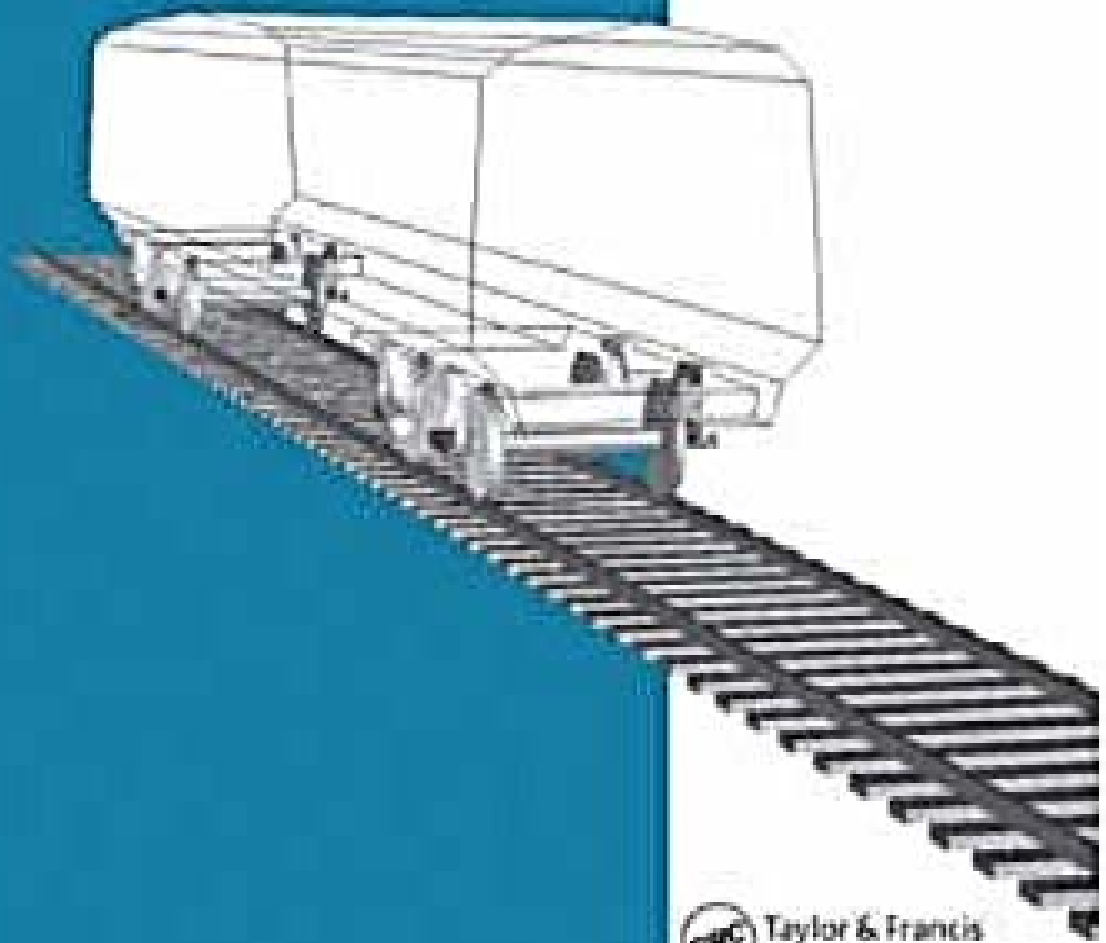


# Handbook of Railway Vehicle Dynamics

Edited by  
**Simon Iwnicki**



Taylor & Francis  
Taylor & Francis Group

# **Handbook of Railway Vehicle Dynamics**

# **Handbook of Railway Vehicle Dynamics**

**Edited by  
Simon Iwnicki**



**Taylor & Francis**

Taylor & Francis Group

Boca Raton London New York

---

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

**Front cover illustration:** A multibody passenger vehicle model developed by Yann Bezin using MSC.Adams and VI-Rail.

Published in 2006 by  
CRC Press  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2006 by Taylor & Francis Group, LLC  
CRC Press is an imprint of Taylor & Francis Group

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

International Standard Book Number-10: 0-8493-3321-0 (Hardcover)  
International Standard Book Number-13: 978-0-8493-3321-7 (Hardcover)  
Library of Congress Card Number 2006002255

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

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

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

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

---

#### Library of Congress Cataloging-in-Publication Data

---

Handbook of railway vehicle dynamics / Simon Iwnicki.  
p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-8493-3321-7

1. Railroads--Cars--Dynamics--Handbooks, manuals, etc. I. Iwnicki, S. (Simon), 1962-

TF550.H37 2006

625.2--dc22

2006002255

---

**informa**  
Taylor & Francis Group  
is the Academic Division of Informa plc.

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



---

# Contributors

**Simon Iwnicki** is a principal lecturer at Manchester Metropolitan University where he teaches applied mechanics and dynamics. Dr. Iwnicki has 15 years of railway vehicle computer modelling experience and has presented papers to international conferences and lectured widely on vehicle–track interaction. He is the editor of the *Journal of Rail and Rapid Transit* and the co-editor (responsible for railway matters) of the international journal *Vehicle System Dynamics*. He is the head of the Rail Technology Unit, a research and consultancy group based in the Department of Engineering and Technology. The unit specialises in the use of computer simulation tools for a variety of railway applications and carries out research and consultancy work on government and industry-funded projects. Dr. Iwnicki is a member of the technical advisory group of the United Kingdom Vehicle/Track System Interface Committee.

**A. H. Wickens** is currently Visiting Industrial Professor at the Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, where his present research interest is in the active guidance and dynamic stability of unconventional railway vehicles. Educated as an aeronautical engineer, he joined British Railways Research in 1962 to carry out research into the dynamics of railway vehicles. He was Director of Research 1971–1983 and Director of Engineering Development and Research 1983–1989. From 1987 to 1990 Professor Wickens was Chairman of the Office for Research and Experiments of the International Union of Railways in Utrecht. He was Professor of Dynamics in the Department of Mechanical Engineering at Loughborough University 1989–1992 and is an honorary member of the Association for Vehicle System Dynamics.

**Anna Orlova** is the Head of the Dynamics and Structural Mechanics Department at the Scientific Research Centre “Vagony” based at the Petersburg State Transport University in St Petersburg, Russia. The department carries out simulation and research testing in the field of railway vehicle dynamics. Dr. Orlova’s special interests include optimisation of running gear parameters for dynamic performance, evaluation of design schemes, and the development of simulation models and testing methods. Dr. Orlova is a supervisor of diploma students at Petersburg State Transport University and the author of several textbooks on bogie design and multibody dynamics simulation.

**Yuri Boronenko** is Professor and Head of the Department of Railway Cars at the Petersburg State Transport University in St Petersburg, Russia. Professor Boronenko is also the Director of the Scientific Research Centre “Vagony” — the independent institution at the Federal Railway Agency of Russia. The centre is involved in many practical fields such as monitoring the fleet of freight wagons in Russia, evaluation of the technical condition of railway vehicles, design of new and modification of existing railcars, and implementation of repair technologies as well as in research and consultancy projects for RZD and industry. Vagony is also a certified testing centre on the Russian Railway Register. Professor Boronenko’s special interests include motion of liquids in railway wagons and vehicle dynamics. For his theoretical and practical contribution in developing railway vehicles Professor Boronenko was made a member of the Transport Academy of Russia.

**Jean-Bernard Ayasse** is Research Director at INRETS (FRANCE). Before joining INRETS he worked at the CEA, Commissariat à l’Energie Atomique, and obtained his PhD from the University of Grenoble in 1970, and a State Thesis in 1977 in solid state physics. He is a specialist

in numerical simulations in the electromagnetic and mechanical domains. His research field goes from the modelling of linear induction motors to railways dynamics. He is the author of several innovations in the modelling of the wheel rail contact and the multibody formalism, implemented in the VOCO code.

**Hugues Chollet** is Researcher at INRETS (FRANCE). He graduated from UTC in 1984 and obtained a PhD in 1991 at Université Pierre et Marie Curie, PARIS 6, on the experimental validation of Kalker's theory for the use in wheel rail contact. He carries out research and consultancy work on guided transportation systems, dealing with wheel rail contact fatigue, derailment situations, instabilities, vibration and comfort problems. He is a referee and board member of several journals and elected member at the IAVSD board.

**Ulf Olofsson** is a senior lecturer and reader in tribology at the Royal Institute of Technology in Stockholm Sweden where he lectures in tribology and machine elements. Before joining the Royal Institute of Technology he worked at the Swedish National Testing and Research Institute with tribological material and component testing and obtained his Lic Eng from Chalmers University of Technology in 1994 and a Ph.D. from the Royal Institute of Technology in 1996. Dr. Olofsson has ten years research experience of the tribology of the wheel–rail contact and supervises postgraduate students and carries out research and consultancy work on government and industry-funded projects. His main research interests include contact mechanics, wear and friction simulation of the wheel–rail contact, wheel and rail wear mapping, and characterisation of the wear mechanisms in the wheel–rail interface.

**Roger Lewis** is a lecturer in the Department of Mechanical Engineering at the University of Sheffield where he teaches design and tribology. He received his Ph.D. in 2000 from the same department before joining the academic staff in 2002. His research has two main areas of focus. The first is in solving industrial wear problems and he is currently working on a number of government and industry-funded projects involving automotive and railway applications. His other area of work is in developing and applying a novel ultrasonic technique for characterising machine element contacts. He has worked on a number of projects related to the wheel–rail interface including the development of a CAE tool for wheel wear prediction, wheel and rail wear mapping, a study of wheel–rail isolation due to track contamination, and characterisation of the wheel–rail interface.

**Tore Dahlberg** is Professor of Solid Mechanics at Linköping University in Linköping, Sweden. He graduated from the School of Mechanical Engineering at Chalmers University of Technology (Gothenburg, Sweden) in 1970 and worked for two years in Burundi, Africa. He gained his Ph.D. at Chalmers in 1979, was appointed “Docent” (associate professor) in 1981, and in 1997 he transferred to Linköping University. Primarily, Professor Dahlberg's research dealt with dynamics of road vehicles, especially random vibration of road vehicles, and dynamics in general. Since the 1980s, his research mainly concerns the dynamic interaction between a train and the track and in recent years focus has been on track dynamics and track deterioration. Professor Dahlberg has published some 30 papers in refereed international journals and he is the author of two textbooks: one in strength of materials and one in fracture mechanics and fatigue. Also, other papers, reports, educational materials, and some popular articles on railway track dynamics have been written by him.

**David M. Johnson** is CEO of Laser Rail Ltd, a technology company specialising in the analysis of the rail vehicle–structure interface. He has been a major contributor to the development of analytical techniques for gauging from empirical methodologies, and in doing so has enabled the running of modern, larger rolling stock on restrictive infrastructure. Mr. Johnson has 28 years of experience in the railway industry, half of which was spent as an engineer with British Rail Research. He has been coauthor of a number of gauging standards and guidance texts, is the

technical advisor to the Vehicle/Structures Systems Interface Committee, and is a visiting reader in mechanical engineering at Imperial College, London.

**Huimin Wu** is a principal investigator at Transportation Technology Center, Inc. (TTCI) in Pueblo, Colorado, U.S.A. Dr. Wu has more than 14 years of experience in the simulation, analysis, and testing of railway vehicles. Her research carries over into areas including vehicle dynamics, vehicle-track interaction, wheel flange climb derailment criteria, computation methodology of studying wheel-rail contact, NUCARS development, wheel-rail profile design, and rail grinding. She has presented papers at international conferences on vehicle-track interaction and published a number of reports on consultancy work carried out for the railways.

**Nicholas Wilson** (BSME, Cornell University, 1980) is a scientist at the Transportation Technology Center, Inc. (TTCI) in Pueblo, Colorado, U.S.A., specialising in rail vehicle dynamics and wheel-rail interaction. He leads the team of engineers developing TTCI's NUCARS multibody vehicle-track dynamic interaction software. Recently, he has also been working on flange climb derailment research, derailment investigations of transit vehicles, wheel-rail wear, and RCF studies. He has also been working on developing rail vehicle dynamic performance specifications for, and analysing performance of trains to carry high-level radioactive material.

**Colin Cole** is a senior lecturer at Central Queensland University where he teaches fluid mechanics and system dynamics. He is also the Deputy Director and Mechanical Group Leader at the Centre for Railway Engineering (CRE), rail research and consultancy group based in the faculty of engineering. His involvement in the rail industry dates back to 1984, including six years experience working in mechanised track maintenance at Queensland Railways. Since 1994 he has conducted a range of rail research and consultancy projects related to field testing of trains, simulation, and development of intelligent systems. He has authored or coauthored 21 technical papers, developed two patents and contributed to numerous commercial research and consulting reports. Dr. Cole is a member of the Engineers Australia and a member of the Queensland Chapter Committee of the Railway Technical Society of Australia.

**David Thompson** is Professor of Railway Noise and Vibration at the Institute of Sound and Vibration Research (ISVR), University of Southampton. Before joining the ISVR in 1996 he worked at British Rail Research in Derby, United Kingdom and TNO Institute of Applied Physics in Delft, the Netherlands and obtained his Ph.D. from the ISVR in 1990. He has written 65 papers in refereed journals and 120 other papers. He is the main author of the TWINS software for railway rolling noise. His research interests include a wide range of aspects of railway noise and vibration as well as noise control, vibroacoustics, structural vibration, modelling rubber, experimental methods for transfer path analysis, and inverse force determination. He teaches undergraduate and masters level courses and is programme organiser for the M.Sc. in sound and vibration.

**Chris Jones** is Reader in Railway Noise and Vibration at the Institute of Sound and Vibration Research (ISVR) at Southampton University. He joined the ISVR from British Rail Research in 1997 where his main interests were in the development and application of numerical models for noise and ground vibration from trains. He continues to work for the railway industry through research and consultancy. He has worked for many railway and tramway projects in the United Kingdom and internationally as well as for track and rolling stock suppliers to develop noise control measures. Dr. Jones teaches numerical methods in acoustics on the bachelors and masters degree programmes at the ISVR and has authored over 80 publications about 30 of which are academic journal papers.

**R. M. Goodall** is Professor of Control Systems Engineering in the Electronic and Electrical Engineering Department at Loughborough University in the United Kingdom. He holds B.A. and

M.A. degrees in engineering from the University of Cambridge, and a Ph.D. degree from Loughborough University. He worked in one of the United Kingdom's GEC companies immediately after graduation before moving to British Rail's Research Division in 1970. He transferred to Loughborough University in 1982 and became a full professor in 1994. His research is concerned with a variety of practical applications of advanced control for high performance electromechanical systems, with a particular specialism related to the railway and aircraft industries. Professor Goodall is a fellow of the IEE and the IMechE in the United Kingdom.

**T. X. Mei** from the University of Leeds is a senior lecturer in control engineering where he leads a research group at the School of Electronic and Electrical Engineering, carrying out leading-edge research in the area of control and systems study for railway vehicles. Dr. Mei has a strong background in railway engineering and substantial expertise in vehicle dynamics and traction control. He has given invited research seminars at an international level and published many papers in leading academic journals and international conferences which explore the application of advanced control techniques and the use of active components. Dr. Mei is one of the most active researchers worldwide in the latest fundamental research into active steering and system integration for railway vehicles, and has made significant contributions to several leading-edge research projects in the field. His educational background includes B.Sc. (1982, Shanghai Tiedao), M.Sc. (1985, Shanghai Tiedao), M.Sc. (1991, Manchester) and Ph.D. (1994, Loughborough).

**Oldrich Polach** is Chief Engineer, Dynamics, at Bombardier Transportation, Winterthur, Switzerland, responsible for sites in Business Unit Bogies Europe. Before joining the railway industry in 1993 he was a research worker and reader at the University of Transport and Communication in Žilina, Czechoslovakia, where he finished his doctoral thesis and obtained his postdoctoral lecture qualification. He teaches railway vehicle dynamics at the ETH Swiss Federal Institute of Technology Zürich and at the Technical University in Berlin. He is a member of the editorial boards of the international journals *Vehicle System Dynamics*, the *International Journal of Vehicle Systems Modelling and Testing*, and the *European Railway Review*, and a member of the working group "Interaction Vehicle-Track" of the European Committee for Standardisation CEN TC 256. His research interests include vehicle dynamics and contact between wheel and rail. His publication activity was awarded the PE Publishing W. A. Agnew/C. N. Goodall Award and the prestigious George Stephenson Prize in 2002.

**Mats Berg** is Professor and Head of Railway Technology at the Royal Institute of Technology (KTH) in Stockholm. Before joining KTH in 1993 he worked at ABB Traction in Västerås, University of California at Berkeley, and obtained his Ph.D. from Lund Institute of Technology in 1987. His main research field is vehicle-track interaction with emphasis on the aspects of structural dynamics, suspension dynamics, track dynamics, and wheel-rail wear. Professor Berg has authored many papers and reports in this field and advised several Ph.D. students. He teaches courses on rail vehicle dynamics and general railway engineering in degree programmes as well as for practising engineers of the railway sector (both in Sweden and internationally).

**Julian Stow** is a research fellow at the Rail Technology Unit (RTU) at Manchester Metropolitan University (MMU). He is a specialist in railway vehicle dynamics and wheel-rail interface engineering. His main area of expertise is in the field of vehicle simulation, safety and design and he has a wide knowledge of testing and instrumentation techniques required to support this work. He has been involved in investigations of rail rolling contact fatigue and a broad range of wheel-rail system problems. He regularly presents work at industry conferences and has organised two conferences on the theme of "railway freight vehicle design". Prior to joining the RTU, he worked in the automotive industry, gaining experience in testing, type approval, and suspension design for commercial vehicles.

**Evert Andersson** is Professor of Railway Technology at the Royal Institute of Technology (KTH) in Stockholm. He started the Division of Railway Technology at KTH in 1987. He also took the initiative of the Railway Group of KTH, a cross-disciplinary centre for research and education in railway engineering. Earlier he was also Chief Mechanical Systems Engineer at ABB Traction (now Bombardier Transportation) in Västerås, Sweden, where he was one of the lead engineers in the development of the high-speed train X2000. He is also an elected member of the Royal Swedish Academy of Engineering Sciences. His main research fields are rail vehicles, vehicle–track interaction and railway energy consumption. Professor Andersson is the author of a vast number of scientific papers and conference contributions in these fields. He also teaches courses on the above topics, as well as on railway engineering in general, and has supervised about ten Ph.D. students for a Ph.D. or licentiate degree.

**Weihua Zhang** is a professor at Southwest Jiaotong University in China. He is the leader of the Traction Power State-Key Laboratory. He holds an M.A. degree in engineering dynamics and a Ph.D. in vehicle engineering from Southwest Jiaotong University (awarded in 1989 and 1996). He teaches vehicle dynamic simulation for graduate students and virtual prototyping for undergraduate students. His research is concerned with railway vehicle dynamic simulation using computers and test rigs including a full-scale roller rig. In recent years he has worked on a high-speed train and tilting train for MOR in China.

**Huanyun Dai** is Professor of Vehicle Engineering at the Traction Power State-Key Laboratory, Southwest Jiaotong University, China. He gained M.A. and Ph.D. degrees in vehicle engineering from Southwest Jiaotong University in 1990 and 1999. He teaches control engineering to post-graduate students. His research is concerned with active suspensions, structural vibration, and roller rig testing for a variety of railway vehicle applications.

**Zhiyun Shen** is a professor at the Traction Power State-Key Laboratory of Southwest Jiaotong University. Professor Shen is an academician of the Academy of Science and the Academy of Engineering of China. He holds a B.A. degree in vehicle engineering from Tangshan Railway University of China, and a Ph.D. degree from Petersburg University of Means of Communication in the former Soviet Union. He has a long experience of teaching and researching and his main research interest has been in dynamics and wheel–rail interactions of vehicle systems.

**Jing Zeng** is a professor at the Traction Power State-Key Laboratory, Southwest Jiaotong University, China. He received MA and PhD degrees in vehicle engineering from Southwest Jiaotong University in 1987 and 1991. He teaches vehicle engineering and vehicle nonlinear system dynamics for undergraduate students and graduate students. He is the vice-director of the laboratory and the leader of research and consultancy group on vehicle dynamics and strength. His research is concerned with the use of computer simulation and roller rig testing for a variety of railway vehicle applications. In recent years he has carried out research and consultancy work on many government and industry-funded projects.

**P. D. Allen** is a research fellow working at Manchester Metropolitan University (MMU), within the Rail Technology Unit. He has a first degree in mechanical engineering and a Ph.D. gained through studies of wheel–rail interaction, particularly the details of errors inherent in scale testing of vehicles on roller rigs. He teaches on postgraduate courses in railway engineering and also carries out varied research and consultancy work under government and industrially funded projects. Dr. Allen is currently involved in working with light rail operators in the United Kingdom, specialising in optimising the wheel–rail interface, together with research work including the integrated modelling of wheel squeal and vehicle dynamics, and neural network prediction of derailment.

---

# Table of Contents

<b>Chapter 1</b> Introduction .....	1
<i>Simon Iwnicki</i>	
<b>Chapter 2</b> A History of Railway Vehicle Dynamics .....	5
<i>A. H. Wickens</i>	
<b>Chapter 3</b> The Anatomy of Railway Vehicle Running Gear .....	39
<i>Anna Orlova and Yuri Boronenko</i>	
<b>Chapter 4</b> Wheel–Rail Contact .....	85
<i>Jean-Bernard Ayasse and Hugues Chollet</i>	
<b>Chapter 5</b> Tribology of the Wheel–Rail Contact .....	121
<i>Ulf Olofsson and Roger Lewis</i>	
<b>Chapter 6</b> Track Issues .....	143
<i>Tore Dahlberg</i>	
<b>Chapter 7</b> Gauging Issues .....	181
<i>David M. Johnson</i>	
<b>Chapter 8</b> Railway Vehicle Derailment and Prevention .....	209
<i>Huimin Wu and Nicholas Wilson</i>	
<b>Chapter 9</b> Longitudinal Train Dynamics .....	239
<i>Colin Cole</i>	
<b>Chapter 10</b> Noise and Vibration from Railway Vehicles .....	279
<i>David Thompson and Chris Jones</i>	
<b>Chapter 11</b> Active Suspensions .....	327
<i>R. M. Goodall and T. X. Mei</i>	
<b>Chapter 12</b> Simulation .....	359
<i>Oldrich Polach, Mats Berg, and Simon Iwnicki</i>	

<b>Chapter 13</b> Field Testing and Instrumentation of Railway Vehicles .....	423
<i>Julian Stow and Evert Andersson</i>	
<b>Chapter 14</b> Roller Rigs .....	457
<i>Weihua Zhang, Huanyun Dai, Zhiyun Shen, and Jing Zeng</i>	
<b>Chapter 15</b> Scale Testing .....	507
<i>P. D. Allen</i>	

---

# 1 Introduction

*Simon Iwnicki*

## CONTENTS

I. Aims .....	1
II. Structure of the Handbook.....	2

### I. AIMS

The principal aim of this handbook is to present a detailed introduction to the main issues influencing the dynamic behaviour of railway vehicles and a summary of the history and state of the art of the analytical and computer tools and techniques that are used in this field around the world. The level of technical detail is intended to be sufficient to allow analysis of common practical situations but references are made to other published material for those who need more detail in specific areas. The main readership will be engineers working in the railway industry worldwide and researchers working on issues connected with railway vehicle behaviour, but it should also prove useful to those wishing to gain a basic knowledge of topics outside their specialist technical area.

Although in the very earliest days of the railways (as described in Chapter 2) an individual was responsible for all aspects of the design of a railway, for most of the historical period of railways the vehicles (or rolling stock) have been under the control of mechanical engineers whereas the track has been seen as the domain of civil engineers. The focus of this book being on the vehicles would tend to put it firmly in the mechanical domain, but in fact, in recent years this rather artificial divide has been lessened as engineers have been forced to consider the railway as a system with the wheel–rail interface at its centre. Increasing use of electrical and electronic components to power, control (or in some cases replace) the basic mechanical components has brought electrical, electronic, mechatronic and control engineers into the teams. The development of equations that represent the complex interactions between a vehicle and the track and of computers able to provide fast solutions to these equations has relied upon the expertise of software engineers and even mathematicians.

The topics covered in this handbook are the main areas which impact on the dynamic behaviour of railway vehicles. These include the analysis of the wheel–rail interface, suspension and suspension component design, simulation and testing of electrical and mechanical systems, interaction with the surrounding infrastructure, and noise generation. Some related areas, such as aerodynamics or crashworthiness, are not covered as they tend to use different techniques and tools and have been extensively developed for road or air transport and are reported on elsewhere.

The handbook is international in scope and draws examples from around the world, but several chapters have a more specific focus where a particular local limitation or need has led to the development of new techniques or tools. For instance, the chapter on longitudinal dynamics mainly uses Australian examples as the issues related to longitudinal dynamics cause most problems in heavy haul lines such as those in Australia where very long trains are used to transport bulk freight with extremely high axle loads, sometimes on narrow gauge track. Similarly, the issue of structure



gauging largely uses the U.K. as a case study, because here the historic lines through dense population centres have resulted in a very restricted loading gauge. The desire to run high-speed trains in this situation has led to the use of highly developed techniques to permit full advantage of the loading gauge to be taken.

The issue of standards has been a tricky one due to the vast number of different organisations who set and control railway standards. It has not been possible to provide comprehensive guidance in this area but typical examples of the application of standards have been brought into the handbook where appropriate. For example, AAR Chapter XI standards for derailment in the U.S. and UIC518 for limits on wheel–rail forces in the E.U. are presented. It should be stressed that these are intended only as illustrative examples of how the results of vehicle dynamic analyses can be used, and those with responsibility for safety should check carefully what the relevant current standards are for their work.

## II. STRUCTURE OF THE HANDBOOK

The history of the field is presented by Alan Wickens in Chapter 2, from the earliest thoughts of George Stephenson about the dynamic behaviour of a wheelset through the development of theoretical principles to the application of modern computing techniques. Professor Wickens was one of the pioneers of these methods and, as director of research at British Rail Research, played a key role in the practical application of vehicle dynamics knowledge to high-speed freight and passenger vehicles. In Chapter 3, Anna Orlova and Yuri Boronenko outline and explain the basic structure of the railway vehicle and the different types of running gear that are commonly used. Each of the relevant components is described and the advantages and disadvantages of the different types explained.

The key area of any study of railway vehicle behaviour is the contact between the wheels and the rails. All the forces that support and guide the vehicle pass through this small contact patch, and an understanding of the nature of these forces is vital to any analysis of the general vehicle behaviour. The equations that govern these forces are developed by Hugues Chollet and Jean-Bernard Ayasse in Chapter 4. They include an analysis of the normal contact that governs the size and shape of the contact patch and the stresses in the wheel and rail and also the tangential problem where slippage or creep in the contact patch produces the creep forces which accelerate, brake, and guide the vehicle. The specific area of tribology applied to the wheel–rail contact is explained by Ulf Olofson and Roger Lewis in Chapter 5. The science of tribology is not a new one but has only recently been linked to vehicle dynamics to allow effective prediction of wheel and rail wear, and examples of this from the Stockholm local railway network are presented.

Although the main focus of railway vehicle dynamics is traditionally on the vehicle, the track is a key part of the system and in Chapter 6 Tore Dahlberg clearly explains the way that track dynamics can be understood. The contribution of each of the main components that make up the track to its overall dynamic behaviour is also presented. Chapter 7 covers the unique railway problem of gauging, where the movement of a railway vehicle means that it sweeps through a space that is larger than it would occupy if it moved in a perfectly straight or curved path. Precise knowledge of this space or envelope is essential to avoid vehicles hitting parts of the surrounding infrastructure or each other. David M. Johnson has developed computer techniques that allow the gauging process to be carried out to permit vehicle designers and operators to ensure safety at the same time as maximising vehicle size and speed, and in this chapter he explains these philosophies and techniques.

Of fundamental concern to all railway engineers is the avoidance of derailment and its potentially catastrophic consequences. Huimin Wu and Nicholas Wilson start Chapter 8 with some statistics from the U.S. that show the main causes of derailment. They go on to summarise the limits

that have been set by standards to try to prevent these occurrences, and cover the special case of independently rotating wheels and several possible preventative measures that can be taken.

Longitudinal train dynamics are covered by Colin Cole in Chapter 9. This is an aspect of vehicle dynamics that is sometimes ignored, but it becomes of major importance in heavy haul railways where very long and heavy trains lead to extremely high coupling forces between vehicles. This chapter also covers rolling resistance and braking systems.

Chapter 10 deals with noise and vibration problems, which have become of greater concern in recent years. David Thompson and Chris Jones explain the key issues including rolling noise caused by rail surface roughness, impact noise, and curve squeal. They outline the basic theory required for a study in this area and also show how computer tools can be used to reduce the problem of noise. The effect of vibrations on human comfort is also discussed and the influence of vehicle design considered.

In Chapter 11, R. M. Goodall and T. X. Mei summarise the possible ways in which active suspensions can allow vehicle designers to provide advantages that are not possible with passive suspensions. The basic concepts from tilting bodies to active secondary and primary suspension components are explained in detail and with examples. Recent tests on a prototype actively controlled bogie are presented and limitations of the current actuators and sensors are explored before conclusions are drawn about the technology that will be seen in future vehicles.

Computer tools are now widely used in vehicle dynamics and some specialist software packages allow all aspects of vehicle–track interaction to be simulated. Oldrich Polach, Mats Berg, and Simon Iwnicki have joined forces in Chapter 12 to explain the historical development and state of the art of the methods that can be used to set up models of railway vehicles and to predict their behaviour as they run on typical track or over specific irregularities or defects. Material in previous chapters is drawn upon to inform the models of suspension elements and wheel–rail contact, and the types of analysis that are typically carried out are described. Typical simulation tasks are presented from the viewpoint of a vehicle designer attempting to optimise suspension performance.

Chapter 13 takes these principles into the field and describes the main test procedures that can be carried out during the design or modification of a vehicle, or as part of an acceptance process to demonstrate safe operation. Julian Stow and Evert Andersson outline the range of transducers available to the test engineer and the ways that these can be most effectively used to obtain valid and useful data. The necessary filtering, corrections, and compensations that are normally made are explained, and data acquisition system requirements are covered. The chapter includes examples of the most commonly carried out laboratory and field tests.

An alternative to field testing is to use a roller rig, on which, a vehicle can be run in relative safety with conditions being varied in a controlled manner and instrumentation can be easily installed. Weihua Zhang and his colleagues at Southwest Jiaotong University in China operate what is probably the most important roller rig in the world today and they outline the characteristics of this and other roller rigs and the ways in which they are used. Chapter 14 also reviews the history of roller rigs, giving summaries of the key details of examples of the main types. Chapter 15 extends the theme to scale testing, which has been used effectively for research into wheel–rail contact. In this chapter P. D. Allen describes the possible scaling philosophies that can be used and how these have been applied to scale roller rigs.

In compiling this handbook I have been fortunate in being able to bring together some of the leading experts in each of the areas that make up the field of railway vehicle dynamics. I and my coauthors hope that this handbook, together with its companion volume, *Road and Off-Road Vehicle Dynamics*, will be a valuable introduction for newcomers and a useful reference text for those working in the field.

Simon Iwnicki  
Manchester

---

# 2 A History of Railway Vehicle Dynamics

*A. H. Wickens*

## CONTENTS

I. Introduction .....	5
II. Coning and the Kinematic Oscillation .....	6
III. Concepts of Curving .....	8
IV. Hunting and the Empirical Development of the Bogie .....	9
V. Interaction between Vehicle and Track.....	10
VI. Innovations for Improved Steering.....	10
VII. Carter .....	12
VIII. Wheel–Rail Geometry.....	15
IX. Matsudaira .....	16
X. The ORE Competition .....	17
XI. Creep .....	18
XII. The Complete Solution of the Hunting Problem .....	19
XIII. Modern Research on Curving .....	22
XIV. Dynamic Response to Track Geometry .....	25
XV. Suspension Design Concepts and Optimisation .....	25
XVI. Derailment .....	28
XVII. The Development of Computer Simulation.....	28
XVIII. Active Suspensions.....	30
XIX. The Expanding Domain of Rail Vehicle Dynamics.....	31
References.....	31

## I. INTRODUCTION

The railway train running along a track is one of the most complicated dynamical systems in engineering. Many bodies comprise the system and so it has many degrees of freedom. The bodies that make up the vehicle can be connected in various ways and a moving interface connects the vehicle with the track. This interface involves the complex geometry of the wheel tread and the rail head and nonconservative frictional forces generated by relative motion in the contact area.

The technology of this complex system rests on a long history. In the late 18th and early 19th century, development concentrated on the prime mover and the possibility of traction using adhesion. Strength of materials presented a major problem. Even though speeds were low, dynamic loads applied to the track were of concern and so the earliest vehicles used elements of suspension adopted from horse carriage practice. Above all, the problem of guidance was resolved by the almost universal adoption of the flanged wheel in the early 19th century, the result of empirical development, and dependent on engineering intuition.

Operation of the early vehicles led to verbal descriptions of their dynamic behaviour, such as Stephenson's description of the kinematic oscillation, discussed below. Later in the 19th century the first simple mathematical models of the action of the coned wheelset were introduced by Redtenbacher and Klingel, but they had virtually no impact on engineering practice. Actually, the balancing of the reciprocating masses of the steam locomotive assumed much greater importance.

A catastrophic bridge failure led to the first analytical model in 1849 of the interaction between vehicle and flexible track.

The growing size of the steam locomotive increased the problem of the forces generated in negotiating curves, and in 1883 Mackenzie gave the first essentially correct description of curving. This became the basis of a standard calculation carried out in design offices throughout the era of the steam locomotive.

As train speeds increased, problems of ride quality, particularly in the lateral direction, became more important. The introduction of the electric locomotive at the end of the 19th century involved Carter, a mathematical electrical engineer, in the problem, with the result that a realistic model of the forces acting between wheel and rail was proposed and the first calculations of lateral stability carried out.

Generally, empirical engineering development was able to keep abreast of the requirements of ride quality and safety until the middle of the 20th century. Then, increasing speeds of trains and the greater potential risks arising from instability stimulated a more scientific approach to vehicle dynamics. Realistic calculations, supported by experiment, on which design decisions were based were achieved in the 1960s and as the power of the digital computer increased so did the scope of engineering calculations, leading to today's powerful modelling tools.

This chapter tells the story of this conceptual and analytical development. It concentrates on the most basic problems associated with stability, response to track geometry, and behaviour in curves of the railway vehicle and most attention is given to the formative stage in which an understanding was gained. Progress in the last 20 years is only sketchily discussed, as the salient points are considered later in the relevant chapters. Moreover, many important aspects such as track dynamics, noise generation, and other high frequency (in this context, above about 15 Hz) phenomena are excluded.

## II. CONING AND THE KINEMATIC OSCILLATION

The conventional railway wheelset, which consists of two wheels mounted on a common axle, has a long history<sup>1</sup> and evolved empirically. In the early days of the railways, speeds were low, and the objectives were the reduction of rolling resistance (so that the useful load that could be hauled by horses could be multiplied) and solving problems of strength and wear.

The flanged wheel running on a rail existed as early as the 17th century. The position of the flanges was on the inside, outside, or even on both sides of the wheels, and was still being debated in the 1820s. Wheels were normally fixed to the axle, although freely rotating wheels were sometimes used in order to reduce friction in curves. To start with, the play allowed between wheel flange and rail was minimal.

Coning was introduced partly to reduce the rubbing of the flange on the rail, and partly to ease the motion of the vehicle around curves. It is not known when coning of the wheel tread was first introduced. It would be natural to provide a smooth curve uniting the flange with the wheel tread, and wear of the tread would contribute to this. Moreover, once wheels were made of cast iron, taper was normal foundry practice. In the early 1830s the flangeway clearance was opened up to reduce the lateral forces between wheel and rail so that, typically, in current practice about 7 to 10 mm of lateral displacement is allowed before flange contact.

Coning of the wheel tread was well-established by 1821. George Stephenson in his “Observations on Edge and Tram Railways”<sup>2</sup> stated that:

It must be understood the form of edge railway wheels are conical that is the outer is rather less than the inner diameter about 3/16 of an inch. Then from a small irregularity of the railway the wheels may be thrown a little to the right or a little to the left, when the former happens the right wheel will expose a larger and the left one a smaller diameter to the bearing surface of the rail which will cause the latter to lose ground of the former but at the same time in moving forward it gradually exposes a greater diameter to the rail while the right one on the contrary is gradually exposing a lesser which will cause it to lose ground of the left one but will regain it on its progress as has been described alternately gaining and losing ground of each other which will cause the wheels to proceed in an oscillatory but easy motion on the rails.

This is a very clear description of what is now called the kinematic oscillation, as shown in Figure 2.1.

The rolling behaviour of the wheelset suggests why it adopted its present form. If the flange is on the inside the conicity is positive and as the flange approaches the rail there will be a strong steering action tending to return the wheelset to the centre of the track. If the flange is on the outside the conicity is negative and the wheelset will simply run into the flange and remain in contact as the wheelset moves along the track. Moreover, consider motion in a sharp curve in which the wheelset is in flange contact. If the flange is on the inside, the lateral force applied by the rail to the leading wheelset is applied to the outer wheel and will be combined with an enhanced vertical load thus diminishing the risk of derailment. If the flange is on the outside, the lateral force applied by the rail is applied to the inner wheel, which has a reduced vertical load, and thus the risk of derailment is increased.

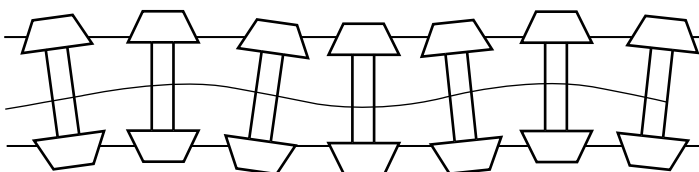
As was explicitly stated by Brunel in 1838 (see Vaughan<sup>3</sup>) it can be seen that for small displacements from the centre of straight or slightly curved track the primary mode of guidance is conicity and it is on sharper curves, switches, and crossings that the flanges become the essential mode of guidance.

Lateral oscillations caused by coning were experienced from the early days of the railways. One solution to the oscillation problem that has been proposed from time to time, even down to modern times, was to fit wheels with cylindrical treads. However, in this case, if the wheels are rigidly mounted on the axle, very slight errors in parallelism would induce large lateral displacements that would be limited by flange contact. Thus, a wheelset with cylindrical treads tends to run in continuous flange contact.

In 1883 Klingel gave the first mathematical analysis of the kinematic oscillation<sup>4</sup> and derived the relationship between the wavelength  $\Lambda$  and the wheelset conicity  $\lambda$ , wheel radius  $r_0$ , and the lateral distance between contact points  $2l$  as

$$\Lambda = 2\pi(r_0l/\lambda)^{1/2} \tag{2.1}$$

Klingel’s formula shows that as the speed is increased, so will the frequency of the kinematic oscillation. Any further aspects of the dynamical behaviour of railway vehicles must be deduced



**FIGURE 2.1** The kinematic oscillation of a wheelset.

from a consideration of the forces acting, and this had to wait for Carter’s much later contribution to the subject.

### III. CONCEPTS OF CURVING

The action of a wheelset with coned wheels in a curve was understood intuitively early in the development of the railways. For example, in 1829 Ross Winans took out a patent that stressed the importance of the axles taking up a radial position on curves,<sup>5</sup> a fundamental objective of running gear designers ever since, and Adams clearly understood the limitations of coning in curves.<sup>6</sup> Redtenbacher<sup>7</sup> provided the first theoretical analysis in 1855 and this is illustrated in Figure 2.2.

From the geometry in this figure it can be seen that there is a simple geometric relationship between the outwards movement of the wheel  $y$ , the radius of the curve  $R$ , the wheel radius  $r_0$ , the distance between the contact points  $2l$  and the the conicity  $\lambda$  of the wheels in order to sustain pure rolling. The application of Redtenbacher’s formula shows that a wheelset will only be able to move outwards to achieve pure rolling if either the radius of curvature or the flangeway clearance is sufficiently large. Otherwise, a realistic consideration of curving requires the analysis of the forces acting between the vehicle and the track. In 1883, Mackenzie<sup>8</sup> supplied the first essentially correct description of curving in a seminal paper (which was subsequently translated and published in both France and Germany). His work was suggested by an unintentional experiment, in which the springs of the driving wheels of a six-wheeled engine were tightened to increase the available adhesion. The leading wheel mounted the rail when the locomotive approached a curve. Mackenzie provided a numerical but nonmathematical treatment of the forces generated in curving. His discussion is based on sliding friction, neglects coning, so that it is appropriate for sharp curves, where guidance is provided by the flanges. Referring to Figure 2.3, Mackenzie explains: If the flange were removed from the outer wheel, the engine would run straight forwards, and this wheel, in making one revolution, would run from A to B; but it is compelled by the flange to move in the direction of the line AC, a tangent to the curve at A, so that it slides sideways through a distance equal to BC. If this wheel were loose on the axle, it would, in making a revolution, run along the rail to F; but the inner wheel, in making a revolution, would run from H to K, the centreline of the axle being KG; so that, if both wheels are keyed on the axle, either the outer wheel must slide forwards or the inner wheel backwards. Assuming that the engine is exerting no tractive force, and that both wheels revolve at the speed due to the inner wheel, then the outer wheel will slide forwards from F to G. Take AL equal to BC, and LM equal to FG, the diagonal AM is the distance which the outer wheel slides in making one revolution.

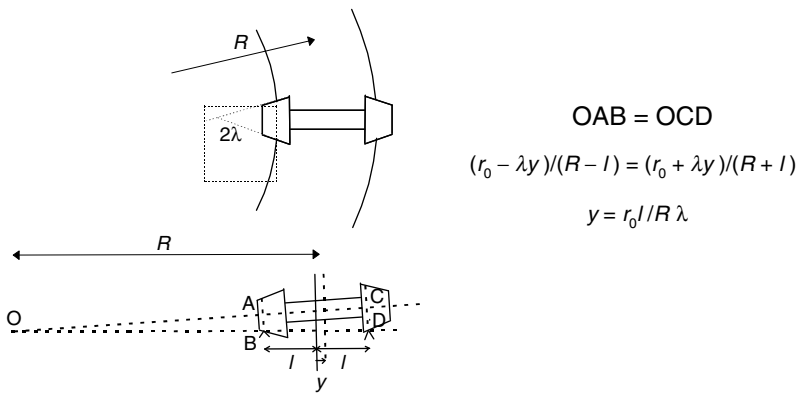
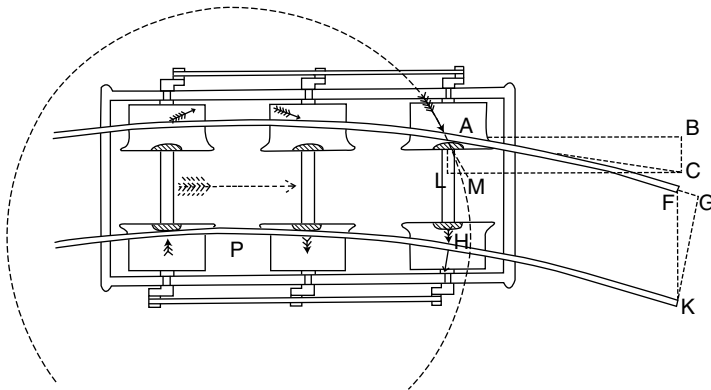


FIGURE 2.2 Redtenbacher’s formula for the rolling of a coned wheelset on a curve.



**FIGURE 2.3** Forces acting on a vehicle in a curve according to Mackenzie.<sup>8</sup>

He then applies similar reasoning to the other wheels, assuming various positions for the wheelsets in relation to the rails. Thus, Mackenzie's calculations showed that the outer wheel flange exerts against the rail a force sufficient to overcome the friction of the wheel treads. Previously, centrifugal forces were regarded as the cause of many derailments. He also made the comment that "the vehicle seems to travel in the direction which causes the smallest amount of sliding," which foresaw a later analytical technique developed by Heumann.

Subsequent work by Boedecker, von Helmholtz, and Uebelacker (described by Gilchrist<sup>9</sup>) was dominated by the need to avoid excessive loads on both vehicle and track caused by steam locomotives with long rigid wheelbases traversing sharp curves. Hence, in these theories, the conicity of the wheelsets is ignored and the wheels are assumed to be in the sliding regime. The corresponding forces are then balanced by a resultant flange force or flange forces. This approach culminated in the work of Heumann in 1913,<sup>10</sup> and Porter in 1934 to 1935.<sup>11</sup>

Superelevation of tracks in curves was introduced on the Liverpool and Manchester Railway, and in the 1830s tables giving the relationship between superelevation of the outer rail and maximum speed were available.

#### IV. HUNTING AND THE EMPIRICAL DEVELOPMENT OF THE BOGIE

The inception of service on the Liverpool and Manchester Railway meant that, for the first time, railway vehicles operated at speeds at which dynamic effects became apparent. The coaches had a very short wheelbase and were reputed to hunt violently at any speed. One measure employed to control this was to close couple the vehicles. The instability of two-axle vehicles was an accepted and often unremarked occurrence throughout their employment on the railways. In the early days of the railways, it had become customary to link together two- and three-axle vehicles not only by couplings but also by side chains to provide yaw restraint between adjacent car bodies in order to stabilise lateral motions.

Two-axle vehicles, and other vehicles with a rigid wheelbase, had obvious limitations in curves. The first known proposal for the bogie was made by William Chapman in 1797, although it was in the United States that the concept was first exploited. As the performance of the rigid wheelbase British locomotives on the lightly built and curvaceous American track was very unsatisfactory, a locomotive with a leading swivelling bogie was introduced in 1832 and this radically improved both stability and curving behaviour. Similarly, the bogie passenger coach became general in North America in the 1840s. These early bogies had very short wheelbases, were free to swivel without restraint and tended to oscillate violently, which was the probable cause of many derailments. In the 1850s the bogie wheelbase was increased, thus improving stability significantly. In Britain,

engineers were cautious about the use of the bogie. Fernihough pointed out the danger of bogie oscillation in his evidence before the Gauge Commission in 1845. He also suggested that it might be controlled by the frictional resistance of a bearing ring of large diameter and this (or, equivalently, friction at the side bearers) became established practice from the 1850s. This pragmatic measure enabled the bogies to follow sharp curves at low speeds while at the same time preventing bogie hunting on straight track.

It was appreciated that the function of the secondary suspension, connecting the bogie frame to the car body, was to isolate the car body from motions of the bogie, for the swing bolster was invented by Davenport in 1841. At a later stage, locomotives often incorporated lateral movement of the bogie pivot restrained by some form of spring, called a centring spring. Experience showed that excessive flexibility of the mounting of the axles gave rise to hunting, and consequently, axles were usually stiffly mounted in bogie frames.

## V. INTERACTION BETWEEN VEHICLE AND TRACK

Although it was thought by some early engineers that the track would be so smooth that no vertical suspension would be necessary, experience soon showed that this was not so. George Stephenson built several locomotives with steam springs as early as 1816, but his first locomotive with steel springs was the Lancashire Witch of 1828. At this stage the suspension of locomotives was provided in order to reduce stresses on the track. Railway carriages followed road practice where laminated steel (leaf) springs had, from about 1770, become normal practice, replacing suspension by leather straps.

A major concern in early locomotive development was the ability to negotiate irregular track, and to maintain contact of all the wheels with the track. A system of equalisation, in which leaf springs connected to levers all attached to the locomotive frame distributed the vertical forces among the wheels, was first used by Timothy Hackworth on his Royal George. However, the most pressing need was, again, in the United States where track was much rougher than in Great Britain, and this resulted in patents covering equalization systems by Eastwick (in 1837) and Harrison (in 1838 and 1842). Such systems became a common feature of the vertical suspension on both locomotives and carriages in the United States.

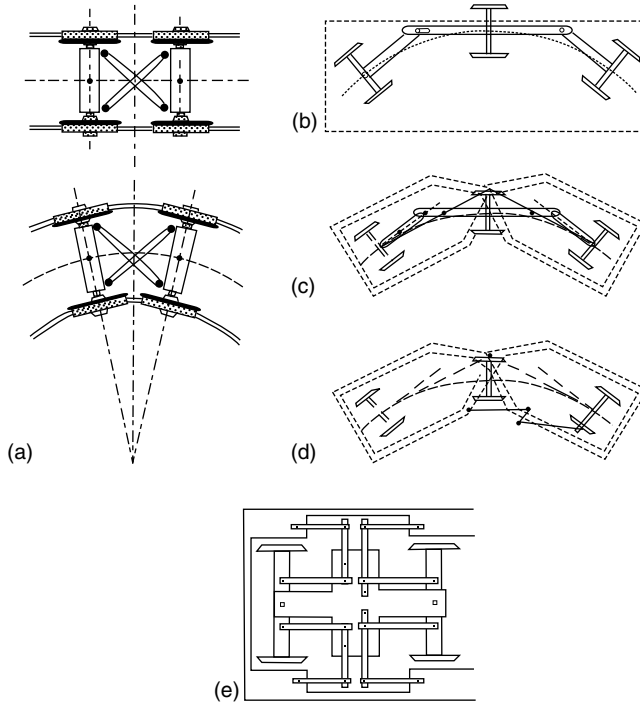
Another source of vertical loading on the track was the impact caused by lack of balance of the revolving and reciprocating parts of the locomotive, which on occasion caused severe track damage. Initially, only the lack of balance of the revolving parts was corrected by balance weights attached to the rims of the driving wheels, and Fernihough appears to be the first to use weights heavy enough to approach complete balance. The theory of balancing was published by Le Chatelier in 1849, who laid down the rules practiced by Clark in his treatise.<sup>12</sup>

Generally, specific problems of interaction between vehicle and track were addressed by empirical and inventive measures. An exception to this was caused by the collapse of Stephenson's bridge across the River Dee at Chester in 1847. At that time, little was known about the dynamic effects of moving loads on bridges. In order to support the inquiry into the accident, a series of experiments was carried out by Willis on a dynamic test rig at Portsmouth dockyard. This was followed by further model tests at Cambridge, and in 1849, G.G. Stokes gave the first analysis of the travelling load problem, albeit with severe simplifying assumptions. This was the beginning of a long history of such investigations.<sup>13</sup>

## VI. INNOVATIONS FOR IMPROVED STEERING

There is a long history of inventions that have attempted to ensure that wheelsets are steered so that they adopt a more or less radial position on curves, and many engineers have tried to improve curving performance by making the vehicle more flexible in plan view. Probably one of the first of





**FIGURE 2.4** Innovations for improved steering: (a) direct connections between wheels by cross-bracing; (b) three-axle vehicle; (c) articulation with a steering beam; (d) articulation with linkage steering driven by angle between adjacent car bodies; (e) bogie with steered wheelsets driven by angle between bogie frame and car body.

these was the arrangement for the Linz-Budweis railway, in which the wheelsets were directly connected by cross-bracing (Figure 2.4a, 1827).

The first articulated locomotive was designed by Horatio Allen in 1832. Although it had a short career, it probably stimulated several of the articulated designs for the Semmering contest in 1851. Thereafter, there was a succession of articulated locomotives, the development of which is described by Weiner.<sup>14</sup> The conflict between the length of the locomotive made necessary by high power and the large curvature of many railway lines was resolved by providing several articulated sections thus reducing the effective wheelbase. The calculation methods described in Section III were applied as these became available.

Various forms of three-axle vehicle have been used widely in the past. In most of these designs the wheelsets were connected to the car body by a conventional suspension similar to that used in two-axle vehicles. Negotiation of curved track was catered for by allowing greater flexibility or clearances for the central wheelset. According to Liechty<sup>15</sup> a three-axle vehicle, in which the lateral displacement of the central axle steered the outer axles through a linkage, was tried out in 1826 on the Linz-Budweis railway. It was argued that three axles, connected by suitable linkages, would assume a radial position on curves and then realign themselves correctly on straight track. Other examples of inventions in which wheelsets are connected so as to achieve radial steering are the three-axle vehicles of Germain (1837), Themor (1844), and Fidler (1868), an example of which is shown in Figure 2.4b. In these schemes the outer wheelsets were pivoted to the car body. More refined arrangements, due to Robinson (1889) and Faye (1898), were much used in trams. More details of these configurations and many others can be found in Ref. 16.

In 1837, as an alternative to the use of the bogie, W.B. Adams proposed an articulated two-axle carriage. Adams invented a form of radial axle in 1863, which had no controlling force, with the result that on straight track there was considerable lateral oscillation of the axle. The idea of a controlling force was suggested by Phipps and this was subsequently applied by Webb (see Ahrons<sup>17</sup>). This was an early example of the fundamental conflict between stability and curving. In the case of locomotives, in 1859, Bissel moved the swivel pin behind the bogie centre in order to allow the wheels to take up a more radial attitude in curves. This idea was adopted quite widely when applied to two-wheel trucks although it was found that a centring spring was required to maintain stability at speed.<sup>18</sup>

Another form of steering exploited the angle between the bogie and the car body in order to steer the wheelsets relative to the bogie frame using a linkage (Figure 2.4e). A similar objective was achieved by mounting the outer wheelset on an arm pivoted on the car body and actuated by a steering beam (Figure 2.4c). An alternative approach was to steer the wheelsets using the angle between adjacent car bodies (Figure 2.4d).

All these developments were based on very simple ideas about the mechanics of vehicles in curves, and depended on systems of rigid linkages and pivots. Not surprisingly, in the light of modern knowledge, there is considerable evidence that when such schemes were built they exhibited an even wider spectrum of various hunting instabilities than the more conventional mainstream designs. This is probably why so few of these inventions achieved widespread adoption.

## VII. CARTER

The configuration of the steam locomotive, originally evolved by Stephenson, was the result of the layout necessary for the boiler, cylinders, and drive to the wheels. This, combined with the guiding bogie at the front of the vehicle, provided a configuration which was unsymmetric fore-and-aft. This configuration was the norm for the steam locomotive intended for main-line operation throughout its history, and which, even if it had riding problems, was usually safe. On the other hand, symmetric configurations were used, but only at low speeds, as at higher speeds they were subject to riding problems, lateral oscillation, and sometimes, derailment. Experience had therefore shown that symmetric configurations were best avoided. This seems to have been forgotten when the first electric locomotives were designed, presumably because they evolved from trams and electric multiple units rather than steam locomotives, and the operational advantages of a symmetric configuration looked attractive. As a result the introduction of the symmetric electric locomotive had been accompanied by many occurrences of lateral instability at high speed, and consequently large lateral forces between vehicle and track. This was how Carter became involved in the problem.

Until then, railway engineering and theory had followed separate paths. The achievements of railway engineers, in the field of running gear at least, largely rested on empirical development and acute mechanical insight. Mackenzie's work in understanding the forces acting on a vehicle in a curve represents an excellent example. It is perhaps not surprising that the seminal development in railway vehicle dynamics was made not by a mechanical engineer but by an electrical engineer who had been exposed to the new analytical techniques necessary to further the application of electrification.

Carter (1870–1952) read mathematics at Cambridge, and after a four-year spell as a lecturer he decided to make electrical engineering his career and spent the following 3 years with General Electric at Schenectady, where he was employed in the testing department working on electric traction. He then returned to England and spent the rest of his career with British Thomson Houston (a company affiliated with General Electric) at Rugby. For most of his career he was consulting engineer to this company, dealing with problems that were beyond the ordinary engineering mathematics of the day. With his mathematical ability and working at the leading edge

of railway electric traction he was able to bridge the gap between science, theory, and railway engineering.<sup>19</sup> After making many significant contributions to electric traction, Carter turned to the mechanical engineering problems of locomotives. The first realistic model of the lateral dynamics of a railway vehicle was that presented by Carter.<sup>20</sup> In this model, Carter introduced the fundamental concept of creep and included the effect of conicity. The paper showed that the combined effects of creep and conicity could lead to a dynamic instability.

Carter stated that the forces acting between wheels and rails can be assumed to be proportional to the creepages, without reference or derivation in this 1916 paper. The concept of creep had first been described by Osborne Reynolds in relation to the transmission of power by belts or straps, and he noted that the concept was equally applicable to rolling wheels.<sup>21</sup>

It was Carter's introduction of the creep mechanism into the theory of lateral dynamics that was the crucial step in identifying the cause of "hunting."

Carter derived equations of motion for the rigid bogie in which two wheelsets were connected by means of a stiff frame. They consist of the two coupled second-order linear differential equations in the variables lateral displacement  $y$  and yaw angle  $\psi$  of the bogie and they are equivalent to

$$\begin{aligned} m\ddot{y} + 4f(\dot{y}/V - \dot{\psi}) &= Y \\ 4f\lambda y/r_0 + I\ddot{\psi} + 4f(l^2 + h^2)\dot{\psi}/V &= G \end{aligned} \quad (2.2)$$

where  $m$  and  $I$  are the mass and yaw moment of inertia of the bogie,  $f$  is the creep coefficient (the creep force per unit creep),  $h$  is the semiwheelbase of the bogie and  $V$  is the forward speed. It can be seen that lateral displacements of the wheelset generate longitudinal creep. The corresponding creep forces are equivalent to a couple that is proportional to the difference in rolling radii or conicity, and which tends to steer the wheelset back into the centre of the track. This is the basic guidance mechanism of the wheelset. In addition, when the wheelset is yawed, a lateral creep force is generated. In effect, this coupling between the lateral displacement and yaw of the wheelset represents a form of feedback, and the achievement of guidance brings with it the possibility of instability. Klingel's solution for pure rolling follows from these equations as a special case in which the wheelset is unrestrained and rolling at low speed.

The theory of dynamic stability had been developed during the 19th century by scientists and mathematicians.<sup>22</sup> The behaviour of governors was analysed by Airy in 1840 in connection with design of a telescope. Maxwell analysed the stability of Saturn's rings in 1856 and derived conditions of stability for governors in 1868. The most significant step forward was Routh's essay for the 1877 Adams Prize,<sup>23</sup> which derived comprehensive conditions for stability of a system in steady motion. Routh incorporated discussion of the stability conditions into the various editions of his textbook.<sup>24</sup> However, in England, mechanical engineers were not familiar with these developments. On the other hand, in 1894 in Switzerland, Stodola studied the stability of steam turbine control systems and encouraged Hurwitz to formulate conditions for stability<sup>25</sup> that are equivalent to Routh's criteria. Carter's work was one of the first engineering applications of Routh's work, and it is interesting to note that Bryan and Williams's<sup>26</sup> pioneering work on the stability of aeroplanes using similar methods had only been published a few years before. Significantly, another early application of stability theory was made by another electrical engineer, Bertram Hopkinson, in his analysis of the hunting of alternating machinery published in 1904.<sup>27</sup> All these early publications refer to Routh's textbook. In contrast, the stability of the bicycle, another system involving rolling wheels, was analysed by Whipple<sup>28</sup> using a solution of the equations of motion.

Moreover, it is interesting that, concurrently with Carter, aeronautical engineers were grappling with the dynamic instabilities of aircraft structures. The first flutter analysis was made in 1916 by Baird and Page<sup>29</sup> and the imperatives of aeronautical progress ensured the development of many of the techniques that were to be brought to bear on the "hunting" problem in the future. By 1927 Frazer and Duncan had laid firm foundations for flutter analysis<sup>30</sup> and a foundation for the

application of matrices in engineering dynamics was established in the text,<sup>31</sup> which eventually was to find application in the railway field.

As Carter's interest was in stability he considered that the flangeway clearance was not taken up and he therefore applied Routh's stability theory, not only to electric bogie locomotives, but also to a variety of steam locomotives. In his mathematical models, a bogie consists of two wheelsets rigidly mounted in a frame, and locomotives comprise wheelsets rigidly mounted in one or more frames. Following Carter's first paper of 1916 the theory was elaborated in a chapter of his book.<sup>32</sup>

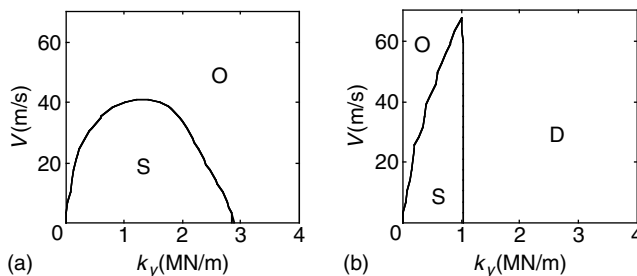
So far, Carter had used an approximation to give the value of the creep coefficient (the constant of proportionality between the creep force and the creepage). In 1926, Carter analysed the creep of a locomotive driving wheel by extending Hertz's theory of elastic contact, as presented by Love.<sup>33</sup> He considered the case of creep in the longitudinal direction, treating the wheel as a two-dimensional cylinder.<sup>34</sup> This not only provided an expression for the creep coefficient but described how the creep force saturated with increasing creepage. Hertz, the German physicist, had become interested in 1881 in the theory of compression of elastic bodies as a result of his work on optics. By making some realistic assumptions he was able to give a theoretical solution for the size of the contact area and the stresses in the two contacting bodies as a function of the normal load between the bodies. This work attracted not only the attention of physicists but of engineers who persuaded Hertz to prepare another version of his paper including experimental results.<sup>35</sup>

Carter's next paper<sup>36</sup> gave a comprehensive analysis of stability within the assumptions mentioned above. As he was concerned with locomotives the emphasis of his analyses was on the lack of fore-and-aft symmetry characteristic of the configurations he was dealing with, and he derived both specific results and design criteria.

His analysis of the 0-6-0 locomotive found that such locomotives were unstable at all speeds if completely symmetric and he comments that this class of locomotive is "much used in working freight trains; but is not employed for high speed running on account of the proclivities indicated in the previous discussion."

Carter analysed the 4-6-0 locomotive both in forward and reverse motion and found that in forward motion beyond the limits shown (i.e., for sufficiently high speed or sufficiently stiff bogie centring spring) the bogie tends to lash the rails; but being comparatively light and connected with the main mass of the locomotive, the impacts are unlikely to be a source of danger at ordinary speeds.

Two of Carter's stability diagrams, the first of their kind in the railway field, are shown in Figure 2.5. As the system considered has four degrees of freedom (lateral translation and yaw of mainframe and bogie), substitution of a trial exponential solution and expansion of the resulting characteristic equation leads to an eighth order polynomial. As Carter writes "expansion of the determinant is long ... but not difficult." Carter examines stability in two ways: first, by extracting the roots of the polynomial and second by Routh's scheme of cross-multiplication [Ref. 24, p. 226].



**FIGURE 2.5** Carter's stability diagram for the 4-6-0 locomotive in: (a) forward motion and (b) reverse motion.  $k_y$  is the centring stiffness. (Recalculated in modern units from Ref. 36.) S = stable; O = oscillatory instability; D = divergence.

Either method involved tedious and lengthy calculations by hand, and tackling more complex cases “becomes, more appropriately, an office undertaking.”

In reverse motion (Figure 2.5b), he found that beyond a certain value of the centring spring stiffness buckling of the wheelbase tending to cause derailment at a fore-wheel and moreover that the impacts of the flanges on the rail when the locomotive is running at speed are backed by the mass of the main frame and are accordingly liable to constitute a source of danger.

This was the explanation for a number of derailments at speed of tank engines such as the Lincoln to Tamworth mail train at Swinderby on June 6, 1928, as discussed in his final paper.<sup>37</sup>

Carter’s analysis of the 2–8–0 with a leading Bissel similarly explained the need for a very strong aligning couple for stability at high speed, while noting that in reverse motion a trailing Bissel has a stabilising effect for a large and useful range of values of aligning couple.

Although Carter had exploited the stabilising influence of elastic elements in his analyses of unsymmetric locomotive configurations, his brief treatment of symmetric vehicles with two-axle bogies (by now a common configuration of passenger rolling stock) assumed that the bogies were pivoted to the car body. That this case received such brief mention is consistent with the fact that railway engineers had, by empirical development, achieved an acceptable standard of ride at the speeds then current. Moreover, as Carter says “the destructive effect of the instability is, however, limited on account of the comparatively small mass of the trucks.”

Carter’s work expressed, in scientific terms, what railway engineers had learnt by hard experience, that stability at speed required rigid-framed locomotives to be unsymmetric and unidirectional. A further practical result of his work was a series of design measures, the subject of various patents,<sup>38</sup> for the stabilisation of symmetric electric bogie locomotives.

The fact that the analyses, although only involving a few degrees of freedom, required heavy algebra and arithmetic, and involved techniques beyond the mechanical engineering training of the day, is perhaps one reason why Carter’s work was not taken up much sooner.

Thus, while the theoretical foundations had been established, the need for vehicle dynamics was not, and practising railway engineers were largely sceptical of theory, particularly when the experimental basis was very limited. As a result, the next 20 years saw only a few significant contributions to the science of railway vehicle dynamics.

Rocard<sup>39</sup> employed the same form of equations of motion as Carter. In addition to covering much of the same ground as Carter, he considered the case of a massless bogie, which is connected by a lateral spring to the car body, and showed that the system could be stabilised. Rocard also considered the case of the unsymmetric bogie in which the wheelsets have different conicities. He found that the distribution of conicity can be arranged to give stability in one direction of motion, but not in both. Rocard states that a successful experiment was made by French National Railways in 1936.

There were also theoretical contributions by Langer and Shamberger<sup>40</sup> and Cain<sup>41</sup> that involved rather severe assumptions, but in general, papers concerned with bogie design published during this period were purely descriptive, reflecting the negligible role played by analysis in this branch of engineering practice. However, in 1939 Davies carried out significant model experiments of instability although stopping short of a complete analysis,<sup>9</sup> and pointed out the importance of worn wheel and rail profiles to wheelset dynamics.<sup>42</sup>

## VIII. WHEEL–RAIL GEOMETRY

Carter assumed that the wheel treads were purely conical. In practice, it had been known from the earliest days of the railways that treads wear rapidly and assume a hollow form. It was also known that there was a connection between ride quality and the amount of wheel wear. An important further step in developing a realistic mathematical model was concerned with the treatment of actual wheel and rail profiles. While new wheel profiles were purely coned on the tread, usually to an angle of 1:20, in 1937 Heumann,<sup>43</sup> emphasised the importance of worn wheel and rail profiles to

wheelset behaviour in curves. Heumann analysed the effect of the mutual wheel and rail geometry on the variation of the rolling radius as the wheelset is displaced laterally, and derived the formula for the effective conicity  $\lambda_0$  of a wheel–rail combination for small displacements from the central running position, defined as the rate of change of the rolling radius with lateral displacement of the wheelset.

$$\lambda_0 = \delta_0 R_w / (R_w - R_r) (1 - r_0 \delta_0 / l) \quad (2.3)$$

where  $R_w$  and  $R_r$  are the wheel and rail radius of curvature and  $\delta_0$  is the slope of the tread at the contact point. Heumann's expression shows clearly that the effective conicity of a worn wheelset can be much greater than that of the corresponding purely coned wheelset. Moreover, Heumann suggested for the first time that profiles approximating to the fully worn should be used rather than the purely coned treads then standard. He argued that after reprofiling to a coned tread, tyre profiles tend to wear rapidly so that the running tread normally in contact with the rail head is worn to a uniform profile. This profile then tends to remain stable during further use, and is largely independent of the original profile and of the tyre steel. Similarly, rail head profiles are developed which also tend to remain stable after the initial period of wear is over. Heumann therefore suggested that vehicles should be designed so as to operate with these naturally worn profiles, as it is only with these profiles that any long-term stability of the wheel–rail geometrical parameters occurs. Moreover, a considerable reduction in the amount of wear would be possible by providing new rails and wheels with an approximation to worn profiles at the outset. Modern wheel and rail profiles are largely based on this concept.

## IX. MATSUDAIRA

Tadashi Matsudaira studied marine engineering at the University of Tokyo and then joined the aircraft development department of the Japanese Imperial Navy where he was concerned with the vibration of aeroplanes. After the end of World War II, he moved to the Railway Technical Research Institute of Japanese National Railways to work on railway vehicle dynamics. During the years 1946 to 1957, Japanese National Railways were attempting to increase the speed of freight trains. The short wheelbase two-axle wagons then in use experienced hunting at low speeds and a high rate of derailment. Matsudaira introduced his experience of the flutter problem in aeroplanes (such as the Japanese Imperial Navy's "Zero" fighter), then using both analysis and scale model experiments on roller rigs, he showed that the hunting problem is one of self-excited vibration and not arising from external factors such as uneven rail geometry. This was the beginning of roller rig testing for vehicle dynamics.<sup>44</sup> In his paper<sup>45</sup> he departed from Carter's model by considering a single wheelset and demonstrated the stabilising effect of elastic restraint. As this paper was in Japanese it had little impact in the West. Subsequently, Matsudaira, for the first time, introduced into the mathematical model of the two-axle vehicle both longitudinal and lateral suspension flexibilities between wheelset and car body, a crucial step in understanding the stability of railway vehicles and based on this was able to suggest an improved suspension design.

In the 1950s planning started for the new Tokaido line or Shinkansen, the first purpose-built dedicated high-speed railway. Shima<sup>46</sup> identifies the bogies as one of the key enabling technologies of the Shinkansen as it made possible the concept of the high-speed multiple unit train in which every bogie is powered. It had been widely assumed that the powered bogies would not run as smoothly as the trailer bogies, but by studying closely the stability of bogies theoretically and experimentally it was possible to improve the riding quality of the powered bogies up to very high speeds. The analysis of these bogies by Matsudaira and his group led to the choice of suspension parameters that were subsequently validated by roller rig and track tests.

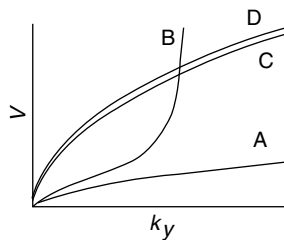
## X. THE ORE COMPETITION

In the 1950s, the newly formed Office for Research and Experiments (ORE) of the International Union of Railways held a competition for the best analysis of the stability of a two-axle railway vehicle. The specification for the competition, drawn up by Committee C9 under the chairmanship of Robert Levi, emphasised worn wheel and rail profiles and nonlinear effects, for it was still widely held, in spite of Carter's work, that the explanation for instability lay in some way in the nonlinearities of the system.<sup>47</sup> The three prize winning papers (by de Possel, Boutefoy, and Matsudaira),<sup>48</sup> in fact, all gave linearised analyses. However, Matsudaira's paper was alone in incorporating both longitudinal and lateral suspension stiffness between wheelsets and frame. Surprisingly, it was awarded only the third prize.

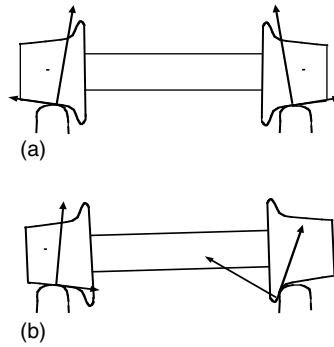
Matsudaira's model has suspension stiffnesses but no suspension damping. It has six degrees of freedom, lateral displacement, and yaw of the wheelsets and car body so that roll of the car body is neglected. Worn wheel and rail profiles were approximated by circular arcs in order to give an approximation for the effective conicity similar to that of Heumann. At the critical state between stable and unstable motion sinusoidal oscillation is possible, and so with this assumption Matsudaira was able to reduce the order of the characteristic equation, making it possible to derive a stability chart by a graphical method (Figure 2.6). In this way Matsudaira avoided the onerous task of calculating the actual eigenvalues for each pair of parameters. In this chart the lines representing an eigenvalue with a zero real part are plotted in the plane of speed vs. lateral suspension stiffness (Figure 2.6). There are four of these lines in the chart on which purely sinusoidal oscillations are possible. Two of these, A and B, relate to relatively large excursions of the car body and two, C and D, relate to relatively large excursions of the wheelsets. Interpreting these as stability boundaries, Matsudaira proposed two approaches to achieve stability for a practical range of speeds. One was to make the lateral stiffness rather large and exploit the lower stable region; the other was to use what was interpreted as an upper stable region with the smallest possible value of the lateral stiffness. In this latter case, Matsudaira suggested that the vehicle goes through an unstable region at a very low speed but the ensuing hunting is not severe.

