front tyres: car with CoG deviating toward the outer side of the turn vs symmetric car

car in comparison with the baseline in the near-saturation and saturation region. This is attributed to the greater inner-outer normal load at each axle of the asymmetrical car.

While Figs. 11.7 and 11.12 show how much laterally asymmetrical cars, with CoG, respectively, deviating toward the inner and the outer sides of rotation center, can produce lateral grips in relation to the bilaterally symmetrical car, Fig. 11.13 compares the lateral accelerations of the two laterally asymmetrical cars. This is actually the comparison between a left turn and a right turn for one particular asymmetrical car; hence, it is of realistic importance. As can be seen from the graph, when a laterally asymmetrical car turns to heavier side it can produce more lateral grip than when it turns to the lighter side. Also, the difference is around 0.5 m/s^2 in the high acceleration region associating with the saturation of lateral force.

Figure 11.14 presents the difference—in percentage—between the lateral grips of that asymmetrical car for turn-to-heavier-side and turn-to-lighter-side cases. As

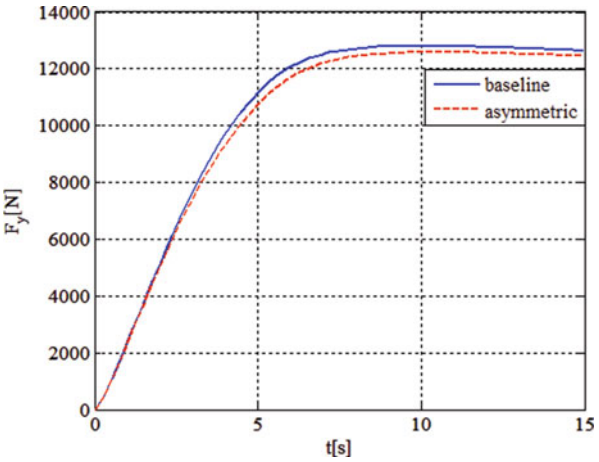


Fig. 11.11 The car resultant lateral force: car with CoG deviating toward the outer side of the turn vs symmetric car

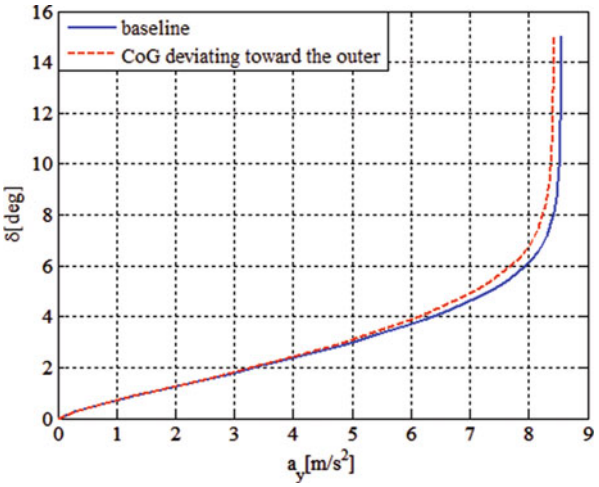


Fig. 11.12 The lateral acceleration: car with CoG deviating toward the outer side of the turn vs symmetric car

can be seen from the graph, the difference increases with the steering angle and reaches the maximum of around 7% in the near-saturation region. This means, if the car is laterally asymmetrical the driver should not treat the car equally when steering to the left and to the right sides, especially in high lateral acceleration region.

11.4 Conclusion

The lateral dynamics of a transversely asymmetrical car was examined in this paper. First, a planar model of steering car that can take the lateral asymmetry into account was constructed. The normal load under each tyre of the asymmetrical car which is an input of tyre lateral force was then determined; and, the unconventional kinematic

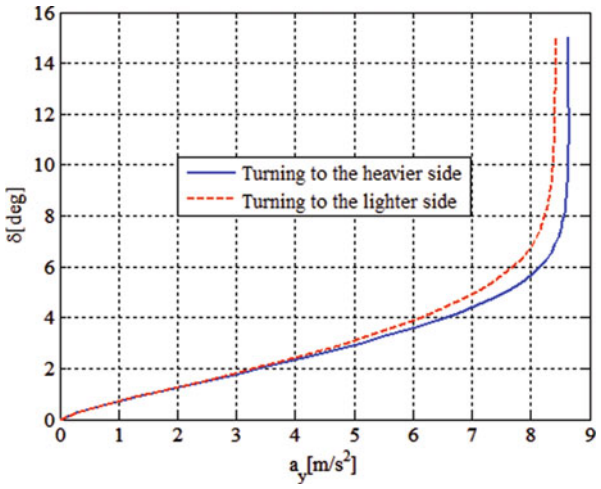


Fig. 11.13 The lateral accelerations of the asymmetrical car when cornering to two sides: heavier side and lighter side

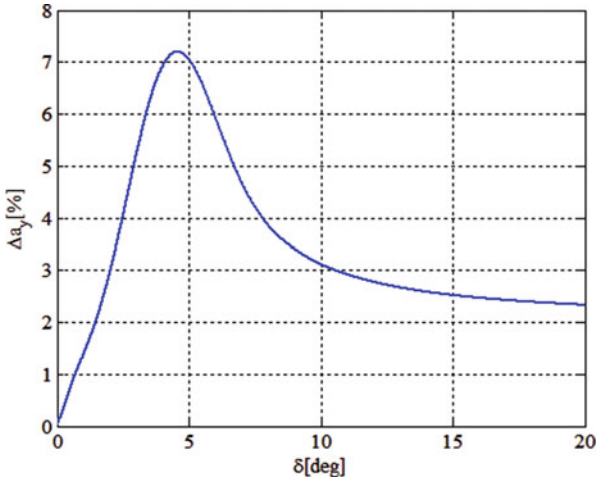


Fig. 11.14 The difference of lateral accelerations of the asymmetrical car in percentage when cornering to two sides: heavier side and lighter side

relation between the car's CoG and each tyre was derived from the asymmetrical model to form the differential equation system. An exemplary set of car parameters was chosen to investigate the effect of lateral position of car's CoG on the dynamics of the steering car. The results demonstrate that the car with CoG deviating the inner side of the turn can produce a few more percent of lateral grip than the symmetrical car. If the CoG is located toward the outer side of the turn, its grip capacity reduces by a few percent compared to the baseline car. The most important thing to note is that the lateral accelerations of the asymmetrical car turning to the heavier side can be 7% higher than that of the car turning to the lighter side. Drivers or autonomous systems being aware of this difference can save lives as it is near the saturation region where lateral grip cannot increase even turning steering wheel more.

11.5 Notations

a	Distance from the CoG of the car to the front wheels
b	Distance from the CoG of the car to the rear wheels
c	Distance from the CoG of the car to the left wheels
d	Distance from the CoG of the car to the right wheels
L	Wheel base
w	The car track
m	Total mass of the car
$h_g(m)$	Height of CoG of the car
I_z	Yaw inertia of the car
δ_{fI}	Steering angle of the inner front wheel
δ_{fO}	Steering angle of the outer front wheel
α_{fI}	Side slip angle of the inner front wheel
α_{fO}	Side slip angle of the outer front wheel
α_{rI}	Side slip angle of the inner rear wheel
α_{rO}	Side slip angle of the outer rear wheel
β	The angle between Cx and velocity vector of the car at CoG
β_{fI}	The angle between Cx and velocity vector of the car at inner front tyre
β_{fO}	The angle between Cx and velocity vector of the car at outer front tyre
β_{rI}	The angle between Cx and velocity vector of the car at inner rear tyre
β_{rO}	The angle between Cx and velocity vector of the car at outer rear tyre
F_{xfI}	The longitudinal force of front inner tyre
F_{yfI}	The lateral force of front inner tyre
F_{zfI}	The normal force of front inner tyre
M_{zfI}	The aligning moment of front inner tyre
F_{xfO}	The longitudinal force of front outer tyre
F_{yfO}	The lateral force of front outer tyre
F_{zfO}	The normal force of front outer tyre
M_{zfO}	The aligning moment of front outer tyre
F_{xrI}	The longitudinal force of rear inner tyre
F_{yrI}	The lateral force of rear inner tyre
F_{zrI}	The normal force of rear inner tyre
M_{zrI}	The aligning moment of rear inner tyre
F_{xrO}	The longitudinal force of rear outer tyre
F_{yrO}	The lateral force of rear outer tyre
F_{zrO}	The normal force of rear outer tyre
M_{zrO}	The aligning moment of rear outer tyre

References

1. Egbert Bakker, Lars Nyborg, and Hans B. Pacejka. Tyre modelling for use in vehicle dynamics studies. *SAE Technical Paper*, 1987. Paper number 870421.
2. Jon D. Demerly and Kamal Youcef-Toumi. *Non-Linear Analysis of Vehicle Dynamics (NAVDyn): A Reduced Order Model for Vehicle Handling Analysis*. SAE Paper 2000-01-1621, 2000.
3. T. D. Gillespie. *Fundamentals of vehicle dynamics*. Society of Automotive Engineers, Warrendale, PA, 1992.
4. R.N. Jazar. *Vehicle Dynamics*. Springer New York, 2017.
5. R.N. Jazar. *Vehicle Planar Dynamics*, pages 115–213. Springer International Publishing, Cham, 2019.
6. H. Marzbani, H. Khayyam, C. N. TO, . V. Quoc, and R. N. Jazar. Autonomous vehicles: Autodriver algorithm and vehicle dynamics. *IEEE Transactions on Vehicular Technology*, 68(4):3201–3211, 2019.
7. Ardashir Mohammadzadeh and Hamid Taghavifar. A novel adaptive control approach for path tracking control of autonomous vehicles subject to uncertain dynamics. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 2020.
8. Hari M Nair and C Sujatha. Prevention of vehicle rollover after wheel lift-off using energy-based controller with proportional gain augmentation. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 234(4):963–980, 2020.
9. Mohamed Salaani. *Development and validation of a vehicle model for the National Advanced Driving Simulator*. PhD thesis, The Ohio State University, 1996.
10. Dai Q Vo, Hormoz Marzbani, Mohammad Fard, and Reza N Jazar. Variable caster steering in vehicle dynamics. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 232(9):1270–1284, 2018.
11. F. Wang and Y. Chen. Vehicle rollover propensity detection based on a mass-center-position metric: A continuous and completed method. *IEEE Transactions on Vehicular Technology*, 68(9):8652–8662, 2019.
12. Jo Yung Wong. *Theory of ground vehicles*. John Wiley & Sons, 2001.

Chapter 12
Roll Model Control of Autonomous Vehicle

Catter Ching Nok To, Hormoz Marzbani, and Reza N. Jazar

Abstract The dynamic vehicle roll study requires incorporating ideas from both the physics and mathematics. Applying artificial intelligence (AI) driver algorithms in autonomously controlled vehicles enables them to determine and negotiate corners effectively. The relational concepts vary when comparing the value of vehicle turn angles and their combined active contribution in turning the vehicle automatically.

The planar mathematical theory model for autonomous vehicles was initially developed. This theory was developed for use with 4-wheel steering vehicles but also works for 2-wheel steering vehicles. This theory extends the roll model by using the parameters of the angular velocity of a vehicle; that is, roll φ, pitch θ, yaw ψ, roll rate p, pitch rate q and yaw rate r. However, a roll model that uses forward, lateral, yaw and roll velocities is more exact and effective compared to this planar model.

Autodriver algorithm was introduced as a path-following algorithm for autonomous vehicles which is using road geometry data and planar vehicle dynamics. An autonomous vehicle can follow a given road if it turns about its centre of curvature at a correct moving position equal to the radius of the curvature of the path. The autodriver algorithm is improved according to practical implications, while a more realistic vehicle model (roll mode) is used, which considers roll degree of freedom in addition to a planar motion. A ghost-car path-following approach is introduced to define the desired location of the car at every instance. Finally, simulations are performed to analyse the path-following performance of the proposed scheme. The results show promising performance of the controller both in terms of error minimisation and passenger comfort.

C. C. N. To (✉)
Vocational Training Council (VTC), Tsing Yi Island, New Territories, Hong Kong
e-mail: catter@vtc.edu.hk

H. Marzbani · R. N. Jazar
School of Engineering RMIT University, Melbourne, VIC, Australia

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_12

12.1 Introduction

A given road can be expressed mathematically in a global (or world) coordinate frame. Following the road can be substituted by following the loci of its curvature centre and turning at the right circle of curvature. Considering that a vehicle in motion is always in turn about an instantaneous rotation centre relative to the ground, an autonomous vehicle capable of following a given path by coinciding the rotation centre of the vehicle at every moment on the curvature centre of the road could be designed. The dynamic reactions of the vehicle influence its path of motion and make its rotation centre depart from the desired path of the curvature centre of the road. In this study, the autodriver algorithm control strategy for front-wheel steering vehicles has been developed, and a control loop is introduced to compensate for the present errors generated by the difference between the real position of the vehicle and its desired location on the road.

A full study on the autodriver algorithms on the ground vehicle emerge and pursue in the area of improvement of vehicle roll dynamics by mathematical theory and make it applicable to a practical model. On the other hand, the study is to design and implement a control system which makes the autodriver theoretical roll model applicable to a real vehicle.

Implementing a feedback PID control to eliminate the autodriver algorithm errors has been expanded in 2018. The algorithm is based on a well-defined road expressed mathematically in a global coordinate frame space such that the curvature centre of the curve is known at any point, aimed to follow the loci of its curvature centre and to turn at the right circle of curvature. The steady-state and transient responses of turning vehicles and the resulted path of motion of the vehicle have been investigated. The effect of acceleration can be changed by the forward velocity. By comparing the two responses of the vehicle on the roll model, the result could be used to prove that the difference between the steady-state and transient centre of rotation of the vehicle in engineering applications. The engineering approximations are well presented and proved of the dynamic vehicle theory using a parameter of steering angle and variable forwarding velocity. It has been shown that the vehicles using their steady-state responses within an acceptable range. More specifically, the dynamic rotation centre of vehicles is determined and compares them with steady-state values. The result would be essential to design the autodriver algorithm for autonomous vehicles. Moreover, a roll model can be more accurate to predict the actual phenomenon of the vehicle [1–10].

In the context of autonomous vehicles, the roll angle constituents assume significant and critical importance, and such vehicles are well governed by linear/bicycle equations of motion. Roll movement, before pitch, yaw and bounce, is the most unpleasant movement for the passengers. This chapter objective is a minimization of roll motion and its fluctuation by applying vehicle roll dynamics application.

12.2 Roll Model Equation of Motion

Consider a rigid vehicle following a bicycle model with a dynamic equation that possesses forward, roll and yaw motion (Fig. 12.1). The model can roll and is related to the bicycle model. Through such a roll model, the roll behaviour of a vehicle can be analysed.

To relate the position and orientation of a vehicle, a principal body coordinate frame B will be attached to the mass centre of the vehicle. The vehicle body will be positioned with respect to a globally fixed coordinate frame G (Fig. 12.2). The equation of motion for a rigid vehicle possessing yaw and roll motion has a movement of four-degrees-of-freedom, located at the x and y axes and the rotation of the x and z axes:

$$F_x = m\dot{v}_x - mrv_y \tag{12.1}$$

$$F_y = m\dot{v}_y + mrv_x \tag{12.2}$$

$$M_z = I_z\dot{\omega}_z = I_z\dot{r} \tag{12.3}$$

$$M_x = I_x\dot{\omega}_x = I_x\dot{p} \tag{12.4}$$

The rigid body equations of motion in the body coordinate frame are:

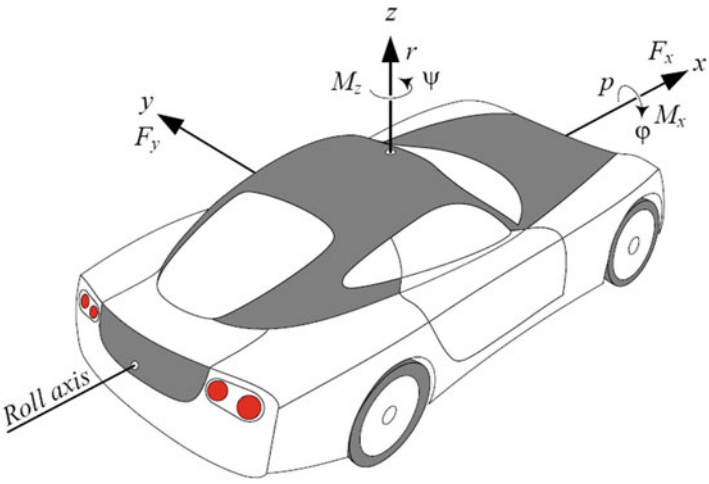


Fig. 12.1 The degree-of-freedom of a roll model of a rigid vehicle

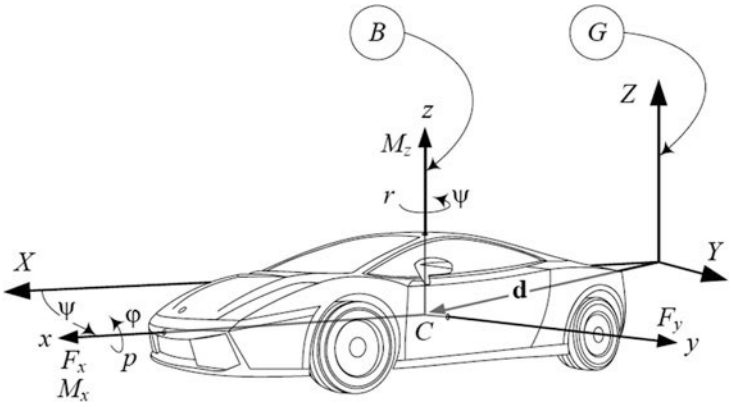


Fig. 12.2 A vehicle with roll and yaw rotation

$$\begin{aligned} {}^{B}F &= {}^{B}R_{G}\,{}^{G}F = {}^{B}R_{G}\left(m^{G}a_{B}\right) = m_{G}{}^{B}a_{B} \\ &= m^{B}\dot{v}_{B} + m_{G}{}^{B}\omega_{B} *{}^{B} v_{B} \end{aligned} \tag{12.5}$$

$$\begin{aligned} {}^{B}M &= \tfrac{{}^{G}d}{dt}{}^{B}L = {}_{G}{}^{B}\dot{L}_{B} = {}^{B}\dot{L} + {}_{G}{}^{B}\omega_{B} *{}^{B} L \\ &= {}^{B}I_{G}{}^{B}\dot{\omega}_{B} + {}_{G}{}^{B}\omega_{B} * \left({}^{B}I_{G}{}^{B}\omega_{B}\right) \end{aligned} \tag{12.6}$$

By substituting the vector and matrices in the equation of motion:

$${}^{B}F = m^{B}\dot{v}_{B} + m_{G}{}^{B}\omega_{B} * {}^{B}v_{B}$$

$$\begin{bmatrix} F_x \\ F_y \\ 0 \end{bmatrix} = m \begin{bmatrix} \dot{v}_x \\ \dot{v}_y \\ 0 \end{bmatrix} + m \begin{bmatrix} \omega_x \\ 0 \\ \omega_z \end{bmatrix} * \begin{bmatrix} v_x \\ v_y \\ 0 \end{bmatrix} = \begin{bmatrix} m\dot{v}_x - m\omega_z v_y \\ m\dot{v}_y + m\omega_z v_x \\ m\omega_x v_y \end{bmatrix} \tag{12.7}$$

$$\begin{aligned} {}^{B}M &= {}^{B}I_{G}^{B}\omega_{B} + {}_{G}{}^{B}\omega_{B} * \left({}^{B}I_{G}{}^{B}\omega_{B}\right) \\ \begin{bmatrix} M_x \\ 0 \\ M_z \end{bmatrix} &= \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \begin{bmatrix} \dot{\omega}_x \\ 0 \\ \dot{\omega}_z \end{bmatrix} + \begin{bmatrix} \omega_x \\ 0 \\ \omega_z \end{bmatrix} * \left(\begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \begin{bmatrix} \omega_x \\ 0 \\ \omega_z \end{bmatrix} \right) \\ &= \begin{bmatrix} I_1\dot{\omega}_x \\ I_1\omega_x\omega_z - I_3\omega_x\omega_z \\ I_3\dot{\omega}_z \end{bmatrix} \end{aligned} \tag{12.8}$$

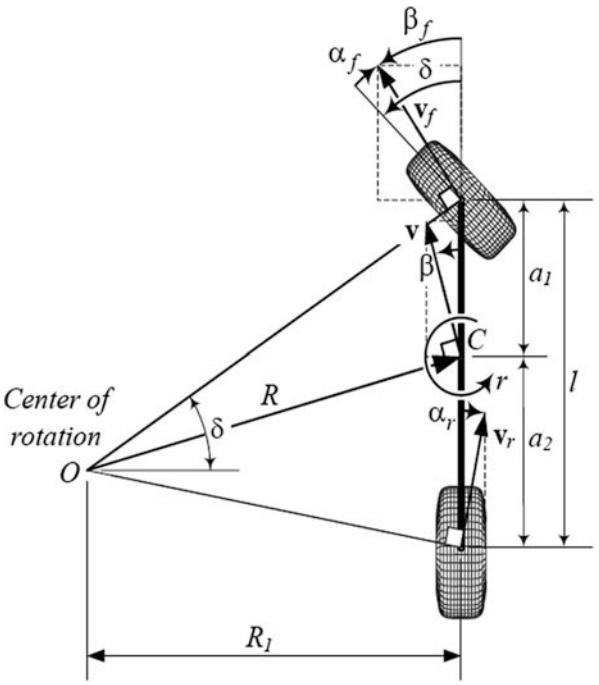


Fig. 12.3 Bicycle model for a positive front-wheel steering vehicle

The equations for a roll model of bicycle vehicle model (Fig. 12.3) can be finalised as:

$$\begin{bmatrix} F_x \\ F_y \\ 0 \end{bmatrix} = m \begin{bmatrix} \dot{v}_x - \omega_z v_y \\ \dot{v}_y + \omega_z v_x \\ \omega_x v_y \end{bmatrix} \tag{12.9}$$

$$\begin{bmatrix} M_x \\ 0 \\ M_z \end{bmatrix} = \begin{bmatrix} \dot{\omega}_x I_1 \\ \omega_x \omega_z I_1 - \omega_x \omega_z I_3 \\ \dot{\omega}_z I_3 \end{bmatrix} \tag{12.10}$$

where:

$$v_z = 0$$

$$\dot{v}_z = 0$$

$$\omega_z = 0$$

$$\dot{\omega}_z = 0$$

$$F_z = 0$$

$$M_y = 0$$

The force system exerted on a rigid vehicle in the roll model is:

$$\begin{aligned} F_y &= F_y(r, p, \beta, \varphi, \delta) \\ &= \tfrac{\partial F_y}{\partial r} r + \tfrac{\partial F_y}{\partial p} p + \tfrac{\partial F_y}{\partial \beta} \beta + \tfrac{\partial F_y}{\partial \varphi} \varphi + \tfrac{\partial F_y}{\partial \delta} \delta \\ &= C_r r + C_p p + C_\beta \beta + C_\varphi \varphi + C_\delta \delta \end{aligned} \tag{12.11}$$

$$\begin{aligned} M_x &= M_x(r, p, \beta, \varphi, \delta) \\ &= \tfrac{\partial M_x}{\partial r} r + \tfrac{\partial M_x}{\partial p} p + \tfrac{\partial M_x}{\partial \beta} \beta + \tfrac{\partial M_x}{\partial \varphi} \varphi + \tfrac{\partial M_x}{\partial \delta} \delta \\ &= E_r r + E_p p + E_\beta \beta + E_\varphi \varphi + E_\delta \delta \end{aligned} \tag{12.12}$$

$$\begin{aligned} M_z &= M_z(r, p, \beta, \varphi, \delta) \\ &= \tfrac{\partial M_z}{\partial r} r + \tfrac{\partial M_z}{\partial p} p + \tfrac{\partial M_z}{\partial \beta} \beta + \tfrac{\partial M_z}{\partial \varphi} \varphi + \tfrac{\partial M_z}{\partial \delta} \delta \\ &= D_r r + D_p p + D_\beta \beta + D_\varphi \varphi + D_\delta \delta \end{aligned} \tag{12.13}$$

The above equations depend on five parameters: β, δ, γ, p and φ.
The force system coefficients are:

$$C_r = \frac{\partial F_y}{\partial r} = -\frac{a_1}{v_x} C_{af} + \frac{a_2}{v_x} C_{ar}$$

$$C_p = \frac{\partial F_y}{\partial p} = \frac{C_{af} C_{\beta f}}{v_x} + \frac{C_{ar} C_{\beta r}}{v_x}$$

$$C_\beta = \frac{\partial F_y}{\partial \beta} = -\left(C_{af} + C_{ar}\right)$$

$$C_\varphi = \frac{\partial F_y}{\partial \varphi} = C_{ar} C_{\delta\varphi r} + C_{af} C_{\delta\varphi f} - C_{\varphi f} - C_{\varphi r}$$

$$C_\delta = \frac{\partial F_y}{\partial \delta} = C_{af}$$

$$D_r = \frac{\partial M_z}{\partial r} = -\frac{a_1^2}{v_x} C_{af} - \frac{a_2^2}{v_x} C_{ar}$$

$$D_p = \frac{\partial M_z}{\partial p} = \frac{a_1}{v_x} C_{\beta f} C_{af} - \frac{a_2}{v_x} C_{\beta r} C_{ar}$$

$$D_\beta = \frac{\partial M_z}{\partial \beta} = -\left(a_1 C_{af} - a_2 C_{ar}\right)$$

$$D_\varphi = \frac{\partial M_z}{\partial \varphi} = -a_1\left(C_{\varphi f} - C_{af} C_{\delta\varphi f}\right) + a_2\left(C_{\varphi r} - C_{ar} C_{\delta\varphi r}\right)$$

$$D_\delta = \frac{\partial M_z}{\partial \delta} = a_1 C_{af}$$

$$E_r = \frac{\partial M_z}{\partial r} = -\frac{a_1}{v_z} C_{Tf} C_{af} + \frac{a_2}{v_z} C_{Tr} C_{ar}$$

$$E_p = \frac{\partial M_z}{\partial p} = \frac{1}{v_z} C_{\beta f} C_{Tf} C_{af} + \frac{1}{v_z} C_{\beta r} C_{Tr} C_{ar} - c_\varphi$$

$$E_\beta = \frac{\partial M_z}{\partial \beta} = -C_{Tf} C_{af} - C_{Tr} C_{ar}$$

$$E_\varphi = \frac{\partial M_z}{\partial \varphi} = -C_{Tf}\left(C_{\varphi f} - C_{af} C_{\delta\varphi f}\right) - k_\varphi - C_{Tr}\left(C_{\varphi r} - C_{ar} C_{\delta\varphi r}\right)$$

$$E_\delta = \frac{\partial M_z}{\partial \delta} = C_{Tf} C_{af}$$

Employing:

$$\beta = \frac{v_y}{v_x}$$

The equation is then transformed to:

$$\dot{v}_x = \frac{F_x}{m} + r v_y \tag{12.14}$$

$$\dot{v}_y = \left(\frac{C_r}{m} - v_x\right) r + \frac{C_p}{m} p + \frac{C_\beta}{m}\frac{v_y}{v_x} + \frac{C_\varphi}{m}\varphi + \frac{C_\delta}{m}\delta \tag{12.15}$$

$$\dot{p} = \frac{1}{I_x}\left(E_r r + E_p p + E_\beta \frac{v_y}{v_x} + E_\varphi \varphi + E_\delta \delta\right) \tag{12.16}$$

$$\dot{r} = \frac{1}{I_z}\left(D_r r + D_p p + D_\beta \frac{v_y}{v_x} + D_\varphi \varphi + D_\delta \delta\right. \tag{12.17}$$

12.3 Steady-State Motion for Roll Model

The steady-state behaviour of a vehicle achieved when the steering angle and forward velocity of the vehicle remain constant long enough. The output variables such as lateral velocity and acceleration, yaw rate, etc. approach a constant and steady-state value. The output steady-state as functions of inputs is the reference for transient analysis of the vehicle dynamics. Such functions are cast in several dimensionless parameters to be applied on all vehicles. The negotiation in an undefined direction of a front-wheel steering, two-wheeled, rollable rigid vehicle at steady-state is governed by the following [11–13]:

$$F_x = -mrv_y \tag{12.18}$$

$$C_r r + C_p p + C_\beta \beta + C_\varphi \varphi + C_\delta \delta = mrv_x \tag{12.19}$$

$$E_r r + E_p p + E_\beta \beta + E_\varphi \varphi + E_\delta \delta = 0 \tag{12.20}$$

$$D_r r + D_p p + D_\beta \beta + D_\varphi \varphi + D_\delta \delta = 0 \tag{12.21}$$

$$F_x = -\frac{m}{R} v_x v_y \tag{12.22}$$

$$\left(C_r v_x - m v_x^2\right) \frac{1}{R} + C_\beta \beta + C_p p + C_\varphi \varphi = -C_\delta \delta \tag{12.23}$$

$$E_r v_x \frac{1}{R} + E_\beta \beta + E_p p + E_\varphi \varphi = -E_\delta \delta \tag{12.24}$$

$$D_r v_x \frac{1}{R} + D_\beta \beta + D_p p + D_\varphi \varphi = -D_\delta \delta \tag{12.25}$$

$$\begin{aligned} Z_0 &= E_\beta \left(D_r C_\varphi - C_r D_\varphi + m v_x D_\varphi\right) + \\ &E_\varphi \left(C_r D_\beta - D_r C_\beta - m v_x D_\beta\right) + E_r \left(C_\beta D_\varphi - D_\beta C_\varphi\right) \\ Z_1 &= E_\beta \left(C_\varphi D_\delta - v_x C_\delta D_\varphi\right) - E_\varphi \left(C_\beta D_\delta - v_x C_\delta D_\beta\right) \\ &+ E_r \left(C_\beta D_\varphi - D_\beta C_\varphi\right) \end{aligned} \tag{12.26}$$

$$\begin{aligned} Z_2 = E_\varphi \left(m v_x D_\delta - C_r D_\delta + D_r v_x C_\delta\right) + E_r \left(C_\varphi D_\delta - v_x C_\delta D_\varphi\right) \\ - E_\delta \left(D_r C_\varphi - C_r D_\varphi + m v_x D_\varphi\right) \end{aligned} \tag{12.27}$$

$$Z_3 = E_\beta \left(m v_x D_\delta - C_r D_\delta + D_r v_x C_\delta\right) + E_r \left(C_\beta D_\delta - v_x C_\delta D_\beta\right) - E_\delta \left(D_r C_\beta - C_r D_\beta + m v_x D_\beta\right) \tag{12.28}$$

1. Curvature response,

$$S_k = \frac{k}{\delta} = \frac{1}{R\delta} = -\frac{Z_1}{v_x Z_0} \tag{12.29}$$

2. Sideslip response,

$$S_\beta = \frac{\beta}{\delta} = \frac{Z_2}{Z_0} \tag{12.30}$$

3. Yaw amount reaction,

$$S_r = \frac{r}{\delta} = \frac{k}{\delta} v_x = S_k v_x = -\frac{Z_1}{Z_0} \tag{12.31}$$

4. Lateral acceleration response,

$$S_a = \frac{v_x^2/R}{\delta} = \frac{k}{\delta} v_x^2 = S_k v_x^2 = \frac{v_x Z_1}{Z_0} \tag{12.32}$$

5. Roll angle response,

$$S_\varphi = \frac{\varphi}{\delta} = -\frac{Z_3}{Z_0} \tag{12.33}$$

In steady-state motion conditions, all the variables are constant, and hence the derivatives are zero. By using the steady-state motion of a roll model, the curvature response, S_k, sideslip response, S_β, and Yaw rate response, S_r demonstrated to show the change as a function of forwarding velocity v_x in Figs. 12.4, 12.5 and 12.6. The equation aimed to smoothen the curve as an end result, including lateral acceleration response and roll angle response.

12.4 Control of Autonomous Vehicles, Autodriver Algorithm and Vehicle Dynamics

Autonomous vehicles have been a major focus of research in automotive engineering within recent decades. One of the enormous technical challenges of autonomous vehicles is a lateral controller for designing dynamic path tracking as a key component of the control system. A development method to achieve lateral control

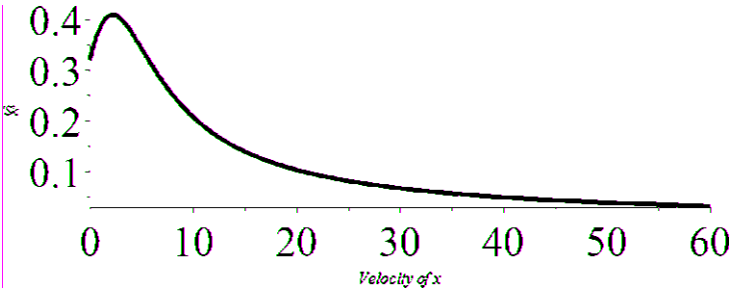


Fig. 12.4 Curvature response, S_k, as a function of forwarding velocity v_x

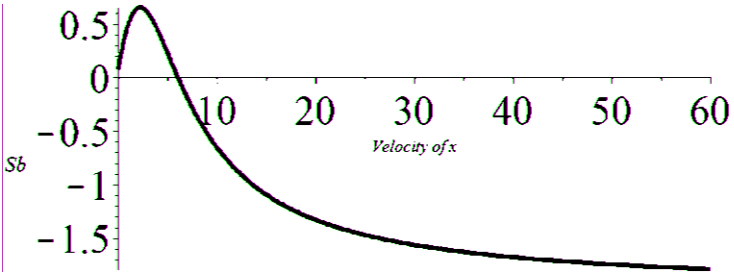


Fig. 12.5 Sideslip response, S_β, as a function of forwarding velocity v_x

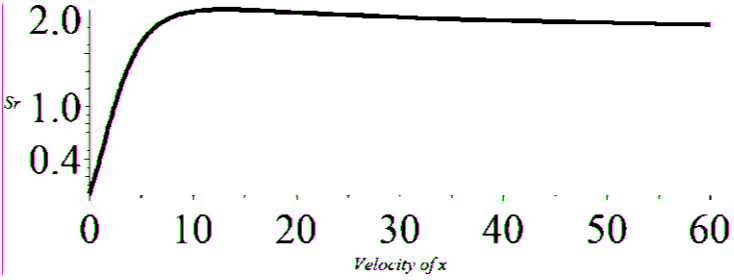


Fig. 12.6 Yaw rate response, S_r, as a function of forwarding velocity v_x

of autonomous vehicles is a steering control system. The output of most of the controllers for lateral direction reported previously is the determination of the steering angle, which is based on path planning approaches and by sensing state parameters of the vehicle. There are many mechanisms by which the optimal trajectories can be determined for steering mechanisms, using cost function optimisation techniques. For steering control execution, the cost map method is used to calculate the optimal path. Cost maps enable fusing of the information collected using several sensors on the vehicle, allowing for a more universal and highly accurate assessment of the reliability of the generated path. However, a control method of autonomous vehicles which are based on steering angle results in inaccurate tracking concerning

the position on the desired path, because as a vehicle moves in real time, the steering wheel direction changes. A circular path with constant curvature is followed when the steering angle is fixed. In the case of varied steering angle, however, there will be a change in curvature. In this case, using a clothoid geometry function, the vehicle motion can be estimated. The curve used to transition from a straight road to a circular road or vice versa is called the clothoid, or Euler spiral.

To avoid the challenges mentioned earlier, as an alternative steering angle-based control, use a steering rate control approach. It is to have a driver steering model which analyses vehicle test data subject to different driver steering scenarios during a standard double-lane-change movement by capturing key driver steering mechanisms. In this approach, drivers control the steering rate in an amount to their error for the target angle which suggests the implementation of a steering rate-based control, instead of the conventional steering angle control. Bae et al. proposed a steering rate-based controller in which at the planning stage, the curvature of the desired trajectory was determined before calculating the time derivative of the steering angle required and steering rate for accurate control. A linear "bicycle model" was used to model the lateral kinematics of the vehicle using the steering angle of both the front wheels. Various other research efforts in this area have enhanced the design of roads intended to be used by autonomous vehicles.

Recent research has proposed to construct the intelligent transportation systems, variable smarter suspensions, steering systems, torque distribution, steering by wire, and vehicle dynamic modelling improvement on designing safer, more intelligent vehicles. As an example, a combination of GPS, sensory systems, mathematical and smart algorithms show to be the ultimate answer to fully automated and driverless vehicles. Besides the technical aspects of autonomous vehicles, a variety of studies are conducted about social impacts, regulations, human-machine interfaces and implementation methods of autonomous vehicles. Most of the algorithms introduced for autonomous control of a vehicle have been introduced, rely on vision systems and sensory equipment in general and are quite functional for searching in unknown, structured environments. There exist other systems designed to guide a vehicle from a starting point or posture to a destination with a known posture by the use of smart decision-making algorithms in conjunction with some obstacle avoidance strategy. Controlling the vehicle's position using GPS is also an applicable alternative method for some applications. The most suitable or easy to use methods for road vehicles are based on paths of motion which are previously mapped and planned. GPS can then be used to determine the best path to be taken by the vehicle. However, roads are known, and they can be well defined using mathematical equations in a coordinate frame attached to the ground called the global frame. Curvature centre and a curvature radius of such roads, as well as all the other geometrical characteristics, can be determined and employed for designing the required control system. A novel algorithm called autodriver was introduced in 2010 which presented the theory mathematically and is used as the proof of concept for the current investigation. The theory is also adopted, developed farther and applied by other investigators. The original autodriver algorithm manuscript introduces the concept of replacement of given roads with the path of their curvature centres. Any moving vehicle on the

road is continuously considered to be in turn, near the curvature centre of the road at any instant in time, which results in three points of importance: (1) road curvature centre, or in other words the geometrical information of the road independent of the vehicles; (2) the rotation centre found kinematical which is the intersection of the perpendicular line to the wheels of the vehicle; the kinematic rotation centre is ideally a single point about which the vehicle tends to turn at no speed; and (3) dynamic turning point or the real rotation centre for the vehicle in motion. The location of the dynamic centre of rotation is a dynamic properties function of the vehicle such as the position of mass centre, mass moments, tire-road interaction coefficients, steer angle and vehicle velocity. The autodriver algorithm is coinciding the dynamic centre of rotation and the road curvature centre at all times during the motion by adjusting the steer angle and velocity of the vehicle.

12.4.1 Road Geometry

The three main factors include driver, environment and vehicle (DEV) and are involved in energy management and vehicle performance. The modelling approach has become a vital tool for automotive engineers and mechanical researchers to reduce time consumed and improve the efficiency of the vehicle design. The modelling results have environmental benefits as well as significant cost saving. Among the factors that are involved in the vehicle system, the environmental conditions, and road geometry are often unknown and uncertain during driving.

12.4.1.1 Horizontal Curve

A horizontal road can be assumed to be several straight lines and circular curves. Transition curve is the name chosen for the curves smoothly attaching straight lines to circles on the road. A curve of sufficient length is usually used whenever there is a change of direction in a road alignment in order to avoid the appearance of a sudden change in direction for a vehicle travelling on the road.

The two significant reasons for vehicle instability on the road are:

- Sliding: When a vehicle is travelling around a curve, lateral friction directly related to the square of the vehicle's speed is developed at the tire-road interface. The force required maintaining a circular path exceeds the force which can be developed by friction and superelevation, as speed is increased.
- Overturning: Often a heavy vehicle issue or in other words vehicles with a high centre of gravity. An overturning moment is formed by the forces acting on the vehicle, which then cause rollovers for a vehicle in turn (State highway geometric design manual-Section 4: Horizontal alignment).

12.4.1.2 Road Curvature Modelling

In order to produce a credible vehicle on the road simulation, realistic environmental modelling is necessary. The stochastic models proposed are used for creating artificial environmental conditions for the present study.

12.4.1.3 Road Curvature Centre

Consider a road to be expressed as a parametric curve in 3 dimensions:

$$x = x(s), y = y(s), z = z(s) \tag{12.34}$$

where s is the distance travelled on the road measured from a point fixed on the road chosen to be the initial location of the vehicle.

$$x_0 = x(0), y_0 = y(0), z_0 = z(0) \tag{12.35}$$

Three principal planes exist for any point on the road: the osculating plane, the perpendicular to the road plane and the rectifying plane defined, respectively, as:

$$(x - x_0)\frac{dx}{ds} + (y - y_0)\frac{dy}{ds} + (z - z_0)\frac{dz}{ds} = 0 \tag{12.36}$$

$$\left(\frac{dy}{ds}\frac{d^2z}{ds^2} - \frac{dz}{ds}\frac{d^2y}{ds^2}\right)(x - x_0) + \left(\frac{dz}{ds}\frac{d^2x}{ds^2} - \frac{dx}{ds}\frac{d^2z}{ds^2}\right)(y - y_0) + \left(\frac{dx}{ds}\frac{d^2y}{ds^2} - \frac{dy}{ds}\frac{d^2x}{ds^2}\right)(z - z_0) = 0 \tag{12.37}$$

$$(x - x_0)\frac{d^2x}{ds^2} + (y - y_0)\frac{d^2y}{ds^2} + (z - z_0)\frac{d^2z}{ds^2} = 0 \tag{12.38}$$

The osculating plane includes the tangent line and point *P,* which is the centre of curvature of the road. The rectifying plane is vertical concerning both osculating and regular planes. The tangential unit vector indicates the perpendicular plane $\hat{u_t} = d_r/d_s$, bi-vector identifies the osculating plane $\hat{u_b} = \hat{u_t} \times \hat{u_n}$, and the rectifying plane is identified by the normal unit vector $\hat{u_n} = d^2r/ds^2$. Figure 12.7 depicts the road centre, radius of road curvature and the three plans at a point P of a road.

Curvature κ of a road at point *P* which has a curvature radius ρ can be calculated by

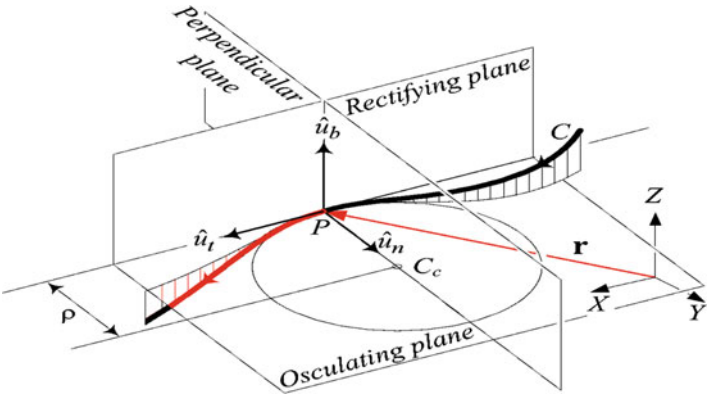


Fig. 12.7 Osculating plane, curvature centre and road's radius of curvature at point P

$$k = \frac{1}{R_k} = \left|\frac{d^2r}{ds^2}\right| = \frac{|v \times a|}{|v|^2} = \sqrt{\left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2 + \left(\frac{d^2z}{ds^2}\right)^2} \tag{12.39}$$

where **v**, **a and r** are the vehicle's velocity, acceleration and position vectors, respectively. A vector can show the position of the centre of curvature of a space curve $r_c = R_k \hat{u_n}$.

These equations are significantly less complicated in the case of a planar road model. If a road with a given equation $Y=f(X)$ is considered to exist in a global coordinate frame G, the radius of curvature R_κ of such a road at any point X of the road is:

$$R_k = \frac{\left(1 + Y'^2\right)^{3/2}}{Y'}, Y' = \frac{dY}{dX}, Y'' = \frac{d^2Y}{dX^2} \tag{12.40}$$

Figure 12.8a illustrates a road constructed by stochastic models of road geometry method used for this study, and Fig. 12.8b shows an example of a two-dimensional road ($y=x^3$) and its centre of the road. The arrows indicate the path of travelling along the road and its associated path of road curvature centre. Figure 12.9 also depicts a three-dimensional road with the following parametric equation.

$$x = (a + b\sin\theta)\cos\theta \tag{12.41}$$

$$y = (a + b\sin\theta)\sin\theta \tag{12.42}$$

$$z = a + b\cos\theta \tag{12.43}$$

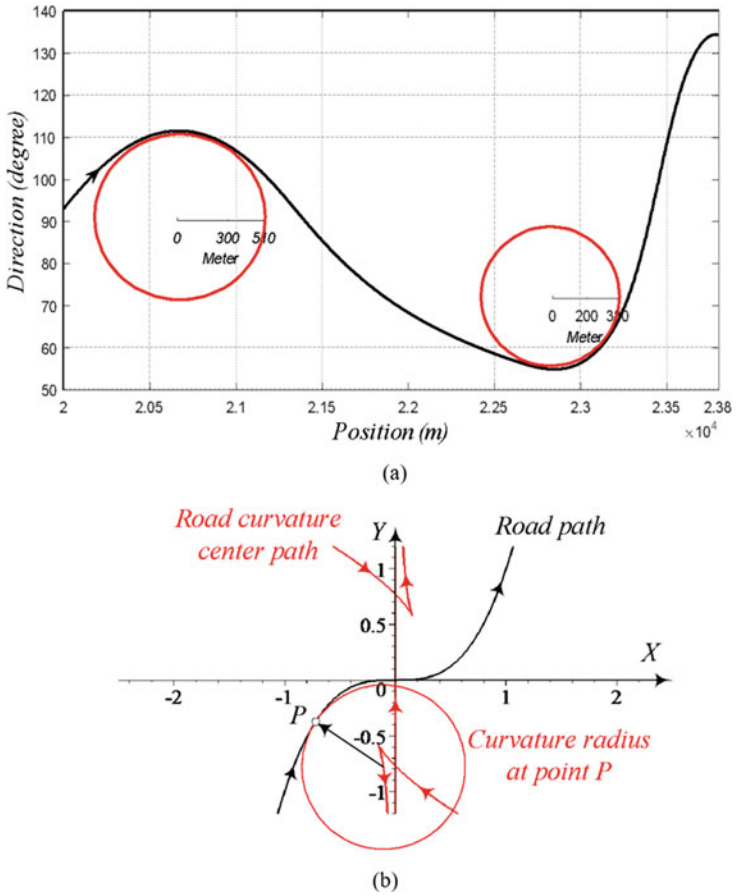


Fig. 12.8 (**a**) Constructed road direction by stochastic models of road geometry method, (**b**) a two-dimensional road and its road curvature centre paths

$$a = 250m, \qquad b = 200m$$

where x and y are the coordinates of the location of the road presented in a parametric format, a and b are magnification factors and θ is the heading angle of the road.