In fact, a later examination of the root locus shows that each eigenvalue has a positive real part above its critical speed, although the magnitude of the real part corresponding to the two lowest critical speeds decreases as the speed increases. The reason that this prescription may work for the two-axle vehicle is that inclusion of suspension damping in the lateral direction can, under certain conditions, eliminate the instability at low speeds completely as was discovered later.

Another factor, which emerged for the first time in de Possel's and Boutefoy's papers but was neglected in Matsudaira's paper, was that of the gravitational stiffness, the lateral resultant of the resolved normal forces at the contact points between wheel and rail (Figure 2.7). On its own, this effect would be strongly stabilising but, in fact, it is largely counteracted by the lateral force due to



**FIGURE 2.6** Stability chart in which the lines represent an eigenvalue with a zero real part.  $V$  = vehicle speed;  $k_y$  = lateral suspension stiffness; A and B correspond to mainly oscillations of the car body on the suspension and C and D correspond to wheelset oscillations.



**FIGURE 2.7** Normal and lateral tangential forces acting on wheelset: (a) in central position, (b) in laterally displaced position, illustrating the gravitational stiffness effect.

spin creep (discovered later, see below) for small lateral displacements. The resulting contact stiffness is therefore often ignored for motions within the flangeway clearance, although the correct representation of these forces in the case of flange contact, or in the case of freely rotating wheels is, of course, vital.

Thus, by the early 1960s, the basic ingredients of an analytical model of the lateral dynamics of a railway vehicle had been identified. Of these, few values of the creep coefficients had been measured, and in any case were not under the control of the designer; conicities could only be controlled within a narrow range by reprofiling, and the leading dimensions and number of wheelsets were largely dictated by the proposed duty of the vehicle. However, Matsudaira recognized that the designer could vary both the way in which wheelsets were connected and the corresponding stiffness properties and this pointed the way to future progress.

One of the members of Levi's ORE committee was de Pater, who considered the hunting problem and formulated it as a nonlinear problem.<sup>49,50</sup> Even though severe assumptions were made, interesting theoretical results emerged. In 1964, one of de Pater's students, P. van Bommel, published nonlinear calculations for a two-axle vehicle using wheel and rail profiles and a creep force-creepage law measured by Müller for the ORE C-9 committee.<sup>51</sup> However, lateral and longitudinal suspension flexibility was not considered.

## XI. CREEP

Following Carter's analysis of creep, similar results were obtained by Poritsky,<sup>52</sup> and in the discussion therein Cain<sup>53</sup> pointed out that the region of adhesion must lie at the leading edge of the contact area. A three-dimensional case was solved approximately by Johnson,<sup>54</sup> who considered an elastic sphere rolling on an elastic plane. This solution was based on the assumption that the area of adhesion is circular and tangential to the area of contact (which is also circular) at the leading edge. Good agreement with experiment was obtained. The influence of spin about an axis normal to the contact area was first studied by Johnson,<sup>55</sup> who showed that spin could generate significant lateral force owing to the curvature of the strain field in the vicinity of the contact patch (the couple about the common normal is small and may be safely neglected). The general case where the contact area is elliptical was considered by Haines and Ollerton,<sup>56</sup> who confined their attention to creep in the direction of motion and assumed that Carter's two-dimensional stress distribution held in strips parallel to the direction of motion. A general theory for the elliptical contact area, based on similar assumptions to those made in Ref. 54, was developed by Vermeulen and Johnson,<sup>57</sup> yielding the relationship between creepage and tangential forces for arbitrary values of the semi-axes of the contact area. This, suitably modified to take account of spin, was later much used in vehicle



dynamics studies. De Pater<sup>58</sup> initiated the complete solution of the problem by considering the case where the contact area is circular, and derived solutions for both small and large creepages, without making assumptions about the shape of the area of adhesion. However, this analysis was confined to the case where Poisson's ratio was zero; Kalker<sup>59</sup> gave a complete analytical treatment for the case in which Poisson's ratio is not zero. The agreement between these theoretical results and the experimental results of Johnson<sup>54</sup> is very good. Kalker gave a full solution of the general three-dimensional case in Ref. 60, covering arbitrary creepage and spin for the case of dry friction and ideal elastic bodies, and subsequently gave simpler approximate solution methods.<sup>61</sup> Kalker's theory is described in Ref. 62.

## XII. THE COMPLETE SOLUTION OF THE HUNTING PROBLEM

Matsudaira's work was not well known in Europe and the results of the ORE competition were inconclusive. This was partly due to the fact that railway engineers were reluctant to accept the results of theory, perhaps justifiably so in view of the draconian assumptions that at that date were necessary to complete an analysis. In the early 1960s, British Railways, similar to Japanese National Railways, faced an increasing incidence of derailments of short wheelbase two-axle wagons as freight train speeds increased. Various modifications to wagon suspensions were made and tested but to no avail and it was obvious that there was a lack of understanding of the basic dynamics of these vehicles. As a result, a team was formed at British Rail Research Department, Derby to undertake research into railway vehicle dynamics. (Editor's note: the writer was recruited to lead this team, having been an aeroelastician in the aerospace industry.) Armed with a description of Carter's work in Rocard's textbook,<sup>39</sup> a ride on a test coach with hunting bogies<sup>63</sup> made it clear that the origin of hunting was dynamic instability. It was decided to mount a combined theoretical and experimental attack on the problem in order to understand it, find practical solutions, and to convince railway engineers of the relevance of the results.

Although some simple studies were made of a rigid bogie it was soon realised that the influence of the primary suspension, as Matsudaira had shown in his 1960 paper, was all important. Therefore, the two-axle vehicle was chosen for detailed study. Not only was this important in its own right, and relevant to the derailment problem mentioned above, but it could represent a bogie and was the simplest complete realistic vehicle. Relying on symmetry considerations, stability is governed by a set of seven differential equations, involving lateral displacement and yaw of the wheelsets and car body, and roll of the car body. Solutions of these equations were obtained numerically using digital computers, by simulation using analogue computers, and by a variety of analytical techniques. Insight into the behaviour of the system was gained by the derivation of solutions for special cases and for particular subsystems.

For example, comprehensive details of the behaviour of a simple elastically restrained wheelset were derived. Not only is this a good approximation at high speeds for certain vehicles, but as it is a system with only two degrees of freedom analytical analysis is straightforward, and the basic mechanics are revealed. As the equations of motion are not symmetric and the system is nonconservative, the wheelset is able to convert energy from the forward motion to the energy of the lateral motion.<sup>64</sup> Moreover, the representation and analysis of the wheelset as a feedback system was introduced.

A significant result of these studies was the scope for improvement of the stability of the two-axle vehicle by appropriate selection of the suspension parameters. Two approaches are possible, both of which make it possible to achieve quite high critical speeds. In the first of these, as exploited by Matsudaira in the Shinkansen bogies, both lateral and yaw primary stiffnesses are increased, there being an optimum at which stability is a maximum. This optimum depends in a complex way on the creep coefficients and conicity. This approach is most appropriate for bogies where the wheelbase is comparatively small.

The second approach employs a relatively flexible suspension. It depends on extending the analysis of stability by introducing a new feature, lateral suspension damping, and by reintroducing the gravitational stiffness effect which de Possel and Boutefoy had already used.<sup>48</sup> As a result it was shown that, with a careful choice of lateral suspension damping and the lateral and longitudinal stiffnesses so that they satisfied certain inequalities and were neither too small nor too large, it was possible to eliminate the low-speed body instability (a strongly contributory factor in wagon derailments) so that the vehicle operating speed was only limited by the wheelset instability.

So far, in the studies of stability, only the creepage due to longitudinal and lateral relative motion between wheel and rail had been considered, and the relative angular motion about the normal to the contact plane, the spin, had been neglected. When the effect of spin creep was included in the equations,<sup>65</sup> the stabilising influence of gravitational stiffness was found to be much reduced and had to be counteracted by an increase in yaw stiffness.

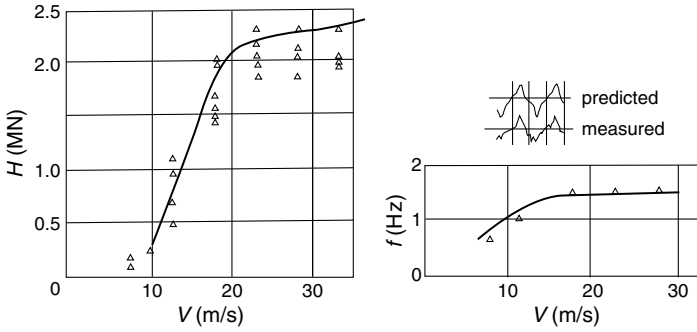
In this work, the application of analytical and both analogue and digital computer techniques marched hand-in-hand with experimental work on models and full-scale vehicles. It was fortunate that this was the heyday of corporate research, with centralised facilities and, most importantly, at British Rail Research, the ability to carry out full-scale experiments on a real railway.

On the full-scale experimental side, the earliest measurements of critical speeds and mode shapes associated with the hunting limit cycle of a range of vehicles were made by King<sup>63</sup> and by Pooley.<sup>66</sup>

The earliest model experiments were at one fifth scale on a roller rig. The model was dynamically scaled, longitudinal and lateral stiffnesses were provided by adjustable independent swing links. Satisfactory agreement was obtained between theory and experiment. The same model was tested on a model test track.

A striking validation of the theory came from a series of full-scale experiments with two kinds of standard two-axle vehicles (Gilchrist et al.).<sup>67</sup> These experiments were sponsored by the chief civil engineer of British Railways, but significantly, not the chief mechanical engineer. The dynamic response of the vehicles was measured on a full-scale test track that featured a series of track imperfections. In addition, as the linear critical speeds of these vehicles were low, it was possible to measure the fully developed hunting limit cycle. Careful measurement of the vehicle parameters followed by a nonlinear analogue simulation led to a successful replication of the fully developed hunting motion of these vehicles by Hobbs (published in Ref. 67). Quite apart from the highly nonlinear suspension characteristics, which were realistically modelled, two major limitations of linear theory were faced. These were creep saturation and wheel–rail geometry. Crude representations of creep saturation had long been available, but consideration was given to the representation of the graph of rolling radius difference against lateral displacement of the wheelset, which determines the yaw moment of the longitudinal creep forces in the equations of motion, as mentioned above. Hence the concept of “equivalent conicity” was introduced. For a coned wheelset the equivalent conicity is simply the cone angle of the tread. For a wheelset with worn or profiled treads the equivalent conicity is defined as that cone angle which for purely coned wheels would produce the same wavelength of kinematic oscillation and is approximately equal to the mean slope out to the amplitude in question. In other words, the equivalent conicity is a “describing function,” a method of dealing with nonlinear control system components introduced by Kochenburger.<sup>68</sup> The circular arc theory of Matsudaira and others is accurate only for extremely small lateral displacements of the wheelset. An example of the comparative results by Gilchrist et al. are shown in [Figure 2.8](#), and demonstrate that both the onset of instability and the fully developed hunting limit cycle were satisfactorily modelled.

The next step was to design and build the full-scale variable parameter test vehicle HSFV-1, which was designed in accordance with the flexible suspension concept described above. Experimental verification of the stability boundaries predicted by the above prescription was obtained from full-scale roller rig and track testing of the specially constructed vehicle HSFV-1.



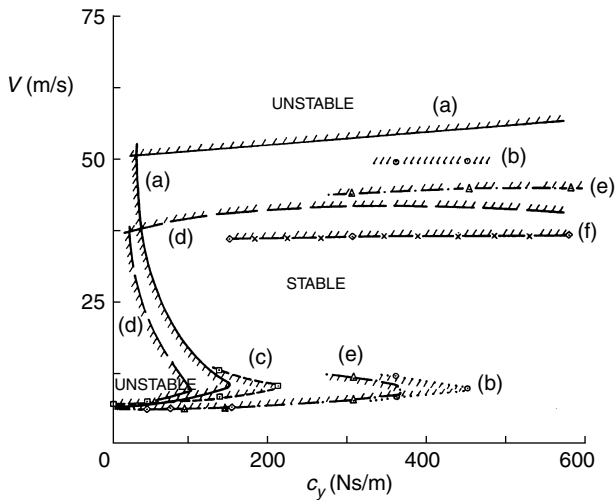
**FIGURE 2.8** Measured and predicted lateral forces  $H$  and frequency  $f$  during hunting of a two-axle vehicle as a function of forward speed  $V$ . Inset shows waveforms of  $H$ . Predicted results indicated by full lines. From Ref. 67, in modern units.

Some results are shown in Figure 2.9,<sup>69</sup> showing the elimination of the low-speed body instability, for a suitable choice of the parameters.

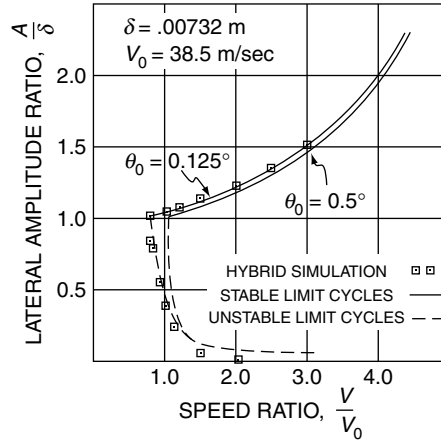
The operating speed was then limited only by the high-speed wheelset instability so that the critical speed was 225 km/h. The results of Figure 2.9 show that qualitative agreement is excellent, but it is clear that refinements to the theory to take account of the influence of the rollers on the joint wheel–rail geometry are necessary.

Although a larger roller rig was built at Derby, the use of roller rigs there was soon entirely superseded by track tests using random process methods of analysis.

Various approaches to the analysis of nonlinear hunting motions have been developed. Cooperrider et al.<sup>70</sup> introduced the more formal method of “quasilinearisation” in which the nonlinear functions are replaced by linear functions so chosen to minimise the mean-square error between the nonlinear and the quasilinear response. They also introduced the limit cycle or bifurcation diagram, an example of which is shown in Figure 2.10. This procedure was extended by



**FIGURE 2.9** Comparison between predicted stability boundaries and roller rig and track test measurements as lateral suspension damping  $c_y$  is varied (from Ref. 69 in modern units). (a) Predicted (equivalent conicity  $\lambda = 0.27$ , yaw stiffness  $k_\psi = 5.88$  MNm), (b) roller rig ( $k_\psi = 5.88$  MNm), (c) track tests ( $k_\psi = 5.88$  MNm), (d) predicted ( $\lambda = 0.27$ ,  $k_\psi = 3.06$  MNm), (e) roller rig ( $k_\psi = 5.20$  MNm), (f) roller rig ( $k_\psi = 3.06$  MNm).



**FIGURE 2.10** Limit cycle or bifurcation diagram.<sup>70</sup>  $\delta$  = nominal flangeway clearance;  $A$  = lateral wheelset amplitude of oscillation;  $V_0$  = nonlinear critical speed;  $\theta_0$  = breakaway yaw angle in yaw spring in series with dry friction.

Gasche, Moelle, and Knothe<sup>71,72</sup> who approximated the limit cycle by a Fourier series and used a Galerkin method to solve the equations. This made it possible to establish much detail about the limit cycle.

Developments in nonlinear dynamics revealed that apparently simple dynamical systems with strong nonlinearities can respond to a disturbance in complex ways. In fact, for certain ranges of parameters no periodic solution may exist. Moreover, systems with large nonlinearities may respond to a disturbance in an apparently random way. In this case, the response is deterministic but is very sensitive to the initial conditions. Such chaotic motions have been studied for railway vehicles by True.<sup>73,74</sup>

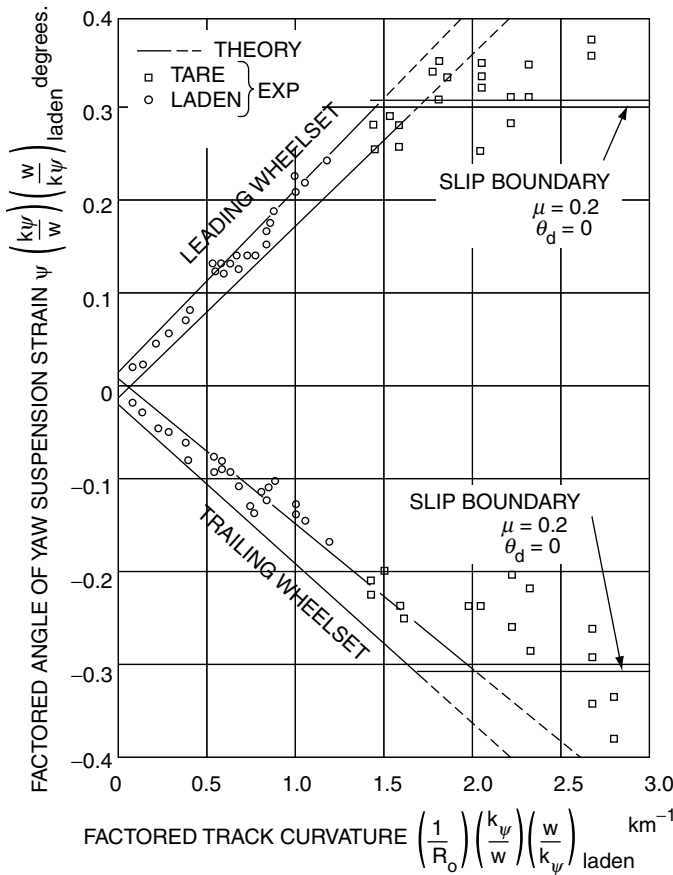
### XIII. MODERN RESEARCH ON CURVING

In the late sixties, Boocock (at British Rail Research)<sup>75</sup> and Newland<sup>76</sup> independently considered the curving of a vehicle using the same equations of motion used in stability analyses but with terms on the right-hand side representing the input due to curvature and cant deficiency. As the wheelsets are constrained by the longitudinal and lateral springs connecting them to the rest of the vehicle, the wheelsets are not able to take up the radial attitude of perfect steering envisaged by Redtenbacher. Instead, a wheelset will balance a yaw couple applied to it by the suspension by moving further in a radial direction so as to generate equal and opposite longitudinal creep forces, and it will balance a lateral force by yawing further. For the complete vehicle, the attitude of the vehicle in the curve and the set of forces acting on it, are obtained by solving the equations of equilibrium. Newland’s model made useful simplifications, but Boocock analysed several configurations including a complete bogie vehicle, a two-axle vehicle and vehicles with cross-braced bogies. The bogie vehicle had 14 degrees of freedom representing lateral displacement and yaw of the wheelsets, bogie frames, and car body. He also included the effects of gravitational stiffness and spin creep. Most important of all, Boocock was able to obtain experimental full-scale confirmation of his theory using the two-axle research vehicle HSFV-1 (Figure 2.11).

These linear theories are valid only for large radius curves. On most curves, the curving of conventional vehicles involves the same nonlinearities due to creep saturation and wheel–rail geometry that were noted in the case of hunting. The first comprehensive nonlinear treatment of practical vehicles in curves was given by Elkins and Gostling (also at British Rail Research).<sup>77</sup> Their

treatment covers the movement of the contact patch across the wheel tread through the flange root and on to the flange and its subsequent change in shape, assuming a single point of contact, appropriate for worn or profiled wheels. At this stage the complication of two-point contact was to be the subject of much future research. As the contact moves across wheel and rail, account is taken of the increasing inclination of the normal force and the lateral creep force generated by spin. They used Kalker's results for the tangential creep forces for arbitrary values of creepage and spin and for a wide range of contact ellipticities, Kalker having issued his results numerically in a tablebook. Elkins and Gostling installed this table in their computer program so that values could be read by interpolation as needed. The resulting equations were solved by iterative numerical procedures, of which two alternatives were given. Elkins and Gostling's program required input in numerical form of the wheel and rail cross-sectional profiles, and much research was carried out on the measurement and analysis of profiles.

The accuracy of the predictions of Elkins and Gostling was demonstrated by experiments carried out on HSFV-1 and the tilting train research vehicle APT-E, an example of their results being given in Figure 2.12. Both vehicles were heavily instrumented, including load-measuring wheels,<sup>78</sup> allowing individual wheel forces to be measured and the limiting value of friction to be identified.



**FIGURE 2.11** Comparison between theory and experiment for the steering behaviour of a two-axle vehicle. (HSFV-1,<sup>75</sup> results are factored to account for two loading conditions).

This comprehensive theory encompassed the two extremes envisaged in the earlier theories. Large radius curves, high conicity, and a high coefficient of friction produced agreement with the linear theory, and small radius curves and flange contacts produced agreement with Porter's results.

The importance of wheel and rail profiles was, as discussed above, recognised in the 1930s. Müller gave a detailed analysis of the wheel-rail contact geometry<sup>79</sup> and he tabulated geometric data,<sup>80</sup> which was measured for a combination of worn wheels and rails. In the early 1960s, King evaluated the contact conditions between a pair of worn wheels and worn rails,<sup>81</sup> between worn wheels and new rails<sup>82</sup> and on the basis of this work designed a new standard wheel profile for British Railways which was subsequently adopted.<sup>83</sup> In this work it was shown that the graph of rolling radius difference vs. wheelset lateral displacement was extremely sensitive to the gauge of the track, rail inclination, and small variations in profile geometry. As a result more refined measuring and computational techniques were developed.<sup>84</sup> Such methods were subsequently used very widely (see for example<sup>85,86</sup>).

The neglect of the effect of wheelset yaw on the wheel-rail geometry is a realistic assumption apart from in the case of flange contact at large angles of wheelset yaw. In addition to Müller's pioneering analysis, three-dimensional geometry analyses were developed in the 1970s by Cooperrider et al.<sup>86</sup> and Hauschild.<sup>87</sup> Research on the topic continues (see for example Duffek,<sup>88</sup> de Pater<sup>89</sup> and Guang<sup>90</sup>).

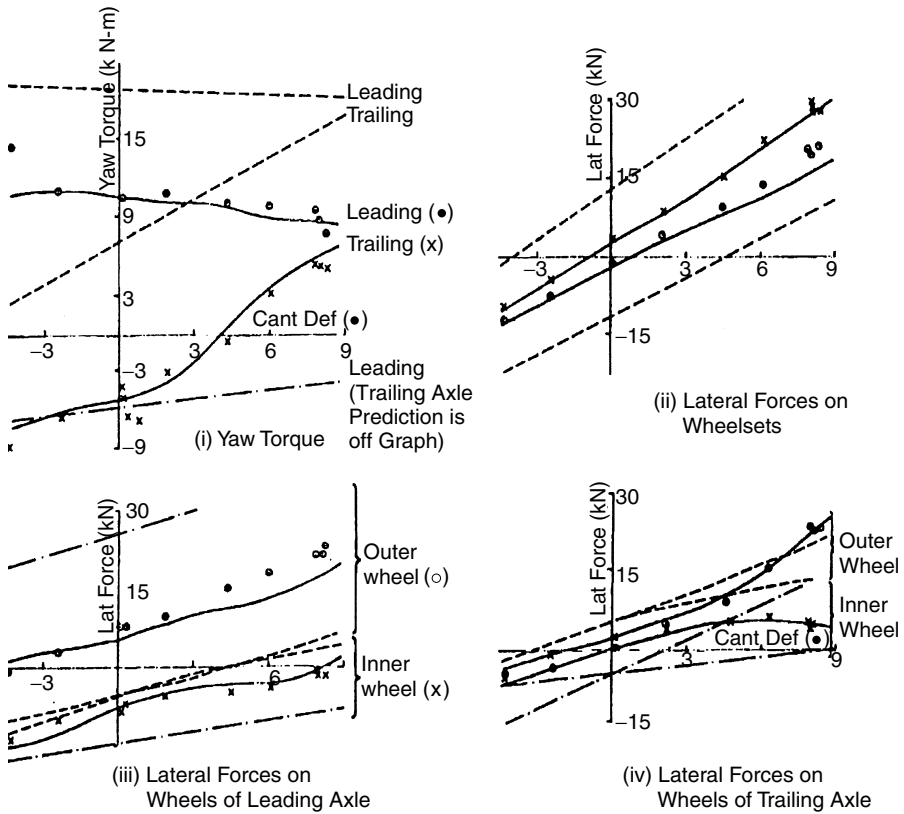


FIGURE 2.12 Curving test results for APT-E<sup>77</sup> for curve radius  $R_0 = 650$  m and cant  $\phi_0 = 150$  mm. Linear theory - - - - -; Porter's theory — — —; Nonlinear theory — — —.

Many wheel–rail combinations experience contact at two points on one wheel for certain values of the lateral wheelset displacement. This commonly occurs, for example, when contact is made between the throat of the flange and the gauge corner of the rail. If the wheels and rails are considered to be rigid, as in the case of single-point contact, discontinuities occur in the geometric characteristics such as the rolling radius difference and slope difference graphs. The mathematical aspects of two-point contact in this case have been considered by Guang.<sup>90</sup> However, in this case the distribution of forces between the points of contact depends on the elasticity in the contact areas and the formulation of the equations of motion becomes more complicated.<sup>91,92</sup>

#### **XIV. DYNAMIC RESPONSE TO TRACK GEOMETRY**

Before vehicle dynamics became established, it was engineering practice to carry out a simple static analysis and tests to measure the amount of wheel unloading on track with a defined degree of twist. In 1964, Gilchrist et al.<sup>67</sup> computed the dynamic response of two-axle vehicles to a dipped rail joint, and compared the results with experiment. Jenkins et al.<sup>93</sup> analysed the vertical response of a vehicle to a dipped rail joint in 1974 and showed that the response involved two distinct peaks. The first fast transient involves the rail mass and the contact stiffness and the second slower transient involves the unsprung mass and the track stiffness. Subsequently, quite complex models of track and vehicle have been used to establish transient stresses resulting from geometric defects in both track and wheels.

For the lateral motions of railway vehicles, the excitation terms in the equations of motion that had been derived for stability analysis first originated with Hobbs.<sup>94</sup> These were validated by Illingworth, using a model roller rig.<sup>95</sup> A comprehensive approach to the dynamic response to large discrete inputs, including both suspension and wheel–rail contact nonlinearities, was carried out by Clark et al.,<sup>96</sup> in 1980 and includes full-scale experimental validation. Similar calculations have been carried out on the response of vehicles to switch and crossing work.

In the case where it can be considered that irregularities are distributed continuously along the track the approach offered by stochastic process theory is appropriate, and this was first applied by Hobbs<sup>94</sup> to the lateral motions of a restrained wheelset. The response of complete vehicles became established as an indication of ride quality, and for passenger comfort assessments existing international standards, which define frequency weighting characteristics, were extended to lower frequencies to cover the railway case. The first measurements of power spectral density using a specially developed trolley-based measurement system were carried out by Gilchrist.<sup>97</sup> Subsequently, extensive measurements of the power spectra of irregularities of track have been made on railways in many countries, resulting in inputs used in design.

With the advent of research into innovative transport systems involving air cushions and magnetic levitation in the late 1960s, the problem of interaction between a flexible track and the vehicle received renewed attention. A review paper by Kortüm and Wormley<sup>98</sup> indicates the progress achieved by 1981 in developing appropriate computer models.

The development of the heavy-haul railway with extremely long trains, often with locomotives attached at various points along the train, introduced serious problems arising from the longitudinal response of trains to hills and to braking. The availability of the digital computer made it possible to develop dynamic models and also to support the development of train-driving simulators.

#### **XV. SUSPENSION DESIGN CONCEPTS AND OPTIMISATION**

It has been discussed above that in the period before adequate mathematical models existed, evolution of the suspension of the railway vehicle had been based on rather general ideas. Based on empirical development and simple calculations, surprisingly good results had been obtained before

1960, providing that conicities were kept low by re-turning wheel treads and speeds were moderate, for example below 160 km/h.

When linear theories of the curving of railway vehicles became available it became possible for the first time to consider the best compromise between the requirements of stability and curving on a numerate basis. This, of course, had been a continuing design consideration from the earliest days of the railways but attempts to resolve it often resulted in intuitive schemes of articulation, as discussed in Section VI. The derivation of configurations and sets of suspension parameters, which are optimal with respect to the needs of stability and curving, are a continuing task for the vehicle designer. Not only are the parameters that are associated with wheel–rail contact, both geometrical and frictional, not under the control of the designer or operator, but they are not known exactly and can vary over a wide range. It follows that practical designs must be very robust in relation to such parameters. On the other hand, there is enormous scope for the design of the suspension system in terms of the way in which the wheelsets and car bodies in a train are connected.

Boocock<sup>75</sup> defined the bending and shear stiffnesses that characterise the elastic properties of the connections (actual or equivalent) between the wheelsets of a two-axle vehicle. It was shown<sup>99</sup> that such a vehicle should possess zero bending stiffness to achieve radial steering, but would be dynamically unstable at low speeds. The design of a two-axle vehicle with a purely elastic suspension therefore requires a compromise between stability and curving. However, Hobbs<sup>100</sup> showed that the use of yaw relaxation dampers could provide sufficient flexibility at low frequencies in curves and sufficient elastic restraint at high frequencies to prevent wheelset instability.

It was also shown by Boocock that for conventional bogies, in which there are primary longitudinal and lateral springs connecting the wheelsets to a frame, there is a limit to the overall shear stiffness that can be provided in relation to the bending stiffness and therefore the stability/curving trade-off in which the bending stiffness must be minimised is constrained. This limitation is removed if the wheelsets are connected directly by diagonal elastic elements or cross-bracing, or interconnections which are structurally equivalent. Such an arrangement is termed a self-steering bogie. Superficially, this arrangement is similar to the systems of articulation between axles by means of rigid linkages which have been discussed in Section VI. In order to discover optimal configurations, various representations of generic two-axle vehicles or bogies have been considered,<sup>101–106</sup> and an example is shown in Figure 2.13.

In the 1970s the self-steering bogie was successfully developed and put into service, notably by Scheffel.<sup>107</sup> Self-steering bogies have been applied to locomotives (with benefits to the maximum exploitation of adhesion), passenger vehicles, and freight vehicles.<sup>108</sup> It should be noted that inter-wheelset connections can be provided by means other than springs and dampers. In Ref. 109 the equivalent of cross-bracing was provided by means of a passive hydrostatic circuit which has a number of potential design advantages.

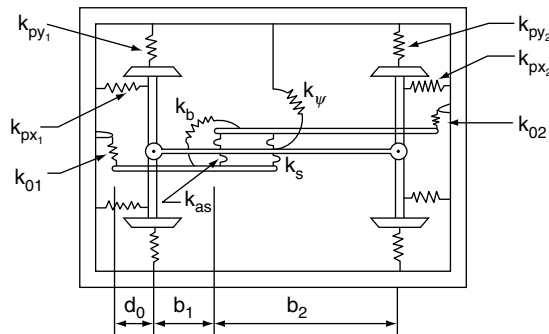


FIGURE 2.13 A generic bogie configuration.<sup>102</sup>



An alternative to providing self-steering by means of elastic or rigid linkages directly between wheelsets is to use a linkage system which allows the wheelsets to take up a radial position but provides stabilising elastic restraint from the vehicle body. This is so-called forced steering as it can be considered that the vehicle body imposes a radial position on the wheelsets. It was early in the 20th century that significant and successful development was carried out by Liechty.<sup>110,111</sup> Schwanck<sup>112</sup> reported on service experience with a particular design of body-steered bogie and its advantages of reduced wheel and rail wear, reduced energy consumption, and increased safety against derailment. Analytical studies of forced-steered bogie vehicles were initiated in 1981 by Bell and Hedrick<sup>113</sup> and Gilmore,<sup>114</sup> who identified various instabilities which were promoted by low conicities and reduced creep coefficients. A considerable body of work by Anderson and Smith and colleagues<sup>115–120</sup> covers the analysis of a vehicle with bogies having separately steered wheelsets. Weeks<sup>121</sup> described dynamic modelling and track testing of vehicles with steered bogies, noting the enhanced sensitivity of this type of configuration to constructional misalignments. Many examples of body steering are in current use.

It was shown<sup>122</sup> that for a vehicle with three or more axles it is possible, in principle, to arrange the suspension so that radial steering and dynamic stability are both achieved. The three-axle vehicle was examined in this context in a series of papers.<sup>122–124</sup> A similar approach, although with slightly different assumptions, has been followed by de Pater<sup>125,126</sup> and Keizer.<sup>127</sup> Although various forms of instability became evident, no radical avenue for improvement emerged from this configuration.

All the configurations discussed so far have been symmetric fore-and-aft. Unsymmetric configurations make it possible, in principle, to achieve a better compromise between curving and dynamic stability, at least in one direction of motion. However, additional forms of instability can occur.

A general theory for the stability of unsymmetric vehicles and the derivation of theorems relating the stability characteristics in forward motion with those in reverse motion has been provided.<sup>128,129</sup> In the case of articulated two-axle vehicles at low speeds it was shown that a suitable choice of elastic restraint in the inter-wheelset connections results in static and dynamic stability in forward and reverse motion, which will steer perfectly, without any modification dependent on the direction of motion. However, the margin of stability is small.

In the case of the three-piece freight truck, calculations taking account of the effects of unsymmetric wear have been described by Tuten et al.<sup>130</sup> Illingworth<sup>131</sup> suggested the use of unsymmetric stiffness in steering bogies. Elkins<sup>132</sup> showed both by calculation and experiment that a configuration of bogie, with the trailing axle having independently rotating wheels and the leading axle conventional, significantly improved stability and curving performance and reduced rolling resistance.

Suda<sup>133,134</sup> studied bogies with unsymmetric stiffnesses and symmetric conicity, and this development work has led to application in service. The concept has been extended to include lack of symmetry of the wheelsets by equipping the trailing axle with freely rotating wheels. This provides an example of a reconfigurable design as the wheelsets are provided with a lock which is released on the trailing wheelset (allowing free rotation of the wheels) and locked on the leading wheelset (providing a solid axle). The lock is switched depending on the direction of motion.

Independently rotating wheels have been frequently proposed as they eliminate the classic hunting problem. Some of the possibilities have been surveyed by Frederick.<sup>135</sup> The essential difference between a conventional wheelset and independent wheels lies in the ability of the two wheels to rotate at different speeds and thus the kinematic oscillation of a conventional wheelset is therefore eliminated. A measure of guidance is then provided by the lateral component of the gravitational stiffness (reduced by the lateral force due to spin creep) which becomes the flange force when the flangeway clearance is taken up, but this leads to slow self-centring action. Extensive experimental experience has shown that, indeed, the kinematic oscillation is absent but

that one or other of the wheels tends to run in continuous flange contact.<sup>136</sup> Good agreement between calculation and experiment is demonstrated.<sup>137–139</sup>

An attempt to increase the effect of the lateral resultant gravitational force but reduce the amount of spin is to incline substantially from the horizontal the axis of rotation of the wheels, as suggested by Wiesinger (see de Pater<sup>140</sup>). A generic wheelset model including the effect of modest amounts of camber has been studied theoretically and experimentally by Jaschinski and Netter.<sup>141</sup>

Another important modern development is the use of articulation on vehicles with single-axle running gear. A current example of a train which has single-axle running gear is the Copenhagen S-Tog which embodies forced steering of the wheelsets through hydraulic actuators driven by the angle between adjacent car bodies. Extensive design calculations have been carried out on this train<sup>142</sup> and its lateral stability has also been discussed.<sup>143</sup>

These developments demonstrate that the availability of computational tools in vehicle dynamics has made it possible to depart from conventional configurations tried and tested by empirical means.

## XVI. DERAILMENT

The conditions necessary to sustain equilibrium of the forces in flange contact were considered in 1908 by Nadal in a classical analysis<sup>144</sup> which provided a derailment criterion. Gilchrist and Brickle<sup>145</sup> applied Kalker's theory of creep in a reexamination of Nadal's analysis and they have shown that Nadal's formula is correct only for the most pessimistic case when the angle of attack is large and the longitudinal creep on the flange is small.

The dynamics of the derailment process was considered by Matsui<sup>146</sup> and Sweet et al.,<sup>147–150</sup> who obtained good agreement with model experiments. The mechanics of derailment remains a topic for research and presents a challenge to the modelling of vehicle dynamics.

## XVII. THE DEVELOPMENT OF COMPUTER SIMULATION

Carter, as discussed above, formulated equations of motion which he was able to use to examine stability of rigid-framed locomotives. The analysis of the restrained wheelset, perhaps the most fruitful model with which to understand stability, could be carried out analytically. Once the importance of the suspension stiffnesses was recognised, the solution of the equations of motion of a system with many degrees of freedom presented a problem that, if attempted, was time-consuming and error-prone even using the calculating machines of the early 1950s. As already mentioned, Matsudaira resorted to a graphical method to establish stability boundaries for a two-axle vehicle. The advent of the digital computer provided the means to calculate the eigenvalues for complete vehicles, while the general purpose electronic analogue computer could be used to compute stability boundaries. Among the first to use these computers on a large scale was the aircraft industry, and it is interesting that in 1962 the first eigenvalue analysis of a complete two-axle vehicle carried out by British Rail Research used a flutter routine and computer at English Electric Aviation. (The linear equations of motion of a railway vehicle and the aeroelastic equations of an aircraft wing are formally the same, provided that vehicle speed is interpreted appropriately.) The work by Gilchrist et al.,<sup>67</sup> in which the simulation of two-axle vehicles involved severe nonlinearities, was carried out using an analogue computer.

With the application of the computer on a large scale in all branches of engineering,<sup>151</sup> interest in numerical methods quickened and efficient methods of computation became available in the form of libraries of standard routines for the eigenvalue problem and the step-by-step solution of differential equations. Thus, in the 1960s and 1970s, simulations of complex nonlinear railway vehicle models with many degrees of freedom were developed that exploited the increase in

computer power becoming available. These simulations were based on equations of motion derived manually which were then incorporated into computer programs to solve specific problems such as stability, response to track irregularities, and curving.

The derivation of equations of motion, their reduction to first-order differential equations suitable for numerical integration or eigenvalue analysis, and the calculation of the parameters in the equations were lengthy and error prone tasks. Moreover, the application of railway vehicle dynamics was moving from the research laboratory to the industrial design office. Many companies set up groups specialising in bogie and suspension design dependent on dynamics calculations. This motivated the development of complete packages, which covered a range of dynamics calculations, using the same consistent model of the vehicle. In these programs, the input data consist of basic dimensions, masses, and the type of interconnections, such as massless force elements representing springs and dampers that connect the bodies that make up the vehicle. The variety and scope of programs available was reviewed in 1993.<sup>152</sup> Two different approaches to simulation emerged.

In the first approach, the equations embodied in the package are restricted to more-or-less standard configurations, such as symmetric bogie vehicles and to specific situations, such as stability or curving. Approximations are made in the formulation of the equations which are consistent with the expected behaviour of these configurations. The subsequent limits of applicability are typically validated by full-scale experiment for the specific situations for which the package is applicable. Such software is extensively used in industry where the emphasis is on design.

The second approach to simulation involves the use of general vehicle models in general situations, so-called multibody programs. The theoretical basis of the multibody approach had its origin in work carried out on satellite dynamics in the 1960s when satellites became more complex and could no longer be considered as single rigid bodies. The degree of generality varies: for example, some packages cater for large rotational angles. Software has been developed that allows the automatic formulation of the complex equations of motion by the computer. A further development is that the formulation of the equations of motion may be carried symbolically instead of numerically.

Another variation in methods lies in the treatment of constraints associated with the motion of the wheelset rolling along the track. Some packages work with a set of differential equations (involving all the states, i.e., displacements and velocities, of the wheelset) and algebraic equations expressing the constraints, and the constraints are not used to reduce the order of the system. Alternatively, many packages work with generalised coordinates and a minimal set of differential equations. With the choice of a moving axis system orientated with reference to the track, the wheelset reference axes are not fixed in the wheelset and are themselves moving with respect to inertial space. The formulation of the equations of motion therefore requires special care and their logical derivation has been considered by de Pater,<sup>153</sup> Guang,<sup>90</sup> and Schiehlen.<sup>154</sup>

Packages also vary in the extent to which they interface to finite-element structural packages, necessary for studies involving structural flexibility, and to control system analysis software, necessary for studies involving active suspensions.

As computer power has increased and vehicle dynamics computations have been used more widely for engineering design purposes it has become necessary to refine the modelling of the suspension components themselves.<sup>155–157</sup> For example, components such as air springs and dampers are complex and cannot be represented adequately by simple springs and viscous dampers. As an example of the sophistication of modelling now considered necessary, the modelling of an articulated vehicle consisting of three car bodies and four single-axle bogie frames and wheelsets<sup>158</sup> involves 167 states to cover these bodies and the suspension details. The maturity of the subject is indicated by the successful results of the Manchester benchmarks for rail simulation<sup>159</sup> in which the performance of five packages was compared.

## XVIII. ACTIVE SUSPENSIONS

As defined by Goodall<sup>162</sup> the concept of an active suspension is to add sensors, a controller, and actuators to an existing mechanical system, and usually involves feedback action so that the dynamics of the system is modified. The concept of feedback is an ancient one although Watts governor was the first widespread application. In railways there are three principal areas of application of active suspensions — car body tilting systems, secondary, and primary suspensions. Attention in the following will be confined to the earliest developments and their relationship with vehicle dynamics, as modern development is thoroughly reported in a number of reviews.<sup>162–164</sup>

As mentioned above, superelevation of the track in curves was used at an early date, and tilting the car body to achieve the same effect was demonstrated in monorail systems, such as that at Wuppertal with cars suspended from an overhead rail, and gyroscopically stabilised cars, such as that of Brennan (1906). In the latter case the developments in dynamics exemplified by Routh's textbook had promoted intense scientific and popular interest in gyroscopes in the 1890s. Various demonstrations were made of monorail systems which exploited gyroscopic stabilisation, that of Brennan being perhaps the most successful within the limits of contemporary technology.

Bogies with provision for pendular suspension of the car body were put into service in limited numbers in the United States.<sup>160</sup> Experiments were carried out on a vehicle with full passive tilting in France,<sup>161</sup> and in the late 1960s the United Aircraft turbotrain had full passive tilting. The limited dynamic performance of passive tilt systems, together with advances in control system analysis techniques and the availability of sensors and actuators from other fields of technology, stimulated the development of active tilt starting in 1967. The evolution of the various configurations of tilt systems has been described by Goodall.<sup>162</sup> After a long period of development this technology became established in commercial operation, representing the first service application of active suspensions, the Pendolino having started operation in Italy in 1988. This has been followed by operation of tilting trains, developed by several manufacturers, in a number of countries.<sup>162</sup>

Also starting in the late 1960s, research was carried out on the applications of active systems to vertical and lateral vibration isolation and stability augmentation involving the secondary suspension, initially largely motivated by development activity on air cushion vehicles. The basic theoretical considerations have been outlined by Hedrick.<sup>163</sup> For conventional rail vehicles, a variety of active systems were the subject of full-scale experimentation,<sup>164</sup> but even though significant benefits in terms of ride quality were demonstrated, the additional cost of the equipment has deterred commercial application, so that there are only a few examples of service operation.<sup>162</sup>

The suspension and guidance of vehicles by magnetic suspension provides another example of an active suspension. Magnetic suspension relies on active control for stabilisation. Albertson, Bachelet, and Graeminger all proposed schemes for magnetic levitation in the 1900s, and in 1938 Kemper<sup>165</sup> demonstrated a model showing the feasibility of a wheelless train. Subsequently, with the advent of high power solid-state electronic devices in the mid-1960s, and the application of various forms of linear motors, research and development on the topic has flourished. There has been considerable cross-fertilisation between the dynamics of rail and magnetically levitated vehicles as a result, both in terms of technique and personnel.

Turning to primary suspensions, feedback control system methods were used in the stability analysis of the wheelset in 1962.<sup>166</sup> This approach revealed some of the deficiencies of the wheelset as a guidance element. Bennington<sup>167</sup> used control system techniques to propose an active torque connection between the two wheels of a wheelset, and subsequently, a wheelset with an active torque connection using a magnetic coupling was developed.<sup>168</sup> Various control laws were used with the object of providing a good torque connection between the wheels at low frequencies so that curving ability is maintained but at high frequencies the wheels are more or less uncoupled so that instability does not arise. Pascal and Petit<sup>169</sup> carried out experiments in which active steering, using freely rotating wheels, was achieved by the yaw moment generated by electromagnets which react

against a guide rail. Currently, many possibilities are being considered.<sup>162</sup> It can be seen that, increasingly, control engineering techniques will exercise a strong influence on the dynamics of railway vehicles (see for example<sup>170</sup>), either by improving the dynamics of vehicles using the conventional wheelset, or by supporting the development of more innovative systems.

## XIX. THE EXPANDING DOMAIN OF RAIL VEHICLE DYNAMICS

The proceedings of recent IAVSD symposia reveal the growing range of dynamics studies undertaken in the railway field. This chapter has reviewed the development of ideas about the basic problems associated with stability, response to track geometry, and behaviour in curves of the railway vehicle. A few final examples will indicate how rail vehicle dynamics is growing in scope, both in the range of frequencies considered and the detail of the models used, to match the challenge of higher speeds and the application of new technology.

The extrapolation of conventional railway technology to higher speeds has led in many cases to increased traction forces, increased wheelset mass, and greater track stiffness. New problems of interaction between vehicle and track have emerged, such as irregular ballast settlement and deterioration, increased levels of rail corrugation, and out-of-round wheels. The solution of these problems requires the consideration of structural dynamics of both vehicle and track in the frequency range of about 40 to 400 Hz together with the analysis of the long-term behaviour of wheel and track components.<sup>171</sup>

The analysis of mechanisms of noise generation requires the consideration of an even higher range of frequencies of structural oscillations of wheelsets and track, up to 5 kHz or more. At these frequencies a nonsteady-state analysis of the contact forces is needed. Moreover, for problems such as corrugation and squealing in curves it is necessary to account for the contact forces at large values of the creepages.<sup>172</sup>

Both tractive and guidance forces are provided by the same wheel–rail interface. It has gradually been recognised that the interaction between traction and guidance due to the contact forces at the wheel–rail interface may severely affect overall performance. Although traction and guidance have usually been considered separately, a systems approach is needed for design in which the control of the drive system is combined with the needs of guidance.<sup>173</sup>

At high speeds, particularly in tunnels, aerodynamic forces are significant not only as a generator of drag but in affecting the lateral response of vehicles. With the further increases in speed, and reductions in mass of car bodies, lateral oscillations have been experienced that have been created by pressure fluctuations caused by unsteady flow separations from the car body surface. This requires simulation in which the vehicle dynamics is combined with the aerodynamics of the flow field with moving boundaries and which is dependent on the car body motion.<sup>174</sup>

Increasing use of electronics and active controls makes it necessary to adopt a mechatronic approach in which the mechanical parts and the electronics are seen as integral parts of the system to be analysed and designed concurrently. Industrial applications demand integration of software tools with design and manufacturing systems.<sup>175</sup>

Although there is much to be understood about the behaviour of apparently simple systems with strong nonlinearities,<sup>176</sup> the history of railway vehicle dynamics suggests that, in most cases, the subject rests on a sound conceptual basis with satisfactory full-scale experimental validation.

## REFERENCES

1. Wickens, A. H., The dynamics of railway vehicle — From Stephenson to Carter, *Proc. Instn. Mech. Engrs.*, 212(Part F), 209–217, 1999.
2. Dendy Marshall, C. F., *A History of British Railways down to the Year 1830*, Oxford University Press, London, pp.147–148, 1938.

3. Vaughan, A., *Isambard Kingdom Brunel — Engineering Knight Errant*, John Murray, London, p. 102, 1992.
4. Klingel, W., Über den Lauf der Eisenbahnwagen auf Gerarder Bahn, *Organ Fortsch. Eisenb.-wes.*, 38, 113–123, 1883.
5. Winans, R., British Patent 5796, 1829.
6. Adams, W. B., On the impedimental friction between wheel tire and rails with plans for improvement, *Proc. Instn. Civ. Engrs.*, 23, 411, 1863–1864.
7. Redtenbacher, F. J., *Die Gesetze des Locomotiv-Bauwes*, Verlag von Friedrich Bassermann, Mannheim, p. 22, 1855.
8. Mackenzie, J., Resistance on railway curves as an element of danger, *Proc. Instn. Civ. Engrs.*, 74, 1–57, 1883.
9. Gilchrist, A. O., The long road to solution of the railway hunting and curving problems, *Proc. Instn. Mech. Engrs.*, 212(Part F), 219–226, 1998.
10. Heumann, H., Das Verhalten von Eisenbahnfahrzeugen in Gleisbogen, *Organ Fortsch. Eisenb.-wes.*, 68, 104–108, 1913, see also pp. 118–121, 136–140, 158–163.
11. Porter, S. R. M., *The Mechanics of a Locomotive on Curved Track*, The Railway Gazette, London, 1935.
12. Clark, D. K., *Railway Machinery*, Blackie and Sons, London, 1855.
13. Timoshenko, S. P., *A History of the Strength of Materials*, Mcgraw-Hill, New York, pp. 173–178, 1953.
14. Weiner, L., *Articulated Locomotives*, Constable, London, 1930.
15. Liechty, R., *Das Bogenlaefige Eisenbahn-Fahrzeug*, Schulthess, Zurich, 1934.
16. Wickens, A. H., *Fundamentals of Rail Vehicle Dynamics—Guidance and Stability*, Swets & Zeitlinger, Lisse, 2003.
17. Ahrons, E. L., *The British Steam Railway Locomotive 1825–1925*, The Locomotive Publishing Company, London, pp. 160–162, 1927.
18. White, J. H., *A History of the American Locomotive*, The Johns Hopkins Press, Baltimore, pp. 173–174, 1968.
19. Hopkirk, K. R., Frederick William Carter, 1870–1952, *Obituary Notices of Fellows of The Royal Society*, 8, 373–388, 1952–1953.
20. Carter, F. W., The electric locomotive, *Proc. Instn. Civ. Engrs.*, 221, 221–252, 1916.
21. Reynolds, O., On the efficiency of belts or straps as communicators of work, *The Engineer*, 27 November 1874.
22. Bennett, S., *A History of Control Engineering*, Peter Peregrinus, London, chap. 3, 1979.
23. Routh, E. J., *Stability of a Given State of Motion*, Macmillan, London, 1877, Reprinted London, Taylor and Francis, 1975.
24. Routh, E. J., *Dynamics of a System of Rigid Bodies (Advanced Part)*, 1st ed., Macmillan, London, 1860, 6th ed., 1905.
25. Hurwitz, A., Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt, *Math. Ann.*, 46, 273–284, 1895.
26. Bryan, G. H. and Williams, W. E., The longitudinal stability of aeroplane gliders, *Proc. R. Soc.*, 73, 100–116, 1904.
27. Hopkinson, B., The “Hunting” of alternating-current machinery, *Proc. R. Soc.*, 72, 235–252, 1904.
28. Whipple, F. J. W., The stability of motion of the bicycle, *Q. J. Math.*, 30, 312–348, 1899.
29. Bairstow, L. and Page, A., Oscillations of the tailplane and body of an aeroplane in flight, *Aeronaut. Res. Counc. R. M.*, 276(Part 2)1916.
30. Frazer, R. A. and Duncan, W. J., The flutter of aeroplane wings, *Aeronaut. Res. Counc. R. M.*, 1155, 1928.
31. Frazer, R. A., Duncan, W. J., and Collar, A. R., *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, Cambridge University Press, Cambridge, 1938.
32. Carter, F. W., *Railway Electric Traction*, Edward Arnold, London, 1922.
33. Love, A. E. H., *Mathematical Theory of Elasticity*, 2nd ed., Cambridge University Press, Cambridge, pp. 195–198, 1906.
34. Carter, F. W., On the action of a locomotive driving wheel, *Proc. R. Soc.*, A112, 151–157, 1926.
35. Timoshenko, S. P., *A History of the Strength of Materials*, Mcgraw-Hill, New York, p. 348, 1953.
36. Carter, F. W., On the stability of running of locomotives, *Proc. R. Soc.*, A121, 585–611, 1928.

37. Carter, F. W., *The Running of Locomotives, with Reference to Their Tendency to Derail, Selected Engineering Papers No. 81*, Institute Civil Engineers, 1930.
38. Carter, F. W., British Patents 128106 (1918), 155038 (1919), 163185, (1920).
39. Rocard, Y., La stabilite de route des locomotives, *Actual. Sci. Ind.*, 234, 1935, Part 1. Rocard, Y., *General Dynamics of Vibrations*, Crosby-Lockwood, London, 1960, (translation from French, 1st publication 1943).
40. Langer, B. F. and Shamberger, J. P., Dynamic stability of railway trucks, *Trans. Am. Soc. Mech. Eng.*, 57, 481–493, 1935.
41. Cain, B. S., *Vibration of Road and Rail Vehicles*, Pitman, New York, pp. 149–189, 1940.
42. Davies, R. D., Some experiments on the lateral oscillation of railway vehicles, *J. Instn. Civ. Engrs.*, 11, 224–261, 1939.
43. Heumann, H., Lauf der Drehgestell-Radsätze in der Geraden, *Organ Fortschr. Eisenb.-wes.*, 92, 336–342, 1937.
44. Jaschinski, A., Chollet, H., Iwnicki, S., Wickens, A. H., and Von Würzen, U., The application of roller rigs to railway vehicle dynamics, *Vehicle Syst. Dyn.*, 31, 345–392, 1999.
45. Matsudaira, T., Shimmy of axles with pair of wheels (in Japanese), *J. Rail. Eng. Res.*, 16–26, 1952.
46. Shima, H., *The New Tokaido Line: Brief Notes on the Way the Idea of the Construction Was Developed, Convention on Guided Land Transport*, The Institution of Mechanical Engineers, 1966.
47. Lévi, R., Study of hunting movement, Office of Research and Experiment (ORE) of the International Unions of Railways (UIC), Question C9 Report, 5 May, 1953.
48. de Possel, R., Beaufefoy, J., and Matsudaira, T., Papers awarded prizes in the competition sponsored by Office of Research and Experiment (ORE) of the International Union of Railways (UIC), ORE-Report RP2/SVA-C9, ORE, Utrecht, 1960.
49. de Pater, A. D., Etude du mouvement de lacet d'un vehicule de chemin de fer, *Appl. Sci. Res. A*, 6, 263–316, 1956.
50. de Pater, A. D., The approximate determination of the hunting movement of a railway vehicle by aid of the method of Krylov and Bogoljubov, *Appl. Sci. Res.*, 10, 205–228, 1961.
51. van Bommel, P., Application de la theorie des vibrations nonlineaires sur le problem du mouvement de lacet d'un vehicule de chemin de fer. Doctoral dissertation, Technische Hogeschool Delft, 1964.
52. Poritsky, H., Stresses and deflections of cylindrical bodies in contact with application to contact of gears and of locomotive wheels, *J. Appl. Mech. Trans. ASME*, 72, 191–201, 1950.
53. Cain, B. S., Discussion of Ref. 52, *J. Appl. Mech. Trans. ASME*, 72, 465–466, 1950.
54. Johnson, K. L., The effect of tangential contact force upon the rolling motion of an elastic sphere upon a plane, *J. Appl. Mech. Trans. ASME*, 80, 339–346, 1958.
55. Johnson, K. L., The effect of spin upon the rolling motion of an elastic sphere upon a plane, *J. Appl. Mech. Trans. ASME*, 80, 332–338, 1958.
56. Haines, D. J. and Ollerton, E., Contact stress distributions on elliptical contact surfaces subjected to radial and tangential forces, *Proc. Instn. Mech. Engrs.*, 177, 95–114, 1963.
57. Vermeulen, P. J. and Johnson, K. L., Contact of non-spherical elastic bodies transmitting tangential forces, *J. Appl. Mech. Trans. ASME*, 86, 338–340, 1964.
58. de Pater, A. D., *On the reciprocal pressure between two elastic bodies, Proceedings of Symposium on Rolling Contact Phenomena*, Elsevier, Amsterdam, pp. 29–74, 1962.
59. Kalker, J. J., The transmission of force and couple between two elastically similar rolling spheres, *Proc. Kon. Ned. Akad. Wet. Amsterdam*, B70, 135–177, 1964.
60. Kalker, J. J., On the rolling of two elastic bodies in the presence of dry friction, Doctoral Thesis, Delft University of Technology, 1967.
61. Kalker, J. J., A fast algorithm for the simplified theory of rolling contact, *Vehicle Syst. Dyn.*, 11, 1–13, 1982.
62. Kalker, J. J., *Three-Dimensional Elastic Bodies in Rolling Contact*, Kluwer Academic Publishers, Dordrecht, 1990.
63. King, B. L., The measurement of the mode of hunting of a coach fitted with standard double-bolster bogies, British Railways Research Department Report E439, 1963.
64. Wickens, A. H., The dynamic stability of railway vehicle wheelsets and bogies having profiled wheels, *Int. J. Solids Struct.*, 1, 319–341, 1965.

65. Wickens, A. H., The dynamics of railway vehicles on straight track: Fundamental considerations of lateral stability, *Proc. Instn. Mech. Engrs.*, 180(Part 3F), 29–44, See discussion, p. 150, 1965–1966.
66. Pooley, R. A., Assessment of the critical speeds of various types of four-wheeled vehicles, British Railways Research Department Report E557, 1965.
67. Gilchrist, A. O., Hobbs, A. E. W., King, B. L., and Washby, V., The riding of two particular designs of four wheeled vehicle, *Proc. Instn. Mech. Engrs.*, 180, 99–113, 1965.
68. Kochenburger, R. J., Frequency-response methods for analysis of a relay servomechanism, *Trans. AIEEE*, 69, 270–284, 1950.
69. Hobbs, A. E. W., The lateral stability of HSFV-I. British Railways Research Technical Note DYN/53. September, 1967.
70. Cooperrider, N. K., Hedrick, J. K., Law, E. H., and Malstrom, C. W., The application of quasilinearization techniques to the prediction of nonlinear railway vehicle response, In *The Dynamics of Vehicles on Roads and on Tracks, Proceedings of the IUTAM Symposium, Delft, The Netherlands, August 1975*, Pacejka, H. B., Ed., Swets & Zeitlinger, Lisse, pp. 314–325, 1975.
71. Moelle, D. and Gasch, R., Nonlinear bogie hunting, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Seventh IAVSD Symposium, Cambridge, September 1981*, Wickens, A. H. Ed., Swets & Zeitlinger Publishers, Lisse, pp. 455–467, 1982.
72. Gasch, R., Moelle, D., and Knothe, K., The effects of non-linearities on the limit-cycles of railway vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Eighth IAVSD Symposium, Cambridge, MA., August 1983*, Hedrick, J. K., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 207–224, 1984.
73. True, H., Dynamics of a rolling wheelset, *App. Mech. Rev.*, 46, 438–444, 1993.
74. True, H., Railway vehicle chaos and asymmetric hunting, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 12th IAVSD Symposium, Linköping, Sweden, August 1991*, Sauvage, G., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 625–637, 1992.
75. Boocock, D., Steady-state motion of railway vehicles on curved track, *J. Mech. Eng. Sci.*, 11, 556–566, 1969.
76. Newland, D. E., Steering characteristics of bogies, *Railway Gazette*, 124, 745–750, 1968.
77. Elkins, J. A. and Gostling, R. J., A general quasi-static curving theory for railway vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Fifth IAVSD Symposium, Vienna, Austria, September 1977*, Slibar, A. and Springer, H., Eds., Swets & Zeitlinger Publishers, Lisse, pp. 388–406, 1978.
78. Pocklington, A. R. and Allen, R. A., Improved data from load-measuring wheels, *Railway Eng.*, 2(4), 37–43, 1977.
79. Müller, C. Th., Kinematik, Spurführungsgeometrie und Führungsvermögen der Eisenbahnradatz, *Glaser's Annalen*, 77, 264–281, 1953.
80. Muller, C. Th., Wear profiles of wheels and rails, Office of Research and Experiment (ORE) of the International Union of Railways (UIC), ORE-Report C9/RP6, Utrecht, 1960.
81. King, B. L., An evaluation of the contact conditions between a pair of worn wheels and worn rails in straight track, British Railways Research Technical Note DYN/37, September, 1966.
82. King, B. L., An evaluation of the contact conditions between a pair of worn wheels and new rails in straight track, British Railways Research Technical Note DYN/42, December, 1966.
83. King, B. L., The design of new tyre profiles for use on British Railways, British Railways Research Technical Note DYN/38, October, 1966.
84. Gostling, R. J., The measurement of real wheel and track profiles and their use in finding contact conditions, equivalent conicity and equilibrium rolling line, British Rail Research Technical Note TN DA 22, 1971.
85. Nefzger, A., Geometrie der Berührung zwischen Radsatz und Gleis, *ETR*, 23, 113–122, 1974.
86. Cooperrider, N. K., Hedrick, J. K., Law, E. H., Kadala, P. S., and Tuten, J. M., Analytical and experimental determination of nonlinear wheel/rail geometric constraints, Report FRA-O&RD 76-244, US Department of Transportation, Washington, 1975.
87. Hauschild, W., Die Kinematik des Rad-Scheine Systems, Institut für Mechanik, Technische Universität Berlin, 1977.
88. Duffek, W., *Contact Geometry in Wheel Rail Mechanics, Proceedings of Symposium Contact Mechanics and Wear of Rail/Wheel Systems*, Kalousek, J. et al., Eds., University of Waterloo Press, pp. 161–179, 1982.



89. de Pater, A. D., The geometric contact between wheel and rail, *Vehicle Syst. Dyn.*, 17(3), 127–140, 1988.
90. Guang, Y., Dynamic analysis of railway wheelsets and complete vehicle systems, Delft University of Technology, Faculty for Mechanical Engineering and Marine Technology, Doctoral Thesis, pp. 42–50, 1993.
91. Netter, H., Schupp, G., Rulka, W., and Schroeder, K., New aspects of contact modelling and validation within multibody system simulation of railway vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 15th IAVSD Symposium, Budapest, August 1997*, Palkovics, L., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 246–269, 1992.
92. Pascal, J.-P., About multi-Hertzian contact hypothesis and equivalent conicity in the case of S1002 and UIC60 analytical wheel/rail profiles, *Vehicle Syst. Dyn.*, 22, 263–275, 1993.
93. Jenkins, H. H., Stephenson, J. E., Clayton, G. A., Moreland, G. W., and Lyon, D., The effect of track and vehicle parameters on wheel/rail vertical dynamic forces, *Railway Eng. J.*, 2–16, 1974, January.
94. Hobbs, A. E. W., The response of a restrained wheelset to variations in the alignment of an ideally straight track, British Railways Research Department Report E542, 1964.
95. Illingworth, R., The mechanism of railway vehicle excitation by track irregularities, Doctoral Thesis, Oxford University, 1973.
96. Clark, R. A., Eickhoff, B. M., and Hunt, G. A., Prediction of the dynamic response of vehicles to lateral track irregularities, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Seventh IAVSD Symposium, Cambridge, U.K. September 1981*, Wickens, A. H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 535–548, 1982.
97. Gilchrist, A. O., Power spectral measurements by TMM 1: Proving trials and three site measurements. British Railways Research Technical Note DYN/67. September, 1967.
98. Kortüm, W. and Wormley, D. N., Dynamic interactions between traveling vehicles and guideway systems, *Vehicle Syst. Dyn.*, 10, 285–317, 1981.
99. Wickens, A. H., Steering and dynamic stability of railway vehicles, *Vehicle Syst. Dyn.*, 5, 15–46, 1978.
100. Hobbs, A. E. W., Improvements in or relating to railway vehicles, British Patent Specification 1261896, 1972.
101. Horak, D., Bell, C. E., and Hedrick, J. K., A comparison of the stability performance of radial and conventional rail vehicle trucks, *ASME J. Dyn. Syst. Meas. Control*, 103, 181, 1981.
102. Kar, A. K., Wormley, D. N., and Hedrick, J. K., Generic rail truck characteristics, In *The Dynamics of Vehicles on Roads and on Railway Tracks, Proceedings of Sixth IAVSD Symposium, Berlin, September 1979*, Willumeit, H.-P., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 198–210, 1980.
103. Kar, A. K. and Wormley, D. N., Generic properties and performance characteristics of passenger rail vehicles, In *The Dynamics of Vehicles on Roads and on Railway Tracks, Proceedings of Seventh IAVSD Symposium, Cambridge, September 1981*, Wickens, A. H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 329–341, 1982.
104. Fujioka, T., Generic representation of primary suspensions of rail vehicles, In *The Dynamics of Vehicles on Roads and on Railway Tracks, Proceedings of 11th IAVSD Symposium, Kingston, August 1989*, Wickens, A. H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 233–247, 1989.
105. Fujioka, T., Suda, Y., and Iguchi, M., Representation of primary suspensions of rail vehicles and performance of radial trucks, *Bull. JSME*, 27(232), 2249–2257, 1984.
106. Hedrick, J. K., Wormley, D. N., Kim, A. K., Kar, A. K., and Baum, W., Performance limits of rail passenger vehicles: Conventional, radial and innovative trucks, U.S. Department of Transportation Report DOT/RSPA/DPB-50/81/28, 1982.
107. Scheffel, H., A new design approach for railway vehicle suspensions, *Rail Int.*, 5, 638–651, 1974.
108. Scheffel, H., Unconventional bogie designs — Their practical basis and historical background, *Vehicle Syst. Dyn.*, 24(6–7), 497–524, 1995.
109. Wickens, A. H., British Patent 1179723, 1967.
110. Liechty, R., Studie über die Spurführung von Eisenbahnfahrzeugen, *Schweizer Archiv f. Angewandte Wissenschaft und Technik*, 3, 81–100, 1937.
111. Liechty, R., Die Bewegungen der Eisenbahnfahrzeuge auf den schienen und die dabei auftretenden Kräfte, *Elektrische Bahnen*, 16, 17–27, 1940.
112. Schwanck, U., Wheelset steering for bogies of railway vehicles, *Rail Eng. Int.*, 4, 352–359, 1974.

113. Bell, C. E. and Hedrick, J. K., Forced steering of rail vehicles: Stability and curving mechanics, *Vehicle Syst. Dyn.*, 10, 357–385, 1981.
114. Gilmore, D. C., The application of linear modelling to the development of a light steerable transit truck, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Seventh IAVSD Symposium, Cambridge, September 1981*, Wickens, A. H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 371–384, 1982.
115. Fortin, J. A. and Anderson, R. J., Steady-state and dynamic predictions of the curving performance of forced-steering rail vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Eighth IAVSD Symposium, Cambridge, MA, August 1983*, Hedrick, J. K., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 179–192, 1984.
116. Fortin, J. A. C., Anderson, R. J., and Gilmore, D. C., Validation of a computer simulation forced-steering rail vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Ninth IAVSD Symposium, Linköping, June 1985*, Nordstrom, O., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 100–111, 1986.
117. Smith, R. E. and Anderson, R. J., Characteristics of guided-steering railway trucks, *Vehicle Syst. Dyn.*, 17, 1–36, 1988.
118. Anderson, R. J. and Fortin, C., Low conicity instabilities in forced-steering railway vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Tenth IAVSD Symposium, Prague, August 1987*, Apetaur, M., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 17–28, 1988.
119. Smith, R. E., Forced-steered truck and vehicle dynamic modes-resonance effects due to car geometry, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Tenth IAVSD Symposium, Prague, August 1987*, Apetaur, A., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 423–424, 1988.
120. Smith, R. E., Dynamic characteristics of steered railway vehicles and implications for design, *Vehicle Syst. Dyn.*, 18, 45–69, 1989.
121. Weeks, R., The design and testing of a bogie with a mechanical steering linkage, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Tenth IAVSD Symposium, Prague, August 1987*, Apetaur, M., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 497–508, 1988.
122. Wickens, A. H., Stability criteria for articulated railway vehicles possessing perfect steering, *Vehicle Syst. Dyn.*, 7(1), 33–48, 1979.
123. Wickens, A. H., Static and dynamic stability of a class of three-axle railway vehicles possessing perfect steering, *Vehicle Syst. Dyn.*, 6(1), 1–19, 1977.
124. Wickens, A. H., Flutter and divergence instabilities in systems of railway vehicles with semi-rigid articulation, *Vehicle Syst. Dyn.*, 8(1), 33–48, 1979.
125. de Pater, A. D., Optimal design of a railway vehicle with regard to cant deficiency forces and stability behaviour, Delft University of Technology, Laboratory for Engineering Mechanics, Report 751, 1984.
126. de Pater, A. D., Optimal design of railway vehicles, *Ingenieur-Archiv*, 57(1), 25–38, 1987.
127. Keizer, C. P., A theory on multi-wheelset systems applied to three wheelsets, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Ninth IAVSD Symposium, Linköping, June 1985*, Nordstrom, O., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 233–249, 1986.
128. Wickens, A. H., Steering and stability of unsymmetric articulated railway vehicles, *Trans. ASME J. Dyn. Syst. Meas. Control.*, 101, 256–262, 1979.
129. Wickens, A. H., Static and dynamic stability of unsymmetric two-axle railway possessing perfect steering, *Vehicle Syst. Dyn.*, 11, 89–106, 1982.
130. Tuten, J. M., Law, E. H., and Cooperrider, N. K., Lateral stability of freight cars with axles having different wheel profiles and asymmetric loading, ASME Paper No. 78-RT-3, 1978.
131. Illingworth, R., The use of unsymmetric plan view suspension in rapid transit steering bogies, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Eighth IAVSD Symposium, Cambridge, MA, August 1983*, Hedrick, J. K., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 252–265, 1984.
132. Elkins, J. A., The performance of three-piece trucks equipped with independently rotating wheels, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 11th IAVSD Symposium, Kingston, Ontario, August 1989*, Anderson, R., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 203–216, 1984.

133. Suda, Y., Improvement of high speed stability and curving performance by parameter control of trucks for rail vehicles considering independently rotating wheelsets and unsymmetric structure, *JSME Int. J. Series III*, 33(2), 176–182, 1990.
134. Suda, Y., High speed stability and curving performance of longitudinally unsymmetric trucks with semi-active control, *Vehicle Syst. Dyn.*, 23, 29–52, 1994.
135. Frederich, F., Possibilities as yet unknown regarding the wheel/rail tracking mechanism, *Rail Int.*, 16, 33–40, 1985.
136. Becker, P., On the use of individual free rolling wheels on railway vehicles, *Eisenbahn Technische Rundschau*, 19, 11, 1970.
137. Eickhoff, B. M. and Harvey, R. F., Theoretical and experimental evaluation of independently rotating wheels for railway vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 11th IAVSD Symposium, Kingston, Ontario, August 1989*, Anderson, R., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 190–202, 1989.
138. Eickhoff, B. M., The application of independently rotating wheels to railway vehicles, *Proc. Instn. Mech. Engrs.*, 205(Part 3F), 43–54, 1991.
139. Elkins, J. A., The performance of three-piece trucks equipped with independently rotating wheels, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 11th IAVSD Symposium, Kingston, Ontario, August 1989*, Anderson, R., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 203–216, 1989.
140. de Pater, A. D., *Analytisch en synthetisch ontwerpen*, Technische Hogeschool, Delft, 1985, pp. 37–38.
141. Jaschinski, A. and Netter, H., Non-linear dynamical investigations by using simplified wheelset models, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 12th IAVSD Symposium, Linköping, Sweden, August 1991*, Sauvage, G., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 284–298, 1992.
142. Rose, R. D., Lenkung und Selbstlenkung von Einzelradsatzfahrwerken am Beispiel des KERF im S-Tog Kopenhagen, *Proceedings of Fourth International Conference on Railway Bogies and Running Gears*, 1998, pp. 123–132.
143. Slivsgaard, E. and Jensen, J. C., On the dynamics of a railway vehicle with a single-axle bogie, *Proceedings of Fourth Mini Conference on Vehicle System Dynamics, Identification and Anomalies*, pp. 197–207, 1994.
144. Nadal, M. J., *Locomotives a Vapeur*, Collection encyclopedie scientifique, bibliotheque de mecanique appliqué et genie, Paris, 1908.
145. Gilchrist, A. O. and Brickle, B. V., A re-examination of the proneness to derailment of a railway wheelset, *J. Mech. Eng. Sci.*, 18, 131–141, 1976.
146. Matsui, N., *On the derailment quotient Q/P*, Railway Technical Research Institute, Japanese National Railways, Japan, 1966.
147. Sweet, L. M., Karmel, A., and Fairley, S. R., Derailment mechanics and safety criteria for complete rail vehicle trucks, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Seventh IAVSD Symposium, Cambridge, August 1983*, Wickens, A. H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 481–494, 1984.
148. Sweet, L. M. and Sivak, J. A., Nonlinear wheelset forces in flange contact — Part I: Steady state analysis and numerical results, *ASME Trans. J. Dyn. Syst. Meas. Control*, September, 1979.
149. Sweet, L. M., Sivak, J. A., Nonlinear wheelset forces in flange contact — Part II: Measurements using dynamically scaled models. *ASME Transactions, J. of Dynamic Systems, Measurement and Control*. September, 1979.
150. Sweet, L. M. and Karmel, A., Evaluation of time-duration dependent wheel load criteria for wheel climb derailment, *ASME Trans. J. Dyn. Syst. Meas. Control*, 103(3), 219–227, 1981.
151. Ceruzzi, P. E., *A History of Modern Computing*, The MIT Press, Cambridge, 1998.
152. Kortum, W., Review of multibody computer codes for vehicle system dynamics, In *Multibody Computer Codes for Vehicle System Dynamics, Supplement to Vehicle System Dynamics*, 22, Kortum, W. and Sharp, R. S., Eds., pp. 3–31, 1993.
153. de Pater, A. D., The equations of motion of a railway wheelset moving over tangent track, LTM Report 985, Delft University of Technology, 1992.