Substituting a road with its centre of curvature helps road following by making the vehicle to always turn on the instantaneous curvature circle about the instantaneous curvature centre. The straight motion of a vehicle is equal to a turn, about a point at infinite distance. This way, the vehicle will follow the given path if it turns about the road centre at the correct distance of the radius of curvature at any time.

The present study has been applied to many of the roads mentioned above and geometries. It was done to test the viability of the methods in different scenarios. The

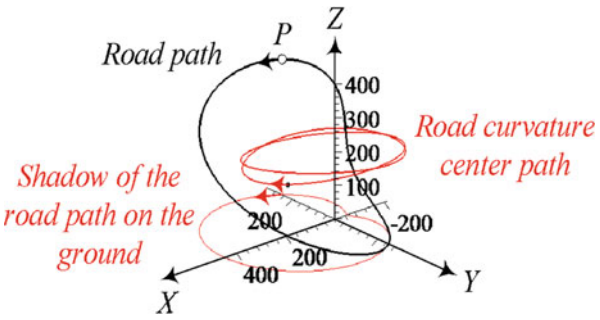


Fig. 12.9 A three-dimensional closed road and its curvature centre path

presented model in this document has been chosen to be a lane change manoeuvre, though. It has been done to provide the readers with the solutions to one of the main activities required by an autonomous vehicle at the moment [15–17].

12.4.2 Kinematic Analysis and the Dynamic Vehicle Rotation Centre

The kinematic centre of rotation for a front-wheel steering car is on a line perpendicular to the rear wheels – in line with the extended line of the rear axle. The kinematic rotation centre is at the intersection of lines which are perpendicular to the wheels, which is an ideal case will intersect at a point at all times. The rotation kinematic centre is theoretically the point that vehicle tends to turn, but this will happen only at low velocities close to zero. When a front-wheel steering vehicle is cruising at low speeds, the kinematic condition between its inner and outer wheels allowing them to turn without slipping on the ground (slip-free) is expressed by Ackerman condition.

$$\cot \delta_0 - \cot \delta_i = \frac{w}{l} \tag{12.44}$$

where δ_i and δ_o are the steer angles of the inner and outer wheels, respectively. Track (*w)* and wheelbase (*l)* are considered and also known as kinematic width and length of a vehicle [18]. The kinematic radius of rotation of such a vehicle, R, is the distance from the vehicle's centre of mass C and the kinematic centre.

$$R = \sqrt{a_2^2 + l^2 \cot^2 \delta} \tag{12.45}$$

where a_2 is the longitudinal distance between the rear axle and point *C* (location of the centre of gravity) and δ is the cot-average of the inner and outer steer angles.

$$\cot \delta = \frac{\cot \delta_0 + \cot \delta_i}{2} \tag{12.46}$$

The angle δ is the equivalent steer angle which is used for any vehicle's bicycle model having a wheelbase l and radius of rotation R. Substituting an equivalent bicycle model of a vehicle is a standard and popular method in vehicle dynamics in order to significantly simplify the resulted mathematical equations.

As soon as the vehicle moves, some side slip happens that makes the reallocation of the vehicle to deviate from the kinematic path of motion. The amount of deviation depends on the dynamic and geometric characteristics of the vehicle, steer angles, speed and tire-road interaction properties. Inputting a constant value for velocity and steer angles of the vehicle would result in a steady-state circular path of motion. The centre of such a steady-state circle is the actual or rotation dynamic centre. To visualise these phenomena, consider a vehicle that is supposed to turn around the origin of a global coordinate frame $G(X, Y)$ on a circle with radius $R=100$ m. A vehicle starts at a location on the X-axis while its local longitudinal x-axis is at $Y=100$ m parallel to the Y-axis and its rear axle is on the X-axis. At the starting instant, the kinematic rotation centre of the vehicle is set to be on the origin. Assume a non-zero constant velocity, and the vehicle will slip out laterally such that a larger circle will be tracked about its dynamic centre, which is illustrated in Fig. 12.10. The mathematical analysis of the kinematic rotation centre of four-wheel steering vehicles provides a set of similar equations involving more parameters [19–22].

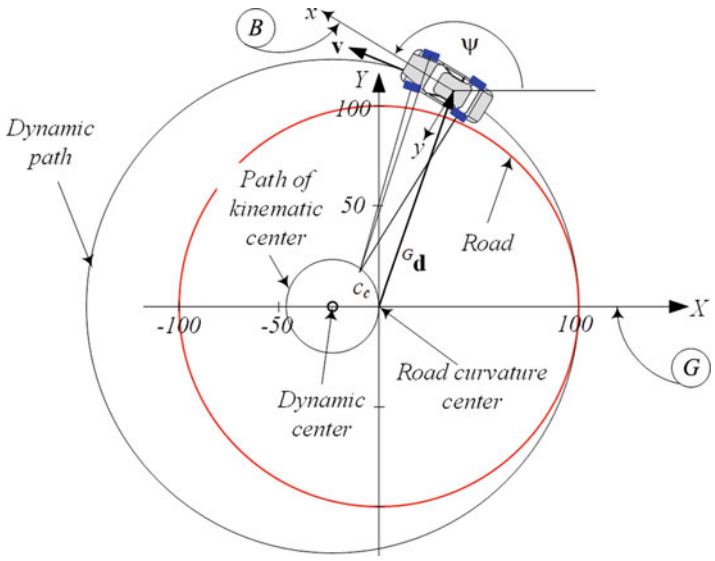


Fig. 12.10 The dynamics of an understeer vehicle makes it go out of the static, designed path. The dynamic curvature centre deviates from the road curvature centre when steering angles kept at the required static values to turn about road curvature (not in scale)

12.4.3 Vehicle Dynamics (High-Velocity Manoeuvres)

Consider a vehicle moving on a plane surface (planar motion). While a global coordinate frame G is attached to the ground, also attached to the mass centre of the vehicle is another local coordinate frame B. The Z- and z-axes are always parallel. The heading angle ψ, which is the angle between the x- and X-axes, indicates the orientation of frame B. ${}^{G}\mathbf{d}$ denotes the global position vector of the mass centre of the vehicle (Fig. 12.11).

A planar vehicle model has 3 degrees of freedom, which are the following:

1- Translation in the x
2- Translation in the y directions
3- Rotation about the z-axis or yaw. The Newton-Euler equations of motion expressed in the body coordinate frame of a planar rigid vehicle are:

$$m\dot{v}_x - mrv_y = F_x \tag{12.47}$$

$$m\dot{v}_y + mrv_x = F_y = \frac{1}{v_x}\left(-a_1C_{af} + a_2C_{ar}\right)r - \frac{1}{v_x}\left(C_{af} + C_{ar}\right)v_y + C_{af}\delta \tag{12.48}$$

$$I_z\dot{r} = M_z = \frac{1}{v_x}\left(-a_1^2C_{af} - a_2^2C_{ar}\right)r - \frac{1}{v_x}\left(a_1C_{af} - a_2C_{ar}\right)v_y + a_1C_{af}\delta \tag{12.49}$$

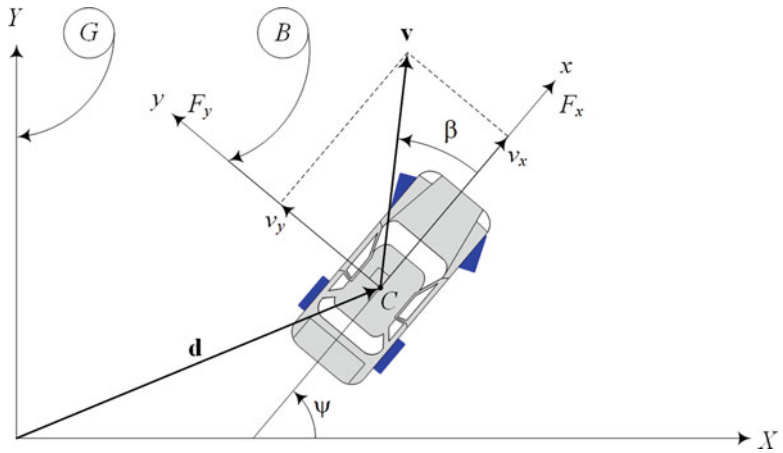


Fig. 12.11 A local body coordinate frame attached to the vehicle's centre of mass C determines the global location d and its orientation

where $r = \dot{\Psi} = \omega_z$, is the vehicle yaw rate, a_1 and a_2 are the distance from the rear and front wheels to the centre of gravity respectively, $C_{\alpha f}$ and $C_{\alpha r}$ are the cornering stiffness of the front and rear tires, respectively, and steer angle δ is the cot-average of the front left and right wheels. The right-hand side of the set of equations indicates that the steer angle δ and the forward traction force F_x are the inputs of the system. The left-hand side of the equations determines that the velocity components v_x and v_y, and the yaw rate r of the vehicle are the outputs of the system.

Let us rewrite the equations in the following form, assuming the forward speed of the vehicle v_x to be the input to the system. It will result in F_x to be an output of the system of equations of motion.

$$\dot{v}_x = \frac{F_x}{m} + r v_y$$

$$\dot{v}_y = \frac{1}{m v_x}\left(-a_1 C_{af} + a_2 C_{ar}\right) r - \frac{1}{m v_x}\left(C_{af} + C_{ar}\right) v_y + \frac{1}{m} C_{af}\delta - r v_x \tag{12.50}$$

$$\dot{r} = \frac{1}{I_z v_x}\left(-a_1^2 C_{af} - a_2^2 C_{ar}\right) r - \frac{1}{I_z v_x}\left(a_1 C_{af} - a_2 C_{ar}\right) v_y + \frac{1}{I_z} a_1 C_{af}\delta$$

The vehicle is assumed to move with constant velocity for the sake of simplifying the equations.

$$v_x = const \tag{12.51}$$

This will cause the steer angles of the vehicle to be the only dynamic system input. Therefore, the two outputs of the system will be the lateral speed v_y, yaw rate r. Traction force F_x is another output which, according to the first equation will be determined by the velocity and the solutions of the equations. Having the function of the steering angles and starting from an initial condition for $v_y(0)$ and $r(0)$ will enable the determination of v_y (lateral velocity) and r (yaw rate) in later times. Integrating v_x , v_y and r results in the body and global (or world) coordinate representations of the position and the vehicle orientation.

Integrating the yaw rate of the vehicle will determine the heading angle or orientation

$$\Psi = \int_0^t r dt \tag{12.52}$$

and now by multiplying the rotation transformation matrix ${}^{G}R_B$, the velocity vector of the vehicle in the global frame can be determined.

$${}^{G}v_c = {}^{G}R_B \, {}^{B}v_c$$

$$\begin{bmatrix} \frac{v_X}{v_Y} \end{bmatrix} = \begin{bmatrix} \cos\Psi & -\sin\Psi & 0 \\ \sin\Psi & \cos\Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12.53}$$

Integrating components of velocity expressed in the global frame and summed with the present location of the mass centre will give the present location of the vehicle.

$${}^{G}\boldsymbol{d} = {}^{G}\boldsymbol{d}_0 + \int {}^{G}v\, dt$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} + \begin{bmatrix} \int_0^t \left(v_x \cos\Psi - v_y \sin\Psi\right) dt \\ \int_0^t \left(v_x \sin\Psi + v_y \cos\Psi\right) dt \end{bmatrix} \tag{12.54}$$

where *Xo* and *Yo* are the coordinates of the initial location which are assumed to be given values.

The actual or dynamic instantaneous curvature radius for such a vehicle would be

$$R = \frac{v}{r} = \frac{\sqrt{v_x^2 + v_y^2}}{r} \approx \frac{v_x}{r}\sqrt{1+\beta^2} \approx \frac{v_x}{r} \tag{12.55}$$

and the body and global frames' representation of the position of the dynamic rotation centre is

$${}^{B}\boldsymbol{r}_c = \begin{bmatrix} R\sin\beta \\ R\cos\beta \end{bmatrix} \approx \begin{bmatrix} 0 \\ R \end{bmatrix} \tag{12.56}$$

$${}^{G}\boldsymbol{r}_c = {}^{G}\boldsymbol{d} + {}^{G}R_B\, {}^{B}\boldsymbol{r}_c$$

$$\begin{bmatrix} X_c \\ Y_c \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} \cos\Psi & -\sin\Psi & 0 \\ \sin\Psi & \cos\Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ R \end{bmatrix} = \begin{bmatrix} X + R\sin\Psi \\ Y + R\cos\Psi \end{bmatrix}$$

$$= \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} + \begin{bmatrix} \int_0^t \left(v_x \cos\left(\int_0^t r dt\right) - v_y \sin\left(\int_0^t r dt\right)\right) dt \\ \int_0^t \left(v_x \sin\left(\int_0^t r dt\right) + v_y \cos\left(\int_0^t r dt\right)\right) dt \end{bmatrix} \tag{12.57}$$

where R is illustrated in Fig. 12.12.

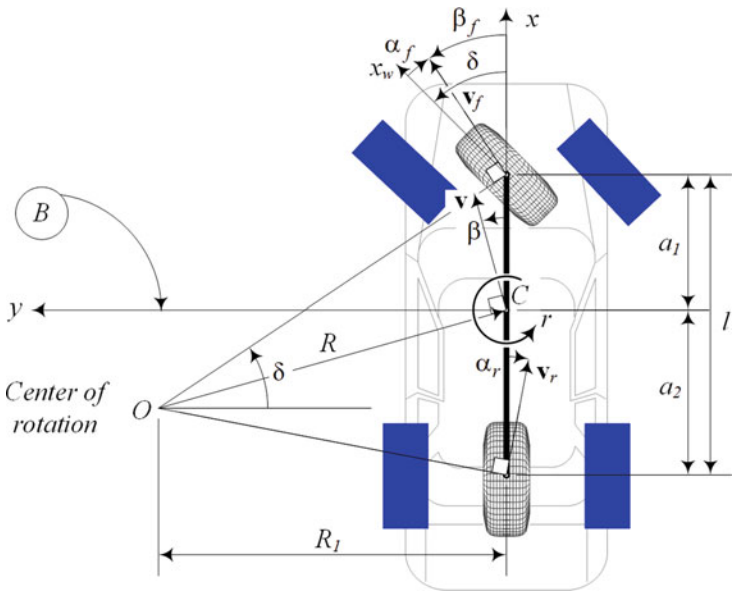


Fig. 12.12 The location of the dynamic rotation centre and the radius of rotation of a vehicle in motion

By solely employing the outputs of the other two equations into the first equation of motion, the required traction force is determined as an output of the system.

$$F_x = -m\frac{v_x}{r} \tag{12.58}$$

12.4.4 Autonomous Control

The input-output dynamics of a vehicle with assuming a constant speed is depicted in Fig. 12.13. The planar bicycle model is the simplest practical mathematical model for vehicle dynamics. Its practicality has been proven by experiment, and it is very well to be used to prove the concept of the autodriver control. A front-wheel steering vehicle is employed: steer angle δ as the input and traction force F_x, generates lateral velocity v_y, and r (yaw rate) as outputs. The interface box is the dynamic equations of motion which uses forward velocity v_x, the geometric characteristics of the vehicle a_1, a_2, and tire-ground dynamic parameters $C_{\alpha f}$, $C_{\alpha r}$, as parameters

The outputs of the dynamic equations are time rate of kinematic variables in the body coordinate frame. Transforming them to the global coordinate frame followed by integration, which determines the vehicle's actual orientation and location on the ground coordinate frame. This step is illustrated in Fig. 12.14. The result would

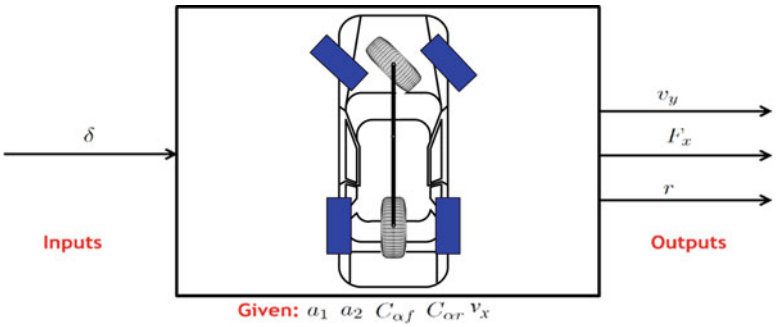


Fig. 12.13 Input-output relationship in vehicle dynamics using planar bicycle model

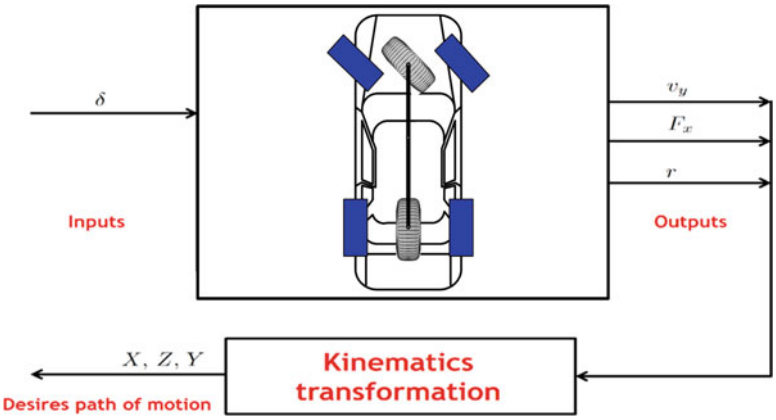


Fig. 12.14 Determination of the global position of a vehicle

be an actual location of the vehicle on the road to be compared with the desired location.

To control the vehicle, inverse dynamic problem must be solved and should determine speed v_x and steer angle δ required such that the actual vehicle position matches with the desired position on the road. Consider Fig. 12.15 shows a loop starting with a desired location on the road, *X, Y, Z*. The mentioned position coordinates will then be fed back to the set of integrations and kinematic transformation mentioned earlier so that the associated kinematics variables of the vehicle under study r, and v_y would be determined. These variables, which are typically outputs of the vehicle dynamics, must be fed back to the vehicle's dynamic equation box theoretically to determine the required steering angle input δ. If this steer angle δ was the correct value then feeding it into the box (dynamic equations) would locate the vehicle in the correct position and the real *X, Y, Z* would be the same as the desired one on the road. However, feeding the kinematic variables r and v_y into the box (dynamic equation) in order to determine the required input δ is not a

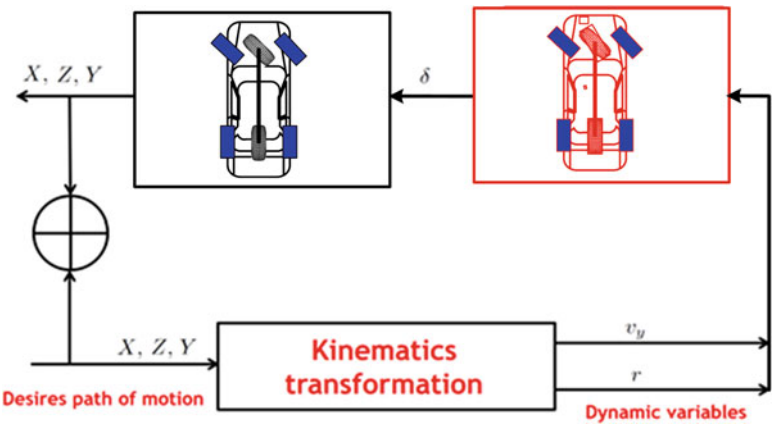


Fig. 12.15 Ideal reverse dynamic to locate a vehicle at a desired position on the road. Reverse calculating (the box in red) the steer angle δ is not straightforward

straightforward step and will not work correctly. This is shown by the red box in Fig. 12.15 to express the idea that there is not an easy way to determine the requirements.

Solving reverse dynamics differential equations, there are several numerical and approximation methods. The shooting method is the usual numerical method that might be a solution to the blocked box in Fig. 12.15. However, this method cannot be applied to an on-time control system. As a result, a quick method is introduced to calculate δ. It has investigated the transient responses of vehicles compared to their steady-state responses. These studies show that under normal conditions, the vehicle's transient response is very close to their steady-state behaviour or in other words that the vehicles are lazy. The same set of equations introduced earlier governs the steady-state conditions of a vehicle. The significant difference is that all-time derivative operators are the equations are set to zero. This results in a set of algebraic equations which gives the steady-state steer angle δ_{ss}.

$$\begin{aligned} 0 &= \frac{F_x}{m} + r v_y \\ 0 &= \frac{1}{m v_x}\left(-a_1 C_{\alpha f} + a_2 C_{\alpha r}\right) r - \frac{1}{m v_x}\left(C_{\alpha f} + C_{\alpha r}\right) v_y + \frac{1}{m} C_{\alpha f}\delta - r v_x \\ 0 &= \frac{1}{I_z v_x}\left(-a_1^2 C_{\alpha f} - a_2^2 C_{\alpha r}\right) r - \frac{1}{I_z v_x}\left(a_1 C_{\alpha f} - a_2 C_{\alpha r}\right) v_y + \frac{1}{I_z} a_1 C_{\alpha f}\delta \end{aligned} \tag{12.59}$$

Therefore, the blocked square with the set of steady-state algebraic equations is substituted. This enables the estimation of the steering angle, which generates the required v_y and r that they are all at their steady-state condition, as illustrated in Fig. 12.16. The calculated steer angle is used as input to the set of transient equations of motion to calculate instant values of v_y and r associated with the scenario under investigation. These variables will be used to calculate the actual road position of the vehicle. However, due to the approximated value of the steering angle, the coordinates of the vehicle on the ground would be potentially different from the

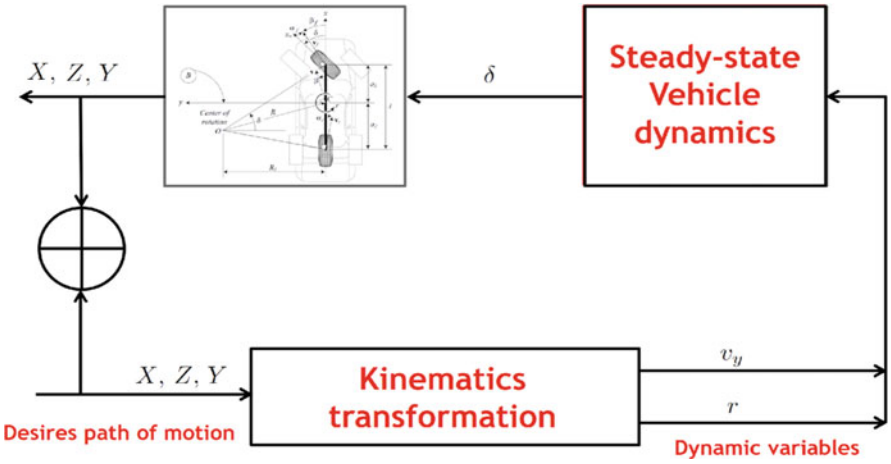


Fig. 12.16 The simplest dynamic model of planar vehicle and planar road

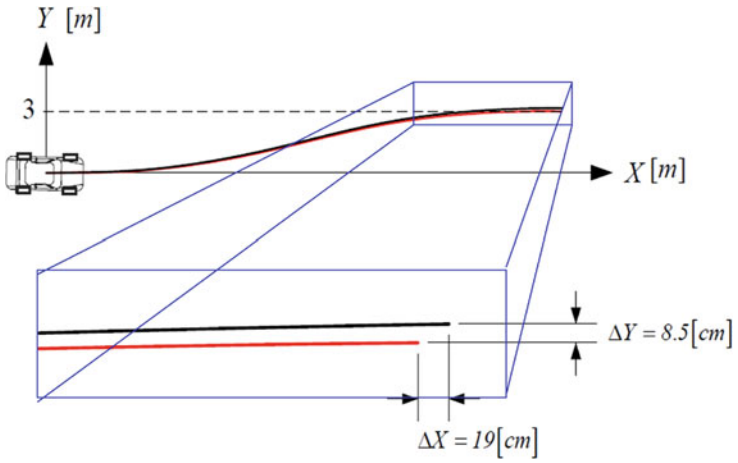


Fig. 12.17 Calculating the required steer angle as the input to the vehicle to provide the desired position on the road using steady-state equations

desired values. The actual variables should be compared with the desired values to close the control loop and generate a control strategy to compensate for the errors.

There would be potential differences between X_d, X, which are the desired and actual locations on the X-axis, respectively, and Y_d, Y which are the desired and actual locations on the Y-axis, respectively, as is illustrated in Fig. 12.17. It shows a 3-meter lateral lane change of the vehicle in a 100-meter-long path. The difference between lateral locations needs a set-out in the value of the steering angle.

A PID controller will eliminate the possible errors between the actual (X, Y) and desired (X_d, Y_d) locations. The difference between longitudinal locations will

have to be adjusted by changing the vehicle speed. The closed-loop control of the proposed system is shown in Fig. 12.16, employing a PID control design to illustrate the concept.

Figure 12.17 illustrates the possible errors that might happen in a manoeuvre of an autonomous vehicle using the autodriver algorithm. The vehicle may have longitudinal and lateral location errors, e_X, e_Y, between the desired and actual positions of the vehicle. The longitudinal error e_X may be a result of many parameters and variables such as aerodynamic forces, estimation of the tire and road characteristics, etc. However, it can be compensated by adjusting the forward velocity of the vehicle using a controller on the longitudinal error signal. Similarly, the lateral position error e_Y also may be a result of many parameters and variables such as lateral wind, bank angle of the road, misalignment of the wheels, etc. The lateral position error can be compensated by adjusting the steering angle of the vehicle.

Figure 12.18 illustrates the overall control system and the loop starting with the desired global coordinates of the vehicle $X_d(s)$, $Y_d(s)$. These coordinates are functions of a parameter, say, which varies when the distance vehicle moves on the road. The desired global (or world) coordinates will be transformed into the vehicle body coordinate frame. As a result, the road equation will be defined in the body coordinate to be used to calculate the loci of the road curvature centre in the body coordinate in which the actual instant centre of rotation of the vehicle will be calculated to compare. Then, the steady-state equations of motion of the vehicle will provide an estimate for the required steer angle to turn the vehicle about the road curvature centre at the current location of the vehicle. The estimated steer angle is one of the two required input set of (δ, v_x) to feed the instantaneous dynamic equations of motion of the vehicle. The output of the dynamic equations of motion is the lateral velocity v_y, traction force F_x and yaw rate r. These outputs are all calculated in the body coordinate frame. A backward transformation will provide us with their values in the global coordinate frame. Integrating velocity components and the yaw rate of the vehicle determines the position and orientation of the vehicle on the road, X, Y. At this stage, the system needs to compare the real position of the vehicle with the desired one. The differences will generate the error signals to be fed into the PID controller and provide adjusting signals to the steering angle and forward velocity.

The optimal values of the three PID gains can be tuned using several methods (i) off-line (practical) such as (1) Ziegler-Nichols (ZN), (2) Cohen-Coon (CC), (3) Chien-Hrones-Reswick (CHR) and (4) ITAE tuning and (ii) on-line (intelligent) such as (1) neural network tuner (NNT), (2) fuzzy logic (FL) and (3) genetic algorithm [23, 24, 33]. PID controllers tuned by ZN method are commonly used in the automotive industry [24, 33]. Based on the values L and T (see Fig. 12.19a and b), ZN finds the first set of the controller. For aperiodic responses, the PID gains are tuned according to the Fig. 12.19c.

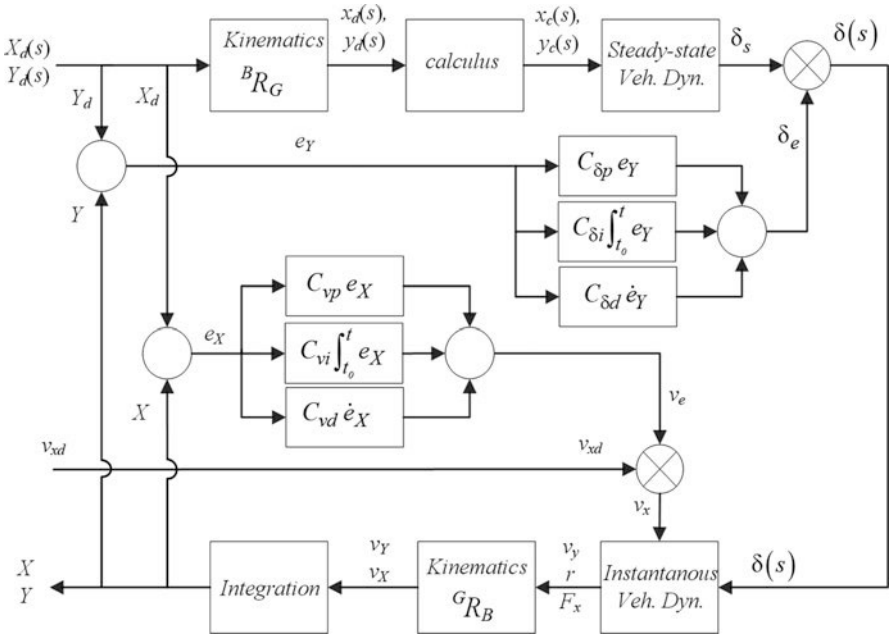


Fig. 12.18 The control loop of the autodriver autonomous vehicle

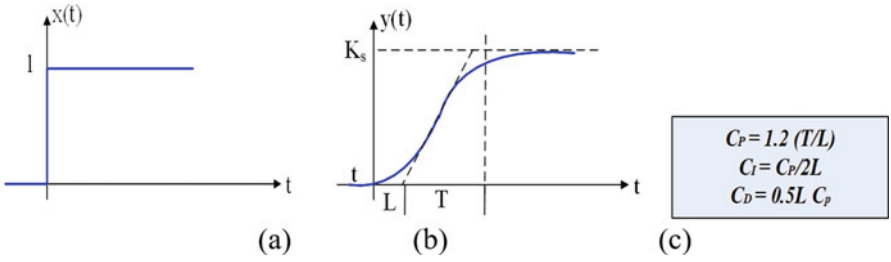


Fig. 12.19 (**a**) Test signal, (**b**) unit-step response, and (**c**) ZN tuning

12.4.5 Case Study Scenarios

12.4.5.1 Constant Velocity

Consider a vehicle with the following characteristics at a constant speed of v_x=20 m/s that is supposed to turn on a flat circle of *R=100* m.

$$\begin{gathered} C_f = 57000\ N/rad, C_r = 52000\ N/rad \\ m = 900\ kg, I_z = 1200\ kgm^2 \\ a_1 = 0.91\ m, a_2 = 1.64\ m \end{gathered} \tag{12.60}$$

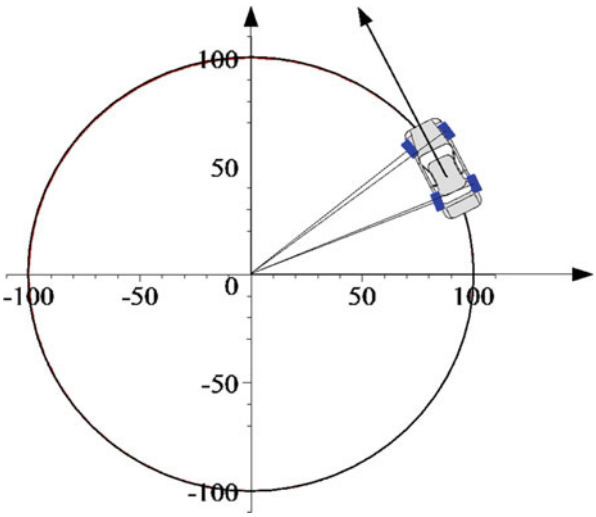


Fig. 12.20 Applying the autodriver autonomous control on a circular path motion

The same analysis was repeated with more complicated nonlinear inputs to the system to evaluate the algorithm's capability in different scenarios.

The proposed PID controller gains using the ZN method are set to an estimated below value (see Eq. 12.58) of and applied to the constant velocity case.

$$\left[C_{P,} C_{D,}\ C_{I,}\right] = [0.12, 0.075, 0.031] \tag{12.61}$$

The path after the application of the controller is labelled as exact compared to the autodriver algorithm resulted in path labelled approximate in Fig. 12.20. The differences in the case of constant velocity and steering are small and hard to see in Fig. 12.20.

12.4.5.2 Nonlinear Varying Steering

The study investigated different nonlinearities introduced to the system. The first example was done by keeping the velocity constant but changing the steering angle according to the following equation and result is in Fig. 12.21.

$$\delta = \delta_0 \left(H\left(t - t_0\right) + \sin^2\left(\frac{t}{2t_0/\pi}\right) H\left(t_0 - t\right)\right) \tag{12.62}$$

where $H(t - t_0)$is the Heaviside function and $t_0 = 1s$, is the response time.

The resultant path of motion with constant velocity and nonlinear varying steering is illustrated in Fig. 12.21. The proposed PID controller gains tuned using the ZN method are set to an estimated below value (see Eq. 12.60) of and applied to

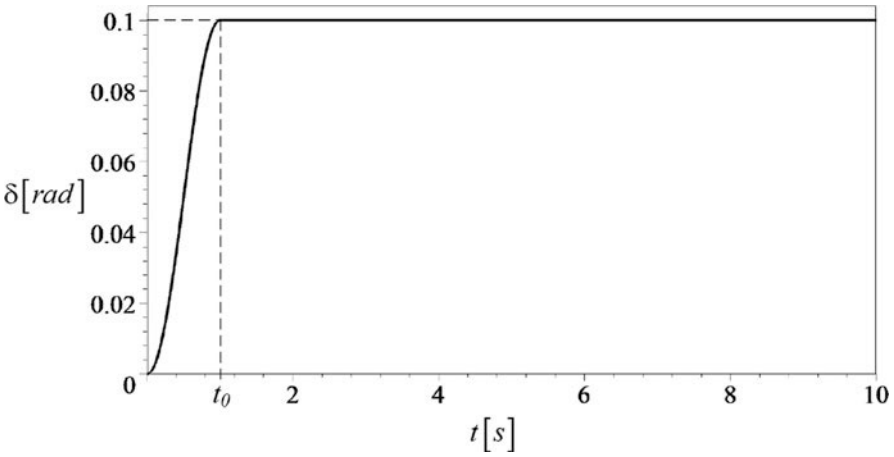


Fig. 12.21 Nonlinear steering input to the autodriver algorithm

constant velocity and nonlinear varying steering case scenario.

$$\left[C_{P,} C_{D,}\ C_{I,}\right] = [0.2, 0.09, 0.04] \tag{12.63}$$

The comparison of the resulted path of motion after and before application of PID controller on the constant velocity with variable steering case is shown in Figs. 12.22 and 12.23 [25–32].

12.4.5.3 Nonlinear Variable Velocity

For a second nonlinear test, a variable velocity was used while the steering input was kept constant during the motion of the vehicle. The velocity was changed according to the below equation.

$$v_x = \frac{20}{t_0} t\ H\ (t_0 - t) + 20\ H\ (t - t_0) \quad m/s \tag{12.64}$$

While the steering angle was kept at 0.1 *rad*/*sec* during the motion, the resultant path of motion is illustrated in Fig. 12.24.

The proposed PID controller gains tuned using the ZN method are set to an estimated below the value of and applied to nonlinear variable velocity case scenario.

$$\left[C_{P,} C_{D,}\ C_{I,}\right] = [0.61, 0.032, 0.021] \tag{12.65}$$

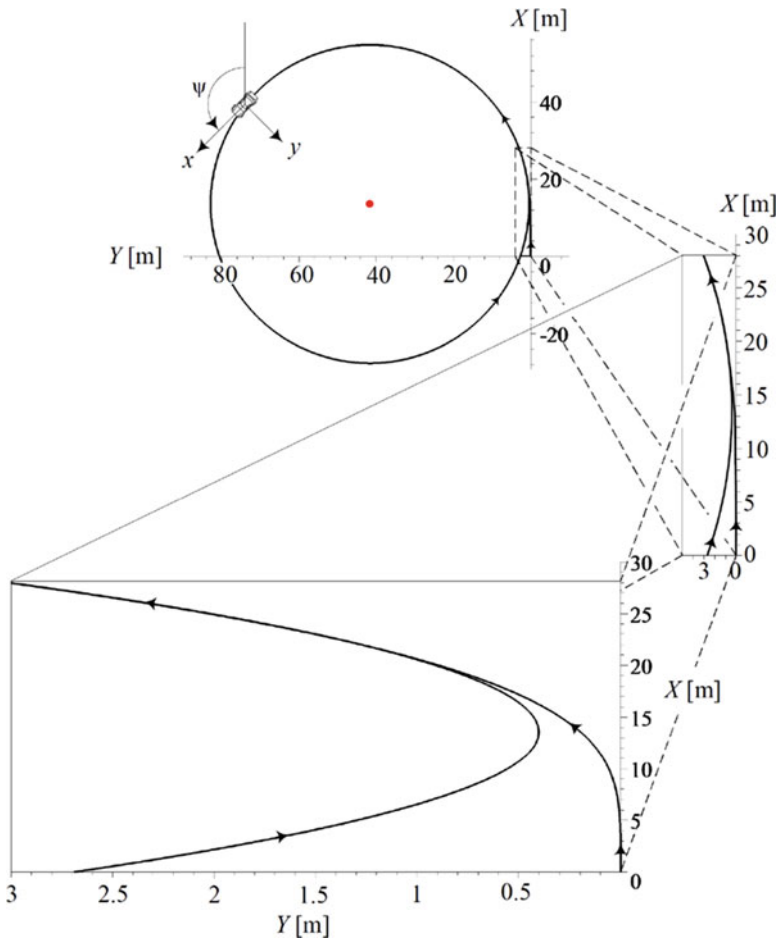


Fig. 12.22 A resulted path of motion for constant velocity and non-linearly varying steering input

However, determination of the optimal gain values, considering variable gains, as well as making the control strategy more complicated is due for further future study. Application of the controller, as mentioned above, resulted in reducing the errors significantly. The path after the application of the controller is labelled as exact compared to the autodriver algorithm resulted in path labelled approximate in Fig. 12.25.

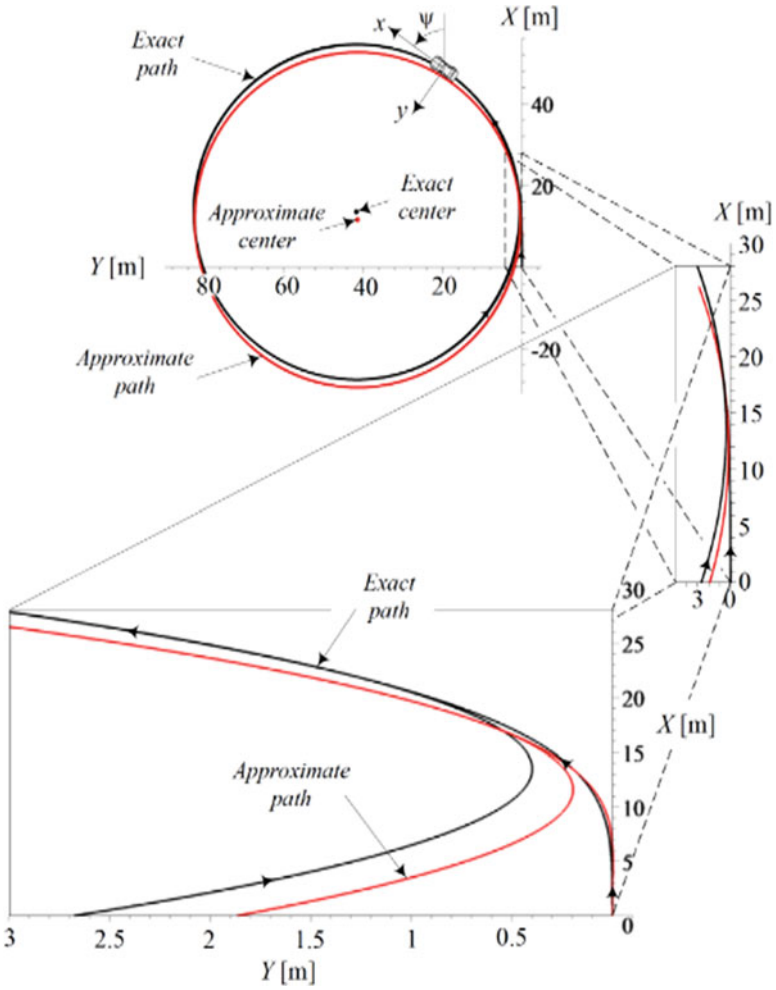


Fig. 12.23 Comparison of the resulted path of motion after and before application of the PID controller in the variable steering case

12.4.6 Planar-Roll Vehicle Dynamics

An autonomous vehicle is a trendy topic and has been a significant focus of research in recent years. The smart system call autopilot is now categorised in different levels which can make semi- to fully self-drive the vehicle by a set of developed theory and algorithm from one destination to the others. Companies around the world work hard on the software and hardware that desire to control the vehicle smarter. We use the lateral controller for designing dynamic path tracking, as a critical component of the control system. This ultimate smart system can guide a vehicle

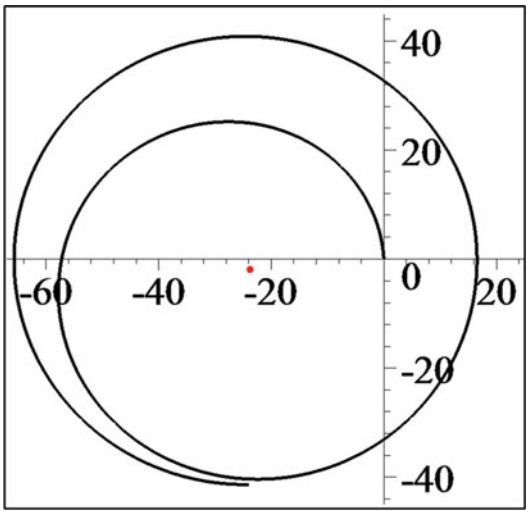


Fig. 12.24 Resulted path of motion with variable non-linear velocity and constant steering input to the autodriver algorithm

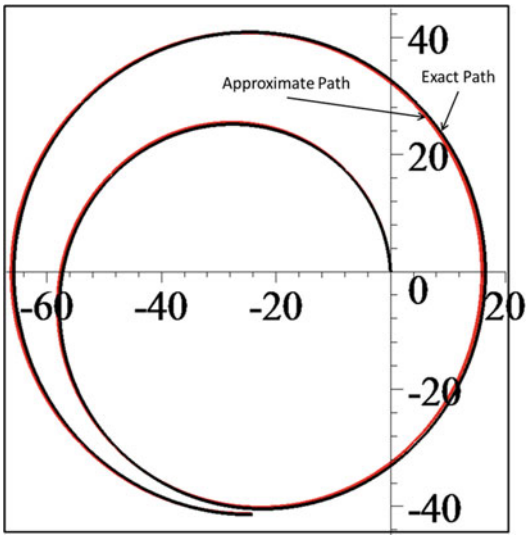


Fig. 12.25 Comparison of resulted paths of motion after and before application of the PID controller in the variable velocity case

with no human control or assistant human to control when driving. Many countries will use autonomous transportation. Application of autonomous vehicle using self-drive system is capable of applying to cargo, naval steering, military, aircraft even space ship. Autopilot function can reduce human fault and error, relaxing the driving stress. Autodriver algorithm was introduced as a path-following algorithm for autonomous vehicles using road geometry data and planar vehicle dynamics. The autodriver algorithm is improved according to practical implications while a more realistic vehicle model (roll model) is used, which considers roll degree of freedom in addition to planar motion. A ghost-car path-following approach

is introduced to define the desired location of the car at every instance. Key steady-state characteristics of turning vehicles, namely, the curvature, yaw rate and side-slip responses, are discussed and used to construct a feed-forward component of a path-following controller based on autodriver algorithm. Feedback control loops are designed and applied to minimise transient errors between the road and vehicle positions. Finally, simulations are performed to analyse the path-following performance of the proposed scheme. The results show promising performance of the controller both in terms of error minimisation and passenger comfort.