154. Schiehlen, W., Modelling of complex vehicle systems, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Eighth IAVSD Symposium, Cambridge, MA., August 1983*, Hedrick, J. K., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 548–563, 1984.
155. Evans, J. R., The modelling of railway passenger vehicles, In *The Dynamics of Vehicles on Roads and on Railway Tracks, Proceedings of 12th IAVSD Symposium, Lyon, August 1991*, Sauvage, G., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 144–156, 1992.
156. Eickhoff, B. M., Evans, J. R., and Minnis, A. J., A review of modelling methods for railway vehicle suspension components, *Vehicle Syst. Dyn.*, 24(6–7), 469–496, 1995.
157. Sauvage, G., Determining the characteristics of helical springs for applications in suspensions of railway vehicles, *Vehicle Syst. Dyn.*, 13, 13–41, 1984.
158. Gretzscel, M. and Vaculin, O., Simulation of an integrated mechatronic train, In *The Dynamics of Vehicles on Roads and on Railway Tracks, Proceedings of 17th IAVSD Symposium, Lyngby, August 2001*, True, H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 137–144, 2002.
159. Iwnicki, S., Ed., *The Manchester Benchmarks for Rail Vehicle Simulation, Supplement to Vehicle System Dynamics*, 31, 1999.
160. Anonymous Pendulum cars for the Sante Fe, Great Northern and Burlington, *Railway Age*, 112, 248–252, 1942.
161. Mauzin, F., Chartet, M., and Lenoir, M., A new pendulum type carriage for high speed traffic, *Revue Generale des Chemin de Fer*, 76, 581–593, 1957.
162. Goodall, R. M., Active railway suspensions: Implementation status and technological trends, *Vehicle Syst. Dyn.*, 28, 87–117, 1997.
163. Hedrick, J. K., Railway vehicle active suspensions, *Vehicle Syst. Dyn.*, 10, 267–283, 1981.
164. Goodall, R. M. and Kortum, W., Active controls in ground transportation — A review of the state-of-the-art and future potential, *Vehicle Syst. Dyn.*, 12, 225–257, 1983.
165. Kemper, H., Schwebende Auffhängung durch elektromagnetische Kräfte: eine Möglichkeit für eine grundsätzlich neue Fortbewegungsart, *Electrotech. Zeus*, 59, 391–395, 1938.
166. Wickens, A. H., Preliminary analytical study of hunting of an idealised railway vehicle bogie, BR Research Department Report E442, 22 March 1963.
167. Bennington, C. K., The railway wheelset and suspension unit as a closed loop guidance control system. A method for performance improvement, *J. Mech. Eng. Sci.*, 10, 91–100, 1966.
168. Geuenich, W., Guenther, C., and Leo, R., Dynamics of fiber composite bogies with creep-controlled wheelsets, In *The Dynamics of Vehicles on Roads and on Railway Tracks, Proceedings of Eighth IAVSD Symposium, Cambridge, MA., August 1983*, Hedrick, J. K., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 225–238, 1984.
169. Pascal, J. P. and Petit, J. M., Dynamique ferroviaire active; vers l'asservissement des bogies. Le Rail, Juillet-Aout, pp. 32–35, 1988.
170. Goodall, R. M. and Li, H., Solid axle and independently-rotating railway wheelsets — A control engineering assessment, *Vehicle Syst. Dyn.*, 33, 57–67, 2000.
171. Popp, K. and Schiehlen, W., Eds, *System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade*. Berlin, Springer-Verlag, 2003.
172. Periard, F., Wheel-noise generation: curve squealing by trams, Doctoral thesis, Delft University of Technology, Delft, 1998.
173. Mei, T. X., Goodall, R. M., and Wickens, A. H., A systems approach for wheelset and traction control, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 17th IAVSD Symposium, Copenhagen, August 2001*, True, H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 257–266, 2003.
174. Ueki, K., Nakade, K., and Fujimoto, H., Lateral vibration of middle cars of Shinkansen train in tunnel section, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 16th IAVSD Symposium, Pretoria, August–September 1999*, Fröhling, R., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 749–761, 2000.
175. Stribersky, A., Moser, F., Rulka, W., and Trautenberg, W., Advances in combined structural dynamics and system dynamics analyses of rail vehicles, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 17th IAVSD Symposium, Copenhagen, August 2001*, True, H., Ed., Swets & Zeitlinger Publishers, Lisse, pp. 465–477, 2003.
176. True, H., On the theory of nonlinear dynamics and its applications in vehicle systems dynamics, *Vehicle Syst. Dyn.*, 31, 393–421, 1999.

---

# 3 The Anatomy of Railway Vehicle Running Gear

*Anna Orlova and Yuri Boronenko*

## CONTENTS

I.	Main Functions of the Running Gear and Terminology.....	40
II.	Bogie Components.....	40
A.	Wheelsets.....	40
B.	Axleboxes.....	44
C.	Wheels.....	46
D.	Suspension.....	46
E.	Elastic Elements (Springs).....	47
F.	Dampers.....	51
G.	Constraints and Bumpstops.....	56
1.	Horn Guides.....	57
2.	Cylindrical Guides.....	57
3.	Beam Links.....	59
4.	Constraints Using Radius Links.....	59
5.	Constraints Using Trailing (Radial) Arms.....	60
6.	Traction Rods.....	60
F.	Car Body to Bogie Connection.....	61
1.	Flat Centre Plate.....	61
2.	Spherical Centre Bowl.....	61
3.	Centre Pivot.....	62
4.	Watts Linkage.....	62
5.	Pendulum Linkage.....	62
6.	Connection of Car Body to Bolsterless Bogies.....	63
III.	Common Passenger Vehicle Bogie Designs.....	65
IV.	Common Freight Wagon Bogie Designs.....	67
V.	Common Tram Bogie Designs.....	71
VI.	Principles of Selecting Suspension Parameters.....	72
A.	Selecting Vertical Suspension Characteristics.....	73
B.	Selecting the Lateral and Longitudinal Primary Suspension Stiffness.....	75
C.	Selecting Suspension Damping.....	77
VII.	Advanced Bogie Designs.....	79
	References.....	83

## I. MAIN FUNCTIONS OF THE RUNNING GEAR AND TERMINOLOGY

The principal difference between a railway vehicle and other types of wheeled transport is the guidance provided by the track. The surface of the rails not only supports the wheels, but also guides them in a lateral direction. The rails and the switches change the rolling direction of wheels and thus determine the travelling direction of the railway vehicle.

The running gear is the system that provides safe motion of the vehicle along railway track. The running gear includes such components as wheelsets with axleboxes, the elastic suspension, the brakes, the traction drive, and the device to transmit traction and braking forces to the car body. Its main functions are:

- Transmission and equalization of the vertical load from the wheels of the vehicle to the rails
- Guidance of vehicle along the track
- Control of the dynamic forces due to motion over track irregularities, in curves, switches and after impacts between the cars
- Efficient damping of excited oscillations
- Application of traction and braking forces

Depending on the running gear, the vehicles may be described as bogied or bogie-less.

In vehicles without bogies the suspension, brakes, and traction equipment are mounted on the car body frame. The traction and braking forces are transmitted through traction rods or axlebox guides (sometimes known as “horn guides”). Conventional two-axle vehicles will generate larger forces in tight curves than the equivalent bogie vehicle; therefore their length is limited.

Running gear mounted on a separate frame that can turn relative to the vehicle body is known as a bogie (or truck). The number of wheelsets that they unite classifies the bogies. The most common type is the two-axle bogie, but three- and four-axle bogies are also encountered, often on locomotives.

Previously, the bogies simply allowed the running gear to turn in a horizontal plane relative to the car body thus making it possible for the wheelsets to have smaller angles of attack in curves. In modern bogies, the bogie frame transmits all the longitudinal, lateral, and vertical forces between the car body and the wheelsets. The frame also carries braking equipment, traction drive, suspension, and dampers. It may also house tilting devices, lubrication devices for wheel-rail contact and mechanisms to provide radial positioning of wheelsets in curves. Bogied vehicles are normally heavier than two-axle vehicles. However, the design of railway vehicles with bogies is often simpler than for two-axle vehicles and this may provide reliability and maintenance benefits.

## II. BOGIE COMPONENTS

### A. WHEELSETS

A wheelset comprises two wheels rigidly connected by a common axle. The wheelset is supported on bearings mounted on the axle journals.

The wheelset provides:

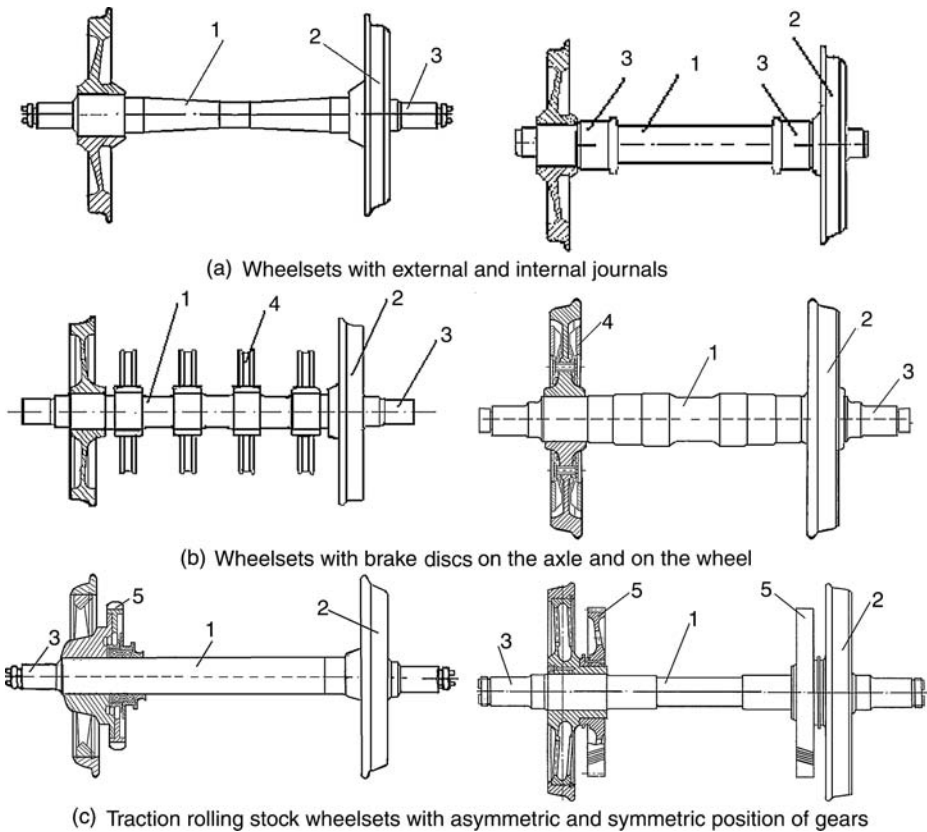
- The necessary distance between the vehicle and the track
- The guidance that determines the motion within the rail gauge, including at curves and switches
- The means of transmitting traction and braking forces to the rails to accelerate and decelerate the vehicle

The design of the wheelset depends on:

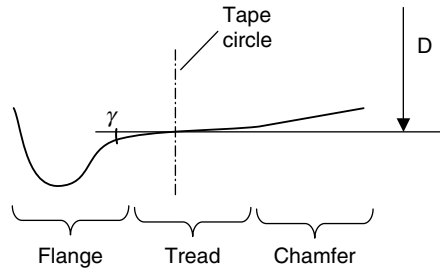
- The type of the vehicle (traction or trailing)
- The type of braking system used (shoe brake, brake disc on the axle, or brake disc on the wheel)
- The construction of the wheel centre and the position of bearings on the axle (inside or outside)
- The desire to limit higher frequency forces by using resilient elements between the wheel centre and the tyre

The main types of wheelset design are shown in Figure 3.1. Despite the variety of designs, all these wheelsets have two common features: the rigid connection between the wheels through the axle and the cross-sectional profile of the wheel rolling surface, named wheel profile.

In curves, the outer rail will be a larger radius than the inner rail. This means that a cylindrical wheel has to travel further on the outer rail than on the inner rail. As the wheels moving on the inner and outer rails must have the same number of rotations per time unit such motion cannot occur by pure rolling. To make the distances travelled by two wheels equal, one or both of them will therefore “slip” thus increasing the rolling resistance and causing wear of wheels and rails. The solution is to machine the rolling surface of wheels to a conical profile with variable inclination angle  $\gamma$  to the axis of the wheelset (Figure 3.2). The position of the contact point when the wheelset



**FIGURE 3.1** Main types of wheelset design: (a) with external and internal journals; (b) with brake discs on the axle and on the wheel; (c) with asymmetric and symmetric position of gears (1, axle; 2, wheel; 3, journal; 4, brake disc; 5, tooth gear).



**FIGURE 3.2** Main elements of a wheel profile.

is at a central position on the rails determines the so-called “tape circle,” where the diameter of the wheel is measured. On the inner side of the wheel, the conical profile has a flange which prevents derailment and guides the vehicle once the available creep forces have been exhausted.

An unrestrained wheelset with conical profiles will move laterally in a curve such that the outer wheel is rolling on a larger radius (due to the cone angle) than the inner one. It can be seen that for each curve radius only one value of conicity exists that eliminates slip. As different railways have varying populations of curve radii the shape of wheel profile that provides minimum slip depends on the features of track. Railway administrations normally specify allowable wheel profiles for their infrastructure and the degree of wear permitted before reprofiling is required.

Figure 3.3 shows several examples of new wheel profiles. For understanding the dynamic behaviour of a railway vehicle the conicity of interface is critical. Conicity is defined as the difference in rolling radii between the wheels for a given lateral shift of the wheelset.

Despite the variety of wheel profiles, they have a number of common features. The width of the profile is typically 125–135 mm and flange height for vehicles is typically 28–30 mm. The flange inclination angle is normally between 65 and 70°. In vicinity of the tape circle the conicity is 1:10 or 1:20 for common rolling stock. For high speed rolling stock, the conicity is reduced to around 1:40 or 1:50 to prevent hunting. It can be seen from Figure 3.3 that the wheel profile has a relief toward the outer side of the wheel. This is intended to lift the outer side of the wheel off the rail and thus ease the motion on switches. Some modern wheel profiles, particularly for passenger rolling stock are not conical but designed instead from a series of radii that approximate a part-worn shape. This is intended to give a more stable shape and prevent the significant changes in conicity that may occur as a conical wheel profile wears. An example of such a profile is the UK P8 wheel profile.

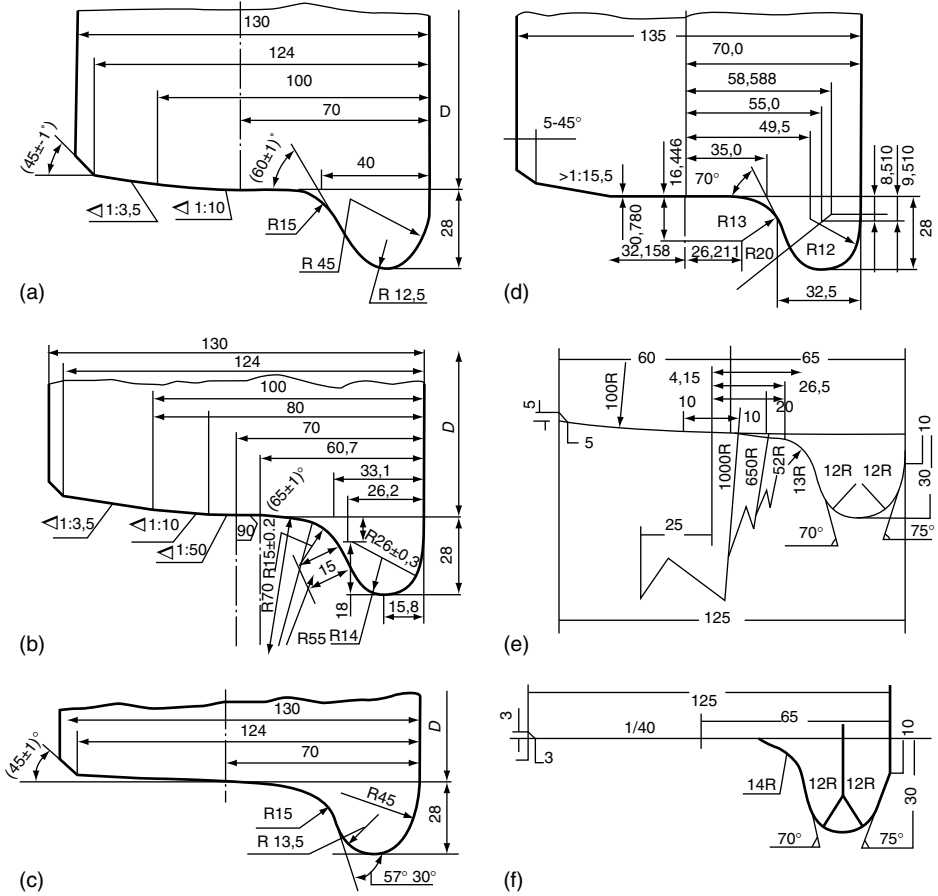
For profiles whose shape is not purely conical (either by design or through wear in service), the term equivalent conicity is applied. This is the ratio of the rolling radius difference to twice the lateral displacement of the wheelset:

$$\gamma_{\text{eq}} = \frac{\Delta R}{2y} \quad (3.1)$$

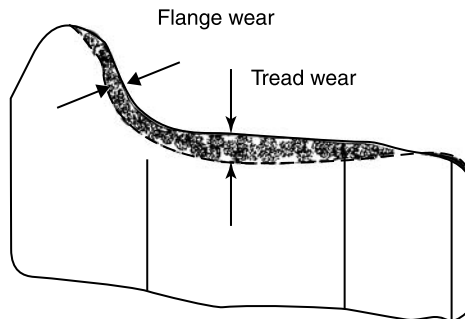
It is important to note that the rolling radius difference is a function of both the wheel and rail shape and hence a wheel profile on its own cannot be described as having an equivalent conicity.

As the wheel wears, the shape of the profile may alter significantly depending upon a large number of factors. These may include the curvature profile of the route, the suspension design, the level of traction and braking forces applied, the average rail profile shape encountered and the lubrication regime. Tread wear (Figure 3.4) will increase the height of the flange and eventually cause it to strike fishplate bolts, etc. If the tread wear causes the profile to become excessively concave damaging stresses may arise at the outer side of the wheel and rail known as false flange damage. Flange wear may lead to increase of the flange angle and reduction of the flange thickness.

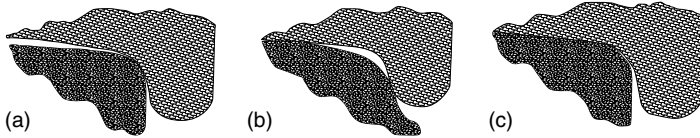




**FIGURE 3.3** Common wheel profiles: (a) for freight and passenger railcars (Russia); (b) for high-speed railcars (Russia); (c) for industrial rolling stock (Russia); (d) for European freight and passenger railcars; (e,f) for high-speed trains (Japan).



**FIGURE 3.4** Tread and flange wear.



**FIGURE 3.5** Possible contact situations between the wheel and the rail: (a) single-point contact; (b) two-point contact; (c) conformal contact.

In extreme conditions, this could increase the risk of switch-splitting derailments. Wheel profiles are generally restored to their design shape by periodic turning on a wheel lathe. This can normally be carried out without the necessity to remove the wheelset from the vehicle.

It is clear that contact conditions will vary considerably depending upon the shape of the wheel and rail profiles. This may take the form of single-point, two-point, or conformal contact as shown in Figure 3.5. One-point contact (a) develops between the conical or tread worn wheel profiles and rounded rail profile. Wheels wear quickly towards the local rail shape. With two-point contact (b) the wheel additionally touches the rail with its flange. In this case, the rolling contact has two different radii which causes intensive slip and fast flange wear. Conformal contact (c) appears when the wheel profile and the gauge side of the railhead wear to the extent that their radii in vicinity of the contact patch become very similar.

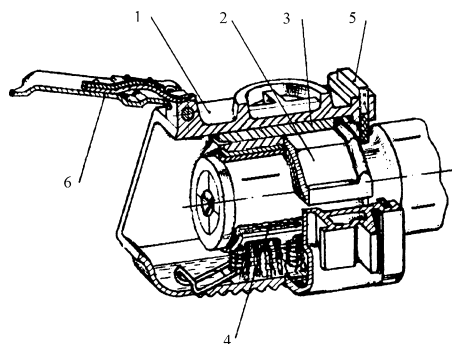
## B. AXLEBOXES

The axlebox is the device that allows the wheelset to rotate by providing the bearing housing and also the mountings for the primary suspension to attach the wheelset to the bogie or vehicle frame. The axlebox transmits longitudinal, lateral, and vertical forces from the wheelset on to the other bogie elements. Axleboxes are classified according to:

- Their position on the axle depending on whether the journals are outside or inside
- The bearing type used, either roller or plain bearings

The external shape of the axlebox is determined by the method of connection between the axlebox and the bogie frame and aims to achieve uniform distribution of forces on the bearing. Internal construction of the axlebox is determined by the bearing and its sealing method.

Axleboxes with plain bearing (Figure 3.6) consist of the housing (1), the bearing itself (2) which is usually made of alloy with low friction coefficient (e.g., bronze or white metal), the bearing shell



**FIGURE 3.6** Construction of an axlebox with friction bearing.

(3) which transmits the forces from the axlebox housing to the bearing, a lubrication device (4) which lubricates the axle journal. Front and rear seals (5 and 6) prevent dirt and foreign bodies entering the axlebox, while the front seal (6) can be removed to monitor the condition of the bearing and add lubricant.

Vertical and longitudinal forces are transmitted through the internal surface of the bearing and lateral forces by its faces.

Plain bearing axleboxes are now largely obsolete as they have several serious disadvantages:

- High friction coefficient when starting from rest
- Poor reliability
- Labour-intensive maintenance
- Environmental pollution

However, from a vehicle dynamic behaviour point of view, axleboxes with plain bearings had certain positive features. In recent years, plain bearing axleboxes that do not require lubrication have been reintroduced on certain types of rolling stock though their use is still rare.

Axleboxes with roller type bearings (Figure 3.7) are classified according to:

- The bearing type (cylindrical, conical, spherical)
- The fitting method (press-fit, shrink-fit, bushing-fit)

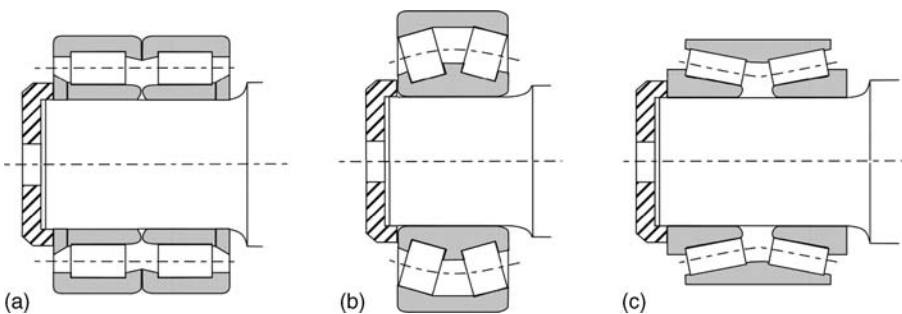
The main factor that determines the construction of the axlebox is the way it experiences the axial forces and distributes the load between the rollers.

Cylindrical roller bearings have high dynamic capacity in the radial direction, but do not transmit axial forces (Figure 3.7a). Experience in operation of railway rolling stock showed that the faces of rollers can resist lateral forces. However, to do this successfully it is necessary to regulate not only the diameter, but also the length of rollers, and the radial, and axial clearances.

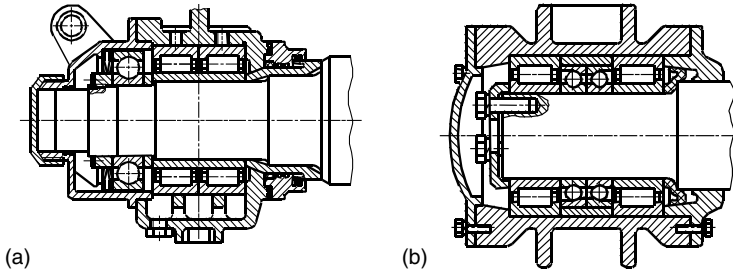
Conical bearings (Figure 3.7b and c) transmit axial forces through the cylindrical surface due to its inclination to the rotation axis. This makes it necessary to keep the tolerances on roller diameters and clearances almost an order of magnitude tighter than for cylindrical bearings. In addition, conical bearings have higher friction coefficients compared to the radial roller bearings and therefore generate more heat. This not only increases traction consumption, but also creates difficulties for diagnostics of axlebox units during motion.

Recently cartridge-type bearings have been widely used. Their special feature is that the bearing is not disassembled for fitting, but is installed as one piece.

Spherical bearings have not been widely applied due to their high cost and lower weight capacity, although they have a significant advantage providing better distribution of load between



**FIGURE 3.7** Constructions of roller bearings: (a) cylindrical double-row; (b) one-row self-alignment; (c) two-row conical.



**FIGURE 3.8** Use of spherical bearings: (a) triple bearing of Japanese high-speed trains; (b) triple bearing of French high-speed trains.

the front and rear rows in case of axle bending. Ball bearings are, however, often combined with cylindrical bearings in railway applications to transmit axial forces.

High speed rolling stock often has three bearings in the axlebox: two transmitting radial forces and one (often a ball bearing) working axially (Figure 3.8).

### C. WHEELS

Wheels and axles are the most critical parts of the railway rolling stock. Mechanical failure or exceedance of design dimensions can cause derailment. Wheels are classified into solid, tyre, and assembly types as shown in Figure 3.9.

Solid wheels (Figure 3.9a) have three major elements: the tyre, the disc, and the hub, and mainly differ in the shape of the disc.

Tyred wheels (Figure 3.9b) have a tyre fitted to the wheel disc that can be removed and replaced when it reaches its maximum turning limit.

Wheels may have straight, conical, S-shaped, spoked, or corrugated type discs when viewed in cross-section. A straight disc reduces the weight of the construction and can be shaped such that the metal thickness corresponds to the level of local stress. The conical and S-shape discs serve to increase the flexibility of the wheel, therefore reducing the interaction forces between the wheels and the rails. Corrugated discs have better resistance to lateral bending.

The desire of reducing wheel-rail interaction forces by reducing the unsprung mass has led to development of resilient wheels (Figure 3.9c) that incorporate a layer of material with low elasticity modulus (rubber, polyurethane). These help to attenuate the higher frequency forces acting at the wheel-rail interface.

Improved bearing reliability aroused interest in independently rotating wheels which provide significant reductions in unsprung mass due to the elimination of the axle. By decoupling the wheels, the independently rotating wheelset inevitably eliminates the majority of wheelset guidance forces. Such wheelsets have found application either on variable gauge rolling stock providing fast transition from one gauge width to another or on urban rail transport where low floor level is necessary.

### D. SUSPENSION

The suspension is the set of elastic elements, dampers and associated components which connect wheelsets to the car body.

If the bogie has a rigid frame, the suspension usually consists of two stages: primary suspension connecting the wheelsets to the bogie frame and secondary suspension between the bogie frame and the bolster or car body. Such bogies are termed double suspended. Sometimes, typically in freight

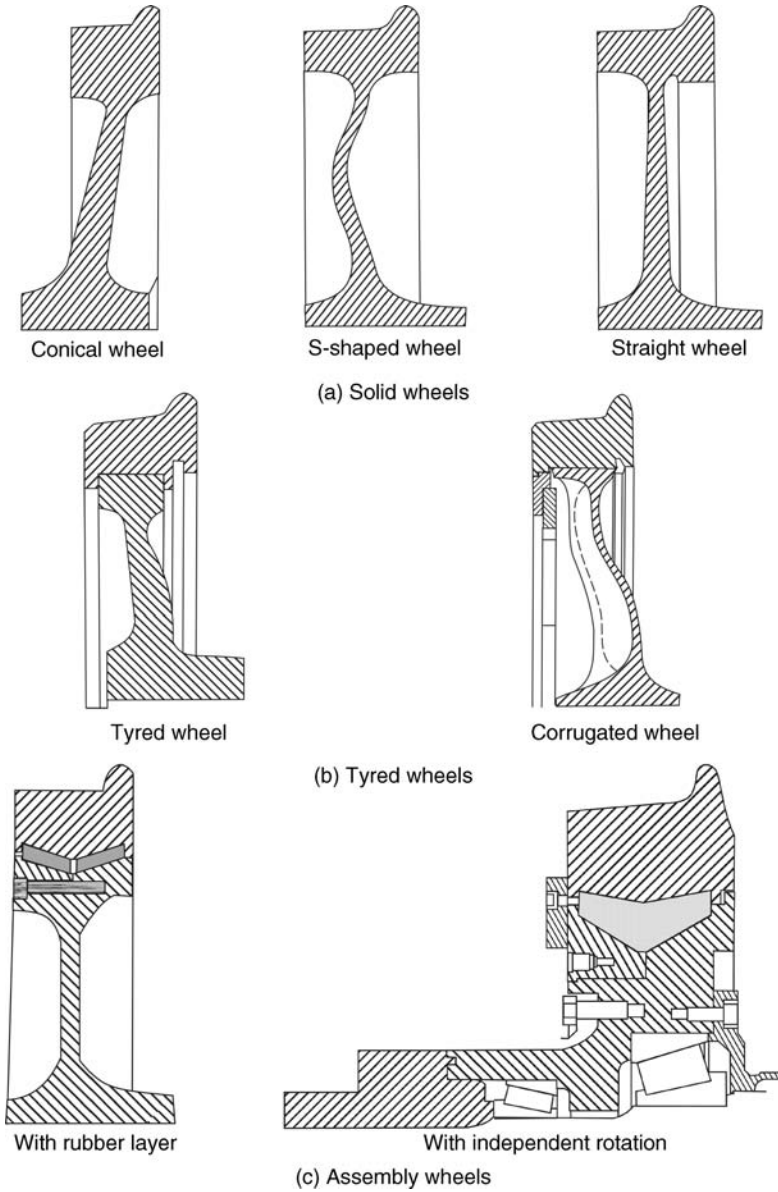


FIGURE 3.9 Major types of railway wheels.

bogies, only a single-stage suspension is used. Where this occupies the primary suspension position it is often termed “axlebox suspension.” In the secondary suspension position it may be termed “central suspension.”

### E. ELASTIC ELEMENTS (SPRINGS)

Elastic elements (springs) are components which return to their original dimensions when forces causing them to deflect are removed. Elastic elements are used to:

- Equalise the vertical loads between wheels (unloading of any wheel is dangerous because it causes a reduction/loss of guidance forces)

- Stabilise the motion of vehicles on track (self-excited lateral oscillations, i.e., hunting of wheelsets is dangerous)
- Reduce the dynamic forces and accelerations due to track irregularities

The capability of elastic elements to provide the above functions is determined by their force characteristic, which is the dependence between the force acting on the elastic element,  $P$ , and its deflection  $z$ :  $P = P(z)$ . Force characteristics can be linear or non-linear. In linear characteristics, the deflection is proportional to the force. For non-linear characteristics the deflection rate increases (or less often for railway applications, decreases) with increase of the load. The principal types of elastic elements are shown below in [Table 3.1](#).

A leaf spring (picture A in [Table 3.1](#)) is an elastic element comprising a number of steel leaves. Leaves work in bending and the “fish-bellied” shape of the beam provides smaller spring stiffness. Depending on their design, leaf springs can be closed (picture A in [Table 3.1](#)), elliptical, or open. They consist of layered leaves 1 and 2 having different length and held together by a buckle 3. The largest leaf (1) is named the master and the other leaves (2) the slaves. Leaf springs also provide damping due to the inter-leaf friction. However, it is difficult to obtain the specific desired damping values and the damping can change considerably due to lubrication or contamination of the rubbing surfaces.

A plate spring (or washer) (picture B in [Table 3.1](#)) consists of a set of elastic steel plates having the conical shape with inclination angle  $\beta$ . Under the load  $P$  the plates flatten and decrease the angle  $\beta$ , thus providing the spring’s deflection. The stiffness of the plate spring depends on the number of plates and their relative arrangement (in series or in parallel).

A ring spring (picture C in [Table 3.1](#)) consists of external and internal steel rings that rest on each other with conical surfaces. Under the load  $P$  the external rings stretch and the internal rings shrink in radial direction, thus providing the vertical deflection of the spring. Deformation causes significant friction forces between the rings.

Coil springs are the most commonly used elastic elements which can either be cylindrical (picture D in [Table 3.1](#)) or conical. Usually they are produced of steel spring wire typically of circular cross-section. Coil springs are cheap and robust, but provide very little damping in suspension applications.

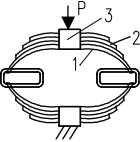
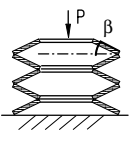
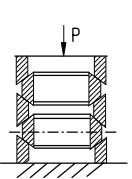
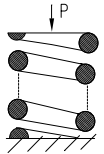
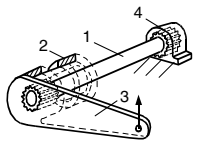
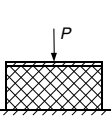
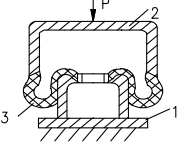
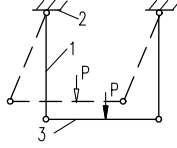
Torsion springs (picture E in [Table 3.1](#)) consist of the torsion bar 1, having its first end in the bearing 2 connected to the arm 3, the second end being fixed in the mounting 4. The force  $P$  causes elastic torsion of rod 1. The most common application of this type of spring in railway vehicles is the roll bar.

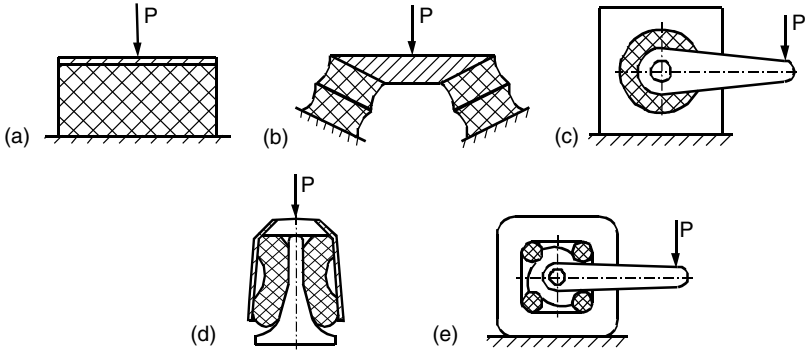
Rubber-metal springs (picture F in [Table 3.1](#)) consist of the rubber blocks 1 interleaved with or reinforced by steel plates 2. This type of spring is widely used in passenger rolling stock, particularly on primary suspensions as it allows damping of high frequency vibrations and reduction of maintenance costs due to the elimination of wearing friction components. Some types of rubber-metal springs are illustrated in [Figure 3.10](#). The elastic properties of rubber can be exploited to make springs that can carry significant loads in both compression and shear ([Figure 3.10b, d, and e](#)).

Air spring (picture G in [Table 3.1](#)) consists of the mounting 1, and rubber-cord elastic chamber 3 filled with compressed gas (usually air). This type of elastic elements is characterised by its small mass, excellent noise and vibration isolation and ability to maintain a constant ride height for different vehicle load conditions. Such springs are found almost universally in the secondary suspension of modern passenger vehicles. Air springs are often arranged in series with a rubber or rubber-interleaved spring to provide some compliance in the suspension if the airspring becomes deflated.

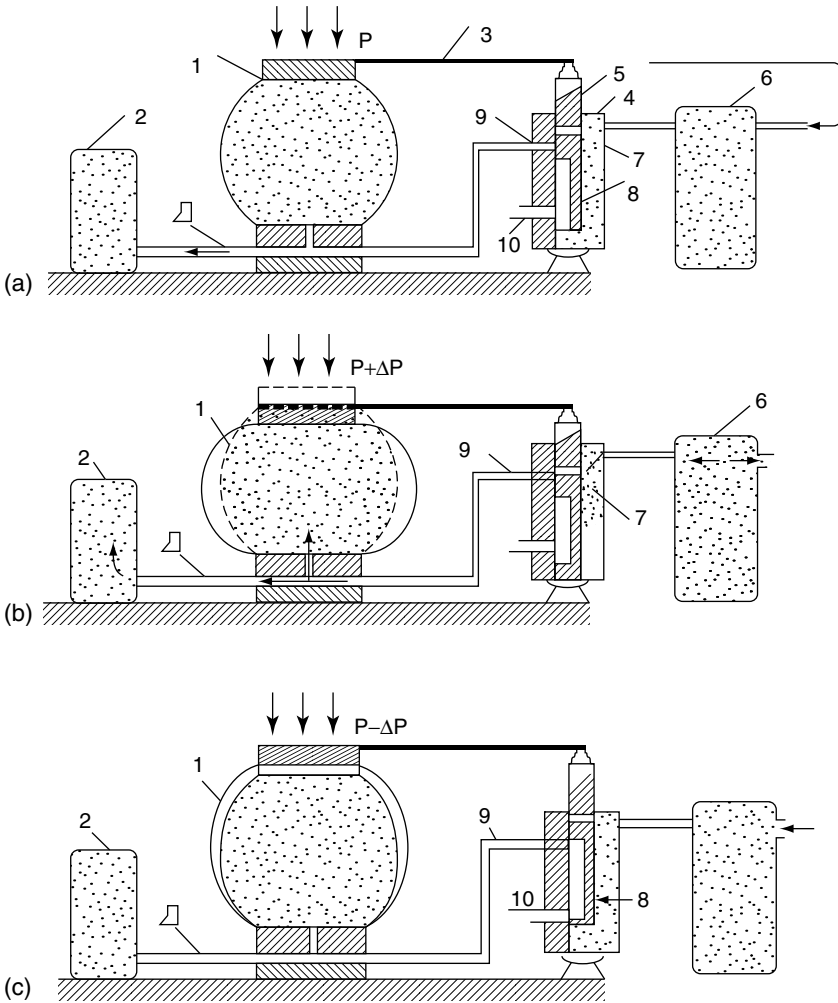
The operation of a typical air suspension with pressure control to maintain constant ride height is shown in schematic form in [Figure 3.11](#).

**TABLE 3.1**  
**Principal Types of Elastic Elements**

Features	Schematic of Elastic Elements							
	A	B	C	D	E	F	G	H
								
Application	Attenuation of vibrations and impacts in one direction	Multidirectional attenuation of vibrations	Multidirectional attenuation of vibrations	Multidirectional attenuation of vibrations and impacts	Attenuation of vibrations and impacts in one direction	Multidirectional attenuation of vibrations and impacts	Multidirectional attenuation of vibrations and impacts, various characteristics are possible	Attenuation of vibrations and impacts in one direction
Advantages	Stiffness of elastic and damping properties in a single component	Good vibration and noise insulation	Good vibration and noise insulation	Easy production and maintenance, small mass and dimensions	Possible to obtain large deflections, small mass and dimensions	Damping of high frequency vibrations, stiffness and damping in a single component	Ability to vary stiffness and maintain constant ride height, good noise and vibration insulation	Easy to maintain and repair
Disadvantages	Unpredictable damping	Complicated control of parameters, production, and maintenance	Complicated control of parameters, production, and maintenance	Inability to vary characteristics in operation	Complicated production, vertical displacements cause longitudinal displacements	Ageing of rubber, inability to vary characteristics in operation	Complicated air supply system and maintenance	Wear, large size



**FIGURE 3.10** Rubber-metal springs: (a) compression; (b) compression and shear; (c) torsion; (d) bell type; (e) cam type.



**FIGURE 3.11** Schematic showing the operation of a typical air suspension: (a) Equilibrium position; (b) Upstroke; (c) Downstroke.



In position (a) above, the system is in static equilibrium when the pressure inside the elastic chamber (airbag) 1 provides the prescribed ride height  $P$ . To reduce the spring stiffness the elastic chamber is connected to the surge reservoir (additional volume) 2. When the load increases (position (b)) the airbag 1 is compressed and moves the valve 5 of control system 4 down. This causes the compressed air from the main reservoir 6 to be admitted to the airspring system through the pipe 9 and orifice 7, thus increasing the pressure. This restores the spring 1 to the equilibrium position (a) again and control valve 4 stops the flow of air from the main chamber 6 into the airbag. Reduction of the load (position (c)) makes the airbag rise and control valve 5 moves up. In this case pipe 9 connects to the atmosphere 10 through orifice 8 and drops the pressure in airbag 1. The spring height reduces and returns to the equilibrium position again.

The surge reservoir 2 and the damping orifice 7 are important features in the operation of the airspring. Increasing the surge reservoir volume leads to decreasing spring stiffness. Reducing the size of the damping orifice increases the damping properties of the spring (by increasing the kinetic energy dissipation), but also increases the stiffness. The lateral stiffness of the pneumatic spring depends on the shape of elastic chamber.

Railway vehicles often use devices whose stiffness is derived from gravitational forces, as for example in the swing link arrangement shown in picture H in Table 3.1. Rollers on inclined planes and various lever systems have also found applications in vehicle suspension arrangements. The swing link suspension is the most common application of those listed above. It consists of swing links 1 that are attached to mountings 2 and connected with a beam or spring plank 3. Swing link suspension effectively acts as a pendulum and is often used in secondary suspensions to give constant lateral frequency.

Typical force characteristics of elastic elements are shown in Table 3.2.

A piecewise linear characteristic (Table 3.2) is typical for coil springs arranged with clearance between springs working in parallel (picture B) or for springs with initial compression (picture C). Parabolic characteristics (picture D) are obtained from coil springs with variable step or wire diameter, conical springs, rubber, or pneumatic springs without pressure control system. An S-shaped force versus deflection characteristic is typical for combinations of elements with a jump between two working modes (picture E). An automatically controlled parabolic characteristic may be obtained from airsprings with automatic pressure control that is dependent on the vehicle loading (picture F). These may be combined with an elastic bump stop acting in compression.

The advantage of coil springs, rubber and pneumatic springs is that they are flexible not only in vertical, but also in longitudinal and lateral direction.

## F. DAMPERS

Damping is usually provided in railway vehicle suspension by the use of viscous or friction damping devices.

Dry friction results from the relative slip between two rigid bodies in contact. The friction force can be constant or dependent on the mass of the car body, but always acts to resist the relative motion. Friction force is proportional to friction coefficient  $\mu$ , pressure between surfaces  $Q$ , and contact surface area  $S$ . This dependence can be represented by the following formula:

$$F_{\text{dry fric}} = -\mu SQ \frac{\dot{z}}{|\dot{z}|} = -F_0 \frac{\dot{z}}{|\dot{z}|} \quad (3.2)$$

where  $F_0$  is the magnitude of friction force;  $\dot{z}$  is the relative velocity of motion;  $|\dot{z}|$  is the magnitude of velocity. The minus sign denotes that the friction force is always in the opposite direction to the velocity.

Viscous damping develops between two parts separated with a layer of viscous liquid (lubricant) or in devices known as hydraulic dampers, where the viscous liquid flows through an

**TABLE 3.2**  
**Typical Force Characteristics of Elastic Elements**

Number	Designation of the Element	Scheme of the Element	Designation of Force Characteristic	View of Force vs. Deflection Characteristic
A	Cylindrical springs with equal height		Linear	
B	Cylindrical springs with different height		Bilinear	
C	Cylindrical springs with different height and initial compression of internal one		Trilinear with a jump	
D	Conical springs, rubber springs		Parabolic	
E	Elements with non-equilibrium characteristics		S-shaped	
F	Controlled pneumatic suspension		Controlled	

orifice and dissipates the energy. The damping force in viscous case is proportional to velocity:

$$F_{\text{hydr fric}} = -\beta z^n \tag{3.3}$$

where  $\beta$  is the coefficient;  $z$  is the velocity of relative motion;  $n$  is the power. Depending on the construction of the device and the liquid properties the power  $n$  can be greater, equal or less than 1.

If the liquid flow is laminar then  $n \approx 1$  and damping is described linear viscous damping:

$$F_{\text{lin visc fric}} = -\beta_1 \dot{z} \tag{3.4}$$

where  $\beta_1$  is the coefficient, named the damping coefficient for the hydraulic damper.

For  $n = 2$  damping is called turbulent or quadratic:

$$F_{\text{turb visc fric}} = -\beta_2 |\dot{z}| \dot{z} \tag{3.5}$$

Gases are also viscous. Therefore, driving the gas through a throttle valve (damper orifice) may also produce sufficient force for damping the oscillations of railway vehicles.

Intermolecular damping (hysteresis) originates mainly in rubber and polyurethane elastic elements. In such cases, the damping force is proportional to oscillations velocity and is inverse to the frequency:

$$F_{\text{molec fric}} = -\frac{\beta_0}{\omega} \dot{z} \tag{3.6}$$

Damping of vibrations can also be obtained by other means such as the introduction of active dampers being controlled proportionally to velocity.

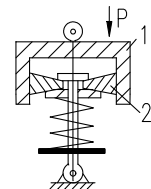
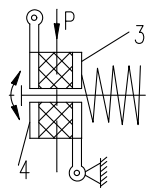
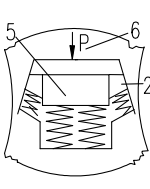
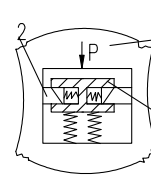
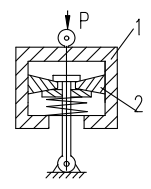
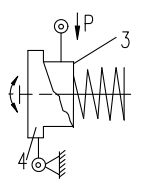
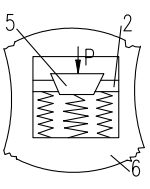
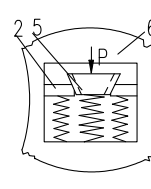
A damper is the device that controls oscillations in the primary or secondary suspension of the vehicle by energy dissipation.

Friction dampers are the devices that transform the energy of oscillations into the heat energy by dry friction. Friction dampers are mainly used in freight vehicle suspensions due to their low cost and simplicity.

Depending on their construction friction dampers may be classified as one of four types: integrated with the elastic element, integrated into the spring suspension, telescopic, and lever (Table 3.3).

Dampers integrated with an elastic element consist of the barrel 1 and friction wedges 2 that are held in contact by a spring. When the elastic element deforms, the friction forces act on the contacting surfaces between the barrel 1 and the wedges 2 transforming the kinetic energy into the heat.

**TABLE 3.3**  
**Classification of Friction Dampers**

	Linear Action		Planar Action Integrated in the Suspension	Spatial Action
	Telescopic	Lever		
Constant friction				
Variable friction				

Telescopic friction dampers consist of the body 1 that contains the piston with the system of friction wedges 2 clamped by a spring.

Dampers integrated in the suspension are mostly used in three-piece bogies and consist of friction wedges 2 that move relative to side frame 6 and bolster 5. Construction of the dampers (Table 3.3) differs by the position of friction wedges 2 (inside the bolster 5 or inside the frame 6), by the number of springs and their inclination angles, as well as by the design of the friction wedges. For example, the Russian CNII-H3 bogie has wedges 2 with inclined faces contacting with the bolster 5 and pressed to the side frame 6 by springs underneath.

Simultaneous and integrated friction dampers are connected to the springs in the suspension, whereas telescopic dampers are independent devices. Friction dampers may be arranged to produce either constant or variable friction force and can be designed to act in one (linear), two (planar), or three (spatial) directions.

Friction dampers integrated in elastic elements have found wide application in freight bogies in Russia, the USA, and many other countries, due to the following advantages: simplicity of design and fabrication, low cost, and easy maintenance. Disadvantages of such dampers include suboptimal damping in the partially laden condition, the difficulties of controlling friction to the desired design values and changes in friction levels as the faces wear or become contaminated in service.

Telescopic dampers have the advantage of being autonomous, protected from the environment (which reduces the likelihood of contamination of the friction surfaces), can be installed at angles other than vertically and hence can be used to damp vertical or horizontal vibrations of sprung elements of the vehicle. They can be inspected and repaired without lifting the car body. One of the reasons that such telescopic dampers are not widely used in freight vehicles using the popular three-piece bogie is that an integrated friction wedge as shown above is required to resist warping in vertical and horizontal planes.

In case of the bogies with a solid frame, friction dampers in the primary suspension must resist wheelset displacements. It is desirable that in primary suspensions the damper has an asymmetric characteristic providing lower damping forces in compression than extension. Hydraulic dampers are superior in this respect.

The main advantage of plane and spatial friction dampers is their ability to damp vibrations in several directions and in certain cases provide friction–elastic connections between parts of bogie frames. Such properties allow significant simplification of the bogies whilst retaining reasonable damping of complex vibrations. They are therefore widely used in freight bogies despite a number of disadvantages including providing unpredictable friction forces, and the fact that repair and adjustment of friction forces may require lifting the car body and disassembling the spring set.

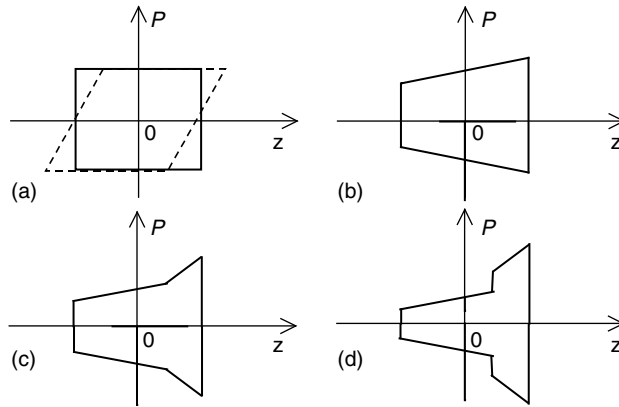
Typical force characteristics for friction dampers are presented in Figure 3.12. Different designs of friction damper have varying arrangements for transmitting the normal force to the friction surfaces. Depending on the design, the damper may provide constant or variable friction. In the latter case, such damper is usually arranged such that a component of the force in one or more of the suspension springs is transmitted via a linkage or wedge to the friction faces.

Force characteristic (a) describes a constant friction damper, where the friction force does not depend on deformation of the spring set and is the same for compression and tension. The dashed line shows the characteristic of the same damper, where the friction pairs are elastically coupled. This can occur, for example, as a result of the friction surface having an elastic layer underneath. Force  $P$  first deforms the elastic pad and when the shear force equals the friction breakout force (i.e.,  $\mu N$ ), then relative displacement of friction pair occurs.

Characteristic (b) is common for most friction dampers used on freight bogies. The friction force depends on the deflection of the suspension, and is different for tension and compression.

Characteristics (c) and (d) are typical for multi-mode dampers, where the friction forces vary according to the given law and depend on the spring set deflection in tension or compression.

It can be seen that the variety of force characteristics available from friction dampers allows freight vehicle to be designed with suspensions providing satisfactory ride qualities.



**FIGURE 3.12** Typical force characteristics of friction dampers.

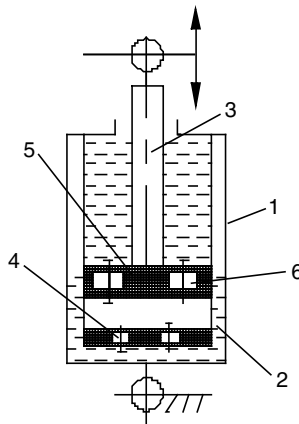
Hydraulic dampers are almost universally used in passenger bogies and are sometimes also used in modern freight bogies.

The energy dissipated in a hydraulic damper is proportional to velocity, and therefore to the amplitude and frequency of vibration. Thus the hydraulic damper is self-tuning to dynamic excitations and provides reliable and predictable damping of vehicle oscillations.

Railway vehicles use the telescopic hydraulic dampers as shown in Figure 3.13. The hydraulic damper operates by forcing the working fluid through an orifice (flow control valve) from one chamber into the other as the vehicle oscillates on the suspension. This produces viscous damping and the kinetic energy of the oscillations is transformed into heat.

Telescopic hydraulic dampers (Figure 3.13) consist of the body 1 with the sealing device, the working cylinder 2 with valves 4 and the shaft 3 with a piston 5 that also has valves 6. When the piston moves relative to the cylinder, the working fluid flows through the valves from the chamber over the piston to the chamber under it and back.

The reliability of hydraulic dampers mostly depends on the sealing between the shaft and the body. Occasionally malfunction of this unit causes excessive pressure in the chamber over the piston resulting in leakage of the working fluid. The capability of hydraulic damper to dissipate



**FIGURE 3.13** Telescopic hydraulic damper.

energy is characterised by its force versus velocity characteristic, which is the dependence between the resistance force developed in the hydraulic damper,  $P$ , and the piston displacement velocity  $\dot{d}$ .

The damper characteristic may be either symmetrical, when the resistance forces are the same for extension and compression, or asymmetric. Dampers with symmetric characteristics are typically used in secondary suspensions. In primary suspensions, asymmetric dampers are often used as the motion of the wheel over a convex irregularity causes larger forces than negotiating a concave one. As a result, dampers may be designed with an asymmetric characteristic providing a smaller force in compression than in extension. However, large damping forces in extension can significantly decrease the vertical wheel load, thus increasing the risk of derailment. Therefore the railway dampers are less asymmetric than the automobile ones.

Common force characteristics of hydraulic dampers are shown in Figure 3.14.

Figure 3.14A shows a hydraulic damper with a resistance force proportional to velocity and not exceeding the “blow-off” (saturation) force. When a predetermined pressure value is reached inside the working chamber, the “blow-off” valve opens to prevent excessive forces being developed by the damper.

Hydraulic dampers having characteristic B have a resistance force proportional to the velocity and the displacement. Such characteristic is obtained by the provision of specially calibrated needles (or other devices) into the flow control valve to change its cross-section. The size of the valve cross-section is controlled depending on static deflection of the suspension.

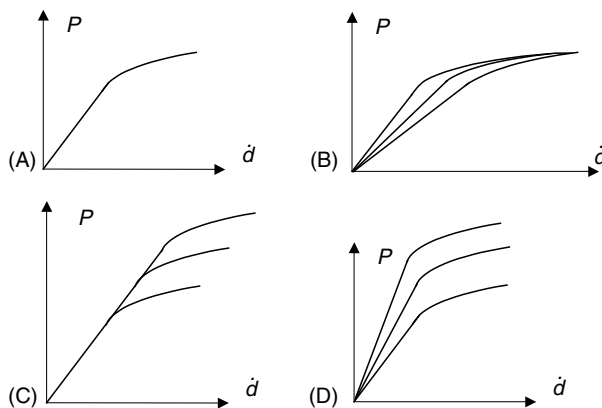
Scheme C is typical for devices where the dissipative force is proportional to velocity, but the operation of “blow-off” valve is controlled depending on the displacement and velocity of the piston. In scheme D, the size of the valve cross-section and the saturation limit for emergency valve are controlled together depending on the relative displacement and velocity of the piston.

Attachment of hydraulic dampers to the vehicle is usually done using the elastic mountings or bushes to prevent the transmission of high frequency vibrations. The internal pressure in the damper often gives it elastic properties. Therefore, hydraulic dampers are often modelled as a spring and viscous damper in series.

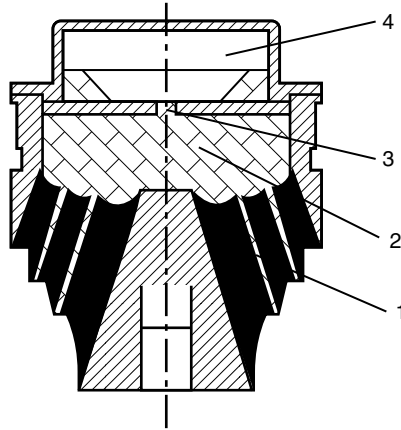
In some designs, the hydraulic dampers are united with the elastic elements. The schematic of a hydraulic damper integrated into coaxial rubber-metal spring is shown in Figure 3.15.

**G. CONSTRAINTS AND BUMPSTOPS**

Constraints are the devices that limit the relative displacements of bogie units in longitudinal and lateral directions.



**FIGURE 3.14** Common force characteristics of hydraulic dampers.



**FIGURE 3.15** Hydraulic spring: (1) rubber-metal conical spring; (2) working fluid; (3) flow control valve; (4) compensation reservoir with rubber diaphragms.

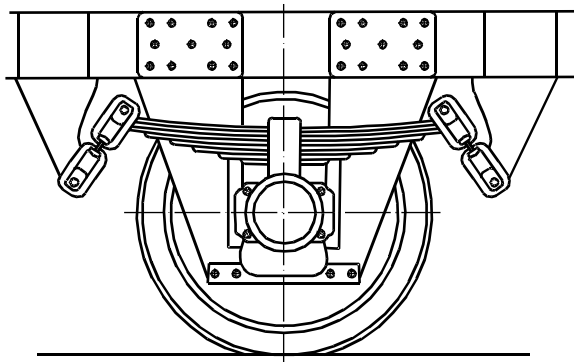
### 1. Horn Guides

A simple primary suspension design uses horn guides to limit the movement of the axlebox (Figure 3.16).

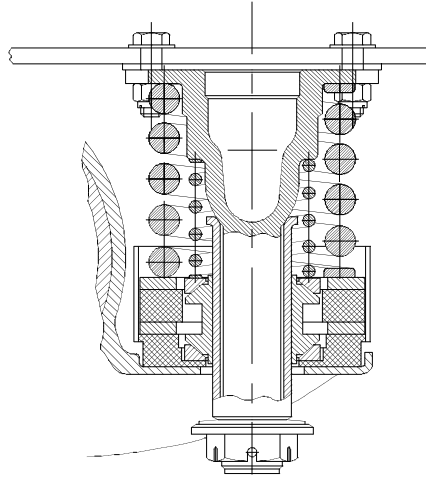
This design has several disadvantages, including fast wear of friction surfaces leading to the increases in clearances, lack of elastic longitudinal and lateral characteristics, and increased friction force in vertical direction in traction and braking modes, when the axlebox is pressed against the slides. The design could be improved by the application of anti-friction materials that do not require lubrication and have high resistance to wear.

### 2. Cylindrical Guides

These comprise two vertical guides and two barrels sliding along them. Typically the vertical guides are attached to the bogie frame and the barrels to the axlebox as shown in Figure 3.17. The barrels are attached to the axlebox through rubber coaxial bushings and therefore provide some flexibility between the wheelset and the bogie frame in the longitudinal and lateral directions. Due to axial symmetry of the rubber bushes, the stiffness in longitudinal and lateral directions is the same, which may limit the provision of optimal suspension characteristics.

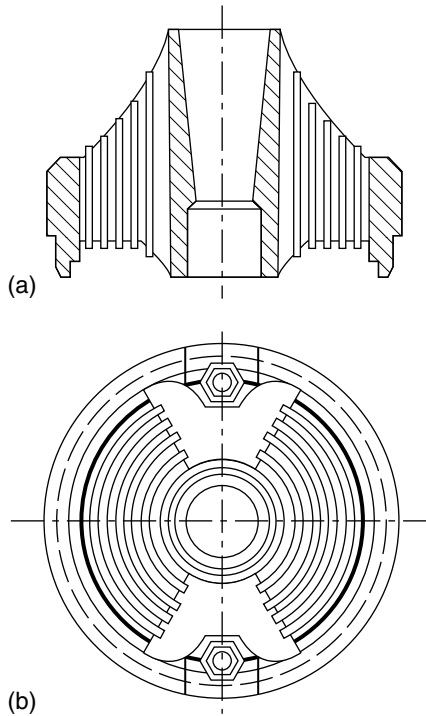


**FIGURE 3.16** Axlebox located by horn guides.



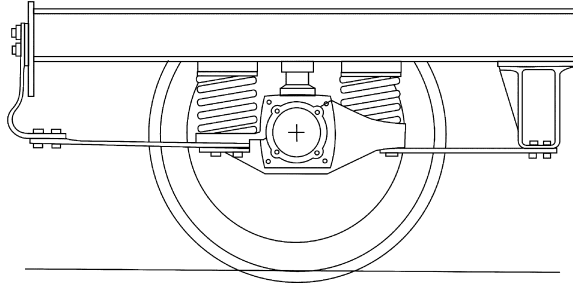
**FIGURE 3.17** Connection between the axlebox and bogie frame using cylindrical guides.

Axlebox constraint with cylindrical guides, where the displacement of the axlebox along the guides occurs by shear deformation of multi-layer rubber-metal block is free from disadvantages of classical construction. Such axlebox designs are used on French TGV Y2-30 bogies. In order to obtain the optimum relationship of horizontal and vertical stiffness this block consists of two longitudinally oriented sections (Figure 3.18).



**FIGURE 3.18** Two-section rubber-metal block used to connect the axlebox and bogie frame.





**FIGURE 3.19** Connection between the axlebox and bogie frame using beam links.

### 3. Beam Links

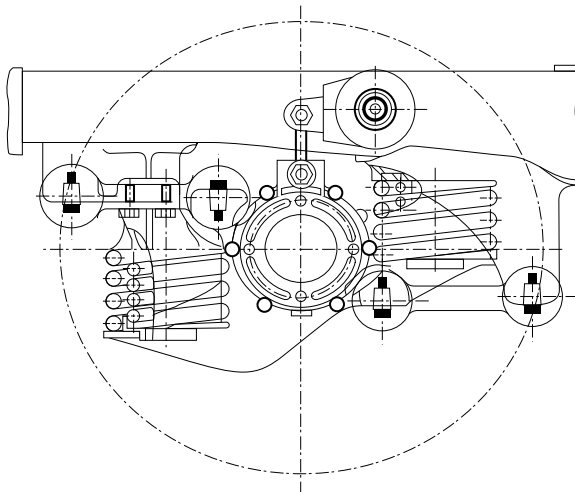
The desire to avoid wear led to the development of links in the form of thin elastic beams that hold the wheelset in the longitudinal direction (Figure 3.19).

When primary suspension springs deflect, the beam links bend, whereas for traction and braking they experience tension or compression. To provide vertical flexibility in such construction it is necessary for at least one of the links to have longitudinal flexibility. This is achieved by attaching the beam to a longitudinally flexible spring support (the Minden Deutz link) or by attaching the links to the frame through radially elastic joints (IS primary suspension of Japanese trains).

The main disadvantage of such designs is high stress which develops around the joints at either end of the beam.

### 4. Constraints Using Radius Links

The use of rubber-metal bushes avoids surface friction and corresponding wear. The main problem with a radius link arrangement is obtaining linear motion of axleboxes when the links rotate. Alstom designed such an arrangement where the links are positioned on different levels in anti-parallelogram configuration (Figure 3.20) and this has found wide application. Links that connect



**FIGURE 3.20** Radius links positioned at different heights in an anti-parallelogram configuration.

the axlebox to the frame provide linear displacement of its centre. By careful choice of size and the material of the rubber elements it is possible to obtain the required stiffness values in different directions. Due to the position of the links, lateral displacements do not cause misalignment of the axlebox therefore providing optimum conditions for the bearings.

Disadvantages of the radius link design include the significant vertical stiffness of the connection due to torsion stiffness of the bushes. Increasing the length of the levers would decrease the vertical stiffness, but it is limited by the space available in the bogie frame.

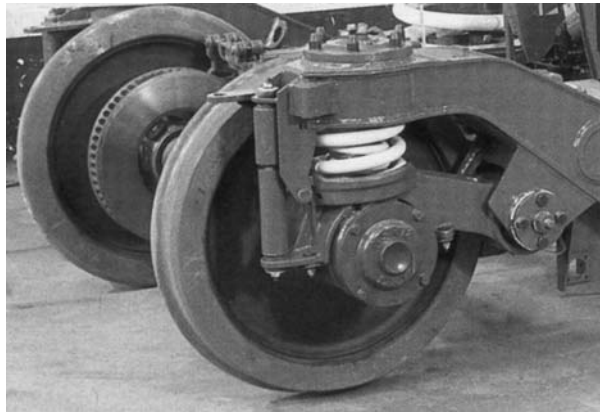
### 5. Constraints Using Trailing (Radial) Arms

Trailing arm suspensions allow the design of shorter and lighter bogie frames. Such designs are now widely used in passenger vehicle primary suspensions, such as the Y32 bogie shown in Figure 3.21.

The disadvantages of such designs include the longitudinal displacement of the axleboxes caused by vertical displacement of the suspension and torque applied to bogie frame due to wheelset lateral displacement.

### 6. Traction Rods

These are normally used to transmit longitudinal (traction and braking) forces in either the primary or secondary suspension. They are typically comprised of a rod with a rubber “doughnut” or bushes at each end. They may be adjustable length to maintain the necessary linear dimensions as wheels or suspension components wear (Figure 3.22).



**FIGURE 3.21** Trailing arm suspension on a Y32 bogie.



**FIGURE 3.22** Traction rod.

## F. CAR BODY TO BOGIE CONNECTION

The connection between the car body and bogie must:

- Allow the bogie to turn relative to the car body in curves
- Transmit the vertical, traction and braking forces
- Provide additional control of lateral suspension inputs
- Assist in maintaining the stability of the bogie
- Provide longitudinal stability of bogie frames and equal distribution of load over the wheelsets (for traction rolling stock)

These problems are solved differently depending on the type of the rolling stock — traction or trailing, passenger or freight, moderate or high speed.

If the vehicle is stable up to the design speed, then introduction of additional yaw resistance torque is not necessary. If the static deflection of the suspension is sufficient, then vertical flexibility in the car body to bogie connection may not be necessary.

Designs generally aim to make the bogie to car body connection as simple as possible by the use of a small number of elements and reduction of the number of elements with surface friction.

### 1. Flat Centre Plate

In three-piece freight bogies the most common connection is the flat circular centre plate, that is secured by pin pivot at the centre (Figure 3.23).

The plate transmits the majority of the car body weight and the longitudinal and lateral interaction forces. The pin pivot has large in-plane gaps to the car body and only provides emergency restraint. When the car body rocks on the flat centre plate a gravitation resistance torque having soft characteristic is produced. The centre plate allows the bogie to rotate in curves and creates a friction torque that resists bogie rotation. Hence the circular centre plate provides a connection between the bogie and the car body in all directions.

Such a unit is of simple construction, but has several disadvantages. Firstly, clearances exist in the lateral and longitudinal directions. Secondly, relative motion occurs under high contact pressure and hence the surfaces are subject to significant wear. In curves, the car body leans on the side bearer creating additional friction torque that resists bogie rotation and increases wheel–rail forces. When the car body rocks on straight track, the contact surface becomes very small and high contact pressures can lead to cracks in the centre plate. To combat these problems, modern designs use a flat centre plate combined with elastic side bearers which resist car body rock and reduce the load on the centre bowl.

### 2. Spherical Centre Bowl

In this case, the car body rests on the spherical centre bowl and elastic side bearers (Figure 3.24).

The advantage of this design is the lack of clearance in the horizontal plane and no edge contact during car body roll. This results in reduced levels of contact stress and increases the centre bowl

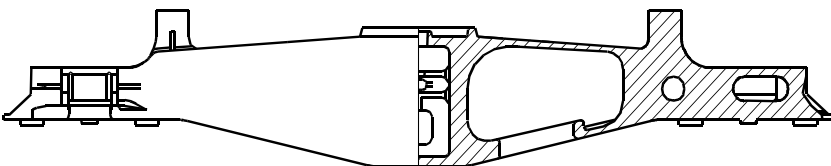


FIGURE 3.23 Flat centre plate.

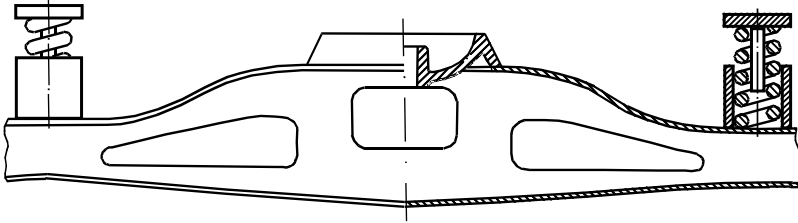


FIGURE 3.24 Spherical centre bowl.

service life. Such centre bowls are widely used in UIC freight bogies, electric trains, and underground cars in Russia.

### 3. Centre Pivot

The desire to exclude edge contact and increase the friction torque to resist bogie yaw led to development of bogies with centre pivots as shown in Figure 3.25. The majority of the car body mass is in this case transmitted to the side bearers and the car body can only turn relative to the bolster about the vertical axis.

This design is widely used in passenger coaches of former USSR. The disadvantages include the clearances in longitudinal and lateral directions. The design provides sufficient ride quality only for bogies having low lateral stiffness of secondary suspension.

### 4. Watts Linkage

This arrangement, illustrated in Figure 3.26, allows the bogie to rotate and move laterally whilst restricting longitudinal movement. It therefore provides a means of transmitting traction and braking forces. Pivots in the linkage are provided with rubber bushes to prevent the transmission of high frequency vibrations through the mechanism.

### 5. Pendulum Linkage

The pendulum linkage consists of a vertical rod connected at each end to the body and bogie frame by conical rubber bushes as shown in Figure 3.27. The mechanism is held in a central position by two precompressed springs. Elastic side supports provide lateral stability to the car body. For the small displacements that are typical of bogie hunting on straight track the pendulum support provides almost infinite stiffness determined by initial compression of springs. When large

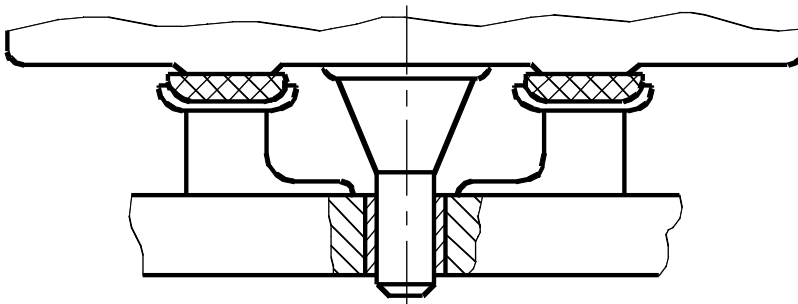


FIGURE 3.25 Centre pivot.

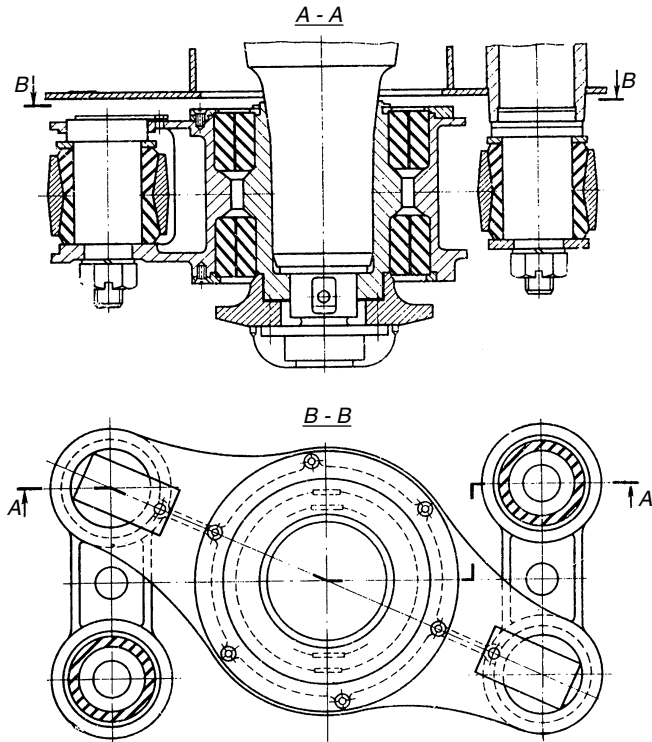


FIGURE 3.26 Watts linkage.

displacements develop in curves, the support provides low stiffness. Thus, the pendulum support has a soft nonlinear characteristic.

The drawback of such an arrangement is the rigid connection with a gap in the longitudinal direction, complex tuning requirements for the precompressed springs and friction forces in the additional sliding supports.

**6. Connection of Car Body to Bolsterless Bogies**

The complexity of the designs described above accounted for the development of modern bolsterless bogies using either flexicoil springs or air springs. In such suspensions the springs can

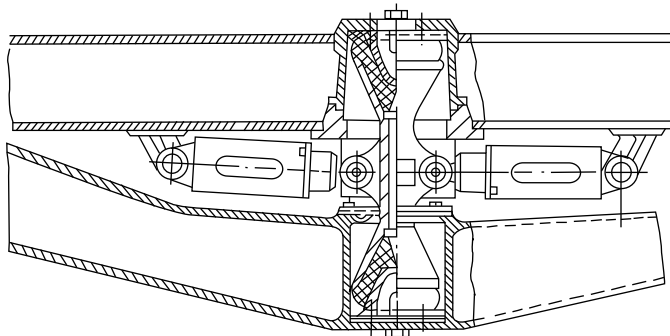


FIGURE 3.27 Pendulum linkage.

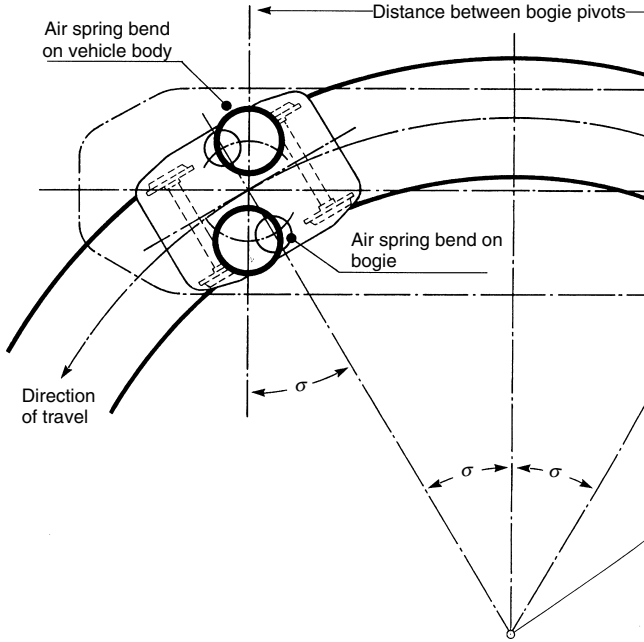


FIGURE 3.28 Schematic showing a bolsterless bogie passing a curve.

achieve large deflections in shear providing sufficiently large longitudinal displacements to allow the bogie to rotate in curves (Figure 3.28).

The top of the flexicoil springs rests on resilient blocks arranged to provide a cylindrical joint with rotation axis perpendicular to the track axis (Figure 3.29).

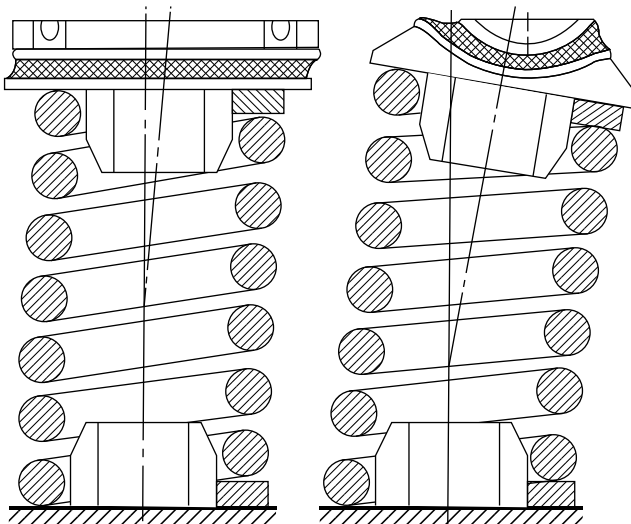


FIGURE 3.29 Spring resting on rubber-metal cylindrical joints.

A similar approach is used in bogies with secondary air suspension. In this case the air spring is often arranged in series with a rubber-metal spring to provide some suspension if the air spring deflates. Transmission of longitudinal forces is done through the centre pivot, Watts linkage, traction rods, or in the case of a Y32 bogie, through the backstay cables. Bolsterless bogie designs typically achieve reductions in bogie mass of around 0.5–1.0 t.

### III. COMMON PASSENGER VEHICLE BOGIE DESIGNS

The most common passenger vehicle designs use a pair of two-axle bogies on each vehicle. However, in articulated trains, for example the French TGV, two-axle bogies are positioned between the car bodies, whilst the Spanish Talgo trains use single-axle articulated bogies.<sup>1</sup>

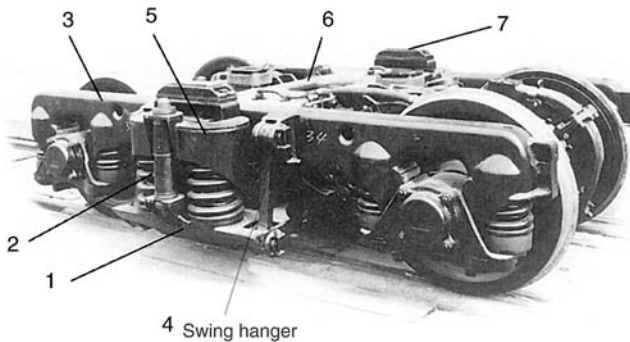
For passenger bogies, the wheelsets are generally mounted in a rigid H-shaped frame that splits the suspension in two stages. The primary suspension transmits forces from the wheelsets to the bogie frame and the secondary suspension transmits forces from the bogie frame to the car body.

The principal functions of the primary suspension are guidance of wheelsets on straight track and in curves, and isolation of the bogie frame from dynamic loads produced by track irregularities. The secondary suspension provides the reduction of dynamic accelerations acting on the car body which determines passenger comfort. The source of these accelerations is excitation from the track irregularity/roughness profile and the natural oscillations of the bogie frame and car body on their suspension elements. It is particularly important to reduce the lateral influences, to which the passengers are more sensitive, and therefore the stiffness of secondary suspension in lateral direction is designed as small as possible.

An example of a traditional type of secondary suspension (used on passenger vehicles for over 100 years) is shown in Figure 3.30. The secondary suspension swing consists of the secondary springs and dampers (2), spring plank (1) that is attached to the bogie frame (3) by swing links (4). This arrangement provides low lateral stiffness, and the height of the secondary springs remains comparatively small.

When curving, the bogie should rotate under the car body to reduce track forces, whereas on straight track it should resist yawing motion. In the case of bogies with swing link secondary suspension, part of the car body mass is transmitted to the bolster (5) through the bogie centre (6) and part through the side bearers (7). The bogie centre serves as the centre of rotation and transmits the traction forces whilst the side bearers provide friction damping to the bogie yaw motion. The traction rod usually limits longitudinal displacements of the bolster relative to the bogie frame.

Swing link secondary suspension may be acceptable for speeds up to 200 km/h. Its disadvantage is the large number of wearing parts that require relatively frequent maintenance to prevent deterioration of ride quality.



**FIGURE 3.30** Bogie with swing link secondary suspension.

Modern bogie designs have a smaller number of parts in the secondary suspension and thus reduce maintenance costs. They typically use elastic elements that have a small stiffness in the horizontal direction. Examples include the ETR-500 bogie (Figure 3.31) which use Flexicoil secondary springs, and the Series E2 Shinkansen (Figure 3.32) which uses an air spring secondary suspension.

In such secondary suspension arrangements, the vehicle body may rest on a bolster (as in the swing link bogie), or directly mount on the secondary suspension, as in the bolsterless bogie in Figure 3.31 and Figure 3.32. In bolsterless bogies the traction forces are transmitted through the centre pivot arrangement, and the bogie rotates under the car body using the flexibility of secondary suspension in longitudinal direction. In such designs, yaw dampers are often fitted longitudinally between the body and bogie to damp hunting motion on straight track.

Modern bogies are normally equipped with separate secondary dampers to damp oscillations in vertical and lateral directions. Lateral damping is normally achieved with a hydraulic damper whilst vertical damping may be hydraulic or orifice damping within the air spring.

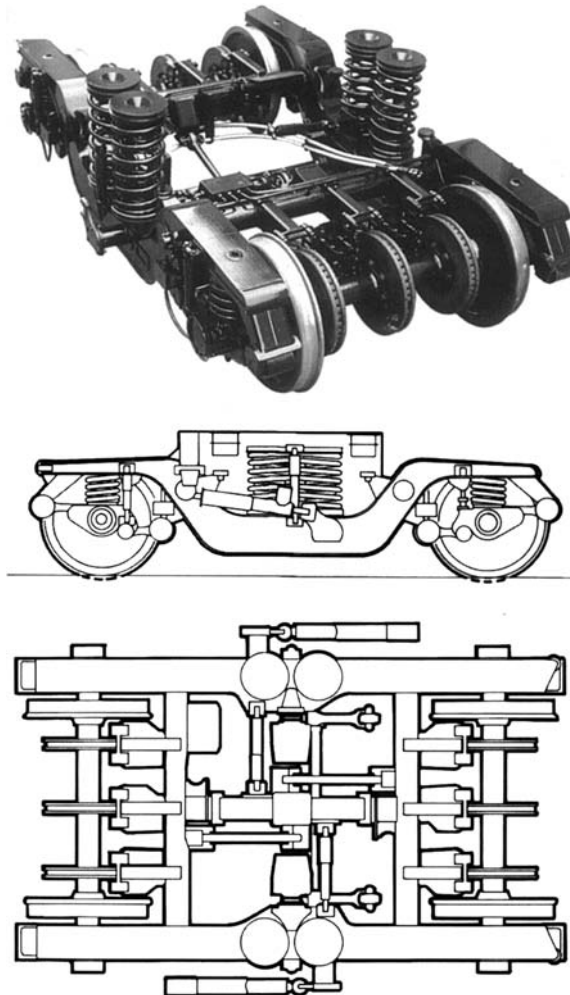
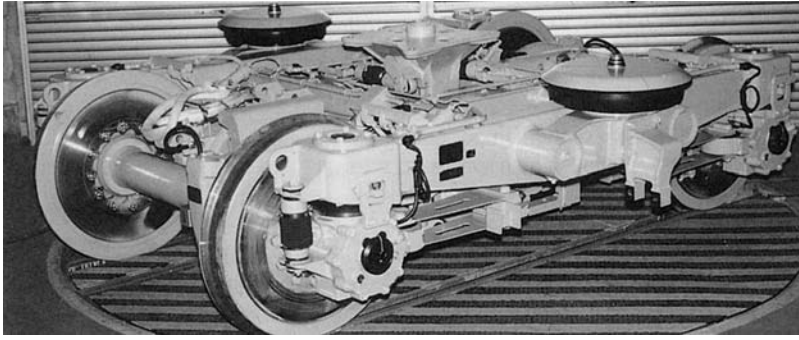


FIGURE 3.31 ETR-500 train bogie (Italy).





**FIGURE 3.32** Series E2 Shinkansen bogie (Japan).

Various types of elastic elements are used in passenger bogie primary suspensions. To achieve high speeds the longitudinal stiffness of primary suspension should be high, whereas the lateral stiffness may be lower. In curves the high longitudinal primary stiffness leads to the increase of contact forces between the wheels and rails causing increased wear. Similarly, high lateral stiffness may lead to increased dynamic force when negotiating lateral track irregularities. For passenger bogies, it is therefore preferable if the suspension design can provide different stiffness in the lateral and longitudinal directions.

The three most common types of primary suspension are those with coil springs and longitudinal traction rods or links (Figure 3.33), coil springs with guide posts (Figure 3.34), and chevron (rubber interleaved) springs (Figure 3.35).

The ETR-460 bogie (Figure 3.33) is an example of a primary suspension using traction links with resilient bushes. The wheelset is guided by two links with spherical joints, and the vertical and lateral loads are mainly reacted by the coil springs.

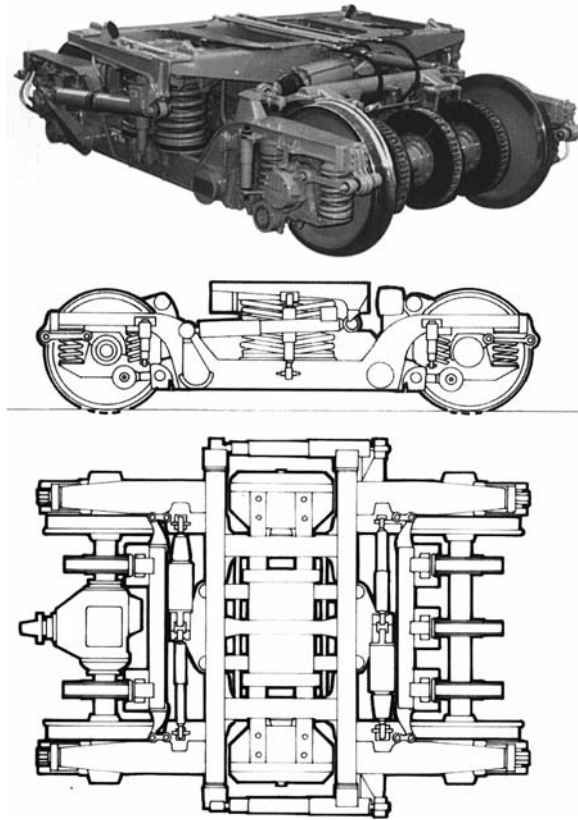
In the primary suspension of Series 300 Shinkansen bogies, coil springs are used together with cylindrical guide posts containing rubber-metal blocks (Figure 3.34). The springs bear the vertical load whilst the rubber-metal block provides different stiffnesses in the longitudinal and lateral directions. It also acts to damp high frequency vibrations.

The X-2000 high speed train bogie primary suspension uses chevron (rubber-interleaved) springs (Figure 3.35). In this type of spring, rubber blocks are separated by steel plates arranged at an inclined position to the vertical. In this way vertical forces on the spring cause both shear and compression forces in the rubber blocks. Depending on the V-angle and material properties of the chevron spring, the longitudinal stiffness can be made three to six times higher than the lateral stiffness. The disadvantage of such design is that the mechanical properties are highly dependent on temperature and this may become a significant factor when operating in climates where extremes of temperature are common.

#### IV. COMMON FREIGHT WAGON BOGIE DESIGNS

In most cases, freight wagons use two two-axle bogies per vehicle. However, some articulated freight vehicles have been designed, principally flat wagons for container transportation.

The majority (Figure 3.36) of freight bogies have single-stage suspension, either between the wheelsets and the bogie frame (similar to passenger bogie primary suspension and often termed “axlebox suspension”), or between the bogie frame and the bolster (similar to passenger bogie secondary suspension and often termed “central” suspension). It can be seen from Figure 3.36 that central suspension make up approximately 6% more of the designs than axlebox suspension. Some



**FIGURE 3.33** ETR-460 bogie (Italy).

wagons use double suspensions similar to passenger bogies to reduce track forces or improve isolation of the load from excess vibrations.

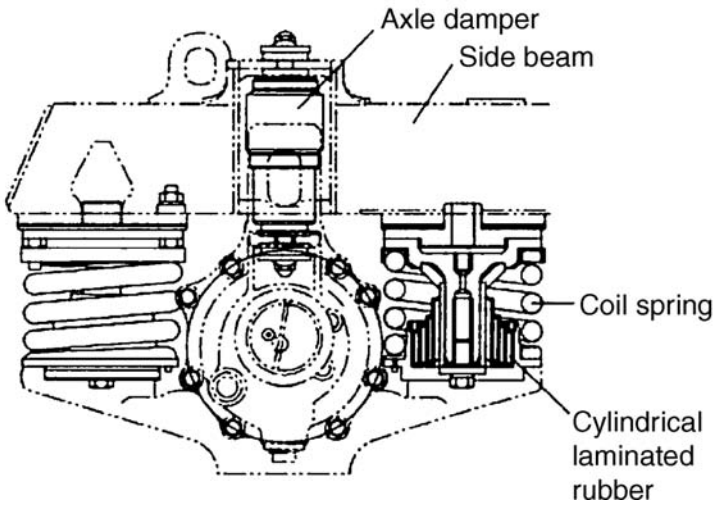
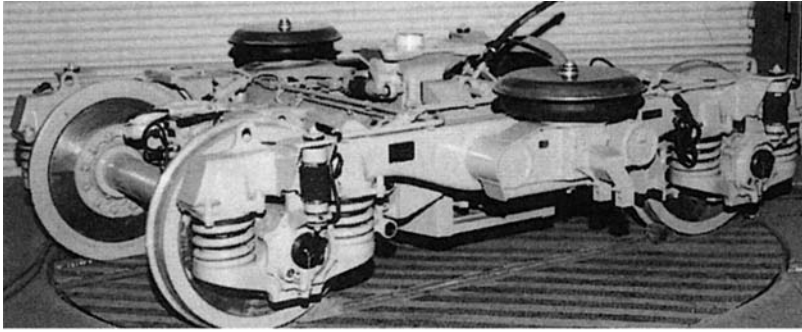
Bogies with central suspension are common in the countries of the former USSR, USA, Canada, China, Australia and most countries in Africa. Examples of the CNII-H3 (type 18-100) Russian bogie and the Barber bogie from USA are shown in [Figure 3.37](#) and [Figure 3.38](#), respectively. Such bogies are often termed “three-piece” bogies.

The frame of a three-piece bogie consists of the bolster and two side frames that are elastically connected by a coil spring and friction wedge-type central suspension system, that beside other functions resists asymmetrical loads and holds the bogie frame square in-plane. Such suspension allows independent pitch of the side frames when negotiating a large vertical irregularity on one rail, allowing the bogie to safely negotiate relatively poor track.

The vehicle body is connected to the bogie with a flat centre bowl and rigid side bearers having clearance in the vertical direction. When moving on straight track, the car body rocks on the centre bowl and does not touch the side bearers, the gravitational force providing recovery to the central position. In curves, the car body contacts the side bearers.

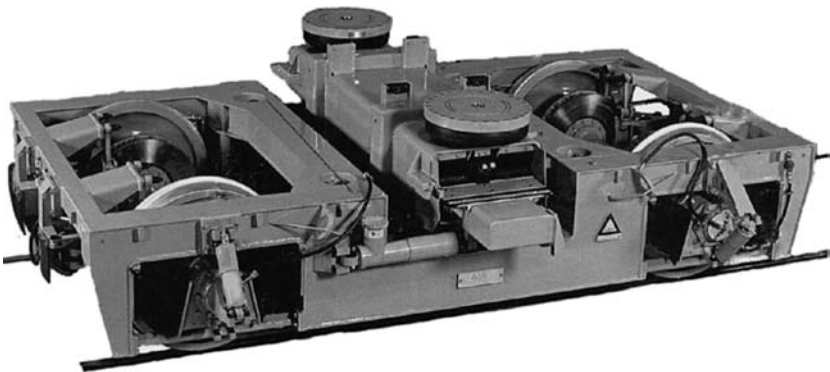
The central suspension consists of a set of nested coil springs and the wedge arrangement that provides friction damping in the vertical and lateral directions. The inclination of the friction wedges may vary between designs: in the 18-100 bogie the angle is  $45^\circ$ , whilst in the Barber bogie it is  $35^\circ$ .

Freight wagons suspensions have to operate under a wide range of load conditions from tare to fully laden, when axle loads can change by more than four times and the load on the spring set more



**FIGURE 3.34** Series 300 Shinkansen bogie (Japan).

than five times. In the 18-100 bogie, the stiffness of the spring set is independent of the load, which leads to poor ride and increased derailment risk due to small deflections of the springs. For the Barber bogie, a range of suspension spring sets are available for axle loads from 7 to 34 t, that include spring sets with bilinear vertical force characteristics.



**FIGURE 3.35** The X-2000 bogie with chevron spring primary suspension (Sweden).

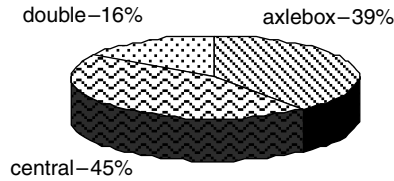


FIGURE 3.36 Proportions of freight bogies by suspension type.

The side frames of a three-piece bogie rest on the wheelsets. In the 18-100 bogie, the bearing is mounted inside the axlebox, whilst the Barber bogie has an adapter between the cylindrical cartridge-type bearing and the flat surface of the side frame. Clearances between the adapter (or the axlebox) and the side frame in the longitudinal and lateral directions allow the wheelsets to move in curves and when passing the large horizontal irregularities. Thus the axlebox unit does not steer the wheelsets, but damps their displacements by friction forces. Due to the absence of primary suspension, such bogies have a large unsprung mass which causes increased track forces.

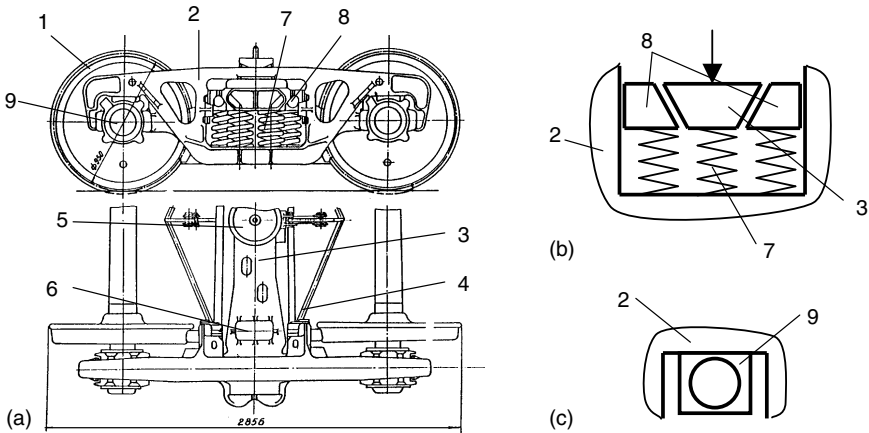


FIGURE 3.37 Type 18-100 bogie: (a) general view; (b) central suspension scheme; (c) primary “suspension” scheme (1, wheelset; 2, side frame; 3, bolster; 4, braking leverage; 5, centre bowl; 6, rigid side bearings; 7, suspension springs; 8, friction wedge; 9, axlebox).

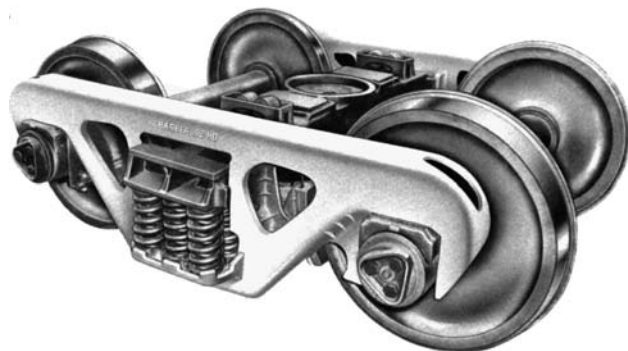
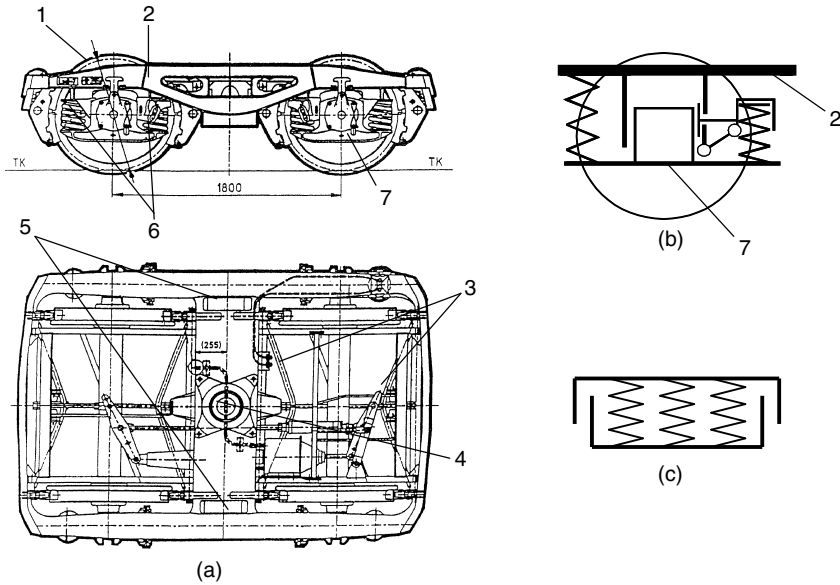


FIGURE 3.38 Barber S-2 bogie.



**FIGURE 3.39** The Y25 bogie: (a) general view; (b) primary suspension scheme (Lenoir damper); (c) elastic side bearing scheme (1, wheelset; 2, rigid H-shaped frame; 3, braking leverage; 4, centre bowl; 5, side bearings; 6, spring set; 7, axlebox).

In curves, the three-piece bogies demonstrate the “lozenging” or “warping” effects, when the two side frames adopt a parallelogram position (in plan view). In this instance, the wheelsets cannot adopt a radial position in the curve, and generate large angles of attack, which leads to constant contact between the high-rail flange of the leading wheelset and the rail causing high levels of wear.

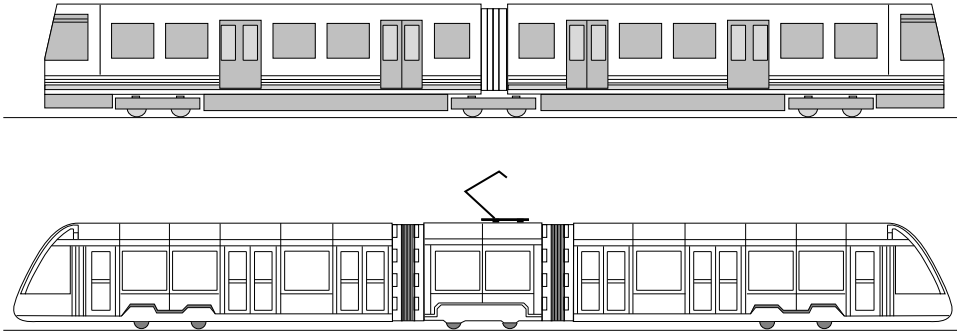
The Y25 (and similar bogies, such as the Y33) are commonly used on European freight vehicles. An example of this bogie is shown in Figure 3.39.

The Y25 bogie has a single-stage primary suspension consisting of a set of pairs of nested coil spring (with a bilinear characteristic for tare/laden ride) and a Lenoir link friction damper (Figure 3.39b) providing vertical and lateral damping. The friction force depends on the vertical load on the spring set, a component of which is transferred to the friction face by the inclined Lenoir link. Derailment safety is improved by the provision of vertical clearance between the inner and outer springs in each pair giving a lower stiffness in tare than in laden. Whilst improving the ride in both conditions, problems may still arise with the part-laden ride, when the bogie is just resting on the inner “load” spring making the suspension relatively stiff for the load being carried.

The bogie has a rigid H-shaped frame that consists of two longitudinal beams, one lateral and two end-beams and may be either cast or fabricated. The connection of the vehicle body is different to the three-piece bogies described above. The centre bowl has a spherical surface to reduce asymmetric forces on the frame and elastic side bearers without clearance resist the body roll motions (Figure 3.39c).

## V. COMMON TRAM BOGIE DESIGNS

Trams and light rail vehicles (LRVs) are generally designed to negotiate very small radius curves and be compact enough for street running with overthrows minimised to avoid contact with cars. Modern trends for the low floor trams to improve accessibility leads to a requirement for very compact running gears.



**FIGURE 3.40** Typical tram configurations.

Traditionally tram bogies have a rigid frame and double suspension similar in design to conventional passenger bogies. Modern tram designs typically use one of the articulated arrangements shown in Figure 3.40.

In the example above, one short body section is rigidly fixed to the centre bogie, such that body section and bogie rotate together when passing a curve (Figure 3.41). Slewing rings are often used to join tram car bodies and bogies if it is necessary to allow for large rotation angles (Figure 3.42).

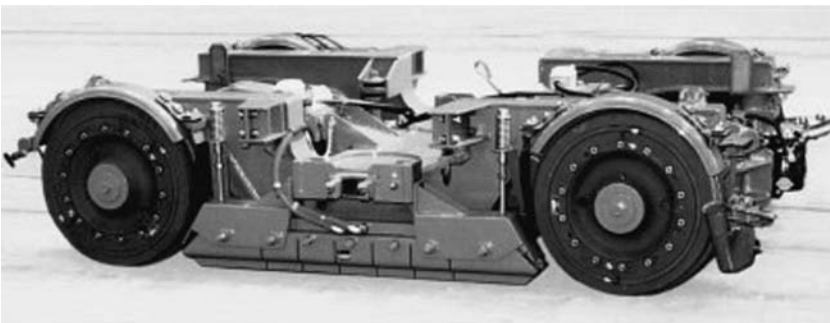
Wheelsets for tram bogies traditionally have a rigid axle (as in Figure 3.41 and Figure 3.42) and a small wheel diameter. However, low floor trams require alternative arrangements and in this case independently rotating wheels are often mounted in a common subframe (Figure 3.43).

Tram bogie primary suspensions typically use small rubber interleaved (chevron) springs. The primary suspension and bearings are generally located inside the wheels to reduce the overall size of the bogie.

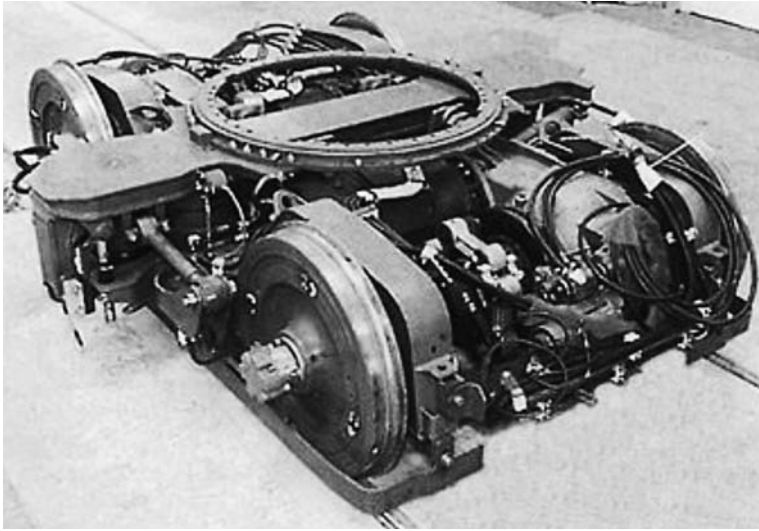
Tram secondary suspensions use coil or air springs. The latter have the advantage of allowing constant floor height to be maintained for various loading conditions. Trams have smaller vertical deflection of the suspension compared to conventional passenger coaches as they operate at lower speeds and may have less stringent ride quality requirements due to shorter passenger journey times.

## VI. PRINCIPLES OF SELECTING SUSPENSION PARAMETERS

The parameters of a rail vehicle may be considered optimal if its dynamic characteristics meet three groups of requirements:



**FIGURE 3.41** Tram bogie with rigid connection to the car body.



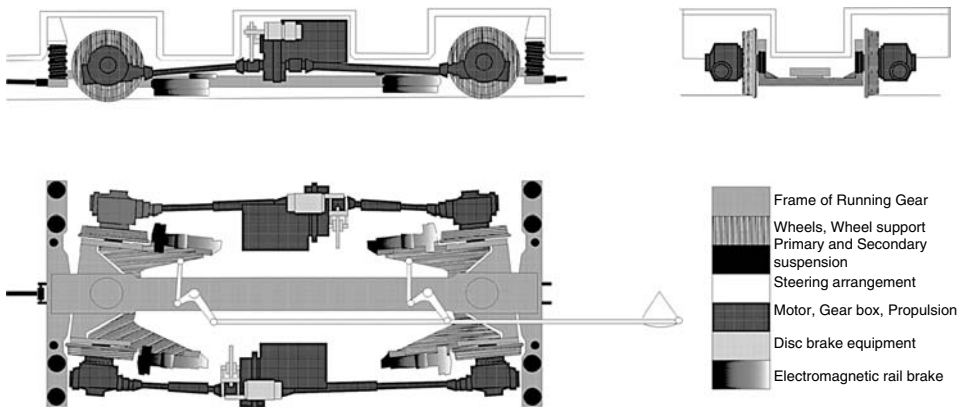
**FIGURE 3.42** Bogie with ball bearing support of car body.

1. There is sufficient reserve of critical speed with respect to design speed
2. Ride quality, track forces, and safety factors satisfy the standards on straight track and in curves in all range of operation speeds
3. Wear of friction elements and wheel profiles is within acceptable limits.

Experience in the development of rail vehicles shows that at the preliminary stage the suspension parameters can be estimated using the simple engineering approaches described below. To make sure that the parameters are optimised, further refinement is usually done using computer simulation.

**A. SELECTING VERTICAL SUSPENSION CHARACTERISTICS**

Suspension should control and damp the motions of both the sprung and unsprung masses in the vehicle to obtain the best possible ride qualities whilst strictly fulfilling the safety requirements



**FIGURE 3.43** Tram bogie with independently rotating wheels.

satisfying specific service limitations such as ensuring that the vehicle remains within the loading gauge.

Bogie elastic elements have various constructions, as described above. In a simple initial analysis of the vertical behaviour, it is not the specific construction of the elastic element that is important, but the force characteristic that it provides, i.e., the dependence of the vertical load on the element  $P$  from its static deflection  $f$ :  $P = P(f)$ .

The static deflection of a suspension with linear characteristics (constant stiffness) is determined by formula:

$$f_{st} = \frac{P_{st}}{c} \tag{3.7}$$

where  $P_{st}$  is the static load on the suspension;  $c$  is the stiffness of the suspension.

For a linear suspension, there is a dependence between the bounce natural frequency and the static deflection:

$$\omega^2 = c/M = g/f_{st}$$

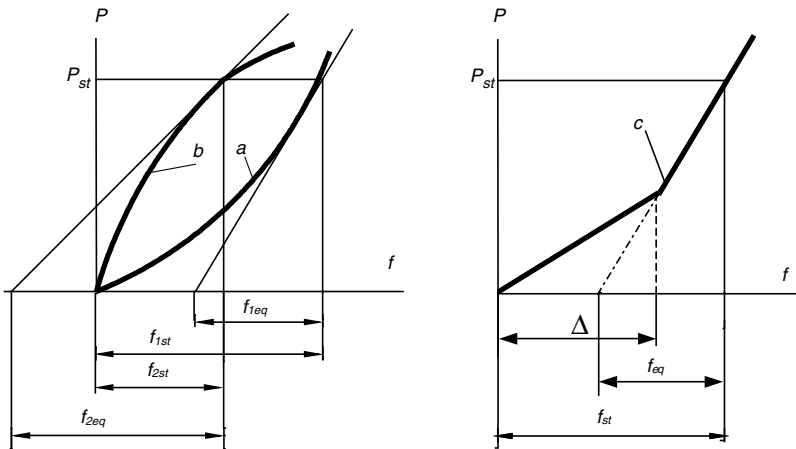
where  $M$  is the sprung mass of the vehicle;  $g$  is the gravity acceleration.

Research has shown that decreasing the suspension stiffness is favourable for the dynamic performance of rail vehicles if other conditions do not change. In general, a low suspension stiffness gives lower accelerations but practical considerations dictate that there must be a relatively small height difference between tare and laden conditions. In addition the human perception of vibrations over a range of frequencies must be considered. For passenger vehicles, the body bounce frequency is generally in the range 0.9–1.2 Hz, whilst for freight wagons this frequency can rise to 2.5 Hz in laden and up to 4 Hz in the tare condition.

In order to avoid excessive suspension deflections, modern suspensions use nonlinear springs to provide optimal stiffness in the vicinity of the static deflection corresponding to required load.

In suspension elements with variable stiffness (Figure 3.44), the dynamic oscillations appear around an equilibrium position given by static force  $P_{st}$ . To estimate the oscillation frequency in this case the equivalent stiffness and equivalent deflection are used:

$$c_{eq} = \left. \frac{dP}{df} \right|_{P=P_{st}} ; \quad f_{eq} = \frac{P_{st}}{c_{eq}} \tag{3.8}$$



**FIGURE 3.44** Nonlinear elastic force characteristics of a suspension: (a) stiff characteristic; (b) soft characteristic; (c) bilinear characteristic.



In a suspension with stiff force characteristic as shown in Figure 3.44a  $f_{eq} < f_{st}$ , and for a soft characteristic (Figure 3.44b)  $f_{eq} > f_{st}$ , where  $f_{st}$  is the total static deflection.

Freight rolling stock often uses a suspension with a bilinear force characteristic as shown in Figure 3.44c. The first part of the characteristic has a constant stiffness  $c_1$  chosen to give the required frequency for the tare condition, whilst the second part with constant stiffness  $c_2$  gives the required frequency for the laden wagon. In this case:

$$c_{eq} = \begin{cases} c_1, & f \leq \Delta \\ c_2, & f > \Delta \end{cases}; \quad f_{eq} = \begin{cases} \frac{P_{st}}{c_1}, & P_{st} \leq c_1 \Delta \\ \frac{P_{st}}{c_2}, & P_{st} > c_1 \Delta \end{cases} \quad (3.9)$$

where  $\Delta$  is the deflection corresponding to the breakpoint of the bilinear characteristic. The value of  $\Delta$  is usually chosen so as to ensure that the breakpoint is not reached during normal suspension movements of the wagon in the tare condition.

Some further limitations on the value of suspension static deflection due to service conditions are discussed below.

An important limitation is imposed by the need to restrict height differences between the couplings of adjacent vehicles. The worst case is calculated from the coupling height of the gross laden car with maximum possible wear of bogie components and the height of the tare vehicle with new bogies (without wear). The difference in the coupler levels is due to static deflection of the suspension under the maximum load, aging of elastic elements and wear of bogie components (for example, wheel profile wear or wear of centre bowls and side bearers).

In service, the car body roll must also be limited to prevent the risk of overturning on highly canted curves and to ensure the vehicle remains within the required loading gauge. Once the maximum allowable roll angle for the vehicle body and the maximum lateral force (centrifugal, wind, and lateral component of the interaction force between the vehicles in curves) has been established, the equilibrium equation gives the minimum acceptable vertical stiffness of the suspension.

The final value of vertical stiffness for the suspension is chosen to be the maximum of the minimum values calculated using the service and design limitations.

**B. SELECTING THE LATERAL AND LONGITUDINAL PRIMARY SUSPENSION STIFFNESS**

Theoretical investigations and experiments show that wheelset stability increases with increasing stiffness of the connection to the bogie frame. However, the character of this dependence is highly nonlinear and the relationship between suspension stiffness and the mass and conicity of the wheels influences the critical speed. Increasing the longitudinal stiffness of the primary suspension impairs the guidance properties of the wheelset in curves whilst increasing the lateral stiffness reduces the ability of the wheelset to safely negotiate large lateral irregularities.

A fundamental conflict therefore exists between the requirements for high speed stability on straight track and good curving with safe negotiation of track irregularities. The “in-plane” (lateral and longitudinal) stiffnesses must therefore be selected to give the best compromise for the conditions under which the vehicle will operate.

In order to make a preliminary choice of bogie in-plane stiffness, it is useful to know the relationship between stiffness and ride quality in an analytical or graphical form. The simplified approach described in Ref. 2 is useful as a starting point.

The natural vibration modes shown in Figure 3.45 and Figure 3.46 can be obtained from the linear equations of motion for two-axle bogie.<sup>2,3</sup>

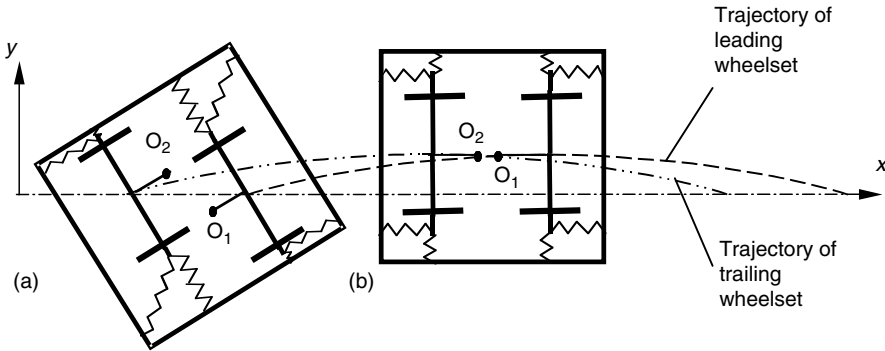


FIGURE 3.45 Wheelset modes for a two-axle bogie: (a) in-phase yaw; (b) in-phase lateral displacement.

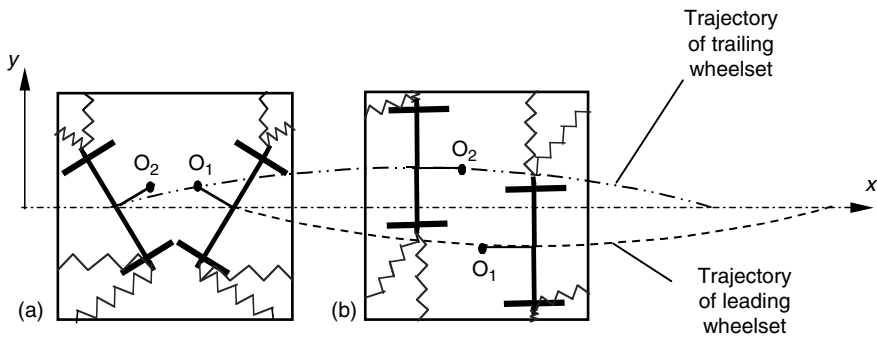


FIGURE 3.46 Wheelset modes for a two-axle bogie: (a) anti-phase yaw; (b) anti-phase lateral displacement.

Analysis of the modes shows:

- For in-phase yaw according (Figure 3.45a) there is a relative lateral displacement between the centres of wheelsets  $O_1$  and  $O_2$  and the bogie centre
- Similar lateral displacements appears for the anti-phase mode shown in Figure 3.46b
- Relative rotation between wheelset centres  $O_1$  and  $O_2$  occur only for anti-phase yaw of wheelsets (Figure 3.46a).

Thus, two generalised parameters can be introduced for the bogie:

1. A stiffness corresponding to relative lateral displacement between the centres of wheelsets referred to as the shear stiffness ( $K_s$ )
2. A stiffness corresponding to the relative yaw angle between the wheelsets referred to as the bending stiffness ( $K_b$ )

The conventional representation of bogie shear and bending stiffness is shown in Figure 3.47 as translational and torsion springs, respectively. The generalised stiffnesses  $K_s$  and  $K_b$  have a particular physical meaning. The shear stiffness  $K_s$  has a greater influence on the critical speed of the vehicle, whilst the bending stiffness  $K_b$  mainly determines the wheelsets' angles of attack in curves.

The use of shear and bending stiffness to give a simplified representation of the primary suspension without consideration of the bogie frame inertia (Figure 3.47) allows the in-plane bogie stiffnesses to be chosen without considering its specific design.

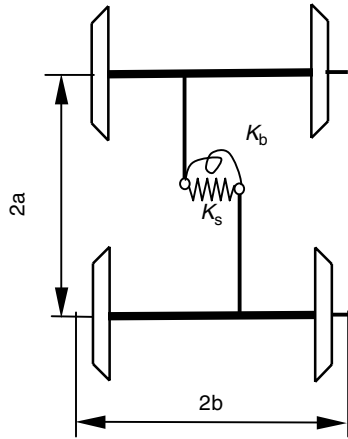


FIGURE 3.47 Representation of the primary suspension using shear and bending stiffness.

Solution of the stability problem<sup>4</sup> shows that the critical speed of a conventional railway vehicle is a function of its shear and bending stiffness as shown in Figure 3.48. The quality of curving can be estimated using the relationship of the wear number (the sum of creep force power for all wheels of the vehicle) to the shear and bending stiffness as shown in Figure 3.49.

These relationships show that the chosen bending stiffness of the bogie should be the minimum that provides the required critical speed and the shear stiffness should be within the critical speed range for the chosen bending stiffness.

### C. SELECTING SUSPENSION DAMPING

Damping is typically provided within the suspension by either friction or hydraulic devices. Some types of elastic elements, such as leaf springs, have sufficient internal friction damping to avoid the necessity of a separate damper.

The selection of the optimum damping levels is more complicated than the choice of suspension stiffness. High levels of damping decrease the amplitudes of vibrations in resonance situations but significantly increase the accelerations acting on the vehicle body for higher frequency inputs such as short wavelength track irregularities.

Hydraulic dampers are almost universally used for passenger vehicles. Let us consider the simplified case of linear dependence between the damper force and the velocity. In this case attenuation of vehicle vibrations is determined by the ratio of the real part of the eigenvalue to the corresponding natural frequency. This is termed the damping coefficient and is different for different natural vibration modes:

$$d_i = \frac{1}{2\omega_i} \frac{\{v_i\}^T [B] \{v_i\}}{\{v_i\}^T [M] \{v_i\}} \tag{3.10}$$

where  $[B]$ ,  $[M]$  are the damping and inertia matrices of the vehicle multi-body model, respectively,  $\{v_i\}$  is the column-vector of  $i$ th eigenmode and  $\omega_i$  is the natural frequency of  $i$ th eigenmode. In a simple multibody model, the vehicle body, wheelsets, and bogie frames are represented by rigid bodies connected with the elastic and damping elements.

Effective damping of the vibrations of railway vehicles is typically obtained with damping coefficients which lie in the following ranges: 0.2–0.3 for vertical oscillations; 0.3–0.4 for horizontal oscillations, and 0.1–0.2 for vehicle body roll.

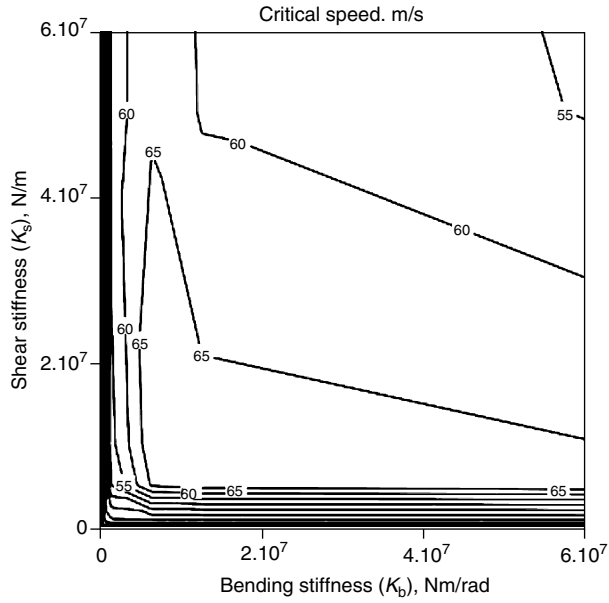


FIGURE 3.48 Critical speed as a function of shear and bending stiffness.

In freight bogies, friction dampers are commonly used. When making the preliminary choice of parameters, the friction force in the damper is estimated on the basis that the amplitude should not increase in the resonance case.

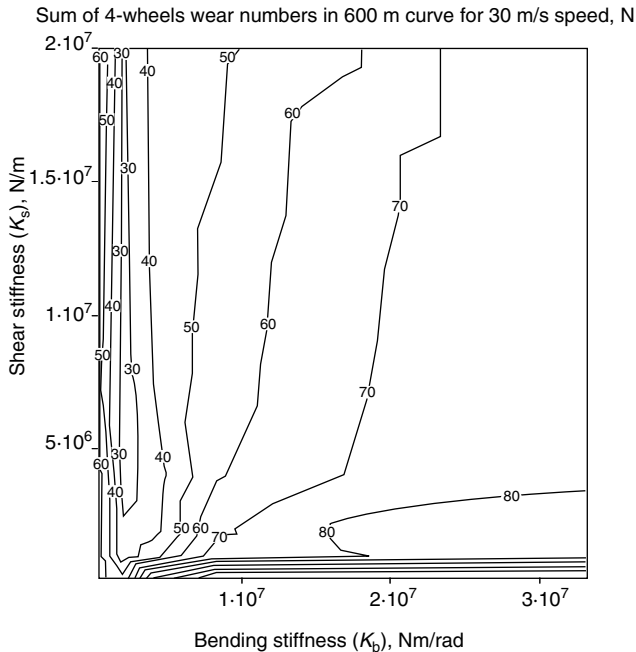


FIGURE 3.49 Wear number as a function of shear and bending stiffness for a 600 m curve at 30 m/sec.

Assuming that the amplitude of oscillations at resonance increases by  $\Delta A'$  during one period, and the friction force  $F$  acting in the suspension reduces it by  $\Delta A''$ , the following conditions must apply to prevent the amplitude increasing in the resonant case:

$$\Delta A'' \geq \Delta A' \quad (3.11)$$

The equations of oscillation for the system with dry friction under periodic excitation give:

$$F \geq \frac{\pi q}{4} c_{\text{eq}} \quad (3.12)$$

where  $q$  is the estimated amplitude of periodic track irregularity (prescribed in regulations) and  $c_{\text{eq}}$  is the equivalent stiffness of the suspension.

Estimating the magnitude of the friction force is easier using relative friction coefficients that equal the ratio of friction force to the static vertical load:

$$\varphi = \frac{F}{P_{\text{st}}} \geq \frac{\pi q}{4f_{\text{eq}}} \quad (3.13)$$

where

$$f_{\text{eq}} = \frac{P_{\text{st}}}{c_{\text{eq}}}$$

For freight cars the recommended value of relative friction coefficient is typically in the range 0.05–0.15.

The relative friction coefficient is a general parameter of the wagon, and the optimal value of friction force depends on the equivalent static deflection of the suspension, or for the case of nonlinear suspension characteristic, the vertical load.

## VII. ADVANCED BOGIE DESIGNS

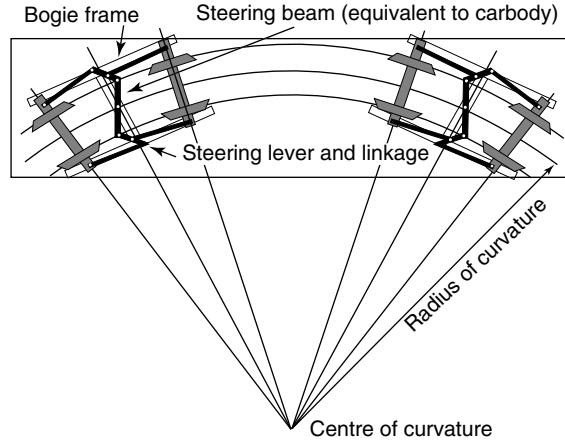
Many novel bogie designs address the fundamental conflict between stability on straight track and good curving described above. It is clear from the foregoing discussion that the bogie should maintain stable conditions on straight track but allow the wheelsets to adopt a radial position in curves.

Bogies where the wheelsets adopt or are forced to take an approximately radial position in curves (Figure 3.50) are called radially steered bogies. Such designs have small angles of attack which leads to significant decrease of flange wear and lower track forces.

Radially steered bogies fall into two groups: those with forced steering of the wheelsets in curves and those with self-steering of the wheelsets. In the first case, the wheelsets are forced to adopt a radial position due to linkages between the wheelsets or linkages from the wheelset to the vehicle body. Various methods of obtaining forced steering for radially steered bogies are shown in Table 3.4. The bogies may be split into three groups depending on the control principle used:

1. Wheelsets yawed by the wheel-rail contact forces
2. Wheelsets yawed by the relative rotation between the bogie frame and vehicle body (either yaw or roll)
3. Wheelsets yawed by an external energy source (electric, hydraulic, or pneumatic actuators)

The first two groups in Table 3.4 have passive control systems that change the kinematic motion of the wheelset depending on the curve radius. Designs where the energy source is provided by



**FIGURE 3.50** Radial position of wheelsets for a bogie with inter-axle steering linkages.

steering force in wheel–rail contact, may be considered preferable as the behaviour of systems relying on interconnection to the car body is dependent upon vehicle speed.

Designs where the wheelsets are forced to adopt a radial position by hydraulic, pneumatic, electric actuators (or a combination of these) are called actively controlled bogies. These are considered in detail in [Chapter 6](#).

Three main groups of executive mechanisms are common: those using links between wheelsets, those using an arrangement of levers, or those using sliders.

[Figure 3.51](#) shows a passenger bogie using a passive steering system with Watts linkage.

An example of a freight bogie with passive control using diagonal links between the axleboxes designed by Scheffel is shown in [Figure 3.52](#).

**TABLE 3.4**  
**Classification of Forced Steering Mechanisms**

Energy Source — Control Mechanism	Steering Mechanisms		
	Links	Levers	Sliders
Wheel–rail interaction			
Car body yaw or centrifugal force action			
Active external force			

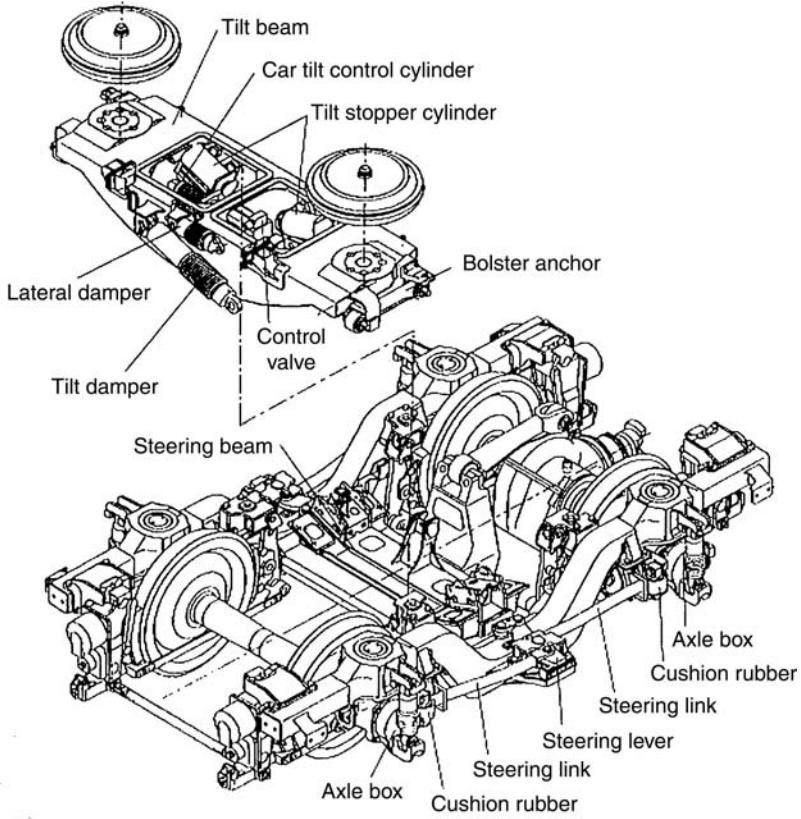


FIGURE 3.51 Passenger car bogie with passive steering system.

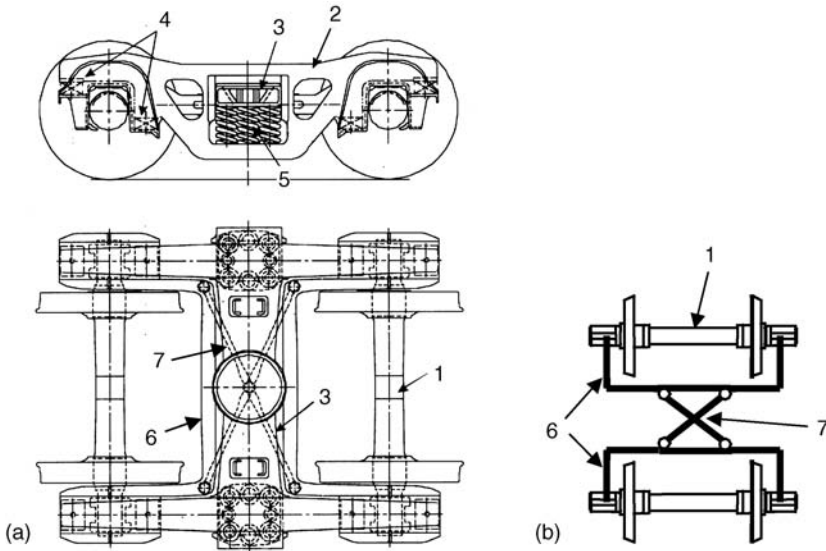
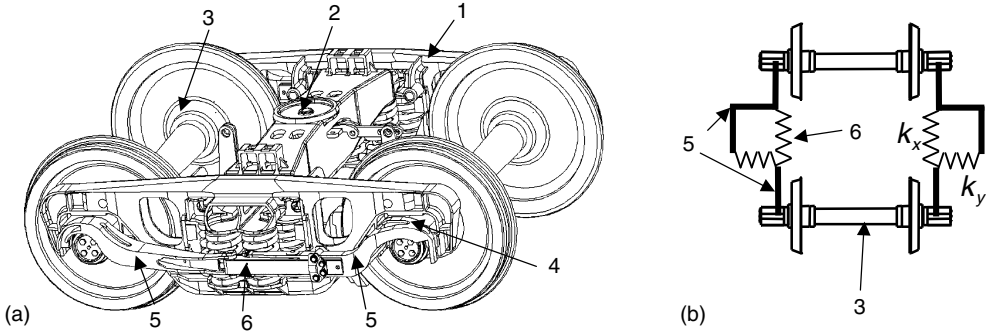


FIGURE 3.52 Scheffel HS bogie with diagonal linkage between wheelsets: (a) general view; (b) the principal scheme of inter-axle linkages (1, wheelset; 2, side frame; 3, bolster; 4, primary suspension; 5, secondary suspension; 6, subframe; 7, diagonal links).



**FIGURE 3.53** Scheffel radial arm bogie: (a) general view; (b) principal scheme of inter-axle linkages (1, side frame; 2, bolster; 3, wheelset; 4, primary suspension; 5, two longitudinal arms; 6, elastic elements between the arms).

The second group of radially steered bogies are those with wheelsets which are self-steering in curves. The design of such bogies is based on selecting the optimum shear and bending stiffnesses. This may be aided by using designs that allow these stiffnesses to be decoupled.

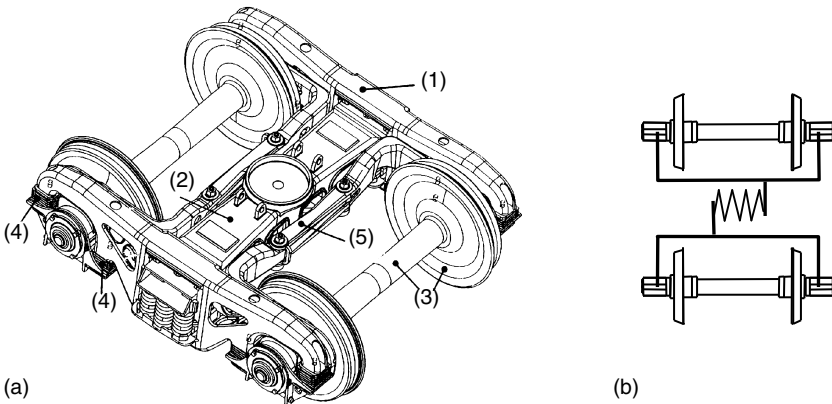
In conventional suspension arrangements, the bending and shear stiffness are not independent. Decreasing the bending stiffness leads to a reduction of shear stiffness, which means that improving the curving qualities leads to reduced stability on straight track. Inevitably, therefore, the bogie in-plane stiffness is chosen to give the best compromise between curving and stability.

In order to resolve the curving–stability controversy, Scheffel proposed several arrangements of inter-axle linkages,<sup>4</sup> two of which are shown in Figure 3.53 and Figure 3.54.

The Scheffel radial arm bogie is shown in Figure 3.53, and for it the generalised bogie stiffness has the following expressions:

Shear stiffness:

$$K_{s\Sigma} = 2k_y + K_s \tag{3.14}$$



**FIGURE 3.54** Scheffel bogie with A-frame inter-axle linkages: (a) general view; (b) principal scheme of inter-axle linkages (1, side frame; 2, bolster; 3, wheelset; 4, primary suspension; 5, elastic connection between the subframes).



where  $k_y$  is the lateral stiffness of inter-axle linkage (per side) and  $K_s$  is the shear stiffness provided by the bogie frame.

Bending stiffness:

$$K_{b\Sigma} = 4b^2k_x + K_b \quad (3.15)$$

where  $k_x$  is the longitudinal stiffness of inter-axle linkage (per side) and  $K_b$  is the bending stiffness provided by the bogie frame.

Thus, the expressions for  $K_{s\Sigma}$  and  $K_{b\Sigma}$  contain two independent parameters  $k_x$  and  $k_y$  that allow optimum shear and bending stiffnesses to be selected.

Such bogie designs are based on the three-piece bogie consisting of a bolster and two side frames. The Scheffel designs retain the advantages of the three-piece bogie when negotiating large track irregularities and carrying the asymmetric loads. However, wheelset steering is provided not by the frame (as in traditional designs), but by the inter-axle links. In order for the inter-axle links to be effective, the bogie must have a low longitudinal and lateral primary suspension stiffness. These bogie designs are effectively therefore double suspended.

## REFERENCES

1. Oriol de, L. M. El Talgo Pendular. *Revista Asociacion de Investigacion del Transporte*, Vol. 53. pp. 1–76 (1983).
2. Wickens, A. H. *Fundamentals of Rail Vehicle Dynamics: Guidance and Stability*. Lisse: Swets & Zeitlinger BV. ISBN 90-265-1946 (2003).
3. Garg, V. K. and Dukkipati, R. V. *Dynamics of Railway Vehicle Systems*. San Diego: Academic Press (Harcourt Brace Jovanovich Publishers). ISBN 0-12-275950-8 (1984).
4. Orlova, A., Boronenko, Y., Scheffel, H., Fröhling, R. and Kik, W. Tuning von Güterwagendrehgestellen durch Radsatzkopplungen. *ZEV-Glasers Annalen* 126: S270–S282 (2002).

---

# 4 Wheel–Rail Contact

*Jean-Bernard Ayasse and Hugues Chollet*

## CONTENTS

I.	Introduction .....	86
A.	Basic Model of a Wheelset, Degrees of Freedom .....	87
II.	The Normal Contact.....	88
A.	Hertzian Contact .....	88
1.	Ratio $A/B$ , Relation with $b/a$ .....	90
2.	Calculation of the Semi Axes .....	90
3.	Convexity, Concavity and Radius Sign .....	91
4.	Particular Case: 2D Contact.....	91
5.	Contact Pressure .....	91
B.	Tables, Polynomial Expressions and Faster Methods .....	92
1.	Continuous Expression of the Tables .....	92
2.	Analytical Approximation of the Tables .....	92
C.	Application to the Railway Field.....	93
1.	Contact Plane Angle, Conicity.....	93
2.	A Note on Conicity .....	93
3.	Normal Load.....	93
4.	Determination of the Longitudinal Curvature of the Wheel.....	93
5.	Contact Point between the Wheel and Rail Profiles .....	93
6.	Direct Comparison of the Antagonist Curvatures .....	94
7.	One or Two Wheel–Rail Pairs? .....	94
III.	The Tangent Problem.....	95
A.	Forces and Couples on a Wheelset.....	95
1.	Approximations .....	95
B.	Tangent Forces: Rolling Friction Simple Models .....	96
1.	Historical Review .....	96
C.	Linear Expressions of the Creep Forces.....	97
D.	Definition of Creepages .....	97
1.	Quasi-Static Creepages.....	97
2.	Quasi-Static Creepages in the Railway Case.....	97
3.	Longitudinal Creepage .....	98
4.	Lateral Creepage .....	98
5.	Spin Creepage.....	99
6.	Dynamic Formulation of the Creepages .....	99
7.	Damping Terms and Stability .....	99
8.	Nondimensional Spin Creepage.....	99
E.	Kalker’s Coefficients $c_{ij}$ .....	100
1.	$c_{ij}$ Values for Simplified Bogie Models.....	101
F.	Creep Forces in the Linear Domain .....	101
1.	Dependence on Load.....	101

2.	Creepages Combinations and Saturation .....	101
3.	Using Linear Models: The C110, C220 Stiffnesses .....	102
4.	Reduced Creepages .....	102
G.	Saturation Laws.....	102
1.	Vermeulen and Johnson .....	102
2.	Kalker's Empirical Proposition.....	103
3.	Exponential Saturation Law: CHOPAYA, Ohyama and Others.....	103
H.	Surface-Based Contact Models.....	103
1.	Discretised Ellipse: FASTSIM from Kalker.....	104
2.	Stresses .....	105
3.	Linear Contact Forces, Elastic Coefficients.....	105
4.	Reduced Creepages in FASTSIM.....	106
5.	Extensions of FASTSIM .....	107
I.	Methods Based on FASTSIM.....	107
IV.	Contact Forces in the Railway Context.....	107
A.	From the Pure Dicone to the Wheelset with Real Profiles.....	107
1.	Equivalent Conicity.....	107
2.	Variable Conicity .....	108
3.	Profile Measurements: Importance of the Angular Reference .....	108
B.	Gravitational Stiffness.....	108
C.	Flange Contact .....	109
1.	Flange Contact Jump and Definition of the Nominal Play in the Track .....	109
2.	Contact Jump and Load Transfer.....	109
D.	Using the Contact Angle Function .....	110
1.	Localisation of Multiple Contacts on the Tread.....	110
2.	Gravitational Centring Ability in Standard Conditions.....	110
E.	The Diverging Effect of Spin, Influence on the Normal Load .....	110
F.	Safety Criteria, Nadal's Formula.....	112
G.	Independent Wheel, Application to Industrial Mechanisms on Rails .....	112
H.	Modelling the Contact Jumps .....	113
1.	Hertzian Multiple Contacts .....	113
I.	Advanced Methods for Non-Hertzian Contacts .....	114
J.	Wheelset Equilibrium, Explicit Scheme.....	115
1.	Wheelset Equilibrium Equations .....	115
2.	Decomposition of Forces .....	115
3.	Adaptation to the Numerical Calculation .....	115
4.	Equilibrium Equations.....	116
Appendix 4.1.	Kinematic Movement: The Klingel Formula .....	116
Appendix 4.2.	Kinematic Hunting and Equivalent Conicity.....	117
Appendix 4.3.	The Circle Theory .....	117
Appendix 4.4.	Analysis of $Y/Q$ and Nadal's Criteria .....	117
Nomenclature.....		119
References.....		120

## I. INTRODUCTION

For more than 150 years, the wheel–rail system has provided a relatively safe system of transport. This safety level is so high that the mechanism is generally neglected and considered as a simple slider by most people.

However, the engineer’s point of view can be different, especially when taking into account responsibilities in a railway network. The wheel–rail contact is actually a complex and imperfect link. Firstly, it is a place of highly concentrated stresses. The conical wheel shape makes the wheelset a mechanical amplifier, limited by the transverse play, with partially sliding surfaces. The contact surfaces are similar to those in a roller bearing but without protection against dust, rain, sand, or even ballast stones.

Looking closer, railway safety can stay at a very high level if some precautions are taken. The aim of this chapter is to provide railways engineers with a basis for understanding and evaluating the wheel–rail contact situation.

Historically, the first theoretical model of the wheel–rail longitudinal contact force is due to Carter in the U.S.A.<sup>1</sup> More recently, Johnson (in the U.K.),<sup>2</sup> and Kalker (in The Netherlands),<sup>3</sup> set out the basis for an accurate description.

In parallel, dynamic software developers decided to model the wheel–rail contact with, at first, a constant conicity limited by the two flange contact springs. Variable conicity was then taken into account followed by the spin effect and contact jumps due to the profile combination. These jumps are a major difficulty in the calculation of wheel–rail forces and this is a major step towards the nonHertzian description of the contact. However, the first step in the wheel–rail contact study is to consider the Hertzian modelisation starting from a simple model of the wheelset.

## A. BASIC MODEL OF A WHEELSET, DEGREES OF FREEDOM

If the track is considered to be rigid, then the railway wheelset has two main degrees of freedom:

- The lateral displacement, or shift,  $y$
- The yaw angle,  $\alpha$

When the behaviour of a wheelset is unstable, the dynamic combination of these two degrees of freedom is called “*hunting*.”

The lateral displacement and the yaw angle must be considered as two small displacements relative to the track. The play will be the limit of the lateral displacement between the two flange contacts. It is generally approximately  $\pm 8$  mm.

The other degrees of freedom are constrained: the displacement along  $Ox$  and the axle rotation speed  $\omega$  around  $Oy$  are determined by the longitudinal speed  $V_x$  and the rolling radius of the wheel  $r_0$  with:  $V_x = \omega r_0$ . The wheelset centre of gravity height  $z$  and the roll angle around  $Ox$  are linked to the rails when there is contact on both rails.

The railway wheelset is basically described by two conical, nearly cylindrical wheels (Figure 4.1 and Figure 4.2), linked together with a rigid axle. Each wheel is equipped with a flange, the role of which is to prevent derailment. In a straight line the flanges are not in contact, but the

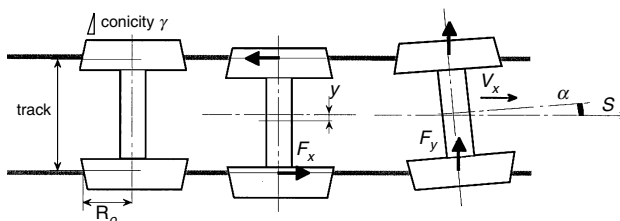
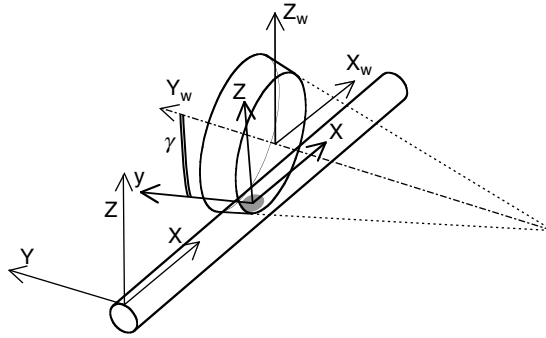


FIGURE 4.1 Wheelset degrees of freedom.



**FIGURE 4.2** Rail, wheel and contact frames.

rigid link between the two wheels suggest that the railway wheelset is designed to go straight ahead, and will go to flange contact only in curves. This is the railway dicone or wheelset.

The interface between the wheel and the rail is a small horizontal contact patch. The contact pressure on this small surface is closer to a stress concentration than in the rest of the bodies. The centre of this surface is also the application point of tangential forces (traction and braking  $F_x$ , guiding or parasite forces  $F_y$ , see Figure 4.1). The knowledge of these forces is necessary to determine the general wheelset equilibrium and its dynamic behaviour.

In order to determine this behaviour and these forces, the first thing to do is to determine some contact parameters: the contact surface, the pressure and the tangential forces. This determination is generally separated into two steps:

1. The normal problem (Hertz theory)
2. The tangential problem (Kalker's theory)

## II. THE NORMAL CONTACT

The study of the contact between bodies is possible today with finite element methods. However, the necessity to calculate as quickly as possible in the dynamic codes leads to the use of analytical methods. This section first describes the classical Hertzian model followed by some particular considerations aimed at speeding up the calculation.