12.4.6.1 Equations of Motion

The vehicle model in this study (the bicycle-roll model) is based on a rigid rollable vehicle moving in the osculating plane of the road. The equations of motion of the bicycle-roll vehicle model (Fig. 12.26), expressed in the principal body coordinate frame B attached to the vehicle body at its mass centre C, are governed by the following set of nonlinear coupled ordinary differential equations.

$$\dot{v}_x = \frac{F_x}{m} + r v_y \tag{12.66}$$

$$\dot{v}_y = \left(\frac{C_r}{m} - v_x\right) + \frac{C_p}{m} p + \frac{C_\beta}{m}\frac{v_x}{v_y} + \frac{C_\varphi}{m}\varphi + \frac{C_\delta}{m}\delta \tag{12.67}$$

$$\dot{p} = \frac{1}{I_x}\left(E_r r + E_p p + E_\beta \frac{v_x}{v_y} + E_\varphi \varphi + E_\delta \delta\right) \tag{12.68}$$

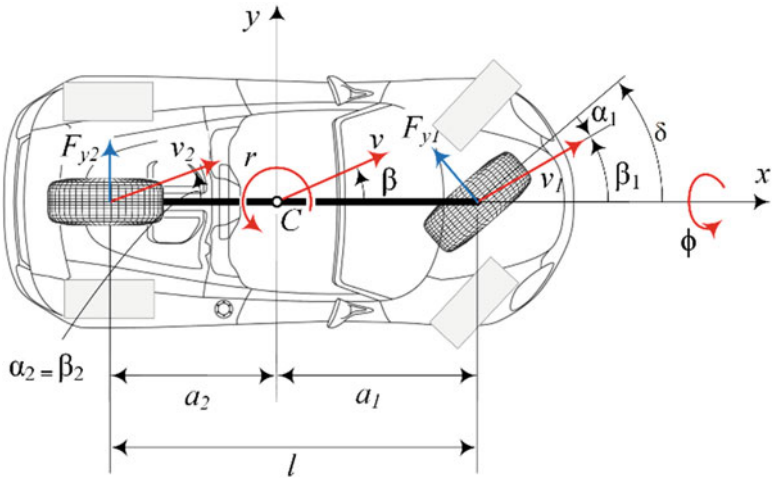


Fig. 12.26 Bicycle-roll vehicle model

$$\dot{r} = \frac{1}{I_Z}\left(D_r r + D_p p + D_\beta \frac{v_x}{v_y} + D_\varphi \varphi + D_\delta \delta\right) \tag{12.69}$$

where,

$$C_r = \frac{\partial F_y}{\partial r} = -\frac{a_1}{v_x} C_{af} + \frac{a_2}{v_x} C_{ar} \tag{12.70}$$

$$C_p = \frac{\partial F_y}{\partial p} = \frac{C_{af} C_{\beta f}}{v_x} + \frac{C_{ar} C_{\beta r}}{v_x} \tag{12.71}$$

$$C_\beta = \frac{\partial F_y}{\partial \beta} = -\left(C_{af} + C_{ar}\right) \tag{12.72}$$

$$C_\varphi = \frac{\partial F_y}{\partial \varphi} = C_{ar} C_{\delta\varphi r} + C_{af} C_{\delta\varphi f} - C_{\varphi f} - C_{\varphi r} \tag{12.73}$$

$$C_\delta = \frac{\partial F_y}{\partial \delta} = C_{af} \tag{12.74}$$

$$D_r = \frac{\partial M_z}{\partial r} = -\frac{a_1^2}{v_x} C_{af} - \frac{a_2^2}{v_x} C_{ar} \tag{12.75}$$

$$D_p = \frac{\partial M_z}{\partial p} = \frac{a_1}{v_x} C_{\beta f} C_{af} - \frac{a_2}{v_x} C_{\beta r} C_{ar} \tag{12.76}$$

$$D_\beta = \frac{\partial M_z}{\partial \beta} = -\left(a_1 C_{af} - a_2 C_{ar}\right) \tag{12.77}$$

$$D_\varphi = \frac{\partial M_z}{\partial \varphi} = -a_1\left(C_{\varphi f} - C_{af} C_{\delta\varphi f}\right) + a_2\left(C_{\varphi r} - C_{ar} C_{\delta\varphi r}\right) \tag{12.78}$$

$$D_\delta = \frac{\partial M_z}{\partial \delta} = a_1 C_{af} \tag{12.79}$$

$$E_r = \frac{\partial M_z}{\partial r} = -\frac{a_1}{v_z} C_{Tf} C_{af} + \frac{a_2}{v_z} C_{Tr} C_{ar} \tag{12.80}$$

$$E_p = \frac{\partial M_z}{\partial p} = \frac{1}{v_z} C_{\beta f} C_{Tf} C_{af} + \frac{1}{v_z} C_{\beta r} C_{Tr} C_{ar} - c_\varphi \tag{12.81}$$

$$E_\beta = \frac{\partial M_z}{\partial \beta} = -C_{Tf} C_{af} - C_{Tr} C_{ar} \tag{12.82}$$

$$E_\varphi = \frac{\partial M_z}{\partial \varphi} = -C_{Tf}\left(C_{\varphi f} - C_{af} C_{\delta\varphi f}\right) - k_\varphi - C_{Tr}\left(C_{\varphi r} - C_{ar} C_{\delta\varphi r}\right) \tag{12.83}$$

$$E_\delta = \frac{\partial M_z}{\partial \delta} = C_{Tf} C_{af} \tag{12.84}$$

In the above equations, the steer angle δ acts as the input, and there is a total of four differential equations of motion. Assume v_x to be a varying parameter for a third-order system constructed by Eqs. 12.63, 12.64, and 12.65. The tire forces are assumed to be proportional to the side-slip angles; also, the right and left tires' cornering stiffnesses are assumed to be similar and equal to half of the entire axles.

12.4.6.2 Steady-State Responses

By using the bicycle-roll vehicle model, the primary vehicle responses at steady-state turning are defined concerning the steer angle input for a front-wheel steering vehicle. At a steady-state turning, the time derivative of the variables are equal to zero, and the equations of motion simplify to a set of algebraic equations. The steady-state curvature, yaw rate and side-slip responses which are used in this study are derived and simplified as:

$$S_\kappa = \frac{k}{\delta} = \frac{1}{\rho\delta} = -\frac{Z_1}{v_x Z_0} \tag{12.85}$$

$$S_r = \frac{r}{\delta} = \frac{k}{\delta} v_x = S_k v_x = -\frac{Z_1}{Z_0} \tag{12.86}$$

$$S_\beta = \frac{\beta}{\delta} = \frac{Z_2}{Z_0} \tag{12.87}$$

where Z_0, Z_1, Z_2 are given as:

$$Z_0 = E_\beta \left(D_r C_\varphi - C_r D_\varphi + m v_x D_\varphi\right) + E_\varphi \left(C_r D_\beta - D_r C_\beta - m v_x D_\beta\right) + E_r \left(C_\beta D_\varphi - D_\beta C_\varphi\right) \tag{12.88}$$

$$Z_1 = E_\beta \left(C_\varphi D_\delta - v_x C_\delta D_\varphi\right) - E_\varphi \left(C_\beta D_\delta - v_x C_\delta D_\beta\right) + E_r \left(C_\beta D_\varphi - D_\beta C_\varphi\right) \tag{12.89}$$

$$Z_2 = E_\phi \left(D_\delta m v_x + C_\delta D_r - C_r D_\delta\right) + \left(C_\phi D_\delta - C_\delta D_\phi\right) E_r + \left(C_r D_\phi - C_\phi D_r - D_\phi m v_x\right) E_\delta \tag{12.90}$$

12.4.7 Vehicle Behaviour

Ground vehicles work in steady-state conditions most of the time. The planar vehicle motion at any instant may be described by being in steady-state or transitioning

between two steady-state conditions. The steady-state response of vehicles is dominant in typical public streets driving conditions when the velocity and steer angle inputs are constant or changing slowly. The transient effect of passenger vehicles is almost negligible in terms of dynamic variables' responses. This feature of the vehicle dynamics allows us to plot a steady-state chart for the local coordinates of the instantaneous centre of rotation (ICR) of the bicycle-roll vehicle model in the body-fixed frame (Fig. 12.26). Figure 12.27 shows that the lateral position of the ICR mainly depends on the steer angle input from the driver and the longitudinal position is mainly affected by the forward velocity.

Note that the constant steer angle and constant velocity curves on Fig. 12.27 are not exactly linear. For a given radius of rotation ρ_0, the side-slip angle β is dictated by the selection of v_x or δ. Theoretically, let δ be free and assume there is no friction limit under the tires; the vehicle may negotiate a given turn with any velocity at the expense of going under some induced side-slip angle which is not necessarily desirable. To negotiate a turn with a given radius, the ICR must be at the same distance from the vehicle body frame; by choosing the desired velocity, a required steer angle is determined. A sample point P illustrated in the figure shows that for the velocity of 20 m/s, it requires a steering angle of around 1.75 degrees to stay on a circular road with radius ρ_0. Considering velocity as the given limiting factor, using Eqs. 12.76 and 12.78, the required steer angle to keep a radius of turning ρ is obtained as a function of velocity as:

$$\delta_\rho = \delta_\rho(v_x) = -\frac{v_x Z_0}{\rho Z_1} \tag{12.91}$$

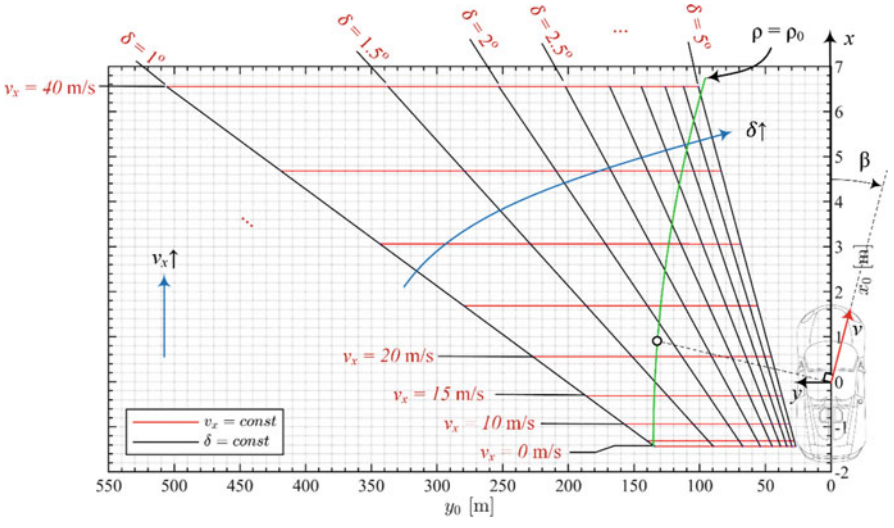


Fig. 12.27 ICR coordinate in body frame (not in scale)

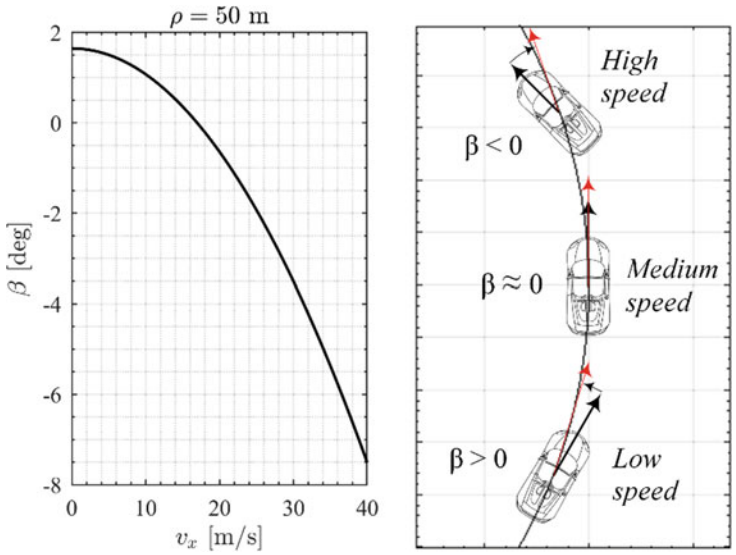


Fig. 12.28 Variation of side-slip angle versus velocity for a constant radius of curvature

$$\beta_\rho = \beta_\rho\left(v_x\right) = \frac{\delta_\rho Z_2}{Z_0} = -\frac{v_x Z_2}{\rho Z_1} \tag{12.92}$$

which is visualized in Fig. 12.28 for the vehicle of interest. The values of vehicle parameters are as follows

$C_{\alpha f} = 52000$ N/rad $C_{\alpha r} = 72000$ N/rad $m = 845.4$ kg $I_z = 1490$ kgm^2 $I_x = 350$ kgm^2 $a_1 = 0.909$ m $a_2 = 1.436$ m $k_\varphi = 26612$ Nm/rad $c_\varphi = 1700$ Nms/rad

12.4.8 Autonomous Control

12.4.8.1 Improved Autodriver Algorithm

Vehicles in motion on the road are always turning about the curvature centre of the road at the right curvature radius. Considering a given road $\mathbf{r} = \mathbf{r}\left(X, Y, Z, \psi\right)$ to be a three-dimensional spatial curve, and the vehicle to be driven in the osculating plane, can calculate the path of curvature centre in the osculation plane, both in global coordinate and vehicle body coordinate frames. The desired location of the road curvature centre must coincide with the vehicle's rotation centre. If correct inputs δ, v_x are selected, there will be no error between the actual position of the vehicle, ${}^G\mathbf{r} = {}^G\mathbf{r}\ \left(X, Y, Z, \psi\right)$ with the desired one ${}^G\mathbf{r}_d = {}^G\mathbf{r}_d\ \left(X, Y, Z, \psi\right)$. The loci of curvature centre for two sample roads are shown in Fig. 12.29.

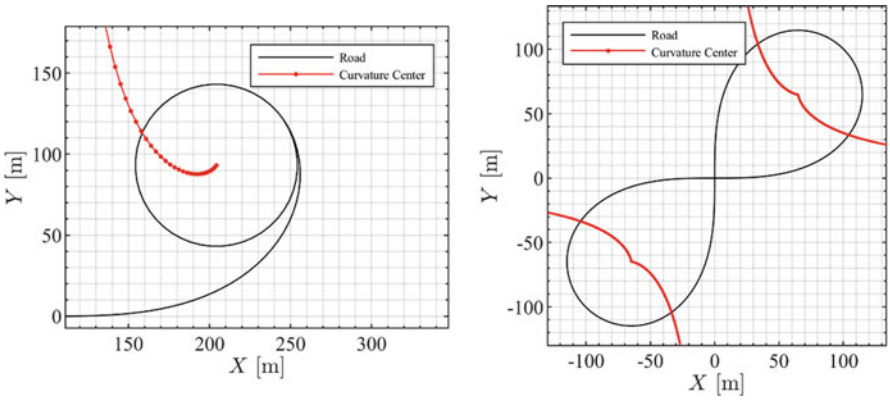


Fig. 12.29 Sample roads and their curvature centres

Based on the discussion in section II, ideally, keep the vehicle on any curve as long as you choose the correct steer angle and velocity from the ICR chart (Fig. 12.27) to coincide the ICR with the curvature centre of the road in vehicle body coordinate frame. Although this solution is correct as of the final value, due to the transient period dynamics, it cannot eliminate the error between the path of motion and the road. In other words, the overall path of motion would have a similar shape to the road curve, but there would be an increasing offset as the vehicle moves forward due to an initial error during the transient stage. Moreover, compensate for the transient error of path-following, a feedback controller is designed in the following section.

12.4.8.2 Calculation of Steady-State Inputs

It can be seen from the axes scales that the steering angle mainly determines the turning radius of the vehicle. Although the velocity does have some contribution to the ICR location, its effect is minor, especially at small side-slip angles. Hence, assume the main contribution on the radius of rotation is coming from the steer angle and $y_0 \approx \rho_0$. Thus, aim to set the steering angle to position the ICR laterally correctly (regulating y_0); then the vehicle running at a given velocity will eventually gain the side-slip angle β dictated by the velocity, which causes the longitudinal position of ICR in body coordinate to match the road curvature centre. The desired velocity is usually given as a velocity profile for different sections of a road.

12.4.8.3 Elimination of Transient Error

In addition to the ICR, the actual global position of the vehicle is calculated as

$$X = \int_0^t v_x \cos\psi - v_y \sin\psi \; dt \tag{12.93}$$

$$Y = \int_0^t v_x \sin\psi + v_y \cos\psi \; dt \tag{12.94}$$

which are solved together with Eqs. 12.67, 12.68, and 12.69, based on the approximated inputs coming from steady-state analysis. Any lateral position error in the vehicle's body frame may be compensated by adjusting the steering angle, and any longitudinal position error in vehicle's body frame may be compensated by adjusting the vehicle's forward acceleration, and hence the velocity. Road constraint implies that there would be no error in Z direction. The overall algorithm may be summarised in the following steps.

A desired path of motion is given in 3D as ${}^G\mathbf{r}_d = {}^G\mathbf{r}_d \; (X_d(s), Y_d(s), Z_d(s))$ as a function of arc length s. Differential geometry provides us with the location of the road curvature centre in global coordinates ${}^G\mathbf{r}_c = {}^G\mathbf{r}_c \; (X_c(s), Y_c(s), Z_c(s))$.

Kinematic transformation provides the coordinates of interest (road and its curvature centre) in the vehicle body coordinate frame as: ${}^B\mathbf{r}_d = [x_d \quad y_d \quad z_d]^T = {}^B R_G \; {}^G\mathbf{r}_d$, ${}^B\mathbf{r}_c = [x_c \quad y_c \quad z_c]^T = {}^B R_G \; {}^G\mathbf{r}_c$,

Steady-state equations determine the required steer angle $\delta_{ss}(s)$ for a given speed v_x and a radius of rotation ρ in order to laterally coincide the rotation centre of the vehicle with curvature centre of the road.

The lateral error between the desired vehicle position and its actual position in the body coordinate frame $e_y = y_d - y = y_d$ is fed into a feedback controller. The controller provides $\delta_e(s)$ to adjust the total steer angle $\delta = \delta_{ss} + \delta_e$ to eliminate the lateral position error

The longitudinal error between the desired vehicle position and its actual position in the body coordinate frame $e_x = x_d - x = x_d$ is fed into a longitudinal feedback controller which provides $\dot{v}_x$ to adjust forward velocity and eliminate the longitudinal position error.

As mentioned in steps 1 and 2 above, the desired vehicle location (the road) is a function of the distance travelled. However, it can assume them to be functions of time, instead:

$${}^G\mathbf{r}_d = {}^G\mathbf{r}_d \; (X_d(t), Y_d(t), Z_d(t)) \, , \, {}^G\mathbf{r}_c = {}^G\mathbf{r}_c \; (X_c(t), Y_c(t), Z_c(t)) \tag{12.95}$$

This assumption defines the ideal vehicle as to the one running at a constant given velocity v_{x0} on the road, without any offset. In other words, the ideal vehicle is considered a kinematic model which can negotiate any turn regardless of the velocity and the road curvature. Hence, the effect of dynamic turning is not included for the ideal vehicle motion. Such a concept for a moving vehicle is called a ghost-car approach. The autodriver algorithm's target to follow the ideal ghost-car on the road and minimise the offset at any time. Figure (a) visualises the ghost-car and the error definitions, all expressed in the vehicle's body coordinate frame.

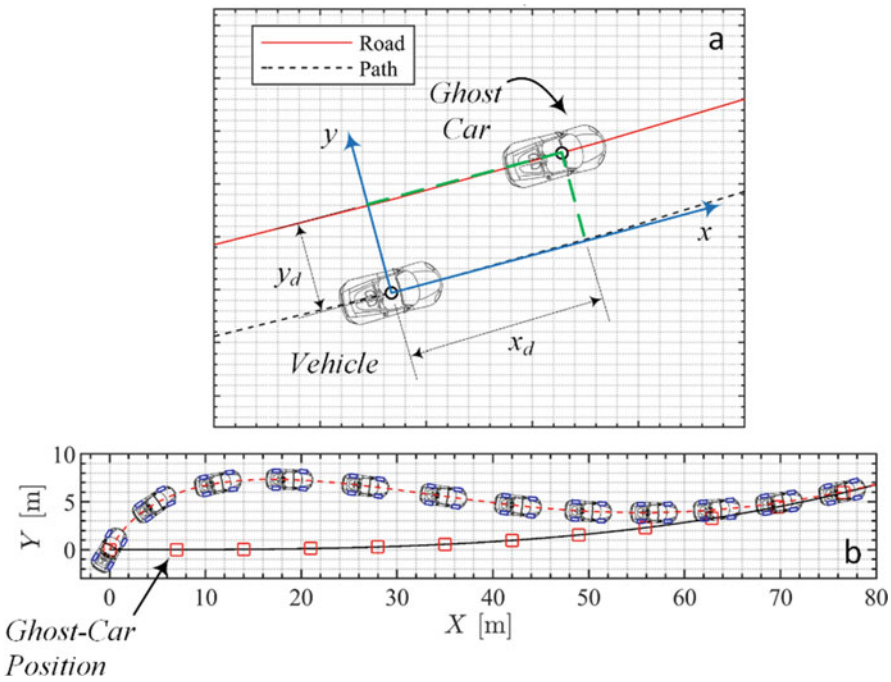


Fig. 12.30 Error values expressed in body frame

An example of a path-following strategy using the ghost-car approach is shown in Fig. 12.30.

12.5 Control

In terms of control terminology, the control layout consists of a feedforward part working based on the steady-state vehicle behaviour, and two feedback loops to compensate for transient errors and keeping the vehicle on the desired point of the road at any time instance. Figure 12.31 shows the proposed control structure.

The feedforward provides a quick response of the control system based on the available road data. The steady-state behaviour may estimate the vehicle response with considerably good accuracy. Hence, the steady-state behaviour is chosen for the feedforward part of the lateral controller. Inclusion of feedforward assures that the central portion of the control input is provided to the vehicle as soon as it reaches a turn. The transient error, however, needs to be fully compensated in order to achieve a reliable autonomous path-following performance. Due to the high level of accuracy of the feedforward, a simple PD feedback control is sufficient to close the control loop. Two errors are to be minimised, namely, the longitudinal and the

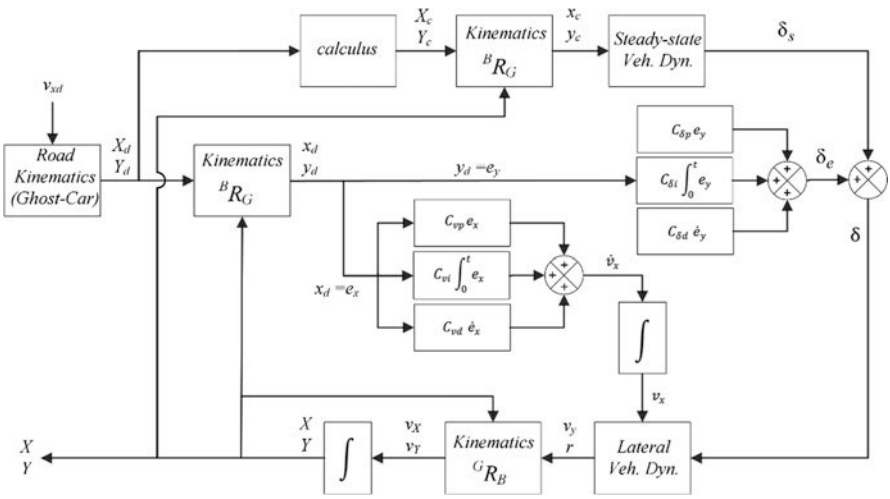


Fig. 12.31 Block diagram of the control system structure

lateral position error expressed in the vehicle body coordinate frame e_x, e_y. The lateral position error is compensated with a correction signal added to the previously calculated steer angle.

Minimisation of the longitudinal position error requires manipulation of the vehicle's longitudinal velocity. Since a sudden change in velocity cannot happen in a real vehicle, the time is taken derivative $\dot{v}_x$ as the control variable for longitudinal adjustment which corresponds to a precise acceleration command to the engine or brakes. The total longitudinal acceleration is given as $a_x = \dot{v}_x - r v_y$, but we only have direct control over $\dot{v}_x$ by throttle/brake command.

Since a position control have problem, the regulated outputs are, in fact, results of the integration of the vehicle's dynamic variables v_x, v_y and r. Hence, a PD controller would be desirable with the equal part having a similar effect as an integrator controller on a first-order system, and the derivative part acting as a proportional controller. This would be equivalent to a PI controller effect on the system. Thus, the PD controller is expected to minimise the steady-state error while the feedforward part provides a quick response. The PD gains are tuned manually. For longitudinal control $C_{vp} = 1$, $C_{vd} = 1$ and for lateral control $C_{\delta p} = 0.01$, $C_{\delta d} = 0.025$.

12.5.1 Simulation Results

Investigating the effectiveness of the proposed path-following controller, two scenarios are simulated with different road geometries. It is worth to note that an ideal transition between two sections of a road with constant curvatures is achieved

using a linearly changing curvature function, rather than connecting the two sections by making them tangent to each other. Such a linearly changing curvature (e.g. clothoid) provides a linearly changing steer angle requirement with no jump, which is physically feasible. The clothoid is defined using Fresnel integrals as follows [14]:

$$X = a\int_0^t \cos\left(\frac{\pi}{2}u^2\right)du = a\int_0^{s/a} \cos\left(\frac{\pi}{2}u^2\right)du \tag{12.96}$$

$$Y = a\int_0^t \sin\left(\frac{\pi}{2}u^2\right)du = a\int_0^{s/a} \sin\left(\frac{\pi}{2}u^2\right)du \tag{12.97}$$

where the parameter a is called the scaling parameter and determines the size of the clothoid and t is the independent variable which may be converted to arc length $s = at$ for simplicity. By using clothoid in road geometry construction, it is expected a smoother performance of the driver's steering control in case of a linear curvature change.

12.5.2 Figure-8 Road

The first scenario investigates the controller's performance on a road which is designed based on transitions between zero and constant curvature ($\kappa = 1/\rho = 0.02$) sections forming an 8-shaped geometry. Such road geometry is a desirable example to evaluate the controller. It includes instantaneous straight parts at the origin while having constant curvature turns (circular arcs) at both ends. The transition between left and right turns as well as driving into turns is examined at the same time using this road geometry. The transitions are achieved using clothoids, keeping the curvature of the road continuous throughout the road. The clothoid is constructed based on the scaling of $a = 147.195$ and total arc length of $s_1 = 137.9$ m.

Figure 12.32 shows the road geometry and the vehicle's path of motion, starting from the origin towards positive x-axis. The initial and desired velocity of the vehicle is set at 20 m/s. The path of motion shows high accuracy of path following achieved by the autodriver algorithm.

The required steer angle and the instantaneous velocity of the vehicle are shown in Fig. 12.32. It is observed that the steering angle is continuous with negligible fluctuations to overcome the transient error. The velocity of the vehicle is also kept very close to the desired value (20 m/s) with minimal variations to compensate for the longitudinal error for the Ghost-Car (Fig. 12.33).

Figure 12.34 shows the time history of the dynamic variables of the system. The continuous and proportional increase in the variables shows a smooth control performance. Negative side-slip angle β and large roll angle of $\phi = 5$ *deg* show relatively high velocity condition and sharpness of the turn corresponding to the

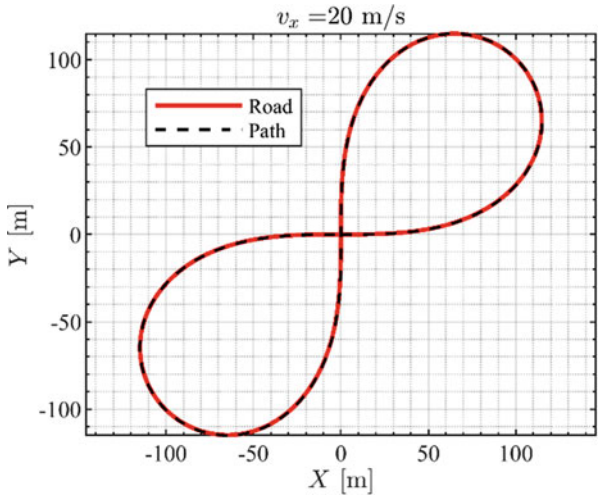


Fig. 12.32 Figure-8 road and vehicle's path of motion

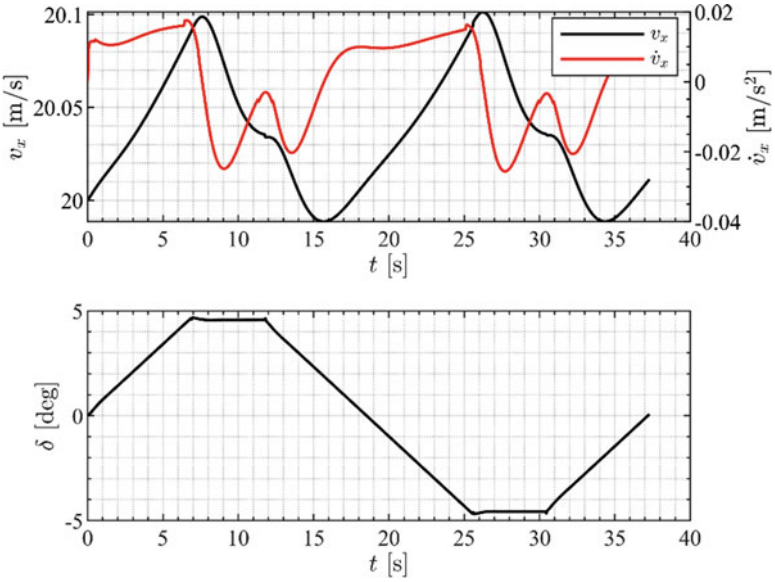


Fig. 12.33 Provided velocity and steer angle for figure-8 road

vehicle running at 20 m/s, respectively. The small roll rate p history also proves the smooth performance of the controller, which provides more comfort to the passengers during such a tight turn. The yaw velocity r is also consistent with the road geometry and do not show fluctuations, maximising passenger comfort.

The longitudinal and lateral error values in the body frame are depicted in Fig. 12.35. The longitudinal error is negligible, which is handled by the velocity controller. It shows that the vehicle successfully followed the ghost-car on the

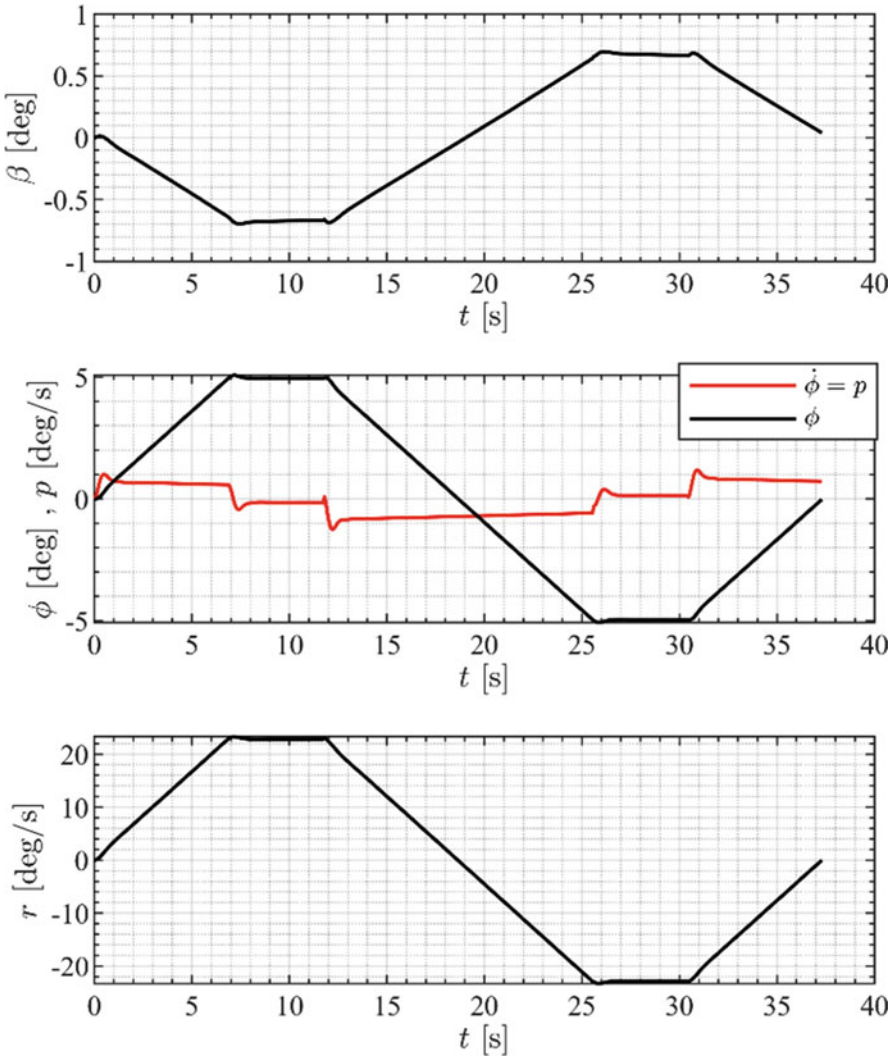


Fig. 12.34 Vehicle's dynamic variables for figure-8 road

road with a minor error. The lateral error Δy gets a maximum of 0.2 m during the manoeuvre which is considered an outstanding path-following performance at a velocity of 20 m/s and a sharp turn with a radius of 50 m. Assuming a lane width of 3–3.5 m, the typical passenger vehicles have approximately 0.5–0.75 m from each side to reach the lane margin which is well above the error range achieved.

The location of ICR in body frame attached to the vehicle is shown in Fig. 12.35c. The green loci show the desired points for ICR and the blue loci indicate the actual

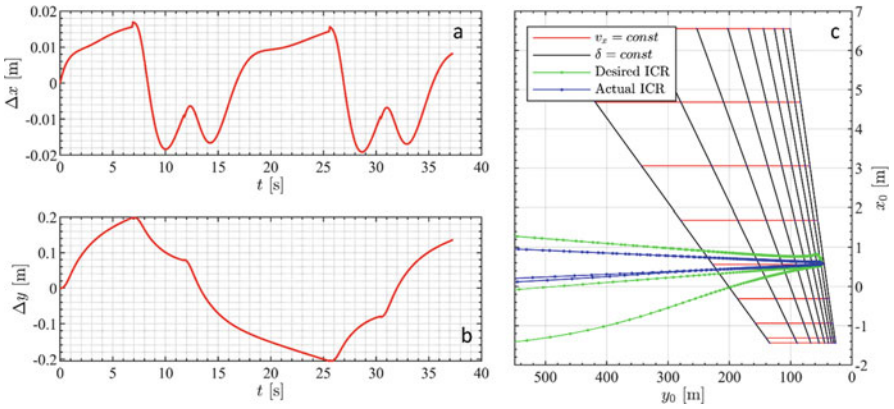


Fig. 12.35 Following errors in body frame for figure-8 road

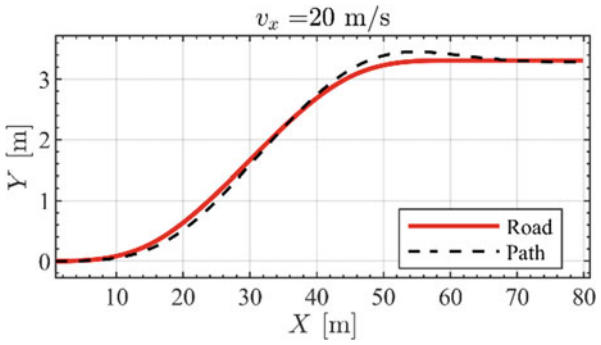


Fig. 12.36 Lane-change virtual road and vehicle path

locations of ICR during the manoeuvre. It is observed that both curves converge to the same location in steady-state with minor deviation.

12.5.3 Lane Change Manoeuvre

Another critical manoeuvre for evaluation of the path-following control is the lane-change manoeuvre. Such a manoeuvre is very popular in the literature and also a quite common practice in every-day driving as well as obstacle-avoidance scenarios. To model the desired path rationally, the lane-change is constituted a virtual road with four different sections created using clothoid and the final section made of a straight line (Fig. 12.36). The clothoid is constructed based on the scaling of $a = 80$ and total arc length of $s_1 = 15$ m. Autodriver algorithm successfully drove the vehicle on the intended path at a longitudinal velocity of 20 m/s, causing the vehicle to displace 3.5 m laterally within 60 m of distance travelled.

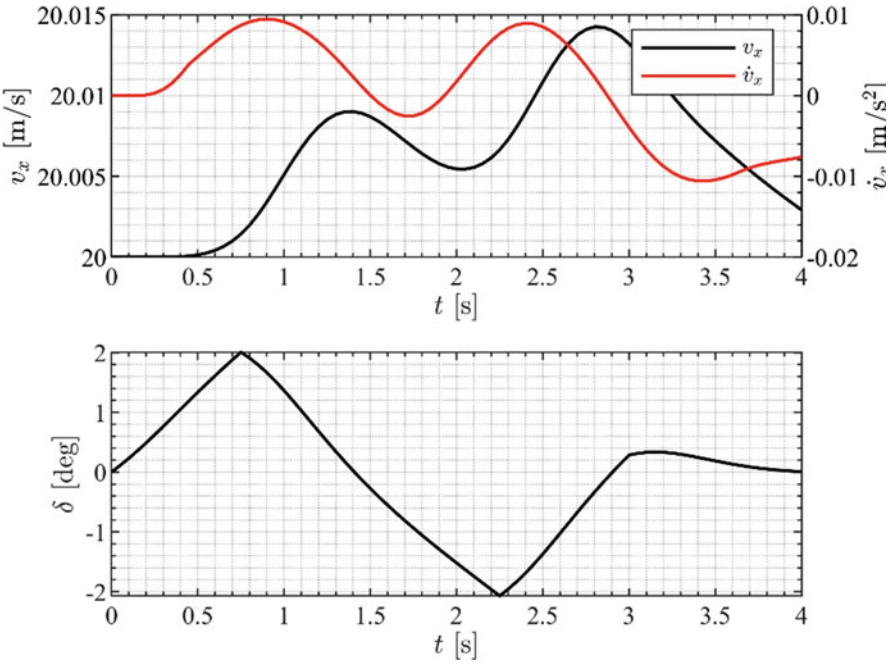


Fig. 12.37 Provided velocity and steer angle for lane-change manoeuvre

Figure 12.36 shows the velocity adjustment by the controller to follow the ghost-car as well as the required steer angle. It is observed that due to short distance travelled and small curvature of the path, the required adjustment in a longitudinal acceleration of the vehicle is negligible, which is a desirable condition from a comfort perspective. The steer angle history shows continuous behaviour proving smooth operation of the system. It is important to note that the selected road geometry imposes the proportional-like behaviour of the steering angle with linearly changing curvature and it does not represent the only way of performing a lane-change.

Vehicle response is depicted in Fig. 12.37 in terms of its dynamic variables. Negative side-slip angle appeared because of relatively high vehicle velocity while the maximum roll angle is relatively small due to small curvature of the path. The roll rate p, however, is larger than the previous manoeuvre because of the sudden nature of the lane-change manoeuvre. In other words, the vehicle is steered back and forth before it settles at a steady-state. Yaw velocity behaviour of the vehicle is smooth, proving a desirable performance in terms of passenger comfort.

Longitudinal and lateral errors between the actual and desired vehicle location throughout the manoeuvre are shown in Fig. 12.38a–b. The longitudinal error is negligible due to small curvature and short travel distance, proving excellent ghost-car following ability of the system. The lateral error is also negligible at such a relatively high speed. The maximum lateral error is less than 0.2 m in the returning

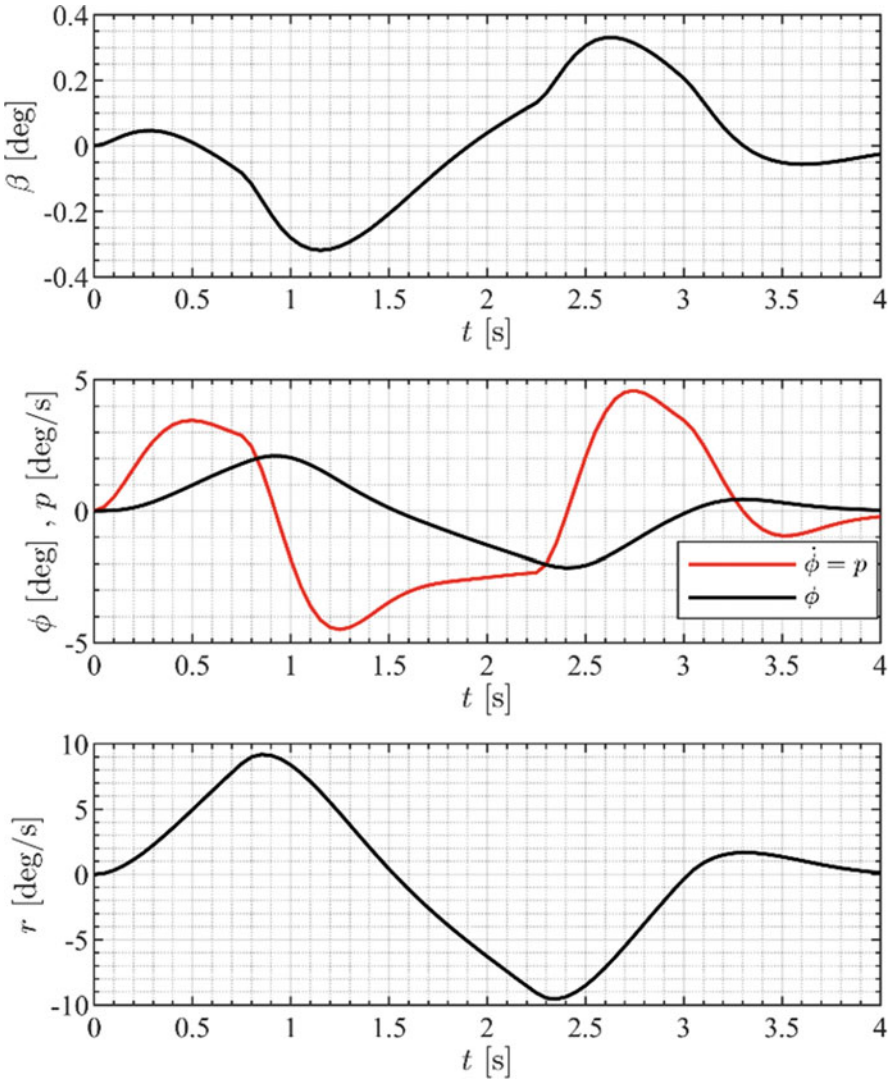


Fig. 12.38 Vehicle's dynamic variables for lane-change manoeuvre

cycle (second half of the lane-change action, approximately between 0 to 30 m travelled). This is due to the nonzero initial yaw and lateral velocities in the return cycle (from 30 m to 60 m). In other words, the second stage of lane-change will result in an overshoot in the following lateral error.

The ICR loci in body coordinate are plotted and shown in Fig. 12.39c. The transient nature of the manoeuvre is observed in this figure, too. The ICR does not reach the desired (steady-state) location during the lane change, but the difference is

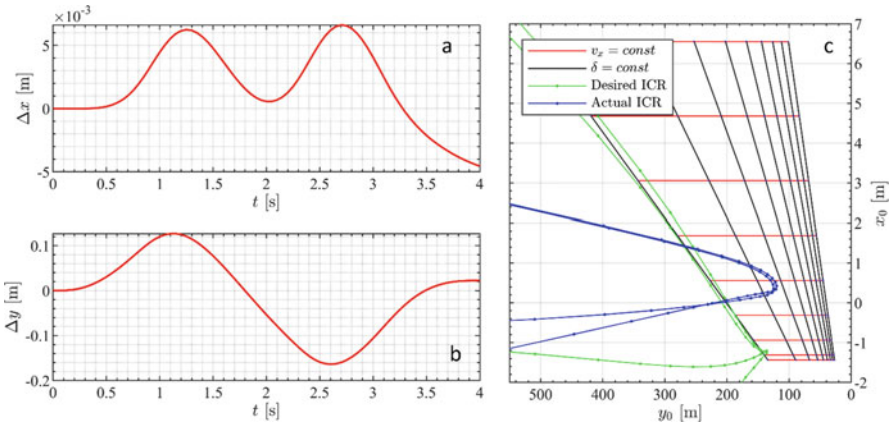


Fig. 12.39 Following errors in body frame for lane-change manoeuvre

small, proving the effectiveness of using the steady-state response as feed-forward and overall control performance, even at sudden manoeuvres.

12.6 Conclusion

In conclusion, in order to cope with the practical implementations, an autodriver algorithm was improved to adjust vehicle position on the desired road with a feed-forward feedback control algorithm. The steady-state lateral response was used to accelerate the controller performance in terms of steering, while the transient error was eliminated with feedback control. Meanwhile, the longitudinal error was determined using a ghost-car concept to follow the road at the desired speed by manipulating the longitudinal acceleration of the vehicle via feedback control.

It is worth to mention that any small error between the locations of the vehicle and the road will be amplified in the location error of ICR; hence, by starting from ICR placement and then adjusting transient position errors, an accurate and quick performance was achieved which is critical for autonomous vehicle control.