### A. HERTZIAN CONTACT

Hertz demonstrates that when two elastic bodies are pressed together in the following conditions:

- Elastic behaviour
- Semi-infinite spaces
- Large curvature radius compared to the contact size
- Constant curvatures inside the contact patch.

then:

- The contact surface is an ellipse
- The contact surface is considered flat
- The contact pressure is a semi-ellipsoid.

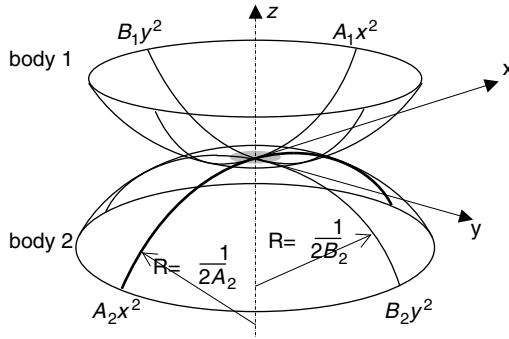


FIGURE 4.3 Hertzian contact: general case.

The main curvatures of the two semispaces are needed for the calculation of surface dimensions and pressure distribution. In the railway case, the four main curvatures can be considered to be in perpendicular planes; their directions correspond to the main axes of the frame:  $xOy$  (Figure 4.3 and Figure 4.4).

Considering the two elastic bodies in contact, they will meet at a single point  $O$  where the normal distance between them is minimal. Near this contact point  $O$ , without load, the bodies surface shapes are represented by two second order polynomials:

$$\begin{aligned} z_1 &= A_1x^2 + B_1y^2 \\ z_2 &= A_2x^2 + B_2y^2 \end{aligned} \tag{4.1}$$

The  $A_{1,2}$  and  $B_{1,2}$  coefficients are assumed constant in the neighbourhood of the contact point  $O$ , and linked to the main local curvatures by the second partial differential expressions, the first being neglected if described in the contact frame.

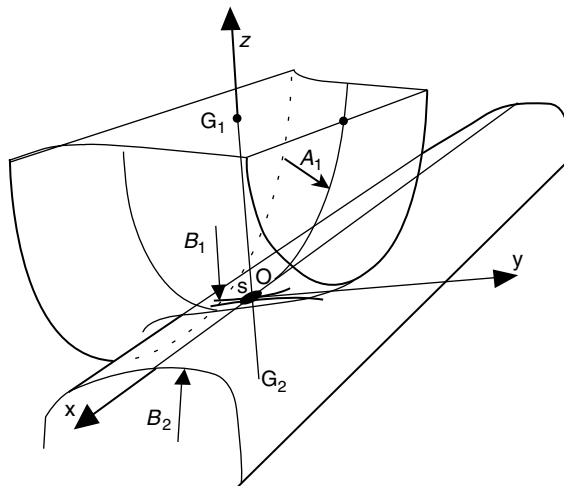


FIGURE 4.4 Hertzian contact: the railway case.

In the railway case (Figure 4.4), these curvatures and radius will be noted:

$$\begin{aligned} \text{Wheel:} \quad & \frac{d^2 z_1}{dx^2} = 2A_1 \approx \frac{1}{r_n} \\ & \frac{d^2 z_1}{dy^2} = 2B_1 \approx \frac{1}{R_{wx}} \\ \text{Rail:} \quad & \frac{d^2 z_2}{dy^2} = 2B_2 \approx \frac{1}{R_{rx}} \end{aligned} \quad (4.2)$$

In the railway case, the curvature  $A_2$  is generally neglected as the rail is straight: the radius is infinite.  $B_1$  and  $B_2$  are deduced from the transverse profiles,  $A_1$  from  $r_n$ , the normal radius of the wheel, itself deduced from  $r_o$ , the rolling radius of the wheel (see Section II.C.4 and Figure 4.5).

### 1. Ratio $A/B$ , Relation with $b/a$

Before being loaded, the vertical relative distance  $d(x,y)$  between the two bodies can be written:

$$z_1 + z_2 = d = Ax^2 + By^2$$

with:

$$A = \frac{1}{2r_n} \quad \text{and} \quad B = \frac{1}{2} \left( \frac{1}{R_{wx}} + \frac{1}{R_{rx}} \right) \quad (4.3)$$

$A$  and  $B$  being strictly positive.

Conventionally,  $a$  is the ellipse longitudinal semi axis in the  $Ox$  direction and  $b$  is in the transversal  $Oy$  direction. The  $A/B$  and  $b/a$  ratios vary in the same way: if  $A > B$ , then  $b > a$ . The equality  $A/B = 1$  leads to a circular contact patch:  $a = b$ .

### 2. Calculation of the Semi Axes

The traditional calculation is based on the determination of the semi axis ratio:  $g < 1$ , ( $g = b/a$  or  $a/b$ ), function of  $B/A$  by using an intermediate parameter, the angle  $\theta$  defined as:

$$\cos \theta = \frac{|B - A|}{B + A} \quad (4.4)$$

The practical values of the semi-axis  $a$  and  $b$ , and  $\delta$  the reduction of the distance between the bodies centres are given by:

If  $a > b$ :

$$\begin{aligned} a &= m \left( \frac{3}{2} N \frac{1 - \nu^2}{E} \frac{1}{A + B} \right)^{1/3} \\ b &= n \left( \frac{3}{2} N \frac{1 - \nu^2}{E} \frac{1}{A + B} \right)^{1/3} \\ \delta &= r \left( \left( \frac{3}{2} N \frac{1 - \nu^2}{E} \right)^2 (A + B) \right)^{1/3} \end{aligned} \quad (4.5)$$

with  $E$  being the Young’s modulus and  $\nu$  the Poisson’s ratio, assuming the same material for the rail and the wheel.

$m$ ,  $n$ , and  $r$  are nondimensional coefficients tabulated as a function of the ratio  $g = n/m$  or the angle  $\theta$  (see Table 4.1).

$\pi ab$  being the surface area of the ellipse, it can be expressed as a function of:

$$ab = mn \left( \frac{3}{2} N \frac{1 - \nu^2}{E} \frac{1}{A + B} \right)^{2/3} N^{2/3} \tag{4.6}$$

in which the first term contains the material and geometrical constants, and the second only the load.

### 3. Convexity, Concavity and Radius Sign

The sign of each radius is important, because one of the calculation method uses  $A - B$  and  $A + B$  values to determinate the shape of the ellipse. Each radius is positive if the curvature centre is inside the body. Most of the time, the wheel rolling radius and the rail transverse radius are positive (convex). However, the wheel transverse radius at the contact point can be positive (convex) or negative (concave).

### 4. Particular Case: 2D Contact

When the wheel and the rail are conformal, they have the same transversal curvature value with opposite sign, giving  $B = 0$ . In this case, the contact is described as 2D, the 3D model being the classical ellipse.

A practical way to manage these slender ellipses in a software algorithm is to limit the  $g$  value, or  $A/B$ , or  $b/a$ . In the later step of the contact tangent forces calculation, analytical expressions are given in engineering books for  $\theta = 0$ , but the  $g$  values extrapolated in the intermediate interval will cause problems.

### 5. Contact Pressure

With an elliptical pressure distribution, the mean pressure being  $N/\pi ab$ , the maximum pressure will be simply:

$$\sigma_{\max} = 1.5 N/\pi ab$$

In the railway field, the maximum contact pressure is frequently over 1000 MPa. This value is over the elastic limit of most steels, but the compression state is more complex than a simple tensile test and the elastic limit is not reached. The determination of the plastification (which is a limit of the Hertzian hypothesis) must be calculated with a criteria based on the hydrostatic stress (Von Mises...).

**TABLE 4.1**  
**Hertz Coefficients ( $A/B < 1$ )**

$\theta^\circ$	90	80	70	60	50	40	30	20	10	0
$g = n/m$	1	0.7916	0.6225	0.4828	0.3652	0.2656	0.1806	0.1080	0.0470	0
$m$	1	1.128	1.285	1.486	1.754	2.136	2.731	3.816	6.612	$\infty$
$n$	1	0.8927	0.8000	0.7171	0.6407	0.5673	0.4931	0.4122	0.3110	0
$r$	1	0.9932	0.9726	0.9376	0.8867	0.8177	0.7263	0.6038	0.4280	0



**TABLE 4.2**  
**Hertz Coefficients for  $\theta = 0$  to  $180^\circ$**

$\theta^\circ$	0	5	10	30	60	90	120	150	170	175	180
$A/B$	0	0.0019	0.0077	0.0717	0.3333	1	3.0	13.93	130.6	524.6	$\infty$
$b/a = n/m$	0	0.0212	0.0470	0.1806	0.4826	1	2.0720	5.5380	21.26	47.20	$\infty$
$m$	$\infty$	11.238	6.612	2.731	1.486	1	0.7171	0.4931	0.311	0.2381	0
$r$	0	0.2969	0.4280	0.7263	0.9376	1	0.9376	0.7263	0.4280	0.2969	0

**B. TABLES, POLYNOMIAL EXPRESSIONS AND FASTER METHODS**

**1. Continuous Expression of the Tables**

The traditional table limited to  $1 > g > 0$  leads to the necessity of determining if the  $A/B$  value is more or less than 1, and to switch between  $g = b/a$  or  $g = alb$ , while at the same time  $g = n/m$  becomes  $g = m/n$ , which is a bit confusing.

This ambiguity can be avoided with a description of the tables in the interval:  $0 < g < \infty$ . Then, in all the cases,  $g = b/a = n/m$

The expression of the elliptic integrals used to calculate the tabulated values of  $n$  and  $m$  have the property:

$$n(A/B) = m\left(\frac{1}{A/B}\right) \tag{4.7}$$

The traditional Hertz table can be rewritten as shown in Table 4.2.

With this presentation,  $A/B$  can be used directly as the input of the tables, instead of  $\cos \theta$ .

**2. Analytical Approximation of the Tables**

Rather reliable expressions of the  $n/m$  and  $mn$  values function of  $A/B$  ( $0, \infty$ ) can be found with:

$$\frac{b}{a} = \frac{n}{m} \approx \left(\frac{A}{B}\right)^{0.63} \quad (mn)^{3/2} \approx \left(\frac{1 + A/B}{2\sqrt{A/B}}\right)^{0.63} \tag{4.8}$$

These expressions have the advantage of being continuous, simple, and fast to calculate.

The exponent 0.63 is a compromise. It is equal to two thirds when  $A/B$  is close to 1 but the value 0.63 will better describe the slender ellipses: the difference from the tabulated values is approximately  $\pm 5\%$  between  $b/a = 1/25$  and  $b/a = 25$  (see Table 4.3).

**TABLE 4.3**  
**Approximation of the  $n/m$  Values**

$\theta^\circ$	0	5	10	30	60	90	120	150	170	175	180
$A/B$	0	0.0019	0.0077	0.0717	0.3333	1	3.0	13.93	130.6	524.6	$\infty$
$b/a = n/m$	0	0.0212	0.0470	0.1806	0.4826	1	2.0720	5.5380	21.26	47.20	$\infty$
$(A/B)^{0.63}$	0	0.0193	0.0466	0.1901	0.5005	1	1.9980	5.2564	21.530	51.700	$\infty$

### C. APPLICATION TO THE RAILWAY FIELD

#### 1. Contact Plane Angle, Conicity

With a perfectly conical wheel, the contact place on the rail appears to be the point with the same slope as on the wheel, measured in the *YOZ* frame (Figure 4.2).

For a perfectly conical wheel of 1:20, and with new rails tilted at 1:20, this gives a contact point in the middle of the rail (see Table 4.4).

#### 2. A Note on Conicity

There is such a large set of wheel profiles that the main cone of the wheel is not able to describe the wheel–rail contact; for a large part of the time, the contact is not on the conical part of the wheel. The mean value is considered to be the conicity. Its calculation does consider the wheel profile, but also the rail profile and the flangeway clearance.

#### 3. Normal Load

On the tread, with a low conicity, the normal load *N* has practically the same value as the vertical load on the wheel, for example with a cone angle of 1:20:

$$N = Q \cos(\text{atan}(0.05)) = 0.9988 Q \tag{4.9}$$

However, a distinct reference frame must be considered and is noted: *z* for the normal and *x*, *y* for the tangent directions (Figure 4.2).

#### 4. Determination of the Longitudinal Curvature of the Wheel

This curvature is referred to as “longitudinal” because the rolling circle of a cylindrical wheel would be in the *XOZ* plane. However, the wheel is generally conical, the contact angle  $\gamma$  is not zero, the *A*<sub>1</sub> curvature in the *xOz* plane differs from the rolling radius *r*<sub>o</sub>. The intersection between *xOz* and the wheel cone is an ellipse with one focus situated on the wheelset axis, and the curvature radius becomes (Figure 4.5):

$$\frac{1}{r_n} = \frac{\cos \gamma}{r_o} \tag{4.10}$$

Note that the *r*<sub>o</sub> value at the contact point is slightly different between the tread and the flange, but this variation (+ 10 to + 15 mm) is a second order compared to the  $\cos \gamma$  influence.

#### 5. Contact Point between the Wheel and Rail Profiles

In order to determine the contact point, the wheel and rail profiles are placed relative to each other, wheelset centred on the track as a function of the wheelset gauge *D*<sub>w</sub> and rail gauge *D*<sub>r</sub>, then moved

**TABLE 4.4**  
**Rail Inclinations and Wheel Cone Values**

	Intercity and Freight	Metros	Tramway
Wheel cone	1/20th to 1/40th for HST	1/20th	1/20th to flat
Rail tilt	1/20th France, England 1/30th Sweden 1/40th Germany	1/20th	Flat
Wheelset load	22.5 t/wheelset but 17 t/wheelset HST	10 t/wheelset	6 t/wheelset

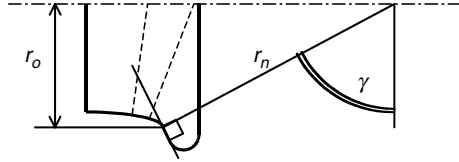


FIGURE 4.5 Determination of the longitudinal curvature from the rolling radius.

laterally with the translation, noted here as  $t_y$ . In this position, the minimal vertical distance between the wheel and the rail defines the contact point  $O$ .

Practically, the profile description, the digitisation, and smoothing operations will have an influence on the accuracy of the position.

6. Direct Comparison of the Antagonist Curvatures

$B$  cannot have a negative value. A simple comparison of the corresponding curvatures (Figure 4.6) gives information on the possibility of having the centre of a Hertzian contact in this zone.

In the transition zone on the wheel between the tread and the flange, the contact is frequently impossible: the sum  $1/R_{wx} + 1/R_{rx}$  must be always positive to have contact.

However, if the centre of a contact cannot be located in this area, it can be covered partially by the contact ellipses.

7. One or Two Wheel–Rail Pairs?

As a first step, the study of a single wheel–rail pair is enough to give a large set of information on the contact, as a function of the lateral relative displacement  $t_y$ .

In a more complex algorithm, it is also possible to take into account the roll effect of the other wheel–rail pair of the wheelset, in which case the tables are functions of  $t_y$  and roll.

In the same way, the yaw angle has an influence on the longitudinal position of the contact on the wheel, this can be taken into account with multiple parameters input.

Generally, these effects do not have a very strong influence on the wheelset equilibrium.

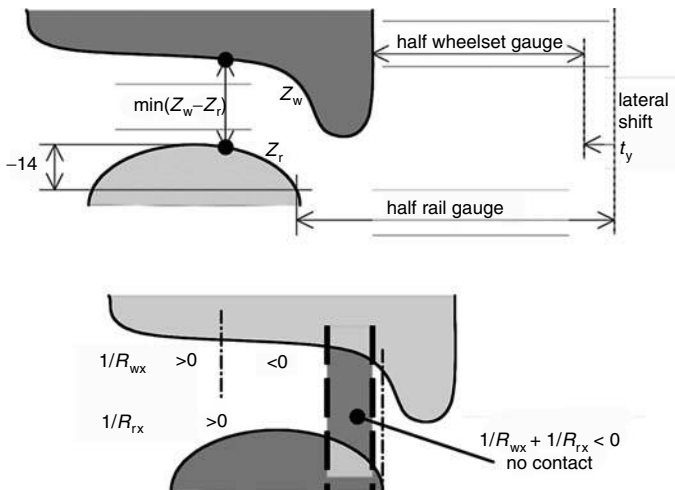


FIGURE 4.6 Corresponding curvatures between the wheel and the rail.

### III. THE TANGENT PROBLEM

#### A. FORCES AND COUPLES ON A WHEELSET

The kinematic representation of the wheelset (Klingel formula, Appendix 4.2) has, for a long time, been used to explain the sinusoidal behaviour of a free wheelset, but the situation is different under a real vehicle.

The real wheelset is strongly linked to the vehicle through flexible suspension elements, and these links creates significant forces when the wheelset is entering a curve or running on a real track with irregularities.

The suspension forces find their reaction forces (normal and tangent) at the rail–wheel contact interface, where the tangent components or creep forces are related to the relative speed between the two bodies: the creepages.

In the contact coordinate systems, the forces are denoted:

- $N$  for the normal forces
- $F_x$  for the longitudinal creep force
- $F_y$  for the lateral creep force in the contact plane

The  $F_y$  forces must be projected on the track plane  $OY$  and summed to give the guiding force.

#### 1. Approximations

The main couple exerted on a rigid wheelset around  $OZ$  is coming from the two nearly opposite longitudinal forces  $F_x$ , separated laterally from the contact distance  $D_c$ :

$$M_z = -\left(F_{xl} \frac{D_c}{2} - F_{xr} \frac{D_c}{2}\right) \tag{4.11}$$

The spin creepage, due to the relative rotation of the contact patches around the normal axis to the contact, generates a couple ( $\varphi$  in Figure 4.7), but it can be neglected in comparison with the

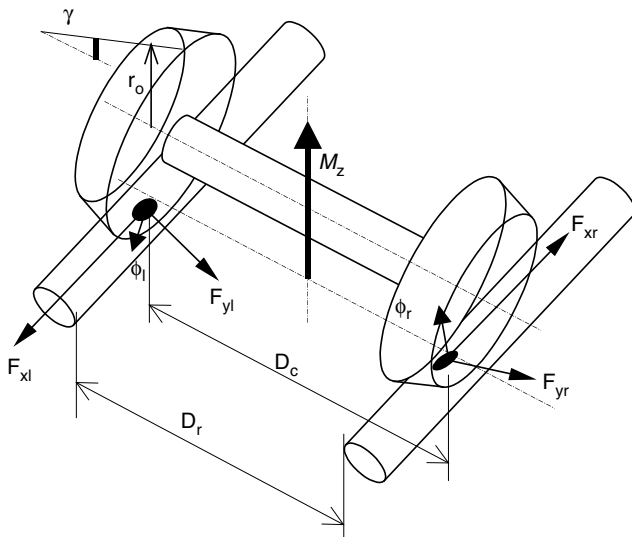


FIGURE 4.7 Wheelset geometry and creep forces.

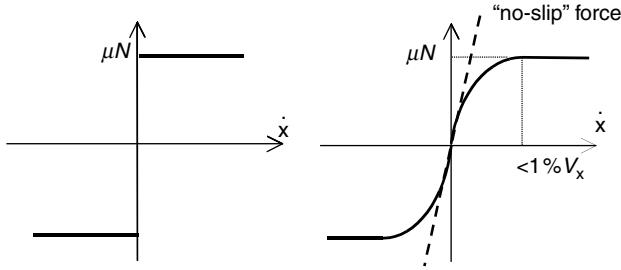


FIGURE 4.8 Coulomb’s model and rolling friction model.

longitudinal forces couple  $M_z$ . However, the spin generates a lateral force which is not negligible when the contact angle becomes higher. This lateral force is described separately or included with the yaw force, depending on the theory used.

The couple  $\Sigma\varphi$  has a component resisting the forward displacement, which can be neglected in a first approximation.

**B. TANGENT FORCES: ROLLING FRICTION SIMPLE MODELS**

The wheel–rail contact is a rolling friction contact. It differs from the sliding friction Coulomb model (which can be found for example in the brake shoes) with an area of adhesion and an area of slip which appears progressively as the slip speed increases (see Figure 4.8).

The transition is characterized by the initial slope or “no-slip force,” the force if the friction coefficient  $\mu$  was infinite, and the “S” saturation curve.

**1. Historical Review**

After Hertz, Boussinescq, and Cerruti, at the beginning of the 20th century three authors were interested by the wheel–rail contact modelisation: Carter and Fromm for longitudinal models,<sup>1</sup> and Rocard for the lateral force.

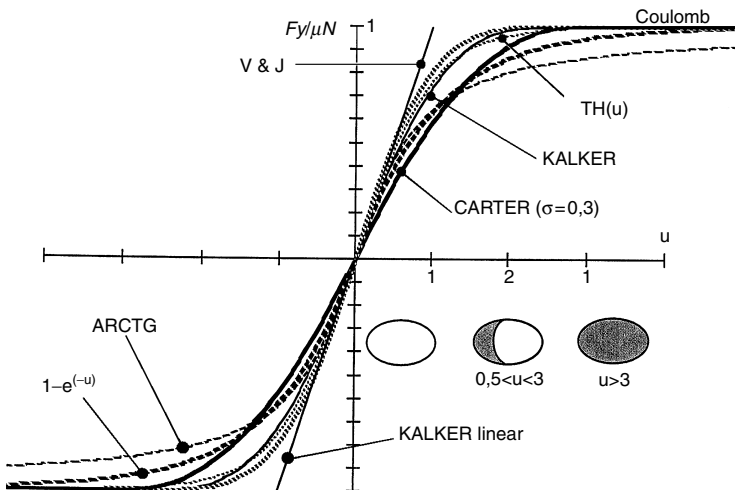


FIGURE 4.9 “Heuristic” expressions used for the saturation and physical meaning of the different parts.

Carter described a simple 2D contact surface, but he was the first to give a rather adequate expression of the force relative to the creepage in the longitudinal direction. His method of describing the stresses in the adhesive zone (see Figure 4.9) was used until the 1960s. Fromm made similar observations. Rocard described the linear relationship between the yaw angle and the guiding force, for rubber tyres and for railway wheels, in the lateral direction. He was particularly interested in the equivalent of the bogie hunting for cars: the shimmy phenomenon.

In the 1960s, more experimental data were available, the definitive expressions were established mainly by Johnson and Kalker, who gave an expression of the creepage stiffness introducing variable coefficients depending on the  $b/a$  ratio of the contact ellipse. This expression is the most common today.

### C. LINEAR EXPRESSIONS OF THE CREEP FORCES

In the case of a Hertzian contact, the creep forces are a function of the relative speeds between rigid bodies near the contact point, the creepages.

The general expression of the creep forces takes into account stiffness coefficients  $c_{ij}$  expressed in the linear theory of Kalker<sup>3</sup> by:

$$\begin{aligned} F_x &= -G abc_{11} \nu_x \\ F_{y_{\text{yaw}}} &= -G ab c_{22} \nu_y \\ F_{y_{\text{spin}}} &= G ab c_{23} c \varphi \quad (\text{with } c = \sqrt{ab}) \end{aligned} \tag{4.12}$$

where  $G$  is the material shear modulus (steel in the railway case);  $\pi ab$  is the contact ellipse surface; and  $c_{ij}$  are the coefficients given in Ref. 4.

### D. DEFINITION OF CREEPAGES

#### 1. Quasi-Static Creepages

A general expression for two rolling bodies can be given by the projection of the speed vectors on  $Ox$ ,  $Oy$ , and  $Oz$ :

$$\begin{aligned} \text{–longitudinal} \quad \nu_x &= \frac{\text{proj.}/x(\vec{V}_0 - \vec{V}_1)}{\frac{1}{2}(\vec{V}_0 + \vec{V}_1)} && \text{(dimensionless)} \\ \text{–lateral} \quad \nu_y &= \frac{\text{proj.}/y(\vec{V}_0 - \vec{V}_1)}{\frac{1}{2}(\vec{V}_0 + \vec{V}_1)} && \text{(dimensionless)} \\ \text{–rotation} \quad \varphi &= \frac{\text{proj.}/z(\vec{\Omega}_0 - \vec{\Omega}_1)}{\frac{1}{2}(\vec{V}_0 + \vec{V}_1)} && (1/\text{m}) \end{aligned} \tag{4.13}$$

$V_0$  and  $V_1$  are the absolute speed at the contact,  $1/2(\vec{V}_0 + \vec{V}_1)$  is the mean speed.

$\omega_0$  and  $\omega_1$  are the angular speed of the two solids  $\Omega_i = V_i/r_i$  projected on the normal to the contact (Figure 4.10).

#### 2. Quasi-Static Creepages in the Railway Case

The above general expressions are useful for the test rigs used in research, however the railway case leads to simplified expressions.

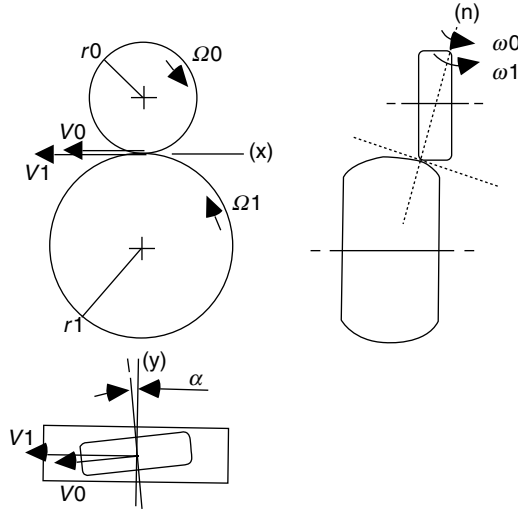


FIGURE 4.10 General geometry and creepages.

### 3. Longitudinal Creepage

When the wheelset is rolling freely without traction or braking, due to the conicity, the two different rolling radius  $r_o \pm \Delta r$  generates the two opposite forces  $f_x$  and  $-f_x$ ;  $r_o$  being the mean rolling radius, the static longitudinal creepage can be described with:

- In the general case in quasi-static conditions with small creepages,  $V_x \approx r_o \omega$  and  $1/2(V_x + r\omega) \approx V_x$  then, for the left wheel (respectively with the opposite sign on the right wheel):

$$v_{xl} = \frac{V_x - r\omega}{\frac{1}{2}(V_x + r\omega)} \approx \frac{V_x - (r_o + \Delta r)\omega}{V_x}$$

$$v_{xl} = -\Delta r/r_o \tag{4.14}$$

- In the case of a perfectly conical wheelset:

$$\text{At left : } \Delta r = \gamma y$$

$$\text{Then } v_{xl} = -\gamma y/r_o \quad v_{xr} = -v_{xl}$$

Rather than the cone value, the equivalent conicity  $\gamma_e$  can be used for simplified models (see [Appendix 4.1](#)).

### 4. Lateral Creepage

The lateral creepage in quasi-static conditions, with small creepages, is simply the yaw angle common to the two wheels:

$$v_y = -\alpha \tag{4.15}$$

## 5. Spin Creepage

In the quasi-static case, the rail speed  $\Omega$  is equal to zero and the general expression is simplified. The spin creepage  $\varphi$  is:

$$\varphi = \sin \gamma / r_o \quad (4.16)$$

( $\gamma$  is an algebraic value, different for the two wheels)

This expression shows that the spin has an important value when flanging, and a larger value with small radius wheels.

## 6. Dynamic Formulation of the Creepages

For the dynamic forces at the contact level, a formulation identical to the other dynamic links is established with an elastic term and a damping term.

The expressions of the dynamic creepages contain the speed terms whose sign is opposed to the elastic deformation, which will contribute to damping:

$$\begin{aligned} \nu_{xl} &= -\left(\frac{\Delta r}{r_o} + \frac{D_c}{2} \frac{\dot{\alpha}}{\dot{x}}\right) \quad \text{for the left wheel} \\ \nu_{xr} &= -\nu_{xl} \quad \text{for the right wheel} \\ \nu_y &= \frac{\dot{y}}{\dot{x}} - \alpha \\ \varphi &= \frac{\sin \gamma}{r_o} - \frac{\dot{\alpha}}{\dot{x}} \cos \gamma \quad (2\text{nd term} \ll 1\text{st term}) \end{aligned} \quad (4.17)$$

where:

- $\alpha$ : is the wheelset yaw angle relative to the rail (radians)
- $r_o$ : is the mean rolling radius of the wheel
- $y$ : is the lateral displacement of the wheel relative to the centred position
- $\dot{\alpha}, \dot{y}$ : is the relative speed in the track reference system
- $\dot{x} = V_x$ : is the running speed of the wheelset along the curvilinear abscissa

## 7. Damping Terms and Stability

In the creepage, the damping terms  $\dot{y}/\dot{x}$  and  $\dot{\alpha}/\dot{x}$  are inversely proportional to the forward speed. This means that these terms are reduced as the speed increases, and the wheelset becomes unstable.

## 8. Nondimensional Spin Creepage

The creepages are relative slips, they have no dimension.

The last spin expression, to become dimensionless, must be multiplied by a distance. The characteristic dimension of the ellipse  $c = \sqrt{ab}$  is used to obtain the spin creepage in a homogeneous form<sup>5</sup>:

$$c\varphi = c\left(\frac{\sin \gamma}{r_o} - \frac{\dot{\alpha}}{\dot{x}} \cos \gamma\right) \quad (4.18)$$



**TABLE 4.5**  
Kalker's coefficient tables (from Ref. 4)

g	C <sub>11</sub>			C <sub>22</sub>			C <sub>23</sub> = -C <sub>32</sub>			C <sub>33</sub>			
	σ = 0	1/4	1/2	σ = 0	1/4	1/2	σ = 0	1/4	1/2	σ = 0	1/4	1/2	
0.0	π <sup>2</sup> /4(1 - σ)			π <sup>2</sup> /4 = 2,47			π√g/3	—	—	π <sup>2</sup> /16(1-σ)g			
a/b	0.1	2.51	3.31	4.85	2.51	2.52	2.53	0.334	0.473	0.731	6.42	8.28	11.7
	0.2	2.59	3.37	4.81	2.59	2.63	2.66	0.483	0.603	0.809	3.46	4.27	5.66
	0.3	2.68	3.44	4.80	2.68	2.75	2.81	0.607	0.715	0.889	2.49	2.96	3.72
	0.4	2.78	3.53	4.82	2.78	2.88	2.98	0.720	0.823	0.977	2.02	2.32	2.77
	0.5	2.88	3.62	4.83	2.88	3.01	3.14	0.827	0.929	1.07	1.74	1.93	2.22
	0.6	2.98	3.72	4.91	2.98	3.14	3.31	0.930	1.03	1.18	1.56	1.68	1.86
	0.7	3.09	3.81	4.97	3.09	3.28	3.48	1.03	1.14	1.29	1.43	1.50	1.60
	0.8	3.19	3.91	5.05	3.19	3.41	3.65	1.13	1.25	1.40	1.34	1.37	1.42
	0.9	3.29	4.01	5.12	3.29	3.54	3.82	1.23	1.36	1.51	1.27	1.27	1.27
b/a	1.0	3.40	4.12	5.20	3.40	3.67	3.98	1.33	1.47	1.63	1.21	1.19	1.16
	0.9	3.51	4.22	5.30	3.51	3.81	4.16	1.44	1.59	1.77	1.16	1.11	1.06
	0.8	3.65	4.36	5.42	3.65	3.99	4.39	1.58	1.75	1.94	1.10	1.04	0.954
	0.7	3.82	4.54	5.58	3.82	4.21	4.67	1.76	1.95	2.18	1.05	0.965	0.852
	0.6	4.06	4.78	5.80	4.06	4.50	5.04	2.01	2.23	2.50	1.01	0.892	0.751
	0.5	4.37	5.10	6.11	4.37	4.90	5.56	2.35	2.62	2.96	0.958	0.819	0.650
	0.4	4.84	5.57	5.57	4.84	5.48	6.31	2.88	3.24	3.70	0.912	0.747	0.549
	0.3	5.57	6.34	7.34	5.57	6.40	7.51	3.79	4.32	5.01	0.868	0.674	0.446
	0.2	6.96	7.78	8.82	6.96	8.14	9.79	5.72	6.63	7.89	0.828	0.601	0.341
0.1	10.7	11.7	12.9	10.7	12.8	16.0	12.2	14.6	18.0	0.795	0.526	0.228	

**E. KALKER'S COEFFICIENTS  $c_{ij}$**

The  $c_{ij}$  coefficients are a given function of the  $b/a$  ratio of the ellipse (Table 4.5). Their values are not far from  $\pi$  for  $b/a$  close to 1 (see Kalker's tables). Initially, Carter uses the value  $\pi$ .

In the bibliography, the  $c_{11}$  and  $c_{22}$  values are given for Poisson's ratio values 0 or 0.25 or 0.5. The typical value of steel is close to 0.27 and the tables must be interpolated.

A polynomial fit is proposed by:

$$c_{11} = 3.2893 + \frac{0.975}{b/a} - \frac{0.012}{(b/a)^2} \tag{4.19}$$

$$c_{22} = 2.4014 + \frac{1.3179}{b/a} - \frac{0.02}{(b/a)^2}$$

$$c_{23} = 0.4147 + \frac{1.0184}{b/a} + \frac{0.0565}{(b/a)^2} - \frac{0.0013}{(b/a)^3}$$

Before using these expressions it is necessary to limit the  $b/a$  ratio to an interval (i.e., 1/25 to 25), the fit becoming wrong for very slender  $A/B$  ratio. It is possible to express the  $c_{ij}$  directly from the curvature ratio  $A/B$  (see Section II.B).

## 1. $c_{ij}$ Values for Simplified Bogie Models

In some simplified bogie models with equivalent conicity (Appendix 4.1), a constant value is taken for  $c_{11}$  and  $c_{22}$ . The contact is assumed to be on the tread at all times, with values varying from 3 to 5.

When the flange contact is considered, the range growth to  $a/b = 20$  ( $c_{22}$  reaches 12.8 when  $a/b = 10$ ). However, the creepages are extremely large and the creep force can be considered saturated.

## F. CREEP FORCES IN THE LINEAR DOMAIN

### 1. Dependence on Load

The three previous expressions, considering a Hertzian contact, can be normalised with the load, implicit in the contact dimension  $a$  and  $b$ :

$$a = a_1 N^{1/3} \quad b = b_1 N^{1/3}$$

giving:

$$ab = a_1 b_1 N^{2/3} \quad abc = a_1 b_1 c_1 N \quad (4.20)$$

$a_1$  and  $b_1$  (respectively  $c_1$ ) are the semi-axes for 1 Newton load. The previous expressions become:

$$F_x = -G a_1 b_1 c_{11} N^{2/3} \nu_x$$

$$F_{y_{yaw}} = -G a_1 b_1 c_{22} N^{2/3} \nu_y$$

$$F_{y_{spin}} = G a_1 b_1 c_{23} c_1 N \varphi \quad (4.21)$$

The spin force is proportional to the load and the pressure  $N/\pi ab$  is proportional to  $N^{1/3}$ .

### 2. Creepages Combinations and Saturation

In the above expression, the transversal force due to the spin  $\varphi$  is separated from the force due to the lateral creepage  $\nu_y$ . The total transversal force is:

$$F_y = F_{y_{yaw}} + F_{y_{spin}} \quad (4.22)$$

In the case of negligible spin, i.e., in models considering mainly the tread contact, these expressions can be used and the lateral force due to the spin can be added to the yaw force, or even neglected.

In the case of combined creepages, when the spin is not negligible, these independent expressions are not adequate because of the nonuniform combined saturation of the shear contact stresses inside the contact area. A model based on the surface description is necessary; the most commonly used being the FASTSIM model due to Kalker.

### 3. Using Linear Models: The C110, C220 Stiffnesses

Some models need to establish the precalculated stiffness coefficients (mean or variable values). These parameters have been standardised by the UIC in the two main directions by the expressions:

$$C110 = G a b c_{11} \quad C220 = G a b c_{22}$$

It is useful to normalise them with the normal load  $N$ :

$$C110 = G a_1 b_1 c_{11} N^{2/3} \quad C220 = G a_1 b_1 c_{22} N^{2/3}$$

Note that these expressions are neglecting the spin effect.

### 4. Reduced Creepages

In the following expressions, it is useful to introduce a reduced parameter which will be used in a great number of equations:

$$u_x = \frac{Gabc_{11}v_x}{\mu N} \quad u_y = \frac{Gabc_{22}v_y}{\mu N} \quad u = \sqrt{u_x^2 + u_y^2} \quad (4.23)$$

These parameters are called the no-slip reduced friction forces, they are characteristic of both the stiffness of the contact, and the creepage.

## G. SATURATION LAWS

The first dynamic software have used the linear stiffnesses expressions in order to calculate the critical speed of a bogie with eigenvalues.

The first time domain resolution programs introduce the saturation as an evolution of the linear expressions with analytical formulations.

### 1. Vermeulen and Johnson

Vermeulen and Johnson's law<sup>4</sup> is a function of a reduced creepage coefficient neglecting the spin:

$$\tau = \sqrt{\tau_x^2 + \tau_y^2}:$$

$$\begin{aligned} \frac{F}{\mu N} &= 1 - (1 - \tau)^3 && \text{from } \tau = 0 \text{ to } \tau = 1 \\ \frac{F}{\mu N} &= 1 && \text{for } \tau > 1 \end{aligned} \quad (4.24)$$

with:

$$\tau_x = \frac{Gabc_{11}v_x}{3\mu N} \quad \tau_y = \frac{Gabc_{22}v_y}{3\mu N}$$

or

$$\tau_x = \frac{u_x}{3} \quad \tau_y = \frac{u_y}{3}$$

Vermeulen and Johnson's proposition is close to Carter's.

## 2. Kalker's Empirical Proposition

If the expression<sup>4</sup> is limited to the lateral creepage and force modulus, a comparison is possible; using the reduced parameter  $\tau_y = u_y/3$  :

$$\frac{F_y}{\mu N} = \left( \frac{3}{2} \tau_y \cos^{-1} \tau_y \right) + \left( 1 - \left( 1 + \frac{\tau_y^2}{2} \right) \sqrt{1 - \tau_y^2} \right) \quad \text{from } \tau_y = 0 \text{ to } \tau_y = 1$$

$$\frac{F_y}{\mu N} = 1 \quad \text{for } \tau_y > 1 \quad (4.25)$$

The expression is more complex if both directions are considered. Separating the two parts of his formulation, Kalker introduces the difference between the forces ( $e_2$ ) and the slip direction ( $e_1$ , see Ref. 6) as the creepage increases.

The original slope C220<sub>K</sub> calculated by Kalker is almost different from Johnson's; both are close to Johnson's experiments.

## 3. Exponential Saturation Law: CHOPAYA, Ohyama and Others

Ohyama,<sup>7</sup> and later Ayasse–Chollet–Pascal (under the name CHOPAYA), proposed a classical exponential saturation starting from forces measurement in railways conditions:

$$F/\mu N = 1 - e^{-u} \quad (4.26)$$

Some other saturation expressions have been proposed, such as the hyperbolic tangent, or the arctangent. All these laws (Figure 4.9) are "heuristic", they are fitted on measured data,<sup>8</sup> respecting the  $c_{ij}$ , but do not correspond to a physical saturation model.

Neglecting the spin, the general mechanism of saturation can be divided into three steps. The linear zone is a full adhesion surface. The saturated case is slipping everywhere with the dry friction Coulomb value, the intermediate zone is a partially saturating surface, where the slipping area is always at the rear side of the ellipse.

Qualitatively, when there is no spin, the slip-stick frontier presented in Figure 4.9 looks like a moon quarter propagating from the rear to the front with increasing creepage,  $\nu_x$  or  $\nu_y$ . Between these creepages, only the stress directions are different.

The situation is more complex with the presence of spin. To determine the shear stresses quantitatively, it is necessary to use a physically based model, the most important for the railway use is FASTSIM from Kalker, it is presented in the following paragraph.

## H. SURFACE-BASED CONTACT MODELS

Kalker proposed several methods to solve the contact problem with models based on the surface description; these methods are widely described in his book<sup>6</sup> and only the simplest one is briefly described here. Both CONTACT and FASTSIM algorithms are based on the "strip theory" originally proposed by Haines and Ollerton (Figure 4.11), and extended to the three creepages.<sup>5</sup>

CONTACT is a program based on the complete theory of elasticity, it can take into account several body shapes, including the railway case. Several methods are available to calculate the tangent stresses and/or the internal stresses. However, the calculation of one case takes several seconds, and it is limited to half space bodies.

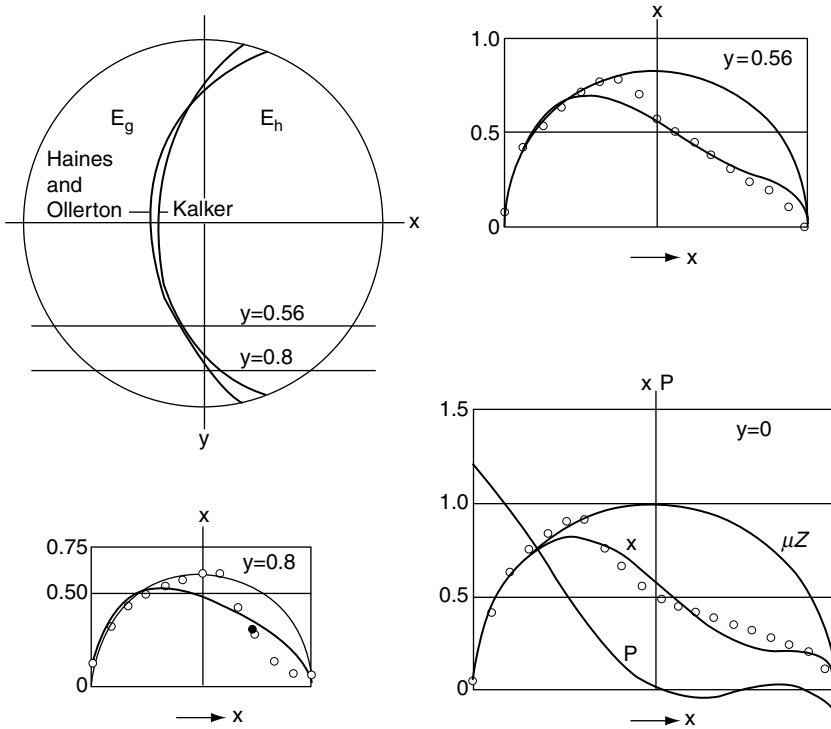


FIGURE 4.11 The separatrix experiments from Haines and Ollerton and proposition from Kalker (from Kalker’s thesis pp. 132–136).

**1. Discretised Ellipse: FASTSIM from Kalker**

FASTSIM, originally a Fortran subroutine,<sup>9</sup> is based on the “simplified theory” and, in the original publication, is limited to ellipses which mean that some assumptions are common to Hertz:

1. The contact surface is elliptic and flat, the pressure  $p_z$  is an ellipsoid
2. The creepages are estimated at the ellipse centre
3. Kalker’s coefficients  $c_{ij}$  are constant everywhere in the ellipse, their values are deduced from  $A/B$  or  $b/a$
4. The elliptic contact surface is divided into independent longitudinal parallel strips of length  $a_i$  –  $a_i$  and width  $\Delta y_i$
5. All the strips are divided into the same number of elements, the stress calculation begins from the leading edge, from element to element
6. The method is simplified: a local deformation corresponds to a local force
7. The saturation is calculated independently for each element loaded by the normal force  $n_{ij}$

Practically, the surface is described by a grid separating parallel strips in the direction of rolling. Due to the elliptical shape, the elements do not have the same length  $a_i/MX$  (Figure 4.12). Internal creepages are computed for each element, starting from the central creepage. The pressure and elementary forces are considered in the centre of each element. The pressure is defined by the ellipsoid value in this point.

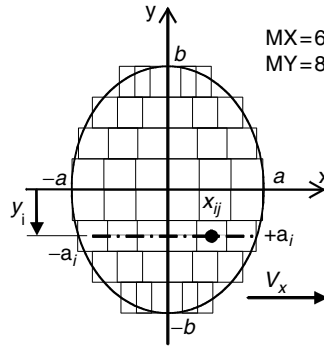


FIGURE 4.12 Strips and elements discretisation with FASTSIM.

**2. Stresses**

The unsaturated stress distribution in the  $x$  and  $y$  directions<sup>a</sup> is the following (simplified or “linear” theory,<sup>9</sup>):

$$p_x(x, y_i) = \left( \frac{v_x}{L_1} - y_i \frac{\varphi}{L_3} \right) (x - a_i) \tag{4.27}$$

The first term represents the mean rigid longitudinal slip, and the second term is the spin effect as a local rigid slip at the point  $(x, y)$  in the strip,  $a_i$  being the leading edge of this strip.

$$p_y(x, y) = \frac{v_y}{L_2} (x - a_i) + \frac{\varphi}{2L_3} (x^2 - a_i^2) \tag{4.28}$$

$L_1, L_2, L_3$  are the elasticity coefficients (or flexibilities) of the contact.<sup>10,9</sup> In the exact complete theory they are:

$$L_1 = \frac{8a}{3c_{11}G} \quad L_2 = \frac{8a}{3c_{22}G} \quad L_3 = \frac{\pi a \sqrt{a/b}}{4c_{23}G} \tag{4.29}$$

**3. Linear Contact Forces, Elastic Coefficients**

The contact forces are the integral distributions of the surface stresses. When there is no saturation:

$$F_x = - \iint p_x(x) dx dy = \frac{-8a^2 b v_x}{3L_1}$$

$$F_y = - \iint p_y(x) dx dy = \frac{-8a^2 b v_y}{3L_2} - \frac{\pi a^3 b \varphi}{4L_3} \tag{4.30}$$

<sup>a</sup> The  $xOy$  reference frame here is the plane containing the ellipse.

### 4. Reduced Creepages in FASTSIM

FASTSIM is composed of two parts; the main program calls the subroutine SR in which the stresses are computed in loops.

Practically, starting from the physical creepages, the four inputs of SR are:

$$\begin{aligned}
 UX &= 0.589 \text{ Gabc}_{11} v_x / \mu N = \frac{3}{8} \frac{\pi}{2} u_x \\
 UY &= 0.589 \text{ Gabc}_{22} v_y / \mu N = \frac{3}{8} \frac{\pi}{2} u_y \\
 FIY &= 2(0.589) \text{ Gab}^{3/2} c_{23} \varphi / \mu N \\
 \text{FIX} &= b/a \text{ FIY}
 \end{aligned}
 \tag{4.31}$$

where it can be seen that 0.589 is a numerical coefficient  $3\pi/16$  coming from the  $L_1, L_2, L_3$  parameters.

The SR algorithm is computed in a normalised form (ellipse reduced to a circle of radius one). The linear contact forces must be normalised by the load and the friction coefficient, giving the reduced creepages already introduced in 2.6:

$$\begin{aligned}
 \frac{F_x}{\mu N} &= \frac{\text{Gabc}_{11}}{\mu N} v_x = u_x \\
 \frac{F_y}{\mu N} &= \frac{\text{Gabc}_{22}}{\mu N} v_y - \text{Gabc}_{23} \sqrt{\frac{b}{a}} \varphi = u_y - u_\varphi
 \end{aligned}
 \tag{4.32}$$

The stresses are independent from one strip to the other, but they are not independent inside a strip. The calculation is shown in Figure 4.13 on a strip; on the right side, the pure spin case is presented without any longitudinal stress, even from the spin effect, in order to simplify the presentation.

The calculation begins with a deformation and stress at zero on the leading edge of each strip:  $+a_i$ . The stress is incremented as a function of the stresses functions, by example in the direction  $O_y$ :

$$\Delta p_y(x, y) = \frac{v_y}{L_2} \Delta x - \frac{\varphi}{L_3} \times \Delta x
 \tag{4.33}$$

This increment is constant for the lateral creepage but variable for the spin.

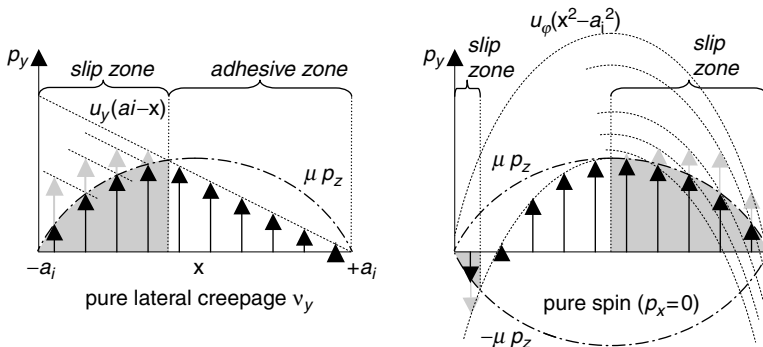


FIGURE 4.13 Internal saturation of the stresses inside FASTSIM; simple cases.

When the incremented stress reaches more than  $\mu p_z$  (shadowed arrows on [Figure 4.13](#)), it is saturated and the new step starts from this correct value (black arrows).

The pure spin calculation shows that an adhesive zone is persistent after the centre of the contact. It can disappear with the other creepages.

## 5. Extensions of FASTSIM

FASTSIM is not limited to Hertzian cases, a contour and the corresponding pressure distribution can be inserted coming from a model other than Hertz. In this case it is necessary to adapt the creepages description and to adapt the  $L_i$  parameters.

In the original FASTSIM algorithm, there is no calculation of the dissipated power because there is no calculation of the slip speed. It is possible to estimate this speed, increasing from the first slipping element. The wear factor:

$$F_x v_x + F_y v_y \quad (4.34)$$

can be replaced by the calculation carried out element by element with FASTSIM, the spin will then be taken into account.

It can be a criteria to compare different vehicles going into the same curve at the same speed. It must be used to give access to wear simulation.

However, FASTSIM is not dedicated to accurate calculations of the dissipated power. The program CONTACT from Kalker is more adapted to such studies. The wear coefficients are themselves very dependent on lubrication and slip speed. They are adjusted from experiments.

## I. METHODS BASED ON FASTSIM

In order to calculate faster than FASTSIM while obtaining similar results, several methods have been developed. Kalker himself proposes to interpolate with tables based on FASTSIM results. The two software NUCARS and VAMPIRE are interpolating large pretabulated tables.

Additional methods to those presented in Section III.F are the Shen-Hedrick-Elkins formulation, and more recently the method proposed in Ref. 11.

## IV. CONTACT FORCES IN THE RAILWAY CONTEXT

### A. FROM THE PURE DICONE TO THE WHEELSET WITH REAL PROFILES

For a complete understanding of the wheel–rail forces, a description of the cross sectional profiles of the wheel and rail are required.

#### 1. Equivalent Conicity

The assimilation of these profiles to a cone rolling on a cylinder has been useful for quasi-static curving but cannot be used when the contact is moving on the wheel tread or the rail head, or to predict stability or derailment.

The first linearisation was to consider only the wheel tread dicone.

The “circle theory” (see [Appendix 4.3](#)) has been an improvement of the dionic formulation in a more realistic case: both rail and wheel profiles are considered as circles, giving an additional term to the yaw couple. Today, it is easier to take into account the real profiles and this second method has become obsolete.



## 2. Variable Conicity

The real conicity to be taken into account has a variable value. On the wheel, the contact with the flange is at least a cone at  $70^\circ$  which supplements the usual 1:20 cone, but in most cases these two slopes are connected by a concave part which can frequently be in contact with the rail. This “intermediate” contact, where the cone angle  $\gamma$  is variable, is very important both for steering and for stability considerations.

Real profiles, measured with different sorts of apparatus or “profile-meters” will be exploited in an operational software for the profiles study.

The theoretical profiles can be described by analytical expressions from the standards, they are generally calculated at some discrete places and stored in a text file as a set of coordinates:  $z(y)$ . The measured profiles are directly given in this form.

## 3. Profile Measurements: Importance of the Angular Reference

The profiles are compared in the same reference frame. This implies that the measurement devices used to measure the wheel and the rail must be correctly referenced to each other.

For the rails, the reference line rests on the top of the two rails. It can be effected by a transverse cylinder.

For the wheel, the reference line is considered to pass on the “middle of the treads,” at a point 70 mm from the inner flange. These two points are not so easy to find. Some devices use the inner flange face, considering that it is a plane. If this face is not correctly machined, it could be a cone, and the profile study can be relatively biased.

It is useful to measure the two wheels of a wheelset, and the two rail profiles at the same abscissa, especially in a curve. However, it must be possible to consider a single wheel on a single rail.

## B. GRAVITATIONAL STIFFNESS

A description closer to the real shape of the wheels and of the rails is necessary to approach the principle of the gravitational centring mechanism.

In a first step, the vertical left and right loads  $Q$  are considered identical, and the profiles of wheels and rails are considered the same on each side.

When a wheelset is perfectly conical, the horizontal reaction forces  $Q \operatorname{tg}(\gamma)$  to the normal loads  $N_g$  and  $N_d$  are compensated as far as there is no flange contact (Figure 4.14).

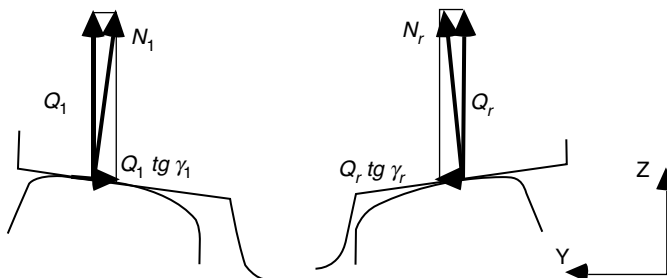


FIGURE 4.14 Conical profiles: equivalence of the horizontal components.

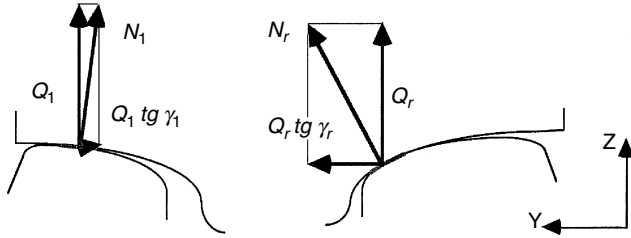


FIGURE 4.15 Progressive centring differential force with concave profiles.

When a wheelset has concave profiles, the normal loads do not stay symmetrical with the lateral displacement. A transversal force appears with the projection of the normal forces in both directions (Figure 4.15).

### C. FLANGE CONTACT

#### 1. Flange Contact Jump and Definition of the Nominal Play in the Track

Even when the profiles are close to two opposite cones, the gravitational effect will appear when flanging (Figure 4.16).

This main contact jump is a way to define the nominal play in the track by two means:

1. For the value where  $\gamma$  is maximum
2. For the value where there is the jump

#### 2. Contact Jump and Load Transfer

From a geometrical point of view, a rigid jump is possible, but from a dynamic one it is unrealistic and from a numerical point of view it is impossible.

A large part of the wheel–rail contact modelisation led in the load transfer when flanging, and more generally when there is a jump between two contact points on the profiles. The load transfer is calculated on the basis of the elastic deformation in the neighbourhood of each contact.

However, the contact stiffness is not the only one elasticity to be taken into account, and the track stiffness itself can be used to smooth the load variation.

The first modelisation of the flange contact has been to consider it as an elastic spring whose value comes from the track and the rail beam deformation.

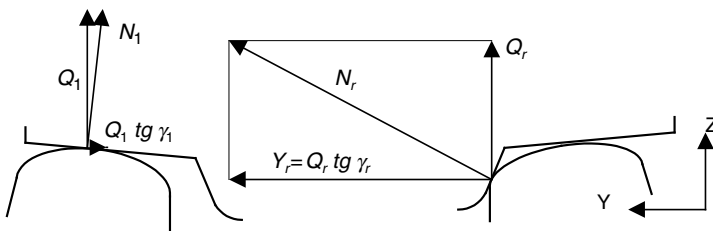
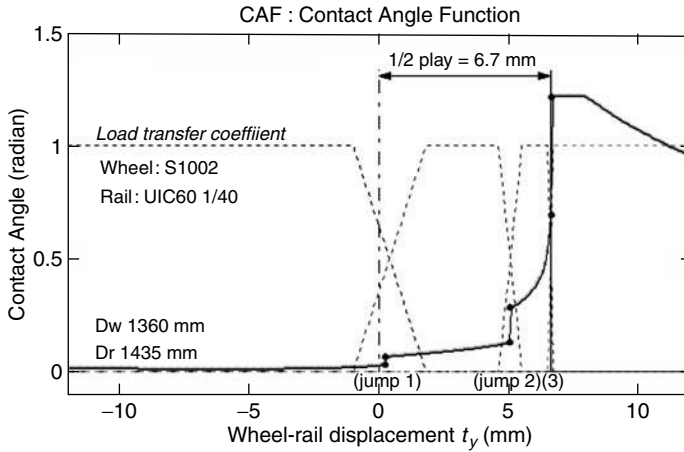


FIGURE 4.16 Gravitational forces and flange contact, without friction.



**FIGURE 4.17** Mono- and multi-contact localisation with the contact angle function, case of the S1002 on UIC60 1:40th.

**D. USING THE CONTACT ANGLE FUNCTION**

**1. Localisation of Multiple Contacts on the Tread**

In order to locate the contact jumps, one of the best contact parameters is the contact angle function (Figure 4.17), i.e., the variation of the contact angle with the transversal displacement  $t_y$ , the variable conicity.

Knowing, for example, the rail profile:

- By integration, it gives the rolling radius (and the wheel profile, in the parts where there is contact)
- By derivation, it gives the transverse curvature of the wheel

As far as the jumps are concerned, any discontinuity in this function can be considered as a jump. An example is given in Figure 4.17 with the well known S1002 wheel profile on the rail UIC60, it appears clearly that the flange contact is not the only possible discontinuity.

If these jumps are giving ellipses too close to each other, the Hertzian assumption of a constant curvature in the contact area is probably no longer valid.

**2. Gravitational Centring Ability in Standard Conditions**

The left wheel and right wheel contact angle values are used to directly calculate the gravitational forces.

When there is a large difference between the angle values, the profiles combination is strongly centring. This is always the case in flange contact. However, the gravitational effect must be efficient even around the central position (Figure 4.18 right).

Routes with many tight curves sometimes use vehicles with concave profiles to improve steering. Worn wheels also generally show this behaviour. However, a strong tendency to instability can be observed with such profiles.

**E. THE DIVERGING EFFECT OF SPIN, INFLUENCE ON THE NORMAL LOAD**

When the friction is not negligible, Figure 4.14 becomes an incorrect representation. The large value of the contact angle generates a large spin creepage value. A large friction value generates a

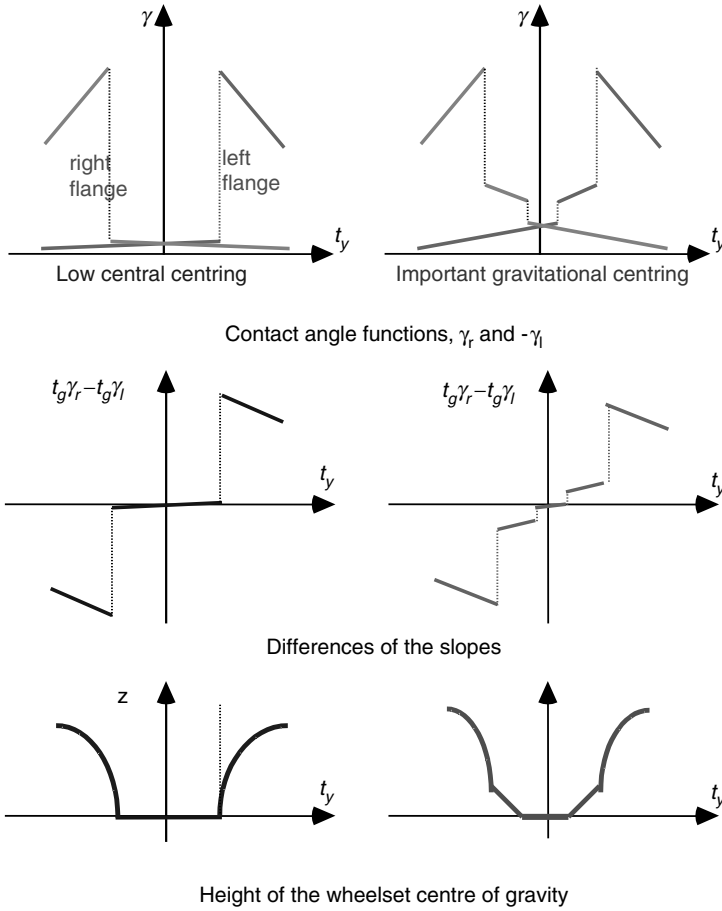


FIGURE 4.18 Gravitational stiffness parameters.

spin torque in the neighbourhood of the contact, generating a lateral force which is always diverging.

Despite the fact that this torsion value is small, it generates the equivalent of a yaw angle offset.

A second effect of this force will be in the wheelset equilibrium: this spin force, due to friction, is directed mainly upward. Considering the equilibrium between the vertical force  $Q$ , this new force  $F_{y\text{spin}}$ , and the reaction force (Figure 4.19), it is found that the normal force  $N$  on a flanging wheel at the equilibrium is reduced by the friction; then the gravitational effect is reduced too. This

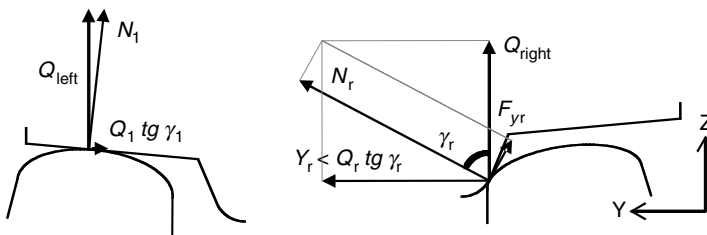


FIGURE 4.19 Gravitational mechanism with the effect of friction and spin.

combined mechanism is a way of explaining some derailments on dry rails. The association with independent wheels increases the friction effect.

## F. SAFETY CRITERIA, NADAL'S FORMULA

The  $Y/Q$  ratio is used as a safety criteria when flanging (see also Chapter 8). In the real case of an attacking wheel, the spin force when flanging is added to the yaw lateral force which together counteract the  $Y$  guiding force  $N \sin \gamma$ . For a right wheel:

$$\begin{aligned} Y &= -N \sin \gamma + F_y \cos \gamma \\ Q &= N \cos \gamma + F_y \sin \gamma \end{aligned} \quad (4.35)$$

When  $F_x$  is negligible,  $F_y \approx \mu N$ , and  $\mu \approx tg \mu$ , a safe level of  $Y/Q$  has been set by Nadal:

$$\left( \frac{Y}{Q} \right)_{\max} = tg(\gamma - \mu) \quad (4.36)$$

where  $\gamma$  is the contact angle (radians) and  $\mu$  is the friction coefficient.

The first way to ensure safety is to give a sufficient contact angle to the flange. The second preference is to reduce the  $Y$  force by a good design of the bogie. Another option is to limit the track twist, limiting the diminution of  $Q$  when flanging. The last way is to reduce the friction coefficient by lubrication. In Europe, this is commonly carried out at the wheel flange, or on the gauge corner of the rail in particular curves.

However, the friction forces are also limited in the lateral direction by the presence of a longitudinal force. The friction coefficient is shared between the two directions (this is the "friction cone"). In the Nadal formula, this sharing effect is not considered, making the formula adequate for independent wheels. This also means that a rigid wheelset, where the longitudinal forces are important on the attack wheel, will be safer than an independent wheel in the same conditions.

## G. INDEPENDENT WHEEL, APPLICATION TO INDUSTRIAL MECHANISMS ON RAILS

Independent wheels are mounted similar to the rear wheelset of a car, on a fixed axle, with roller bearings. This solution is more and more popular for tramways with low floors. A large number of mining trucks and many rolling bridges and rolling cranes are also equipped with independent wheels. In the railway domain, some propositions have even been made to use freely steerable independent wheels, similar to the steering wheelset of the cars, using the gravitational phenomenon to naturally steer the wheelset.<sup>12</sup>

Except due to friction in the bearings, or when braking, there are no longitudinal forces on the independent wheel contact patch. The wheelset cannot be steered by the dicone effect and the gravitational centring effect is the only passive mechanism which can be used to centre the wheelset. This implies that it is helpful to adopt concave profiles for the independent wheels. For large industrial cranes rolling with heavily loaded independent wheels, diabolo profiles can be adopted.

In the case where they are purely conical or cylindrical, the wheelset tends to run with one wheel flange permanently against the rail. This is due to the high sensitivity of the steel wheel to any small perturbation of the yaw angle, as the tangent stiffness is very high.

Because there are no longitudinal friction forces, the adhesion is fully available to generate lateral force and the independent wheels lead more easily to derailment on curves. The diverging force due to the spin must be counterbalanced largely by the gravitational effect, and reduced by lubrication.

Many people do not know about the dicone effect and believe that the gravitational mechanism is the main guiding mode of the railway axle. The gravitational mechanism is an effect of the

normal forces, while the dicone is an effect of tangential forces, the nature of which is not so easy to explain.

### H. MODELLING THE CONTACT JUMPS

During the 1970s, Kalker set out the basis for stress calculations in the contact, followed by Knothe and others. However, the difficult problem of contact jumps has not been managed correctly for several years.

Several methods were proposed using the last parameter calculated from Hertz: the contact deflection  $\delta$ .

#### 1. Hertzian Multiple Contacts

In 1988, Sauvage proposed to interpenetrate the two initial profiles with the  $\delta$  values found in mono-contact conditions. With circular profiles, the normal relative distance is not elliptic but a parabola; FASTSIM used such parabola in the longitudinal direction and it is not so far from the Hertzian ellipsoid pressure.

It was found<sup>13,10</sup> that the intercepted contour (Figure 4.20, lower part) was too large, so only the normal relative distance (upper part) was used.

When the curvatures are irregular, this intersection profile is no longer a parabola (Figure 4.21). This led to the idea of deducing the parabola representing the main ellipse, in order to deduce a second ellipse, then a third one, and so on, with particular criteria to decide how many ellipses are present at the same time.

Unfortunately, at this time, the resulting software was very slow. In order to accelerate the calculation, Pascal proposed making a table giving the equivalent forces with a single “equivalent” ellipse.<sup>14</sup>

At the same time, Ayasse simplified the model and developed analytical equations giving access to the jump width (Figures 4.17 and 4.22) improved with the expressions (4.37).<sup>15,16</sup>

The jump width is defined in two parts, as a function of the angles and of the interpenetration:

$$\Delta t_y = \Delta t_{y_1} + \Delta t_{y_2} \quad \Delta t_{y_1} = \frac{\delta_1}{2} \frac{\cos \gamma_2}{\sin |\gamma_2 - \gamma_1|} \quad \Delta t_{y_2} = \frac{\delta_2}{2} \frac{\cos \gamma_1}{\sin |\gamma_2 - \gamma_1|} \quad (4.37)$$

Between each mono-contact situations, as presented in Figure 4.22, the load transfer is considered to be linear.

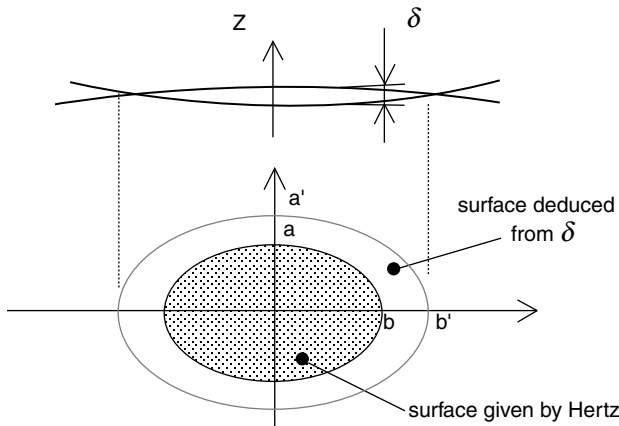


FIGURE 4.20 Interpenetration and contour.

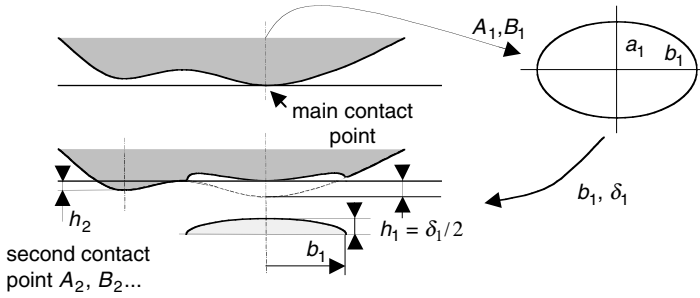


FIGURE 4.21 A possible determination of the secondary ellipses.

When the transfer appears at low angles, between two points on the tread, the transfer width  $\Delta t_y$  is large (see Figure 4.16); on the opposite it is closer to  $\delta$  when flanging.

Both the Sauvage and Ayasse methods have been applied in different versions of the VOCO family codes.

These methods have common hypotheses, one of them is that, when calculating the forces, the different ellipses are independent. In Figure 4.23 for example, the pressure ellipsoids of the multi-elliptic method are not cumulated. However, when they are, the pressure shape is sometimes not very realistic, when there is a large common area between the ellipses. For this reason, more advanced models have been developed.

I. ADVANCED METHODS FOR NONHERTZIAN CONTACTS

At the end of the 1980s, the reference model for normal and tangential contact in Hertzian and nonHertzian cases was the software CONTACT developed by Kalker to simulate his complete theory. This accurate model is generally too slow to be used in a dynamic software loop, and limited to nonHertzian contacts on the tread, it has therefore been mainly used in the railway domain for the validation of other models.

In 1996, a nonHertzian simplified method was proposed<sup>10</sup> in order to make a better stress description in the contact patch and to be used directly in a multibody code. This method is presently used in MEDYNA and ADAMS/Rail, and has inspired the wheel–rail contact of *Universal Mechanism*.

Back to the Hertzian basis, an improved method has been developed recently by the authors.<sup>17</sup> Even in Hertzian slender cases the contact contour is very similar to the Hertzian solution. In nonHertzian cases the results are close to Kalker’s CONTACT, and the multibody simulations can be made in real time.

These methods are semiHertzian, the contact is separated in longitudinal strips as in FASTSIM. The longitudinal pressure can be elliptic. The transversal pressure is nonHertzian, evaluated in different ways from the indentation contour, and the shear stresses are calculated almost as in FASTSIM.<sup>19</sup>

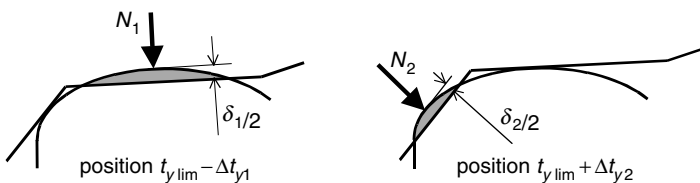
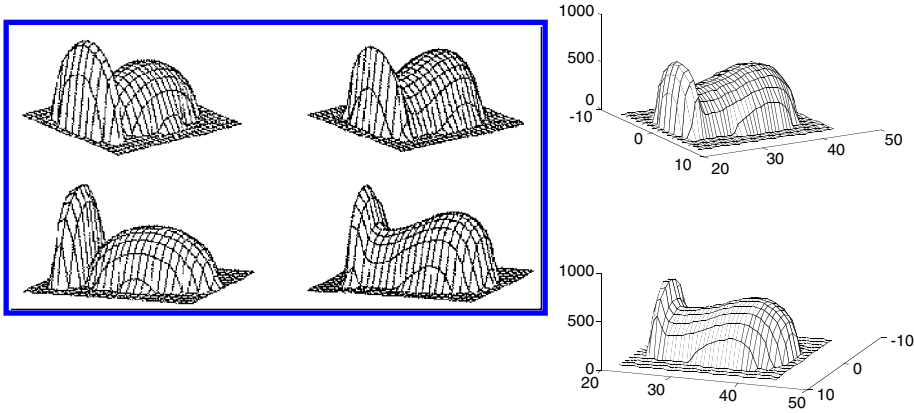


FIGURE 4.22 MultiHertzian contact — CAF method: principle of the determination of the jump limits.