Simulations are performed for a figure-8 and lane-change manoeuvres, which are comprehensive indicators of different driving conditions. The results show promising performance of the algorithm in terms of path-following, even under sudden manoeuvering requirements. The algorithm was evaluated using a more realistic vehicle model (roll model), including the roll motion. Observation of roll behaviour also proved the desirable performance of the controller in terms of passenger comfort, proving the efficiency of the proposed algorithm both for path-following and passenger comfort. Ideally to minimise the roll rate, a vehicle can control the roll so that human can enjoy the comfort inside the vehicle which fulfils the objective of the research.

References

1. R. Bishop, *Intelligent Vehicle Technology and Trend*. Norwood, MA: Artech House, 2005.
2. M. A. Sotelo, "Lateral control strategy for autonomous steering of Ackerman-like vehicles," *Robot. Auton. Syst.,* vol. 45, pp. 223–233, 2003.
3. J. M. Snider, "Automatic Steering Methods for Autonomous Automobile Path Tracking," Ph.D. dissertation, Robotics Institute, Carnegie Mellon University, Pittsburgh, 2009.
4. V. Milanés, J. Pérez, E. Onieva, C. González, and T. de Pedro, "Lateral power controller for unmanned vehicles," *Elect. Rev.,* vol. 86, pp. 207–211, 2010.
5. J. Pérez, V. Milanés, and E. Onieva, "Cascade architecture for lateral control in autonomous vehicles," *IEEE T. Intell. Transp. Syst.,* vol. 12, pp. 73-82, 2011.
6. J. W. Lee and B. Litkouhi, A unified framework of the automated lane centreing/changing control for motion smoothness adaptation, presented at the International IEEE Conference on Intelligent Transportation Systems, Anchorage, Alaska, 2012.
7. M. H. Lee *et al.*, "Lateral controller design for an unmanned vehicle via kalman filtering," *Int. J. Auto. Tech.,* vol. 13, pp. 801–807, 2012.
8. A. Broggi, P. Medici, E. Cardarelli, P. Cerri, and A. Giacomazzo, Development of the control system for the VisLab intercontinental autonomous challenge, presented at the International IEEE Annual Conference on Intelligent Transportation Systems, Madeira Island, 2010.
9. H. Tan and J. Huang, "Experimental development of a new target and control driver steering model based on DLC test data," *IEEE T. Intell. Transp. Syst.,* vol. 13, pp. 375–384, 2012.
10. I. Bae, J. Hyo Kim, and S. Kim, Steering rate controller based on curvature of path for autonomous driving vehicles, presented at the IEEE Intelligent Vehicles Symposium (IV), Australia, 2013.
11. R. Rajamani, *Lateral Vehicle Dynamics*, Springer, 2006.
12. M. Elbanhawi, M. Simic, and R. N. Jazar, "Autonomous robots path planning: An adaptive roadmap approach," *Appl. Mech. Mater.,* vol. 373, pp. 246–254, 2013.
13. M. Elbanhawi, M. Simic, and R. N. Jazar, "Continuous-curvature bounded path planning using parametric splines," *FAIA,* vol. 262, pp. 513–522, 2014.
14. H. Marzbani, M. Simic, M. Fard, and R. N. Jazar, "Better road design for autonomous vehicles using clothoids," *IIMSS,* vol. 40, pp. 265–278, 2015.
15. D.Q. Vo, H. Marzbani, M. Fard, and R. Jazar, "A novel kinematic model of a steerable tire for examining kingpin moment during low-speed-large-steering-angle cornering," *SAE Int. J. Passeng. Cars Mech. Syst.*, vol. 10, 2016.
16. D.Q. Vo, H. Marzbani, M. Fard, and R. Jazar, "Variable caster steering in vehicle dynamics," in *Proc. Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 2017.
17. M. Guiggiani, *The Science of Vehicle Dynamics: Handling, Braking, and Ride of Road and Race Cars*, Springer, 2014.
18. C. Fu, R. Hoseinnezhad, A. Bab-Hadiashar, and G. Nakhaie Jazar, "Electric vehicle side-slip control via electronic differential," *Int. J. Veh. Auton. Syst.,* vol. 6, pp. 108–132, 2014.
19. H. Marzbani, M. H. A. Salahuddin, M. Simic, M. Fard, and R. N. Jazar, "Steady-state dynamic steering," *FAIA,* vol. 262, pp. 493–504, 2014.
20. P. K. Agarwal and H. Wang, "Approximation algorithms for curvature-constrained shortest paths," *SIAM J. Comput.*, vol. 30, pp. 1739–1772, 1996.
21. H. Marzbani, D. Q. Vo, A. Khazaei, M. Fard, and R. N. Jazar, "Transient and steady-state rotation centre of vehicle dynamics," *IJNDC,* vol. 1, pp. 97–113, 2017.
22. A. Lari, F. Douma, and I. Onyiah, "Self-driving vehicles and policy implications: Current status of autonomous vehicle development and Minnesota policy implications," *Minn. J. L. Sci. & Tech.,* vol. 16, 2015.
23. H. Khayyam, A. Z. Kouzani, E. J. Hu, and S. Nahavandi, "Coordinated energy management of vehicle air conditioning system," *Appl. Therm. Eng.*, vol. 31, pp. 750–764, 2011.
24. S. Song, W. Cai, and Y-G. Wang, "Auto-tuning of cascade control systems," *ISA Trans.*, vol. 42, pp. 63–72, 2003.

25. R. N. Jazar, "Mathematical theory of autodriver for autonomous vehicles," *J. Vib. Control*, vol. 16, pp. 253–279, 2010.
26. H. Marzbani, "Application of the mathematical autodriver algorithm for autonomous vehicles," Ph.D dissertation, Engineering, RMIT University, Melbourne, Australia, 2014.
27. A. Bourmistrova, M. Simic, R. Hoseinnezhad, and R. N. Jazar, "Autodriver algorithm," *J. Syst., Cyber. Inform.,* vol. 9, pp. 55–56, 2011.
28. H. Khayyam, A. Kouzani, H. Hamid Abdi, and S. Nahavandi, "Modeling of Highway Heights for Vehicle Modeling and Simulation," in *ASME 2009 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2009, pp. 365–369.
29. H. Khayyam, "Stochastic models of road geometry and wind condition for vehicle energy management and control," *IEEE Trans. Vehic. Tech.,* vol. 62, pp. 61–68, 2012.
30. R. N. Jazar, *Advanced Dynamics: Rigid Body, Multibody, and Aerospace Applications*, New York: Wiley, 2011.
31. H. Marzbani, S. Harithuddin, M. Simic, M. Fard, and R. N. Jazar, "Four wheel steering advantageous for the Autodriver algorithm," *FAIA,* vol. 262, pp. 505–512, 2014.
32. R. N. Jazar, *Vehicle Dynamics: Theory and Application*. New York: Springer, 2018.
33. J. M. S. Ribeiro, M. F. Santos, M. J. Carmo, and M. F. Silva, "Comparison of PID controller tuning methods: Analytical/classical techniques versus optimization algorithms" in *18th International Carpathian Control Conference*, pp. 533–538, 2017.

Chapter 13
Oil Leakage Analysis for an Active Anti-Roll Bar System of Heavy Vehicles

Vu Van Tan, Vo Quoc Dai, Olivier Sename, and Reza Nakhaie Jazar

Abstract Vehicle rollover is an important road safety problem world-wide. Although rollovers are relatively rare events, they are usually deadly accidents when they occur. The roll stability loss is the main cause of rollover accidents in which heavy vehicles are involved. In order to improve the roll stability, most of modern heavy vehicles are equipped with passive anti-roll bars to reduce roll motion during cornering or riding on uneven roads. However, these may be not sufficient to overcome critical situations. The active anti-roll bar system is considered as the most common method in order to improve the roll stability of heavy vehicles. In this chapter, the authors are interested in the effects of the internal leakage inside the electronic servo-valve on the performance of the active anti-roll bar system of heavy vehicles. Hence, the main contents are summarized in the following points:

- The internal leakage inside the electronic servo-valve is analysed in detail and characterized by the total flow pressure coefficient. Thanks to this leakage, it is important that the active anti-roll bar system can act in a self-protection capacity when the controller fails.
- An H_∞/LPV active anti-roll bar controller for the fully integrated model is synthesized by using the grid-based LPV approach. Here, the forward velocity

V. V. Tan (✉)
Department of Automotive Mechanical Engineering, Faculty of Mechanical Engineering, University of Transport and Communications, Hanoi, Vietnam
e-mail: vvtan@utc.edu.vn

V. Q. Dai
Faculty of Vehicle and Energy Engineering, Le Quy Don Technical University, Hanoi, Vietnam
e-mail: dai.voquoc@lqdtu.edu.vn

O. Sename
Université Grenoble Alpes, CNRS, Grenoble INP (Institute of Engineering), GIPSA-lab, Grenoble, France
e-mail: olivier.sename@grenoble-inp.fr

R. N. Jazar
School of Engineering, RMIT University, Melbourne, VIC, Australia
e-mail: reza.jazar@rmit.edu.au

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_13

is considered as the varying parameter to adapt to the different types of heavy vehicle movements.
- The simulation results indicate that the internal leakage inside the electronic servo-valve drastically affects the characteristics of the closed-loop system. The two main objectives (enhancing roll stability and avoiding the saturation of the actuators) are simultaneously satisfied when the total flow pressure coefficient K_P is chosen in the interval $[5 \times 10^{-15}, 4 \times 10^{-10}] \frac{m^5}{Ns}$.
- This analysis is the basis for further studies of the fault tolerant control and fault accommodation on the active anti-roll bar system of heavy vehicles using the ESVH actuators in order to improve the performance of the active anti-roll bar system.

13.1 Introduction

13.1.1 Rollover of Heavy Vehicles

Heavy vehicles are defined as goods vehicles having a maximum permissible gross weight (vehicle and load of over 3.5 tons), and includes trucks and buses [1]. The use of heavy vehicles as a means of transport is of economic importance in most areas of the world. In the United States, for example, the commercial trucking sector employs nearly 10 million people and has annual revenues of more than US$500 billion. Trucks transport over 11 billion tons of goods annually, about 60% of the total domestic tonnage shipped [2]. Additionally, North American intercity and charter buses carry an estimated 860 million passengers annually, which is more than those transported by commercial air carriers or rail [3].

We have to acknowledge the role of heavy vehicles in economic development. However, due to their high mass, there can be severe consequences for other road users when they are involved in accidents. The accidents related to heavy vehicles are also a complex issue not just in developing countries but also in developed countries like the USA and Europe. The rollover phenomenon is the most dangerous type of accident for heavy vehicles, although rollovers are relatively rare events, they can be deadly accidents when they occur. Loss of roll stability is the main cause of rollover accidents involving heavy vehicles. According to the Federal National Highway Traffic Safety Administration (NHTSA), in the United States, there were 333,000 heavy vehicles involved in traffic crashes during 2012. There were 3921 people killed in rollover crashes and 104,000 people injured (an increase of 18% from 2011) [4]. In 2013, more than 4500 persons were killed in road traffic accidents involving heavy vehicles in the EU, constituting almost 18% of all road accident fatalities for that year [1]. While heavy vehicles account for just a small proportion of the vehicle fleet or the total vehicle kms travelled in the EU, they are more often involved in severe road accidents, creating a significant need to better understand the physical and mechanical characteristics specific to this vehicle group. Figure 13.1 illustrates two examples of the heavy vehicle rollover phenomenon.

Fig. 13.1 Rollover of a single unit heavy vehicle [8]

13.1.2 Different Categories of Vehicle Rollover Accidents

The rollover of heavy vehicles is a very serious safety issue, which can result in large financial and environmental consequences. Rollover accidents are classified into four categories [5]:

- **Preventable:** the driver would have been able to avoid the accident if a warning device had been installed on the vehicle. This is closely linked to the "Time-To-Rollover (TTR)" concept [6, 7]. The survey results indicated that this category accounts for about 3.3% of total rollover accidents.
- **Potentially preventable:** the accident might have been avoided, depending on the driver skills and on the performance of the warning device. It accounts about 38.4% of total rollover accidents.
- **Non-preventable:** regardless of the driver skills or the performance of the warning device, vehicle rollover will always happen. This group accounts about 49.7% of total rollover accidents.
- **Preventable unknown:** this category accounts for about 8.6% of total rollover accidents.

It is usually difficult for the driver to feel the rollover behaviour of a heavy vehicle. Investigations have shown that only a minority of rollover accidents could have been avoided with a warning device, potentially more with a skilled driver, but half of the rollover accidents were not preventable by driver action alone. The main cause of rollover accidents in which heavy vehicles are involved is the loss of roll stability when the tyre-road contact force on one of the side wheels becomes zero. Roll stability refers to the ability of a vehicle to overcome overturning moments generated during cornering and lane changing. It is well known that heavy vehicles do have relatively high centres of mass and narrow track widths and can lose roll stability at moderate levels of lateral acceleration [9].

Until now, most of heavy vehicles is using a passive anti-roll bar system in order to reduce body roll motion, the driving safety and roll stability are highly improved. However, passive anti-roll bars do have drawbacks. During cornering manoeuvres, the anti-roll bar will transfer the vertical forces of one side of the suspension to the other one, creating therefore a yaw moment. In order to overcome such drawbacks, several schemes concerned with the possible active intervention onto the vehicle dynamics have been proposed as follows: active steering, active brake, active suspension and active anti-roll bars. Among them, the most common method used to improve the roll stability is the active anti-roll bar.

13.1.3 Active Anti-Roll Bar System

Active roll control systems for heavy vehicles have been studied for more than two decades, and it has been proven that roll stability can be improved and rollover can be prevented in an emergency situation. Among active roll control systems, the most common methods used to improve roll stability are active anti-roll bar systems which are usually made with a pair of hydraulic actuators. Indeed lateral acceleration makes vehicles with a conventional passive suspension tilt out in corners. The centre of the sprung mass shifts outward of the vehicle centreline, which then, on the other hand, creates a destabilizing moment that degrades roll stability. The lateral load response is reduced by an active anti-roll bar system that generates a stabilizing moment to counterbalance the overturning moment in such a way that the control torque leans the vehicle into the corners. Figure 13.2 describes the ideal structure of an active anti-roll bar system applied on one axle of a heavy vehicle.

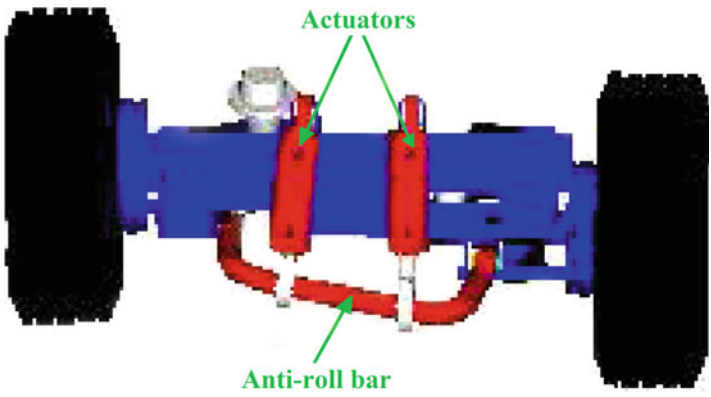


Fig. 13.2 An active anti-roll bar system applied on one axle of heavy vehicles [10]

13.1.4 Oil Leakage of the Electronic Servo-Valve

The Electronic Servo-Valve Hydraulic (ESVH) actuator is the most commonly used in the industrial sector and engineering practices. The ESVH actuator includes the main elements such as: (1) a servo regulator (controller), (2) a servo-valve, (3) a hydraulic cylinder, (4) a feedback position transducer and (5) a power supply, as shown in Fig. 13.3. The dynamic response of the system mainly builds upon the frequency characteristics of the servo-valve and the load, but this overall quality of the servo control could be impaired by faults in the system. The fault usually includes the internal and external leakages [11, 12]. The external leakage can be detected visually, so this leakage can be easily perceived. Conversely, the internal leakage which occurs inside the hydraulic cylinder and the electronic servo-valve, cannot be detected easily, and therefore its detection is much more important. This internal leakage is caused by excessive wear of the piston seal or abrasion. This seal prevents leakage and closes the gap between the piston and the cylinder wall. The internal leakage inside the hydraulic cylinder can be detected when the actuator seal is almost completely destroyed [12, 13]. Notice that, the internal leakage inside the servo-valve always exists for all the ESVH actuators, even if they are completely new [14, 15].

In the literature, many studies have been dedicated to examine the internal leakage inside the electronic servo-valve of the ESVH actuator. In [16], the authors combined the internal leakage and the orifice flows in a hydraulic servo-valve model. It indicated that with small spool valve displacements, the internal leakage flow between the spool valve and the body dominates the orifice flow through the valve. Kalyoncu et al. [15] developed the mathematical model of the ESVH system considering the internal leakage within the electronic servo-valve. The

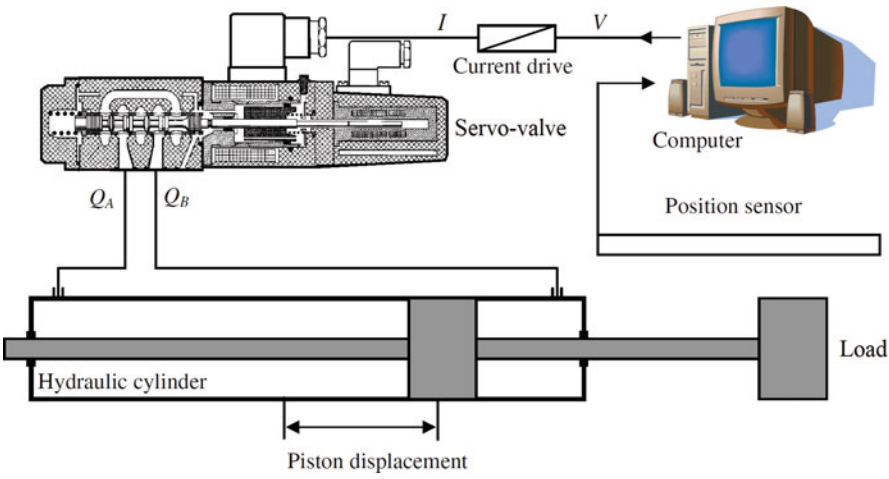


Fig. 13.3 Structure of the general electronic servo-valve hydraulic system [15]

simulation results showed that the leakage occurring during the small spool valve displacements does affect the mathematical model and position control of the system. By considering the internal leakage inside the electronic servo-valve, further improvements in the ESVH system performance are achieved. In [17, 18] and [19], the authors proposed the position controller for the ESVH system. The proposed scheme has the ability to improve the position tracking performance of the actuator in the presence of friction and internal leakage inside the electronic servo-valve.

13.2 An Electronic Servo-Valve Hydraulic Actuator Model

An Electronic Servo-Valve Hydraulic (ESVH) system is one of the most important drive systems in the industrial sector and most engineering practices due to its high power to weight ratio, stiffness response, good performance, and smooth fast action. The range of applications for ESVH systems are diverse, and includes: manufacturing systems, material test machines, active systems on vehicles, mining machinery, fatigue testing, flight simulation, marine engineering, robotics, etc [17].

The two elements of the ESVH actuator (the electronic servo-valve, the hydraulic cylinder) will be modelled in the sequel.

13.2.1 The Electronic Servo-Valve Model

The three-land-four-way spool valve is used in the ESVH actuator with the diagram shown in Fig. 13.4. The displacement of the spool valve X_v is controlled by the electrical current u. The effects of hysteresis and flow forces on the servo-valve are neglected here, then the dynamical behaviour of the electronic servo-valve can be approximated by a first-order model [14, 20], as:

$$\frac{dX_v}{dt} + \frac{1}{\tau}X_v - \frac{K_v}{\tau}u = 0 \tag{13.1}$$

where τ is the time constant and K_v the gain of the servo-valve model.

The diagram of the flows direction of the three-land-four-way spool valve is shown in Fig. 13.5. The four orifices are completely analogous to the four arms of a Wheatstone bridge. Arrows at the ports indicate the assumed direction of the different flows, and the numbers at the ports refer to the subscripts of the flow and the area at the ports. Because the compressibility flows are zero therefore the continuity equations for the two valve chambers are:

$$Q_L = Q_1 - Q_4 = Q_3 - Q_2 \tag{13.2}$$

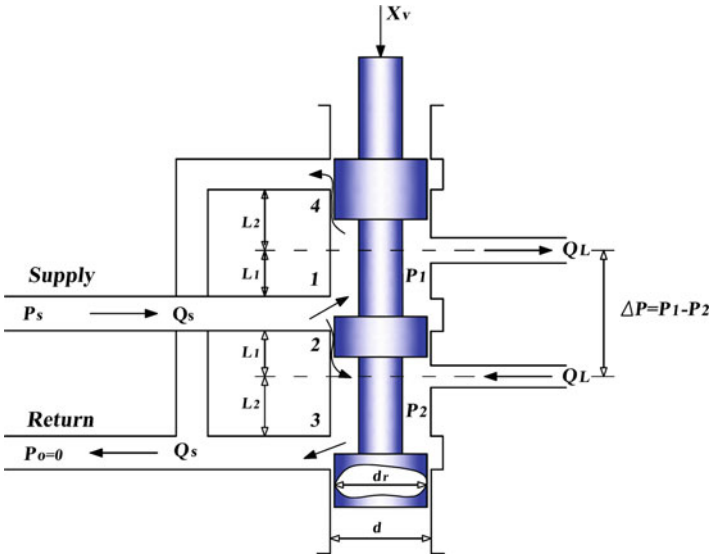


Fig. 13.4 Diagram of the three-land-four-way spool valve [21]

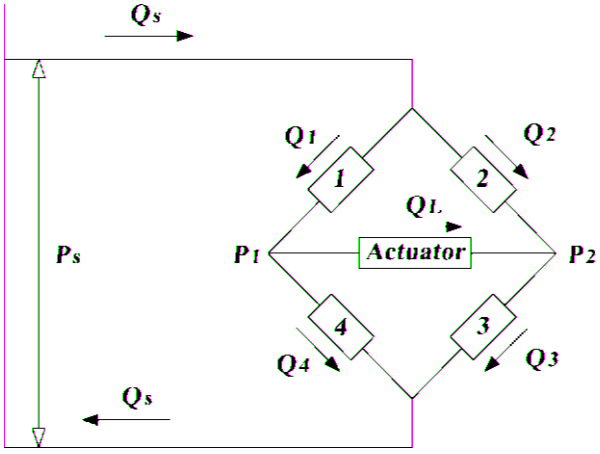


Fig. 13.5 Diagram of the flow directions of the three-land-four-way spool valve [21]

where Q_L is the load flow through the actuator, and $Q_{i(i=1,4)}$ the load flow through the orifices, respectively.

The differential pressure inside the hydraulic cylinder is defined as:

$$\Delta P = P_1 - P_2 \tag{13.3}$$

The different flows through the valve orifices are described as follows [21]:

$$\begin{cases} Q_1 = C_d A_1 \sqrt{\frac{2}{\rho}(P_s - P_1)} \\ Q_2 = C_d A_2 \sqrt{\frac{2}{\rho}(P_s - P_2)} \\ Q_3 = C_d A_3 \sqrt{\frac{2}{\rho} P_2} \\ Q_4 = C_d A_4 \sqrt{\frac{2}{\rho} P_1} \end{cases} \tag{13.4}$$

where ρ is the fluid mass density, A_i the orifice areas, and C_d the discharge coefficient.

The orifice areas (A_1, A_2, A_3, A_4) depend on the valve geometry and they are function of the spool valve displacement X_v as follows:

$$\begin{cases} A_1 = A_1(X_v) \\ A_2 = A_2(-X_v) \\ A_3 = A_3(X_v) \\ A_4 = A_4(-X_v) \end{cases} \tag{13.5}$$

Remark 1 The valve orifices are usually matched and symmetrical. The requirement for the matched orifices is: $A_1 = A_3, A_2 = A_4$; and for the symmetrical orifices are: $A_1(X_v) = A_2(-X_v), A_3(X_v) = A_4(-X_v)$.

Due to the orifices being symmetrically matched, the different flows in the diagonally opposite arms of the bridge in Fig. 13.4 are:

$$\begin{cases} Q_1 = Q_3 \\ Q_2 = Q_4 \end{cases} \tag{13.6}$$

From Eqs. (13.4) and (13.6), the oil supply high pressure is

$$P_s = P_1 + P_2 \tag{13.7}$$

The pressures inside the two chambers of the hydraulic cylinder are determined from Eqs. (13.3) and (13.7) as follows:

$$\begin{cases} P_1 = \frac{P_s + \Delta P}{2} \\ P_2 = \frac{P_s - \Delta P}{2} \end{cases} \tag{13.8}$$

From Eqs. (13.2), (13.4), (13.6) and (13.8), for a matched and symmetrical servovalve, the load flow through the hydraulic cylinder (Q_L) is defined as:

$$Q_L = C_d A_1 \sqrt{\frac{1}{\rho}(P_s - \Delta P)} - C_d A_2 \sqrt{\frac{1}{\rho}(P_s + \Delta P)} \tag{13.9}$$

The nonlinear algebraic equation (13.9) can be written as follows [21, 22]:

$$Q_L = K_x X_v - K_P \Delta P \tag{13.10}$$

where the valve flow gain K_x and the total flow pressure coefficient K_P are defined as:

$$K_x = \frac{\delta Q_L}{\delta X_v}, \quad K_P = \frac{\delta Q_L}{\delta \Delta P} \tag{13.11}$$

In Eq. (13.10), the load flow of the ESVH actuator includes two parts: the first part ($Q_{L1} = K_x X_v$) is the orifice load flow through the servo-valve and is adjusted by the movement of the spool valve displacement X_v. The second part ($Q_{L2} = K_P \Delta P$) is the internal leakage load flow through the contact surface between the spool valve and the body of the servo-valve.

13.2.2 The Hydraulic Cylinder Model

Figure 13.6 illustrates the diagram of a hydraulic cylinder in combination with an electronic servo-valve. The spool valve of the electronic servo-valve is controlled by a current which generates a displacement X_v. The high pressure oil supply P_s is always stored outside the electronic servo-valve and the moving spool valve distributes the high pressure oil into two chambers of the hydraulic cylinder. The

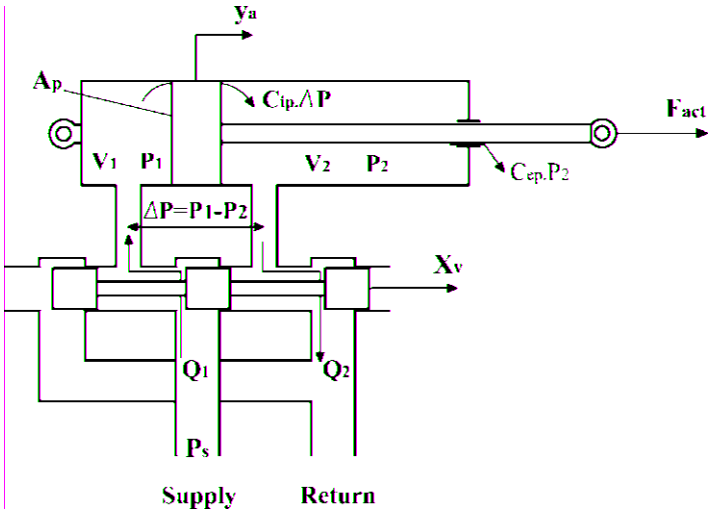


Fig. 13.6 Diagram of the ESVH actuator [10]

difference of pressure $\Delta P = P_1 - P_2$ between the two chambers produces the output force F_{act} given by

$$F_{act} = A_P \Delta P \tag{13.12}$$

where A_P is the area of the piston.

The equations for each chamber of the hydraulic cylinder can be written as:

$$\begin{cases} \frac{dV_1}{dt} + \frac{V_1}{\beta_e}\frac{dP_1}{dt} = Q_1 - C_{ip}(P_1 - P_2) - C_{ep}P_1 \\ \frac{dV_2}{dt} + \frac{V_2}{\beta_e}\frac{dP_2}{dt} = C_{ip}(P_1 - P_2) - C_{ep}P_2 - Q_2 \end{cases} \tag{13.13}$$

where β_e is the effective bulk modulus of the oil, C_{ep} and C_{ip} are the external and internal leakage coefficients of the hydraulic cylinder.

The volume in each chamber varies with the piston displacement y_a as:

$$\begin{cases} V_1 = V_{01} + A_p y_a \\ V_2 = V_{02} - A_p y_a \end{cases} \tag{13.14}$$

where V_{01} and V_{02} are the initial volumes in each chamber. Assuming that $V_{01} = V_{02} = V_0$, the total volume of trapped oil is given by

$$V_t = V_1 + V_2 = V_{01} + V_{02} = 2V_0 \tag{13.15}$$

Therefore, the equations in each chamber become:

$$\begin{cases} A_p \frac{dy_a}{dt} + \frac{V_0 + A_p y_a}{\beta_e}\frac{dP_1}{dt} = Q_1 - C_{ip}(P_1 - P_2) - C_{ep}P_1 \\ -A_p \frac{dy_a}{dt} + \frac{V_0 - A_p y_a}{\beta_e}\frac{dP_2}{dt} = C_{ip}(P_1 - P_2) - C_{ep}P_2 - Q_2 \end{cases} \tag{13.16}$$

Subtracting the second equation from the first one leads to

$$2Q_L = Q_1 + Q_2 = 2C_{tp}\Delta P + 2A_p \frac{dy_a}{dt} + \frac{V_0}{\beta_e}\frac{d\Delta P}{dt} \tag{13.17}$$

where $C_{tp} = 2C_{ip} + C_{ep}$ is the total leakage coefficient of the hydraulic cylinder.

From Eqs. (13.10) and (13.17), the dynamic equation of the servo-valve hydraulic cylinder is obtained as follows:

$$\frac{V_t}{4\beta_e}\frac{d\Delta P}{dt} + (K_P + C_{tp})\Delta P - K_x X_v + A_P \frac{dy_a}{dt} = 0 \tag{13.18}$$

where y_a is the displacement of the piston inside the hydraulic cylinder.

From the Eqs. (13.1), (13.12), (13.18) the dynamical equations of the ESVH actuator are summarized in Eq. (13.19). Here the input signal is the current u and

Table 13.1 Symbols and parameters of the ESVH actuator [5, 14].

Symbols	Description	Value	Unit
A_P	Area of the piston	0.0123	m^2
K_x	Valve flow gain coefficient	2.5	m^2/s
K_P	Total flow pressure coefficient	4.2×10^{-11}	$m^5/(Ns)$
C_{tp}	Total leakage coefficient of the actuator	0	–
V_t	Total volume of trapped oil	0.0014	m^3
β_e	Effective bulk modulus of the oil	6.89×10^6	N/m^2
τ	Time constant of the servo-valve	0.01	s
K_v	Servo-valve gain	0.0239	m/A

the output is the force F_{act}. The symbols and parameters of the model are shown in Table 13.1.

$$\begin{cases} F_{act} = A_P \Delta P \\ \frac{V_t}{4\beta_e}\frac{d\Delta P}{dt} + (K_P + C_{tp})\Delta P - K_x X_v + A_P \frac{dy_a}{dt} = 0 \\ \frac{dX_v}{dt} + \frac{1}{\tau} X_v - \frac{K_v}{\tau} u = 0 \end{cases} \tag{13.19}$$

Remark 2 As presented later, according to the given actuator model, the torque generated by the active anti-roll bar system at each axle is given by $T = -l_{act} F_{actl} + l_{act} F_{actr}$. Here l_{act} is half the distance between the two actuators, $F_{actl,r}$ the actuator forces on the left and on the right.

13.2.3 Internal Leakage Inside the Electronic Servo-Valve

In Sect. 13.2.1, the internal leakage inside the electronic servo-valve is linearized by the total flow pressure coefficient K_P to model the ESVH actuator. In this section, to further clarify this leakage, the structure and nonlinear characteristics of the servo-valve will be considered in detail.

At small spool valve displacements, the leakage load flow between the spool valve and body of the servo-valve dominates the orifice load flow through the valve. In precision positioning applications, where the servo-valve operates around the null region, this load flow, if ignored, may severely degrade the performance of a conventional servo hydraulic design [16, 18].

Here, we consider an accurate model of the leakage load flow [16], by making a smooth transition between the leakage load flow and the orifice load flow, which would likely improve the precision of the ESVH system design and performance. The model used is a nonlinear servo-valve model that accurately captures the servo-valve leakage behaviour over the whole range of the spool valve displacement. The leakage behaviour is modelled as a turbulent load flow with a load flow area

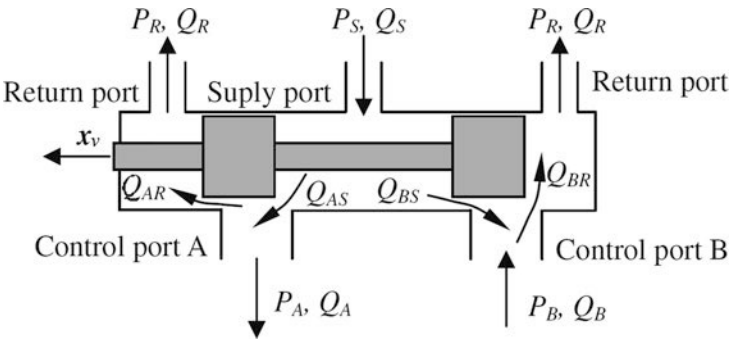


Fig. 13.7 Servo-valve configuration [16]

inversely proportional to the overlap between the spool valve areas and the servo-valve orifices.

Figure 13.7 depicts a servo-valve configuration. This servo-valve consists of two control ports (A, B) with variable orifices which regulate the load flows. The load flow through the control ports of the servo-valve is expressed in Eq. (13.20), and the load flow at the supply and return ports are represented in Eq. (13.21).

$$Q_A = Q_{AS} - Q_{AR} \quad and \quad Q_B = Q_{BR} - Q_{BS} \tag{13.20}$$

$$Q_S = Q_{AS} + Q_{BS} \quad and \quad Q_R = Q_{AR} + Q_{BR} \tag{13.21}$$

By combining the load flow relations in Eqs. (13.20) and (13.21), the supply port load flow in terms of the load flows at the control and return ports is defined as follows:

$$Q_S = Q_R + (Q_A - Q_B) \tag{13.22}$$

The nonlinear load flow relations for the control port **A** are given by the following equations [15]:

$$Q_{AS} = K_{AS}\sqrt{P_S - P_A}.\begin{cases} x_0 + x_v & , x_v \geq 0 \\ x_0^2(x_0 - k_{AS}x_v)^{-1} & , x_v < 0 \end{cases} \tag{13.23}$$

$$Q_{AR} = K_{AR}\sqrt{P_A - P_R}.\begin{cases} x_0^2(x_0 + k_{AR}x_v)^{-1} & , x_v \geq 0 \\ x_0 - x_v & , x_v < 0 \end{cases} \tag{13.24}$$

The nonlinear load flow relations for the control port **B** are given as follows [15]:

$$Q_{BS} = K_{BS}\sqrt{P_S - P_B}.\begin{cases} x_0^2(x_0 + k_{BS}x_v)^{-1} & , x_v \geq 0 \\ x_0 - x_v & , x_v < 0 \end{cases} \tag{13.25}$$

$$Q_{BR} = K_{BR}\sqrt{P_B - P_R}. \begin{cases} x_0 + x_v & , x_v \geq 0 \\ x_0^2 (x_0 - k_{BR} x_v)^{-1} & , x_v < 0 \end{cases} \tag{13.26}$$

where x_0 is equivalent to a spool valve displacement that would result in the same amount of the load flow in a non-leaking servo-valve as the load flow in a leaking servo-valve with a centred spool. Namely, the equivalent orifice opening x_0 is assumed to be the same for matched servo-valve ports. For a symmetric servo-valve with matched control ports, it can be considered as:

$$K_x = K_{AS} = K_{AR} = K_{BS} = K_{BR} \tag{13.27}$$

$$K_P = k_{AS} = k_{AR} = k_{BS} = k_{BR} \tag{13.28}$$

where K_P is the leakage coefficient (total flow pressure coefficient), K_x is the servo-valve flow gain. For any servo-valve, the servo-valve leakage parameters (K_x, x_0, K_P) can be determined from readily available manufacturer data for a symmetric servo-valve with matched ports.

When the **control ports (A, B) are blocked**, the total supply load flow Q_S representing the internal leakage load flow can be expressed as follows [18]:

$$Q_S = 2K_x (P_S - P_R)^2 (x_0 + |x_v|)(1 + f(x_v))^2 \tag{13.29}$$

where

$$f(x_v) = \left[1 + \frac{|x_v|}{x_0}\right]^2 + \left[1 + K_P \frac{|x_v|}{x_0}\right]^2 \tag{13.30}$$

Figure 13.8 shows a typical leakage load flow curve. It has a maximum at the neutral spool valve position and decreases rapidly with the valve stroke.

Equations (13.29), (13.30) and Fig. 13.8 indicate that the internal leakage always exists for any electronic servo-valve, whenever there is a pressure difference $\Delta_P = P_S - P_R$.

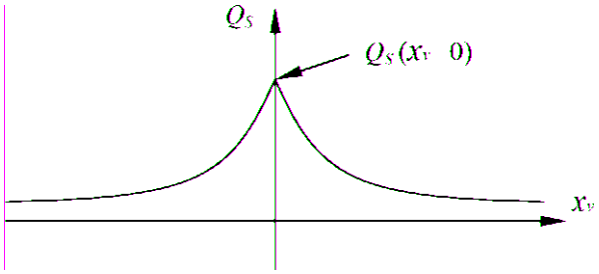


Fig. 13.8 Typical servo-valve leakage flow rate curve [15]

13.3 Vehicle Modelling

13.3.1 The Yaw-Roll Model of a Single Unit Heavy Vehicle

Figure 13.9 illustrates the combined yaw-roll dynamics of the vehicle modelled by a three-body system, in which m_s is the sprung mass, m_{uf} the unsprung mass at the front including the front wheels and axle, and m_{ur} the unsprung mass at the rear for the rear wheels and axle. The vehicle as a whole can translate longitudinally and laterally, and can yaw. The sprung mass can rotate about a horizontal axis (the roll axis). The location of the roll axis depends on the kinematic properties of the front and rear suspensions. The unsprung masses can also rotate in roll, enabling the effect of the vertical compliance of the tyres on the roll performance to be included in the model. The effects of the aerodynamic inputs (wind disturbances) and road inputs (cross-gradients, dips and bumps) are neglected.

The suspension springs, dampers and anti-roll bars generate moments between the sprung and unsprung masses in response to roll motions. The active anti-roll bar control system at each axle consists of a pair of actuators and a series of mechanical linkages, which are in parallel with the existing passive springs and dampers, and these active anti-roll bar systems generate additional (controlled) roll moments between the sprung and unsprung masses. The roll stiffness and damping of the vehicle suspension systems are assumed to be constant for the range of roll motions considered. The symbols of the yaw-roll model are found in Table 13.2.

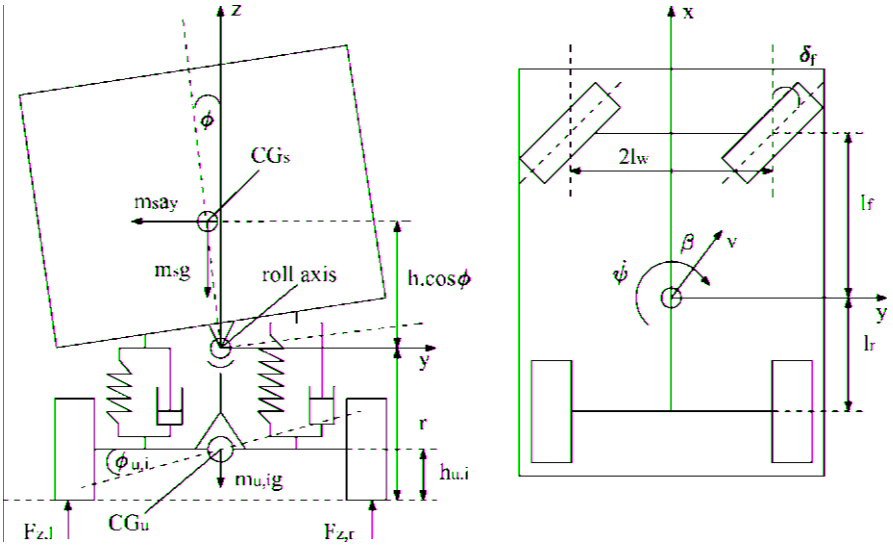


Fig. 13.9 The Yaw-Roll model of a single unit heavy vehicle [23]

Table 13.2 Symbols and parameters of the yaw-roll model [25]. [26]

Symbols	Description	Value	Unit
m_s	Sprung mass	12,487	kg
$m_{u,f}$	Unsprung mass on the front axle	706	kg
$m_{u,r}$	Unsprung mass on the rear axle	1000	kg
m	The total vehicle mass	14,193	kg
v	Forward velocity	–	$\frac{\text{Km}}{\text{h}}$
v_{wi}	Components of the forward velocity	–	$\frac{\text{Km}}{\text{h}}$
h	Height of sprung mass from the roll axis	1.15	m
$h_{u,i}$	Height of unsprung mass from the ground	0.53	m
r	Height of roll axis from the ground	0.83	m
a_y	Lateral acceleration	–	$\frac{\text{m}}{\text{s}^2}$
β	Side-slip angle at the centre of mass	–	rad
ψ	Heading angle	–	rad
$\dot{\psi}$	Yaw rate	–	$\frac{\text{rad}}{\text{s}}$
α	Side-slip angle	–	rad
ϕ	Sprung mass roll angle	–	rad
$\phi_{u,i}$	Unsprung mass roll angle	–	rad
δ_f	Steering angle	–	rad
u_i	Control current	–	A
C_f	Tyre cornering stiffness on the front axle	582	$\frac{\text{kN}}{\text{rad}}$
C_r	Tyre cornering stiffness on the rear axle	783	$\frac{\text{kN}}{\text{rad}}$
k_f	Suspension roll stiffness on the front axle	380	$\frac{\text{kNm}}{\text{rad}}$
k_r	Suspension roll stiffness on the rear axle	684	$\frac{\text{kNm}}{\text{rad}}$
b_f	Suspension roll damping on the front axle	100	$\frac{\text{kN}}{\text{rad}}$
b_r	Suspension roll damping on the rear axle	100	$\frac{\text{kN}}{\text{rad}}$
k_{tf}	Tyre roll stiffness on the front axle	2060	$\frac{\text{kNm}}{\text{rad}}$
k_{tr}	Tyre roll stiffness on the rear axle	3337	$\frac{\text{kNm}}{\text{rad}}$
I_{xx}	Roll moment of inertia of sprung mass	24,201	kgm^2
I_{xz}	Yaw-roll inertial of sprung mass	4200	kgm^2
I_{zz}	Yaw moment of inertia of sprung mass	34,917	kgm^2
l_f	Length of the front axle from the CG	1.95	m
l_r	Length of the rear axle from the CG	1.54	m
l_w	Half of the vehicle width	0.93	m
μ	Road adhesion coefficient	1	–
D_f	Outer diameter of the front anti-roll bar	32	mm
D_r	Outer diameter of the rear anti-roll bar	34	mm
E	Young's modulus of material	206,000	MPa

In the vehicle modelling, the differential equations of motion of the yaw-roll dynamics of the single unit vehicle, i.e., the lateral dynamics, the yaw moment, the roll moment of the sprung mass, the roll moment of the front and the rear unsprung masses, are formalized in the following Eqs. (13.31):

$$\begin{cases} mv(\dot{\beta}+\dot{\psi}) - m_s h\ddot{\phi} = F_{yf} + F_{yr} \\ -I_{xz}\ddot{\phi} + I_{zz}\ddot{\psi} = F_{yf}l_f - F_{yr}l_r \\ (I_{xx} + m_s h^2)\ddot{\phi} - I_{xz}\ddot{\psi} = m_s g h\phi + m_s v h(\dot{\beta}+\dot{\psi}) - k_f(\phi - \phi_{uf}) \\ \quad -b_f(\dot{\phi} - \dot{\phi}_{uf}) + M_{ARf} + T_f - k_r(\phi - \phi_{ur}) - b_r(\dot{\phi} - \dot{\phi}_{ur}) + M_{ARr} + T_r \\ -rF_{yf} = m_{uf} v(r - h_{uf})(\dot{\beta}+\dot{\psi}) + m_{uf} g h_{uf}.\phi_{uf} - k_{tf}\phi_{uf} \\ \quad +k_f(\phi - \phi_{uf}) + b_f(\dot{\phi} - \dot{\phi}_{uf}) + M_{ARf} + T_f \\ -rF_{yr} = m_{ur} v(r - h_{ur})(\dot{\beta}+\dot{\psi}) - m_{ur} g h_{ur}\phi_{ur} - k_{tr}\phi_{ur} \\ \quad +k_r(\phi - \phi_{ur}) + b_r(\dot{\phi} - \dot{\phi}_{ur}) + M_{ARr} + T_r \end{cases} \tag{13.31}$$

where T_f, T_r are the torques generated by the active anti-roll bar system at the each axle. The lateral tyre forces $F_{y;i}$ in the direction of the velocity at the wheel ground contact points are modelled by using linear stiffness coefficients as:

$$\begin{cases} F_{yf} = \mu C_f \alpha_f \\ F_{yr} = \mu C_r \alpha_r \end{cases} \tag{13.32}$$

with the tyre side-slip angles:

$$\begin{cases} \alpha_f = -\beta + \delta_f - \frac{l_f \dot{\psi}}{v} \\ \alpha_r = -\beta + \frac{l_r \dot{\psi}}{v} \end{cases} \tag{13.33}$$

Let us now detail how the moments M_{ARf} and M_{ARr} in (13.31) are computed. When the vertical displacements of the left and the right wheels differ, the passive anti-roll bar with the rotational stiffness k_{AO} creates an anti-roll moment, resulting in the anti-roll forces F_{AU}, see Fig. 13.10, which are acting on the unsprung mass as follows:

$$F_{AUl} = -F_{AUr} = k_{AU}(\Delta Z_{Ar} - \Delta Z_{Al}) \tag{13.34}$$

and the anti-roll forces F_{AS} acting on the sprung mass are

$$F_{ASl} = -F_{ASr} = F_{AUl}\frac{t_A}{t_B} = k_{AS}(\Delta Z_{Ar} - \Delta Z_{Al}) \tag{13.35}$$

where $\Delta Z_{Ar,l}$ are the displacements of the connection point between the anti-roll bars and the wheels, t_A is half the distance between the two suspensions, t_B is half

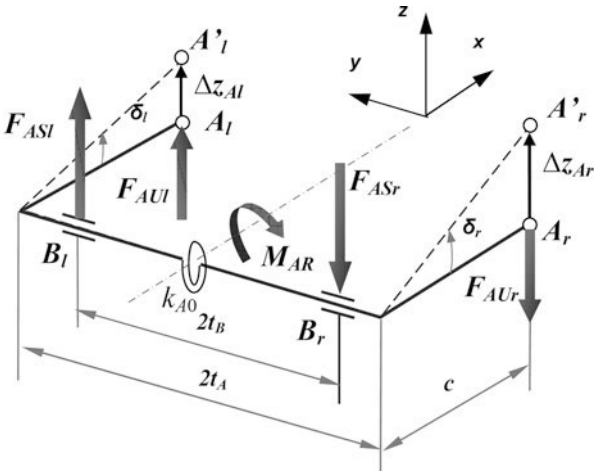


Fig. 13.10 Diagram of the passive anti-roll bars on vehicles [8]

the distance of the chassis, c is the length of the anti-roll bars' arm, k_{AU} and k_{AS} are the modified rotational stiffness corresponding to the unsprung and sprung mass, respectively:

$$k_{AU} = k_{AO}\frac{1}{c^2} \quad and \quad k_{AS} = k_{AO}\frac{t_A}{t_B c^2} \tag{13.36}$$

The moment of the passive anti-roll bar impacts the unsprung and sprung masses at the front axle as follows:

$$M_{ARf} = 4k_{AOf}\frac{t_A t_B}{c^2}\phi - 4k_{AOf}\frac{t_A^2}{c^2}\phi_{uf} \tag{13.37}$$

The moment of the passive anti-roll bar impacts the unsprung and sprung masses at the rear axle as follows:

$$M_{ARr} = 4k_{AOr}\frac{t_A t_B}{c^2}\phi - 4k_{AOr}\frac{t_A^2}{c^2}\phi_{ur} \tag{13.38}$$

13.3.2 The Fully Integrated Model of a Single Unit Heavy Vehicle

Figure 13.11 shows the fully integrated model using an active anti-roll bar system. This model includes a linear single unit heavy vehicle yaw-roll model with the differential equations of motion given in Eq. (13.31) and the four Electronic Servo-

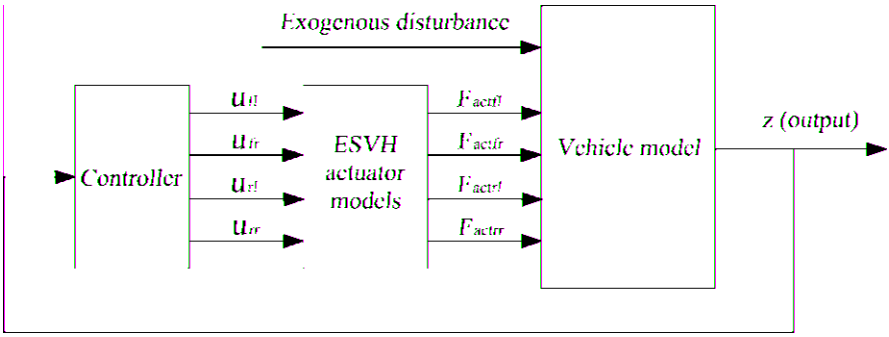


Fig. 13.11 Diagram of the fully integrated model using an active anti-roll bar system

Valve Hydraulic (ESVH) actuators (two at the front axle and two at the rear axle) with the differential equations of motion given in Eqs. (13.1), (13.12), (13.18). The controller receives the information from the output (z) and computes the input currents ($u_{fl}, u_{fr}, u_{rl}, u_{rr}$), where u_{fl} and u_{fr} are, respectively, the input currents of the electronic servo-valves on the left and right at the front axle, u_{rl} and u_{rr} at the rear axle. The forces ($F_{actfl}, F_{actfr}, F_{actrl}, F_{actrr}$) of the hydraulic actuators are applied to the vehicle model, where F_{actfl} and F_{actfr} are, respectively, the forces of the hydraulic actuators on the left and on the right at the front axle, F_{actrl} and F_{actrr} at the rear axle. Therefore the torque generated by the active anti-roll bar system at the front axle is determined by

$$T_f = -l_{act}F_{actfl} + l_{act}F_{actfr} = -l_{act}A_p\Delta_{Pfl} + l_{act}A_p\Delta_{Pfr} \tag{13.39}$$

and the torque generated by the active anti-roll bar system at the rear axle is

$$T_r = -l_{act}F_{actrl} + l_{act}F_{actrr} = -l_{act}A_p\Delta_{Prl} + l_{act}A_p\Delta_{Prr} \tag{13.40}$$

where l_{act} is half the distance between the two actuators, Δ_{Pfl} and Δ_{Pfr} are, respectively, the difference of pressure inside the hydraulic cylinders on the left and right at the front axle, Δ_{Prl} and Δ_{Prr} are, respectively, the difference of pressure inside the hydraulic cylinders on the left and right at the rear axle. They are given by the state equation (13.42).