**FIGURE 4.23** Multi-Hertzian, non-Hertzian, and semi-Hertzian models, contact pressure (S1002 wheel on UIC60 rail at 1:40 in the centered position).

**J. WHEELSET EQUILIBRIUM, EXPLICIT SCHEME**

Considering the mechanical system of the wheelset, the spin is taken into account to determine the normal load  $N$ . However, this normal load is used to determine the creep forces. For this reason, in a numerical model, the explicit resolution of the wheelset equilibrium must use a set of values from two consecutive time steps.

In this paragraph, it is proposed to show how to stabilise the numerical resolution of the wheelset equilibrium equations by an adequate choice of the force expressions.

**1. Wheelset Equilibrium Equations**

Establishing the equilibrium equations for a wheelset is equivalent to establishing the sum of the normal and tangent forces described in Section III, complete with the acceleration terms for each degree of freedom. A numerical method is proposed here to improve the stability of the integration.

**2. Decomposition of the Forces**

Here, the track is supposed to be flat in order to avoid the additional terms due to the cant angle. The left and right wheels are supposed to be in contact with the rails on a single contact point: at the left:

$$\begin{aligned} Q_1 &= N_1 \cos \gamma + F_{y1} \sin \gamma \\ Y_1 &= -N_1 \sin \gamma + F_{y1} \cos \gamma \end{aligned} \tag{4.38}$$

with

$$F_{y1} = Ga_1 b_1 c_{22} N_1^{2/3} v_y + Ga_1 b_1 c_{23} N_1 \varphi$$

**3. Adaptation to the Numerical Calculation**

Initially

$$F_{y1} = \frac{Ga_1 b_1 c_{22} N_1}{N_1^{1/3}} v_y + Ga_1 b_1 c_{23} N_1 \varphi \quad (\text{“1” for “left”}) \tag{4.39}$$



a linearisation is proposed by approximation, as a function of  $N_1$ :

$$Q_1 = N_1 \left[ \cos \gamma + \sin \gamma \left( \frac{Ga_1 b_1 c_{22}}{N_1^{1/3}} v_y + Ga_1 b_1 c_{23} \varphi \right) \right] \quad (4.40a)$$

The order of the normal term is  $< 1$  and the tangent term  $< \mu$ . Similarly for the transversal component  $Y_1$ :

$$Y_1 = N_1 \left[ -\sin \gamma + \cos \gamma \left( \frac{Ga_1 b_1 c_{22}}{N_1^{1/3}} v_y + Ga_1 b_1 c_{23} \varphi \right) \right] \quad (4.40b)$$

The term in parenthesis includes a term in  $N_1^{1/3}$  whose variation will be smoother than the  $N_1$  in a numeric calculation. For the  $N_1^{1/3}$  value, the previous time step value  $N_g(i-1)^{1/3}$  is taken. Such a procedure avoids an iteration process.

The above expression will be used in the equilibrium equations in the form of the  $C_{*g-d} = 1/[\dots]$  coefficients.

#### 4. Equilibrium Equations

Function of the left and right normal loads, the wheelset six degrees of freedom leads to the general formulation:

Longitudinal — not considered here

Lateral  $y$  —  $M\ddot{y}CG = C_{yg}N_1 + C_{yr}N_r + \Sigma F_{y \text{ springs}}$

Bounce  $z$  —  $M\ddot{z}CG = C_{zg}N_1 + C_{zr}N_r + \Sigma F_{z \text{ springs}} - Mg$

Roll  $\psi$  —  $I_{xx}\ddot{\psi} = C\psi_l N_1 + C\psi_r N_r + \Sigma M_{F/x}$

Pitch — considered separately

Yaw  $\alpha$  —  $I_{zz}\ddot{\alpha} = C_{\alpha 1}N_1 + C_{\alpha d}N_d + \Sigma M_{F/z}$  (4.41)

(The last equation being independent of the others).

Note that the three equations in (y), (z) and roll ( $\psi$ ) are coupled as the vertical translation  $t_z$  and the roll are two functions of the lateral wheelset translation  $t_y$  due to the geometry of the wheels and rails.

The  $C_{*g-d}$  coefficients which are presented in these expressions as proportional to  $N_{g-d}$  can be determined as in the previous paragraph. They introduce the effect of the tangent term, which includes the three combined creep forces. Their combination through a common friction coefficient is considered in the friction contact model.

These expressions can be extended to multiple ellipses and to the semi-Hertzian models.

#### APPENDIX 4.1. KINEMATIC MOVEMENT: THE KLINGEL FORMULA

For this simplified formula, the wheels are considered perfectly conic, rolling on a line representing the rail (Figure 4.6).

The Klingel formula is the expression of the hunting wavelength, without any tangential forces:

$$\lambda = 2\pi \sqrt{\frac{r_0 Dc/2}{\gamma}} \quad (A4.1)$$

The amplitude of this sinusoidal movement is an initial condition. However it is limited by the flange contact.

## APPENDIX 4.2. KINEMATIC HUNTING AND EQUIVALENT CONICITY

Even if the wheel profile is not a perfect cone, and the rail is not a line, the periodic movement of a free real wheelset in the track will remain close to a sinusoidal movement. This is called kinematic hunting. Its wavelength is a way to determine the equivalent conicity of the wheel–rail profile combination.

The equivalent conicity notion is well known in the railway field, where it was useful at the time of the linearised dynamic models.

## APPENDIX 4.3. THE CIRCLE THEORY

From the concave shape of the wheel ( $R_{wx}$ ) and the rail ( $R_{rx}$ ), with the approximation that these radii are constant in the wheelset excursion, for a small excursion around the central position, Joly shows that the equivalent conicity  $\gamma_e$  must be used in place of the value  $\gamma_o$  of the cone angle in the central position:

$$\gamma_e = \frac{R_{wx}}{R_{wx} - R_{rx}} \gamma_o \quad (\text{A4.2})$$

If the wheel radius is infinite, the wheel is a cone, and the expression returns to the central value  $\gamma_o$ .

This expression is an improvement of the wavelength expression, in comparison with the cone value. The running safety of the TGV was established before 1980 with this conicity determination.<sup>18</sup>

With 1:20 tapered wheels, this formula can be used for an equivalent conicity of 0.2. Over this value, the contact probably differs from the initial dicone.

## APPENDIX 4.4. ANALYSIS OF $Y/Q$ AND NADAL'S CRITERIA

The Nadal criteria  $Y/Q < tg(\gamma - \mu)$  is critical for the evaluation of the safety of a wheelset to derailment. It can be estimated by measurement, by numerical simulation, and also by an analytical quasistatic model. The following study establishes such an analytical expression of  $Y/Q$ , in the track frame, at the contact point with a set of complete hypothesis.

At the contact point, the tangent plane common to the wheel and the rail makes an angle  $\gamma$  with the track plane. The contact force can be dispatched in a normal force  $N$ , and two tangential forces  $f_x$  and  $f_y$  in this plane, due to the friction (Figure 4.24).

Projecting these forces in the track frame  $OYZ$ , then follows the forces  $Y$  and  $Q$ , in the directions  $OY$  and  $OZ$ , respectively:

$$\begin{aligned} Y &= f_y \cos \gamma - N \sin \gamma \\ Q &= f_y \sin \gamma + N \cos \gamma \end{aligned} \quad (\text{A4.3})$$

Inversely, the forces  $f_y$  and  $N$  can be expressed by:

$$\begin{aligned} f_y &= Y \cos \gamma + Q \sin \gamma \\ N &= -Y \sin \gamma + Q \cos \gamma \end{aligned} \quad (\text{A4.4})$$

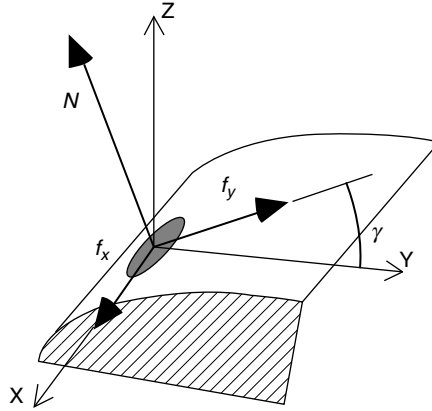


FIGURE 4.24 Normal and tangent forces.

In the particular case when the tangential resultant force is saturated, the additional relation follows:

$$f_x^2 + f_y^2 = (\mu N)^2 \tag{A4.5}$$

where  $\mu$  is the friction coefficient at the contact point.

Replacing  $f_y$  and  $N$  in Equation A4.5 by their expression A4.2, one obtains the relation:

$$\left(\frac{Y}{Q}\right)^2 (\cos^2 \gamma - \mu^2 \sin^2 \gamma) + 2 \frac{Y}{Q} \cos \gamma \sin \gamma (1 + \mu^2) + \sin^2 \gamma - \mu^2 \cos^2 \gamma + \left(\frac{f_x}{Q}\right)^2 = 0 \tag{A4.6}$$

a second order equation in  $Y/Q$  whose determinant  $\Delta$  writes:

$$\Delta = \mu^2 \left[ 1 - \left(\frac{f_x}{\mu Q}\right)^2 (\cos^2 \gamma - \mu^2 \sin^2 \gamma) \right] \tag{A4.7}$$

It is always positive if the relation Equation A4.5 concerning the saturation is respected.

With a limited development at the first order:

$$\left(\frac{f_x}{\mu Q}\right)^2 (\cos^2 \gamma - \mu^2 \sin^2 \gamma) \ll 1 \tag{A4.8}$$

the two roots of the equation can be written:

$$\frac{Y}{Q} \approx -tg(\gamma - \epsilon Atg\mu) - \frac{\epsilon}{2\mu} \left(\frac{f_x}{Q}\right)^2 \tag{A4.9}$$

where  $\epsilon = \pm 1$  depends on the choice for  $\sqrt{\Delta}$

This indetermination comes from the saturation relation A4.5; as  $N$  is always positive and if the longitudinal force  $f_x$  is supposedly small,  $f_y$  is the main responsible of the saturation which is a hypothesis (meaning that the yaw angle is important); then it follows that  $f_y = \epsilon \mu N$ ; the value of  $\epsilon$  is directly linked to the sign of the force  $f_y$ .

The ratio  $Y/Q$  for one contact point depends mainly, respecting the hypothesis, on the contact angle and the friction coefficient.

Neglecting the term in  $f_x$ , for a single contact in the flange, with a contact angle of  $70^\circ$  and a friction coefficient of 0.35, the unfavourable force  $f_y$ , facilitating the wheel climb, is positive:  $\epsilon = +1$ . In this case, the modulus of  $Y/Q$  reaches 1.22, this value is commonly given as critical for derailment.

However, the relation A4.9 shows that, with a reduced friction coefficient,  $Y/Q$  can, with the same other conditions, reach more important values ( $Y/Q = 2$  if  $\mu = 0.1$ ). However, in reality, this value is difficult to reach because, with a low friction coefficient on the flange, there will certainly be a bi-contact on both flange and tread, in contradiction to the proposed hypothesis. On the other hand, if  $\mu = 0.5$ ,  $Y/Q$  decreases to 0.9.

Note that the ratio  $Y/Q$  is close to the friction coefficient on the tread, where the contact angle is low. This situation is also found in the case, during derailment, when the flange top is rolling across the tread.

These considerations show that  $Y/Q$ , for a single wheel–rail contact, can be expressed analytically in a simple way and with clearly defined approximations, and that the limit of 1.2 commonly proposed corresponds to specific values of the contact angle, the friction coefficient, and saturation.

## NOMENCLATURE

$a, b$ :	longitudinal and lateral semi-axis of an ellipse
$c_{11} c_{22} c_{23}$ :	Kalker's coefficients
$f_x, f_y$ :	contact forces in the tangent plane $Oxy$
$m, n, r$ :	Hertz parameters
$p_x p_y p_z$ :	contact pressure, normal, transversal, longitudinal
$r_o$ :	rolling radius of the wheel, around axis $Oy$
$r_n$ :	longitudinal radius of the wheel at the contact point
$y$ or $t_y$ :	lateral displacement of the wheel respective to the rail
$A, B$ :	curvatures at the contact point
$D_r$ :	rail gauge: distance between the inner faces of the rails
$D_w$ :	wheelset gauge: distance between the inner flanges
$D_c$ :	track gauge: distance between the contacts
$E$ :	Young's modulus of the material
$F_x, F_y$ :	longitudinal, lateral force
$G$ :	shear modulus of the material
$L_1 L_2 L_3$ :	Kalker's elastic coefficients
$N$ :	normal load on a contact patch
$OXYZ$ :	reference frame of the track
$Ox$ :	longitudinal axis, direction of rolling
$Oxyz$ :	reference frame at the contact point
$Oy$ :	lateral axis, to the left
$Oz$ :	vertical axis, to the top
$Q$ :	vertical load on the wheel–rail contact, $OXYZ$ frame
$R_{rx}$ :	transversal radius of the rail profile
$R_{wx}$ :	transversal radius of the wheel profile
$V_x$ :	longitudinal speed of the wheelset
$Y$ :	lateral load on the wheel–rail contact, $OXYZ$ frame
$\alpha$ :	yaw angle
$\delta$ :	relative reduction of distance between elastic bodies, Hertz' theory
$\varphi$ :	spin creepage
$\gamma$ :	contact angle, inclination of the profiles at the contact point, for any position of the wheelset
$\gamma_o$ :	contact angle in the central position of the wheelset

$\gamma_e$ :	equivalent conicity
$\mu$ :	friction coefficient
$\nu$ :	Poisson's ratio of the material
$\nu_x$ :	longitudinal creepage
$\nu_y$ :	lateral creepage
$\omega$ :	rotation speed of the wheelset around $O_y$
$\psi$ :	roll angle of the wheelset

## REFERENCES

1. Carter, F. W., On the action of a locomotive driving wheel, Proc. Roy. Soc. Lond. Ser. A, 112, 151–157, 1926.
2. Vermeulen, P. J. and Johnson, K. L., Contact of nonspherical elastic bodies transmitting tangential forces, *Trans. ASME*, 1964, June.
3. Kalker, J. J., A strip theory for rolling with slip and spin, Proc. Kon. Nederlandse Akademie van Wetenschappen, Amsterdam, B70, pp. 10–62, 1966.
4. Kalker, J. J., The tangential force transmitted by two elastic bodies rolling over each other with pure creepage, *Wear*, Vol. 11, pp. 421–430, 1968.
5. Kalker, J.J., On the rolling contact of two elastic bodies in the presence of dry friction, PhD Thesis, DELFT, 155 pages, 1967.
6. Kalker, J. J., *Three Dimensional Elastic Bodies in Rolling Contact*, 1st ed., Kluwer, Dordrecht, 1990.
7. Ohyama, T., Some problems of the fundamental adhesion at higher speeds, *Quart. Rep. RTRI*, 14(4), 181, 1973.
8. Chollet H., Etude en similitude mécanique des efforts tangents au contact roue-rail, Thèse de Doctorat de l'Université PARIS 6, (in French) 1991.
9. Kalker, J. J., *A Fast Algorithm for the Simplified Theory of Rolling Contact (FASTSIM program)*, *Vehicle Systems Dynamics*, Vol. 11, Swets & Zeitlinger B.V., Lisse, pp. 1–13, 1982.
10. Kik, W. and Piotrowski, J., A fast approximate method to calculate normal load at contact between wheel and rail and creep forces during rolling, Second Mini-Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Budapest, July 29–31, 1996.
11. Polach, O., Creep forces in simulations of traction vehicles running on adhesion limit, Proceedings of the Sixth International Conference on Contact Mechanics and Wear of Rail/Wheel Systems (CM2003), Goteburg, 2003.
12. Frederich, F., Possibilities as yet unknown regarding the wheel/rail tracking mechanism, *Rail Int.*, 16, 33–40, 1985.
13. Pascal J. P., About Multi Hertzian contact hypothesis and equivalent conicity in the case of S1002 and UIC60 analytical wheel/rail profiles, *Vehicle System Dynamics (VSD)*, 22(2), Swets & Zeitlinger B.V., Lisse, pp. 57–78, 1993.
14. Pascal, J. P. and Sauvage, G., New method for reducing the multicontact wheel/rail problem to one equivalent rigid contact patch, Proceedings of the 12th IAVSD Symposium, Swets & Zeitlinger B.V., Lisse, Lyon, August 26–30, 1991.
15. Ayasse, J. B., Chollet, H., and Maupu, J. L., Paramètres caractéristiques du contact roue-rail. INRETS report no. 225, ISBN 0768-9756 (in French), 2000.
16. Ayasse, J. B., Chollet, H., Fleuret, J. S., and Lévêque, E., A generalised conicity criteria for the wheel–rail contact. Example of a switch blade safety study, In *Eighth Mini Conference VSDIA, B.U.T.E, Budapest, 11–13 November*, Zobory, I., Ed., ISBN 963 420 817 7, 2002.
17. Ayasse, J. B. and Chollet, H., Determination of the wheel rail contact patch for semi-Hertzian conditions, *Vehicle Syst. Dyn.*, 43(3)2005.
18. Joly R., Etude de la stabilité transversale d'un véhicule ferroviaire, *Revue Francaise de Mécanique*, Vol. 36, 1970.
19. Piotrowski, J. and Chollet, H., Wheel-rail contact models for vehicle system dynamics including multi-point contact, *Vehicle System Dynamics*, Vol. 43(6-7), pp. 455–483, June–July 2005.

---

# 5 Tribology of the Wheel–Rail Contact

*Ulf Olofsson and Roger Lewis*

## CONTENTS

I.	Introduction .....	121
II.	Contact Conditions at the Wheel–Rail Contact.....	122
III.	Wear and Other Surface Damage Mechanisms.....	125
	A. Wear .....	125
	B. Plastic Deformation.....	127
	C. Rolling Contact Fatigue .....	129
IV.	Friction .....	131
	A. Wheel–Rail Friction Conditions .....	132
	B. Friction Modification .....	133
	C. Adhesion Loss .....	134
	D. Increasing Adhesion.....	135
V.	Lubrication and Surface Coatings .....	136
	A. Benefits of Lubrication .....	136
	B. Methods of Lubrication Application .....	136
	C. Problems with Lubrication.....	137
	D. Lubricator System Selection and Positioning .....	137
	E. Surface Coatings .....	138
	Acknowledgments .....	138
	References.....	138

## I. INTRODUCTION

Tribology, the science and technology of friction, wear, and lubrication, is an interdisciplinary subject. It can therefore be addressed from several different viewpoints. This chapter focuses on the friction, wear, and lubrication of the tiny contact zone (roughly  $1\text{ cm}^2$ ), where steel wheel meets steel rail, from a mechanical engineer's viewpoint. In contrast to other well-investigated machinery, such as roller bearings, the wheel–rail contact is an open system. It is exposed to dirt and particles and natural lubrication, such as high humidity, rain, and leaves, all of which can seriously affect the contact conditions and the forces transmitted through the contact. In contrast, in roller bearing the ball–cage contacts are sealed away. The steel rail meets a population of steel wheels from a number of different vehicles and the form of both the wheels and the rail can change due to wear. In contrast, a roller bearing meets the same rollers without any form change of the contacting bodies.

A comprehensive overview of the science of tribology is presented in the ASM handbook,<sup>1</sup> while a closer examination of the material science field is given by Hutchings.<sup>2</sup> The mathematical modelling aspects of tribology, i.e., contact mechanics and fluid film lubrication, are presented by

Johnson<sup>3</sup> and Dowson and Higginson.<sup>4</sup> An excellent historical overview of the field is presented in Dowson.<sup>5</sup>

In the contact zone between wheel and rail, normal and tangential loads are transmitted. How the steel wheel meets the steel rail and the size of the forces transmitted in the contact zone influence damage mechanisms, such as wear and surface cracking, are discussed. The contact conditions of the wheel–rail contact are discussed in Section II (Contact conditions at the wheel–rail contact).

When two surfaces under load move relative to each other, wear will occur. Wear is often defined as damage to one or both surfaces, involving loss of material. Wear and other surface damage mechanisms are discussed in Section III (Wear and other surface damage mechanisms).

The friction force can be defined as the resistance encountered by one body moving over another body. This definition covers both sliding and rolling bodies. Note that even pure rolling nearly always involves some sliding and that the two classes of motion are not mutually exclusive. Any substance between the contacting surfaces may affect the friction force. The contact conditions may cause the substance to be wiped away quickly and its effect will be minimal. On the other hand, surface films formed between interposed substances have a major effect on the frictional behaviour. The friction of the wheel–rail contact is discussed in Section IV (Friction), as well as causes of friction loss and methods for increasing the friction.

Lubricant application to the wheel–rail contact as well as surface coatings are used to reduce friction and damage due to wear etc. This is discussed in Section V (Lubrication and surface coatings).

What one always should bear in mind when studying and using tribological data is that friction and wear are system parameters and not material parameters like modulus of elasticity or fracture toughness. This means that frictional and wear data taken from one system, such as a roller bearing, cannot be directly applied to another system such as the wheel–rail contact. This also highlights the need for a special study of the tribology of the wheel–rail contact.

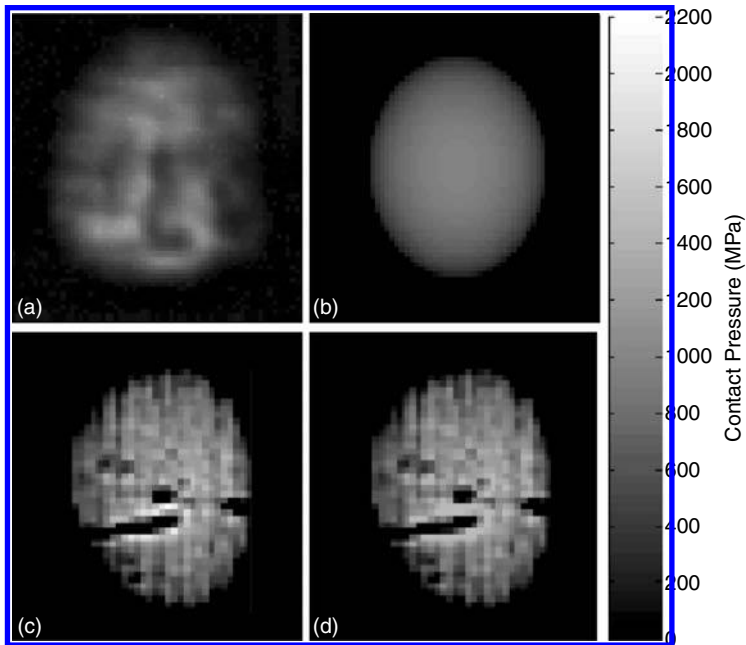
## II. CONTACT CONDITIONS AT THE WHEEL–RAIL CONTACT

In the contact zone between railway wheel and rail the surfaces and bulk material must be strong enough to resist the normal (vertical) forces introduced by heavy loads and the dynamic response induced by track and wheel irregularities. The tangential forces in the contact zone must be low enough to allow moving heavy loads with little resistance, at the same time the tangential loads must be high enough to provide traction, braking, and steering of the trains.

The contact zone (roughly 1 cm<sup>2</sup>) between a railway wheel and rail is small compared with their overall dimensions and its shape depends not only on the rail and wheel geometry but also on how the wheel meets the rail influence, i.e., lateral position and angle of wheel relative to the rail, as shown by Le The Hung.<sup>6</sup>

It is difficult to make direct measurements of the contact area between the wheel and the rail. An interesting approach for measuring the contact area for full-scale worn wheel and rail pieces is presented by Marshall et al.<sup>7</sup> They used an ultrasonic reflection technique and the results were compared with calculated contact areas showing good agreement, as shown in [Figure 5.1](#). The surface topographies of the ultrasonic measured surfaces were measured with a stylus instrument and used as input to a contact mechanics method for rough surfaces (for details see Björklund et al.<sup>8</sup>). Poole<sup>9</sup> used low-pressure air passing through 1 mm diameter holes drilled into the rail head to measure the contact area as the holes being blocked by the passing wheel. Measurement of these pressure variations allows studying of the contact area shape under dynamic conditions.

The size and shape of the contact zone where the railway wheel meets the rail can be calculated with different techniques. Traditionally, the Hertz theory of elliptical contacts<sup>3</sup> has been used implying the following assumptions: the contact surfaces are smooth and can be described by second degree surfaces; the material model is linear elastic and there is no friction



**FIGURE 5.1** Contact pressure maps for a load of 80 kN: (a) ultrasonic measurement; (b) Hertzian; (c) elastic model; (d) elastic-plastic mode (from Marshall et al.<sup>7</sup>).

between the contacting surfaces; and the contacting bodies are assumed to deform as infinite half spaces. The half space assumption puts geometrical limitations on the contact, i.e., the significant dimensions of the contact area must be small compared with the relative radii of the curvature of each body. Especially in the gauge corner of the rail profile, the half plane assumption is questionable since the contact radius here can be as small as 10 mm. Due to its simple closed form solutions, the Hertz method is the most commonly used approach in vehicle dynamics simulation. However, other methods are used for simulation of wear and surface fatigue due to the overestimation of the contact stresses attributed to the nonvalidity of the half plane assumption and nonlinear material behaviour. Kalker's numerical program *Contact*<sup>10</sup> still depends on the half space assumption, but is not restricted to elliptical contact zones. The contact surfaces are meshed into rectangular elements with constant normal and tangential stresses in each rectangular element. Telliskivi and Olofsson<sup>11</sup> developed a finite element model, including plastic deformation, of the wheel–rail contact using measured wheel and rail profiles as input data. They compared the traditional methods (Hertz and *Contact*) with their detailed finite element solutions of the wheel in contact with the rail gauge (Case 1 in Figure 5.2) and the wheel in contact with the rail head (Case 2 in Figure 5.2). The results in terms of contact zone shape and size, as well as stress distribution, are presented in Figure 5.3. The results from two test cases show that the difference in maximum contact pressure between *Contact*/Hertz and the model was small for test case 2 when the minimum contact radius is large compared with the significant dimensions of the contact area (half space assumptions valid). However, in test case 1 where the minimum contact radius not was small compared with the significant dimensions of the contact area, the difference between the model and *Contact*/Hertz was as large as 3 GPa. Here, the difference was probably due to both the half space assumption and the material model.

The Stockholm local network has been the subject of a national Swedish transport programme (the Stockholm test case)<sup>12–15</sup> in which the wear, surface cracks, plastic deformation, and friction



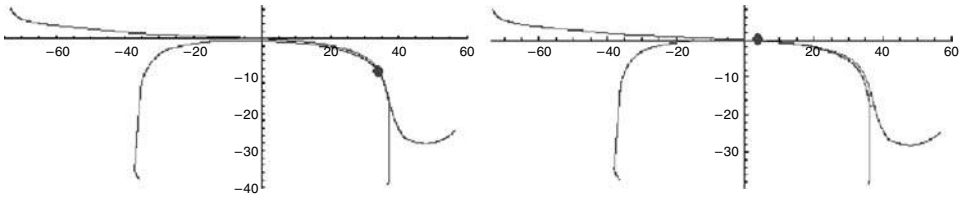


FIGURE 5.2 Contact point location for the two load cases. Load case 1 (left) and load case 2 (right).

of rail and wheel have been observed for a period of 2 years. The data from the Stockholm test case has been used for validation of different wear models, see<sup>16–18</sup> and also surface crack models.<sup>19</sup>

Furthermore, the trains used in this study have been modelled with train dynamic simulation software such as GENSYS<sup>17</sup> and Medyna.<sup>20</sup> A parametric study<sup>17</sup> was performed on curves with different radii representative of Stockholm local traffic. The results are presented here in the form of a contact pressure sliding velocity diagram (Figure 5.4). A clear difference could be found between the rail head-wheel tread contact and the rail gauge-wheel flange contacts in terms of sliding velocity and contact pressure. For the rail head-wheel tread contact, the sliding velocity and the contact pressure was never above 0.1 m/sec and 1.5 GPa, respectively, but for the rail gauge-wheel flange the maximum sliding velocities reached 0.9 m/sec, and maximum contact pressure was observed up to 2.7 GPa. Also shown in Figure 5.4 are simulation results from a curve with a 303 m radius, for the Stockholm test case using the software Medyna. This is a sharp curve with one of the smallest radii in the network and one can note a very high contact pressure for the first wheel on the

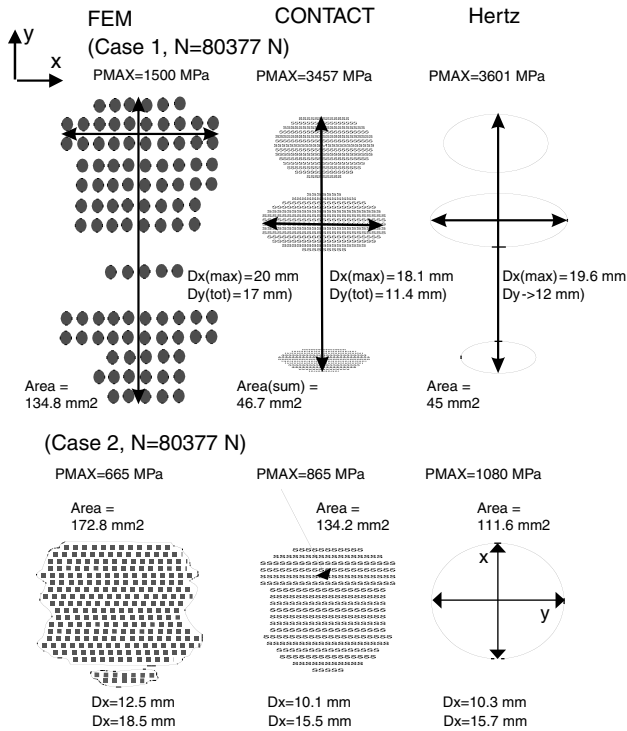
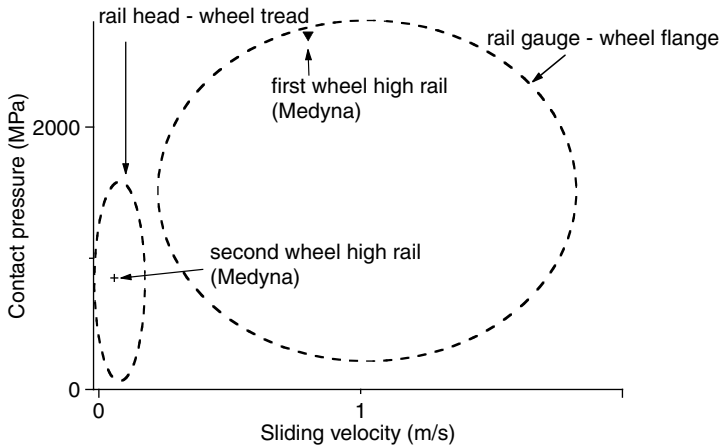


FIGURE 5.3 Comparison, with respect to maximum contact pressure and the contact area, between three different contact mechanics analysis methods.<sup>11</sup>



**FIGURE 5.4** Sliding velocity contact pressure chart from the Stockholm test case.<sup>13</sup> The elliptical areas show typical regions where rail head-wheel tread and rail gauge-wheel flange occur.<sup>17</sup> Also shown in the figure are simulation results from a small radius curve in the Stockholm test case the using the software Medyna.<sup>20</sup>

leading bogie in contact with the rail gauge. Other examples of modern railway operation which led to high contact stresses that are significantly over the yield strength of the material are presented in Kumar<sup>21</sup> and Cassidy.<sup>22</sup>

### III. WEAR AND OTHER SURFACE DAMAGE MECHANISMS

The profile change of rails on curves makes a large contribution to track maintenance cost. The profile change on wheels can also be significant, especially on a curved track. Damage mechanisms such as wear and plastic deformation are the main contributors to profile change. Another growing problem for many railways is rolling contact fatigue.<sup>23</sup> In Europe, there are more than one hundred broken rails each year due to rolling contact fatigue. In 1995, rail maintenance costs within the European Union were estimated to total 300 million Euro annually.<sup>23</sup>

#### A. WEAR

Wear is the loss or displacement of material from a contacting surface. Material loss may be in the form of debris. Material displacement may occur by transfer of material from one surface to another by adhesion or by local plastic deformation. There are many different wear mechanisms that can occur between contacting bodies, each of them producing different wear rates. The simplest classification of the different types of wear that produce different wear rates is “mild wear” and “severe wear”. Mild wear results in a smooth surface that often is smoother than the original surface. On the other hand, severe wear results in a rough surface that often is rougher than the original surface.<sup>24</sup> Mild wear is a form of wear characterised by the removal of materials in very small fragments. Mild wear is favourable in many cases for the wear life of the contact as it causes a smooth run-in of the contacting surfaces. However, in some cases it has been observed that it worsens the contact condition and the mild wear can change the form of the contacting surfaces in an unfavourable way.<sup>25</sup> Another wear process that results in a smooth surface is the oxidative wear process characterised by the removal of the oxide layer on the contacting surfaces. In this case the contact temperature and asperity level influence the wear rate.<sup>26</sup> Abrasive wear caused by hard particles between the contacting surfaces can also cause significant wear and reduce the life of the contacting bodies.<sup>27</sup>

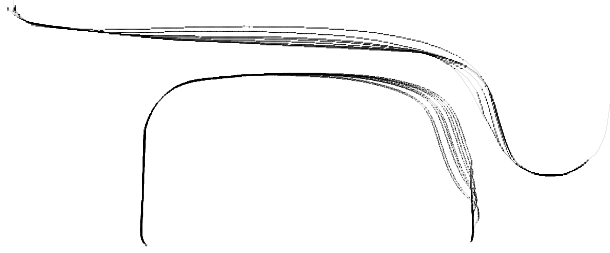


FIGURE 5.5 Form change of wheel and rail from the Stockholm test case.<sup>15</sup>

In wheel–rail contact, both rolling and sliding occur in the contacting zone. Especially in curves, there can be a large sliding component on the contact patch at the track side of the rail head (gauge corner). Due to this sliding, wear occurs in the contact under the poorly lubricated condition that is typical of wheel–rail contact, as shown in Figure 5.5. An observation that can be made on sliding wear is that an increase of the severity of loading (normal load, sliding velocity, or bulk temperature) leads, at some stage, to a sudden change in the wear rate (volume loss per sliding distance). The severe wear form is often associated with seizure. The transfer from mild acceptable wear to severe/catastrophic wear depends strongly on the surface topography. The loading capability of a sliding contact may be increased considerably by smoothing the surface.<sup>28</sup> Chemically reacted boundary layers imposed by additives in the lubricant can improve the properties of lubricated contacting surfaces and reduce the risk of seizure.<sup>29</sup> Also, as shown by Lewis and Dwyer-Joyce,<sup>30</sup> the surface temperature influences the transition from mild to severe wear.

In addition to the contact pressure and the size of the sliding component, natural and applied lubrication strongly influenced the wear rate<sup>13–15</sup> for the full-scale test results from the Stockholm test case. Both lubricated and nonlubricated, as well as seasonal variations, were studied. In addition, two different rail hardnesses were studied in the same test curves. Track side lubrication reduced the wear significantly, and a lubrication benefit factor 9 for small radius curves (300 m) was reported. For 600–800 m radius curves the lubrication benefit factor was approximately 4. The variation seen in wear rates over the year was probably due to natural lubrication caused by changing weather conditions. An analysis of the relationship between weather conditions and measured rail wear shows that the precipitation has a significant effect on rail wear as shown in Figure 5.6. Waara<sup>31</sup> reports that gauge face wear in a northern Sweden heavy haul application can be reduced 3–6 times with correct full year lubrication. Engel<sup>32</sup> also reports significant reduction of wear by lubrication, here the lubricant benefit factor was 4 in a twin-disc test. An on-board

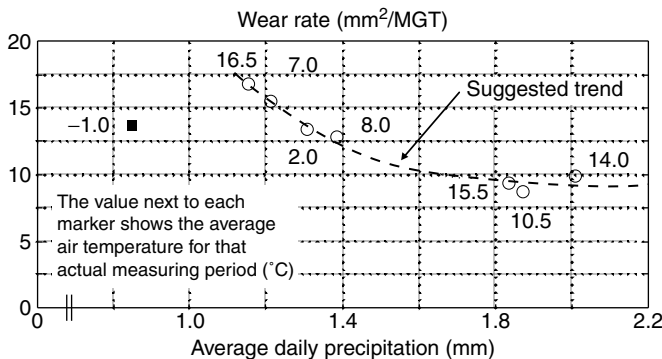
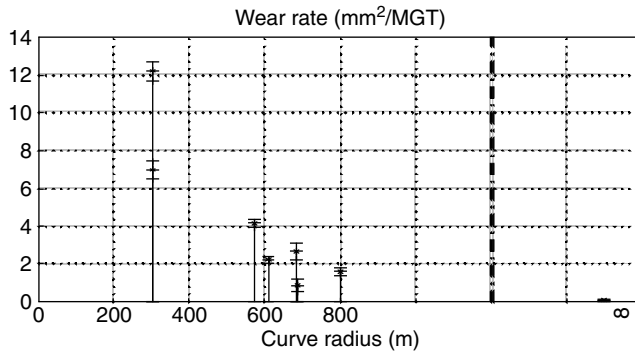


FIGURE 5.6 Rail wear rate vs. average daily precipitation, from Nilsson.<sup>15</sup> MGT = mega gross tonne traffic.



**FIGURE 5.7** Wear rate for high rail as function of curve radius in the Stockholm test case (from Nilsson<sup>15</sup>). MGT = mega gross tonne traffic.

lubrication system was evaluated by Cantara<sup>33</sup> in a Spanish study. The results were that the flange wear was reduced by a factor of 4.5 for wheels equipped with the on-board lubrication device.

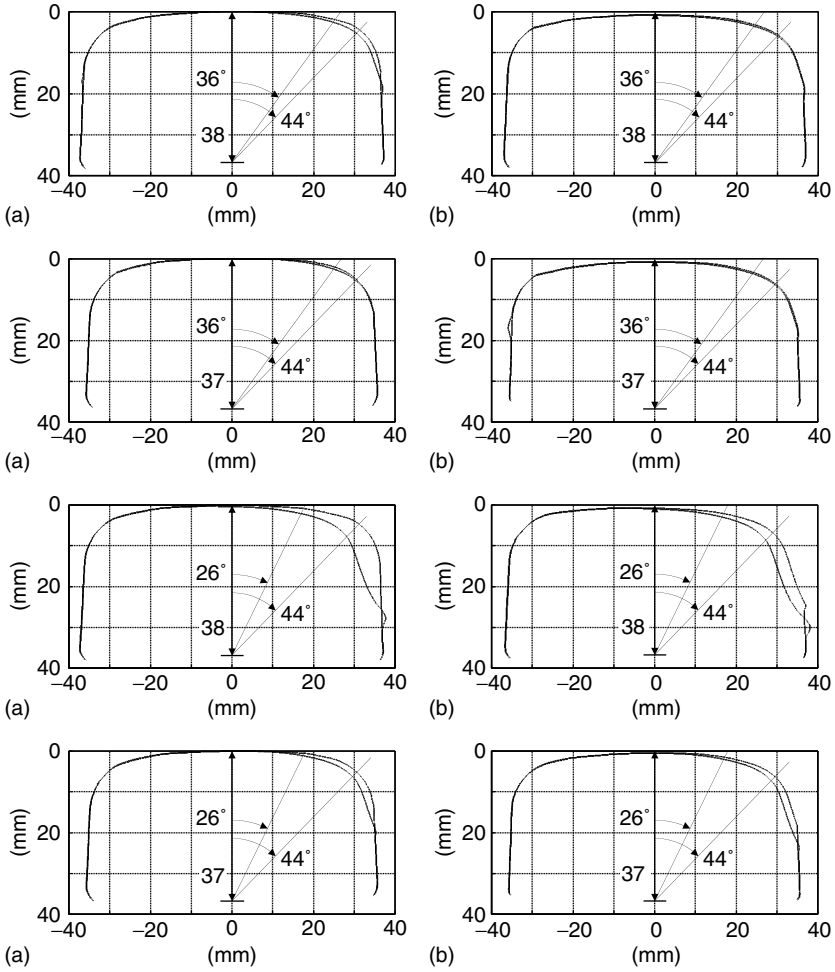
The curve radius of the track has a strong influence on rail wear. The influence also strongly depends on the vehicles and their behaviour. In the Stockholm test case all vehicles were of the same type and passed over all the test sites with the same frequency. In this case the influence of curve radius can be clearly seen when comparing rail wear rate as a function of curve radius. The rail wear rate seems to increase exponentially for decreasing curve radius, as shown in Figure 5.7.

For a given situation a higher steel grade usually reduces rail wear. This effect is shown in Figure 5.8 for two different high rails with steel grade UIC 900A and UIC 1100, respectively, within the same lubricated, as well as a parallel nonlubricated, 300 m radius curve. For the nonlubricated curve the ratio between rail wear rate for the 900A grade rail compared to that of the 1100 grade rail is approximately 2. This can be compared with the lubricant benefit factor that was approximately 9 in this curve, as can be seen in Figure 5.8a–d, when comparing the nonlubricated and lubricated cases. The difference between rail head wear (low sliding velocities and contact pressure) and rail gauge wear (high contact pressure and sliding velocities) was seen to be a factor 10. This is also comparably higher than the rail grade benefit for modern rail steels as UIC 900A and UIC 1100. The observation that the contact conditions in terms of contact pressure and sliding velocity are more important than the grade of steel (900A and 1100) has also been verified in two-roller tests.<sup>13</sup> However, when Lewis and Olofsson<sup>34</sup> compared rail steel wear coefficients taken from laboratory tests run on twin disc and pin-on-disc machines, as well as those derived from measurements taken in the field, they found that the introduction of more modern rail materials had reduced wear rates by up to an order of magnitude in the last 20 years.

Fully pearlitic rail steels are still the most common and are used by most railways. Pearlite is a lamellar product of eutectoid composition that is formed in steel during transformation under isothermal continuous cooling. It consists of ferrite and cementite. Perez-Uzeta and Beynon<sup>35</sup> have shown that the wear rate of pearlitic rail steel decreases with lower interlamellar spacing between the cementite lamella giving a corresponding increase in hardness. Steels with a bainitic microstructure are the other main rail steels. They have shown better rolling contact fatigue resistance than pearlitic rail steels. However, the wear resistance of bainitic rail steels is inferior to that of pearlitic rail steels at a fixed tensile strength, as shown by Graham and Beynon<sup>36</sup> and Mitao et al.<sup>37</sup>

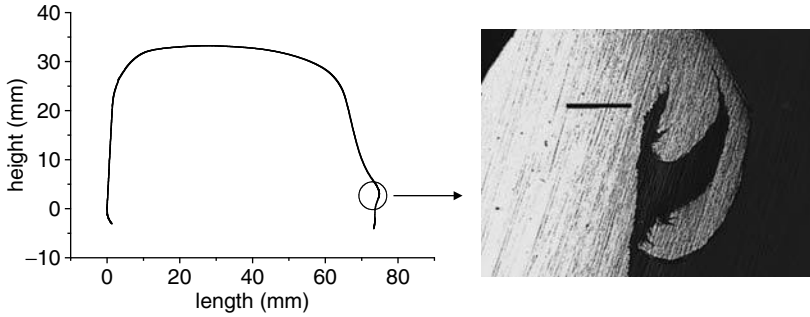
## B. PLASTIC DEFORMATION

On a straight track, the wheel is in contact with the top of the rail, but in curves, the wheel flange may be in contact with the gauge corner of the rail. The wheel load is transmitted to the rail through



**FIGURE 5.8** (a) Results from form measurements of high rail at test start and after 2 years of traffic: (i) new rail at test start, (ii) worn rail at test start. The curve radius is 346 m, the rail steel grade is UIC 900A and the curve was lubricated during the measuring period (from Ref.14), (b) Results from form measurements of high rail at test start and after 2 years of traffic: (i) new rail at test start, (ii) worn rail at test start. The curve radius is 346 m, the rail steel grade is UIC 1100 and the curve was lubricated during the measuring period (from Ref.14), (c) Results from form measurements of high rail at test start and after 2 years of traffic: (i) new rail at test start, (ii) worn rail at test start. The curve radius is 303 m, the rail steel grade is UIC 900A and the curve was not lubricated during the measuring period (from Ref.14), (d) Results from form measurements of high rail at test start and after 2 years of traffic: (i) new rail at test start, (ii) worn rail at test start. The curve radius is 303 m, the rail steel grade is UIC 1100 and the curve was not lubricated during the measuring period (from Ref.14).

a tiny contact area under high contact stresses. This results in repeated loading above the elastic limit, which leads to plastic deformation. The depth of plastic flow depends on the hardness of the rail and the severity of the curves; it can be as much as 15 mm.<sup>38,39</sup> When a material is subjected to repeat loading, its response depends on the ratio of the amplitude of the maximum stress to the yield stress of the material. When the load increases above the elastic limit, the contact stresses exceed yield and the material flow plastically. After the wheel has passed, residual stresses will develop. These residual stresses are protective in nature in that they reduce the tendency of plastic flow in the subsequent passes of the wheel. This, together with any effect of strain hardening, makes it possible

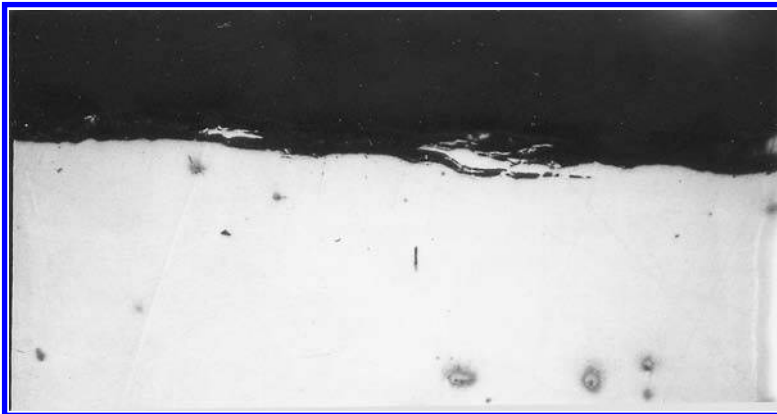


**FIGURE 5.9** Lip down of rail from the Stockholm test case showing plastic ratchetting (from Olofsson and Telliskivi<sup>13</sup>).

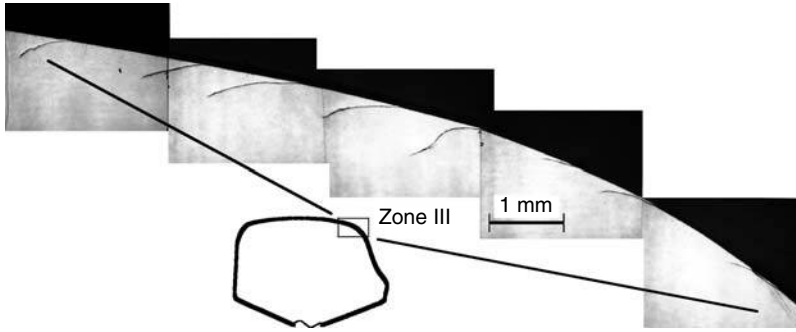
for the rail material to support stresses that are much higher than its elastic limit. This process is called elastic shakedown and the contact pressure limit below which this process is possible is known as the elastic shakedown limit. There is also a plastic shakedown limit. Loads between the elastic and plastic shakedown limit will lead to cyclic plasticity of the rail. If repeated, cyclic plastic deformation takes place and the rail material can cyclically harden, which leads to an increase in the yield stress and reduces the tendency of plastic flow.<sup>40</sup> For loads above the plastic shakedown limit, plastic ratchetting will occur, i.e., small increments of plastic deformation accumulate with each pass of the wheel.<sup>41</sup> Plastic ratchetting can be found in a curved track as a lip down of the rail gauge corner, as shown in Figure 5.9. Plastic ratchetting is the main cause of headcheck surface cracks.<sup>19</sup> The consequences of ratchetting are wear and the initiating of fatigue cracks as the material accumulates strain up to its limiting ductility. Beyond this limit, failed materials can separate from the surface as wear debris or forms crack like flaws, as shown in Figure 5.10.

### C. ROLLING CONTACT FATIGUE

Rolling contact fatigue cracks on the rail can be classified into those that are subsurface-initiated and surface-initiated. Subsurface-initiated cracks are often caused by metallurgical defects. On the other hand, surface initiated cracks seem to be the result of traffic intensity and axle load. A more specific division can be made into shelling, head checks, tache ovale, and squats. Shelling



**FIGURE 5.10** Micrograph from the Stockholm test case showing wear debris formation and crack like flaw Length of wear debris = 50  $\mu\text{m}$  (micrograph U. Olofsson).



**FIGURE 5.11** Head check cracks at gauge corner from the Stockholm test case (from Olofsson and Nilsson<sup>14</sup>).

(see Grassie and Kalousec<sup>42</sup>) is a subsurface defect that occurs at the gauge corner of the high rail in curves on railways with a high axle load. An elliptical shell-like crack propagates predominantly parallel to the surface. In many cases the shell causes metal to spall from the gauge corner. However, when the crack length reaches a critical value, the crack may turn down into the rail, giving rise to fracture of the rail. Head checks (Boulanger et al.<sup>43</sup>) generally occur as a surface initiated crack on or near the gauge corner in curves, as shown in Figure 5.11. Head checks may branch up towards the surface of the rail, giving rise to spalls. However, for reasons still not clearly understood, cracks can turn down into the rail and, if not detected, cause the rail to break. These events are rare, but are dangerous since surface cracks tend to form continuously.<sup>14</sup> Frederick<sup>44</sup> discusses the effect of train speed and wheel–rail forces as a result of surface roughness. Furthermore, he discusses whether hard rail or soft rails should be used in curves and also the relationship between wear rate and surface crack propagation. The conclusion was that hard rails are more prone to surface cracking. This was also seen in the Stockholm test case,<sup>14</sup> where UIC 900A rail material was compared against UIC 1100 rail material. Both materials seemed to be similarly sensitive to crack initiation, but the 1100 grade rail was more sensitive to crack propagation and also more sensitive to the formation of headcheck cracks. More information on the initiation mechanisms and growth of rolling contact fatigue cracks can be found in Beynon et al.<sup>45</sup> Tache ovale, or shatter cracks from hydrogen,<sup>42</sup> are defects that develop approximately 10–15 mm below the railhead from cavities caused by hydrogen. They can occur in the rail or in welds from poor welding practice. Development of tache ovale is influenced by thermal or residual stresses from roller straightening. Squats<sup>42,43</sup> occur on tangent tracks and in curves of large radius on the railhead and are characterised by the darkened area on the rail. Squats are surface initiated defects that can initiate from a white etching martensitic layer on the surface of the rail. Other mechanisms of squat formation are linked to longitudinal traction by wheels, which cause the surface layer of material to plastic ratchetting until a crack develops at the rail head.

Rolling contact fatigue cracks on wheels can be classified as shelling and spalling. Shelling is a subsurface rolling contact fatigue defect that occurs on the wheel thread and the mechanism is similar to the formation of shelling in rails. Spalling (Bartley<sup>46</sup>) can be initiated on the wheel thread surface when the wheel experiences gross sliding on the rail (braking). Large wheel surface temperatures above the austenization limit (720°C) can form martensite, a hard brittle steel phase. This brittle phase will easily fracture under following wheel passages and eventually result in spalling.

Surface coating of the track has been shown to reduce the advent of RCF cracking in the laboratory and full-scale tests are currently underway to establish if this behaviour is replicated in the field.<sup>47</sup>

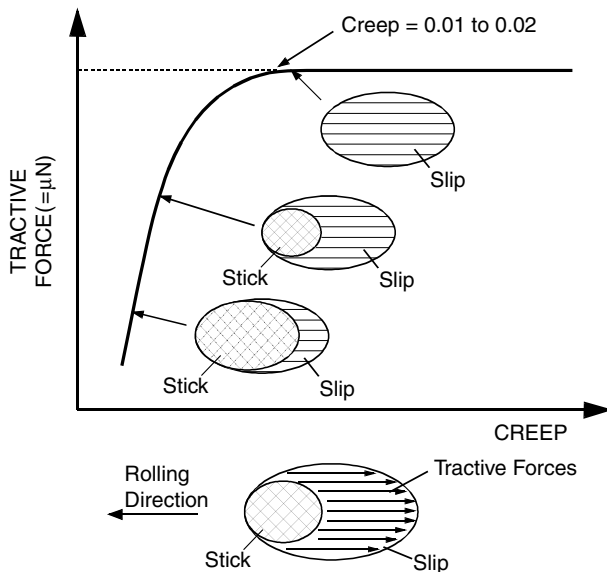
### IV. FRICTION

The friction force can be defined as the resistance encountered by one body moving over another body. This definition covers both sliding and rolling bodies. Note that even pure rolling nearly always involves some sliding and that the two classes of motion are not mutually exclusive. The resistive force, which is parallel to the direction of motion, is called the friction force. If the solid bodies are loaded together, the static friction force is equal to the tangential force required to initiate sliding between the bodies. The kinetic friction force is then the tangential force required to maintain sliding. Kinetic friction is generally lower than static friction.

For sliding bodies, the friction force, and thereby the coefficient of friction (friction force divided by normal force), depends on three different mechanisms in dry and mixed lubricated conditions: deformation of asperities, adhesion of the sliding surfaces, and ploughing caused by deterioration particles and hard asperities.<sup>48</sup> For most metal pairs, the maximum value of the coefficient of friction ranges from 0.3 to 1.0.<sup>49</sup> The ploughing component of the coefficient varies from 0 to 1.0 and the adhesion component varies from 0 to 0.4.<sup>50</sup> It is generally recognised that friction due to rolling of nonlubricated surfaces over each other is considerably less than dry sliding friction of the same surfaces.<sup>51</sup> For the steel wheel–steel rail contact, the rolling coefficient of friction is of the order of  $1 \times 10^{-4}$ .

As shown in Figure 5.12, the contact area between a wheel and rail can be divided into stick (no slip) and slip regions. Longitudinal creep and tangential (tractive) forces arise due to the slip that occurs in the trailing region of the contact patch. With increasing tractive force, the slip region increases and the stick region decreases, resulting in a rolling and sliding contact. When the tractive force reaches its saturation value, the stick region disappears, and the entire contact area is in a state of pure sliding. The maximum level of tractive force depends on the capability of the contact patch to absorb traction. This is expressed in the form of the friction coefficient,  $\mu$  (ratio of tractive force to normal load,  $N$ ). Normally, wheel–rail traction reaches a maximum at creep levels of 0.01 to 0.02.

The traction/creep curve can be dramatically affected by the presence of a third body layer in the wheel–rail contact. This could be formed either by a substance applied to increase/decrease



**FIGURE 5.12** Relationship between traction and creep in the wheel–rail contact.



friction (friction modifier or lubricant), or by a naturally occurring substance acting to decrease friction (water or leaves etc.). Hou et al.<sup>52</sup> have proposed a frictional model for rolling-sling contacts separated by an interfacial layer, which is based on the three rheological parameters: the shear moduli of elasticity ( $G$ ) and plasticity ( $k$ ) and the critical shear stress ( $\tau_c$ ). It shows that the friction is greatly affected by the rheology of the third body, slip distance and load with the shear stress vs. slip distance relationship exhibiting the dominant influence.

### A. WHEEL–RAIL FRICTION CONDITIONS

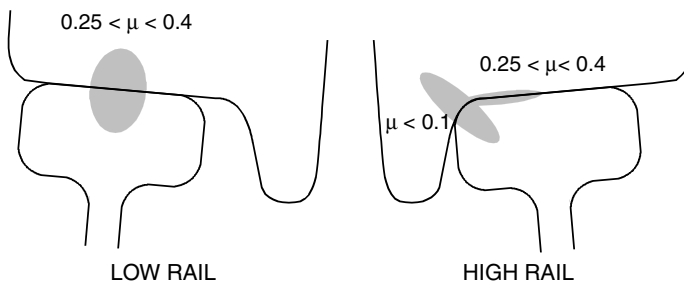
The friction between the wheels and rail is extremely important as it plays a major role in the wheel–rail interface process such as adhesion, wear, rolling contact fatigue, and noise generation. Effective control of friction through the application of friction modifiers to the wheel–rail contact is therefore clearly advantageous, although the process has to be carefully managed. The aim of friction management is to maintain friction levels in the wheel–rail contact to give<sup>53</sup>:

- Low friction in the wheel flange–rail gauge corner contact.
- Intermediate friction wheel tread–rail top contact (especially for freight trucks).
- High friction at the wheel tread–rail top contact for locomotives (especially where adhesion loss problems occur).

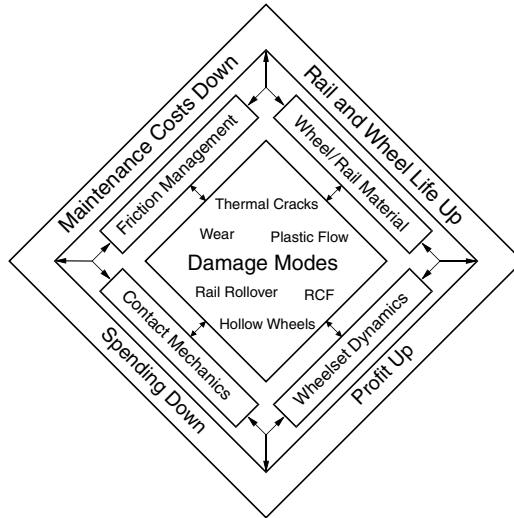
Ideal friction conditions in these contact regions for high and low rails are shown in Figure 5.13.<sup>54</sup> These are similar to values quoted for Canadian Pacific.<sup>55</sup>

Olofsson and Telliskivi<sup>13</sup> compared coefficients of friction measured on track and in the laboratory. For pure nonlubricated sliding tests the level is roughly the same, varying between 0.5 and 0.6. For a full-scale lubricated rail, the coefficient of friction was lower and varied between 0.2 and 0.4. Other results found in the literature support the measured coefficients of friction from the full-scale tests. In another project, the Swedish National Rail Administration studied how leaves on the track influenced the coefficient of friction using a special friction measurement train.<sup>56</sup> The reported coefficient of friction varied between 0.1 and 0.4. Harrison et al.<sup>57</sup> compared a hand-pushed rail tribometer and a TriboRailer that operated from a companion vehicle. For the hand-pushed tribometer, the coefficient of friction was typically 0.7 under dry conditions and varied between 0.25 and 0.45 under lubricated conditions. The Triborailer presented lower values of the coefficient of friction. Under dry conditions, the coefficient of friction was approximately 0.5, and varied between 0.05–0.3 under lubricated conditions.

To gain the greatest benefit from friction management and to ensure efficient train operation, the coefficient of friction needs to be integrated with the overall wheel–rail management system. It has been noted that, for example, it should be closely tied in with grinding schedules used in the maintenance of wheels and rails.<sup>53</sup> A scheme for the systematic approach to wheel–rail interface



**FIGURE 5.13** Ideal friction coefficients in the wheel–rail contact.



**FIGURE 5.14** Systems approach to wheel–rail interface research and development (adapted from Kalousec and Magel<sup>58</sup>).

research and development is shown in Figure 5.14, which emphasises the consideration of all aspects including materials, dynamics, etc., as well as friction. Nothing can really be treated in isolation.

## B. FRICTION MODIFICATION

Friction modifiers can be applied to the wheel–rail contact to generate the required coefficients of friction. These can be divided into three categories<sup>59</sup>:

- Low coefficient friction modifiers (lubricants) are used to give friction coefficients less than 0.2 at the wheel flange–gauge corner interface.
- High friction modifiers with intermediate friction coefficients of 0.2–0.4 are used in wheel tread–rail top applications.
- Very high friction modifiers (friction enhancers) are used to increase adhesion for both traction and braking.

Low friction modifiers can be solid or liquid (greases), the main difference between the two being the thickness of the film they form in the wheel/rail contact (solid lubricants will give a film of 10–30  $\mu\text{m}$  and grease lubricants less than 5  $\mu\text{m}$ ).<sup>53</sup> The primary application of these modifiers is in reducing friction in the wheel flange–rail gauge corner contacts, particularly in curves, where the contact conditions can be quite severe. The main focus of the remainder of this section is on low friction conditions and how to deal with them. Further discussion relating to reduction of friction can be found in the subsequent section on lubrication.

Friction modifiers are classified according to their influence after full slip conditions have been reached in the wheel–rail contact, as shown in Figure 5.15.<sup>60</sup> If friction increases after the saturation point, the modifiers have positive friction properties, if friction reduces, the modifier has negative friction properties. Positive friction modifiers can be described as high positive friction (HPF) or very high positive friction (VHPF), depending on the rate of increase in friction.

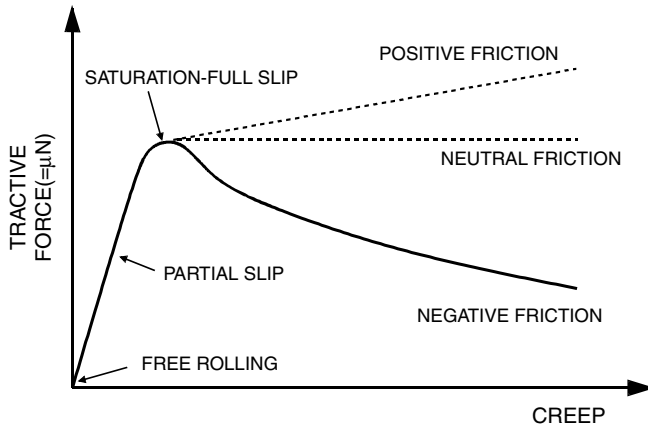


FIGURE 5.15 Behaviour of friction modifiers.

### C. ADHESION LOSS

Loss of friction or adhesion between the wheel and rail is particularly important as this has implications for both braking and traction. Poor adhesion in braking is a safety issue as it leads to extended stopping distances, and also in traction as it may lead to reduced acceleration which will increase the risk of a rear collision from a following train. In traction, however, it is also a performance issue. If a train experiences poor adhesion when pulling away from a station and a delay is enforced, the train operator will incur costs. Similar delays will occur if a train passes over areas of poor adhesion while in service.

A great deal of research was carried out on adhesion loss in the U.K. during the 1970s using both laboratory and field tests.<sup>61–65</sup> This identified the major causes of adhesion as being: water (from rainfall or dew), humidity, leaves, wear debris, and oil contamination.

Relative humidity has been shown to influence the frictional behaviour of a wide variety of materials.<sup>66</sup> By increasing the relative humidity, an absorbed layer of water molecules can be produced that can modify frictional behaviour. Relative humidity effects may also produce new chemical reactions on the surface together with other added substances.

The problems caused by leaves on the line remain prevalent today and each autumn can cause considerable delays to trains on the U.K. rail network. They are also a problem in Sweden, where it has been estimated by the Swedish National Railroad Administration that the cost of leaves on the rails is 100 million SEK annually (9 SEK  $\approx$  1 EUR).<sup>56</sup>

Work carried out on Japanese, American, and Canadian railways has re-emphasised the effect of the problems outlined above and identified further causes of adhesion loss, such as frost and mud deposited on rails by car wheels passing over level-crossings.<sup>67–69</sup> This work also showed the varying effects on adhesion of different types of leaves. Oily leaves, such as pine and cedar, caused a larger decrease in adhesion. Tunnels were also highlighted as being a problem, especially where water was leaking onto the track. Full-scale testing has also shown that weather conditions affect both the coefficient of friction and wear rates.<sup>13,15</sup>

Most of the work carried out in the U.K. was at relatively low speeds. Work on adhesion issues related to high speed lines, using both full-scale roller rigs and field measurements, has shown that adhesion decreases with train velocity and wheel–rail contact force.<sup>70,71</sup>

A number of experimental and theoretical investigations have revealed other significant parameters affecting adhesion. Chen et al.<sup>72</sup> carried out a detailed theoretical investigation of a water lubricated contact, studying the effect of rolling speed, slip, load, surface roughness, and

water temperature. The results indicated that the biggest influence on adhesion was the roughness of the wheel and rail surfaces (with adhesion rising with increased roughness). Third body effects due to material generated within the wheel–rail contact have been characterised by Niccolini and Bertier<sup>73</sup> and due to externally applied materials by Hou et al.<sup>52</sup> These can have a large influence on the adhesion, which is heavily dependent on the rheological properties of the layer formed in the contact. There is only limited data to validate these studies, but they give an important insight to aspects of the problem that are harder to evaluate in the field.

#### D. INCREASING ADHESION

While conditions leading to poor adhesion have been well investigated, methods for addressing the problems have not. The main adhesion enhancer used on railway networks world wide is sand. Sanding is used in train operations to improve adhesion in both braking and traction. In braking it is used to ensure that the train stops in as short a distance as possible. It usually occurs automatically when the train driver selects emergency braking. Sanding in traction, however, is a manual process. The train driver must determine when to apply the sand and how long the application should last.

The sand is supplied from a hopper mounted under the train. Compressed air is used to blow the sand out of a nozzle attached to the bogie and directed at the wheel–rail contact region (see Figure 5.16). In most systems the sand is blown at a constant flow rate, but some can provide a variable flow rate.

While sanding is effective and easy to use, it can potentially cause complex and costly problems relating to both rolling stock and track infrastructure. Sand application has been shown to increase wear rates of both wheel and rail materials by up to an order of magnitude.<sup>69,74,75</sup> Maintenance of sanders and control of sand build-up around track adhesion trouble spots are also issues that require particular attention.

Very high positive friction modifiers to enhance the coefficient of friction to 0.4–0.6 are available, but are really only in the development stage. There are a number of different products available, but most involve a solid stick of material that is applied directly to the wheel tread.

During autumn, when leaf fall occurs, leaf mulch is compressed in the wheel–rail contact and forms an extremely hard layer on the rail surface. This layer can cause adhesion loss problems, as already mentioned, but is also extremely hard to remove. A number of methods are used including using high pressure water-jets and blasting with Sandite (a mixture of sand and aluminum oxide particles), and a new system has now been developed that involves using a high power laser to burn away the layer. All of these, however, in the U.K., are applied by maintenance trains, of which there

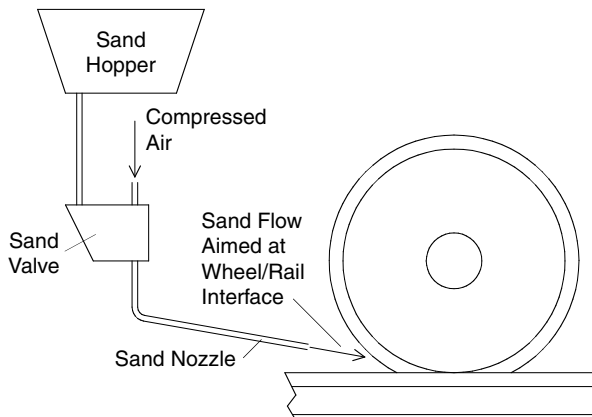


FIGURE 5.16 Sanding apparatus.

are very few, and gaining track access is extremely difficult. Water-jets and Sandite also have knock-on effects, which may be detrimental to the track infrastructure.

## V. LUBRICATION AND SURFACE COATINGS

This section focuses on the problems of high friction coefficients and how to reduce them using lubrication. High friction coefficients are most prevalent at the wheel flange–rail gauge corner contact, particularly in curves. Load and slip conditions are also high, which means that wear and rolling contact fatigue are more likely to occur at these sites. In order to reduce the wear problems, lubrication can be applied to reduce friction and alter the load bearing capacity. Lubrication, however, is also applied to alleviate other problems as will be shown. Surface coatings have also been applied to the track to address the problem of high friction.

### A. BENEFITS OF LUBRICATION

The benefits of lubrication have been well documented<sup>76,53</sup> and are concerned with the reduction of:

- wheel flange and rail gauge corner wear
- energy consumption
- noise generation.

Laboratory<sup>77–79</sup> and field tests<sup>14,31,80</sup> have all shown the wear reducing benefits of lubrication in the wheel–rail contact.

Fuel savings of approximately 30% (compared to dry conditions) have been reported for measurements taken on test tracks.<sup>81</sup> Other studies carried out in the field have shown improvements of a similar order of magnitude.<sup>82,83</sup>

### B. METHODS OF LUBRICATION APPLICATION

There are a number of different ways to apply lubricant:

- *Mobile lubricators*: these are basically railway vehicles designed to apply lubricant to the gauge corner of the track.
- *Wayside lubricators*: these are mounted next to the track and apply lubricant to the rail gauge corner. There are three types: mechanical, hydraulic, and electronic.
- *On-board lubricators*: these apply grease or solid lubricant or spray oil on to the wheel flange, which is then transferred to the gauge corner of the rail. Complex control systems are used in the application process to avoid the application of lubricant at inappropriate locations.

Mechanical wayside lubricators rely on the wheel making contact with a plunger, which operates a pump. The pump supplies lubricant from a reservoir to a distribution unit. The lubricant is then picked up by the wheel flange and distributed along the rail. Problems exist because there is only a single circuit so if a failure occurs the lubricant supply is ineffective. Mechanical lubricators have a low initial cost because of their simple design, but require good maintenance to remain effective. Hydraulic lubricators have been found to more reliable, but have some of the same problems as their mechanical counterparts.

Electronic lubricators use sensors to detect the approach of a train and activate electric pumps to deliver the lubricant. They are inherently more reliable than mechanical or hydraulic lubricators and can also be adjusted away from the track.

On-board lubricators supply lubricant to the wheel flange–rail gauge corner. In most designs the lubricant is deposited on the wheel flange and spread along the rail, although in some the

lubricant is directly applied to the rail. Grease or oil spray systems are used that employ complex control strategies using sensors measuring vehicle speed and track curvature to govern lubricant application. Solid stick lubricators are also available, in which a stick of lubricant is spring loaded against the wheel flange.

On board systems have a number of advantages over wayside lubricators<sup>54</sup>:

- reduced safety risk exposure to staff during installation, inspection and maintenance.
- easier inspection and maintenance (carried out in more controlled conditions).
- the rail will continue to receive some friction control protection in the event of the failure of an individual on board lubricator.

Despite these advantages, at problem tracks, site wayside lubricators will still be a necessity.

### C. PROBLEMS WITH LUBRICATION

Problems with lubrication systems have been found to be related to both technical and human issues.<sup>84</sup> The main technical problems with wayside lubricators have been highlighted as: blocked applicator openings; leaking holes; ineffective pumps and trigger mechanisms; and poor choice of lubricant. Human related problems can result from the technical issues. If over lubrication occurs and lubricant migrates onto the rail top, adhesion loss can occur. Train drivers may then be tempted to apply sand to compensate and increase friction, however, this will lead to increased wear and could cause the applicators to become blocked. The thought that the application of lubricant will lead to wheel slip can also lead train drivers to switch off on board lubrication systems.

Some of the consequences of poor wayside lubrication have been listed as<sup>55</sup>:

- wheel slip and loss of braking (and potentially, wheel flats and rail burn)
- poor train handling
- prevention of ultrasonic flaw detection
- wastage of lubricant
- high lateral forces in curves and subsequent increase in wear.

Other than adhesion problems, over-lubrication can cause an increase in rolling contact fatigue crack growth on the rail gauge corner.<sup>22,23</sup> This can be due to pressurisation of the crack leading to increased growth rates or because reduced wear means that cracks are truncated less. However, full-scale test results from narrow curves show that well-maintained lubrication could reduce both the wear rate and the propagation rate of surface cracks.<sup>14</sup>

### D. LUBRICATOR SYSTEM SELECTION AND POSITIONING

The effectiveness of a lubrication system is affected by a number of parameters including the climate, the railway operating conditions, the dispensing mechanism, and the maintenance of the lubricating equipment. Clearly, selection of the most appropriate type of lubricator and lubricant is very important, but also the positioning of the lubricator is critical to its successful operation.

The key characteristics required of a lubricant are<sup>55</sup>:

- Lubricity or the ability of the lubricant to reduce friction (although of greater importance is the effect on wear).
- Retentivity or the measure of time over which the lubricant retains its lubricity. Flash temperatures in the wheel–rail contact can be as high as 600 to 800°C, these lead to the lubricant in the contact being burned up. The retentivity is therefore a function of the

loads and creepages seen at the lubrication site as these dictate the temperature in the contact.

- Pumpability or how easily the lubricant can be applied to the track. The temperature is an issue here as some track locations will experience a wide range across which some lubricants may not maintain their pumpability. Some networks use different lubricants in the winter and summer for this reason.

Laboratory tests have been developed to assess the wear reducing capacity of lubricants and energy saving potential.<sup>77–79</sup> These are good for screening and ranking purposes and selecting those lubricants suitable to take forward for field trials.

Monitoring the effectiveness of lubricants in the field, either during trials or in actual practice, is clearly essential. This will provide information necessary to decide on a lubrication strategy during trials or in maintaining performance once implemented. Measurements of friction can be taken using tribometers, either hand-propelled along the track or train mounted. The hand-propelled equipment is useful for monitoring short stretches of track. Obviously, for long stretches, a train or vehicle mounted system is preferable. It should be noted here that it has been shown that the benefits of lubrication may take some time to become evident on installation of a lubrication system.<sup>85</sup>

Correct positioning of a wayside lubricator is critical to providing effective lubrication. Each site will require something different, which makes this task quite complex. Controlled field testing has been used to assess the reliability and efficiency of wayside lubricators based on a number of factors related to the lubricant including: waste prevention; burn up; distance covered; washing off by rain or snow; and migration to the rail top. This data and factors related to the track, such as length of curve, gradient and applicator configuration, and traffic, including direction, types of bogie, axle loads and speeds, have been combined to develop criteria and a model for positioning wayside lubricators.<sup>86</sup>

Ultimately, however, the most critical element in preserving effective lubrication is maintenance. Once in place, wayside lubricators need regular maintenance to prevent the problems outlined occurring.

## E. SURFACE COATINGS

Coating the rail surface is now being investigated as a means to control friction and reduce wear, rolling contact fatigue and noise problems.<sup>47</sup> This is quite new technology, coatings have previously been applied to wheels to address these problems. If successful, this could address some of the problems evident with lubrication supply and reduce maintenance requirements.

## ACKNOWLEDGMENTS

This work formed part of the activities of the Railway group KTH at the Department of Machine Design KTH in Stockholm Sweden. The partners in the Railway group KTH are Banverket, Green Cargo, SL Infrateknik, Bombardier Transportation, SJ and Traintech Engineering.

## REFERENCES

1. *ASM Handbook, Friction, Lubrication and Wear Technology*, 10th ed., Vol. 18, 1992.
2. Hutchings, I. M., *Tribology Friction and Wear of Engineering Materials*, Edward Arnold, London, 1992.
3. Johnson, K. L., *Contact Mechanics*, Cambridge University Press, Cambridge, 1985.
4. Dowson, D. and Higginson, G. R., *Elasto-Hydrodynamic Lubrication*, Pergamon, New York, 1977.
5. Dowson, D., *History of Tribology*, Longman, London, 1979.

6. Le The Hung, *Normal und tangential spannungsberuchnung beim rollen den kontakt fur rotationskörper mit nichtelliptischen kontakt flächen*, Fortschrifts berichte VDI 12, 1987.
7. Marshall, M., Lewis, R., Dwyer-Joyce, R., Olofsson, U., and Björklund, S., *Ultrasonic Charaterisation of a Wheel Contact, Leeds–Lyon Symposium on Tribology*, Lyon, 2–5 September, 2003.
8. Björklund, S., Olofsson, U., Marshall, M., Lewis, R., and Dwyer-Joyce, R., *Contact Pressure Calculations on Rough Measured Surfaces, Metrology and Properties of Engineering Surfaces*, Halmstad, Sweden, September, pp. 10–11, 2003.
9. Poole, W., *The measurement of contact area between opaque objects under static and dynamic rolling conditions, Contact mechanics and wear of the wheel/rail system*, Wv press, University of Rhode Island, pp. 59–72, 1987.
10. Kalker, J. J., *Three-Dimensional Elastic Bodies in Rolling Contact*, Kluwer, Dordrecht, 1990.
11. Telliskivi, T. and Olofsson, U., Contact mechanics analysis of measured wheel–rail profiles using the finite element method, *J. Rail Rapid Transit*, 215, 65–72, 2001.
12. Olofsson, U. and Nilsson, R., *Initial Wear of a Commuter Train Track, Nordtrib' 98 Ebeltoft*, Denmark, June 7–10, 1998.
13. Olofsson, U. and Telliskivi, T., Wear, friction and plastic deformation of two rail steels — full scale test and laboratory study, *Wear*, 254, 80–93, 2003.
14. Olofsson, U. and Nilsson, R., Surface cracks and wear of rail: a full scale test and laboratory study, *J. Rail Rapid Transit*, 216, 249–264, 2002.
15. Nilsson, R., *Wheel/Rail Wear and Surface Cracks*, Licentiate Thesis, KTH, Stockholm 2003.
16. Telliskivi, T., *Wheel–Rail Interaction Analysis*, Doctoral Thesis, KTH, Stockholm 2003.
17. Jendel, T., *Prediction of Wheel Profile Wear — Methodology and Verification*, Licentiate thesis, KTH, Vehicle Engineering, Sweden, 2000.
18. Alwahdi, F., *Wear and Rolling Contact Fatigue of Ductile Material*, Doctoral Thesis, Sheffield University, Sheffield, UK 2004.
19. Ringsberg, J., *Rolling Contact Fatigue of Railway Rails with Emphasis on Crack Initiation*, Doctoral Thesis, Chalmers University of Technology, Göteborg, 2000.
20. Knothe, K., Theiler, A., and Güney, S., *Investigation of Contact Stresses on the Wheel/Rail-System at Steady State Curving, 16th IAVSD Conference*, Pretoria, South Africa, 30 August–3 September, 1999.
21. Kumar, S., Adenwale, Y. S., and Rajkumar, B. R., Experimental investigation of contact stresses between a U.S. locomotive wheel and rail, *J. Eng. Ind.*, 105, 64–70, 1983.
22. Cassidy, P. D., *Variation of Normal Contact Stresses for Different Wheel/Rail Profile Combinations*, European Rail Research Institute, ERRI D 173/DT 336, 1996.
23. Cannon, D. F. and Pradier, H., Rail rolling contact fatigue — research by the European rail research institute, *Wear*, 191, 1–13, 1996.
24. Williams, J. A., Wear modelling: analytical, computing and mapping: a continuum mechanics approach, *Wear*, 225–229, 1–17, 1999.
25. Olofsson, U., Andersson, S., and Björklund, S., Simulation of mild wear in boundary lubricated spherical roller thrust bearings, *Wear*, 241, 180–185, 2000.
26. Sullivan, J. L., Boundary lubrication and oxidational wear, *J. Phys. Appl. Phys.*, 19, 1999–2011, 1986.
27. Olofsson, U. and Svedberg, G., *Low Concentration Level Contaminant-Related Wear in Sliding and Rolling Contacts*, 2nd World Tribology Congress Vienna September 3–7, 2001.
28. Andersson, S. and Salas-Russo, E., The influence of surface roughness and oil viscosity on the transition in mixed lubricated sliding steel contacts, *Wear*, 174, 71–79, 1994.
29. Dizdar, S., Wear transition of a lubricated sliding steel contact as a function of surface texture anisotropy and formation of boundary layers, *Wear*, 237, 205–210, 2000.
30. Lewis, R. and Dwyer-Joyce, R., Wear mechanisms and transitions in railway wheel steels, *J. Eng. Tribol., Proceedings of the IMechE Part J*, 218, 467–478, 2004.
31. Waara, P., *Wear Reduction Performance of Rail Flange Lubrication*. Licentiate thesis LTU, Mechanical Engineering, 2000.
32. Engel, S., *Reibungs- und Erdmudungsverhalten des Rad-Schiene-System mit und ohne Schmierung*. Ph.D. thesis Otto von Guericke Universität Magdeburg, 2002.
33. Cantara, F., Investigation of wheel flange wear on the Santander FEVE rail — a case study, *Wear*, 162, 975–979, 1993.



34. Lewis, R. and Olofsson, U., Mapping rail wear transitions, *Wear*, 257, 721–729, 2004.
35. Perz-Unzueta, A. J. and Beynon, J., Microstructure and wear resistance of perlitic rail steels, *Wear*, 162, 173–182, 1993.
36. Garnham, J. E. and Beynon, J. H., Dry rolling-sliding wear of bainitic and perlitic steels, *Wear*, 157, 81–109, 1992.
37. Mitao, S., Yokoyama, H., and Yamamoto, S., *High Strength Bainitic Steel Rails for Heavy Haul Railways with Superior Damage Resistance*, IHHA'99 STS Conference, 1999.
38. Johnson, K. L., *The Mechanism of Plastic Deformation of Surface and Subsurface Layers in Rolling and Sliding Contact*, *Material Science Forum: The Role of Subsurface Zone in Wear of Materials*, Trans Tech Publications, Switzerland, pp. 33–40, 1988.
39. Jones, C. P., Tyfor, W. R., Beynon, J. H., and Kapoor, A., The effect of strain hardening on shakedown limits of a pearlitic rail steel, *J. Rail Rapid Transit*, 211, 131–140, 1997.
40. Schleinzner, G., *Residual Stress Formation During the Roller Straightening of Rail*. Ph.D. thesis, Montanuniversität Leoben Austria 2000.
41. Kapoor, A. and Johnson, K. L., Plastic ratchetting as a mechanism of metallic wear, *Proc. R. Soc. (Lond.) A*, 445, 367–381, 1994.
42. Grassie, S. and Kalousek, J., *Rolling Contact Fatigue on Rails: Characteristics, Causes and Treatments*, *Proceedings of the 6th IHHA Conference*, Capetown, South Africa, pp. 381–404, 1997.
43. Boulanger, D., Girardi, L., Galtier, G., and Baudry, G., *Prediction and Prevention of Rail Contact Fatigue*, *IHHA'99 STS Conference*, Moscow, Russia, June, pp. 14–17, 1999.
44. Frederick, C. O., *Future Rail Requirements*, *Proceedings of Rail Quality and Maintenance for Modern Railway Operation*, Kluwer Academic Publishers, The Netherlands, pp. 3–14, 1993.
45. Beynon, J. H., Brown, M. W., and Kapoor, A., *Initiation, Growth and Branching of Cracks in Railway Track*, *Proceedings of Engineering Against Fatigue*, Balkema, Rotterdam, The Netherlands, pp. 461–472, 1999.
46. Bartley, G. W., *A Practical View of Wheel Thread Shelling*, Ninth International Wheelset Congress, Montreal, 1966.
47. Heinsch, E. J. M., Franklin, F. J., Nielson, J. C. O., Ringsberg, J. W., Weeda, G. J., Kapoor, A., and Josefson, B. L., Prevention of RCF damage in curved rail through development of the INFRA–STAR two-material rail, *Fatigue Fract. Eng. Mater. Struct.*, 26, 1007–1017, 2003.
48. Suh, N. P. and Sin, H. C., The genesis of friction, *Wear*, 69, 91–114, 1981.
49. Czichos, H., Presentation of friction and wear data, In *Friction, Lubrication and Wear Technology*, ASM Handbook, Vol. 18, Blau, P. J., Ed., 1992.
50. Suh, N. P., *Tribophysics*, Prentice-Hall International, New Jersey, USA, 1986.
51. Harris, T. A., *Rolling Bearing Analysis*, Wiley, New York, 1991.
52. Hou, K., Kalousek, J., and Magel, E., Rheological model of solid layer in rolling contact, *Wear*, 211, 134–140, 1997.
53. Zakharov, S., *Wheel/Rail Performance, Guidelines to Best Practice for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia, USA, 2001.
54. Sinclair, J., Friction Modifiers, in “Vehicle Track Interaction: Identifying and Implementing Solutions”, IMechE Seminar, February 17th, 2004.
55. Roney, M. D., *Maintaining Optimal Wheel and Rail Performance, Guidelines to Best Practice for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, Virginia, USA, 2001.
56. Forslöv, L., *Wheel slip due to leaf contamination*. Swedish National Rail Administration TM 1996 03 19, Borlänge, Sweden (published in Swedish), 1996.
57. Harrison, H., McCanney, T., and Cotter, J., *Recent Development in COF Measurements at the Rail/Wheel Interface, Contact Mechanics and Wear of Rail/Wheel Systems*, Tokyo, Japan, 25–27 July, 2000.
58. Kalousek, J. and Magel, E., Optimising the wheel rail system, *Rail. Track Struct.*, January 1997.
59. Kalousek, J. and Magel, E., Modifying and managing friction, *Rail. Track Struct.*, May 1999.
60. Eadie, D. T., Kalousek, J., and Chiddick, K. C., *The Role of High Positive Friction (HPF) Modifier in the Control of Short Pitch Corrugation and Related Phenomena*, *Proceedings of the 5th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems*, Tokyo, pp. 36–41, 2000.

61. Collins, A. H. and Pritchard, C., Recent research on adhesion, *Rail. Eng. J.*, 1(1), 19–29, 1972.
62. Broster, M., Pritchard, C., and Smith, D. A., Wheel/rail adhesion: it's relation to rail contamination on British railways, *Wear*, 29, 309–321, 1974.
63. Beagley, T. M. and Pritchard, C., Wheel/rail adhesion — the overriding influence of water, *Wear*, 35, 299–313, 1975.
64. Beagley, T. M., McEwen, I. J., and Pritchard, C., Wheel/rail adhesion — the influence of railhead debris, *Wear*, 33, 141–152, 1975.
65. Beagley, T. M., McEwen, I. J., and Pritchard, C., Wheel/rail adhesion — boundary lubrication by oily fluids, *Wear*, 33, 77–88, 1975.
66. Demizu, K., Wadabayashi, R., and Ishigaki, H., Dry friction of oxide ceramics against metals: the effect of humidity, *Tribol. Trans.*, 33, 505–510, 1990.
67. Nagase, K., A study of adhesion between the rails and running wheels on main lines: results of investigations by slipping adhesion test bogie, *J. Rail Rapid Transit, Proceedings of the IMechE Part F*, 203, 33–43, 1989.
68. Logston, C. F. and Itami, G. S., Locomotive friction-creep studies, *Trans. ASME, J. Eng. Ind.*, 102, 275–281, 1980.
69. Jenks, C.W., *Improved Methods for Increasing Wheel/Rail Adhesion in the Presence of Natural Contaminants*, Transit Co-operative Research Program, Research Results Digest, No. 17, 1997.
70. Chen, W., Wu, J., and Jin, X., Wheel/rail adhesion and analysis by using full scale roller rig, *Wear*, 253, 82–88, 2002.
71. Ohyama, T., Tribological studies on adhesion phenomena between wheel and rail at high speeds, *Wear*, 144, 263–275, 1991.
72. Chen, H., Ban, T., Ishida, I., and Nakahara, T., Adhesion between rail/wheel under water lubricated contact, *Wear*, 253, 75–81, 2002.
73. Niccolini, E. and Bertier, Y., *Progression of the Stick/Slip Zones in a Dry Wheel/Rail Contact: Updating Theories on the Basis of Tribological Reality*, *Proceedings of the 29th Leeds–Lyon Symposium on Tribology*, Elsevier Tribology Series No. 41, pp. 845–853, 2003.
74. Lewis, R. and Dwyer-Joyce, R.S., *Wheel–Rail Wear and Surface Damage Caused by Adhesion Sanding*, *Proceedings of the 30th Leeds–Lyon Symposium on Tribology*, Elsevier Tribology Series No. 43, pp. 731–741, 2004.
75. Kumar, S., Krishnamoorthy, P. K., and Prasanna Rao, D. L., Wheel–rail wear and adhesion with and without sand for a North American locomotive, *J. Eng. Ind., Trans. ASME*, 108, 141–147, 1986.
76. Marich, S., Makie, S., and Fogary, R., *The Optimisation of Rail/Wheel Lubrication Practice in the Hunter Valley*, *Proceedings of the RTSA Technical Conference*, Adelaide, 2000.
77. Clayton, P., Danks, D., and Steele, R. K., Laboratory assessment of lubricants for wheel/rail applications, *Lubr. Eng.*, 45(8), 501–506, 1989.
78. Zhao, X. Z., Zhu, B. L., and Wang, C. Y., Laboratory assessment of lubricants for wheel/rail lubrication, *J. Mater. Sci. Technol.*, 13(1), 57–60, 1997.
79. Alp, A., Erdemir, A., and Kumar, S., Energy and wear analysis in lubricated sliding contact, *Wear*, 191, 261–264, 1996.
80. Waara, P., Lubricant influence on flange wear in sharp railroad curves, *Ind. Lubr. Tribol.*, 53(4), 161–168, 2001.
81. Reiff, R. and Creggor, D., Systems Approach to best Practice for Wheel and Rail Friction Control, International Heavy Haul Conference, 1999.
82. Allen, R. A., Mims, W. E., Rownd, R. C., and Singh, S. P., Energy savings due to wheel rail lubrication — seaboard system test and other investigations, *J. Eng. Ind., Trans. ASME*, 107, 190–196, 1985.
83. Samuels, J. M. and Tharp, D. B., *Reducing Train Rolling Resistance by On-Board Lubrication*, *Proceedings of the 2nd Rail and Wheel Lubrication Symposium*, Memphis, USA, 1987.
84. Thelen, G. and Lovette, M., A parametric study of the lubrication transport mechanism at the rail–wheel interface, *Wear*, 191, 113–120, 1996.
85. Reiff, R. P., *Rail–Wheel Lubrication, A Strategy for Improving Wear and Energy Efficiency*, *Proceedings of the 3rd International Heavy Haul Conference*, Vancouver, 1986.
86. de Koker, *Development of a Formula to Place Rail Lubricators*, *Proceedings of the 5th International Tribology Conference*, 1994.

---

# 6 Track Issues

*Tore Dahlberg*

## CONTENTS

I.	The Railway Track as a Dynamic System .....	144
A.	The Track and Its Components.....	144
B.	Rails.....	145
C.	Railpads .....	145
D.	Sleepers (or Crosssties) .....	146
E.	Ballast.....	146
F.	Subballast .....	146
G.	Geotextiles.....	146
H.	Subgrade .....	146
II.	Function of the Track.....	147
A.	To Guide the Train.....	147
B.	To Carry the Load.....	147
III.	Dynamic Properties of the Track.....	147
A.	Nonlinear Track .....	149
B.	Train Moving on Track–Excitation Sources of Train and Track Vibrations .....	150
C.	Excitation Frequencies .....	151
D.	Railhead Corrugation (Short Wavelength Periodic Irregularities) .....	151
1.	Classification of Railhead Irregularities .....	152
2.	Consequences of Rail Corrugation .....	152
3.	Origin of Rail Corrugation.....	152
E.	Long Wavelength Irregularities .....	154
1.	Sleeper Spacing .....	154
2.	Wheel Out-of-Roundness .....	154
3.	Rail Manufacturing .....	154
4.	Track Stiffness Irregularities .....	154
5.	Track Embankment Settlements .....	155
F.	Impact Loads.....	157
1.	Wheel Flats.....	157
2.	Measurements with Wheelset Having Wheel Flats.....	157
3.	Rail Joints.....	157
4.	Switches.....	158
G.	Mathematical Modelling of Track Dynamics .....	159
1.	Beam (Rail) on Continuous Elastic Foundation (Winkler Beam) .....	159
2.	Vehicle–Bridge Interaction (Moving Mass on Simply Supported Beam).....	160
3.	Beam (Rail) on Discrete Supports .....	161
4.	Discretely Supported Track Including Ballast Mass .....	162

- 5. Rails on Sleepers Embedded in Continuum.
  - Three-Dimensional Finite Element Models..... 163
- H. Modelling of Dynamic Train–Track Interaction and Computer Program Developments..... 163
  - 1. Frequency-Domain Techniques ..... 163
  - 2. Time-Domain Techniques..... 164
  - 3. Computer Program Developments..... 165
- IV. Dynamic Properties of Track Components ..... 165
  - A. The Rail..... 166
  - B. Mathematical Modelling of Rails ..... 166
  - C. Railpads and Fastenings..... 168
  - D. The Sleepers..... 169
    - 1. Sleeper Vibrations ..... 169
    - 2. Elastic Foundation..... 170
    - 3. Measurements and Calculations..... 170
  - E. Ballast, Subballast and Subgrade..... 171
    - 1. Track Settlement ..... 172
    - 2. Research on Ballast..... 173
    - 3. Modelling Track Settlement..... 174
- V. Summary ..... 175
- Acknowledgments ..... 175
- References..... 175

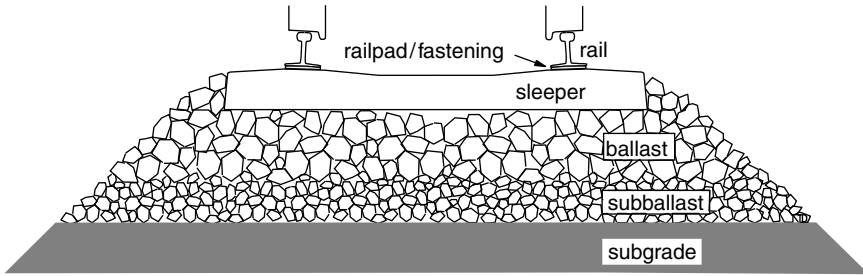
**I. THE RAILWAY TRACK AS A DYNAMIC SYSTEM**

The purpose of a railway track is to guide the trains in a safe and economic manner. The track and the switches should allow smooth passage of the trains. If the track is not perfectly levelled and aligned, the irregularities will cause oscillations or vibrations of the train, and this may induce discomfort for passengers and damage for goods. Long wavelength undulations of the track will give rise to low-frequency oscillations of the train and short wavelength irregularities cause vibrations and noise, both in the train and in the environment. Oscillations, vibrations, and noise may become unpleasant for the train passengers and for people in the vicinity of the railway line.

This chapter focuses on different aspects of track dynamics and train–track interaction. Dynamic properties of the track as a whole and the different components of the track will be investigated. In the first part of this chapter the dynamics of the complete track structure will be discussed including the dynamics of the compound train–track system, followed by how dynamic properties of some track components are treated.

**A. THE TRACK AND ITS COMPONENTS**

A railway track structure consists of rails, sleepers, railpads, fastenings, ballast, subballast, and subgrade, see [Figure 6.1](#). Sometimes, for example, in tunnels, the ballast bed is omitted and the rails are fastened to concrete slabs resting on the track foundation. Two subsystems of a ballasted track structure can be distinguished: the superstructure, composed of rails, sleepers, ballast, and subballast, and the substructure (subgrade, subground) composed of a formation layer and the ground. First, a short description of the different parts of the ballasted track and their functions will be given. Then, the function of the complete track structure will be discussed.



**FIGURE 6.1** Track with its different components: rails, railpads, and fastenings (fastenings not shown in this figure, see Figure 6.2), sleepers, ballast, subballast, and subgrade (subground).

## B. RAILS

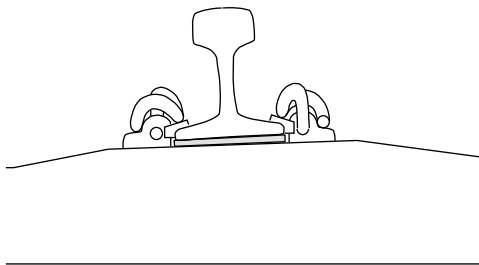
A modern steel rail has a flat bottom and its cross section is derived from an I-profile. The upper flanges of the I-profile have been converted to form the railhead. The English engineer Charles Vignoles has been credited the invention of this design in the 1830s. In Europe, one commonly used rail profile is the UIC60 rail (with a Vignoles profile), where 60 refers to the mass of the rail in kg per meter.

The rails should provide smooth running surfaces for the train wheels and they should guide the wheelsets in the direction of the track. The rails also carry the vertical load of the train and distribute the load over the sleepers. Lateral forces from the wheelsets, and longitudinal forces due to traction and braking of the train should also be transmitted to the sleepers and further down into the track bed. The rails also act as electrical conductors for the signalling system.

## C. RAILPADS

In a railway track with concrete sleepers, railpads are placed between the steel rails and the sleepers, see Figure 6.2. The railpads protect the sleepers from wear and impact damage, and they provide electrical insulation of the rails. Wooden sleepered tracks may not have rail pads.

From a track dynamics point of view, the railpads play an important role. They influence the overall track stiffness. When the track is loaded by the train, a soft railpad permits a larger deflection of the rails and the axle load from the train is distributed over more sleepers. Also, soft railpads isolate high-frequency vibrations. They suppress the transmission of high-frequency vibrations down to the sleepers and further down into the ballast. A stiff railpad, on the other hand,



**FIGURE 6.2** Rail fastened to sleeper with railpad inserted between the rail and the sleeper. Rail is fastened to the sleeper by fastening clips, with an electrical insulation between the rail and the clip.

gives a more direct transmission of the axle load, including the high-frequency load variations, down to the sleepers below the wheels.

#### **D. SLEEPERS (OR CROSSTIES)**

The sleepers provide support of the rails and preserve gauge, level, and alignment of the track. The sleepers transmit vertical, lateral, and longitudinal forces from the rail down to the ballast bed. They should also provide electrical insulation between the two rails.

Nowadays, ballasted railway tracks are usually constructed with monobloc concrete sleepers (for example, twin-bloc concrete sleepers are used in France). Timber sleepers have been used almost since the beginning of railway construction. In countries where the timber price is acceptable, timber sleepers are still frequently used. Sometimes, it is suitable to use steel sleepers.

#### **E. BALLAST**

Coarse stones are used to form the bed (a ballast bed) of the railway track. The sleepers to which the rails are fastened are embedded in the ballast. The ballast layer supports the track (the rails and the sleepers) against vertical and lateral forces from the trains. It is tightly compacted or tamped around the sleepers to keep the track precisely levelled and aligned. The standard depth of ballast is 0.3 m, but it is packed to 0.5 m around the sleeper ends to ensure lateral stability. Traditionally, angular, crushed, uniformly graded hard stones and rocks (granite, limestone, slag, or other crushed stone) have been considered good ballast materials. However, availability and economic motives have often been prime factors considered in the selection of ballast materials. From a physical point of view, the ballast materials and their interactions are complex. Constitutive laws of ballast materials are under development.

#### **F. SUBBALLAST**

Subballast is material chosen as a transition layer between the upper layer of large-particle, good quality ballast and the lower layer of fine-graded subgrade. The subballast used in most new constructions is intended to prevent the mutual penetration of the subgrade and the ballast and to reduce frost penetration. Any sand or gravel materials may serve as subballast material as long as they meet necessary filtering requirements.

#### **G. GEOTEXTILES**

Sometimes, geotextiles are used to prevent the intermixing of the subgrade and the subballast. Geotextiles are permeable geomembranes of synthetic fibers. They are used to separate two consecutive layers of granular materials and/or to reinforce a soil layer of insufficient mechanical strength. They can also be used as filters or drainage.

#### **H. SUBGRADE**

Subgrade, or formation, is a surface of earth or rock levelled off to receive a foundation for the track bed. Sometimes, an extra layer, a formation layer, is put on the earth so as to give the correct profile of the track bed. The subballast and ballast layers rest on this material. The subgrade is a very important component in the track structure and has been the cause of track failure and poor track quality, Li and Selig.<sup>1</sup> Unfortunately, in existing tracks the subgrade is not involved in the maintenance operations, and once the track has been laid, little can be done to alter its characteristics, Chrismer and Read.<sup>2</sup>

## II. FUNCTION OF THE TRACK

The track should guide the train. It should bring the train safely along the track and through switches to the destination. It should also carry the load of the train and distribute the load over an area that is as large as possible.

### A. TO GUIDE THE TRAIN

The main function of the track is to guide the train. The two wheels of a wheelset are rigidly connected to the wheel axle. The wheel treads are conical in order to steer the wheelset. If the wheelset is not exactly on the centreline of the track, one wheel will have a larger rolling radius than the other. Due to the rigid connection of the two wheels (via the axle), the wheelset turns towards the centre of the track, and, having passed the centreline, the second wheel will have the larger rolling radius forcing the wheelset back towards the centre. This leads to a sinusoidal movement of the wheelset along the track. It also promotes better radial adjustment of the wheelset in curves. If the lateral movement of the wheelset becomes too large, the flanges of the wheels will prevent derailment. This movement will induce a low-frequency lateral movement of the railway vehicle and lateral forces on the track.

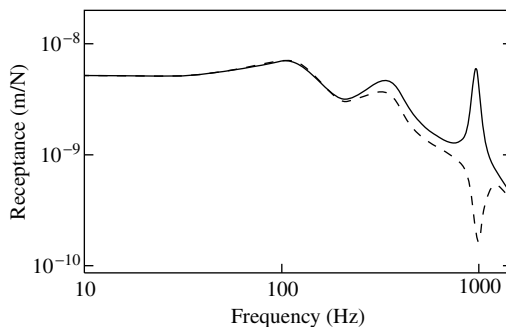
### B. TO CARRY THE LOAD

Another function of the track is to carry the load of the train and to distribute the load over an area of the subgrade that is as large as possible. In a conventional railway track the rails distribute the wheel–rail contact forces over several sleepers. The sleepers, supported by the ballast, transmit the load via the sleeper base area to the ballast, and the ballast disperses the load over a larger area of the subballast and the subgrade. A wheel–rail contact force of, say, 10 t, that is applied over a few square centimetres at the wheel–rail contact patch, is distributed over an area of the subgrade that is, probably, more than one square metre.

## III. DYNAMIC PROPERTIES OF THE TRACK

In this section, the dynamics of the entire track structure will be investigated. In Section IV the dynamics of some components of the track will be examined in more detail.

One way to investigate the dynamic properties of a railway track is to load the track with a sinusoidal force. At frequencies up to approximately 200 Hz, this can be carried out by using hydraulic cylinders. If one wants to investigate the track response at higher frequencies, the track may be excited by an impact load, for example, from a sledgehammer. Figure 6.3 shows a typical



**FIGURE 6.3** Typical track receptances when rail is loaded with a sinusoidally varying force. Track receptance when rail is loaded between two sleepers (full-line curve) and above one sleeper (dashed curve) versus loading frequency. The maxima indicate resonance frequencies in the track structure.

receptance curve of a track. The receptance is the ratio of the track deflection and the force put on the track, thus giving deflection in metres per Newton of the load. The receptance is the inverse of the track stiffness, cf. Figure 6.6, where the track stiffness variation along a track is shown. As can be seen in Figure 6.3, the receptance, and thus also the track stiffness, will depend on the frequency of the load. Normally, the receptance also depends on the preload on the track (preload is a static load superimposed to the dynamic load), because most tracks have a nonlinear relationship between load and deflection. (In Figure 6.3 the scale on the ordinate is given just to indicate an overall dynamic range. The curves give the principal behaviour only; these curves differ from one track to another.)

Several well-damped resonances can be found in a track structure. Sometimes, when the track is built on a soft ground, one resonance may appear in the frequency range 20 to 40 Hz, Oscarsson<sup>3</sup> (this resonance is not shown in the figure). This is a resonance when the track, and a great deal of the track substructure, vibrates on, for example, a layered structure of the ground. The track structure itself thus plays a minor role for this resonance. These vibrations can be perceived several metres away from the track. To bring this resonance into a track model, Oscarsson needed to introduce inertia terms from the subground into the dynamic model, see Section III.G.4.

One track resonance is usually obtained in the frequency range 50 to 300 Hz. This resonance is obtained when the track structure (rails and sleepers) vibrates on the ballast bed. The rails and the sleepers provide the “mass” and the ballast provides the “spring” for this resonance vibration. The ballast also provides a large amount of damping, so this resonance is very well damped, see the first, very flat maximum in Figure 6.3 at a frequency slightly above 100 Hz.

Another resonance can often be found in the frequency range 200 to 600 Hz. This resonance is explained by the rail bouncing on the railpads. The railpad acts as a spring inserted between two masses: the rail and the sleeper. Here, the ballast provides most of the damping.

The highest resonance frequency discussed here is the so-called pinned–pinned resonance frequency. This is the resonance that can be seen at approximately 1000 Hz in Figure 6.3. The resonance peak is narrow, indicating that the resonance at this frequency is very lightly damped. The pinned–pinned frequency occurs when the wavelength of the bending waves of the rail is twice the sleeper spacing. In this case, the bending vibration of the rail has nodes at the supports, i.e., at the sleepers. This explains why the pinned–pinned frequency is so lightly damped; mainly the material damping of the steel itself (the rail material) is involved in damping this vibration and very little vibration energy is transmitted to the surroundings (and then mainly as propagating waves along the rail and almost nothing to the ballast).

In Figure 6.3 it can be seen that the track behaves quite differently if the receptance is measured between two sleepers (the full-line curve) or above a sleeper (the dashed curve). When loading the track halfway between two sleepers, and when the track is excited at the pinned–pinned frequency, the pinned–pinned resonance is easily excited. There will be a large deflection of the rail at a unit amplitude force excitation. This gives a peak at 800 to 1000 Hz in the receptance curve (the peak shown by the full-line curve in Figure 6.3 at approximately 1000 Hz). On the other hand, if the rail is excited above a sleeper, i.e., at a node of the pinned–pinned vibration, then the track will be very stiff at this point. Evidently, the track has an anti-resonance at this point (and this frequency). At this excitation, the track deflection must be symmetric with respect to a vertical line through the point of loading. This implies that the slope of the rail is zero at the loaded sleeper. On either side of the load, the rail vibrates with what is approximately the pinned–pinned vibration mode (only very close to the loading point, where the slope is zero, the vibration mode differs from the pinned–pinned mode). This rail vibration then acts as a dynamic vibration absorber making it difficult for the load to excite that point of the track. (Thus, if the pinned–pinned vibration mode had not been there, neither the peak nor the dip would have appeared in the receptance curves at the pinned–pinned frequency.)

If the Euler–Bernoulli beam theory is used, the pinned–pinned frequency can easily be estimated. The frequency  $f$  (Hz) (or angular frequency  $\omega$  in radians per second) is the same as the



fundamental frequency of a simply supported beam of length  $L$ , and it is:

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi} \pi^2 \sqrt{\frac{EI}{mL^4}} \quad (6.1)$$

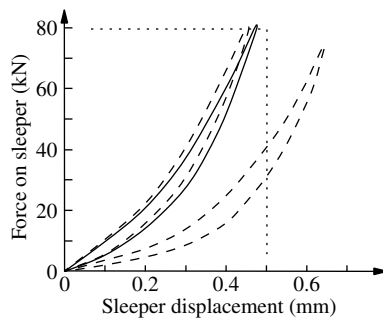
where  $EI$  is the bending stiffness of the rail,  $m$  is the mass of the rail per unit length, and  $L$  is the sleeper spacing. For a UIC60 rail one has, approximately,  $m = 60$  kg/m and  $EI = 6.4$  MNm<sup>2</sup>. If sleeper spacing  $L$  is  $L = 0.65$  m, one obtains  $f = 1214$  Hz. This is, as said, an estimation of the pinned–pinned frequency. In practice, this resonance frequency will be lower than the one calculated here.

The reason why the real pinned–pinned frequency is lower than the frequency given by the formula above is that the Euler–Bernoulli beam theory is not very accurate when the wavelength  $\lambda$  of the vibration is short. The Euler–Bernoulli theory is accurate only when the wavelength of the vibration is much longer than the height of the beam, and here this condition is not fulfilled. The height  $h$  of the rail cross section of a UIC60 rail is  $h = 172$  mm, and this is not much shorter than the wavelength  $\lambda$  of the pinned–pinned frequency (here  $\lambda = 2L = 1.3$  m = 7.5  $h$ ). In practice, shear deformation of the rail and rotatory inertia should also be included to give a better estimation of the pinned–pinned frequency. This is done in the Rayleigh–Timoshenko beam theory. Using this theory, a lower resonance frequency is obtained (this is discussed further in Section IV.A below). The Rayleigh–Timoshenko beam theory gives, for a UIC60 rail and in this frequency range, a frequency that is 20 to 25% lower than the frequency given by the Euler–Bernoulli theory.

Higher resonance frequencies also occur in the track structure, see Section IV below. These frequencies are of importance for the noise emission of the track, and they are treated in the chapter dealing with noise in this handbook.

## A. NONLINEAR TRACK

The discussion above concerns linear tracks, i.e., the relationship between the track displacement and the force loading the track is linear. In practice, however, many tracks are nonlinear, giving a nonlinear relationship between the load on the track and the track deflection. Figure 6.4 shows measurements performed by Banverket (the Swedish National Rail Administration) on a newly built track in Sweden. The rail has been loaded with a sinusoidal force and the part of the force that was transmitted to the sleeper below the load was measured (this part is little more than 50%). Then, the load on the sleeper was plotted as a function of sleeper displacement. The measurements were repeated three times at three adjacent sleepers giving the three curves in Figure 6.4. It is



**FIGURE 6.4** Measured sleeper displacement when track is loaded with a sinusoidal force (measurements performed by Banverket).

clearly seen in the figure that the relationship between the load on a sleeper and the sleeper displacement is nonlinear, and it is also seen that the ballast stiffness is different below the different sleepers. An extreme case of nonlinearity is when a sleeper is hanging in the rail with a gap between the sleeper and the ballast. One of the sleepers in Figure 6.4 could be such a sleeper that is not fully supported by the ballast.

The nonlinearity implies that the receptance of a track (Figure 6.3) looks different at different load levels. If the track is loaded by a static force (a static preload), and superimposed to that of a dynamic (sinusoidal) force, then the receptance of the track will depend on the preload; a large preload makes the track stiffer, i.e., the curve in Figure 6.3 moves downwards, and many resonance frequencies shift to a higher frequency.

## B. TRAIN MOVING ON TRACK—EXCITATION SOURCES OF TRAIN AND TRACK VIBRATIONS

The vibrations and the frequencies discussed so far refer to the track alone. In the case of a train loading the track, the springs and masses of the train (wheelsets, suspensions, bogie frames, and so on) contribute to the dynamic behaviour of the compound train–track system and other resonance frequencies will appear. One such resonance is when the wheelset vibrates as a mass using the track stiffness as the “spring.” This resonance may fall in the frequency range 30 to 100 Hz.

One “spring” in the train–track system is obtained at the wheel–rail contact. The elastic deformation at the contact patch can be seen as deformation of a spring with nonlinear characteristics. The stiffness of the “contact spring” increases with increasing wheel–rail contact force. This spring will influence the high-frequency behaviour of the wheel–rail system. (The spring is called a Hertzian spring, named after Hertz who investigated the contact between elastic bodies.)

When the train moves along the track, especially at high speed, other dynamic phenomena appear. Heavier and faster trains induce ground vibrations that are transmitted to buildings along the railway lines. This brings discomfort to people due to noise and vibrations generated in the buildings, Jonsson.<sup>4</sup>

Also, when train speed increases, the intensity of railway-generated noise and vibration generally becomes higher. If the track support is soft, for example, when the track is built on a layer of clay, then the train speed may exceed the velocity of the Rayleigh surface wave in the ground. This will induce a tremendous increase of the vibration level, and train speed must be limited (this happened on the Swedish West Coast Line, where the track had to be removed and the subground strengthened before the trains were allowed to pass at high speed). This topic has been thoroughly investigated by, for example, Krylov.<sup>5</sup>

A large number of excitation sources exist that may induce oscillations, vibrations, and noise in the train and in the track and its surroundings. For the track, four different (and independent) geometric errors, or irregularities, may develop. These errors can be in track alignment, track level, track gauge, and the cant of the track (cant is the superelevation of the outer rail in relation to the inner rail of the track in curves). Long wavelength geometric irregularities in the track alignment will induce lateral displacements of the railway cars, and this will induce travelling discomfort for the passengers. Short wavelength irregularities will induce vibrations and noise. The same can be said for long and short wavelength irregularities of the track level (vertical profile). Irregularities in the track gauge will also induce lateral displacements of the railway cars. Deficiencies or excesses in the cant will induce lateral forces in curves.

Several sources contribute to make the track geometry deteriorate. Track settlement will induce long wavelength irregularities of the track both vertically and laterally. A rail may not be absolutely straight when it comes from the manufacturer and the railhead top surface may not be absolutely plane. Also, when in use, short wavelength periodic irregularities (also called corrugations) may develop on the railhead due to the train traffic. The corrugation generates high-frequency vibration of the rail and the wheel, and noise is emitted.

One source generating train and track vibrations is wheel out-of-roundness. A wheel is seldom perfectly round because when it is used, periodic irregularities may develop on the wheel tread. These irregularities may be of either long wavelengths (so-called polygonalisation with one to five wavelengths around the wheel) or of short wavelength. Especially, short wavelength irregularities often appear on block-braked wheels.

Another source of vibration excitation is the irregular track stiffness. For example, it has already been mentioned that a track is stiffer at a sleeper and more flexible between two sleepers. Thus, there will be a larger deflection of the rail when the wheel is between two sleepers than when it is above one. This induces an excitation of the train and the track at the “sleeper passage frequency.” This frequency depends on the sleeper distance and the train speed, and it may, of course, induce resonance vibrations in the train and in the track.

### C. EXCITATION FREQUENCIES

A wheel running with speed  $v$  (m/s) over a sinusoidal rail irregularity of wavelength  $\lambda$  (m) will perceive an excitation frequency  $f$  (in hertz, Hz) that is:

$$f = \frac{v}{\lambda} \quad (6.2)$$

With vehicle speeds  $v = 5$  to 60 m/s and with irregularity wavelengths  $\lambda = 0.030$  to 0.300 m one obtains excitation frequencies in the frequency interval  $f = 17$  to 2000 Hz. Longer wavelengths give lower frequencies. The excitation will then induce vibrations and noise in the train and in the track structure and the environment. Resonances in the train–track system will also amplify these vibrations at the resonance frequencies.

Three different sources generating train and track vibrations will now be discussed in some detail. These are railhead corrugation (short wavelength irregularities) in Section III.D, long wavelength irregularities in Section III.E and impact loads in Section III.F.

### D. RAILHEAD CORRUGATION (SHORT WAVELENGTH PERIODIC IRREGULARITIES)

Periodic irregularities of various wavelengths sometimes develop on the railhead and on the wheel tread. Railhead irregularities seem to appear on almost every kind of railway track, from heavy haul tracks to lightweight metropolitan lines. The development of irregularities on the railhead is explained by the dynamics of the train–track system and linked to resonance effects in the combined rail–wheelset system. It is assumed to be so, because the irregularities that develop are periodic of a certain wavelength; mostly in the range 30 to 300 mm. Such periodic irregularities are named corrugation. When the train speed increases, the dynamic interaction between the vehicle and the track becomes more and more pronounced and this gives rise to larger dynamic forces between the wheels and the rails. Once the rail corrugation has begun to develop, the dynamic forces in the train–track system will be further magnified, and the deterioration rate of the track will increase.

Following the wavelength of the periodic irregularities, these are classified in different ways. Here, the irregularities of very short wavelengths are called corrugation; then come short wavelength irregularities and long wavelength irregularities. Unfortunately, different authors use different classifications, and sometimes the word corrugation is used for all wavelengths. In this section, the word corrugation is used for short wavelength railhead periodic irregularities of lengths 300 mm or less. A recent review of studies on rail corrugation was presented by Sato et al.,<sup>6</sup> where a short historical survey on rail corrugation is also given.

## 1. Classification of Railhead Irregularities

Alias<sup>7</sup> gives an overview of different types of wave formations on the railheads. The irregularities are divided into three categories: corrugation, with wavelength 30 to 80 mm and with amplitudes of a few hundredths of a mm; short waves, with wavelength 150 to 300 mm and with amplitudes up to 1 mm; and long waves, with wavelength up to 2 m. Only the longer wavelengths can be explained by the manufacturing process. Lévy<sup>8</sup> classifies the lengths of the waves as: short waves, 0 to 250 mm; medium waves, 150 to 600 mm; and long waves, 0.3 to 2 m. Lévy also presents a system to measure and analyse undulatory rail wear. Grassie and Kalousek<sup>9</sup> classify the corrugation by mechanisms. Two mechanisms are identified in their paper. First the wavelength-fixing mechanism, i.e., a mechanism that establishes a certain wavelength of the corrugation, and secondly, the damage mechanism, which is a mechanism that generates the unevenness (wear, plastic deformation, etc.). Six classes of corrugation are identified and the authors name the different types of corrugation after where they appear: heavy haul, light rail, booted sleeper, contact fatigue, rutting, and roaring rails. The authors claim that the wavelength fixing mechanism is known in all but one class; the roaring rails. Treatments to avoid or reduce the problem are proposed, Kalousek and Grassie.<sup>10</sup>

Alias<sup>7</sup> refers to the UIC (l'Union Internationale des Chemins de Fer, International Union of Railways) catalogue of rail defects. Two definitions of corrugation are given. One says that "corrugation is characterised by an almost regular sequence of shiny peaks and dark troughs generally spaced approximately 30 to 80 mm apart". For another type of corrugation it is stated that "in this wave effect there is no difference in appearance between the peaks and troughs of the waves", and "the wavelength generally varies between about 80 and 300 mm". Of the latter, the longer waves (300 mm) occur preferentially on the low rails in curves. Thus, railhead irregularities are divided into two groups here: corrugation (30 to 80 mm) and short wavelength irregularities (80 to 300 mm). This corrugation thus develops on the railhead due to the train traffic.

## 2. Consequences of Rail Corrugation

Clark et al.<sup>11</sup> developed a mathematical model to describe the dynamics of the vehicle-track system. They also performed experiments with two specially prepared test rails; this involved grinding simulated corrugations 60 mm long over a 6 m length of the rails. Their findings were that the rail vibrates out of phase with the corrugations so that the cyclic irregularity seen by the wheel is minimised. This means that all the motion is being induced into the track because the inertia of the rail is so much less than that of the wheel. At certain speeds, sleeper resonances will be excited by the corrugation wavelength. Likely effects of sleeper resonances are ballast degradation and track settlement with loss of vertical track profile. Also, railpads and fastenings (clips) may deteriorate due to the many alternating load cycles, and sleepers may be exposed to fatigue damage due to the cyclic stresses set up in the sleepers in the resonance situation.

## 3. Origin of Rail Corrugation

The origin of corrugations on rails has not yet been fully explained. Resonance effects between rail, wheel, and axle are believed to be involved. Also, resonance vibration of the portion of the rail between the front and rear wheel of a bogie has been found to be involved in the corrugation growth process, Igeland.<sup>12</sup> Only the long waves (order of metres) can be explained by the rail manufacturing process.

Several hypotheses have been suggested to explain the development of rail corrugation. So far, however, no generally accepted explanation has been given. Probably, different phenomena are involved in creating corrugation of different wavelengths. What makes the corrugation phenomenon especially remarkable is that there does not seem to be any direct relationship between the train

speed and the corrugation wavelength. One idea put forward by some authors is that the short-term dynamic behaviour of the train–track system causes long-term wear to develop. Other authors have tried to refine dynamic models of wheels and track wishing to find a corrugation-initiating mechanism in terms of the short-term behaviour of the wheel–track system. Thus, many authors have tried many different hypotheses to explain the origin of rail corrugation. Some hypotheses will now be presented. (Note that this section is intended as a presentation of different hypotheses put forward to explain the rail corrugation phenomenon. The references given are samples only; for a complete review, please refer to the paper by Sato et al.<sup>6</sup>)

Clark and Foster<sup>13</sup> put forward the theory that self-excited vibration characteristics of a flexible wheelset and track system under high creepage conditions may provide possible explanations for the formation of corrugations on the running surface of the rails.

Alias<sup>7</sup> gives some characteristics of wave formation in rails. Some rails have a succession of polished shiny areas which contrast with a duller base on their running surfaces. The shiny patches are 30 to 80 mm apart and are generally quite regularly spaced. The shiny, raised parts correspond to the highest points of the bumps; the dark parts, i.e., the troughs, are oxidised. Metallographic analyses reveal that the peaks have a hardened martensitic structure to a very slight depth (from a few hundredths of a millimetre to 0.2 to 0.3 mm). This structure is explained by slippage of the wheels. The slippage brings about a sharp increase in temperature of the rail surface followed by rapid cooling due to thermal conduction to the surrounding material. As martensite occupies a greater volume than a structure in equilibrium, this is assumed to explain the formation of the peaks. This explanation of the formation of the peaks can be compared with the explanation of the formation of the troughs given by other authors, see for example Igeland.<sup>14</sup> The formation of the troughs is explained by slippage and wear. The wheel–rail friction force and the irregular normal force create the troughs by periodic slippage and wear.

Werner<sup>15</sup> pointed out that the corrugation wavelengths are predominantly integer fractions of the wheel width, and thus can be related to standing surface shear waves in the wheel tread surface.

In a paper by Frederick,<sup>16</sup> it was shown how measured relationships to quantify the dynamic response of the wheel and rail to vertical, lateral, and longitudinal forces can be combined with formulae for the rate of rail wear and wheel–rail creepage to predict whether or not a small wave in the surface of the rail will be deepened or erased by passing axles.

In an approach by Brockley and Ko,<sup>17</sup> it was proposed that corrugations are formed by wear resulting from torsional vibration of the drive wheels and by longitudinal vibration of the rail.

Valdivia,<sup>18</sup> Knothe and Valdivia,<sup>19</sup> and Knothe and Ripke,<sup>20</sup> suggested that corrugation formation can be explained as a feedback process between (a) the wheel and rail high frequency oscillations and (b) long-term wear phenomena. This work has been continued in, for example, Hempelmann and Knothe<sup>21</sup> and in Müller.<sup>22</sup> Müller also introduced a “contact mechanical filter” effect indicating that the wheel–rail contact itself could promote corrugation.

Suda and Iguchi<sup>23</sup> stated that the corrugation growth conditions depend on the natural frequencies of the system. By appropriate selection of the natural frequencies, it would be possible to reduce the development of corrugation of one wavelength without inducing growth of corrugation of other wavelengths. They also stated that the rolling direction and the speed have a great influence on the corrugation growth. Scaled and full-scale roller test stands have been used, Sato et al.<sup>6</sup>

Kalousek<sup>24</sup> suggested that corrugations develop from tiny surface cracks in the rail and that rail life can be substantially increased by preventive grinding.

A linear model giving wear rate as a function of frequency was suggested by Tassilly and Vincent.<sup>25,26</sup> A transfer function between the initial wheel and rail roughness and the wear rate spectrum in the contact patch was presented. Under some conditions, the initial roughness on the rail was shown to degenerate into corrugation in some frequency bands. The model has been used as a tool to design track modifications to prevent the growth of corrugations.

Ideas including plastic deformation of the railhead are put forward by Cervoni and Vincent,<sup>27</sup> Bogacz et al.,<sup>28</sup> and Suda and Iguchi.<sup>23</sup> Böhmer and Klimpel<sup>29</sup> used mathematical models to study the influence of work hardening and residual stresses on corrugation.

Igeland<sup>12</sup> based the corrugation calculation on a semiempirical contact mechanics relationship between creep, friction force, and normal force. The vertical contact forces between the moving wheels and the rails were calculated in the time domain, allowing for a nonlinear Hertzian wheel–rail contact stiffness. It was found that the two wheelsets of a travelling bogie interact via the rail between the wheels, and that both the front and rear wheel of the bogie should be considered in the corrugation growth analysis. In Nielsen<sup>30</sup> a validation of an integrated mathematical model to field observations was made and the agreement between simulations and field observations was good.

## E. LONG WAVELENGTH IRREGULARITIES

Long wavelength irregularities are defined here as irregularities of wavelengths of 300 mm or longer. These irregularities may be either geometric irregularities in the track or on the wheel, or irregularities of the track stiffness.

### 1. Sleeper Spacing

Even though the rail itself may not have an irregularity with a wavelength corresponding to the sleeper spacing, the wheel will be influenced by the varying stiffness of the track. As the track is stiffer at the sleepers, and softer in-between, the wheel will be excited with a frequency that corresponds to the speed  $v$  of the train and the sleeper spacing  $\lambda$ . The excitation frequency  $f$  becomes  $f = v/\lambda$ .

### 2. Wheel Out-of-Roundness

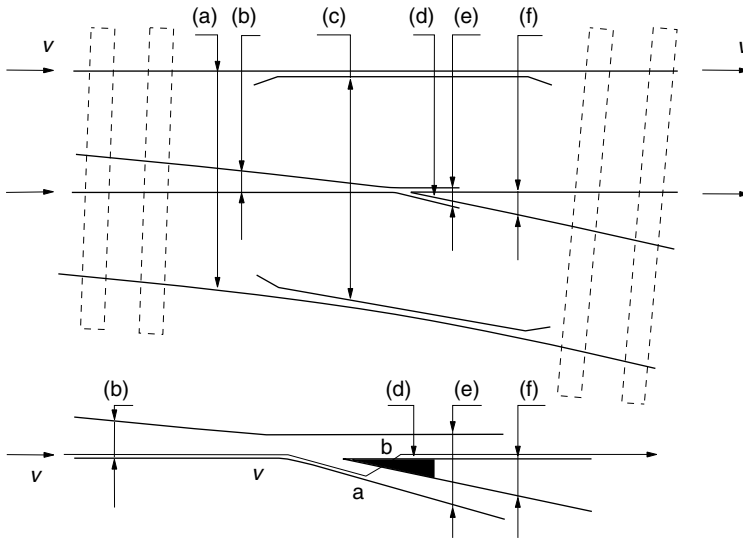
An out-of-round wheel will excite the train–track structure with one or several frequencies corresponding to the wavelengths of the wheel out-of-roundness. An eccentric wheel will induce the frequency corresponding to the number of wheel revolutions per second (a wheel of radius  $r = 0.45$  m gives an irregularity of wavelength  $\lambda = 2.8$  m). An oval wheel will induce twice that frequency (wavelength  $\lambda/2$ ), and so on. Also, nonperiodic irregularities may appear on the wheel tread. Especially if the wheel is equipped with block brakes, short wavelength corrugation may develop. A survey is given in Nielsen and Johansson.<sup>31</sup> In Johansson and Nielsen,<sup>32</sup> the influence of different types of out-of-roundness on the vertical dynamic wheel–rail contact force and track response is investigated through extensive field tests and numerical simulations.

### 3. Rail Manufacturing

The rail manufacturing process may induce rail irregularities. The wavelength of this irregularity is of the order of one to several metres. These irregularities will induce a low-frequency excitation of the train and the track.

### 4. Track Stiffness Irregularities

Track stiffness variation due to the sleeper spacing has already been mentioned. Other places along the track having variable stiffness are at switches and turnouts. At switches, the sleepers have different lengths and different spacings, and this influences the track stiffness. The symmetry of the track is lost at switches, implying that the left and right rail will have different stiffnesses (the stock rail keeps the stiffness of the track, whereas the switch rail becomes stiffer because of the longer sleepers supporting that rail, see [Figure 6.5](#)).



**FIGURE 6.5** Railway turnout: (a) stock rails, (b) switch rails, (c) guard rails, (d) nose, (e) wing rails, and (f) nose rails. Some sleepers are indicated by dashed lines. The entrance and the exit of a wheelset passing straight through the turnout are indicated by arrows. The wheel path through the crossing is shown in the lower, close-up figure. The wheel leaves the wing rail at point a and enters the nose at point b.

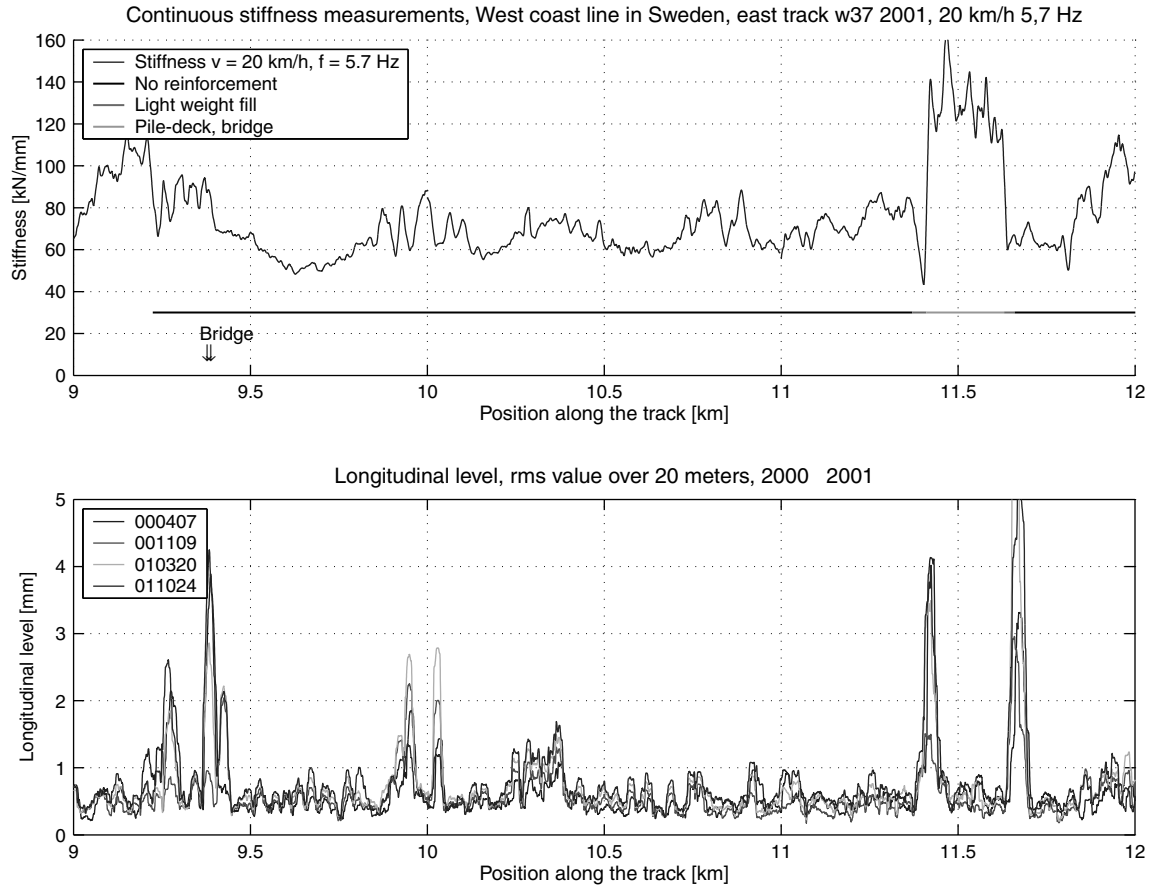
Especially if a switch is equipped with a manganese crossing (frog), the rail bending stiffness ( $EI$ ) will change dramatically at the crossing. Such a sudden change of the stiffness will induce transient and high-frequency vibrations in the train and in the track. The mass is also larger at the crossing making inertia forces larger, see Andersson and Dahlberg.<sup>33</sup>

The track superstructure is seldom built on a homogeneous substructure. Due to a variable stiffness of the track subgrade, the track stiffness experienced by the train will vary along the track. As this stiffness is more or less random along the track, it will induce low-frequency random oscillations of the train.

Nowadays, it is possible to measure the vertical track stiffness continuously along the track. Banverket, the Swedish National Rail Administration, has developed a trolley by which the track can be loaded and track stiffness measured while moving along the track at a speed of up to 30 km/h, (see Figure 6.6). A new track stiffness measurement car, by which measurements can be performed at speeds of 70 km/h, has been developed, Berggren.<sup>34</sup> Measured track stiffness along a distance of 3000 m is shown in Figure 6.6. It is seen in the figure that there may be rather rapid changes of the track stiffness, and it is also seen that the track subground has a large influence on the track stiffness. It is seen in the figure (km 11.4 to 11.65) that a pile deck and a bridge make the track very stiff, whereas the light-weight fill makes it very flexible (at km 11.4). A transition area from an embankment to a bridge is also a place where large and rapid changes of the track stiffness may occur.

## 5. Track Embankment Settlements

When a track is loaded by the weight of the train and, superimposed on that, high-frequency load variations, the ballast and subground may undergo nonelastic deformations. After the load has passed, the track will not return exactly to its original position but to a position very close to the original one. After thousands and thousands of train passages, all these small nonelastic deformations will add, differently in different parts of the track, to give a new track position.



**FIGURE 6.6** Track stiffness (axle load divided by track deflection) along railway track as measured by the Banverket track stiffness measurement trolley, and four measurements (during two years) of the longitudinal level of the track (positive downwards, meaning that a large peak in the curve indicates a local settlement of the track). Figure provided by Eric Berggren at Banverket.



This phenomenon is called differential track settlement. The track alignment and the track level will change with time. Depending on the subground, the wavelength of these track irregularities will be of the order of metres up to hundreds of metres. The uneven track will induce low-frequency oscillations of the train. Gradually, the track load variations will increase and so will the track settlement. Especially, the transition area from an embankment to a bridge is a place where track settlements tend to occur, see [Figure 6.6](#). In the lower figure (of [Figure 6.6](#)) it can be seen that large track settlements occur at 9.4 km due to a bridge and at 11.4 and 11.7 km due to an embankment and light-weight fill.

## F. IMPACT LOADS

In the above two sections (Section III.D and Section III.E), short wavelength irregularities on the railhead and long wave-length irregularities of the track have been discussed. In this section, distinct single impact loadings of the track will be discussed. Such loadings could come from, for example, wheel flats or other defects on the wheel tread, rail joints, and switch crossing (frog) passages.

### 1. Wheel Flats

It sometimes happens that due to problems with the brakes at a wheel (or due to low adhesion), a wheel does not start to rotate immediately as the train starts or the wheel stops rotating before the train comes to a standstill. Instead, the wheel slides along the rail for a while. When the wheel slides, part of the wheel tread is removed and a flat surface is formed. Later, when the wheel starts to rotate, an impact load will hit the rail at every revolution of the wheel. This load will induce high-frequency vibrations of the track and noise will be emitted. Sometimes the impact load may be so high that the rail is cut off (this may occur at low temperatures). Wheel flat impacts have been investigated by, among others, Fermér and Nielsen.<sup>35</sup>

### 2. Measurements with Wheelset Having Wheel Flats

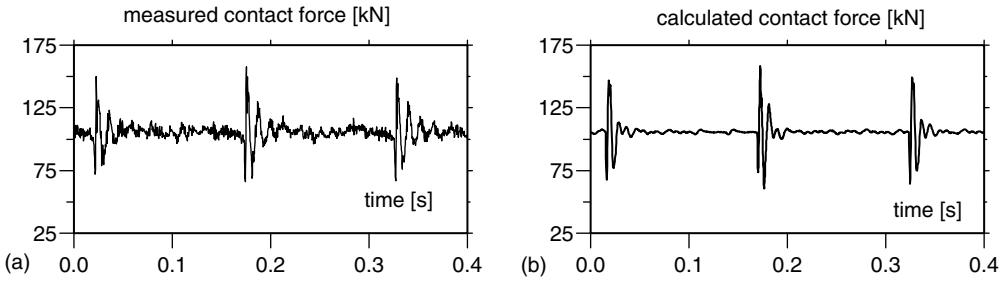
Measurements with an instrumented wheelset prepared with wheel flats have been performed by the Swedish National Rail Administration (Banverket). Some of these measurements were reported in Fermér and Nielsen.<sup>35</sup> The wheelset had standard wheels with solid wheel discs. Strain gauges on the wheel discs were used to measure the wheel–rail contact force. After a first test series without wheel flats, artificial wheel flats of 40 mm length were ground on the two wheel treads. To maintain symmetry, the wheel flats were ground with the same position on the left and the right wheel.

In [Figure 6.7](#), measured (figure to the left) and calculated (figure to the right) contact forces between the wheel and the rail are shown. The calculations were performed using the computer program DIFF, developed for calculation of the dynamic interaction between the train and the track, Nielsen<sup>36</sup> (see also [Section III.H](#) below). The results shown in [Figure 6.7](#) are for train speed  $v = 70$  km/h and for axle load 22 t. It is seen that this relatively small wheel flat induces a large impact load on the rail even if the train speed is not very high.

In [Figure 6.8](#), the bending stress in a sleeper due to the load impact from the wheel flat is shown (measured and calculated, respectively). The wheel flat hits the rail above the sleeper in this case. It is concluded from the figures that the wheel flat causes large dynamic contributions to the contact force and to the stresses in the sleeper.

### 3. Rail Joints

In the case of welded rail joints, there may be a small difference in level of the two rails connected at the joint. When a wheel, especially at high speed, “climbs” this difference in level, large dynamic forces will arise (or more correctly: as the wheel mass is larger than the rail mass, the wheel will

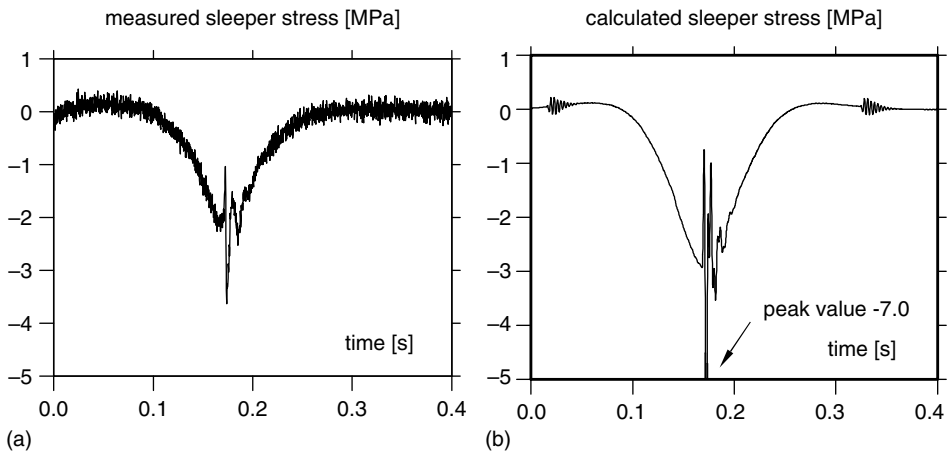


**FIGURE 6.7** (a) Measured contact force between wheel and rail when wheel has a wheel flat (left figure), and (b) calculated contact force between wheel and rail (right figure) (Source: From Fermér, M., and Nielsen, J.C.O., *Vehicle Syst. Dyn.*, 23, pp. 142–157, August 23–27, 1993).

“press down” the rail, but this will, of course, also create large dynamic forces). In the case of fishplated joints, or other expansion joints, not only the problem of difference in level appears, but also an irregularity of stiffness will occur at the joint.

#### 4. Switches

When a train passes a switch, dynamic effects will appear due to several factors. As mentioned above, sleepers have another spacing and a different length at the switch, giving a change of track stiffness. When a wheel passes the switch blade and the crossing (the frog, see Figure 6.5), the wheel–rail contact patch will suddenly move laterally on the wheel tread. This implies that the wheel will move laterally to find its new equilibrium position, inducing a longitudinal force on the wheel and on the track and thus a yaw moment on the wheelset. At the crossing, where the wheel moves over from the wing rail to the nose rail (or vice versa) the geometry of the crossing and the wheel profile seldom fit exactly to give a smooth transition. Instead, due to different heights of the wing rail and the nose, this transition often induces a vertical impact load on the wheel and on the crossing, Andersson and Dahlberg.<sup>33</sup>



**FIGURE 6.8** (a) Measured bending stress in sleeper when load from the wheelflat hits the rail above sleeper (left figure), and (b) calculated bending stress in sleeper (right figure) (Source: From Fermér M., and Nielsen, J.C.O., *Vehicle Syst. Dyn.*, 23, pp. 142–157, August 23–27, 1993).

Another source that induces dynamic effects in switches is the sudden change of rail bending stiffness at the crossing. Especially for cast manganese crossings, two rails meet in one large block containing both the wing rails and the nose rail. At the transition from the rail to the manganese crossing, the stiffness changes dramatically because the manganese block is very stiff. The block is also extremely heavy and the large mass of the crossing contributes to make the wheel–rail impact force at the crossing even worse.

## G. MATHEMATICAL MODELLING OF TRACK DYNAMICS

In this section, mathematical models used to simulate the train–track dynamic interaction will be presented. The survey mainly focuses on models describing the track deflection in the vertical plane. The following topics will be discussed:

- beam (rail) on continuous elastic foundation (Section III.G.1)
- vehicle-bridge interaction (moving mass on simply supported beam) (Section III.G.2)
- beam (rail) on discrete supports (Section III.G.3)
- beam (rail) on discrete supports including ballast model (Section III.G.4), and
- beams (rails) on sleepers embedded in continuum, including three-dimensional FEM models (Section III.G.5).

### 1. Beam (Rail) on Continuous Elastic Foundation (Winkler Beam)

The railway track structure consists of the rails, sleepers, railpads, fastenings, ballast, and subgrade. Depending on what one wants to investigate, these components may be modelled in a simpler or a more sophisticated manner. The rail may be modelled either as an ordinary Euler–Bernoulli beam (the conventional beam theory is used) or as a Rayleigh–Timoshenko beam (see [Section IV.A](#)). The Rayleigh–Timoshenko beam theory includes the rotatory inertia of the beam cross section and beam deformations due to the shear force. Also, a longitudinal (axial) force in the rail may be included in these models.

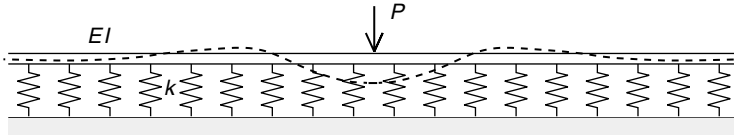
In the most simple track model, a beam (that is a model of the rail) rests on a continuous elastic foundation. The foundation is modelled by an evenly distributed linear spring stiffness. The distributed force supporting the beam is then proportional to the beam deflection. This model was introduced by Winkler in 1867 and is still in use for easy and quick track deflection calculations. The only track parameters needed for this model are the beam bending stiffness  $EI$  ( $\text{Nm}^2$ ) and the foundation stiffness (the bed modulus)  $k$  ( $\text{N/m}^2$ , i.e.,  $\text{N/m}$  per meter of rail). The rail deflection  $w(x)$  ( $x$  is the length coordinate) is then obtained from the differential equation:

$$EI \frac{d^4 w}{dx^4} + kw(x) = q(x) \quad (6.3)$$

where  $q(x)$  is the distributed load on the rail.

This model may be acceptable only for static loading of a track on a soft support, for example, a track with wooden sleepers. No dynamic effects can be analysed using this model as it contains no mass. Especially, the pinned–pinned frequency at approximately 800 to 1000 Hz, when the rail vibrates with nodes at the sleepers, is not contained in this model. On the other hand, using this model it can be seen that the track is lifted in front of the wheel and behind the wheel, see the dashed line in [Figure 6.9](#). This model assumes, however, that there is a tensile force between the rail and the foundation where the track is lifted. In practice, it is only the dead load of the rail and sleepers that will close the “gap” between the track structure and the bed.

To introduce models of railpads, sleepers, and ballast into the track model, the beam may be placed on a bed consisting of several continuous layers instead of the single layer shown in [Figure 6.9](#). The discrete rail supports (railpads and sleepers) are then “smeared out” along the rail



**FIGURE 6.9** Beam (bending stiffness  $EI$ ) on elastic foundation (bed modulus  $k$ ). The beam is loaded with a point force  $P$  from the wheel. The dashed line indicates beam (rail) deflection due to the wheel load  $P$ .

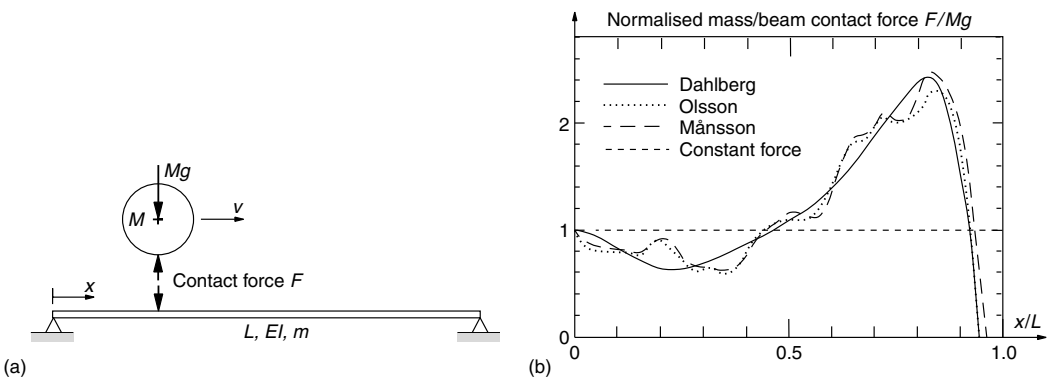
to give several distributed layers. Nor can this model reproduce the pinned–pinned frequency. A continuous model like this one was used by Grassie and Cox.<sup>37</sup> They examined the behaviour of the track support and it was concluded that large sleeper strains were associated with poorly damped sleeper resonances. Comparisons with British Rail experiments were made. Effects of changes in the support (the railpads) were studied, and it was found that a softer railpad isolates the sleeper more effectively and it will significantly reduce the sleeper strains.