The displacements of the piston inside the hydraulic cylinders on the left and right at the front axle (y_{afl}, y_{afr}) and at the rear axle (y_{arl}, y_{arr}) are approximately calculated as follows [5]:

$$\begin{cases} y_{afl} = -l_{act}(\phi - \phi_{uf}) \\ y_{afr} = l_{act}(\phi - \phi_{uf}) \\ y_{arl} = -l_{act}(\phi - \phi_{ur}) \\ y_{arr} = l_{act}(\phi - \phi_{ur}) \end{cases} \tag{13.41}$$

From Eqs. (13.19) and (13.41), the dynamical equations of these ESVH actuators are shown as

$$\begin{cases} \frac{V_t}{4\beta_e}\dot{\Delta}_{Pfl} + (K_P + C_{tp})\Delta_{Pfl} - K_x X_{vfl} - A_p l_{act}\dot{\phi} + A_p l_{act}\dot{\phi}_{uf} = 0 \\ \dot{X}_{vfl} + \frac{1}{\tau}X_{vfl} - \frac{K_v}{\tau}u_{fl} = 0 \\ \frac{V_t}{4\beta_e}\dot{\Delta}_{Pfr} + (K_P + C_{tp})\Delta_{Pfr} - K_x X_{vfr} + A_p l_{act}\dot{\phi} - A_p l_{act}\dot{\phi}_{uf} = 0 \\ \dot{X}_{vfr} + \frac{1}{\tau}X_{vfr} - \frac{K_v}{\tau}u_{fr} = 0 \\ \frac{V_t}{4\beta_e}\dot{\Delta}_{Prl} + (K_P + C_{tp})\Delta_{Prl} - K_x X_{vrl} - A_p l_{act}\dot{\phi} + A_p l_{act}\dot{\phi}_{ur} = 0 \\ \dot{X}_{vrl} + \frac{1}{\tau}X_{vrl} - \frac{K_v}{\tau}u_{rl} = 0 \\ \frac{V_t}{4\beta_e}\dot{\Delta}_{Prr} + (K_P + C_{tp})\Delta_{Prr} - K_x X_{vrr} + A_p l_{act}\dot{\phi} - A_p l_{act}\dot{\phi}_{ur} = 0 \\ \dot{X}_{vrr} + \frac{1}{\tau}X_{vrr} - \frac{K_v}{\tau}u_{rr} = 0 \end{cases} \tag{13.42}$$

The combination of equations from (13.31) to (13.42) is the motion differential equations of the fully integrated model.

The fully integrated model is written in the LTI state-space representation form:

$$\dot{x} = A^f.x + B_1^f.w + B_2^f.u \tag{13.43}$$

where the state vector is given by

$$x = \left[\beta\ \dot{\psi}\ \phi\ \dot{\phi}\ \phi_{uf}\ \phi_{ur}\ \Delta_{Pfl}\ X_{vfl}\ \Delta_{Pfr}\ X_{vfr}\ \Delta_{Prl}\ X_{vrl}\ \Delta_{Prr}\ X_{vrr}\right]^T$$

The exogenous disturbance (steering angle) is

$$w = \left[\delta_f\right]^T$$

and the control inputs (input currents):

$$u = \left[u_{fl}\ u_{fr}\ u_{rl}\ u_{rr}\right]^T$$

Here, the matrices A^f, B_1^f and B_2^f are matrices of appropriate dimensions.

13.4 Effect of the Internal Leakage Inside the Electronic Servo-Valve on the Open-Loop System

The Electronic Servo-Valve Hydraulic (ESVH) actuator is used for the active anti-roll bar system in the proposed fully integrated model. This actuator is not the same as with the active damper which is often used for the active system in cars, such as an active or semi-active suspension system, because the piston of the hydraulic cylinder

is solid. From Fig. 13.6 and Table 13.1 we can see that the total leakage coefficient of the hydraulic cylinder is zero, which means that oil cannot pass through the contact surfaces between the cylinder and piston of the hydraulic cylinder.

The problem we want to tackle is: **What happens, if when the vehicle is running, the active anti-roll bar system suddenly stops working?** The reasons may be due to broken wiring from the controller to the servo-valve, or the controller is not working, etc. In this situation, the spool valve of the servo-valve is at the neutral position, but the suspension roll angles always exist due to the effect of the lateral inertia force. If there is no internal oil leakage inside the electronic servo-valve, the oil cannot pass between the two chambers of the hydraulic cylinder, therefore the movement of the piston inside the hydraulic cylinder is impossible. This leads to the sprung mass and unsprung mass becoming one block and the role of the suspension will be removed, which would lead to a dangerous situation for the vehicle. In this section, we will give an answer to the above question by investigating the internal leakage inside the electronic servo-valve when the spool valve is at the neutral position.

13.4.1 Neutral Position of the Spool Valve

In Sect. 13.2.3, it was indicated that the internal leakage always exists for any electronic servo-valve. It has a maximum at the neutral spool valve position when the two control ports are blocked.

To evaluate the effect of the internal leakage inside the electronic servo-valve on an ESVH actuator, let us consider the servo-valve structural characteristics illustrated in Sect. 13.2.1. Due to the orifices of the servo-valve being symmetrically matched, from Eq. (13.5), at the neutral position of the spool valve, the areas of the 1st and 2nd orifices are equal and defined as follows:

$$A_1(0) = A_2(0) = A_0 \tag{13.44}$$

In this case, the load flow through the ESVH actuator in Eq. (13.9) is as follows:

$$Q_L = C_d A_0 \sqrt{\frac{1}{\rho}(P_s - \Delta P)} - C_d A_0 \sqrt{\frac{1}{\rho}(P_s + \Delta P)} \tag{13.45}$$

Considering Eq. (13.45) and Fig. 13.4, there are three cases of interest for the differential pressure between the two chambers of the hydraulic cylinder as follows:

- If $P_1 < P_2$, then $\Delta P < 0$ and $Q_L > 0$. This means that the direction of the load flow is in the same direction with the direction as illustrated in Fig. 13.4.
- If $P_1 = P_2$, then $\Delta P = 0$ and $Q_L = 0$. This means that there is no the load flow though the actuator.

- If $P_1 > P_2$, then $\Delta P > 0$ and $Q_L < 0$. This means that the direction of the load flow will reverse with the direction as illustrated in Fig. 13.4.

Thus, although the spool valve is at the neutral position, thanks to the internal leakage inside the electronic servo-valve, the oil can still pass from the higher pressure chamber to the lower pressure chamber of the hydraulic cylinder with a certain level. This allows the piston to move inside the hydraulic cylinder. Therefore when the active anti-roll bar system suddenly stops working, the sprung and unsprung masses will not be blocked. This can be called the self-protection capacity of the active anti-roll bar system when the controller does not work. To clarify this issue, in the next section the author considers the absence of the input currents from the controller to the ESVH actuators and it is compared with the two cases of the passive and "without anti-roll bar".

13.4.2 Effect of the Internal Leakage on the Open-Loop System

In this section, the effect of the internal leakage inside the servo-valve on the vehicle behaviour in the frequency and time domains will be assessed by using the proposed fully integrated model. The parameter values of the ESVH actuators and the yaw-roll model are detailed in Tables 13.1 and 13.2. The forward velocity is constant at 70 km/h. Here, we compare the three following cases:

- **First case: "Open-loop system"** is defined when the ESVH actuators are installed in the vehicle but there is no input current entering these actuators (i.e., the faulty case);
- **Second case: "Passive anti-roll bar"** is determined in [8]. This is typical of the modern vehicles;
- **Third case: "Without anti-roll bar"**, this is the case where heavy vehicles do not have any anti-roll bar system.

13.4.2.1 Effect of the Internal Leakage Inside the Servo-Valve in the Frequency Domain

In the active anti-roll bar system of heavy vehicles, the frequency response up to 4 rad/s is considered to represent the limited bandwidth of the driver [27]. Figures 13.12 and 13.13 show the comparison of the transfer function magnitude of the normalized load transfer at both axles, for three cases: (1) open-loop system (continued line), (2) passive anti-roll bar (dash line) and (3) without anti-roll bar (dash-dot line). We can see that in the case of the passive anti-roll bar, the normalized load transfers at the two axles are reduced by about 3.5 dB in the frequency range to over 4 rad/s, compared to the case without anti-roll bar.

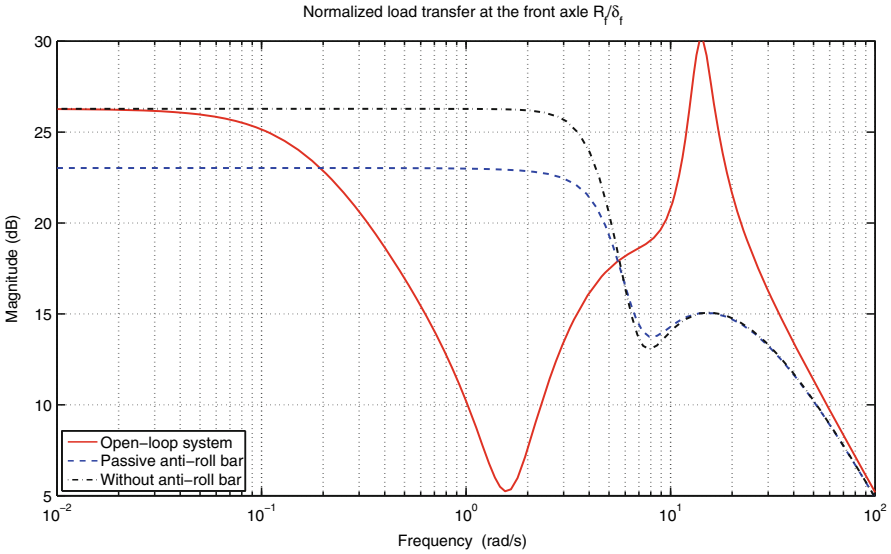


Fig. 13.12 The transfer function magnitude of the normalized load transfer ($\frac{R_f}{\delta_f}$) at the front axle

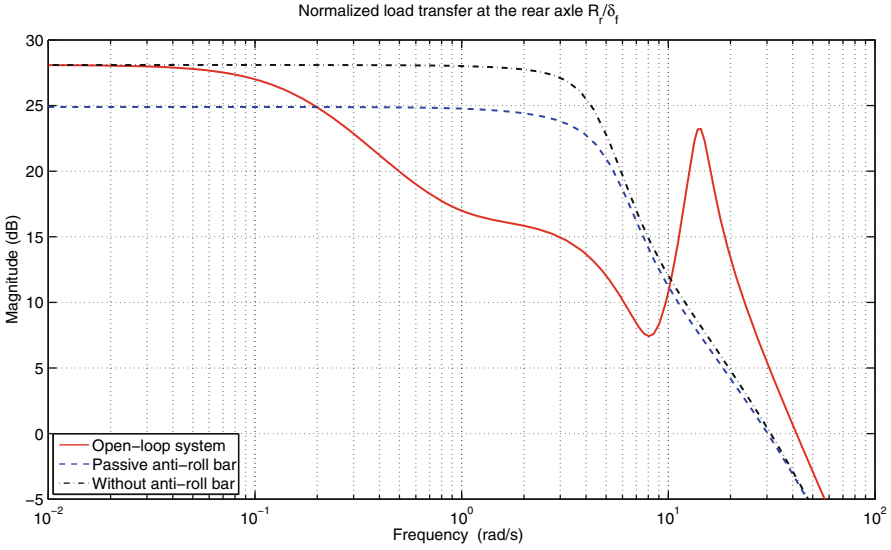


Fig. 13.13 Transfer function magnitude of normalized load transfer ($\frac{R_r}{\delta_f}$) at the rear axle

In the case of the open-loop system, there are two frequency ranges of interest:

- **The frequency ranges from** 0.1 rad/s **to over** 4 rad/s: the normalized load transfers at both axles are reduced significantly, compared with the case without anti-roll bar. Hence, these reductions show that even if the controller does not work (no input currents entering to the ESVH actuators), thanks to the internal leakage inside the servo-valve, the active anti-roll bar system still has a positive effect on improving the roll stability of heavy vehicles.
- **The frequency range less than** 0.1 rad/s (**the steady state manoeuvre**): the open-loop system cannot change the roll stability of heavy vehicles, compared with the case without anti-roll bar.

In the next section we will consider the heavy vehicle behaviour in the steady state manoeuvre case (the frequency range less than 0.1 rad/s).

13.4.2.2 Effect of the Internal Leakage Inside the Servo-Valve in the Time Domain

To survey the effect of the internal leakage inside the servo-valve in the steady state manoeuvre, we consider that the steering angle is in a step signal from 0.5s to 2.5s and then remains constant at 2.5 deg [23, 24] as showed in Fig. 13.15a. The vehicle's trajectory is shown in Fig. 13.14.

Figure 13.15 shows the time response of the roll angles of the sprung mass, the normalized load transfers, differential pressures inside the hydraulic cylinder and load flows of the oil through the servo-valve at both axles. From 0.5s, the driver starts to change the steering angle. Due to the impact of the inertial force, the roll angles of the suspension are altered and they generate the relative displacement of the piston and cylinder inside each hydraulic cylinder. This increases the differential pressure up to 2.5s, and the oil also starts to flow through the servo-valve. But from 2.5s, the inertial force is held constant and the impact of the internal leakage inside the servo-valve makes the differential pressure rapidly drop up to 35s; then there is no more differential pressure inside the hydraulic cylinder and the oil will not pass through the servo-valve. This greatly affects the roll stability of the vehicle.

Figures 13.15c,d show the normalized load transfers at the two axles. They indicate that from 0.5s the normalized load transfers at both axles start to increase. The normalized load transfers in the open-loop system case will intersect that of the passive anti-roll bar case at 6.5s (point B in Fig. 13.14). In the open-loop system case, at 12.5s the normalized load transfer at the rear axle equals 1 ($R_r = 1$) and at this time, rollover will occur (point C in Fig. 13.14). Additionally the time response of the heavy vehicle as well as of the ESVH actuators on the right at the front and rear axles in the case of the open-loop system are synthesized in Table 13.3.

Due to the constant forward velocity of 70 km/h, at 6.5s the vehicle will travel the distance (S) of 126m (from point A to point B in Fig. 13.14) and at 12.5s the distance travelled would be 243m (from point A to point C in Fig. 13.14). Although at 12.5s the rollover will occur ($R_r = 1$), but in fact with the distance of 243m

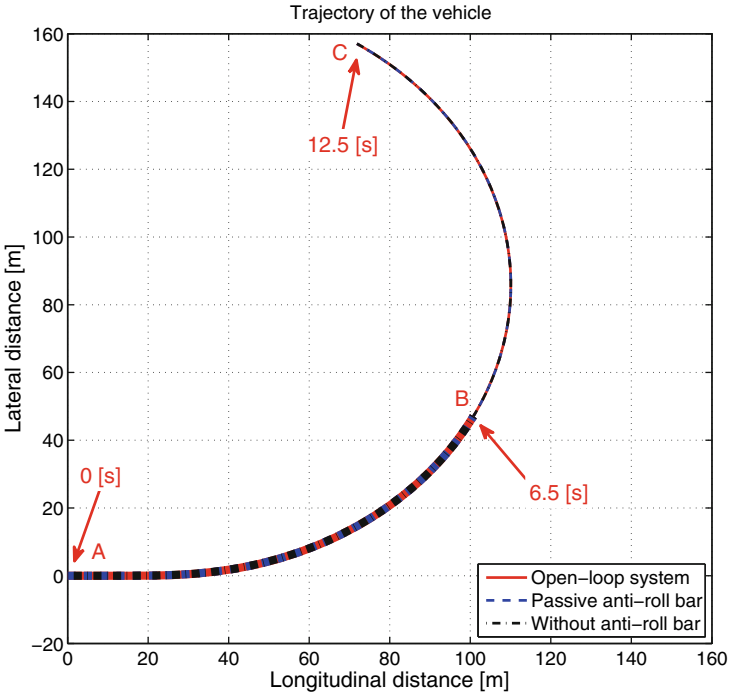


Fig. 13.14 Trajectory of the vehicle

the driver usually will have some impacts on the steering angle to accommodate the vehicle's trajectory or there is the effect of the road on the front wheels. This will allow the open-loop system to improve the roll stability by the internal leakage inside the servo-valve in the frequency range of [0.1, 4] rad/s. Therefore even in the steady state manoeuvre, roll stability of the vehicle is always guaranteed in practice.

From the analysis and simulation results in the frequency and time domains shown above we have answers to the previous questions about roll stability of the vehicle when the active anti-roll bar system suddenly fails. We can confirm that the installation of the active anti-roll system using ESVH actuators does not reduce the operational performance of the vehicle for every situation.

Survey results also indicate that:

- If the total flow pressure coefficient K_P is less than $2.2 \times 10^{-16}\ \frac{m^5}{Ns}$, the role of the internal leakage inside the servo-valve will be ignored. The oil cannot move between the two chambers of the hydraulic cylinder and the movement of the piston inside the hydraulic cylinder is impossible. In this case, the sprung mass and unsprung mass will become one block. This is a dangerous situation for heavy vehicles.

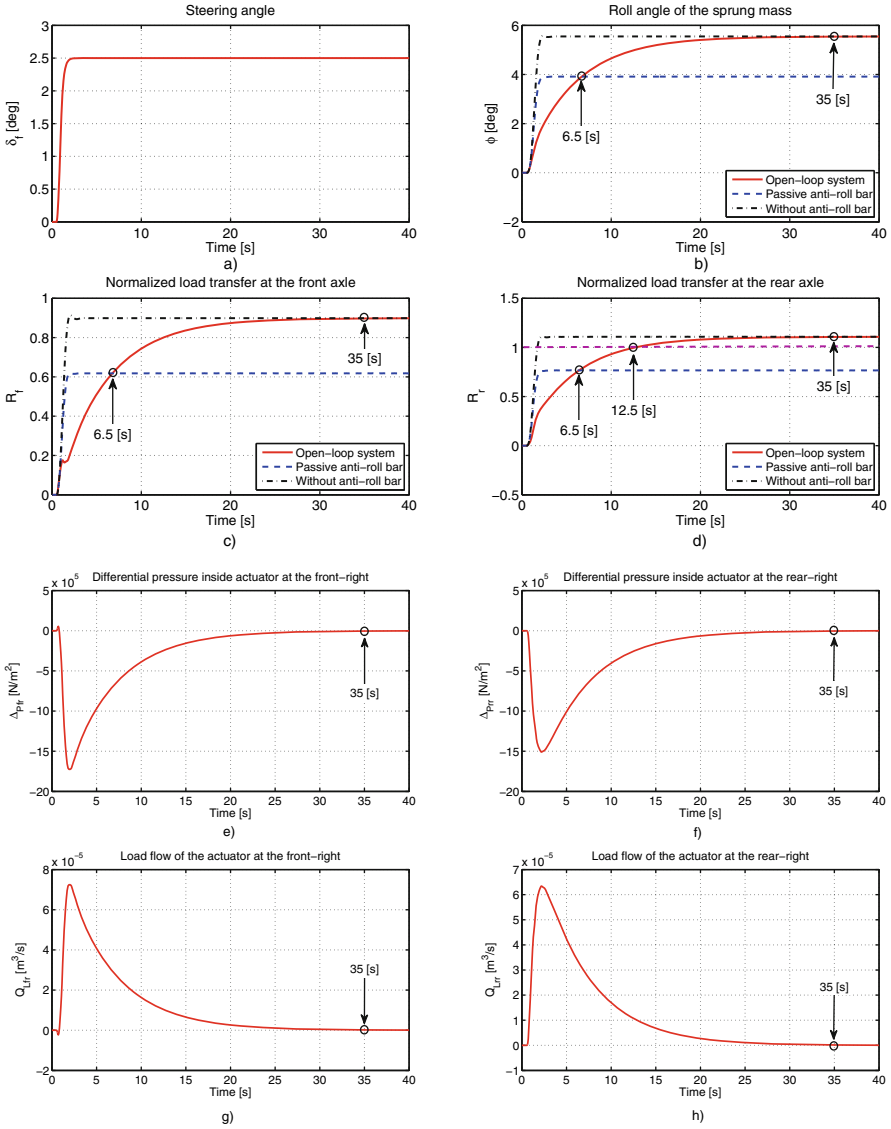


Fig. 13.15 Time responses of the heavy vehicle and the actuators in the case of the steady state manoeuvre

- If the total flow pressure coefficient K_P is higher than $5.2 \times 10^{-10}\ \frac{m^5}{Ns}$, the internal leakage inside the servo-valve does not produce any effect on roll stability. Therefore the vehicle behaviour will be the same as in the case without anti-roll bar.

Table 13.3 Synthesis of the time response in the steady state manoeuvre

Time (s)	0.5	2.5	6.5	12.5	35
δ_f (deg)	0	2.5	2.5	2.5	2.5
$\Delta P_{fr}(\times 10^5 \mathrm{N/m^2})$	0	−17	−8	−2.5	0
$Q_{Lfr}(\times 10^{-5} \mathrm{N/m^2})$	0	7	3	1	0
R_f	0	0.3	0.6	0.8	0.9
$\Delta P_{rr}(\times 10^5 \mathrm{N/m^2})$	0	−15	−8	−2.5	0
$Q_{Lrr}(\times 10^{-5} \mathrm{N/m^2})$	0	6.3	3.5	1	0
R_r	0	0.4	0.8	1 (rollover)	1.1
$S(m)$	9.7	49	126	243	680

13.5 Effect of the Internal Leakage Inside the Electronic Servo-Valve on the Closed-Loop System

13.5.1 H_∞/LPV Control Design for the Fully Integrated Model

The aim is to design the control so that the active anti-roll bars are operating all the time, thus improving roll stability. Roll stability is achieved by limiting the lateral load transfers to below the levels required for wheel lift-off. In order to describe the control objective, the nonlinear model (13.43) has a partitioned representation in the following way:

$$\begin{bmatrix} \dot{x}(t) \\ z(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A(\rho) & B_1(\rho) & B_2(\rho) \\ C_1(\rho) & D_{11}(\rho) & D_{12}(\rho) \\ C_2(\rho) & D_{21}(\rho) & D_{22}(\rho) \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \\ u(t) \end{bmatrix} \tag{13.46}$$

with the exogenous input $w(t) = \begin{bmatrix} \delta_f \end{bmatrix}$, the control input $u(t) = \begin{bmatrix} u_{fl} & u_{fr} & u_{rl} & u_{rr} \end{bmatrix}^T$, the performance output vector $z(t) = \begin{bmatrix} u_{fl} & u_{fr} & u_{rl} & u_{rr} & R_f & R_r \end{bmatrix}^T$ and the measured output vector $y(t) = \begin{bmatrix} a_y & \dot{\phi} \end{bmatrix}^T$. Here, the lateral acceleration is defined as follows: $a_y = v\dot{\beta} + v\dot{\psi} - h\ddot{\phi}$ [27, 28].

In Fig. 13.16, the closed-loop system includes the feedback structure of the nominal model $G(\rho)$, the controller $K(\rho)$, the weighting functions and the performance objectives. In this diagram, u is the control input, y the measured output, n the measurement noise and z the performance output. The steering angle δ_f is the disturbance signal set by the driver. The weighting functions W_δ, W_n, W_z are, respectively, characterized for the steering angle, sensor noise and performance output. The weighting functions W_δ and W_n are chosen as in Table 13.4.

The weighting function W_z represents the performance output and is chosen as a diagonal matrix $W_z = diag[W_{zu}, W_{zR}]$. The weighting function $W_{zu} = diag[W_{zu1}, W_{zu2}, W_{zu3}, W_{zu4}]$ is adapted to the four input currents entering the four ESVH actuators. The weighting function $W_{zR} = diag[W_{zRf}, W_{zRr}]$, corresponds

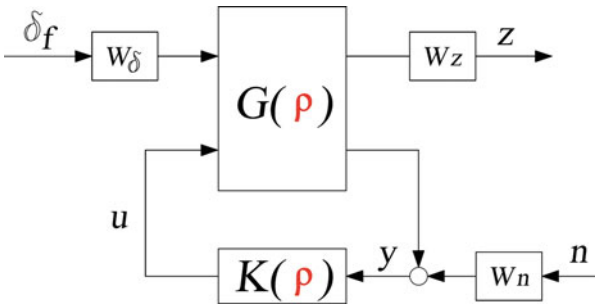


Fig. 13.16 The closed-loop interconnection structure of the fully integrated model

Table 13.4 The weighting functions of the H_∞/LPV synthesis for the fully integrated model

W_δ	W_{n1}	W_{n2}	W_{zu1}	W_{zu2}	W_{zu3}	W_{zu4}	W_{zRf}	W_{zRr}
$\frac{\pi}{180}$	0.01	0.01	$\frac{1}{0.4}$	$\frac{1}{0.4}$	$\frac{1}{0.4}$	$\frac{1}{0.4}$	$\frac{\frac{s}{20}+2}{\frac{s}{100}+15}$	$\frac{\frac{s}{20}+2}{\frac{s}{100}+15}$

to the normalized load transfers at the front and rear axles. The elements of the weighting function W_z are selected in Table 13.4.

In the LPV model of an active anti-roll bar system (13.46), the forward velocity v is selected as the varying parameter $\rho = v$ which can be measured directly by the sensors. The quadratic LPV γ-performance problem is to choose the parameter-varying control matrices $A_K(\rho)$, $B_K(\rho)$, $C_K(\rho)$, $D_K(\rho)$ in such a way that the resulting closed-loop system is quadratically stable and the induced $\mathcal{L}_2$ norm from w to z is less than γ. The structure of the LPV controller $K(\rho)$ is defined as follows [29]:

$$\begin{bmatrix} \dot{x}_K(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} A_K(\rho) & B_K(\rho) \\ C_K(\rho) & D_K(\rho) \end{bmatrix} \begin{bmatrix} x_K(t) \\ y(t) \end{bmatrix} \tag{13.47}$$

where $A_K(\rho)$, $B_K(\rho)$, $C_K(\rho)$, $D_K(\rho)$ are continuous bounded matrix functions.

For the interconnection structure shown in Fig. 13.16, the H_∞ controllers are synthesized for 10 values of the forward velocity v in the range $[30, \ldots, 130]$ km/h. The spacing of the grid points is based upon how well the H_∞ point design performs for the plant around the design point.

At all the grid points, the proposed weighting functions are applied to the entire grid parameter space and the effect of the scheduling parameter is ignored. In the H_∞ control design, the γ iteration results in an optimal γ value and an optimal controller. However, if the weighting functions were changed, another optimal γ and another optimal controller would be gained.

The following commands are used to make the grid points as well as the LPV controller synthesis by using LPVTools^{TM}:
rho = pgrid('rho',linspace(30/3.6,130/3.6,10));
and *[Klpv,normlpv] = lpvsyn(H,nmeas,ncont).*
The optimal γ of the controller is $\gamma_{opt} = normlpv = 0.8096$.

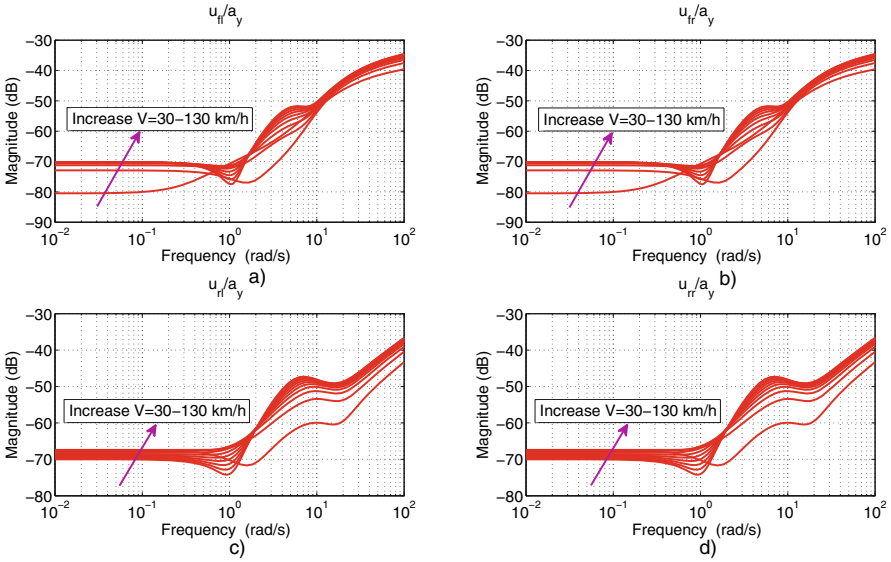


Fig. 13.17 Transfer function magnitude of controller: from lateral acceleration to (**a**) input current at front-left $\frac{u_{fl}}{a_y}$, (**b**) input current at front-right $\frac{u_{fr}}{a_y}$, (**c**) input current at rear-left $\frac{u_{rl}}{a_y}$, (**d**) input current at rear-right $\frac{u_{rr}}{a_y}$

Figures 13.17, 13.18 show the transfer function magnitude of the H_∞/LPV active anti-roll bar controller, with two inputs (the lateral acceleration a_y, the roll rate $\dot{\phi}$) and the four outputs (the input currents entering the ESVH actuators at the front-left u_{fl}, at the front-right u_{fr}, at the rear-left u_{rl}, at the rear-right u_{rr}). We can see that, due to the ESVH actuators at each axle being identical and symmetrical, the transfer functions of the H_∞/LPV controller from the inputs to the outputs have the same magnitude at the front and rear axles.

13.5.2 Simulation Results Analysis with the Nominal Value of the Total Flow Pressure Coefficient

In this section, the simulation results of the fully integrated model with the H_∞/LPV controller are shown in both frequency and time domains. The nominal value of the total flow pressure coefficient is $K_P = 4.2 \times 10^{-11}$ m^5/(Ns). The parameters values of the ESVH actuators and of the yaw-roll model are those given in Tables 13.1 and 13.2.

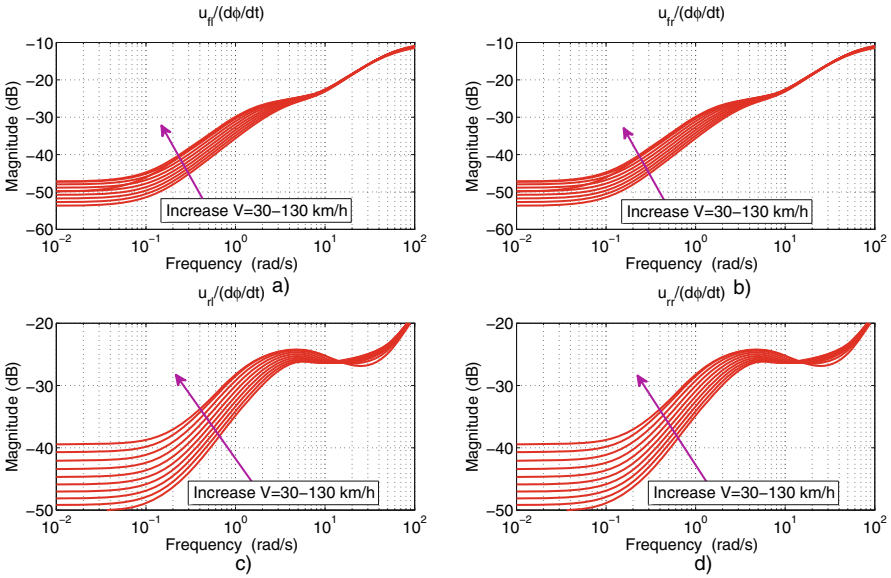


Fig. 13.18 Transfer function magnitude of controller: from roll rate to (**a**) input current at front-left $\frac{u_{fl}}{\dot{\phi}}$, (**b**) input current at front-right $\frac{u_{fr}}{\dot{\phi}}$, (**c**) input current at rear-left $\frac{u_{rl}}{\dot{\phi}}$, (**d**) input current at rear-right $\frac{u_{rr}}{\dot{\phi}}$

13.5.2.1 Analysis in the Frequency Domain

The main objective of the active anti-roll bar system is to reduce the normalized load transfers at all axles. Figure 13.19 shows the transfer function magnitude of the normalized load transfers and input currents of the ESVH actuators on the right at both axles. The varying parameter $\rho = v$ is considered in the interval [30, 130] km/h. Figures 13.19a,b indicates that in the case of the H_∞/LPV active anti-roll bar control, the normalized load transfers at the two axles are significantly reduced when compared to the passive anti-roll bar. This reduction is achieved in the frequency range to over 4 rad/s, which represents the limited bandwidth of the driver [31]. Therefore, we can confirm that the H_∞/LPV active anti-roll bar control improves roll stability of heavy vehicles in reducing the rollover risk in all of the desired forward velocity range.

Figures 13.19c,d show the transfer functions magnitude of the input currents on the right at the front axle ($\frac{u_{fr}}{\delta_f}$) and at the rear axle ($\frac{u_{rr}}{\delta_f}$), respectively. When the forward velocity increases, the controller input currents ($u_{fr,rr}$) also increase. This indicates that the active anti-roll bar system requires more input current (i.e., energy) at high forward velocity than at low forward velocity.

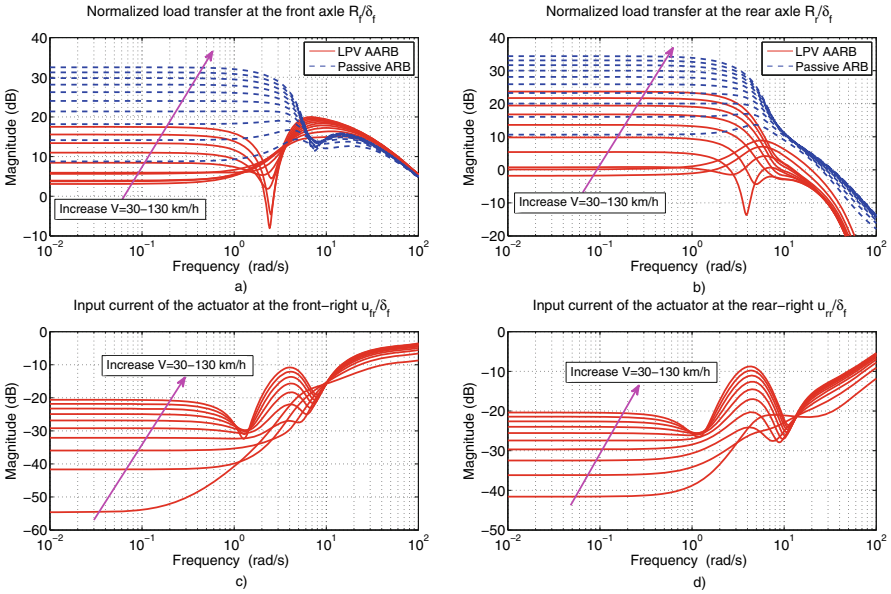


Fig. 13.19 Transfer function magnitude of (**a**, **b**) the normalized load transfers at the two axles $\frac{R_{f,r}}{\delta_f}$, (**c**, **d**) the input currents of the ESVH actuators at the front-right $\frac{u_{fr}}{\delta_f}$ and at the rear-right $\frac{u_{rr}}{\delta_f}$

13.5.2.2 Analysis in the Time Domain

In this section, the simulation results of the single unit heavy vehicle in the time domain are shown in the steady state manoeuvre. The steering angle (a step signal) is shown in Fig. 13.15a and the forward velocity is held constant at 70 km/h.

Figure 13.20 shows the normalized load transfers and suspension roll angles at the two axles in comparison between the H_∞/LPV active anti-roll bar control and the passive anti-roll bar. The H_∞/LPV active anti-roll bar control reduces the normalized load transfers by about 60%, compared to the passive anti-roll bar. On the other hand, when compared with Figs. 13.15c,d, it indicates that the H_∞/LPV active anti-roll bar control system has provided a stabilization of the normalized load transfers in the steady state manoeuvre. Indeed in the case of the H_∞/LPV active anti-roll bar control, the normalized load transfers at both axles increase with the increase of the steering angle from 0.5s to 2s, and then they are kept stable at 0.32 for the front axle, and 0.42 for the rear axle. Meanwhile, in the open-loop system, these values increase constantly and they reach stability at 35s, then the vehicle shows the same results as the case without anti-roll bar.

Figures 13.21 and 13.22 show the characteristics of the four ESVH actuators of the fully integrated model. If the right and left ESVH actuators are identical and symmetrically mounted at each axle, then the characteristics of the two ESVH actuators do have the same magnitude and the opposite direction.

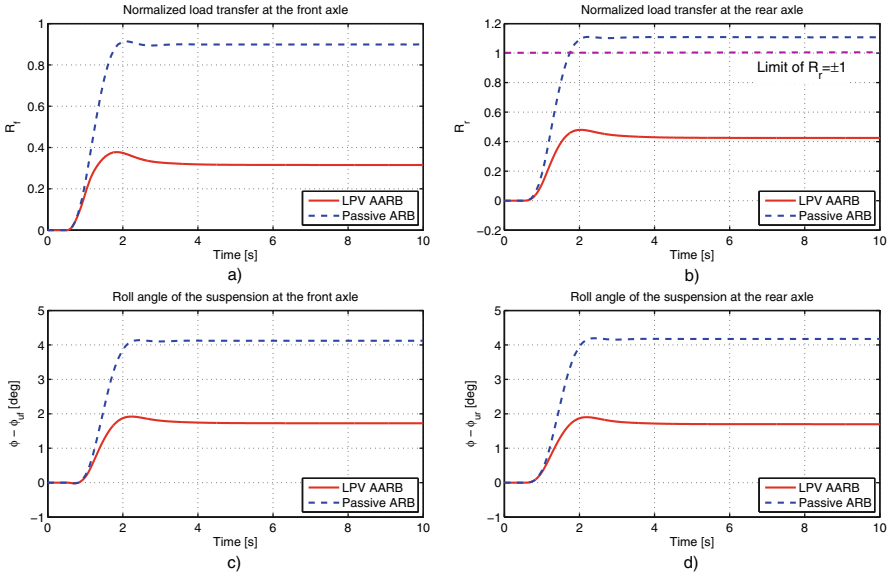


Fig. 13.20 The normalized load transfers and roll angle of the suspensions in the steady state manoeuvre

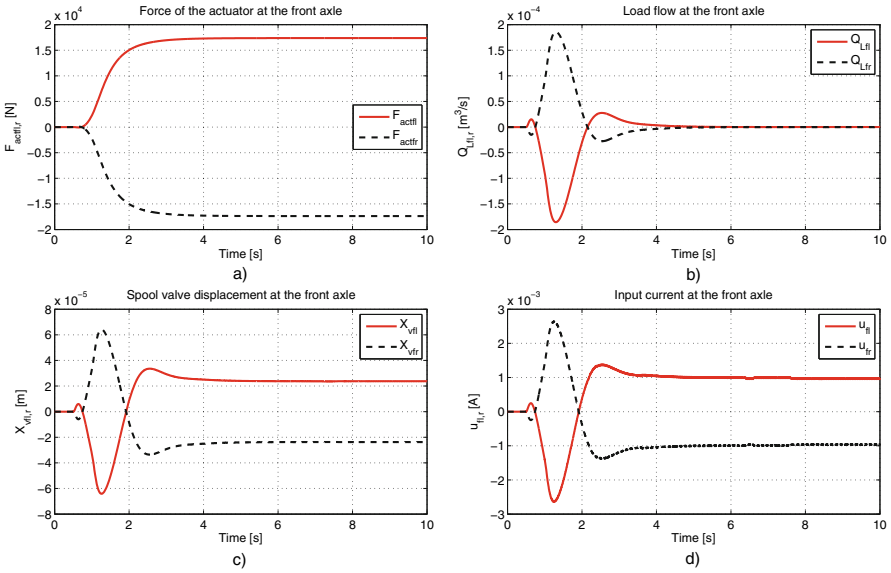


Fig. 13.21 The characteristics of the ESVH actuators at the front axle

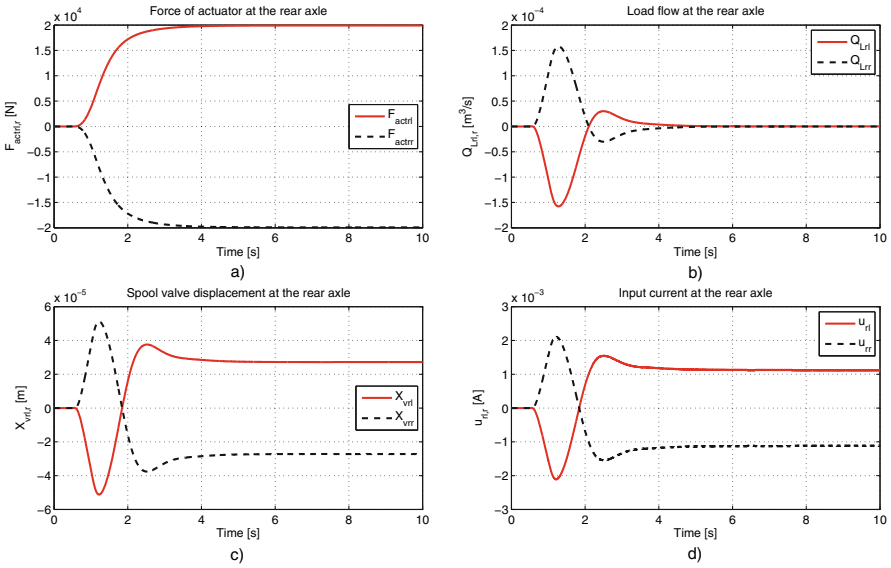


Fig. 13.22 The characteristics of the ESVH actuators at the rear axle

As mentioned in Sect. 13.2.1, the load flow through the hydraulic cylinder (Q_L) includes:

- **The first part** ($Q_{L1} = K_x X_v$) is the orifice load flow through the servo-valve and is adjusted by the movement of the spool valve displacement X_v.
- **The second part** ($Q_{L2} = K_P \Delta P$) is the internal leakage load flow inside the electronic servo-valve, which passes through the contact surface between the spool valve and the body of the servo-valve.

Therefore, the actual load flow between the two chambers of the ESVH actuator is as follows:

$$Q_L = Q_{L1} - Q_{L2} \tag{13.48}$$

The analysis in Fig. 13.23 is the detail of the load flow of the four ESVH actuators. Here, the first part Q_{L1} is shown in the dash-dot line, the second part Q_{L2} in the dash line and the actual load flow between the two chambers of the ESVH actuator in the solid line. We can see that in the case of the H_∞/LPV active anti-roll bar control system, due to the displacement of the spool valve (shown in Figs. 13.21c and 13.22c), the orifice load flows through the servo-valve $Q_{L1} = K_x X_v$ are balanced with the oil leakage inside the electronic servo-valve $Q_{L2} = K_P \Delta_P$. Therefore the actual load flow between the two chambers of the ESVH actuator Q_L becomes zero in the steady state manoeuvre.

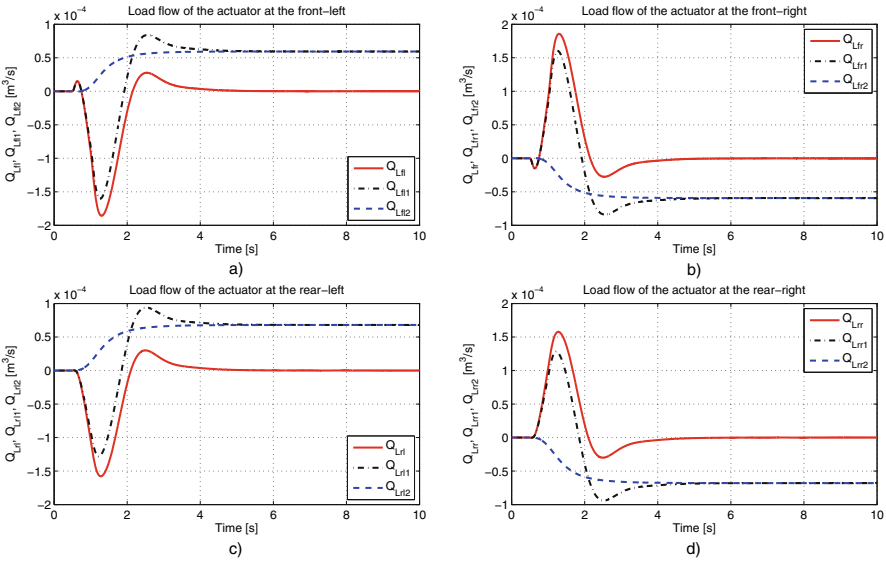


Fig. 13.23 The load flow of the four ESVH actuators in the steady state manoeuvre

The simulation results in the frequency and time domains, allow to conclude that with the nominal value of the total flow of pressure coefficient ($K_{PNominal} = 4.2 \times 10^{-11}\ \frac{m^5}{Ns}$), the proposed H_∞/LPV active anti-roll bar control system can guarantee the objective of improving the roll stability of heavy vehicles. It is obtained by creating a balance between the load flows generated by the controller with that of the oil leakage. However, in practice, the total flow of pressure coefficient K_P often changes depending on the working conditions and the life of the ESVH actuator. In the next section, we will consider the effect of the internal leakage inside the electronic servo-valve on the proposed H_∞/LPV active anti-roll bar control of the fully integrated model.

13.5.3 Effect of the Internal Leakage on the Performance of the H_∞/LPV Active Anti-Roll Bar Control System

To assess the influence of the internal leakage inside electronic servo-valve on the performance of the proposed H_∞/LPV active anti-roll bar control system, three cases of the total flow pressure coefficient (K_P) are considered and detailed as follows:

- First case: The 1st H_∞/LPV active anti-roll bar control system with the **nominal value** of the total flow pressure coefficient, $K_{PNominal} = 4.2 \times 10^{-11}\ \frac{m^5}{Ns}$.

- Second case: The 2nd H_∞/LPV active anti-roll bar control system with the **low** total flow pressure coefficient, $K_{PLow} = 5 \times 10^{-13}\ \frac{m^5}{Ns}$, (small internal oil leakage).
- Third case: The 3rd H_∞/LPV active anti-roll bar control system with the **hight** total flow pressure coefficient, $K_{PHight} = 1 \times 10^{-8}\ \frac{m^5}{Ns}$, (high internal oil leakage).

The three cases of the H_∞/LPV active anti-roll bar control systems presented above will be compared with the passive anti-roll bar case.

13.5.3.1 Analysis in the Frequency Domain

In this section, we consider the forward velocity at 90 km/h. As mentioned above, the main objective of the active anti-roll bar system is to reduce the normalized load transfers at all the axles. But we need also to pay attention to the actuator's saturation.

Figure 13.24 shows the transfer function magnitude of the normalized load transfers and input currents on the right at both axles. When the total flow pressure coefficient is low (K_{PLow}), the internal leakage inside the electronic servo-valve is small, so the 2nd H_∞/LPV active anti-roll bar control system needs lower

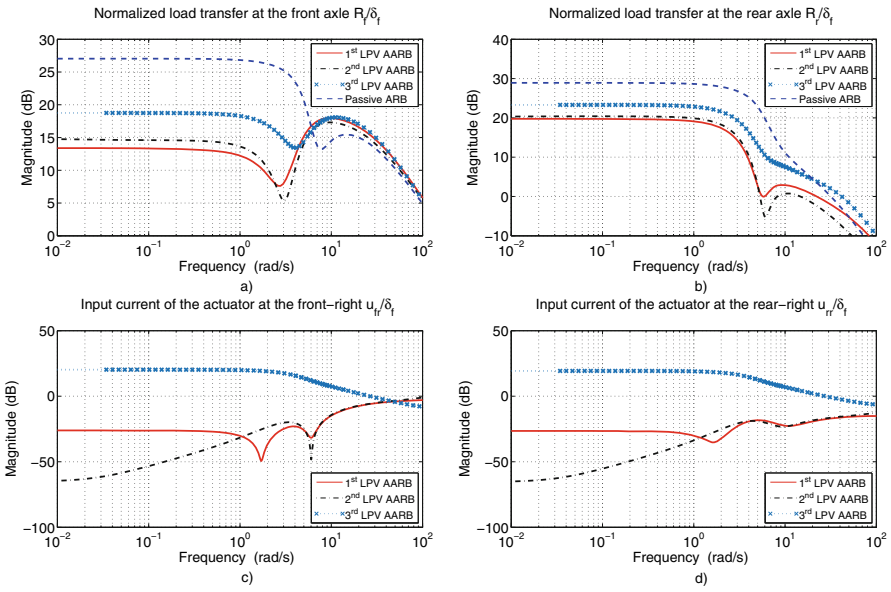


Fig. 13.24 The transfer function magnitude of (**a**, **b**) the normalized load transfers at the two axles $\frac{R_{fr}}{\delta_f}$, (**c**, **d**) the input currents of the actuators at the front-right $\frac{u_{fr}}{\delta_f}$ and at the rear-right $\frac{u_{rr}}{\delta_f}$

input current. In the high total flow pressure coefficient case (K_{PHight}), the internal leakage inside the electronic servo-valve is high, so the 3rd H_∞/LPV active anti-roll bar control system needs a huge input current, without having the ability to greatly improve the roll stability of heavy vehicles compared to the nominal case ($K_{PNominal}$).

The simulation results indicate that, if the total flow pressure coefficient (K_P) is greater than $3 \times 10^{-8}\ \frac{m^5}{Ns}$, the H_∞/LPV active anti-roll bar control system does not provide any improvement in the roll stability of heavy vehicles, despite the input current being very high. Indeed in these cases, the orifice load flow through the servo-valve generated by the movement of the spool valve cannot adapt to the oil leakage inside the electronic servo-valve.

13.5.3.2 Analysis in the Time Domain

In this section, some simulation results in the time domain are shown for four different cases: the three H_∞/LPV active anti-roll bar control systems and the passive anti-roll bar. The vehicle manoeuvre is a double lane change to overtake. The steering angle δ_f is shown in Fig. 13.25a.

The following scenario is used to evaluate the effect of the internal leakage inside the electronic servo-valve on the closed-loop system with the varying parameter $\rho = v$:

- The initial forward velocity is 80 km/h, the vehicle is running on a dry road ($\mu = 1$).
- When the driver wishes to overtake another vehicle, the driver will increase the throttle, so the total tractive force will increase from 0.5s to 2s in order to increase the forward velocity from 80 km/h up to 108 km/h. By ignoring the total rolling resistance and aerodynamic resistance forces, the forward velocity is kept constant at 108 km/h. At 6s the driver starts to brake to reduce the forward velocity of the vehicle to 90 km/h as shown in Fig. 13.25b. The total brake force will increase from 6s to 6.8s and then the driver will reduce the pressure on the brake pedal. The total tractive and braking forces are shown in Figs. 13.25c,d.

The differential equation for the forward velocity is determined as [30]:

$$m\dot{v} = \sum_{i=1}^{2} F_{ti} - \sum_{i=1}^{4} F_{bi} \tag{13.49}$$

where F_{ti} is the tractive force at each wheel drive, F_{bi} the braking force at each wheel.

Figures 13.25e,f show the normalized load transfers at both axles. In the cases of the three H_∞/LPV active anti-roll bar controllers, roll stability of the heavy vehicle is greatly improved, compared to the passive anti-roll bar (this is perfectly consistent with the results on the frequency domain). Figures 13.25g,h show the

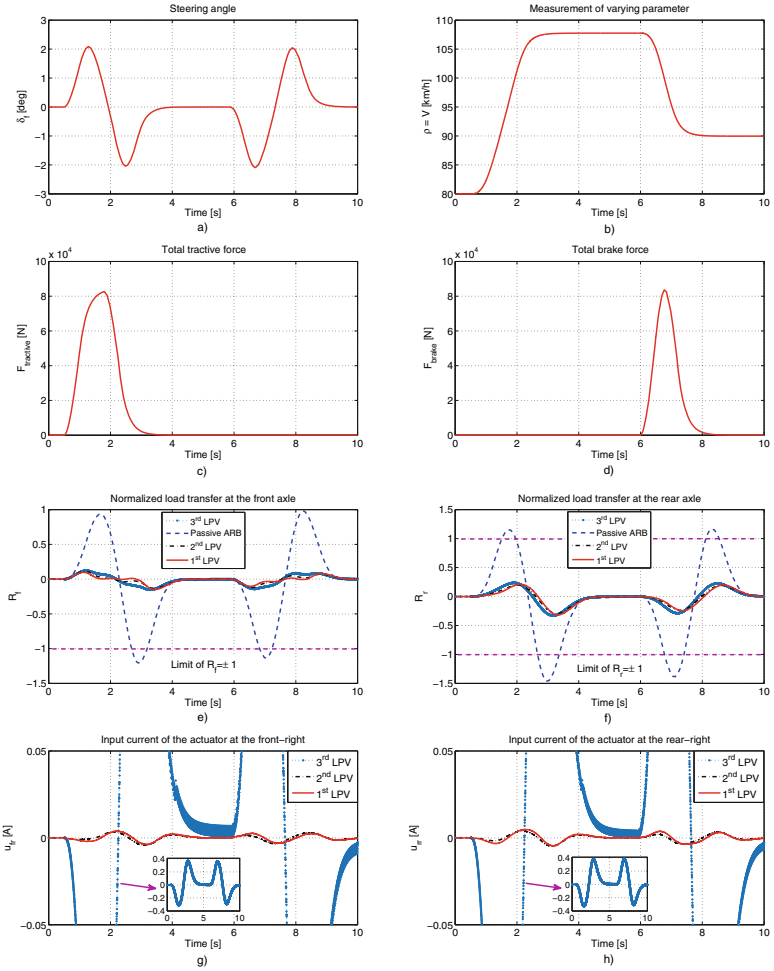


Fig. 13.25 Time responses of a single unit heavy vehicle in a double lane change manoeuvre to overtake

input currents at both axles. To create a good efficiency for roll stability (shown in Figs. 13.25e,f) in the case of the 3rd H_∞/LPV active anti-roll bar controller with the hight total flow pressure coefficient (K_{PHight}), one needs a huge input current (reaches 362 [mA] for the front axle, 382 [mA] for the rear axle), this value largely exceeds the limit allowed 20 [mA] [14]. Whereas in the cases of the 2nd H_∞/LPV controller with the low total flow pressure coefficient (K_{PLow}) and of the 1st H_∞/LPV controller with the nominal total flow pressure coefficient ($K_{PNominal}$), the maximum input currents required are less than 5 [mA]. The maximum absolute value of the normalized load transfers and the input currents are summarized in Table 13.5.

Table 13.5 Maximum absolute value of the signals in the double lane change manoeuvre to overtake

Anti-roll bar system	$\|R_f\|_{max}$	$\|R_r\|_{max}$	$\|u_{fr}\|_{max}$ [mA]	$\|u_{rr}\|_{max}$ [mA]
1st H_∞/LPV	0.149	0.317	4	4.7
2nd H_∞/LPV	0.132	0.294	3.5	4
3rd H_∞/LPV	0.153	0.329	362	382
Passive	1.207 (rollover)	1.46 (rollover)	0	0

As mentioned in the control objective, the main objective of the active anti-roll bar system is to improve roll stability, and vehicle rollover will occur if the normalized load transfer takes on the limit of ± 1 [32, 33]. However, it is also necessary to avoid the actuators' saturation, with 20 s[mA] being the maximum absolute value of the input current recommended by [14]. The survey results in both the frequency and time domains have shown that, to satisfy simultaneously the two goals of enhancing roll stability and avoiding the saturation of the actuators, the total flow pressure coefficient has to be kept in the range of $K_P = [5 \times 10^{-15}, 4 \times 10^{-10}] \frac{m^5}{Ns}$.

13.6 Conclusion

In this chapter, to evaluate the effect of the internal leakage inside the electronic servo-valve on the active anti-roll bar system of heavy vehicles, the authors used the fully integrated model including four ESVH actuators (two at the front and two at the rear axles) on a linear single unit heavy vehicle yaw-roll model. The nonlinear model of a single unit heavy vehicle is considered as a LPV model, where the forward velocity is considered as a scheduling parameter. The H_∞/LPV active anti-roll bar controller is synthesized in order to improve roll stability of heavy vehicles. The effect of the internal leakage inside the electronic servo-valve on the open-loop and closed-loop systems is analysed in detail. The survey results have shown that with the total flow pressure coefficient $K_P = [5 \times 10^{-15}, 4 \times 10^{-10}] \frac{m^5}{Ns}$, the two objectives of enhancing roll stability and avoiding the saturation of the actuators are simultaneously satisfied.

In actual use, the total flow pressure coefficient K_P will decline over time, because it depends on the quality of the contact surfaces between the spool valve and the body of the servo-valve. From the simulation results by using the H_∞/LPV active anti-roll bar controller in this chapter, we can confirm that the internal leakage inside the electronic servo-valve will affect the performance quality of the active anti-roll bar system of heavy vehicles. Therefore an LPV approach for fault tolerant control design will be also an interesting area for further research.

Acknowledgments This work has been supported by the University of Transport and Communications through the key project T2019-CK-012TD.

References

1. Evgenikos, P., Yannis, G., Folla, K., Bauer, R., et al. (2016). Characteristics and Causes of Heavy Goods Vehicles and Buses Accidents in Europe. *Transportation Research Procedia*, 14, 2158–2167.
2. Schwartz, SH., Fleming, SA. (2007). Motor Carrier Safety: A statistical approach will better identify commercial carriers that pose high crash risks than does the current federal approach. *USA Publication GAO-07-585*. US Government Accountability Office, Washington DC, US.
3. Knipling, P.R. (2007). The domain of truck and bus safety research. *Transportation Research Circular, No. E-C117*. Transportation Research Board, National Research Council, Washington DC, US.
4. NHTSA. (2012). https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812150
5. Miège, A., Cebon, D. (2002). Design and implementation of an active roll control system for heavy vehicles. 6th *International Symposium on Advanced Vehicle Control*. Hiroshima, Japan.
6. Chen, Bo., Peng, H. (2001). Differential-Braking-Based Rollover Prevention for Sport Utility Vehicles with Human-in-the-loop Evaluations. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 36(4), 359–389.
7. Yu, H., Guvenc, L., Ozguner, U. (2008). Heavy duty vehicle rollover detection and active roll control. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 46(6), 451–470.
8. Van Tan, Vu. (2017). Enhancing the roll stability of heavy vehicles by using an active anti-roll bar system. *PhD thesis*. University Grenoble Alpes, France.
9. Boada, M., Boada, B., Quesada, A., Gauchía, A., Díaz, V. (2007). Active roll control using reinforcement learning for a single unit heavy vehicle. 12th *IFToMM World Congress*. Besancon, France.
10. Miège Arnaud, J.P. (2000). Development of Active Anti-Roll Control for Heavy Vehicles. *PhD thesis*. University of Cambridge, UK.
11. Vladimir M., Zeljko S., Mario E. (2010). Robust H_∞ position control synthesis of an electro-hydraulic servo system. *ISA Transactions*, 49(4), 535–542.
12. Martin Choux. (2011). Nonlinear, Adaptive and Fault-Tolerant Control for Electro Hydraulic Servo Systems. *PhD thesis*. Technical University of Denmark, Denmark.
13. Kovari, Attila. (2009). Influence of cylinder leakage on dynamic behavior of electro-hydraulic servo system. 7th *International Symposium on Intelligent Systems and Informatics*. Subotica, Serbia.
14. Rafa, A., Yahya, A., Rawand, E. (2009). A Study on the Effects of Servo-valve Lap on the Performance of a Closed - Loop Electrohydraulic Position Control System. *Al-Rafidain Engineering*, 17(5), 1–14.
15. Mete K., Mustafa H. (2009). Mathematical modelling and fuzzy logic based position control of an electrohydraulic servosystem with internal leakage. *Mechatronics*, 19(6), 847–858.
16. Erylmaz, B., Wilson, B. (2000). Combining leakage and orifice flows in a hydraulic servo-valve model. *Journal of Dynamic Systems Measurement Control*, 122(3), 576–579.
17. Zulfatman, H., Rahmat, M., Husain, A., Ishaque, K., Ghazali, R., Ahmad, et al. (2014). Robust Position Tracking Control of an Electro-Hydraulic Actuator in the Presence of Friction and Internal Leakage. *Arabian Journal for Science and Engineering*, 39(4), 2965–2978.
18. Daniel, M., Kothapalli, Ganesh, M., et al. (2013). Position control of electro-hydraulic actuator system using fuzzy logic controller optimized by particle swarm optimization. *International Journal of Automation and Computing*, 10(3), 181–193.
19. Rahmat, M., Zulfatman, A., Husain, K., et al. (2011). Modeling and controller design of an industrial hydraulic actuator system in the presence of friction and internal leakage. *International Journal of the Physical Sciences*, 6(14), 3502–3517.
20. Renn, J., Wu, T. (2007). Modeling and control of a new 1/4T servo-hydraulic vehicle active suspension system. *Journal of Marine Science and Technology*, 15(3), 265–272.
21. Merritt, E. (1967). Hydraulic control systems. 3th *edition*. John Wiley & Sons. Isbn = 978-0-471-59617-2.

22. Rydberg, K. (2016). Hydraulic Servo Systems: Dynamic Properties and Control. 3rd *edition*. Linköping University Electronic Press. Isbn = 9789176856208.
23. Gaspar, P., Bokor, J., Szaszi, I. (2004). The Design of a Combined Control Structure to Prevent the Rollover of Heavy Vehicles. *European Journal of Control*, 10(2), 148–162.
24. Van Tan V., Sename, O., Dugard, L., Gaspar, P. (2017). H_∞/LPV controller design for an active anti-roll bar system of heavy vehicles using parameter dependent weighting functions. *Heliyon*, 5(2019), 1–11.
25. Gaspar, P., Bokor, J., Szaszi, I. (2005). Reconfigurable control structure to prevent the rollover of heavy vehicles. *Control Engineering Practice*, 13(6), 699–711.
26. Van Tan V., Sename, O., Dugard, L., Gaspar, P. (2016). Active anti-roll bar control using electronic servo-valve hydraulic damper on single unit heavy vehicle. 8th *IFAC Symposium on Advances in Automotive Control*. Norrköping, Sweden.
27. Gaspar, P., Szabo, Z., Bokor, J. (2005). Prediction based combined control to prevent the rollover of heavy vehicles. 13th *Mediterranean Conference on Control and Automation*. Limassol, Cyprus.
28. Huang, H., Rama, K., Dennis, A. (2012). Active roll control for rollover prevention of heavy articulated vehicles with multiple-rollover-index minimisation. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 50(3), 471–493.
29. Van Tan V., Sename, O., Dugard, L., Gaspar, P. (2019). An Investigation into the Oil Leakage Effect Inside the Electronic Servovalve for an H_∞/LPV Active Anti-roll Bar System. *International Journal of Control, Automation and Systems*, 17(X), 1–12.
30. Wong, J.Y. (2008). Theory of ground vehicles. 4th *edition*. John Wiley & Sons. Isbn = 978-0-470-17038-0.
31. Sampson, D., Cebon, D. (2003). Achievable roll stability of heavy road vehicles. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 217, 269–287.
32. Sampson, D., Cebon, D. (2010). Active Roll Control of Single Unit Heavy Road Vehicles. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 40(4), 229–270.
33. Van Tan V., Sename, O., Dugard, L., Gaspar, P. (2017). Enhancing roll stability of heavy vehicle by LQR active anti-roll bar control using electronic servo-valve hydraulic actuators. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 55(9), 1405–1429.

Chapter 14
Thermal Comfort and Game Theory

Javad Khazaii, Ali Khazaei, Hamid Khayyam, and Reza N. Jazar

Abstract Creating and maintaining thermal comfort for people who are working and living inside the buildings is the primary target of engineers who are designing building environment control system. Since thermal comfort by its definition is the state of mind and is different for different people, calculating, measuring, and controlling thermal comfort for everybody with similar and unique values are not sufficient. Following the current applicable thermal comfort standard as it is done in the professional industry will lead to setting and using similar temperature and relative humidity values for all people. Therefore, this is done without implementing direct comfort preferences of actual people who are working and living in the space. In this chapter, we introduce a novel method for calculating and maintaining the maximum possible thermal comfort for all the occupants in a building. We use the rules of game theory to propose a new way to collect, analyze, calculate, and maintain the maximum possible thermal comfort for all the actual occupants in one building with a holistic approach. In this approach, we categorize all the building occupants into a few subgroups of the people with similar thermal comfort preferences and use these actual preferences along with the game theory rules to find best responses of all occupants. The environment control system then will be programmed to operate based on these best response values and not based on same temperature and humidity ratio at all time as it is being done currently.

Keywords Game theory · Thermar comfort · Fanger method · HVAC system · Thermal conditions

C. C. N. To (✉)
Vocational Training Council (VTC), Tsing Yi Island, New Territories, Hong Kong
e-mail: catter@vtc.edu.hk

H. Marzbani · R. N. Jazar
School of Engineering RMIT University, Melbourne, VIC, Australia

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_14

14.1 Introduction

Human thermal comfort in buildings is addressed by ASHRAE Standard 55 [1], and it is measured based on a relatively old approach developed by Fanger which defines two concepts of PMV (predicted mean vote) and PPD (percentage people dissatisfied). One of the notable issues with this approach is that the data which this method is built based on is developed using responses of a specific group of people at that time epoch to Fanger's questionnaire. The structure of the work force in buildings (e.g., office buildings) has been changed since then considerably. New collected data shows the predicted results of Fanger method lack considerable accuracy.

In this work we will present an innovative method which is good for using for different office buildings, based on the questionnaire issued to the current occupants of each building. The collected information then will be used along with rules of game theory to maximize the possible comfort level for all the occupants of each building. In fact a custom algorithm will be developed for each building and its specific group of peoples. The logic will be added into control sequences of the air handling units. These units then will be able to switch important thermal comfort parameters between different set points to maintain the maximum possible comfort level for all the building occupants at the same time. The method can be further advanced in the future to enable the control algorithm of the building to function as a real-time thermal comfort control algorithm.

The target of this work is to utilize game theory concept in order to customize the control sequence operation of air handling units based on the real preference of all the occupants, living and working in building that are served with these units, and not based on a globally agreed value that may or may not satisfy the actual occupants of the building.

As it is stated above, this work is built upon investigating the application of game theory in occupant thermal comfort controls. Therefore, we've dedicated the next few paragraphs to introducing both of these concepts briefly.

14.2 Thermal Comfort

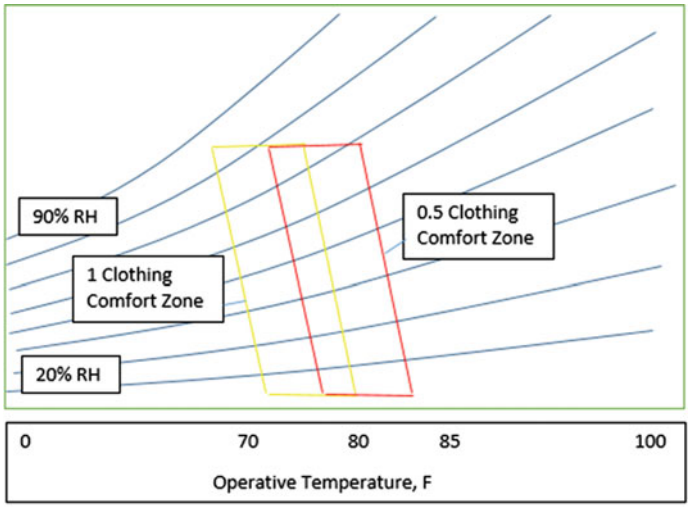
Human thermal comfort is measured based on personal and environmental factors. Personal factors are metabolic rate and clothing insulation level, and environmental factors are temperature, mean radiant temperature, humidity ratio, and air velocity. Human body continuously generates heat. This generated heat is due to biological processes that happen inside the body of person which can happen without the person consciously knowing, and it can be created even without presence of any visible physical activity of the person. So even if the person seats in rest and is without any visible activity, this person still generates some heat. When a person engages in a specific activity, part of this heat will be consumed in the form of

work. The remainder of the generated heat (that part that was not consumed for performing activities) has to be exchanged with the environment. If this remaining heat generation cannot be balanced through exchange with environment, it will cause the body temperature to either rise or fall. Both events create situations of discomfort at first and if the condition continues can be dangerous and even fatal. In general, the body regulates the generated heat in order to keep the core parts of the body in the comfortable condition of 37F, while the temperature of the skin in comfort condition is a few degrees less than that. When the body is exposed to cold environment conditions, it will act such that it collects the heat from the noncore parts of the body in order to keep the core temperature at 37F. This will cause the temperature of the rest of the body including skin temperature to drop below the comfortable level. To the contrary, if the body is exposed to the high temperature condition, it acts in such a manner to protect the core sections of the body at 37F by increasing the temperature of the noncore parts of the body. These two conditions change the proper level of stored heat in the body. If the uncomfortable cold and heat environment persists, the levels of stored heat in the body move in two extreme conditions that can be dangerous and even fatal.

The main duty of the HVAC system is to keep the indoor conditions such that it helps people live and work indoor comfortably and maintain the proper balance between the heat generated in the person's body and the environment. Therefore, in order to set the space condition to a level that keeps people in comfortable conditions, a detailed analysis of the heat exchange types and quantities should be done. Three main body heat analyses that have been done and have been used for this purpose are Fanger's method [2], Gagge's two-node method, [2] and finally ASHRAE method [3]. It should be noted that all three methods are very similar in basic approach to the evaluation of the comfort level, with slight difference in the detail calculation assumptions. Part of the metabolic heat generated in the body will be used to perform work. The rest of the heat generated should be balanced with the environment through heat exchange to keep the body temperature constant. Heat exchange between the body and the environment happens in all the form of conduction, convection, radiation, and evaporation. In general, all the research provided about the heat exchange between the body and the environment ignore the conduction heat exchange due to its negligible quantity. Part of the heat exchange is due to sensible and latent heat exchange of inhaled and exhaled air in and out of our lungs, as ambient air with its temperature and humidity content gets in the person's lungs and leaves the lungs with its new temperature and humidity content due to the lung temperature and humidity contents. The remainder of the heat will reach the skin and will be transferred to the air through evaporative effects of sweating and also sensible heat transfer due to radiation and convection.

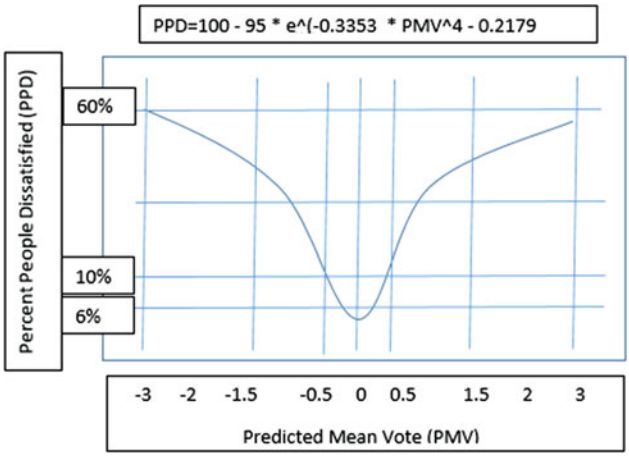
Based on the level of activity (e.g., office work) and standard type of clothing in winter and summer, ASHRAE Standard 55 [1] has designated a specific area in a psychometric chart that majority of people are claimed to reach a state of mind that is called thermal comfort when the environmental condition lies within the borders of this area. The issue is that these results are not very accurate. "The PMV/PPD model has a low prediction accuracy [4]. Using the world largest thermal comfort

field survey database [5], the accuracy of PMV in predicting occupant's thermal sensation was only 34%, meaning that the thermal sensation is correctly predicted one out of three times" [6]. In addition, HVAC engineers are generally designing the building air conditioning systems to provide and maintain a constant set point of 75 degree F dry-bulb and 50% relative humidity in summer and 70 degree F and a relative humidity close to 50% everywhere. In these systems, temperature is being controlled via operation of cooling or heating coils, and relative humidity is being controlled via humidifiers and dehumidification. As it is said earlier, since the thermal comfort is a subjective concept, and it is different for different people, with the generalization represented in ASHRAE Standard 55 [1] and typical engineering design universal values, it is well possible that many people living and working in the same office building, which have different preferences regarding reaching thermal comfort, feel uncomfortable while the air conditioning system is operating even more than it really needs to. Presented method in this work is meant to use more accurate engineering values as plugged in the sequence of controls of the air handling unit to address satisfaction of all of the people living and working in the building based on their own preferences and not a preselected universal value.



[h] Thermal Comfort zone [1]

Thermal Comfort zone [1]



[h] Predicted Mean Vote vs Percent People Dissatisfied [1]

Predicted Mean Vote vs Percent People Dissatisfied [1]

14.3 Game Theory

The Merriam-Webster definition of game theory is "game theory is a tool used to analyze strategic behavior by taking into account how participants expect others to behave. Game theory is used to find the optimal outcome from a set of choices by analyzing the costs and benefits to each independent party as they compete with each other." In general for performing the game, we need players, definite number of strategies for each player, and quantified incentives for each strategy for each player. Game theory can be played in both competitive and cooperative modes, when in competitive games players tend to reach the maximum benefit individually while in cooperative games players intend to make the maximum benefit for both (all) sides. Example of a competitive game is competition between two industrial giants to come up with strategies to make maximum benefit from their products in expense of the other, and example of cooperative game is cooperation between different teams within the same company to provide the maximum benefit for each of them and therefore for the company as whole. Within the structure of games, concepts such as Nash equilibrium, "combination actions for all players with the characteristics that no single player can obtain a higher outcome by deviating unilaterally from this strategy"; Pareto optimal, "when the problem only have one Nash equilibrium"; best response, "best strategy for one player, knowing all the other players' strategies"; mixed strategy, "when the players will not choose one definite strategy but, instead, choose according to a probability distribution over his strategies"; etc. are used to reach an equilibrium point. Such equilibrium point is the best outcome that all the players can achieve. The benefits are usually represented by utility functions (values) that represent the degree of satisfaction of the player from that outcome.

Even though the majority of the usage of game theory have been social and financial applications, in recent years, some attempts have been done to utilize this tool for engineering applications as well.

"In a good attempt for utilizing game theory method in HVAC and energy conservation, Forouzandehmehr et al. [7] presented a solution for controlling HVAC systems in smart buildings using concepts of game theory. Their proposed learning algorithm makes it possible to minimize the cost of keeping maximum zones in the proper temperature range" [8]. Also "Coogan et al. [9] provided a framework for building energy management setting to incentivize desirable control policies among the building occupants based on Nash equilibrium and showed the technique on an eight zone, physics-based building model can save around 3–4% in energy cost" [8]. Khazaii in his book [8] looked at applications of game theory in competition between energy provider and energy consumer and also thermal comfort in lab environment.

"Game theory is concerned with the actions of decision-makers who are conscious that their actions affect each other" [10]. Game theory is played in two forms, extensive and normal. In this work, we have utilized the normal form analysis, and therefore in this section we will only focus on describing concepts relevant to our work (normal form). A game in a normal form usually consists of three components. The first part is a finite set of players, the second part is a set of available strategies for each player, and the third component is a set of payoffs for each player. For example, "If there are two players, player 1 has m strategies, and the player 2 has n strategies, and then the game in normal form can be represented by an m $\times n$ matrix of ordered pairs of numbers" [11]. "Standard equilibrium analyses assume that all players (1) form beliefs based on an analysis of what others might do (strategic thinking); (2) choose the best response given those beliefs (optimization); and (3) adjust best responses and beliefs until they are mutually consistent (equilibrium)" [12].

Two basic types of strategies are pure and mixed strategies. A pure strategy is a strategy that one player will choose for any situation he/she faces. A player's strategy set is the set of pure strategies available to that player. On the other hand, a mixed strategy is an assignment of probability to each pure strategy. Game theorist defines two types of dominated strategies that are strictly dominated and weakly dominated strategies. A strictly dominated strategy for one player is a strategy that for every choice of strategies from the other players, player one's selected strategy gives him/her a higher payoff than any other one of his/her strategies. A weakly dominated strategy for one player is a strategy that for all the choice of strategies by the other players, player one's strategy gives him/her a greater or equal payoff relative to his/her other strategies. By game theory definition, players are defined to be rational, and therefore a rational player will not use a strategy that is weakly dominated by another one of his strategies.

The best response is a strategy (or strategies) which produces the most favorable outcome for a player, taking other players' strategies as given [13]. Best response is the central definition in a Nash equilibrium that is the one that "No player can profitably deviate, given the actions of the other players" [14].

There is a plethora of books and resources for learning and using game theory. In the following paragraphs, we have utilized a small section of this extremely powerful tool to create a novel method for the advancement of science of thermal comfort control.

14.4 Proposed Method

The proposed method in this work requires that every year at the beginning of the cooling season (we only focus on cooling season in this work), the building owner/manager performs 9 days of test for its entire building and employees. Each day the building air conditioning system will be set to one of the nine combination of 72, 75, and 78 F and lower than 50% and 50% (controlled) and higher than 50% relative humidity. The building management in these 9 days will distribute surveys that question each employee to express his/her level of thermal comfort in each specific date (on a 0 to 10 level with 0 being the least comfortable and 10 the most comfortable condition) with its associate combination of temperature and humidity ratio produced by the air conditioning system. At the end of this 9 days of testing, there will be a pool of collected values representing the level of comfort of all employees in this building under different air delivery condition.

The next step is to group the occupants of a building to two random and equal groups, each group with its averaged specific preferences (utilities) to temperature and relative humidity. The next step is to use these preferences and evaluate them with game theory different techniques in order to customize the control sequence of operation for the air conditioning system in that building for the upcoming cooling season. The preferences (utilities) used in this work are assumptions only, and the purpose of the work is to introduce an innovate method of thermal comfort control, proven by the basic rules of game theory. This method if used in control of air handling units can lead to a better comfort for the majority of people and also saving energy by operating air handling unit based on this custom control sequence.

14.5 Results and Discussion

Let's assume each person has specified his/her comfort level (utility) when exposed to a combination of 72, 75, and 78 F and 40, 50, and 60% relative humidity and we have compartmented people in two groups as shown in Table 14.1 with utility of zero representative of the least and one representative of the most comfortable condition, with first number in each box representative of the first group of people comfort level and the second number in that box representative of the comfort level of the second group of people. The intention here is to control temperature and humidity simultaneously to provide the optimum comfort level for both groups of occupants.

Table 14.1 Game theory (1) setup of preferences of people temperature and relative humidity

T (F)↓ /RH→	40%		50%		60%	
72	0.3	1	0.4	0.8	0.6	0.7
75	1	0.3	0.8	0.6	0.4	0.3
78	0.4	0.6	0.6	0.7	0.5	0.8

Table 14.2 Best response analysis for game represented by Table 14.1

T (F)↓ /RH→	40%		50%		60%	
72	0.3	1	0.4	0.8	0.6	0.7
75	1	0.3	0.8	0.6	0.4	0.3
78	0.4	0.6	0.6	0.7	0.5	0.8

From here on, we use different techniques of game theory to find optimum control points for our air conditioning system. Easily looking into Table 14.2, one can say that the best outcome for first group of people for thermal comfort is achieved at 75 F and 40% RH due to allocation of the utility equal 1 to this group at this point. Similarly it can be seen that the most favorite condition for people in group 2 is 72 F and 40% RH since this group utility is equal to 1 at this point, but the challenging work is to find the best simultaneous outcome for both group. By investigation, we can find that there is not a "dominant strategy" for neither of the groups (i.e., there is not one single strategy for group 1 that by changing that irrespective of either of the strategies of group 2, group 1 can get a better utility and vice versa). But by using "best response" method, we can find that strategy (0.8, 0.6) is a best response for both groups and therefore a Nash equilibrium (Table 14.2). It means that at this point, neither group can individually change its strategy and get a better outcome. This point is also a Pareto optimal point since there is not existed another point in this table that both groups of people simultaneously can get a higher outcome from 0.8 and 0.6 (If such point existed because of the nature of this problem, we can change the game to a cooperative form of the game and choose that point for point of control.).

In this simple case control system for this building should use 75 F, 50% strategy all the time, and that provides the best level of comfort possible for both groups of people with 0.8 and 0.6 comfort payoffs (please do not forget that we are using hypothetical values of preferences in this work, not real values).

Let's change the utilities just a little amount to avoid a simple solution similar to Table 14.2 and discuss the game in more detail with another table (Table 14.3). In this case, let's assume the responses of two groups of occupants to different combinations of temperature and humidity ratio to be tabulated as follows, so by primary investigation we cannot find any point of "best response" or Nash Equilibrium with the data as it is presented here.

In this case, the control system should be programmed for a mixed condition in which either of the occupants will be worse off if control system changes either

Table 14.3 Game theory (2) setup of preferences of people temperature and relative humidity

T (F)↓ /RH→	40%		50%		60%	
72	0.8	1	0.7	0.8	0.6	0.7
75	1	0.6	0.6	0.7	0.7	0.8
78	0.7	0.3	0.7	0.6	0.8	0.5

Table 14.4 Partial analysis (1) for game represented by Table 14.3

T (F)↓ /RH→	40%		50%		60%	
72	0.8	1	0.7	0.8	0.6	0.7
75	1	0.6	0.6	0.7	0.5	0.8
78	0.7	0.3	0.7	0.6	0.8	0.5

temperature or relative humidity levels independently from that (these) pivotal point(s). To do that let's examine all possible strategies.

First as it was noted earlier, try to search for a pure Nash equilibrium point or intersection of "best responses" for both groups. It is obvious that such point is not existed in Table 14.3. At this point, we need to move to find a mixed strategy for the game, and we start our investigation within subsets of the main table. Let's assume the strategies for temperature are 72 and 75 and not 78 F and strategies for relative humidity are 40% and 50% and not 60% (Table 14.4). In this case, since there is neither a dominated nor a pure Nash equilibrium strategy here, we need to check for a possible mixed strategy for this subset. If we assign p, 1-p, q, and 1-q probabilities for first and second strategies of group 1 and 2, respectively, then for group 1, probability results should comply with:

$$1 \times p + 0.6 \times (1 - p) = 0.8 \times p + 0.7 \times (1 - p) \geq 0.7 \times p + 0.8 \times (1 - p) \tag{14.1}$$

This means group 1 should set its probabilities of p and 1-p such that group 2 first and second strategies should provide same outcome, and this outcome should also be equal and larger than utility that it gets from the third strategy. Of course by solving these three equations, we can find out that there is no answer for this situation and therefore this condition is impossible and there is no NE in this subset. If there was an answer for this, we should then investigate the same condition for group 2 as well.

Now let's assume the strategies for temperature are 72 and 75 and not 78 F and strategies for relative humidity are 40% and 60% and not 50% (Table 14.5).

In this case, since there is neither a dominated nor a pure Nash equilibrium strategy, we need to check for a possible mixed strategy. If we assign p, 1-p, q, and 1-q probability distribution for first and second strategies of group 1 and 2, respectively, then for group 1 probability distribution, we should comply with:

Table 14.5 Partial analysis (2) for game represented by Table 14.3

T (F)↓ /RH→	40%		50%		60%	
72	0.8	1	0.7	0.8	0.6	0.7
75	1	0.6	0.6	0.7	0.5	0.8
78	0.7	0.3	0.7	0.6	0.8	0.5

Table 14.6 Full option analysis for game represented by Table 14.3

T (F)↓ /RH→	40%		50%		60%	
72	0.8	1	0.7	0.8	0.6	0.7
75	1	0.6	0.6	0.7	0.7	0.8
78	0.7	0.3	0.7	0.6	0.8	0.5

$$1 \times p + 0.6 \times (1 - p) = 0.7 \times p + 0.8 \times (1 - p) \geq 0.8 \times p + 0.7 \times (1 - p) \tag{14.2}$$

That leads to p = 2/5 and 1-p = 3/5.
And for player 2 probability distribution, we should comply with:

$$0.8 \times q + 0.6 \times (1 - q) = 1 \times q + 0.5 \times (1 - q) \geq 0.7 \times q + 0.8 \times (1 - q) \tag{14.3}$$

which the equality part of the equation leads to q = 1/3 and 1-q = 2/3. But the inequality part of the equation is impossible, and therefore there is no Nash equilibrium here either.

We can research the rest of the subsets to see if there is a Nash equilibrium in either of them. If there is a Nash equilibrium in any of the subsets, then that point can be selected to be used for controlling the mixture of temperature and humidity ratio for this group of occupants. If not we can continue our search for mix strategy in the full game table (Table 14.6) as follows:

In this case and in order to find the equilibrium point, let's assign p, q, and 1-p-q probability ratios to the first, second, and third strategy for the temperature and r, t, and 1-r-t probability ratio to the first, second, and third strategy of the humidity ratio. The idea is to find temperature probability ratios p, q, and 1-p-q in a way that all three options of humidity ratio provide the same overall utility outcome. In return humidity ratio probability ratios r, t, and 1-r-t should be selected in such a manner that temperature strategies' overall utility for its all three strategies be the same as well. Therefore, the solution to the problem will be achieved by solving the following two sets of equations: all equations must be done by equation editor

$$1 \times p + 0.6 \times q + 0.3 \times (1 - p - q) = 0.8 \times p + 0.7 \times q + 0.6 \times (1 - p - q) = 0.7 \times p + 0.8 \times q + 0.5 \times (1 - p - q) \tag{14.4}$$

and

$$0.8 \times r + 0.7 \times t + 0.6 \times (1 - r - t) = 1 \times r + 0.6 \times t + 0.7 \times (1 - r - t) = 0.7 \times r + 0.7 \times t + 0.8 \times (1 - r - t) \tag{14.5}$$

By solving first set and second set separately, we will be able to calculate p = 4/10, q = 5/10, 1-p-q = 1/10, and r = 1/4, t = 5/8, and 1-r-t = 1/8 and satisfy the equations completely. Therefore we have found a mixed strategy for this game.

That means the optimum control for these two occupant groups can be achieved when the control system randomly switches between occupant group 1's strategies 4/10, 5/10, and 1/10 of the time and occupant group 2's strategies 1/4, 5/8, and 1/8 time, respectively. In other words, the control system shall switch between 72, 75, and 78 F and 40%, 50%, and 60% relative humidity with the abovementioned probabilities.

Finally, as the result of these calculated probability ratios, the utility of each of them will be calculated to be 0.53 and 0.73, respectively, as follows:

$$\left(\frac{4}{10}\right) \times \left(\frac{1}{4}\right) \times 0.8 + \left(\frac{4}{10}\right) \times \left(\frac{5}{8}\right) \times 0.7 + \left(\frac{4}{10}\right) \times \left(\frac{1}{8}\right) \times 0.6 + \left(\frac{5}{10}\right) \times \left(\frac{1}{4}\right) \times 0.6 + \left(\frac{5}{10}\right) \times \left(\frac{5}{8}\right) \times 0.6 + \left(\frac{5}{10}\right) \times \left(\frac{1}{8}\right) \times 0.8 + \left(\frac{1}{10}\right) \times \left(\frac{1}{4}\right) \times 0.7 + \left(\frac{1}{10}\right) \times \left(\frac{5}{8}\right) \times 0.7 + \left(\frac{1}{10}\right) \times \left(\frac{1}{8}\right) \times 0.8 = 0.712 \tag{14.6}$$

and:

$$\left(\frac{4}{10}\right) \times \left(\frac{1}{4}\right) \times 1 + \left(\frac{4}{10}\right) \times \left(\frac{5}{8}\right) \times 0.8 + \left(\frac{4}{10}\right) \times \left(\frac{1}{8}\right) \times 0.7 + \left(\frac{5}{10}\right) \times \left(\frac{1}{4}\right) \times 0.6 + \left(\frac{5}{10}\right) \times \left(\frac{5}{8}\right) \times 0.7 + \left(\frac{5}{10}\right) \times \left(\frac{1}{8}\right) \times 0.8 + \left(\frac{1}{10}\right) \times \left(\frac{1}{4}\right) \times 0.3 + \left(\frac{1}{10}\right) \times \left(\frac{5}{8}\right) \times 0.6 + \left(\frac{1}{10}\right) \times \left(\frac{1}{8}\right) \times 0.5 = 0.73 \tag{14.7}$$

These are best possible outcomes that either of the occupants can achieve. It means that by following this control algorithm, occupant group 1 can achieve 0.712 and occupant group 2 can achieve 0.73 level of satisfaction. There is no other combination that can provide higher satisfaction for both groups simultaneously.

A typical sequence of operation for controlling air handling units in practice currently reads similar to this "building control system shall modulate control valves of cooling/heating coils and enable humidifier to provide a discharge temperature and humidity ratio to maintain temperature (75 F) and humidity ratio (50%) as it is scheduled in building control system in the space." With the new presented method here and based on the calculated values in this exercise, this sequence can be revised to read "building control system shall modulate control valves of cooling/heating coils and enable humidifier to provide discharge temperature and humidity ratio to maintain temperature (72, 75, 78 F) and humidity ratio (40%, 50%, 60%) and changes these settings randomly with a ratio of 0.4,0.5,0.1 for temperature and ratio of 1/4,5/8,1/8 for humidity ratio."

14.6 Conclusion and Future Work

A novel method for controlling the thermal comfort in buildings has been presented here. We used randomly generated utilities (preferences) for a typical sample of occupants in the building and used the basics of the game theory to specify the optimum level of comfort for all the people in the building with a holistic view. We have shown that all the occupants of a building can be categorized under two main groups with similar preferences toward two of the main factors that can contribute to thermal comfort. These two factors are temperature and relative humidity. We then suggested that the control sequences for the building air handling unit shall follow algorithms generated based on the game theory rules for controlling these factors instead of the generic current method of setting up these values the same for every occupant and every occasion. By moving away from generic set points for operation of the air handling unit (usually set at 75 F and 50% relative humidity as industry standard) to custom-made set points, there will be opportunity for saving energy as well, by partially deviating from the most energy-consuming options of thermal comfort. This work represents a holistic approach for the whole building occupant thermal comfort. The future work should be done in a more detailed depth such as strategies for room by room environment control strategies. This work is done based on categorizing all occupants in two groups, but more work can be done by dividing occupants to three or more group as well and even larger combination of temperature and humidity ratio options. We can also include other factors of thermal comfort such as air velocity delivered to the space in our calculations.

Game theory can also be used for selecting strategies for energy purchase from providers with variable electricity rates. Game theory has offered groundbreaking innovations in different fields of science, and this work provides a first step in utilizing this method for the thermal comfort and energy savings research.

References

1. Standard 55-2017 – Thermal Environmental Conditions for Human Occupancy (ANSI-ASHRAE); American Society of Heating, Ventilation and Refrigeration Engineers
2. Auliciems Andris and Szokolay, S.V.; Thermal Comfort, Passive and low energy architecture, in association with department of architecture, The university of Queensland Brisbane 4072, Second revision 2007.
3. ASHRAE Applications Handbook, 2017, American Society of Heating, Ventilation and Refrigeration Engineers.
4. Cheung, Toby; Schiavon, Stefano; Parkinson, Thomas; Li, Peixian; Brager, Gail (2019-04-15). "Analysis of the accuracy on PMV – PPD model using the ASHRAE Global Thermal Comfort Database II".
5. Földváry Ličina, Veronika; Cheung, Toby; Zhang, Hui; de Dear, Richard; Parkinson, Thomas; Arens, Edward; Chun, Chungyoon; Schiavon, Stefano; Luo, Maohui (2018-09-01). "Development of the ASHRAE Global Thermal Comfort Database II".
6. https://en.wikipedia.org/wiki/Thermal_comfort
7. Forouzandehmehr, N., Perlaza, S.M., Han, Z, Poor, H.V.; 2014; Distributed control of heating, ventilation and air conditioning systems in smart buildings.
8. Khazaii, Javad; 2016; Advanced decision making for HVAC engineers, Creating energy efficient smart buildings. Springer
9. Coogan S., Ratliff, L.J., Caldrone, D., Tomlin, C., Sastry, S., 2013; Energy Management via pricing in LQ dynamic games.
10. Rasmusen, E., Blackwell, B.; 2006; Games and information, An Introduction to Game Theory; Fourth Edition.
11. Schecter, S., Gintis,H. ; 2017; Introduction to Game Theory
12. Camere, C.F., Ho, T.H., Chong, J.K.; 2004; Behavioral Game Theory: Thinking, Learning and Teaching. Springer
13. Fudenberg, D., Tirole, J.; 1991; Game Theory, Cambridge, Massachusetts; MIT Press
14. Osborne, M.J., Rubinstein, A.; 1994; A course in Game Theory, Cambridge, Massachusetts; MIT Press

Chapter 15
Wind Resource Assessment

Stefania Castelletto and Alberto Boretti

Abstract In this chapter, we will review the fundamental science related to wind origin and variability and the modeling of such variability based on empirical and fundamental laws. We will describe the currently accepted methodologies to assess wind energy potentials based on the literature review of key past research findings on this topic with implications on the current practices. We will focus on wind variability on the long and medium time scale of years and wind statistics based on the medium-short-term average (1 h). We will describe the use of specific software as well as current statistical and physical models together with common assumptions, which underpin the wind energy sector to determine wind farm siting adequacy. We will show the applications of this wealth of knowledge to specific case studies and provide insight into the currently accepted methods to determine wind energy generation based on medium-term time series data average. We will identify current gaps in this knowledge about short-term time variability (minutes) and implications to wind energy production for grid uptake in the conclusions.

Keywords Renewable energy · Wind energy · Wind statistics · Wind energy variability · Wind energy production

15.1 Introduction to Wind Speed and Energy

The origin of the wind is due to the primary source of energy on earth – and as such the primary source of renewable energy – which is the Sun. The Sun unevenly warms up different areas of the globe creating pressure differences and thus large air current circulations. This combined with the rotational motion of the Earth is

C. C. N. To (✉)
Vocational Training Council (VTC), Tsing Yi Island, New Territories, Hong Kong
e-mail: catter@vtc.edu.hk

H. Marzbani · R. N. Jazar
School of Engineering RMIT University, Melbourne, VIC, Australia

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9_15

the origin of the global wind patterns. Winds blow across the Earth from high-pressure areas to low-pressure areas, as such local variation during seasons, land/sea distribution, valleys, and hills, contribute to further variability of wind patterns at local levels.

The physics of wind energy is described by the kinetic energy associated with a moving mass of air, which can be expressed in terms of wind power or most commonly as wind power density. This energy is captured by a wind turbine, consisting of a rotor with one or more blades, which is mechanically attached to an electrical generator to provide electrical energy. A wind turbine converts the mechanical energy in the wind to electricity by harvesting wind energy via rotating blades around a rotor. First, the wind energy is converted to mechanical rotational energy due to the differential pressure exerted along with the blades by the upstream and downstream wind speeds, which produces resultant forces along the plane of rotation due to the projection of lift and drag forces. These resultant forces applied to the distance of the blade from the rotational axis provide torque or momentum at a specific point in the blade, which applied all along the blades determine the mechanical power output of the turbine.

The wind turbine transmits the mechanical power using a gear-box or other power transmission systems to an electrical generator, whose power is then conditioned to the utility grid and delivered to the consumers. The rotor turns the driveshaft, which turns an electric generator either using a gearbox or direct drive. The concept is very old, goes back to windmills, where wind power was used for producing mechanical work.

The kinetic energy in an air of mass m moving with speed U is given by the following in units of joules:

$$Ke = \frac{1}{2} \cdot m \cdot U^2.$$

The mechanical power in the moving air is given by the mass flow rate, $\dot{m}$, of the kinetic energy (watts):

$$P_{\text{wind}} = \frac{1}{2} \cdot \dot{m} \cdot U^2.$$

If ρ is the air density (kg/m^3), and A is the area covered by the rotor blades (m^2), the volumetric flow rate is determined by $A \cdot U$, and the mass flow rate of the air is

$$\dot{m} = \rho \cdot A \cdot U$$

(units in kilograms per second). The mechanical power associated with the upstream wind is represented by the following formula in units of watts:

$$P_{\text{wind}} = \frac{1}{2} \cdot \dot{m} \cdot U^2 = \frac{1}{2} \cdot \rho A U \cdot U^2 = \frac{1}{2} \cdot \rho \cdot A \cdot U^3.$$

Here U is referred to as short-term average wind speed (as it will be discussed in the section below).

We often refer to wind power density as in W/m^2,

$$F = \frac{P_{\text{wind}}}{A} = \frac{1}{2} \cdot \rho \cdot U^3,$$

to compare different wind sites independently on the type of rotor area in place and deduce the average power potential in specific sites. Based on the wind power density in different locations, maps of wind potentials have been created in many countries. These maps are relevant and known as wind resources assessment, essential before siting a wind farm.

As the amount of power in the wind varies with the cube of the wind speed, the wind speed is the most relevant element to be considered in the evaluation of the energy production by a wind turbine and a wind farm. The variability of wind speed due to locations and topography are relevant elements to consider even before the wind speed recording and analysis. As such a study of global and local wind patterns is essential elements determining wind energy resources assessment. The wind speed measurements and analysis are further needed to investigate the location before siting on a wind farm. In particular, the variability of wind speed across time, i.e., over long periods such as 10 years or more, 1 year and short periods, are relevant to establish the energy production by a wind farm at a specific location. The wind speed correlation with seasons and daily hours are relevant to determine the production of energy. Moreover, the wind power varies also with the air density, and as such atmospheric considerations associated with temperature and pressure are to be included. In particular, the temperature is a key element to modify the wind energy at a site during diurnal and seasonal periods. As temperature and pressure also depend on altitude, models to estimate wind speed at the rotor heights constitute important tools to properly account for the wind pattern for an estimate of wind power production. These models are based on theoretical assumptions (boundary layer) and phenomenological observations, or both.

In this chapter, we first discuss the type of wind patterns around the globe to broadly identify areas suitable for wind power generation. We provide elements to perform local wind resources assessment based on local winds patterns. We then provide the necessary information on known patterns of wind speed variability over inter-annual, annual, diurnal, and short-term intervals and their assessment in various parts of the world. We provide information about wind speed variability with locations. We will discuss the effect of atmospheric conditions on air density and the wind speed gradient with height, and the relevant models adopted in case studies.

We will provide statistical methods to analyze wind data, statistical and empirical methods to determine wind potentials at a site from data monitoring information.

15.2 Wind Patterns Around the Globe

Wind farm siting requires an understanding of global- and local-scale spatial variations in the wind, which for its nature is an extremely variable element. The temporal and spatial variation of the wind speed is different on different timescales and different locations, and its magnitude varies dramatically depending on local climatology, terrain, and topography. The wind variability is primarily linked to the so-called global circulation (geostrophic winds) and local or small-scale circulations, also related to topology (sea and land breezes, mountain and valley winds, tunnel effects). The uncertainties in the wind resource assessment can occur, and it can be remarkable when siting wind turbines at the edges of the complex non-flat terrains landscape [6].

15.2.1 Global Effects

Irregular heating of the earth combined to its rotation is responsible for global wind patterns [19]. The warm air rises near the equator, and the surface air moves toward the equator to replace the rising air, giving rise to air circulation.

Geostrophic winds originate from ununiformed heating of the atmosphere which gives rise to a pressure gradient and thus a force. As the heating is different at various latitudes on the globe, these winds are associated with specific latitudes. This generated gradient force is then combined with the effect of the apparent Coriolis force, depending on the angular rotation of the earth a rotating reference frame at the specific latitude. Consequently, major wind zones are identified, namely, polar easterlies, westerlies, trade winds, and doldrums.

Trade Winds, divided into Northeast and Southeast (NE and SE), is from 0 to 30 degrees' latitude. Westerlies are from 30 to 60 degrees' latitude north or South and blow from the west at mid-latitudes, strongest in the winter when the pressure over the pole is low. Polar Easterlies or Westerlies are from 60 to 90 degrees' latitude north or south. The Doldrums are low potential winds at the intersection between SE and NE Trade Winds (see Fig. 15.1).

Friction at the surface of the earth within the atmospheric boundary layer is another systematic effect on the wind speed. The boundary layer is defined as the bottom layer of the troposphere that is in contact with the surface of the earth, where viscous forces are not negligible and turbulence is present. Viscous forces on the moving air slow down the wind flow. These forces decrease with height and become negligible above the boundary layer. As most wind turbines are operating within the boundary layer region, this effect must be considered in the model. It is producing a specific wind gradient profile along with the boundary layer, which is modeled by estimating the terrain roughness length where the wind speed tends to be zero.

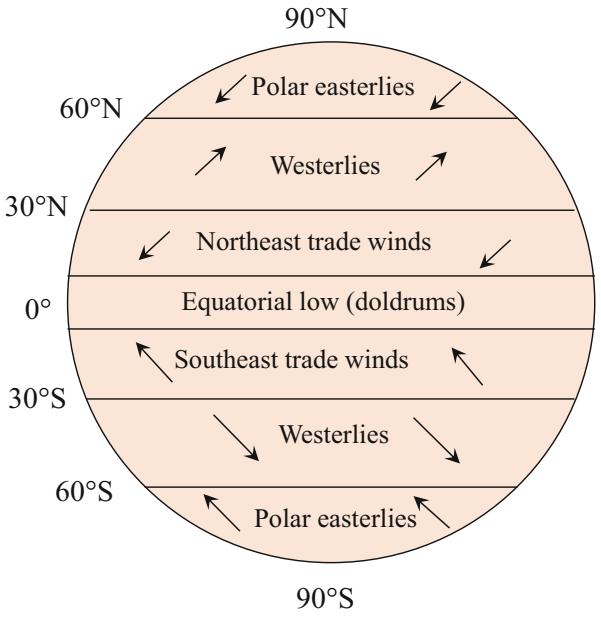


Fig. 15.1 Global wind patterns

15.2.2 Local Effects

Additionally, to these major effects on wind speed, other large-scale effects such as monsoons circulation, hurricanes, and extra-tropical cyclones, together with other small-scale effects such as tornadoes and thunderstorms, can alter the wind speeds.

Smaller-scale atmospheric circulation is distinguished in secondary and tertiary circulation [20]. Secondary circulations comprise monsoon circulation, hurricanes, and extra-tropical cyclones. Tertiary circulations are associated with local winds linked to topography and comprise of land, sea, valley, and mountain winds [7]. These also are combined with thermal effects. Coastal sea breezes are caused by the temperature difference between land and sea. The hot air over land heats up and rises. It is replaced by cooler air from over the sea. While not usually very strong, they are quite regular.

Mountain winds are valley winds that originate on north-facing slopes (south-facing in the northern hemisphere). When the air in the slopes and the neighboring region is heated, its density decreases, and the air ascends towards the top following the surface of the slope. These effects are difficult to resolve in numerical models.

In these complex terrains, the wind changes are pronounced and the flow has additional and variable turbulent components. These terrains have generally typical wind characteristics such as wind acceleration, reduced velocity, and variable wind shear which contribute to turbulence.

15.2.3 Topographic Speedup

There are effects due to topography that can induce wind speed acceleration and consequently turbulence. Acceleration effects in wind flow are often referred to as wind tunneling [5].

It is an undesirable condition to have high turbulence for a wind turbine, as turbulent wind impairs on the wind turbine's loading and its performance. As an example, it can cause an increase of the dynamic load on the wind turbine blades, which consequently increased wear and tear with an eventual reduction of the lifetime of the wind turbine associated with a reduction in wind power generation in time. We will analyze some topography that can induce turbulence in the wind pattern, such as the so-called hill effect, natural tunnel effects, and elevated ridges.

The wind flow pattern results in all directions more accelerated over the top of the hills and mountains. If the hill is very high (1000 m), it can simply block and steer the wind flow around it. Smaller-scale hills can act as aerodynamic objects, such as aircraft wings (See. Here the wind will generally be accelerated from the ground to the top of the hill (with a maximum over the crest). The magnitude of the acceleration depends on the shape of the hill and its orientation compared to the wind. The slope of the hill introduces wind flow separation when the slope is greater than 17° [7], and it can occur at two locations such as lee side flow and hilltop flow [23]. The separation will create a wake of highly turbulent air which is chaotic and detrimental for the fatigue of wind turbines blades.

For artificial obstacles such as buildings and storage towers, they have been commonly studied as a rectangular block while the flow is considered two-dimensional [23]. On local scales, some isolated features such as buildings, windbreaks, and large isolated trees resemble discrete obstacles to the flow. The effect of such obstacles can be sensed at a considerable distance, up to ten times the dimensions of the obstacle [7]. The amount of turbulence will depend on the porosity of the features and the angle of the wind direction to the obstacle [26].

In case of the wind potential of wind from the ocean to the land, the energy potential from the sea with low friction falls as the first feature of a rough surface is encountered by the wind over the island, the potential is then recovering on the water again and dropping again in the proximity of the sea area in at the edge of the cliff, with a potential drop due to turbulence. After the wind potential is showing wind speed acceleration along the cliff up to a maximum at the top of the cliff, followed by again a very low potential well after the cliff, which is considered a no-go zone. After the cliff, the land introduces roughness and again a potential reduction. The hill provides a major boost, enhancing the potential above the one available on the land. However, if the hill is steep enough, it causes flow separation and excessive turbulence after the hill. In this simple case, the wind turbine could be placed on top of the hill but not behind it.

Mathematical models are very reliable in describing the flow over moderate slope hills or obstacles, using linear models, and are at the basis of modern wind industry practices. Most of these models of flow over topography and roughness changes,

currently used as the industry standard, are originated from The European Wind Atlas Analysis and Application Program, Wa^{s}P [27]. Wa^{s}P helps in predicting and modeling of wind flows over terrain at a site. There are several software packages which use computational tools for the assessment of wind resources and to predict energy produced by wind farms as well as to study the layout of the wind farm, as shown in Table 15.1.

The procedure and documentation related to Wa^{s}P is described in Mortensen [21], and it is based on wind data and wind farm calculations measured at or nearby the wind farm site.

The procedure to provide wind data atlas is shown in Fig. 15.2a. Wa^{s}P is specialized to make the wind resource representative over a broader area, given, for instance, the measurements of wind speed from a mast located in an accelerated flow at the crest of a hill. These measurements include accelerated wind that is not representative of wind speed in the surrounding topography; therefore, using the above model, the wind acceleration is calculated and subtracted to the data to provide representative wind speed information for a broader area than just the measurement tower location. Using this method in reverse, thus using a wind profile background from the region and modeling the terrain and topography acceleration in another location, the wind speed at this location can be determined.

However, these models are linear, so for very steep hills, nonlinear models, that are few, would be required; by applying a linear model considerable errors can be introduced. Wa^{s}P can also be used to calculate wind potentials and wind energy production around a wind farm by using meteorological wind directions and intensity from a nearby station as in Fig. 15.2b.

15.3 Estimating Wind Speed with Height, Atmospheric Condition, and Terrain

In wind power engineering, it is a relevant design parameter the accurate determination of wind profile as it directly determines the energy production of a wind turbine for a given tower at a certain height and influences the lifetime of the turbine blades. However, most of the wind speed monitoring occurs at 10 m or below, and rarely monitoring of wind speeds at hub height is done before siting a wind farm. As such, models to estimate the wind potentials at 50 m or 100 m are often necessary to evaluate the production of wind energy or to estimate the wind potentials in nearby sites to the ones already known.

In determining the energy potentials at a specific location and height from the ground, the following elements should be included: a model of the air density, which is affected by temperature, pressure, and height; model of the atmospheric boundary layer to determine the wind shear profile, e.g., the variation of the wind speed with height above the ground; the determination of the boundary layer stability and the evaluation of the terrain roughness.

Table 15.1 Software tools list for modeling and design of wind and solar system using advanced methods

Software name	Application	Other special characteristics	Developer/founder
RETScreen	RETScreen is used to assess the technical, financial viability, and the energy efficiency of a potential renewable energy system. It permits to analyze energy production, cost, emissions	Wind energy resource maps and other renewable energy projects sizing can be produced	Government of Canada
RETScreen plus	An improved version of RETScreen for energy management to enable performance verification on ongoing projects	The software package integrates many databases and includes global climatic conditions	Government of Canada
Wind data generator	The software is used for generating high-resolution wind data	–	Massachusetts Institute of Technology
FOCUS	This software is used to design wind turbines and turbine components such as rotor blades	–	Netherlands Energy Research Centre
FAST	This software is an aero-elastic simulator used for the calculation of onshore wind turbine loads	–	Germanischer Lloyd (National Renewable Energy Laboratory, NREL)
QBlade	Software used for the design of airfoil shapes, wind turbine rotor and blades using BEM theory	Built with the quantity framework with a graphical user interface	Hermann Fottinger Institute of Technical University, Berlin
Vortexje	Used for the simulation of vertical and horizontal axis wind turbines	Aerodynamic optimization, fluid-structure interaction problems, system simulation, and unsteady control	Baayen & Heinz GmbH in Berlin
WAsP	Used to predict and model the wind flow over terrain at a site	Wind flow modeling	Denmarks' Risø National Laboratory

(continued)

Table 15.1 (continued)

Software name	Application	Other special characteristics	Developer/founder
OpenWind, WindPRO, and WindSim	Simulates the wind farm energy output with input wind data, the height of wind turbines, roughness factor, and wind turbine power curve	Windfarm modeling for maximizing the energy output of a wind farm	–
Farm visualization	Graphical representation of a wind farm. This software includes Openwind and WindPRO	The three-dimensional visualization of wind farms and turbines	–
HOMER	Microgrid design and optimization for the advance of hybrid systems, such as a combination of solar, wind, and small hydro systems, with and without diesel generators	Determines the technical and economic analysis for systems ON and OFF the grid	NREL, USA (1993)
Metodyn wind turbine	Computational fluid dynamics software for wind farm optimization over any kind of terrains	Three-dimensional visualization is possible.	Metodyn
System Advisor Model (SAM)	SAM is a techno-economic software model used in the renewable energy industry for most of the renewable energy type		NREL [4]

Adapted from Ref. [23], and references therein

The atmospheric boundary layer is the lowest part of the atmosphere directly influenced by contact with the earth's surface. Within the boundary layer, temperature, relative humidity, and air velocity can change rapidly in both space and time.

Two separate effects are inducing the temporal variation of the vertical wind profiles which need to be considered for wind energy applications, primarily the wind speeds variation with seasonal time scale, which is also depending on the height, together with the instantaneous variation with a time scale of the order of second in wind speeds as a function of height. It is important to consider the

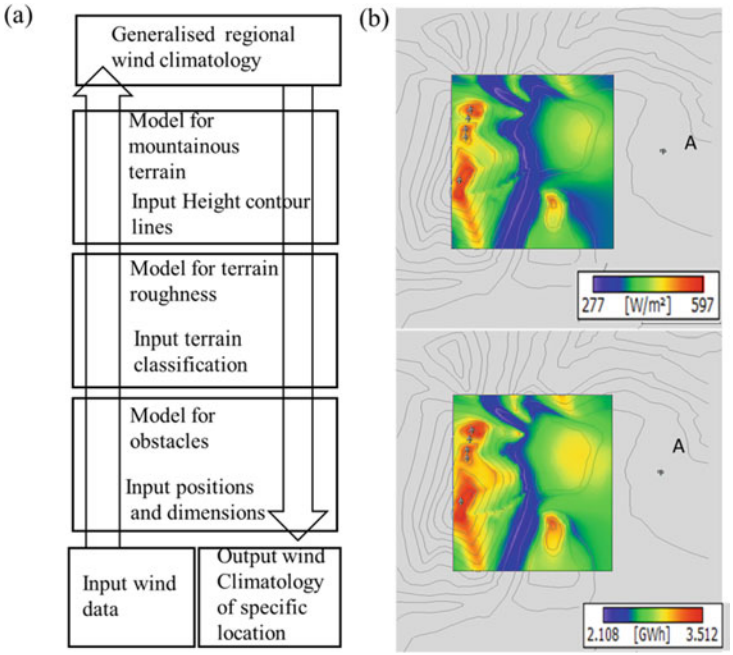


Fig. 15.2 (**a**) Description of the wind atlas methodology of WA^{s}P [27]. Meteorological models are used to calculate the generalized wind climatology from the input measured data. In the reverse process, the wind climate at any specific site may be calculated from the generalized wind climatology. (**b**) Example of Spatial Map of Wind power density and Annual wind energy generation at a wind farm, based on meteorological data at the location of the anemometer, (*A*). The calculation considers wind farm wake effects

monthly or annual wind speed averages trend at a specific location to determine the potentials.

The variation of "instantaneous" wind profiles is modeled using the similarity theory of boundary layers, while the changes in long-term averages as a function of height are related to the regular manifestation of various influencing factors underlying the atmospheric stability conditions. For these second types of variability, more empirical approaches are followed.

The above differences are attributed to the atmospheric stability, while the variation of wind speed with elevation depends on the surface roughness and terrain. These last factors will be presented in the following sections.

15.3.1 Air Density Model

As first the air density, ρ, a model should be considered as a relevant component for the evaluation of wind potential, as it is depending on temperature and pressure,

as such it depends on the height. Denser the air provides more energy that can be received by the turbine. Air density varies with elevation and temperature. The density of dry air can be modeled applying the perfect gas law, $\rho = \frac{p}{R\,T}$, where the density is in units of kg/m^3, the pressure p in Pa and R is the specific gas constant for dry air equal to 287.058 J/(kg·K). Additional corrections for moist air are rarely used. The international standard atmosphere adopts sea-level temperature and pressure equal to $T_o = 288.15$ K and $p_o = 101.325$ kPa, respectively, which gives a standard sea-level density of $\rho_o = 1.225$ kg/m^3 [20]. Air pressure is known to decrease with altitude or elevation above sea level. The pressure in the international standard atmosphere up to an elevation of 5000 m is very closely approximated by

$$p = 101.29 - 0.011837\, z + 4.793\, 10^{-7} z^2$$

where z is the altitude in SI units of meters and the pressure expressed kPa. The real pressure may vary considerably with weather patterns modification if compared to the standard pressure.

The temperature variation with elevation is approximated by

$$T = T_o - \left(\frac{dT}{dz}\right) z$$

where $\left(\frac{dT}{dz}\right)$ is known as known are temperature lapse rate, described in the next section.

In practice, it has been observed that for a given location, the daily and seasonal temperature fluctuations account for a larger influence on the air density than the daily and seasonal variation in pressure. As such a variation of pressure and temperature in a location should be considered in estimating wind energy potentials by using measured values.

If measurements of atmospheric pressure and temperature at given locations are not carried out and are not available, the engineering practice recommends the use of numerical models based on various software commercially available. As an example of a commercial tool, the air density calculator of WA^{s}P [21] can be used to estimate the air density versus height at the site altitude together with the annual average air temperature at the site. We report an example where the measured and WA^{s}P-derived mean air densities are compared in the case studies of ten sites in South Africa and eight sites in NE China [22], with a very good agreement as it is shown in Fig. 15.3.

15.3.2 Atmospheric Boundary Layer

The general trend for wind speed is to approach zero at the surface of the earth and to increase with altitude within the atmospheric boundary layer. This variation

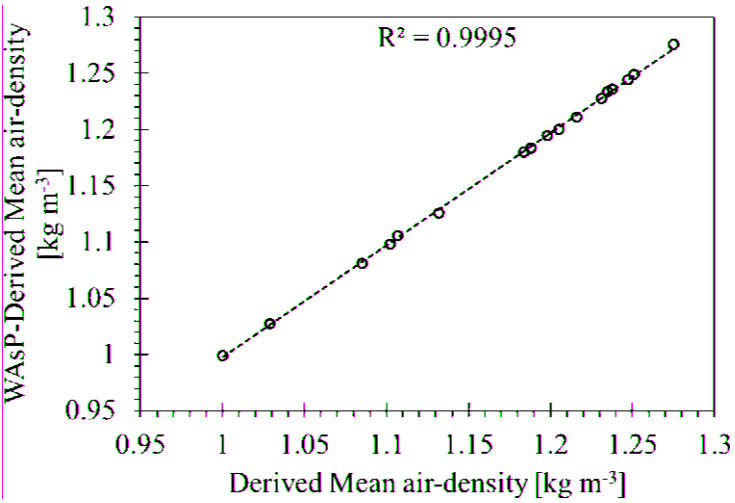


Fig. 15.3 Measured (open circled) and WA^{S}P estimated mean air densities (dashed line) with high statistical linear correlation. (Adapted from Ref. [22])

or gradient of wind speed with the elevation is called vertical wind shear (or vertical profile). Wind experiences increase of speed in the first few hundred meters above the ground. The profile depends on the roughness and orography, on the vertical temperature profile. This second is also referred to as atmospheric stability or boundary layer stability, which is the tendency of the atmosphere to resist the vertical motion of air. Consequently, the boundary layer stability affects the modeling of wind speed gradient with height. The boundary layer stability is related to the rate of change of the temperature with the height $\left(\frac{dT}{dz}\right)$, known as the lapse rate. The stability conditions in the atmosphere are variable when different temperature gradients occur. Three general stability cases can be categorized, neutral, stable, and unstable. In neutral conditions, the temperature profile is adiabatic corresponding to equilibrium between cooling/heating and expansion/contraction with no vertical heat exchange. Based on thermodynamic assumptions, the lapse rate of a system with no heat transfer (dry and adiabatic conditions) is 1 °C/100 m [20], corresponding to a temperature decrease of around 1 °C per 100 m. This is also known as the adiabatic lapse rate. Neutral conditions are characteristic of high wind speeds conditions and the vertical wind profile, in this case, depends only on the terrain roughness and orography.

In unstable conditions, the temperature decreases with altitude faster than in the neutral condition, that is, the temperature lapse rate is greater than the adiabatic lapse rate. This is typically occurring during the summertime because of the ground being heated, the air close to the ground is rising and a lower air density is vertically established. Therefore, a vertical exchange of air mass is occurring leading to a

higher level of turbulence. The vertical wind shear is generally small in these situations due to the heavy mixing.

In stable conditions, the temperature lapse rate is less than the dry adiabatic lapse rate, and the temperature does decrease with altitude. The air at ground level does not rise and if the air is forced to rise, it will sink back towards the ground again. Stable conditions are observed in winter or during the night time. As the air density becomes higher with increasing height, this effect suppresses all vertical exchange of air as well as turbulence. The wind shear in this case, however, can be substantial with large wind direction gradients, since there is little vertical exchange. The standard lapse time based on international convection is in this case is 0.66 °C/100 m [20]. Any atmosphere whose lapse rate is greater than 0.66 °C/100 m is considered a stable atmosphere. One should note that the standard international lapse rate rarely occurs in practice, which motivates the need for the daily balloon soundings, used at major airports worldwide, to determine the actual lapse rate.

15.3.3 Wind Profile Models

In wind energy assessment, to model the vertical profile of wind speed, a neutral atmosphere is assumed in addition to assuming surrounding regions of homogenous flat terrain, which include deserts, prairies, and fields. Commonly two mathematical approaches have been used. The first approach, based on the log law, has its origins in fluid mechanics of the boundary layer flow and atmospheric research, resulting in a combination of theoretical (the first) and empirical (the second) research. The second approach is known as the power law. Both approaches are both approximated and thus responsible for a certain degree of uncertainty due to the unpredictable and complex nature of turbulent flows. Additionally, other numerical methods can be also used. We will analyze all these approaches in the following paragraphs.

According to mixing length theory, eddy viscosity theory, and similarity theory, the wind speed gradient prediction provides a logarithmic wind profile varying with height or elevation [20], given by

$$U(z) = \frac{u^*}{k}\left[\ln\left(\frac{z-d}{z_o}\right) + \psi\left(z, z_o, L\right)\right].$$

Here, $U(z)$ is the instantaneous wind speed and z_o is the surface roughness that is the distance from the surface when the instantaneous wind speed becomes zero, $k = 0.41$ is the von Karman's constant, and u^* is the friction velocity [23]. This is a more general equation where d is the zero-plane displacement, i.e., the height above the ground at which zero wind speed is achieved due to obstacles such as trees or buildings. The function $\psi(z, z_o, L)$ accounts for the stability of the atmospheric boundary, with L is the Obukhov's length from the Monin-Obukhov similarity theory. Under neutral stability conditions $\psi(z, z_o, L)$ is dropped, and the wind profile formula is simplified as follows:

Table 15.2 Terrain roughness reproduced from [20]

Terrain description	z_o (mm)
Very smooth, ice or mud	0.01
Calm open sea	0.2
Blown sea	0.5
Snow surface	3
Lawn grass	8
Rough pasture	10
Fallow field	30
Crops	50
Few trees	100
Many trees, hedges, few buildings	250
Forest and woodlands	500
Suburbs	1500
Centers of cities with tall buildings	3000

$$U(z) = \frac{u^*}{k} \ln\left(\frac{z}{z_o}\right)$$

The short-term average wind speed gradient is normally written relatively to a reference height z_R, where measured data of wind speed are available such as:

$$\frac{U(z)}{U(z_R)} = \frac{\ln(z - z_o)}{\ln(z_R - z_o)},$$

where $U(z_R)$ are the measured wind speed. This is also known as the logarithm law of wind profile variation with height. Using this model, the wind speed is measured at a certain height z_R and then extrapolated at the desired height. This simplified model is commonly used, combined with other air density considerations, to estimate wind energy potentials and production, provided the roughness of the terrain is known. Values of approximated surface roughness length for different terrains have been calculated (Table 15.2).

The wind speed gradient is however reduced in unstable conditions due to heating of the surface and increased vertical mixing of air, while it is amplified during stable conditions, where cooling of the surface occurs and there is the suppression of air vertical mixing. It is important to note that no Coriolis' effect is accounted for in this model.

An alternative empirical model, known as the power law, is also used for estimating the wind speed variability with height, its basic form is:

$$\frac{U(z)}{U(z_R)} = \left(\frac{z}{z_R}\right)^{\alpha},$$

where α is the exponent in the power law. Regardless of the empiricism of the model, often it has shown to fit well the data [3]. Under certain conditions,

Table 15.3 Wind shear parameters [9, 13]

Terrain description	α, wind shear exponent	z_0 (m), winter/summer
Open water	0.10	0.001/0.001
Pasture/hay	0.19	0.01/0.15
Deciduous forest	0.43	0.5/1.3
Mixed forest	0.43	0.9/1.3
Evergreen forest	0.43	1.3/1.3
Lake, ocean, and smooth hard ground	0.1	
Foot high grass on level ground	0.15	
Tall crops, hedges, and shrubs	0.2	
A wooded country with many trees	0.25	
A small town with some trees and shrubs	0.3	
City area with tall buildings	0.4	

$\alpha = 1/7 = 0.143$, indicates a correspondence between wind profiles and flow over flat plates. The constant value of 1/7 for α is commonly assumed in wind resource assessments. However, using a constant exponent, the model does not take into account the roughness of the surface, the presence of obstacles displacing calm winds from the surface due, or the stability conditions of the atmosphere. In locations where trees or structures obstruct the near-surface wind, the use of a constant 1/7 exponent may yield quite erroneous estimates. For these cases, the log law for wind profile is preferred. Under neutral stability conditions, an exponent of 0.10 is more appropriate for open water for example for offshore wind farms, while an exponent of $1/7 = 0.143$ [15] is instead more applicable over open land surfaces. In practice, α varies substantially with elevation, nature of the terrain, wind speed, temperature. Some empirical values can be suitable for certain terrain as reported in Table 15.3, based on [9].

A statistical analysis of the values of wind shear coefficients calculated from the measurements performed in several different regions was carried out. It was shown that wind shear coefficients were assuming values between 0 and 0.14 for 7.3% of the cases, while 91.9% of them were showing values above 0.14; the rest 0.8% of wind shear coefficients were calculated to be negative [13].

Some studies have provided some approaches to deduce α from the parameters in the log law.

In Ref. Counihan [8], the correlation between α and z_0 was found to be

$$\alpha = 0.096 \log_{10} z_0 + 0.16 \left(\log_{10} z_0\right)^2 + 0.24.$$

In Ref. Manwell [20], the correlation of α with wind speed, $U(z_R)$ at reference height, z_R, was also found:

$$\alpha = \frac{0.37 - 0.088 \ \ln \ (U \ (z_R))}{1 - 0.088 \ \ln \ \left(\frac{z_R}{10}\right)}.$$

However, wind energy professionals should ponder on the experimental origin of the power law and choose consequently values of α that are obtained from the best fitting of the available wind data or refer to some tables in the literature.

The variation power law exponent in the wind profile due to variability in surface roughness and atmospheric stability is studied and detailed in Irwin [16]. The theoretical estimates of the power law exponent compare satisfactorily with the power law exponent data from measured sources. Both the log and the power laws may yield the same level of accuracy or inaccuracy depending on the terrain. It was determined that for flat terrain with the absence of trees, forested terrain, and hilly terrain without trees, the variance between the wind speed as predicted by these models at hub height and the available experimental values ranged from 1% to 13% [12]. Thus, one should include in the variability of wind speed, wind shear changes due to variable atmospheric stability, changes in surface conditions and terrain shape, surface roughness. As the atmospheric conditions are rarely neutral and diurnal, using the above laws with only one constant parameter is not possible to predict seasonal or stability-dependent variations of the atmospheric conditions.

Alternatively, methods based on least-squares fitting (LS) have been applied to improve the agreement of the extrapolated data with the observed ones at specific heights [2, 11]. By applying the proposed models log law and power law to a set of measured data of wind speed at different heights (N number of data), taken at least at three measurement heights, better estimates of the α and z_0 can be obtained. In Archer Cristina and Jacobson Mark [2] using LS an equation for the residual R of the squares of the error in wind speed was written as

$$R = \sum_{i=1}^{N} \left(U_i^{\text{meas}} - U\left(z_i\right)\right)^2$$

where U_i^{meas} are the wind speed measured at a certain height z_i, while $U(z_i)$ is the calculated wind speed at z_i based on the log law or the power law, with power coefficient and terrain roughness to be determined as fitting parameters. The partial derivative of R is performed concerning the fitting parameters α and z_0, respectively, and set them to zero, and then the resulting equation solved for the fitting parameter. This gives two LS fitting parameters as:

$$\alpha^{LS} = \frac{\sum_{i=1}^{N} \ln\left(\frac{U_i^{\text{meas}}}{U(z_R)}\right) \ln\left(\frac{z_i}{z_R}\right)}{\sum_{i=1}^{N} \ln\left(\frac{z_i}{z_R}\right)^2},$$

$$\ln\left({z_0}^{LS}\right) = \frac{U\left(z_R\right)\left\{\sum_{i=1}^{N} \ln\left(z_i\right)^2 - \ln\left(z_R\right)\sum_{i=1}^{N} \ln\left(z_i\right)\right\} - \ln\left(z_R\right)\sum_{i=1}^{N} U_i^{\text{meas}} \ln\left(\frac{z_i}{z_R}\right)}{\left\{U\left(z_R\right)\sum_{i=1}^{N} \ln\left(z_i\right) - \sum_{i=1}^{N} U_i^{\text{meas}} \ln\left(\frac{z_i}{z_R}\right) - NU\left(z_R\right)\sum_{i=1}^{N} \ln\left(z_R\right)\right\}}.$$

Here the first reference point is at $z_1 = z_R = 10$ m. The fitting parameters are based on twice-a-day (at 00 and 12) wind profiles from different surface stations (at 10 m) and sounding stations. Eventually, the best fit is used, which is the one that

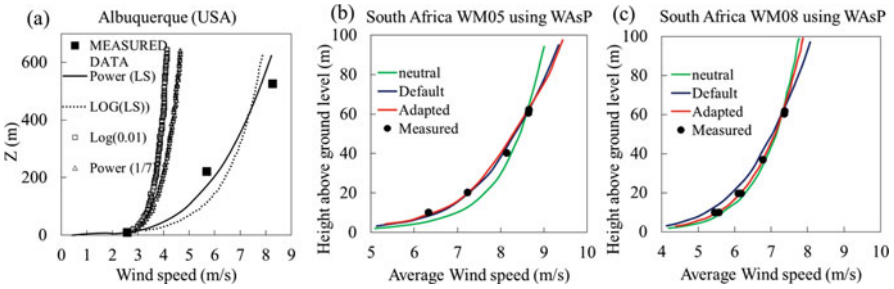


Fig. 15.4 (**a**) Measured data (large filled squares) and interpolated profiles of wind speed with triangles representing the power law with 1/7 exponent coefficient; the squares representing the logarithmic law with 0.01 m roughness length; the solid line representing the power law with LS determined exponent coefficient; the dashed line representing the logarithmic law with LS determined roughness length. Adapted from Archer Cristina and Jacobson Mark [2]. (**b** and **c**) Measured data (solid circles) and modeled vertical wind profiles at two sites in South Africa. The blue profile was modeled using the default setup of WA^{s}P 10.1; the green profile is a strictly neutral atmosphere with a logarithmic wind profile, and the red profile represents the logarithmic law which corrected for the non-neutral atmosphere by slightly changing the heat flux values in WA^{s}P. (Adapted from Ref. [21])

gives the lowest residuals. An iteration between the best fit of surface and sounding stations is used to finally determine the wind speeds at 80 m. This method has been applied to 1327 surface stations and 87 sounding stations in the USA in the year 2000. It was shown that the 80-m wind speeds were, on average, 1.3–1.7 m/s faster than those obtained from wind power maps, logarithmic law, and a power law where constant coefficients were used. In Fig. 15.4a, a comparison for a specific location in Albuquerque in NM (USA) of the three measured points of wind speed with the predictions from the power law with $\alpha = 1/7$, the log law with $z_o = 0.01$ m and LS fitting of the two laws, are shown. It is evident in this case that the models are not predicting the actual wind speed shear at the specific location, and wind speed is increasing faster, which indicates that the neutral atmosphere assumption is quite inaccurate.

WA^{s}P is also used to model the wind speed at different heights. From wind speed measured at two or more heights along the mast, it is possible to adjust the WA^{s}P model to predict the vertical wind profile at the site. The adjustments are related to the terrain descriptions and the atmospheric stability settings. The wind speed profile graph can be adjusted by modifying the terrain descriptions which is changing the roughness length map and the elevation map. Further, the wind profile can be slightly adjusted by changing the heat flux values in the software, which is equivalent to change the atmospheric stability conditions. In Fig. 15.4b, it is shown how the measured wind profile at a specific mast can be modeled well by using the default heat flux values for atmospheric stability. In contrast, the measured wind profile at another mast is better described by neutral atmospheric conditions. To match the neutral condition, the mean heat flux was changed to get a good model description of the profile. In general, the wind speed measurements rarely fit exactly

to the wind speed profile, in part also because the lower level wind speed is disturbed by the flow distortion induced by the mast itself. Selected sectors only where the flow distortion is expected to be small could be used to model the wind speed profile. By doing this, however, the data may not result in a good representation for the average and annual site surroundings.

In Đurišić and Mikulović [11], the LS method is applied to wind speeds at different heights z_j with each data measured with a 10 min interval i, for 1 year period and three locations. The different locations were characterized by different topographies of the terrain and different climatic conditions. By using the Wa^{s}P program, a spatial extrapolation is then carried out. By applying the LS method to the power law, the optimal estimates of the coefficient α_i for each interval are determined from:

$$\alpha_i = \frac{M\sum_{j=1}^{M} \ln\left(U_{ij}\right) \ln\left(z_j\right) - \sum_{j=1}^{M} \ln\left(z_j\right) \cdot \sum_{j=1}^{M} \ln\left(U_{ij}\right)}{M\sum_{j=1}^{M} \ln\left(z_j\right)^2 - \left(\sum_{j=1}^{M} \ln\left(z_j\right)\right)^2}$$

where U_{ij} is the wind speed in i-th 10 min measurement interval ($i = 1, 2, \ldots N$), where N is the total number of observed 10 min measurement intervals, corresponding to the height z_j ($j = 1, 2, \ldots M$). M is the total number of heights on the measurement mast. From the available wind speed measurement data at heights z_j ($j = 1, 2, \ldots M$), one can estimate the wind speed U_i at the desired height z for each 10 min interval i:

$$U_i = U_{M,i}\left(\frac{z}{z_M}\right)^{\alpha_i},$$

where $U_{M,\,i}$ is the wind speed over i-th 10 min interval at the reference height z_M, which corresponds to the highest measurement point. The average hourly value of α calculated over 1 year for each location, shown in Fig. 15.5a, is then determined and used to extrapolate wind speed at different heights (shown in Fig. 15.5b and compared to the measured data). The performed analyses show that by analyzing the measurement data by the LS method results in a better approximation of the wind power potential at heights even greater than the measurement heights.

15.3.4 Characteristics of Terrain

The approaches described above using the log profile and the power law for extracting the wind speed variation with elevation from the ground are based on the assumptions of flat and homogenous terrain. It is, however, important to consider also the terrain topographies on the wind characteristics as presented in various literature dealing with the siting of wind systems [27, 29]. The terrain properties can induce velocity reduction, unusual wind profile, and wind acceleration that are

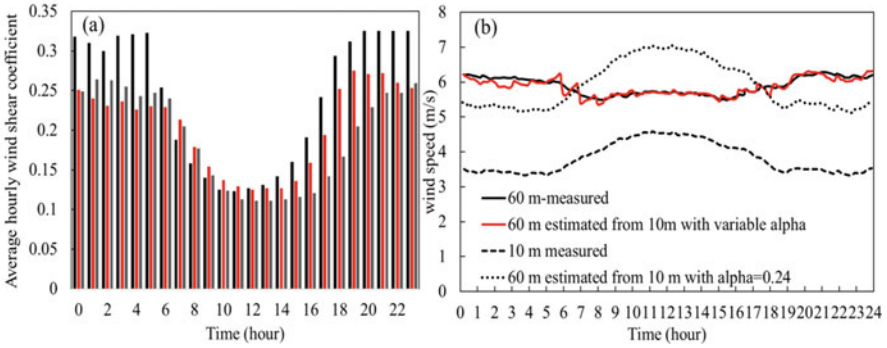


Fig. 15.5 (**a**) Average hourly values of the coefficient α calculated over 1 year measuring period at three locations. (**b**) The hourly variation in wind speed measured at 60 and 10 m and estimated from 10 m to 60 m for one site one of the sites in (**a**). (Data adapted from Ref. [13])

not accounted for by the above laws. The energy output from a turbine is influenced by the terrain topographies very significantly in certain locations as in the majority of natural terrains, where the surface of the earth appears not uniform and changes considerably from location to location. This affects the local wind profile from the examined models. As example, the wind profile changes significantly for the wind moving from a rough to a smooth surface. The terrain smoothness can be partially included in wind shear modeling based on an empirical approach related to the log law. However, other types of terrain, such as non-flat terrain, can induce major errors in estimating wind speed if the above models alone are used.

Terrains are classified as flat and non-flat terrain. The flat terrain is a terrain with small irregularities such as forest and shelterbelts. Non-flat terrain has large-scale elevations or depressions such as hills, valleys, ridges, and canyons, as such, they are landscapes where terrain effects on the wind flow over the land area are significant and should be considered.

The flat terrain is assumed when small elevation differences between the wind turbine site and the surroundings are present; the surrounding estimated as a circular area around the turbine site of approximately 5.75 km of radius. The elevation feature difference should be less than about 60 m. The flat terrain is also classified based on not having any hill with a height to width ratio greater than 0.02 if considering 4 km around the turbine site upstream and downstream wind directions [20].

Non-flat terrain is conversely classified as having hills, depressions, and mountainous features, within the 1 km assumed planetary boundary layer. Non-flat terrains are further divided into small- and large-scale flows landscapes. As an example, a small hill consisting of 10% of the planetary boundary layer is classified as a small-scale terrain feature, whereas mountains are associated with large-scale features.

For small-scale flows landscape, depressions and elevations are additional features to be considered.

Flow over elevated terrain features bears a resemblance to flow around obstacles. Ridges are considered the equivalent of elongated hills of more or less of 600 m elevation above the surrounding terrain, showing almost a top flat area. A ridge should have a ratio of its length to its height of at least 10. A wind turbine should be placed in ridges considering that the optimal prevailing wind direction is perpendicular to the ridge axis. For wind velocity, the slope of a ridge has to be considered as steeper slopes are responsible for higher wind speed, while steeper slopes produce also higher turbulence on the shade of ridges. On a flat-topped ridge, there is a region of high wind shear due to flow separation [29].

Depressions are characterized by a terrain feature lower than the features of the surrounding comprising of valleys, basins, canyons, and passes. Valleys and canyons are considered shallow or small-scale features when they are characterized by a depression estimated to be less than 50 m. Other depressions such as basins and gaps, on the other hand, are considered large-scale features. Large and small depressions can effectively channel the wind inducing an increased change in wind speed. There are a large number of parameters affecting the wind characteristics in a valley, in addition to their variability from a valley to another valley, as such it is almost impossible to determine specific general flow characterization recommendations.

Petersen Erik et al. [24] is an example of a detailed characterization of terrain in terms of surface roughness, presence of nearby obstacles, orographic features for wind resource estimation, and wind turbine siting. Besides, comprehensive mesoscale/microscale meteorological models have been also applied for the assessment and prediction of short-term wind data.

15.4 Wind Variability

All the above-discussed effects combined provide temporal and spatial wind variability. The temporal variability is characterized based on the methodology of the data analysis and their sampling availability and thus classified as inter-annual, annual, diurnal, and short-terms variability [1, 18, 25]. Spatial variability is classified in terms of topography, wind direction changes, and altitude. Average wind speed on different time scales is determined or measured. Variability of wind speed is also associated with local topographical and ground cover variations, so that the wind speed variability between two sites close to each other can vary significantly of decades of percentage [14].

15.4.1 Inter-Annual Wind Variability

Inter-annual variability has been used for long-term wind potential estimates and generally, at least 5 years of data are needed to estimate the variability of average wind speed at a location. Generally, 1-year data at a location are the most common and are regarded to provide accuracy at the best of 10% with a confidence level of 90% [20]. Climate variability is usually computed over a much longer time window such as decadal and multi-decadal, often showing many-years oscillation concerning the analysis parameters. Unfortunately, no wind data are spanning more than 6 decades, and by far less than that, wind farm operation data are very limited to less than 1 decade. Prediction models for inter-annual wind speed are still sought after due to their quite complex nature. The most noticeable way to assess the long-term variability of wind speeds is to inspect historic long-term records of surface station observations.

In Ref. Krakauer and Cohan [18] one of the few global studies has been reported on both the inter-annual wind and solar resources variability. An average coefficient of variation (CV) was defined as

$$CV = \frac{\sum_{i=1}^{12} \sigma_i}{\sum_{i=1}^{12} M_i}$$

where σ_i is the inter-annual standard deviation of the average wind speed in the month M_i. The CV was found to be 11% in the period 1980–2016 when considering monthly variation between 5 years. The estimated CV varies significantly with the latitudes, with a maximum variation at high latitudes on the land and mid-latitudes on water. Seasonal CV differences are characterized by rainy seasons occurring in the winter in Mediterranean climates and the summer in tropical and monsoon regions, where higher CVs are observed compared to CVs observed during the dry seasons. Dry seasons are also associated with lower wind potentials due to lower air density. Europe tends to have more variability in winter as compared to summer, while South and East Asia tend to be the opposite, and North America is mixed. This approach links variability to global-scale climate but suffers from poor accuracy when specific locations must be considered. In Ref. [17] the close to-ground wind speed variability is investigated at seven stations in the surrounding of Minnesota made possible by the availability of climate records of 22–35 years (see Fig. 15.6a). Most of the seven stations showed a trend of reducing mean annual wind speeds with time; this trend is unexpected and not fully understood, whereas it has been observed in some other cases.

In Ref. Wan [28], the wind power generation data from four wind farms in different parts of the USA are shown with relatively large inter-annual changes. It is quite clear that the climate and regional weather pattern is the driving main cause shaping the variability of wind plant energy outputs. A standard deviation value of 13% of its 8-year average is the largest inter-annual variation among the four wind plants here considered. The monthly energy production at these four plants shows

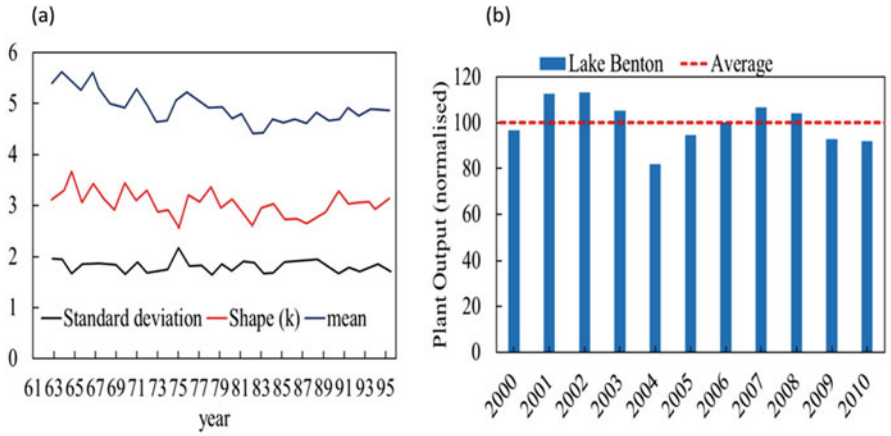


Fig. 15.6 (**a**) Inter-annual average wind speed variability and standard deviation in one station in Minnesota as retrieved from [17]. (**b**) Inter-annual variation of wind energy production in Lake Benton [28]

also clearly noticeable patterns in the 10-year period, where low energy production during the summer months is observed as expected by the reduced wind potential. Figure 15.6b an example of wind energy production variability of one specific power plant from ref. Wan [28] is shown over 10 years.

15.4.2 Annual and Diurnal Wind Variability

The annual wind speed variability is known in most parts of the world where wind turbines are erected. As an example in the eastern, Midwest, upper, and lower plains of the USA, a minimum of average wind speed is found during June and July, with maximum potential typically in the winter and early spring, yielding a variability on the average wind speed of 65% [20]. In California and north-west states, it has been found that the seasonal variability is mixed or the opposite.

Figure 15.7 shows exemplary diurnal wind speed profiles in selected USA regions, shown as of January, July, and Annual. A diurnal variation is very typical within all locations, with the lowest wind speed during hours around middays and high potentials during the night [14]. The largest wind speeds diurnal variation occurs during spring and summer, and the smallest during winter. The diurnal pattern also strongly depends on altitude, being diurnal patterns in mountains and ridges quite different from flat locations.

This typical seasonal wind potential is determined using a few years' data to identify a clear pattern. However, even in seasonal variability, the relative standard deviation on the average wind speed determined can still be as large as 40%. This is also observed in power plant production as shown in [28]

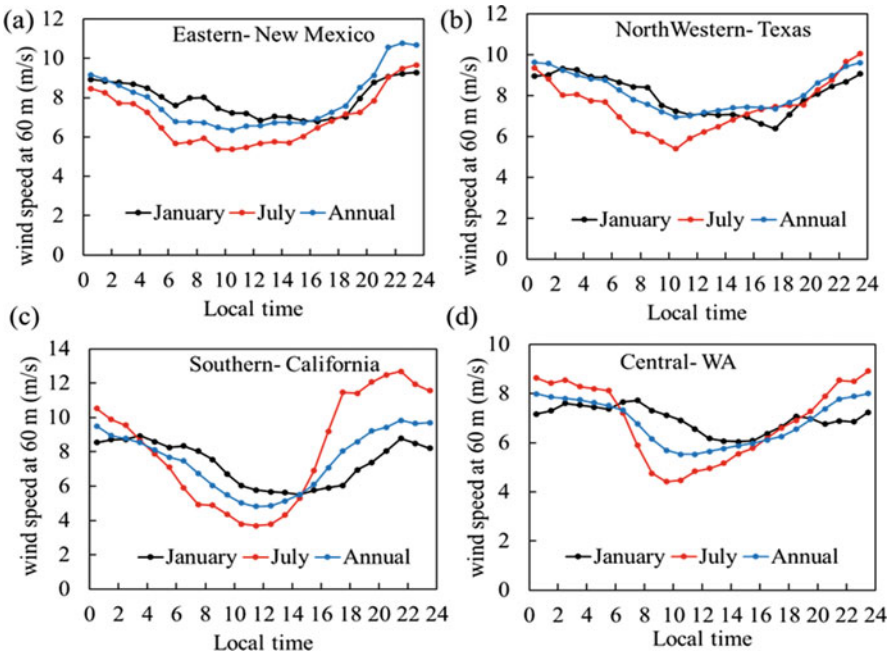


Fig. 15.7 Diurnal wind speed profiles in selected locations in the USA shown in January, July, and Annual average. (**a** and **b**) flatlands and (**c** and **d**) rolling hills. (Data retrieved from NREL)

In Fig. 15.8, the wind speed diurnal and seasonal variability of three locations in the USA is shown, together with an example of the seasonal variability reflected in the energy production of two plants in Lake Benton and Blue Canyon. Lake Benton is in Minnesota and it is a 104 MW plant with data available from 2000–2010, Blue Canyon (BC) is in Oklahoma a 75 MW plant with data from 2003–2010. The highest production from a specific month was more than three times higher than the lowest production in the other months. Even in the same year, the highest production of a specific month could be almost three times more than the lowest production at the minimum month.

15.4.3 Short-Term Wind Variability and Turbulence

The short-term wind variability includes turbulence and gusts. Turbulence is a random wind speed fluctuation imposed on the average wind speed. It is observable in fact in 100 seconds providing a superimposed fluctuating wind speed, $\widetilde{u}$, with mean zero to the short-term wind speed U,

$$u = U + \widetilde{u}$$

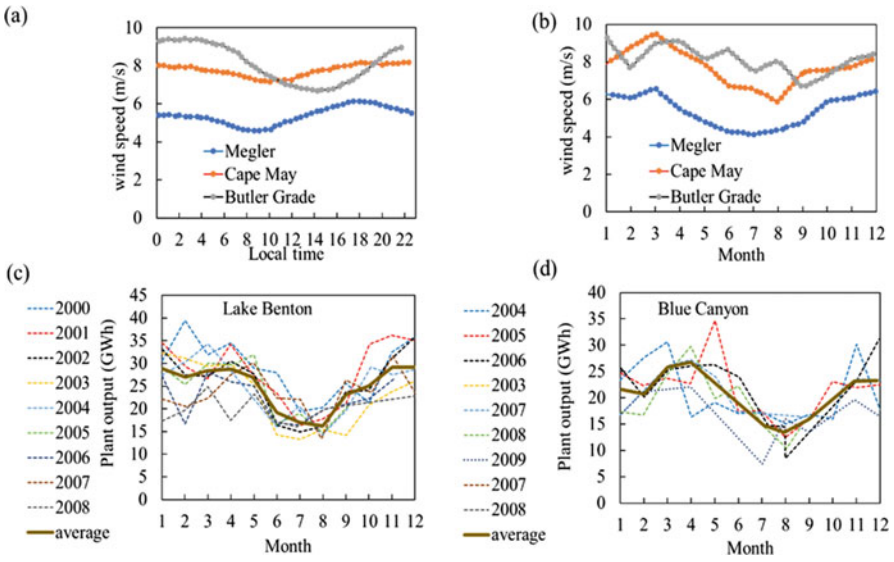


Fig. 15.8 Measured wind speed variability (**a**) diurnal and (**b**) seasonal in three locations in the USA: Cape May at the southern tip of Cape May Peninsula, located by the Atlantic Ocean, New Jersey; Batler Grade, undeveloped terrain where atmospheric pressure gradients in the east and west of the Cascade Range lead to channeling of the flow through the Columbia River Gorge, leading to high wind speeds; Megler a tree-rich and hilly terrain, located on the northern side of the Columbia River on the U.S. West Coast. Data adapted from Ref. [10]. (**c** and **d**) Plants energy outputs by month in several years of two locations in the USA. (Data from Ref. [28])

u is the instantaneous longitudinal wind speed. Here U is an average wind speed over a short period that can be from 10 min up to 1 h. Typically, it is averaged every 1 h time which is much longer of the characteristic time of fluctuations of the turbulence. The short-term average wind speed is therefore

$$U = \frac{1}{N_s} \sum_{i=1}^{N_s} u_i$$

where N_s is the number of data during the chosen short-term period. Typically, on this short term average, the turbulence fluctuations are averaged out.

15.5 Wind Data Analysis

In the analysis of wind data, the variability due to the direction and intensity of wind speed has to be considered. Wind direction can vary on the same time scale of wind speed variability, with seasonal wind speed direction changes of 30°, and average

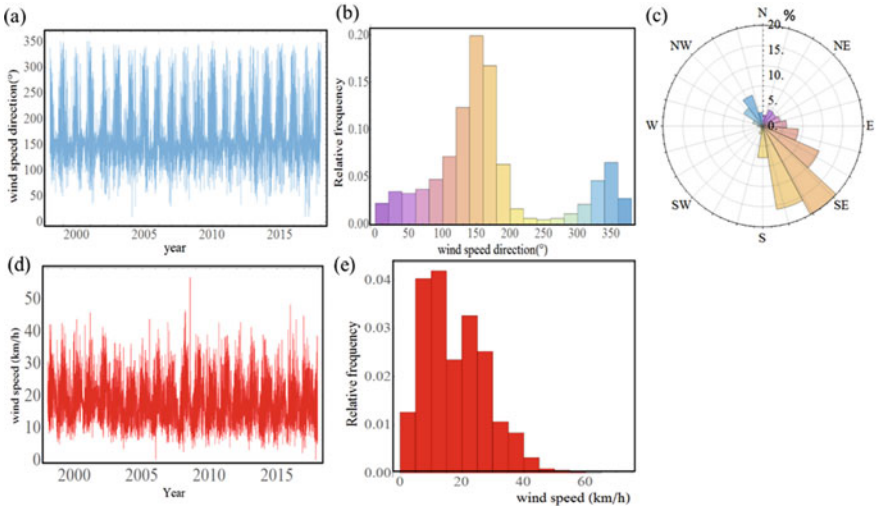


Fig. 15.9 Location wind farm Los Vientos *Latitude*: 26° 20′ 7.8″; *Longitude*: -97° 38′ 51.4″ for the period 1998–2017. (**a**) Time series of wind speed over the period 1998–2017. (**b**) Histogram of annual wind speed with relative frequencies. (**c**) The wind rose, indicating prevailing wind directions SE. (**d**) Time series of wind speed direction (km/h) (0° indicates north). (**e**) Histogram of annual wind direction with relative frequencies. Reference data are at 10 m height and retrieved from Mathematica

monthly wind direction change up to 180°. An example of wind speed direction variability is shown in Fig. 15.9a, c. Horizontal axis wind turbines, which are the majority of installations, must rotate with changes of wind direction (a yawing mechanism). Both turbulence and wind direction changes do impact on the fatigue life of the components such as blades and yaw drives, with a potential reduction of energy produced over time due to the turbine loss of efficiency. Excessive turbulence or change of wind direction at a site should be considered in installing wind turbines, but they do not directly affect wind speed statistics.

In general, a statistical analysis of 1-year data, with average wind speeds measured over $\Delta t = 1$ h for 8760 yours, is used for seasonal wind speed determination. By combining 5 or 10 years (minimum) of wind speed statistics information, wind speed variability over a long time can be determined. This information combined with the extrapolation of wind speed at hub height can eventually provide a more accurate prediction of a wind turbine's annual energy production. The data analysis using the data set is referred to as the bin method.

15.5.1 Statistical Analysis with Direct Use of Data

Given the N data set of short-term average wind speed, U_i, measured during a time Δt, the average long-term wind speed is given by

$$\langle U \rangle = \frac{1}{N} \sum_{i=1}^{N} U_i$$

Δt is typically 1 h, but it depends on the set of data available. In case Δt is more than 1 h the direct data use is the more accurate method. The standard deviation of individual wind speed averages is

$$\sigma_{U_i} = \sqrt{\frac{\sum_{i=1}^{N} (U_i - \langle U \rangle)^2}{N-1}}$$

the average wind power density is

$$\langle F \rangle = \frac{\langle P_{\text{wind}} \rangle}{A} = \frac{1}{2} \cdot \rho \frac{\sum_{i=1}^{N} {U_i}^3}{N}$$

and the average energy in the wind

$$\frac{\langle E_{\text{wind}} \rangle}{A} = \frac{1}{2} \cdot \rho \Delta t \sum_{i=1}^{N} {U_i}^3 = \Delta t \; N \frac{\langle P_{\text{wind}} \rangle}{A}.$$

The average wind power density is also referred to as the wind potential.

If a turbine with power output is given by $P_{\text{Turbine}}(U)$, the average power output of the turbine in a specific wind regime is given by

$$\langle P_{\text{Turbine}} \rangle = \frac{\sum_{i=1}^{N} P_{\text{Turbine}}(U_i)}{N}.$$

The energy produced by this turbine is therefore

$$\langle E_{\text{Turbine}} \rangle = \Delta t \sum_{i=1}^{N} P_{\text{Turbine}}(U_i).$$

As an example, we consider the location with wind farm Los Vientos USA, *Latitude*: 26° 20′ 25.5″; *Longitude*: −97° 41′ 19.6″, with 87 turbines Siemens 2.3 MW and 133 rotor diameter at 100 m hub height (data at 10 m). The terrain is modeled as a flat terrain in neutral atmospheric conditions, so the log law is used with roughness' length of 0.01 m, which corresponds to $\alpha \approx 0.125$. For this site considering 10 years data, we determine by direct data use a $\langle F \rangle = 371$ W/m^2/year and

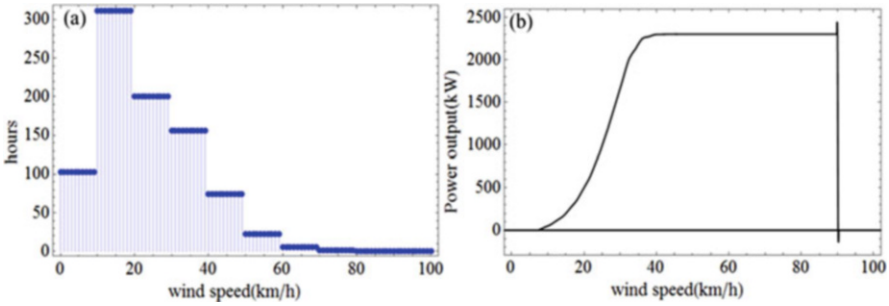


Fig. 15.10 (**a**) Histogram of the frequency in hours of wind speed (in km/h) at the wind farm Los Vientos (Latitude: 26°20′8″ Longitude: −97° 38′ 51″) for the period 1998–2017 of wind speeds converted to hub height. (**b**) The power curve of the installed 87 turbines

average energy in the wind of $\frac{\langle E_{\text{wind}}\rangle}{A} = 3.3\ \text{MWh/m}^2\text{/year}$, an average power output of $\langle P_{\text{Turbine}}\rangle = 990$ kW and annual energy produced by 87 turbines over 10 years of $\langle E_{87,\,\text{Turbine}}\rangle = 755$ GWh/year (in 2016 the $\langle E_{87,\,\text{Turbine}}\rangle = 620$ GWh/year). The histogram of the frequency distribution in hours of the wind data (converted at hub height) and wind turbine power curve are shown in Fig. 15.10.

15.5.2 Statistical Analysis Using Bin Methods

This method is suggested when $N = 8760$ data in 1 year. We will assume as before N data set over 1 year (or more years) of short-term average wind speed, U_i, with the average long-term wind speed given by

$$\langle U\rangle = \frac{1}{N}\sum_{i=1}^{N} U_i$$

Wind speed is also converted to values at hub height for analysis. By separating the wind speed U_i into N_b bins of width w_i, midpoint m_i and f_i the frequency of occurrence in each bin of certain wind speed,

$$N = \sum\nolimits_{i=1}^{N_b} f_i$$

It is necessary to use the wind speed in m/s to have the midpoints as integers. A histogram showing the frequency f_i of each integer mid-point, m_i, can be obtained as shown in Fig. 15.11a. Thus, the average wind speed can be rewritten as

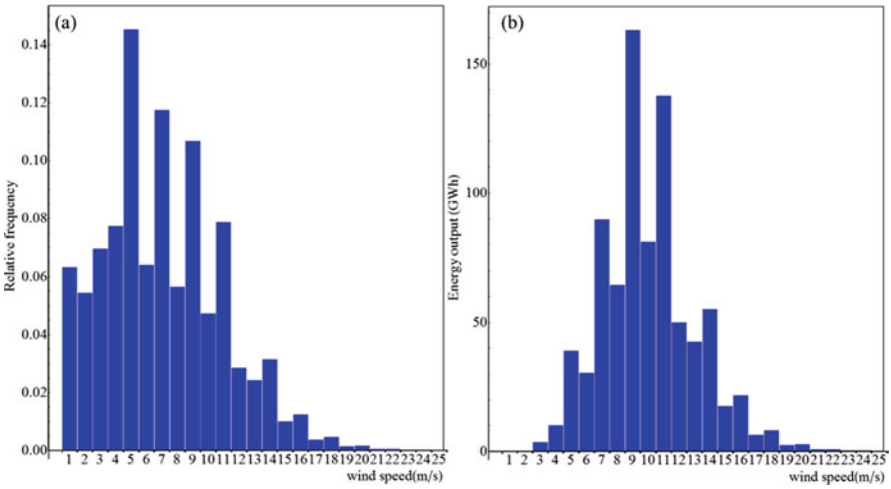


Fig. 15.11 (**a**) Histogram of wind speed and bin with midpoints. (**b**) Energy produced by 87 wind turbines using 10 years of wind data. Latitude: 26°20′8″ Longitude: −97° 38′ 51″, for the period 1998–2017. Data are converted to hub height

$$\langle U \rangle = \frac{1}{N} \sum_{i=1}^{N_b} m_i f_i$$

and the standard deviation is

$$\sigma_{U_i} = \sqrt{\frac{\sum_{i=1}^{N_b} \left({m_i}^2 f_i - N \langle U \rangle \right)^2}{N-1}}.$$

The average wind power density is, therefore:

$$\langle F \rangle = \langle P_{\text{wind}} \rangle / A = \frac{1}{2} \rho \frac{1}{N} \sum\nolimits_{i=1}^{N_b} m_i^3 f_i.$$

The average power output of the turbine at the site is

$$\langle P_{\text{Turbine}} \rangle = \frac{\sum_{i=1}^{N N_b} f_i P_{\text{Turbine}} (m_i)}{N}.$$

Similarly, the wind energy density for the unit area is

$$\frac{\langle E_{\text{wind}} \rangle}{A} = \frac{1}{2} \rho \Delta t \sum\nolimits_{i=1}^{N_b} m_i^3 f_i,$$

and the energy output of the turbine is

$$\langle E_{\text{Turbine}} \rangle = \Delta t \sum_{i=1}^{N_b} f_i \times P_{\text{Turbine}}(m_i).$$

By using data at Latitude 26°20′8″, Longitude: −97° 38′ 51″ for the period 1998–2017, 87 turbines Siemens 2.3 MW and 133 rotor diameter at 100 m hub height, we can calculate:

$$\frac{\langle P_{\text{wind}} \rangle^{\text{bin}}}{A} = 425 \text{ W/m}^2\text{/year}, \quad \frac{\langle E_{\text{wind}} \rangle^{\text{bin}}}{A} = 3.8 \text{ MWh/m}^2\text{/year},$$

$\langle E_{87,\text{ Turbine, 10 years}} \rangle^{\text{bin}}$= 827 GWh/year (in 2016 the $\langle E_{87,\text{ Turbine}} \rangle^{\text{bin}}$= 680 GWh/year). The $\langle P_{\text{Turbine}} \rangle^{\text{bin}}$= 1086 kW. This method is overestimating the energy produced over 10 years as well as the average power compared to the direct data use.

15.5.3 Statistical Analysis Using Wind Speed Probability Density Function

An alternative method relies on using the theoretical wind speed probability density function $\boldsymbol{p}(\boldsymbol{U})$. The probability density function is used to express the probability of wind speed between $\boldsymbol{U_1}$ and $\boldsymbol{U_2}$, as

$$\text{prob } (U_1 \leq U \leq U_2) = \int_{U_1}^{U_2} p(U) dU$$

With

$$\int_0^{\infty} p(U) dU = 1$$

The mean wind speed is

$$\langle U \rangle = \int_0^{\infty} U p(U) dU$$

and the average power density

$$\frac{\langle P_{\text{wind}} \rangle}{A} = \frac{1}{2} \rho \int_0^{\infty} U^3 p(U) dU$$

It is commonly used also the cumulative distribution function, $\boldsymbol{Q(U)}$, as the probability of wind speed smaller or equal to a value $\boldsymbol{U}$

$$Q(U) = \int_0^U p\,(U_b)\,dU_b$$

Common density probability functions for wind are Rayleigh, Weibull, and Gumbel distributions; this last one for extreme wind speeds, respectively given by

$$p^{\text{Rayleigh}}(U) = \frac{\pi}{2}\frac{U}{\langle U\rangle^2}e^{-\frac{\pi}{4}\left(\frac{U}{\langle U\rangle}\right)^2}$$

$$p^{\text{Weibull}}(U) = \frac{k}{c}\left(\frac{U}{c}\right)^{k-1}e^{-\left(\frac{U}{c}\right)^k}$$

$$Q^{\text{Weibull}}(U) = 1 - e^{-\left(\frac{U}{c}\right)^k}$$

$$\langle U\rangle^{\text{Weibull}} = c\,\Gamma\left(1+\frac{1}{k}\right)$$

$$\frac{\langle P_{\text{wind}}\rangle^{\text{Weibull}}}{A} = \frac{1}{2}\rho c^3\Gamma\left(1+\frac{3}{k}\right)$$

$$p^{\text{Gumbel}}(U) = \frac{1}{\beta}e^{-\frac{U-\mu}{\beta}}\exp\left(-e^{-\frac{U-\mu}{\beta}}\right)$$

The average power of the turbine is

$$\langle P_{\text{Turbine}}\rangle = \int_0^\infty P_{\text{Turbine}}(U)p(U)dU$$

and the energy produced is

$$\langle E_{\text{Turbine}}\rangle = \text{duration} * \int_0^\infty P_{\text{Turbine}}(U)p(U)dU.$$

By using data at Latitude 26°20′8″, Longitude: −97° 38′ 51″ for the period 1998–2017, 87 turbines Siemens 2.3 MW and 133 rotor diameter at 100 m hub height, the following average power is calculated for the Weibull, Extreme values and Rayleigh distributions, respectively $\langle P_{\text{Turbine}} \rangle =$ 1078,1018,1056 kW corresponding to total energy produced by $\langle E_{87,\text{ Turbine, 10 years}} \rangle^{\text{probability density}} =$ 821,776,805 GWh/year, respectively. In this case, the extreme values (Gumbel) probability density function provides the closest results to the method with the direct use of data. The fit of the statistical distribution should be done at the height of the measured available data, and fitting parameters adjusted to hub height based on the model for the terrain used. Examples of data fitting with various statistical distributions as described above of the wind data at 10 m are shown in Fig. 15.12.

15.6 Conclusions

A large amount of literature has been discussing wind patterns around the globe as well as its variability and modeling. This knowledge has been translated into classifying wind resources in many parts of the world for harnessing wind energy using wind farms. Wind siting is still a relevant part of the planning for wind farm construction as it is a relevant component for maximizing production with a given wind turbine power curve. Models of terrain and air density with height are very relevant due to the strong dependence of wind speed with a height from ground level. A large amount of literature can be found on modeling the wind gradient at a given location and the annual variability of the parameters used in semi-empirical models are also been investigated in some case studies. The wind speed inter-annual variability is known to a certain extent, but the recording of wind speeds for more than 60 years is scarce. Most of the wind data available and object of case studies span over periods of 1 or 3 decades at the most with at the best of 1-h average. Current modeling of wind statistics and its energy production relies on at the best 1-h wind speed averaged over 10 years. Very short-term variability data (less than 1 h, for example, 5–10 min) are very scarce and prediction of energy production based on a 1-h average tends to provide a good agreement with average monthly energy production. It is to be noted that the energy production data are available even less than wind data, as power plants have been operating sometimes for less than 10 years. The current gap in the literature indicates limited availability of very short-term wind data and related energy production, which can provide a better view of the actual performance if grid integration for a higher contribution of wind in the energy mix is the goal.

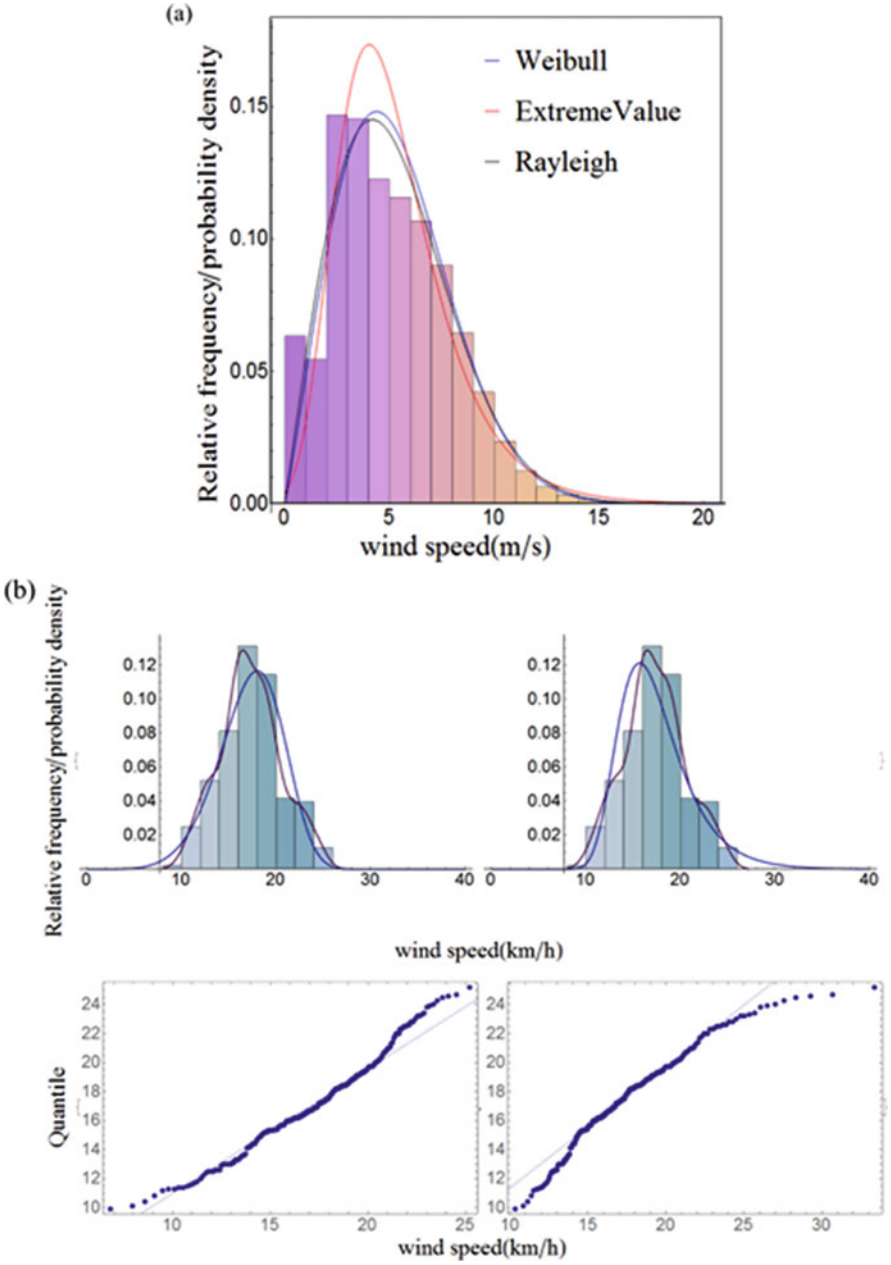


Fig. 15.12 (**a**) Histogram of wind speed (Latitude: 26°20′8″ Longitude: −97° 38′ 51″) compared to three fitted probability density function (Weibull, Extreme values and Rayleigh using fine binning of 1 m/s. Estimated distribution parameters are for Weibull $k = 2.09$, $c = 5.97$ m/s $\langle U \rangle = 5.29$ m/s, for Extreme Values $\mu = 4.02$ m/s, $\beta = 2.12$ m/s, Rayleigh Distribution $\langle U \rangle = 4.18$ m/s. (**b**) Histogram using average monthly data and quantiles compared to Gaussian distribution of the fitting probability density functions

References

1. Archer CL, Jacobson MZ (2013) Geographical and seasonal variability of the global "practical" wind resources. Applied Geography 45:119–130. doi: https://doi.org/10.1016/j.apgeog.2013.07.006
2. Archer Cristina L, Jacobson Mark Z (2003) Spatial and temporal distributions of U.S. winds and wind power at 80 m derived from measurements. Journal of Geophysical Research: Atmospheres 108 (D9). doi:https://doi.org/10.1029/2002JD002076
3. Arya SP (2001) Introduction to Micrometeorology. Second edition edn. elsevier,
4. Blair N, Diorio N, Freeman J, Gilman P, Janzou S, Neises T, Wagner M (2018) System Advisor Model (SAM) General Description (Version 2017.9.5).
5. Bukala J, Damaziak K, Kroszczynski K, Krzeszowiec M, Malachowski J (2015) Investigation of parameters influencing the efficiency of small wind turbines. Journal of Wind Engineering and Industrial Aerodynamics 146:29–38. doi: https://doi.org/10.1016/j.jweia.2015.06.017
6. Cherry NJ (1980) Wind energy resource survey methodology. Journal of Wind Engineering and Industrial Aerodynamics 5 (3):247–280. doi: https://doi.org/10.1016/0167-6105(80)90037-9
7. Coppin PA (2003) Wind resources assessment in Australia: A Planner guide. CSIRO Land and Water
8. Counihan J (1975) Adiabatic atmospheric boundary layers: A review and analysis of data from the period 1880–1972. Atmospheric Environment (1967) 9 (10):871–905. doi: https://doi.org/10.1016/0004-6981(75)90088-8
9. Delgado A, Gertig C, Blesa E, Loza A, Hidalgo C, Ron R (2016) Evaluation of the variability of wind speed at different heights and its impact on the receiver efficiency of central receiver systems. AIP Conference Proceedings 1734 (1):030011. doi:https://doi.org/10.1063/1.4949063
10. Draxl C, Clifton A, Hodge B-M, McCaa J (2015) The Wind Integration National Dataset (WIND) Toolkit. Applied Energy 151:355–366. doi: https://doi.org/10.1016/j.apenergy.2015.03.121
11. Đurišić Ž, Mikulović J (2012) A model for vertical wind speed data extrapolation for improving wind resource assessment using WAsP. Renewable Energy 41:407–411. doi: https://doi.org/10.1016/j.renene.2011.11.016
12. Elkinton MR, Rogers AL, McGowan JG (2006) An Investigation of Wind-Shear Models and Experimental Data Trends for Different Terrains. Wind Engineering 30 (4):341–350. doi:https://doi.org/10.1260/030952406779295417
13. Fırtın E, Güler Ö, Akdağ SA (2011) Investigation of wind shear coefficients and their effect on electrical energy generation. Applied Energy 88 (11):4097–4105. doi: https://doi.org/10.1016/j.apenergy.2011.05.025
14. Hiester TR, Pennell wind turbine (1981) Meteorological aspects of siting large wind turbines. Battelle Pacific Northwest Labs., Richland, WA (USA)
15. Hsu SA, Meindl EA, Gilhousen DB (1994) Determining the Power-Law Wind-Profile Exponent under Near-Neutral Stability Conditions at Sea. Journal of Applied Meteorology 33 (6):757–765. doi:https://doi.org/10.1175/1520-0450(1994)033<0757:DTPLWP>2.0.CO;2
16. Irwin JS (1979) A theoretical variation of the wind profile power-law exponent as a function of surface roughness and stability. Atmospheric Environment (1967) 13 (1):191–194. doi: https://doi.org/10.1016/0004-6981(79)90260-9
17. Klink K (2002) Trends and Interannual Variability of Wind Speed Distributions in Minnesota. Journal of Climate 15 (22):3311–3317. doi:https://doi.org/10.1175/1520-0442(2002)015<3311:TAIVOW>2.0.CO;2
18. Krakauer N, Cohan D (2017) Interannual Variability and Seasonal Predictability of Wind and Solar Resources. Resources 6 (3):29
19. Makarieva AM, Gorshkov VG, Sheil D, Nobre AD, Li BL (2013) Where do winds come from? A new theory on how water vapor condensation influences atmospheric pressure and dynamics. Atmos Chem Phys 13 (2):1039–1056. doi:https://doi.org/10.5194/acp-13-1039-2013

20. Manwell JF., McGowan J.G, Rogers A. L. (2011) Wind Energy Explained: Theory, Design and Application. John Wiley & Sons, Ltd. doi:https://doi.org/10.1002/9781119994367
21. Mortensen NG (2013) Planning and Development of wind farms: Wind Resource Assessment and Siting. DTU Wind Energy,
22. Mortensen NG, Yang Z, Hansen JC, Rathmann O, Kelly MC (2010) Meso- and Micro-scale Modelling in China: Wind atlas analysis for 12 meteorological stations in NE China (Dongbei). Risø National Laboratory for Sustainable Energy, Technical University of Denmark,
23. Murthy KSR, Rahi OP (2017) A comprehensive review of wind resource assessment. Renewable and Sustainable Energy Reviews 72:1320–1342. doi: https://doi.org/10.1016/j.rser.2016.10.038
24. Petersen Erik L, Mortensen Niels G, Landberg L, Højstrup J, Frank Helmut P (1999) Wind power meteorology. Part II: siting and models. Wind Energy 1 (2):55–72. doi:https://doi.org/10.1002/(SICI)1099-1824(199812)1:2<55::AID-WE5>3.0.CO;2-R
25. Simon W (2014) Quantifying the variability of wind energy. Wiley Interdisciplinary Reviews: Energy and Environment 3 (4):330–342. doi:https://doi.org/10.1002/wene.95
26. Taylor PA, Salmon JR (1993) A Model for the Correction of Surface Wind Data for Sheltering by Upwind Obstacles. Journal of Applied Meteorology 32 (11):1683–1694. doi:https://doi.org/10.1175/1520-0450(1993)032<1683:AMFTCO>2.0.CO;2
27. Troen IaP, E.L. (1989) European Wind Atlas. Riso National LaboratoryDenmark
28. Wan YH (2012) Long-term wind power variability National Renewable Energy Laboratory (NREL)
29. Wegley HL, Ramsdell JV, Orgill MM, Drake RL (1980) Siting handbook for small wind energy conversion systems. United States. doi: https://doi.org/10.2172/5490541

Index

L. Dai, R. N. Jazar (eds.), *Nonlinear Approaches in Engineering Application*,
https://doi.org/10.1007/978-3-030-82719-9