**2. Vehicle–Bridge Interaction (Moving Mass on Simply Supported Beam)**

So far in this chapter, only the railway track has been considered a dynamic system. A railway bridge is also a dynamic system that interacts with a train passing the bridge. The dynamic interaction of a wheel (a rigid mass) and a simple bridge model (a simply supported beam) will now be investigated.

The well known “moving mass” problem will be discussed. This problem was reviewed and reinvestigated by Månsson.<sup>38</sup> The dynamic interaction between a rigid mass  $M$  moving with a constant speed  $v$  along a simply supported beam (length  $L$ , mass per unit length  $m$ , bending stiffness  $EI$ ) was examined. The system parameters  $M/mL = 0.5$  (mass ratio) and vehicle speed  $v/v_{ref} = 0.5$  were used. The reference speed  $v_{ref}$  is  $v_{ref} = 2L/T_e$ , where the eigenperiod  $T_e$  of the fundamental (cyclic) eigenfrequency  $f_e$  of the beam is  $T_e = 1/f_e = 2\pi/\omega_e$  and  $\omega_e$  is the fundamental angular eigenfrequency:  $\omega_e = \pi^2(EI/mL^4)^{1/2}$ . Results obtained by Månsson were then compared with results found in the literature.

The wheel–rail contact force (or, if the simply supported beam is a model of a bridge, the wheel–bridge contact force), as calculated by Månsson with the commercial software LS-DYNA, is shown in Figure 6.10 together with results from two other investigations. The fourth curve in



**FIGURE 6.10** Contact force between a moving mass  $M$  and a simply supported beam. The contact force is calculated by three different methods. The force is normalised with respect to the static load (the weight  $Mg$ ) of the moving mass. Here mass ratio is  $M/mL = 0.5$  and vehicle speed  $v/v_{ref} = 0.5$ , where  $v_{ref} = 2L/T_e$  and  $T_e$  is the eigenperiod of the fundamental frequency of the beam, i.e.,  $T_e = 2\pi/[\pi^2(EI/mL^4)^{1/2}]$ .

Figure 6.10, the straight line, gives the constant force  $Mg$ , which is the dead load of the mass. If no dynamic effects were present, then the wheel–bridge force would have been  $Mg$ .

The full-line curve in Figure 6.10 was obtained by Dahlberg.<sup>39</sup> Modal analysis (three modes) was used to calculate the beam deflection and it was assumed that the contact force could be written in the form of a polynomial. Due to the low number of modes included, and due to the restriction that the contact force should be written in the form of a polynomial, no high-frequency variations of the contact force could be modelled by that method.

The dotted curve in Figure 6.10, from Olsson,<sup>40</sup> was obtained by use of the finite element method and modal analysis for the beam. Five modes were included in that analysis, which, again, means that high-frequency variations cannot be modelled very accurately (thus, variations with a frequency higher than the fifth eigenfrequency of the beam cannot be modelled with a high accuracy).

The dashed curve, from Månsson,<sup>38</sup> gives the contact force obtained by using the finite element method and a direct time-stepping algorithm for the moving mass problem (no modal analysis is involved). It is noted that the two finite element methods give very similar results apart from a minor deviation around  $x/L = 0.80$  to  $0.87$ . One explanation of this discrepancy could be that the modal analysis used in the two first cases (Dahlberg and Olsson) will smooth the variations of the curve and thereby make the maximum at  $x/L = 0.82$  slightly too small.

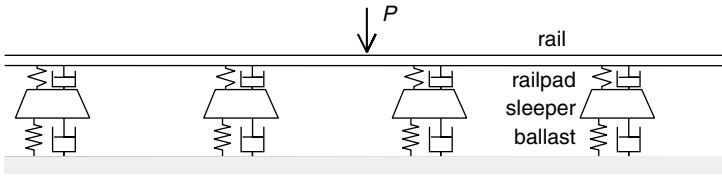
In Figure 6.10, it is seen that the overall behaviour of the wheel–bridge contact force  $F = F(x)$  is that initially, when the wheel enters the bridge, the contact force equals the static load of the wheel, i.e.,  $F(0) = Mg$ , but immediately when the wheel starts to load the bridge, the contact force decreases. This happens when the bridge (the beam) starts to deflect under the horizontally moving mass (the wheel). Because of the decrease of the contact force  $F(x)$ , the dead load  $Mg$  of the mass becomes larger than the reaction force  $F(x)$  and also the mass will start to move downwards (it follows the deflected beam). After a while the beam deflection will change direction and the beam moves upwards, also pushing the mass upwards. This induces a large contact force between the mass and the beam, and  $F(x)$  becomes larger than  $Mg$ . It is seen in the figure that when the wheel reaches the position  $x = 0.82L$ , the wheel–beam contact force has its maximum, which is almost 2.5 times the static load  $Mg$ . This causes the wheel to continue moving upwards, and, as can be seen in the figure, the wheel will lose contact with the beam just before it reaches the support to the right.

This is the overall behaviour of the beam–mass system when the beam vibrates in its fundamental vibration mode. The more high-frequency variations of the contact force in Figure 6.10 are due to beam vibrations at higher modes. It should also be mentioned here that the speed  $v$  used is such that the time it takes for the wheel to pass the beam is half the eigenperiod of the beam at its fundamental frequency. The beam thus makes half a cycle in its fundamental mode during the time it takes for the wheel to pass the bridge. For an ordinary bridge, this normally gives a very high speed.

From this discussion it is concluded that in this case, with this mass ratio and speed, the “dynamic impact factor” is as large as 2.5, implying that due to dynamic effects the maximum mass–beam contact force is 2.5 times the static load. The dynamic impact factor depends on the mass ratio  $M/mL$  and vehicle speed  $v/v_{\text{ref}}$ , so other values than those used here will, of course, give other dynamic impact factors.

### 3. Beam (Rail) on Discrete Supports

To include also the pinned–pinned frequency in the track model, the continuous rail should be supported at discrete points. The supports could then be either discrete spring–damper systems or spring–mass–spring systems, modelling railpads, sleepers and ballast bed. One commonly used method to model this is to place the rail (a beam) on a spring and a damper in parallel. This spring–damper system models the railpad, below which is placed a rigid mass modelling the sleeper.



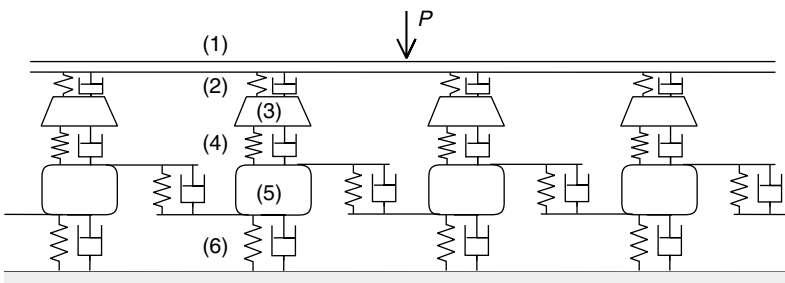
**FIGURE 6.11** Rail on discrete supports. Rail is modelled as a beam (Euler–Bernoulli or Rayleigh–Timoshenko beam theory), the railpads are modelled by spring–damper systems, the sleepers are rigid masses, and the ballast is modelled by spring–damper systems.

The sleeper rests on an elastic foundation, i.e., another spring–damper system, see Figure 6.11. Using this track model, the three resonance frequencies described in Figure 6.3 can be captured, namely the track structure bouncing on the ballast, the rail and sleeper vibration with the railpad as a spring between the two masses, and the pinned–pinned frequency.

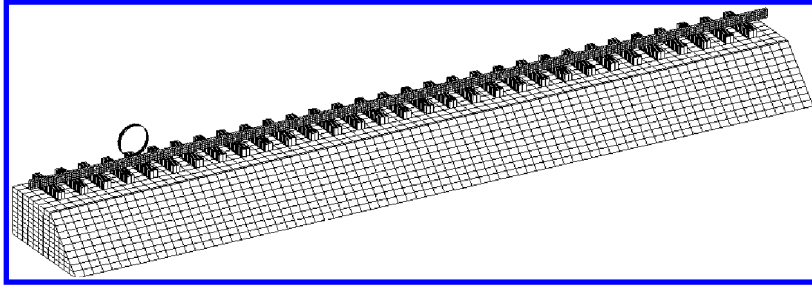
Sometimes, the rigid sleeper mass is replaced by a beam on an elastic foundation. The beam then extends perpendicularly to the rail, and the track model becomes three-dimensional. The sleeper in an elastic embedment is further discussed in Section IV.D.

**4. Discretely Supported Track Including Ballast Mass**

To be able to add a resonance frequency at low frequency (20 to 40 Hz) to the model described above, Oscarsson<sup>41</sup> incorporated more masses into the model, see Figure 6.12. By making the ballast and subgrade mass large (much larger than the sleeper and rail mass) and by adjusting the subgrade stiffness, a resonance at low frequency can be achieved. Then, essentially, the ballast-subgrade masses vibrate on the subgrade stiffness. It is noted in Figure 6.12 that there are connections between the ballast and subgrade masses, implying that a deflection at one point (at one sleeper) will influence the deflection at the adjacent sleepers. This phenomenon (which exists in a real track) cannot be modelled with the simpler models such as the track model in Figure 6.11. Zhai and Cai<sup>42</sup> used a similar model to investigate (among other things) the influence of the ballast density on the wheel–rail contact force at a rail joint and the ballast acceleration.



**FIGURE 6.12** Rail on discrete supports with rigid masses modelling the sleepers. (1) Rail, (2) railpad stiffness and damping, (3) rigid sleepers, (4) ballast stiffness and damping. Rigid masses (5) below the sleepers represent the mass of the ballast and the subgrade, (6) subgrade stiffness and damping. By this model, the four resonance vibration modes (a) embankment vibration, (b) track-on-the-ballast vibration, (c) rail-on-railpad vibration, and (d) pinned–pinned vibration of the rail, may be captured (for (b), (c) and (d), see Figure 6.3).



**FIGURE 6.13** Rail on sleepers and the sleepers are embedded in a continuous ballast and subgrade medium. For the ballast and subgrade, three-dimensional finite elements are required. Symmetry with respect to the track centre-line is used here, implying that only half of the track needs to be modelled.

### 5. Rails on Sleepers Embedded in Continuum. Three-Dimensional Finite Element Models

The most realistic track model, and the model that normally requires the most computer capacity, is the model where rails and sleepers are modelled as beams (or possibly as three-dimensional bodies) with elastic elements modelling the railpads between the rails and the sleepers. The sleepers are embedded in a continuous medium. This requires that the track bed is modelled by three-dimensional finite elements. Figure 6.13 shows such a track model. Using a model like this, and also modelling a larger part of the surroundings, wave propagation from the track to the surroundings can be simulated. In the small model in Figure 6.13, nonreflecting boundary conditions must be used to avoid wave reflections at the boundaries.

## H. MODELLING OF DYNAMIC TRAIN-TRACK INTERACTION AND COMPUTER PROGRAM DEVELOPMENTS

The dynamics of the compound train and track system plays an important role when investigating vehicle and track dynamics, ground-borne vibrations, and noise from the train traffic. Low-frequency (less than 20 Hz) motion of the train is crucial for assessment of safety and riding quality. High-frequency vibrations cause discomfort to passengers and emit noise and vibration to the surroundings. Also, vibrations may lead to track deterioration, such as railhead corrugation growth, damage to track components (railpads, fastenings, sleepers, ballast), track settlement, and so on.

In the early 1900s, Timoshenko published papers on the strength of rails, and later Inglis was active in this field. Early studies of the train-track interaction problem have been reviewed in the book by Fryba<sup>43</sup> (1972, 3rd ed. 1999). This book contains investigations on the vibrations of solids and structures under moving loads; the train (or wheel) then being modelled as a moving force. Knothe and Grassie<sup>44</sup> and Popp et al.<sup>45</sup> have presented state-of-the-art reviews in the field of train-track interaction.

Techniques to study the train-track interaction can be divided into two groups: frequency-domain techniques and time-domain techniques. These two techniques will be discussed in Section III.H.1 and Section III.H.2, respectively, and in Section III.H.3 examples of institutions developing computer programs for calculation of train-track interaction will be given.

### 1. Frequency-Domain Techniques

In the frequency-domain technique, receptances of the track are required (cf. Figure 6.3). If a stationary (not moving) wheel is loading the track, then the track receptance is needed only at

the point where the wheel is situated. The receptance (vertical or lateral, depending on what is studied) may be measured *in situ* on the track or it can be calculated using a track model. If a harmonically varying stationary load excites the track, then the direct receptance provides the track response.

Using the frequency-domain technique, it is possible to investigate the track and wheel response to a “moving irregularity.” Instead of having a wheel moving on an uneven rail, one investigates a stationary wheel. The rail and the (stationary) wheel are then excited at the wheel–rail contact patch by a prescribed displacement. One may think of this excitation as if a strip of irregular thickness were inserted between the wheel and the rail. The strip is then forced to move between the wheel and the rail so that the irregularity of the strip will excite both wheel and rail. The response of the wheel and the track is obtained in the frequency domain. If the strip thickness irregularity is sinusoidal, the response is found from a receptance function. A nonsinusoidal irregularity (as from a wheel-flat) must first be transformed into the frequency domain by the Fourier transform. The track and wheel receptances and the wheel–rail contact stiffness are combined to form the appropriate transfer function. Together with the Fourier transform of the irregularity, the Fourier transform of the response is obtained, and the inverse transform provides the time-domain response. Several authors have used this technique to investigate the development of short wavelength corrugation on the railhead, see [Section III.D](#).

If continuously supported rail is excited by a harmonically varying moving load, then the track response can be determined in a coordinate system following the load. The response is then assumed to be stationary. This topic has been thoroughly investigated in the book mentioned above, Fryba.<sup>43</sup> One method to treat a discretely supported rail is to develop the support reactions into Fourier series (making the support continuous but nonuniform) and then the moving load problem is solved with respect to a coordinate system following the load.

In the frequency-domain technique, only fully linear systems can be treated. The track responses are also assumed to be stationary, implying that singular events along the track, such as a rail joint, a sleeper hanging in the rail (not supported by the ballast), varying track stiffness, and so on, cannot be treated.

## 2. Time-Domain Techniques

When train–track dynamics are investigated in the time domain, deflections of the track and displacements of the vehicle are calculated by numerical time integration as the vehicle moves along the track. The vertical motion of the wheelset should then coincide with the vertical deflection of the rail, while taking the wheel–rail contact deformation into account. The wheel–rail contact force is unknown and has to be determined in the calculations.

The track can be modelled by finite elements, and in many cases a modal analysis of the track is performed. This requires that the track is modelled over a finite length. The track is then described through its modal parameters, and the physical deflections of the track are determined by modal superposition (this requires a linear track model). Often the vehicle is modelled by use of rigid masses, springs (linear or nonlinear) and viscous dampers. If a more detailed response of the vehicle is of interest, then a better vehicle model should be used and it could be convenient to also use modal analysis for the vehicle deformations (the vehicle is no longer composed of rigid bodies). Modal analysis of a wheelset makes it possible to include elastic deformations of the wheelset without a large increase of the number of degrees of freedom of the compound train–track system, Ripke<sup>46</sup> and Andersson.<sup>47</sup>

The modal analysis technique requires linear models. A finite element track model may also comprise nonlinear track elements. In such a model, the material properties may be selected to display the physical behaviour of the nonlinear track elements. Normally, nonlinearities can be found in railpads and in the ballast-subgrade material, see Oscarsson.<sup>41</sup> Nonlinearities in the track



have been treated as extra loads giving a force–displacement relationship for the track comparable with the nonlinear characteristics of the real track, Oscarsson.<sup>41</sup>

### 3. Computer Program Developments

A benchmark test of different train–track interaction models has been carried out, Grassie.<sup>48</sup> Both frequency-domain and time-domain models contributed to the test. In the test, eight sets of calculations were contributed, of which one was essentially from a low-frequency model, two contributions were from finite element models with calculations undertaken in the time domain, two were from finite element models with calculations undertaken in the frequency domain, and the remaining three were essentially analytical models with calculations undertaken in the frequency domain. In the time domain models and in two of the frequency domain models it was assumed that the rail was supported at discrete sleepers, while in the other frequency domain models a continuous support was assumed. The organisations or individuals who submitted results were:

- Dr S Grassie (private, program TRACK), U.K.
- Technical University, Berlin (two models), Germany
- Pandrol International Ltd, U.K.
- TNO (Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek) in Delft (two models, program TWINS), The Netherlands
- Chalmers University of Technology, Gothenburg (program DIFF), Sweden, and
- University of Transportation and Communication, Zilina, Slovakia.

From the results of the benchmark test it was concluded that it was at least possible to obtain substantially similar results from both time domain and frequency domain models of vehicle–track interaction in the majority of circumstances considered in the test. Both types of models have attractions and disadvantages. It was concluded that the different types of model can be used with confidence in those areas where their attractions can be exploited in full. It can be mentioned here that the Swedish program DIFF has been further developed to comprise general three-dimensional motion of a train traversing the track, Andersson and Abrahamsson.<sup>49</sup>

Among other organisations developing train–track interaction models are:

- Spoornet, South Africa, see for example Fröhling,<sup>50</sup>
- Institute of Sound and Vibration Research, University of Southampton, U.K., see for example Wu and Thompson,<sup>51</sup>
- Technical University in Munich, Germany, see for example Dinkel,<sup>52</sup>
- Technical University in Milano, Italy, Diana et al.,<sup>53</sup>
- In the U.S.A., research is performed at the University of Massachusetts, Amherst, and at the Association of American Railroads, Transportation Technology at Pueblo, CO, Li and Selig<sup>54</sup>
- Train and Track Research Institute at Southwest Jiaotong University, Chengdu, P R of China, see for example Zhai W and Cai Z<sup>42</sup>
- Central Queensland University in Australia, Sun and Dhanasekar<sup>55</sup>
- Railway Technical Research Institute in Japan, Sato et al.<sup>56</sup>
- Linköping University, Sweden (author of this chapter)

## IV. DYNAMIC PROPERTIES OF TRACK COMPONENTS

The most important parts of a track structure, from a track dynamics point of view, are the rails, sleepers or ties, railpads, fastenings, ballast, and the subgrade. Depending on what is to be studied,

these components may be modelled in a simpler or a more advanced manner. In this section, some track components will be investigated. The role and dynamic behaviour of the four components rail, railpad, sleeper, and ballast-subgrade will be examined in some detail.

## A. THE RAIL

Here, the mathematical modelling of bending vibration of a free rail (and only the rail) will be treated. Therefore, there is no support along the rail; the rail is supported only at the boundaries. For bending vibration of a rail in the vertical plane, frequencies up to 3000 Hz are discussed. Below this frequency, beam theories can be used and the rail cross section can be considered not to change too much. At higher frequencies, the web and the flanges (the foot) of the rail start to vibrate. This is treated in this handbook in the chapter on noise. For lateral vibration, cross-section deformation becomes significant above 1000 Hz.

## B. MATHEMATICAL MODELLING OF RAILS

The rail may be modelled either as an ordinary Euler–Bernoulli (E–B) beam or as a Rayleigh–Timoshenko (R–T) beam. In the E–B beam theory, only the bending of the rail is taken into account, and in case of vibrations, only the mass inertia in translation of the beam is included. The differential equation describing the beam deflection  $w(x, t)$  reads

$$EI \frac{\partial^4 w(x, t)}{\partial x^4} + \rho A \frac{\partial^2 w(x, t)}{\partial t^2} = q(x, t) \quad (6.4)$$

where  $EI$  is the bending stiffness of the beam,  $\rho$  is density,  $A$  is cross-sectional area, giving  $\rho A = m$ , which is the mass of the beam per metre (kg/m), and  $q(x, t)$  is the load on the beam ( $t$  is time). The beam is supported at the ends only, i.e., at  $x = 0$  and  $x = L$  (beam length  $L$  is assumed). Damping of the beam is not included in this model. For stationary vibrations of the undamped beam, the solution to (the homogeneous part of) this equation may be written in the form

$$w_{\text{hom}}(x, t) = X(x)T(t) = X(x)\sin \omega t \quad (6.5)$$

where  $X(x)$  gives the form of the beam deflection when it vibrates (the vibration mode) and  $\omega$  is the vibration angular frequency.

The R–T beam theory includes rotatory inertia and shear deformation of the beam. In this case, two differential equations are needed to describe the vibrations. The deflection  $w(x, t)$  and the shear deformation  $\psi(x, t)$  are unknown functions. The differential equation for the deflection  $w(x, t)$  becomes

$$\begin{aligned} EI \frac{\partial^4 w(x, t)}{\partial x^4} + \rho A \frac{\partial^2 w(x, t)}{\partial t^2} - \rho I \left( 1 + \frac{E}{kG} \right) \frac{\partial^4 w(x, t)}{\partial x^2 \partial t^2} + \frac{\rho^2 I}{kG} \frac{\partial^4 w(x, t)}{\partial t^4} \\ = q(x, t) + \frac{\rho I}{kGA} \frac{\partial^2 q}{\partial t^2} - \frac{EI}{kGA} \frac{\partial^2 q}{\partial x^2} \end{aligned} \quad (6.6)$$

where  $EI$ ,  $\rho$ ,  $A$ , and  $q(x, t)$  are the same as in the E–B case,  $G$  is the shear modulus, and  $k$  is the “shear factor.” Also, this equation (like the E–B one) describes the behaviour of the beam between the supports at the beam ends. A similar equation is obtained for the shear deformation  $\psi(x, t)$ .

If the shear deformation of the beam is suppressed, i.e., if one gives  $k$  a very large value, then the two last terms on both sides tend to zero. Further, if the mass inertia in rotation of the beam cross section is eliminated (noting that  $\rho I = \rho r^2 A = mr^2$ , and let  $r$  tend to zero), then the third term also tends to zero and the E–B differential equation is obtained.

In a state-of-the-art paper by Knothe and Grassie,<sup>44</sup> the authors state that shear deformation of the rail can be neglected only for frequencies below 500 Hz. Dahlberg<sup>57</sup> showed that at this

**TABLE 6.1**  
Eigenfrequencies of a UIC60 Rail of Length  $l = 4.2$  m

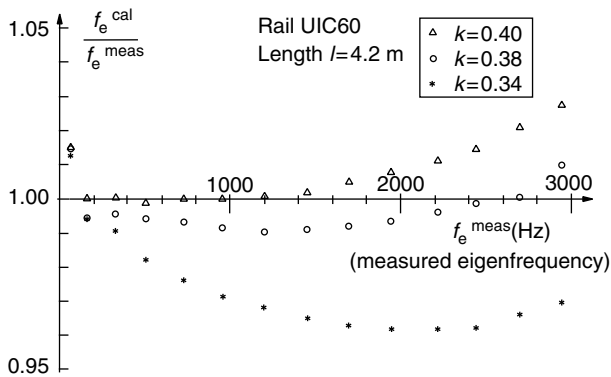
Measured	Calculated			E–B
	$k = 0.34$	$k = 0.38$	$k = 0.40$	
64	64.8	64.9	64.9	65.8
174	173	173	174	181
327	324	326	327	356
515	506	512	514	588
728	711	723	728	878
959	932	951	959	1227
1201	1163	1190	1202	1633
1450	1400	1437	1453	2098
1702	1640	1688	1710	2620
1955	1881	1942	1969	3201
2206	2123	2197	2230	3840
2456	2365	2452	2492	4537
2697	2606	2707	2753	—
2933	2846	2961	3013	—

Measured eigenfrequencies (in Hz) and calculated for some different values of the Timoshenko shear factor  $k$  (in factor  $kG$ ). Last column provides eigenfrequencies obtained by the Euler–Bernoulli (E–B) beam theory.

frequency (500 Hz), and for a UIC60 rail, the Euler–Bernoulli beam theory provides a vibration frequency that is 10 to 15% too high, see Table 6.1 above.

An investigation of the significance of the rotatory inertia and the shear deformation of the rail was performed by Dahlberg.<sup>57</sup> Eigenfrequencies of some pieces of rails were measured and also calculated by use of Euler–Bernoulli and Rayleigh–Timoshenko beam theory. Different values of Timoshenko’s shear factor  $k$  (in  $kG$ ) were tried and results were compared to the measured eigenfrequencies and to the frequencies calculated by the Euler–Bernoulli beam theory. Results are presented in Table 6.1 and in Figure 6.14.

From this investigation, it can be concluded that for a UIC60 rail (a rail with mass 60 kg per meter) the Euler–Bernoulli beam theory yields approximately 15% too high frequency at 500 Hz



**FIGURE 6.14** Influence of Timoshenko’s shear factor  $k$  (in  $kG$ ) for rail UIC60. Ratio of calculated eigenfrequency  $f_e^{cal}$  and measured eigenfrequency  $f_e^{meas}$  for rail of length  $l = 4.2$  m.

and almost 30% too high frequency at 1000 Hz. For a 50 kg/m rail (the Swedish BV50 rail profile) the corresponding numbers are 12 and 30%, respectively.

In [Figure 6.14](#), and also in [Table 6.1](#), it can be seen that in order to obtain a good agreement between measured and calculated eigenfrequencies, the Timoshenko shear factor  $k$  should be different in different frequency intervals. The factor also depends on which rail is used. For the UIC60 rail and for frequencies below 1500 Hz, a reasonable value of  $k$  should be  $k = 0.40$ . For frequencies up to 3000 Hz, say, the  $k$  value could be somewhat lower ( $k = 0.38$ ). For the BV50 rail, the value  $k = 0.34$  gives a good fit over a large frequency range, see [Dahlberg](#).<sup>57</sup> (It should be mentioned here that the widely used method to calculate the factor  $k$  proposed in [Cowper](#)<sup>58</sup> refers to static loading, and hence to static deflection, of the beam.)

Similar measurements and calculations of eigenfrequencies of concrete sleepers have been reported in [Dahlberg and Nielsen](#),<sup>59</sup> see [Section IV.D](#) on sleeper vibration.

### C. RAILPADS AND FASTENINGS

In a track, the rails are fastened onto the sleepers. Usually, railpads are inserted between the sleepers and the rails, see [Figure 6.2](#). The railpads provide electrical insulation of the rails and they protect the sleepers from wear. The railpads also affect the dynamic behaviour of the track. [Grassie et al.](#)<sup>60</sup> compared different track models with measurements and showed that it is important to include the railpads to obtain an accurate track model. [Grassie and Cox](#)<sup>37</sup> pointed out the importance of railpads when calculating the strains in the sleepers. The influence of soft and stiff railpads on the wheel–rail contact force and the track dynamics has been measured in full-scale experiments with a moving train, as reported by [Fermér and Nielsen](#).<sup>35</sup> In those measurements, the train was equipped with an instrumented wheelset (with and without wheel flats) and the track was also equipped with measurement devices (strain gauges and accelerometers). Soft railpads were found to result in a lower sleeper acceleration and a higher railhead acceleration than the stiff railpads.

Laboratory measurements of stiffness and damping of studded rubber railpads have been performed by [Thompson et al.](#)<sup>61</sup> Also, [Fenander](#)<sup>62</sup> reported these measurements and compared them with field measurements on railpads. The stiffness of the railpads was found to be weakly dependent on frequency, but to increase strongly with preload. The damping loss factor of the railpads was found to be nearly independent of the preload and to increase only slightly with frequency.

The most commonly used physical model of a railpad is the spring–damper system, see [Figure 6.11](#). The spring is usually assumed to be linear, and the damping is assumed to be proportional to the deformation rate of the railpad. It should be pointed out, however, that this viscous damping model does not always agree well with measured data (what may give a good fit at one train speed need not give a good fit at another speed). In practice, the damping is more or less independent of frequency, implying that a constant loss factor model agrees better.

As the stiffness of railpads has been found to increase strongly with preload (especially for studded railpads), a more accurate track model should display this behaviour, [Wu and Thompson](#).<sup>63</sup>

[Oscarsson](#)<sup>41</sup> expanded a linear dynamic train–track interaction model to encompass nonlinear stiffnesses in the railpads and in the ballast. He undertook an investigation on the influence of the nonlinear railpad stiffness on rail corrugation growth (the growth of short wavelength periodic irregularities), and he found that a weak nonlinearity of the railpad stiffness does not influence the corrugation growth very much, whereas a strong nonlinearity may lead to an increased wear of the rail. Also, in [Dahlberg](#),<sup>64</sup> a comparison of linear and nonlinear track models containing linear and nonlinear railpad stiffnesses was performed. It was shown that the axle load from the train distributes differently into the track in the linear and nonlinear track models.

A more advanced railpad model was set up by [Fenander](#).<sup>62</sup> A fractional calculus model of the dynamic behaviour of railpads was proposed. Earlier, it has been shown that molecular theories for viscoelastic polymers can be used to derive the fractional derivative model, and that fractional derivative models accurately describe many viscoelastic materials by use of only few parameters.

This was exploited by Fenander, who used a four-parameter model to fit the measured stiffness and loss factor of the railpad. The fractional derivative model is linear. Therefore, for each preload level a new set of parameters had to be determined.

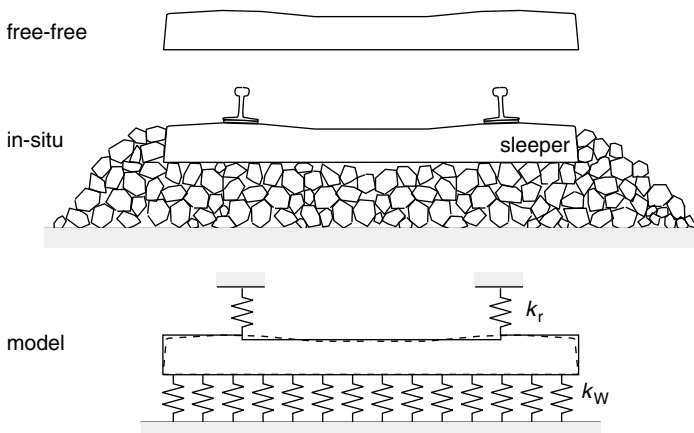
The role of the fastenings is normally neglected when investigating track dynamics. The fastenings should keep the rails in place (attached to the sleepers) and prevent the rail from rolling over when lateral forces are applied to it. The stiffness of the fastening is normally much less than that of the railpad. Therefore, when a wheelset loads the track, only the railpad stiffness is important. For an unloaded rail, however, the fastening will induce a certain preload (static load) on the railpad, and, knowing that the railpad stiffness is nonlinear, this may influence the dynamics of the unloaded track. The dynamic behaviour of rail fasteners at high frequencies has been investigated by Thompson and Verheij.<sup>65</sup>

## D. THE SLEEPERS

Railway track models sometimes contain beams (sleepers) on an elastic foundation. In this section the dynamic behaviour of free–free and *in situ* concrete sleepers will be treated. Two problems are faced: how to model the sleeper vibrations and how the elastic foundation supporting the *in situ* sleeper will influence the dynamic behaviour of it.

### 1. Sleeper Vibrations

Dynamic investigations of concrete sleepers are often made on sleepers with free–free boundary conditions, i.e., the sleeper is suspended in very soft springs, not influencing the vibration of the sleeper, so that the sleeper can be considered totally free. The free–free boundary conditions thus means that the two sleeper ends are free, see Figure 6.15. If Euler–Bernoulli beam theory is used to describe the sleeper vibration, good agreement between measured and calculated eigenfrequencies and eigenmodes is normally obtained only for the lowest two or three eigenfrequencies. To obtain a better agreement at higher frequencies, the more advanced Rayleigh–Timoshenko beam theory is required for the sleeper vibration calculations, Grassie.<sup>66</sup>



**FIGURE 6.15** Free–free sleeper (top figure), *in situ* sleeper (sleeper in track), and model of the *in situ* sleeper. In the sleeper model, three beam elements were used. The sleeper is connected (via the rail) to the surrounding structure and is supported by distributed springs along the sleeper.

## 2. Elastic Foundation

The elastic bed of a sleeper in the track will influence the sleeper vibration. Therefore, eigenfrequencies of a beam on an elastic foundation will also be discussed. The eigenfrequencies (measured and calculated) of a free–free sleeper are compared with the frequencies of the same sleeper when it is placed in the track on the elastic foundation.

## 3. Measurements and Calculations

Measurements performed on free–free precast concrete railway sleepers give eigenfrequencies of the sleepers. Measured eigenfrequencies will be compared here with calculated frequencies obtained with the Euler–Bernoulli beam theory and with the Rayleigh–Timoshenko theory. The investigation summarised here was first reported in Dahlberg and Nielsen.<sup>59</sup>

Figure 6.15 shows the free–free sleeper, the sleeper in track, and the modelling of the sleeper. The physical, nonuniform sleeper used in the measurements was modelled as three uniform beams. The *in situ* sleeper rests on an elastic foundation (the ballast bed) and is connected to the surrounding structure via the rails. In the sleeper model, the elastic foundation is modelled by distributed springs along the sleeper, with spring stiffness  $k_W$  (N/m<sup>2</sup>, i.e., spring stiffness N/m per metre of the sleeper). The connection of the sleeper to the rails and to other parts of the track (via the rail) is modelled by discrete springs, stiffness  $k_r$  (N/m), see Figure 6.15. For the central part of the sleeper, length 1.5 m, and for the two end parts, length 0.5 m each, beam elements were used with the following properties: bending stiffness  $EI = 4.60$  and  $6.41$  MNm<sup>2</sup>, mass  $m = 91.2$  and  $114$  kg/m (giving a total sleeper mass of 251 kg), shear factor  $kGA = 502$  and  $628$  MN, and radius of gyration  $r = 0.0568$  and  $0.0600$  m, respectively ( $kGA$  and  $r$  are used in the R–T beam theory). Two different track beds were used, namely a soft bed, called *In situ* 1, with bed stiffness  $k_W = 13$  MN/m<sup>2</sup> and “rail spring” stiffness  $k_r = 17$  MN/m. For the stiffer bed, *In situ* 2, bed stiffness  $k_W = 26$  MN/m<sup>2</sup> and “rail spring” stiffness  $k_r = 34$  MN/m were used. Measured and calculated results are presented in Table 6.2.

The eigenfrequencies of a beam on an elastic foundation (i.e., the *in situ* sleeper) will differ from those obtained for the same beam if the beam is totally free. As can be seen in Table 6.2, the foundation induces two new eigenfrequencies (close to 80 Hz for the soft track and a little more than 110 Hz for the stiff track). These frequencies are marked 0 for the free–free sleeper and are two rigid-body motions of the free–free sleeper (rigid-body motion in translation and rotation).

**TABLE 6.2**  
Eigenfrequencies of a Concrete Sleeper

Measured free–free (Hz)	Calculated			
	free–free E–B	free–free R–T	<i>In situ</i> 1 (soft)	<i>In situ</i> 2 (stiff)
0	0	0	79 (rotation)	111 (rotation)
0	0	0	82 (translation)	115 (translation)
113	121	118	131	143
335	347	327	335	344
639	696	624	631	637
995	1167	986	989	993
1390	1752	1391	1393	1394

Measured eigenfrequencies for free–free sleeper (in Hz) and calculated (also for free–free sleeper), using Euler–Bernoulli (E–B) and Rayleigh–Timoshenko (R–T) beam theory. Columns four and five (*in situ* sleeper) give eigenfrequencies of the same sleeper when it is placed in track, see Figure 6.15. Two different bed stiffnesses and rail stiffnesses were used (soft and stiff). The *in situ* frequencies have been calculated by use of the R–T theory.

The rigid-body motions of the free–free sleeper become two almost rigid-body eigenmodes (translation and rotation) of the *in situ* sleeper. It is also seen in the table that the lower eigenfrequencies in bending are influenced by the foundation stiffness and the “rail spring” stiffness, whereas higher eigenfrequencies are almost unaffected by these stiffnesses.

In Table 6.2, the two rigid-body motions of the free–free sleeper (frequency 0) become almost rigid-body motions of the same sleeper placed in the ballast bed (the stiffer bed makes these frequencies higher, of course). It is seen that the lowest bending mode, measured frequency 113 Hz, is obtained (fairly closely) both from the E–B theory and the R–T theory. This frequency increases some 10 to 15 Hz when the sleeper is placed in the track. The second bending mode, measured frequency 335 Hz, is also obtained (fairly closely) by the two theories, but it is seen that the E–B theory gives a slightly too high value here. This frequency is affected only a small amount by the track bed. At the third resonance frequency, the E–B theory gives almost 10% too high a value whereas the R–T theory is close to the measured one. This frequency is not greatly affected by the track bed. For the fourth and fifth frequencies, the E–B theory gives too large values whereas the R–T theory gives values very close to the measured ones. These frequencies are (more or less) unaffected by the support.

Thus, depending on which frequency interval is of interest, the concrete sleeper can be modelled as either a rigid mass (at frequencies below 100 Hz) or as a flexible beam. For frequencies up to 300 or 400 Hz, the Euler–Bernoulli beam theory may suffice. At higher frequencies, the Rayleigh–Timoshenko beam theory should be used for an accurate description of the sleeper vibration.

## E. BALLAST, SUBBALLAST AND SUBGRADE

Placed under and around the sleepers, the ballast layer has several important functions:

- It limits sleeper movement by resisting vertical, transverse, and longitudinal forces from the trains
- It distributes the load from the sleepers to protect the subgrade from high stresses, thereby limiting permanent settlement of the track
- It provides necessary resilience to absorb shock from dynamic loading
- It facilitates maintenance surfacing and lining operations
- It provides immediate water drainage from the track structure
- It helps alleviate frost problems and
- It retards the growth of vegetation and resists the effects of fouling from surface-deposited materials.

At present, the state-of-the-art of track design concerning the ballast and the subgrade is mostly empirical. The factors that control the performance of the ballast are poorly understood. To assess the reasons why a particular section of track requires maintenance, it is necessary to know:

1. The characteristics of the ballast and subgrade
2. The maintenance history
3. The environmental history, and
4. The traffic history

Usually, only the last three items can be estimated from records. Information on the characteristics of the ballast and subgrade of an existing track is, in most cases, nonexistent. To gain information on the present conditions of a site, field examination is all that is possible.

As mentioned, the factors that control the performance of the ballast are poorly understood. In Knothe,<sup>67</sup> the long-term behaviour of the railroad track, including the ballast behaviour and the damage mechanisms underlying the ballast settlement, is discussed. Knothe states that any generally accepted damage and settlement equations do not exist, and hardly any material equations for the ballast itself. Only different suggestions to describe the ballast settlement from a phenomenological point of view are available; the settlement then being a function of number of loading cycles and/or magnitude of the loading.

## 1. Track Settlement

Railway track will settle (change its position) as a result of permanent deformation in the ballast and underlying soil. After having been used some time, the track will not be so straight and so well levelled as it was when it was new. The settlement is caused by the repeated traffic loading and the severity of the settlement depends on the quality and the behaviour of the ballast, the subballast, and the subgrade.

Track settlement occurs in two major phases:

- Directly after tamping, when the track position has been adjusted to a straight level, the settlement is relatively fast until the gaps between the ballast particles have been reduced and the ballast is consolidated.
- The second phase of settlement is slower and there is a more or less linear relationship between settlement and time (or load).

The second phase of settlement is caused by several basic mechanisms of ballast and subgrade behaviour:

- Continued (after the first phase) volume reduction, i.e., densification of the ballast and subgrade, caused by particle rearrangement produced by repeated train loading.
- Subballast and/or subgrade penetration into ballast voids. This causes the ballast to sink into the subballast and subgrade and the track level will change accordingly.
- Volume reduction caused by particle breakdown from train loading or environmental factors; i.e., ballast particles may fracture (divide into two or more pieces) due to the loading.
- Volume reduction caused by abrasive wear. A particle may diminish in volume due to abrasive wear at points in contact with other particles, i.e., originally cornered stones become rounded; then occupying less space.
- Inelastic recovery on unloading. Due to micro-slip between ballast particles at loading, all deformations will not be fully recovered upon unloading the track. In this case, the permanent deformation is a function of both stress history and stress state.
- Movement of ballast and subgrade particles away from under the sleepers. This causes the sleepers to sink into the ballast and subgrade.
- Lateral, and possibly also longitudinal (in the rail direction), movement of sleepers causing the ballast beneath the sleepers to be “pushed away,” and the sleepers will sink deeper into the ballast.

Here, the first four items concern densification of ballast and subgrade, whereas the three last-mentioned items concern inelastic behaviour of the ballast and subgrade materials.

Concerning the volume reduction or densification caused by particle rearrangement produced by repeated train loading, it could be mentioned that the train load may also have an opposite effect. Due to the elastic foundation, the train load will lift the track (rails and sleepers) in front of and



behind the loading point, thus reducing or eliminating the preload (the dead load) caused by the rails and sleepers on the ballast. At the same time, due to the dynamic high-frequency train-track interaction forces, waves will propagate from the wheel-rail contact patches, either through the ballast and subgrade or through the track structure, to the region with the unloaded ballast. These waves will normally propagate faster than the train, giving vibrations in the unloaded ballast. This, in turn, may cause a rearrangement of the ballast particles so that the density decreases. As a result, this may cause a lift, at least temporarily, of the track.

## 2. Research on Ballast

A review of research on railroad ballast used as track substructure has been presented by Peplow et al.<sup>68</sup> The ballast materials and their interactions are complex from a physical point of view. Hence, appropriate constitutive laws for the (short-term) response of the materials have been developed including resilient modulus and variable modulus approaches. The laws are verified with respect to laboratory tests.

A historical method for assessing track performance is the use of track modulus. Its value for static and dynamic loading for track structure-ballast interaction is reviewed and discussed in Peplow et al.<sup>68</sup> For static loading, a comprehensive review of one approach is given. This approach uses multi-layer linear elastic static theory to represent the ballast and the subgrade layers. A number of finite element models have been developed and compared, and stresses incurred within the ballast and subgrade for various configurations are discussed.

For modelling the dynamic interaction between the track structure and the ballast, a simple beam on elastic foundation model (the Winkler foundation) is used in many analyses. In a number of studies a discrete support model with a finite number of parameters describing the rail, railpads, sleepers, and ballast mass and stiffness is used, see for example, Oscarsson<sup>41</sup> and Zhai et al.<sup>69</sup> In the final part of the survey by Peplow et al., modelling of the dynamic interaction between the train and the track structures is reviewed. The review also presents some mathematical and numerical methods dealing with the static and dynamic interaction of the train-track system and the substructure.

In Jacobsson,<sup>70</sup> a literature review of research on railroad ballast is presented, with emphasis on constitutive and mathematical modelling of the behaviour of ballast in tests. Different loading conditions in the tests are compression-tension, triaxial, and shear. Different mathematical descriptions of the constitutive behaviour of the material are given. In particular, descriptions of the resilient material properties and the evolution of permanent deformations, as a function of stress state, stress history, and the degree of compaction, are summarised.

The effect of different vehicles on track deterioration (and consequent maintenance costs) has been examined by Iwnicki et al.<sup>71</sup> A number of track settlement models were investigated. It was noted, for example, that the ORE (Office de Recherches et d'Essais de l'Union Internationale des Chemins de Fer) deterioration model contains no track parameters at all but only loading parameters such as traffic volume, dynamic axle load, and speed. Such a model, containing no track parameters, would imply that two different tracks, one stiff and one soft, would undergo the same deterioration if they were subjected to the same loading. This can be questioned, of course, since in such a case the quality of the ballast and subgrade material would have no influence on the track deterioration.

In Suiker,<sup>72</sup> advanced models were developed in order to provide detailed insight into short-term and long-term mechanical processes in a railway track. One of the purposes of the work was to derive enhanced continuum models from the discrete microstructure of a granular material (the ballast). The long-term mechanical process concerns the evolution of track deterioration as a result of a large number of train axle passages. A model that simulates the plastic deformation behaviour of the track bed during each loading cycle (wheel passage) would be unattractive.

Instead, a model is employed that captures only the envelope of the maximum plastic deformations generated during the cyclic loading of the track.

In a research programme in Germany, aimed at a better understanding of the dynamic interaction of vehicle and track and the long-term behaviour of the components of the entire system, settlement and deterioration of the ballast and the subgrade were examined. Nonlinear behaviour of the ballast was investigated experimentally in laboratories and simulated by new material laws, and the coupling of the track model and the subgrade was defined using various models, see Popp and Schiehlen.<sup>73</sup>

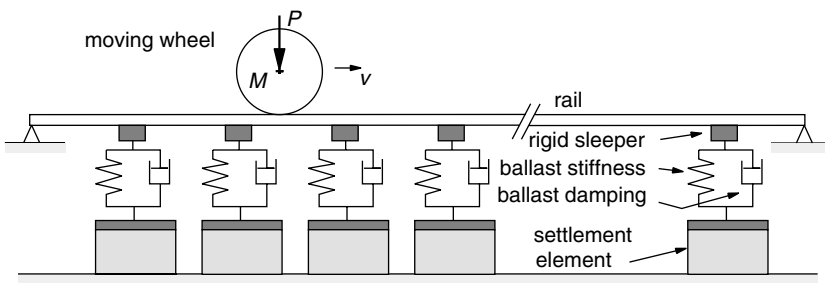
### 3. Modelling Track Settlement

The mathematical modelling of railway track settlement caused by inelastic behaviour of ballast and subground will now be discussed. In many track models the track stiffness and the track mass are modelled by discrete elements. The mass of the track is modelled with rigid bodies and the track stiffness is modelled with springs, see Figure 6.11 and Figure 6.12. In the model that will be discussed here, it is also assumed that the track settlement can be “discretized”. Thus, the settlement of the track is collected in a “settlement element” in the track model.

For the purpose of modelling track settlement, a track model with a “settlement element” has been created, Dahlberg.<sup>74</sup> The model contains one rail (symmetry with respect to the centre line of the track is assumed), sleepers, ballast stiffness (spring elements), ballast damping, and an element beneath the ballast stiffness to take care of the permanent deformations in the ballast. The permanent deformation element thus models the track settlement (like a spring models the track stiffness), see Figure 6.16. The rail is modelled by finite elements, the (half)sleepers are rigid masses, the ballast springs have stiffnesses that may be linear or nonlinear, and the ballast damping is linear (modelled by viscous dampers).

The element taking care of the track settlement is modelled as a solid block (three-dimensional finite elements used) made of a linear elastic-ideally plastic material. This means that if the loading is not too high (i.e., below a threshold value, here the yield limit of the material), no settlement will occur. If, on the other hand, the threshold value (the yield limit) is exceeded, then plastic deformation will occur in the solid block. (It should be noted that this simple model was created to explore if it was possible to use a computer software to simulate the dynamic train–track interaction problem and to consider the track settlement in the same model. In a practical situation, a better model for the track settlement should be used, of course.)

The loading of the track comes from the moving wheel, which is loaded by a constant force: the dead load of the car body. The wheel mass and half of the axle mass are included, implying that the inertia force from the unsprung mass, i.e., the wheel and the axle, is taken into account.



**FIGURE 6.16** Model of railroad track containing rail, sleepers, ballast, and an element to model the track settlement. The model is loaded by the dead weight of the track structure (rail and sleepers) and by a moving mass (a wheel, mass  $M$ ) subjected to a constant force  $P$  from the car body. Also, inertia forces of wheel and track are taken into account.

It was found that a large value of the yield limit of the settlement element material (all elements were elastic) would imply that the track returned to its original position after the wheel had passed (as expected, of course). A low value of the yield limit in some of the settlement elements resulted in a permanent deformation of the track after the wheel had passed. The level of the rail thus changed, giving a dip in the track at the place where the elastic–plastic settlement elements were placed.

## V. SUMMARY

Dynamic properties of railway tracks have been surveyed. When a train moves on the track, the train, which is one dynamic system, interacts with the track, which is another dynamic system. Dynamic effects in the compound train–track system become more pronounced when train speed increases and when the axle load gets higher.

After the function of the track was introduced and dynamic properties of the track were discussed as a whole, several sources of dynamic excitation of the train and the track systems were discussed. Short wavelength irregularities on wheel and railhead induce high-frequency vibrations into the two systems (train and track), and so do impact loadings. Long wavelength irregularities, both in geometry and track stiffness, induce low-frequency oscillations and vibrations. The oscillations and vibrations then induce decreased ride comfort, increased track deterioration and noise emission.

Several mathematical models of the track structure were presented. These extended from the simple beam on Winkler foundation model to the more sophisticated three-dimensional finite element models, including rails, railpads, sleepers, ballast, and so on.

Modelling of the dynamic train–track interaction can be performed either in the frequency domain or in the time domain. Organisations and institutions developing computer programs dealing with the interaction problem were listed.

In the last part of this survey, dynamic properties of some track components (rail, railpad, sleeper, and ballast-subgrade), and mathematical modelling of them, were examined.

Finally, in addition to the textbooks and other books mentioned here, recent textbooks on railway engineering have been written by, for example, Esveld<sup>75</sup> and Profillidis.<sup>76</sup>

## ACKNOWLEDGMENTS

Support by the following persons is gratefully acknowledged: calculations were performed by F. Månsson and A. Lundqvist at Linköping University. E. Berggren at Banverket (the Swedish National Rail Administration) provided [Figure 6.6](#) (measurements performed by Banverket). Dr J. Nielsen at Chalmers University in Gothenburg has read the manuscript and commented upon it. Discussions with Prof David Thompson at the Institute of Sound and Vibration Research in Southampton, U.K., especially on the antiresonance effect in [Figure 6.3](#), is gratefully acknowledged. This work was partly performed within the European FP5 programme GROWTH, project SUPERTRACK (Sustained Performance of Railway Tracks), contract No G1RD-CT-2002-00777.

## REFERENCES

1. Li, D. and Selig, E. T., Evaluation of railway subgrade problems, *Transport. Res. Rec.*, 1489, 17–25, 1995.
2. Chrismer, S. and Read, D. M., Examining ballast and subgrade conditions, *Rail. Track and Struct.*, 90(6), 39–42, 1994.
3. Oscarsson, J., Dynamic train–track–ballast interaction: linear and state-dependent track models, *Thesis for the Degree of Licentiate of Engineering*, Department of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, ISSN 0283-8672, 1999.

4. Jonsson, J., On ground and structural vibrations related to railway traffic, PhD thesis, Department of Structural Engineering, Chalmers University of Technology, Gothenburg, Sweden, 172 pp., 2000.
5. Krylov, V. V., Effects of the embankment topography and track curvature on ground vibration boom from high-speed trains, In *Proceedings of the Symposium EURO-DYN2002*, Munich, Germany, Grundmann and Schueller, Eds., Swetz & Zeitlinger, Lisse, pp. 473–478, ISBN 90 5809 510X, 2002.
6. Sato, Y., Matsumoto, A., and Knothe, K., Review on rail corrugation studies, *Wear*, 253(1–2), 130–139, 2002.
7. Alias, J., Characteristics of wave formation in rails, *Rail Int.*, 17(11), 17–23, 1986.
8. Lévy, D., Conception d'un système de mesure et d'analyse de l'usure ondulatoire des rails, *Revue Général des Chemins de Fer*, 108, 51–57, 1989.
9. Grassie, S. L. and Kalousek, J., Rail corrugation: characteristics, causes and treatments. Proceedings of the Institution of Mechanical Engineers, Part F, *J. of Rail and Rapid Transit*, 207(F1), pp. 57–68, 1993.
10. Kalousek, J. and Grassie, S., Rail corrugation: causes and cures, *Int. Rail. J.*, July, 24–26, 2000.
11. Clark, R. A., Dean, P. A., Elkins, J. A., and Newton, S. G., An investigation into the dynamic effects of railway vehicles running on corrugated rails, *J. Mech. Eng. Sci.*, 24(2), 65–76, 1982.
12. Igeland, A., Railhead corrugation growth explained by dynamic interaction between track and bogie wheelsets, Proceedings of the Institution of Mechanical Engineers, Part F, *J. of Rail and Rapid Transit*, 210(F1), pp. 11–20, 1996.
13. Clark, R. A. and Foster, P., On the mechanics of rail corrugation formation, *Proceedings of the Eighth IAVSD-Symposium the Dynamics of Vehicles on Roads and on Tracks*, held at the Massachusetts Institute of Technology, Cambridge, Massachusetts, August 14–19, 1983, Hedrick, J. K. Ed., Swets & Zeitlinger B V, Lisse, pp. 72–85, 1983.
14. Igeland, A., *Dynamic train-track interaction: simulation of railhead corrugation growth under a moving bogie using mathematical models combined with full-scale measurements*, PhD thesis, Division of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, 1997.
15. Werner, K., Diskrete Riffelabstände und die Suche nach den Ursachen der Schienenriffeln, *Zeitschrift für Eisenbahnwesen und Verkehrstechnik (ZEV) Glasers Annalen*, 110(10), 353–359, 1986.
16. Frederick, C. O., A rail corrugation theory, Contact Mechanics and Wear of Rail–Wheel Systems II, *Proceedings of the International Symposium*, Gladwell et al., Ed., University of Rhode Island, Kingston, R I, July 8–11, 1986, University of Waterloo Press, Waterloo, Ontario, Canada, pp. 181–211, 1987.
17. Brockley, C. A. and Ko, P. L., An investigation of rail corrugation using friction-induced vibration theory, *Wear*, 128, 99–106, 1988.
18. Valdivia, A. R., A linear dynamic wear model to explain the initiating mechanism of corrugation, *Proceedings of the Tenth IAVSD-Symposium the Dynamics of Vehicles on Roads and on Tracks*, held in Prague, August 24–28, 1987, Apetaur, M., Ed., Swets & Zeitlinger B V, Lisse, pp. 493–496, 1988.
19. Knothe, K. and Valdivia, A., Riffelbildung auf Eisenbahnschienen — Wechselspiel zwischen Kurzzeitdynamik und Langzeit-Verschleissverhalten, *Zeitschrift für Eisenbahnwesen und Verkehrstechnik (ZEV) Glasers Annalen*, 112(2), 50–57, 1988.
20. Knothe, K. and Ripke, B., The effects of the parameters of wheelset, track and running conditions on the growth rate of rail corrugations, *Proceedings of the 11th IAVSD-Symposium the Dynamics of Vehicles on Roads and on Tracks*, held in Kingston, Ontario, August 21–25, 1989, Anderson, R., Ed., Swets & Zeitlinger B V, Lisse, pp. 345–356, 1989.
21. Hempelmann, K. and Knothe, K., An extended linear model for the prediction of short-pitch corrugation, *Wear*, 191, 161–169, 1996.
22. Müller, S., *Linearized Wheel–Rail Dynamics — Stability and Corrugation*, Fortschritt-Berichte VDI, No 12-369, VDI-Verlag, Düsseldorf, 1998.
23. Suda, Y. and Iguchi, M., Basic study of corrugation mechanism on rolling contact in order to control rail surfaces, *Proceedings of the 11th IAVSD-Symposium the Dynamics of Vehicles on Roads and on Tracks*, held in Kingston, Ontario, August 21–25, 1989, Anderson, R., Ed., Swets & Zeitlinger B V, Lisse, pp. 345–356, 1989.
24. Kalousek, J., Keeping heavy haul track corrugation-free, *Rail. Gaz. Int.*, 145, 545–547, 1989.
25. Tassilly, E. and Vincent, N., A linear model for the corrugation of rails, *J. of Sound Vib.*, 150(1), 25–45, 1991.

26. Tassilly, E. and Vincent, N., Rail corrugations: analytical model and field tests, *Wear*, 144, 163–178, 1991.
27. Cervoni, G. and Vincent, N., Lutte contre l'usure ondulatoire des rails. Proposition d'une approche globale liant la voie et le matériel roulant, *Revue Général des Chemins de Fer*, 105(10), 559–570, 1986.
28. Bogacz, R., Brzozowski, M., Mahrenholtz, O., and Ronda, J., Dynamic effects in a rolling contact problem, *Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM)*, 67(4), T176–T179, 1987.
29. Böhmer, A. and Klimpel, T., Plastic deformations of corrugated rails: a numerical approach using material data of rail steel, Proceedings of the Fifth International Conference on Contact Mechanics and Wear, Tokyo, Japan 2000, *Wear*, 253(1–2), pp. 150–161, 2001.
30. Nielsen, J. C. O., Numerical prediction of rail roughness growth on tangent railway track, *J. of Sound Vib.*, 267, 537–548, 2003.
31. Nielsen, J. C. O. and Johansson, A., Out-of round railway wheels — a literature survey, Proceedings of the Institution of Mechanical Engineers, Part F, *J. of Rail and Rapid Transit*, 214(F2), pp. 79–91, 2000.
32. Johansson, A., and Nielsen, J. C. O., Out-of-round railway wheels — Wheel–rail contact forces and track response derived from field tests and numerical simulations, Proceedings of the Institution of Mechanical Engineers, Part F, *J. of Rail and Rapid Transit*, 217(F2), pp. 135–145, 2003.
33. Andersson, C. and Dahlberg, T., Load impacts at railway turnout crossing, *Vehicle System Dynamics*, 33(Suppl.), 131–142, 1999.
34. Berggren, E., *Dynamic Track Stiffness Measurement – A New Tool for Condition Monitoring of Track Substructure*, Licentiate Thesis, Report TRITA AVE 2005:14, Railway Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 2005.
35. Fermér, M. and Nielsen, J. C. O., Wheel–rail contact forces for flexible versus solid wheels due to tread irregularities, Proceedings of the 13th IAVSD Symposium on Dynamics of Vehicles on Roads and on Tracks, Chengdu, Sichuan, China, August 23–27, 1993, *Vehicle Syst. Dyn.*, 23(suppl.), pp. 142–157, 1994.
36. Nielsen, J. C. O., *Train–track interaction, Coupling of moving and stationary systems — theoretical and experimental analysis of railway structures considering wheel and track imperfections*, PhD thesis, Division of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, 1993.
37. Grassie, S. L. and Cox, S. J., The dynamic response of railway track with flexible sleepers to high frequency vertical excitation, *Proc. of Inst. Mech. Eng., IMechE*, 198(D7), 117–123, 1984.
38. Månsson, F., *Simulation of dynamic train–track interaction considering track settlement*, MSc thesis, Department of Mechanical Engineering, Linköping University, Linköping, Sweden, Report LiTH-IKP-EX-1668, 2000.
39. Dahlberg, T., Vehicle–bridge interaction, *Vehicle Syst. Dyn.*, 13, 187–206, 1984.
40. Olsson, M., *Analysis of structures subjected to moving loads*, PhD thesis, Lund Institute of Technology, Division of Structural Mechanics, Report TVSM-1003, Lund, 1986.
41. Oscarsson, J., *Dynamic train–track interaction: linear and non-linear track models with property scatter*, PhD thesis, Department of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, ISSN 0346-718X, 2001.
42. Zhai, W. and Cai, Z., Dynamic interaction between a lumped mass vehicle and a discretely supported continuous rail track, *Comput. & Struct.*, 63(5), 987–997, 1997.
43. Fryba, L., *Vibration of Solids and Structures Under Moving Loads*, 3rd ed., Thomas Telford, London, ISBN 0-7277-2741-9, 1999.
44. Knothe, K. L. and Grassie, S. L., Modelling of railway track and vehicle–track interaction at high frequencies, *Vehicle Syst. Dyn.*, 22, 209–262, 1993.
45. Popp, K., Kruse, H., and Kaiser, I., Vehicle–track dynamics in the mid-frequency range, *Vehicle Syst. Dyn.*, 31(5–6), 423–464, 1999.
46. Ripke, B., Hochfrequente Gleismodellierung und Simulation der Fahrzeug-Gleis-Dynamik unter Verwendung einer nichtlinearen Kontaktdynamik, *Fortschritt-Berichte VDI*, Reihe 12 Nr 249, VDI-Verlag, Düsseldorf, 1995.
47. Andersson, C., *Modelling and simulation of train–track interaction including wear prediction*, PhD thesis, Department of Applied Mechanics, Chalmers University of Technology, Gothenburg, Sweden, 2003.

48. Grassie, S. L., Models of railway track and train-track interaction at high frequencies: Results of benchmark test, *Vehicle Syst. Dyn.*, 25(Suppl.), 243–262, 1996.
49. Andersson, C. and Abrahamsson, T., Simulation of interaction between a train in general motion and a track, *Vehicle Syst. Dyn.*, 38(6), 433–455, 2002.
50. Fröhling, R. D., *Deterioration of railway track due to dynamic vehicle loading and spatially varying track stiffness*, PhD thesis, Faculty of Engineering, University of Pretoria, Pretoria, South Africa, 1997.
51. Wu, T. X. and Thompson, D. J., Vibration analysis of railway track with multiple wheels on the rail, *J. Sound Vib.*, 239(1), 69–97, 2001.
52. Dinkel, J., *Ein semi-analytisches Modell zur dynamischen Berechnung des gekoppelten Systems Fahrzeug-Fahrweg-Untergrund für das Oberbausystem Feste Fahrbahn*, PhD thesis, Technische Universität, München, Germany, ISSN 0941-925X, 2001.
53. Diana, G., Cheli, F., Bruni, S., and Collina, A., Interaction between railroad superstructure and railway vehicles, *Vehicle Syst. Dyn.*, 23(Suppl.), 75–86, 1994.
54. Li, D. and Selig, E. T., Wheel/track dynamic interaction: track substructure perspective, *Vehicle Syst. Dyn.*, 24(Suppl.), 183–196, 1995.
55. Sun, Y. Q. and Dhanasekar, M., A dynamic model for the vertical interaction of the rail track and wagon system, *Int. J. of Solids Struct.*, 39, 1337–1359, 2002.
56. Sato, Y., Odaka, T., and Takai, H., Theoretical analyses on vibration of ballasted track, *Q. Rep. Rail. Tech. Res. Inst. (Jpn.)*, 29(1), 30–32, 1988.
57. Dahlberg, T., Vertical dynamic train-track interaction — verifying a theoretical model by full-scale experiments, *Vehicle Syst. Dyn.*, 24(Suppl.), 45–57, 1995.
58. Cowper, G. R., The shear coefficient in Timoshenko's beam theory, *Trans. ASME, J. of Appl. Mech.*, 33, 335–340, 1966.
59. Dahlberg, T. and Nielsen, J. C. O., Dynamic behaviour of free-free and in-situ concrete railway sleepers, *International Symposium on Precast Concrete Railway Sleepers*, Madrid, Spain, pp. 393–416, 1991.
60. Grassie, S. L., Gregory, R. W., Harrison, D., and Johnson, K. L., The dynamic response of railway track to high frequency vertical excitation, Proceedings of the Institution of Mechanical Engineers, Part C, *J. Mech. Eng. Sci.*, 24(2), pp. 77–90, 1982.
61. Thompson, D. J., van Vliet, W. J., and Verheij, J. W., Development of the indirect method for measuring the high frequency dynamic stiffness of resilient elements, *J. of Sound Vib.*, 213, 169–188, 1998.
62. Fenander, Å., *Modelling stiffness and damping by use of fractional calculus, with application to railpads*, PhD thesis, Division of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, 1997.
63. Wu, T. X. and Thompson, D. J., Effects of local preload on the foundation stiffness and vertical vibration of railway track, *J. of Sound Vib.*, 219(5), 881–904, 1999.
64. Dahlberg, T., Dynamic interaction between train and non-linear railway track model, *Proceedings of the Fourth International Conference on Structural Dynamics, EURO-DYN2002*, Munich, Germany, 2–5 September, 2002, Grundmann, and Schuëller, Ed., Swetz & Zeitlinger, Lisse, ISBN 90 5809 510X, 2002.
65. Thompson, D. J. and Verheij, J. W., The dynamic behaviour of rail fasteners at high frequencies, *Appl. Acoust.*, 52, 1–17, 1997.
66. Grassie, S. L., Dynamic modelling of concrete railway sleepers, *J. Sound Vib.*, 187, 799–813, 1995.
67. Knothe, K., Wechselwirkung zwischen Fahrzeug und Fahrweg, *Zeitschrift für Eisenbahnwesen und Verkehrstechnik*, 122(5), 173–180, 1998.
68. Peplow, A. T., Oscarsson, J., and Dahlberg, T., *Review of research on ballast as track substructure*, Report F189, Department of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, ISSN 0349-8107, 1996.
69. Zhai, W. M., Wang, K. Y., and Lin, J. H., Modelling and experiment of railway ballast vibrations, *J. Sound Vib.*, 270, 673–683, 2004.
70. Jacobsson, L., *Review of research on railway ballast behaviour — Experimental findings and constitutive models*, Report F208, Department of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden, ISSN 0349-8107, 1998.

71. Iwnicki, S., Grassie, S., and Kik, W., Track settlement prediction using computer simulation tools, *Vehicle Syst. Dyn.*, 33(Suppl.), 2–12, 2000.
72. Suiker, A. S. L., *The mechanical behaviour of ballasted railway tracks*, PhD thesis, Delft Technical University, Delft, The Netherlands. Delft University Press, ISBN 90-407-2307-9, 2002.
73. Popp, K. and Schiehlen, W., *System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade*, Springer, Berlin, 2003, ISBN 3-540-43892-0.
74. Dahlberg, T., Some railroad settlement models — a critical review, Proceedings of the Institution of Mechanical Engineers, Part F, *J. of Rail and Rapid Transit*, 215(F4), pp. 289–300, 2001.
75. Esveld, C., *Modern railway track*, MRT-Productions, Duisburg, Germany, 1989, ISBN 90-800324-1-7, (new edition 2001).
76. Profillidis, V. A., *Railway Engineering*, 2nd ed., Ashgate, Aldershot UK, 2000.

---

# 7 Gauging Issues

*David M. Johnson*

## CONTENTS

I.	Philosophy and History of Gauging .....	182
A.	Gauges .....	183
1.	Static Gauges .....	183
2.	Geometric or Swept Gauges .....	183
B.	Swept Envelopes .....	184
1.	Kinematic Envelopes .....	184
2.	Dynamic Envelopes .....	184
C.	Hybrid Gauges .....	184
1.	Pseudokinematic Gauges .....	184
2.	Kinematic Gauges .....	184
D.	International Methods .....	185
1.	UIC .....	185
II.	Components of Gauging .....	187
A.	Structure .....	187
1.	Shape .....	187
2.	Accuracy of Measurement .....	187
B.	Track .....	189
1.	Track Position .....	189
2.	Track Geometry .....	189
C.	Vehicle .....	191
1.	Geometric Considerations — Overthrow on Curves .....	191
2.	Kinematic Considerations .....	192
a.	Movement from Curving Forces .....	192
b.	Movement due to Motion .....	193
c.	Critical Speeds .....	193
d.	Effect of Loading .....	195
e.	Time Factors .....	196
f.	Vehicle Height .....	196
3.	Vehicle Tolerances .....	196
a.	Tilting Trains .....	197
i.	Effects of Speed .....	199
ii.	Critical Speeds .....	199
iii.	Time Factors .....	200
b.	Advanced Modelling .....	201
III.	Interaction between Gauging Components .....	202
A.	Vehicle–Track Interaction .....	202
1.	Wheelset Movement .....	202
B.	Track–Structure Interaction .....	204



1. Track Tolerances .....	204
a. Lateral Track Positional Tolerance.....	204
b. Vertical Track Positional Tolerance .....	204
c. Cross-Level Error .....	204
d. Sidewear .....	204
C. Structure–Vehicle Interaction .....	205
1. Clearances.....	205
2. Stepping Distances .....	206
References.....	206

## I. PHILOSOPHY AND HISTORY OF GAUGING

Gauging is the name given to the techniques used to ensure that rail vehicles fit through the infrastructure and pass by each other in safety. Increasingly, there is emphasis on maximising the capacity of the railway corridor through a more thorough understanding of the gauging system, and reducing conservatism in the processes that ensured adequate space was available when the railways were first built.<sup>1</sup>

This chapter is intended to provide readers with an insight into the techniques used in Britain, where an infrastructure of up to 200 years old is now required to deliver the capability of running large, intermodal freight trains and passenger trains of increased capacity and comfort for which they were not designed. Internationally, capacity constraints are also real, but invariably infrastructure that has been built later requires less incremental change to cope with the near standardised loads being transported. Virtually every country has its own methodologies through which loading gauge is managed, and it is beyond the scope of this chapter to provide anything but a look at gauging principles and approaches that form the basis of most of these gauging practices. Beyond Britain, western Europe has adopted a near-standard UIC (Union Internationale des Chemins de Fer) approach, which is described briefly for information.

Railways were originally built to *gauges* — vehicles to a maximum vehicle (or load) gauge and infrastructure (structures) to a minimum structure gauge. A *clearance* was included between the vehicle gauge and the structure gauge to allow for unknowns, or those items that were known but had not been included in the gauge.

At the turn of the last century, the U.K. Board of Trade (whose job it was to monitor rail traffic) had registered 127 different load gauges from (private) railway companies. No load gauge was universal — except, perhaps, the smallest. Many railway administrations still work by these simple gauging methods, indeed the methodology used in most of Europe is a derivative of the earlier fixed gauge approach. Much of the original railway infrastructure built to accommodate these load gauges still exists, but the trend is to increase vehicle size. The challenge is to develop new gauging methodologies that enable this to happen. The original methods provide a good starting point.

British engineers, forced to make increasingly better use of small, Victorian (predominantly arched) infrastructure, have been at the forefront of developing gauging systems that analyse the vehicle-infrastructure interaction on a case-by-case basis to minimise the cost of upgrade works needed to run these larger passenger coaches and bigger freight loads.

New developments also introduced additional factors that had to be taken into account. Early railways used short wagons, and their swept envelopes were not significantly different to their static size. The introduction of long coaches rather than short carriages generated a new vehicle-infrastructure interaction. Overthrow complicated the basic interface between mechanical engineer

and civil engineer as it related to both the curve geometry and the arrangement of the vehicle. Railways that had worked well with short vehicles now exhibited weaknesses in certain situations, having restricted clearances on curves. The first significant development in the trend towards gauging analysis was the adjustment of gauges to include vehicle overthrows associated with curvature of the track.

Increasingly, the understanding of vehicle dynamics has led to techniques that predict suspension movements (and hence the local swept envelope) in response to curvature and speed. Those techniques, largely developed by British Rail Research, became invaluable in the acceptance processes for air-suspension rolling stock (with implicitly softer suspension) in the 1980s. However, although it became increasingly possible to calculate vehicle movements with precision, gauging standards were slow to react to these improved methodologies and for a while failed to allow all of the benefits that the techniques could offer.

Significant advances in vehicle dynamics and the introduction of computerised techniques have allowed tolerances, clearances, and “unknowns” to be defined in a more robust manner. Tolerances may now be calculated accurately, and clearances provided for the fewer remaining unknown or incalculable effects. An important factor is that as unknowns are understood, they may be removed from mandated clearances and analysed as appropriate tolerances. Conservatism is thus being progressively removed from the system.

Modern gauging technology is far removed from the simple pen and paper solutions of 100 years ago. This chapter aims to give an insight into the factors considered and the calculations performed in modern gauging methods.

In simple terms, gauging has moved on from being the technique for simply deciding whether something will fit to what can be done to enable something to fit.

## A. GAUGES

### 1. Static Gauges

In what may be described as “simple gauging” the mechanical engineer built vehicles to a “vehicle gauge,” being the maximum cross section of the train, and the civil engineer ensured that structures were always larger than the “structure gauge.” A separation between the two, known as clearance, allowed for any variations of track position (track being anything but the “permanent way” that it is traditionally called) and the suspension movements of the vehicle. These are known as static gauges.

### 2. Geometric or Swept Gauges

Geometric or swept gauges represented a development of the above, where the vehicle was substantially affected by the geometry of the track. On curves, vehicles sweep a larger path than on straight track, a phenomenon known as “overthrow.” The amount depends on the tightness of the curve, the vehicle bogie (or axle) centres and the overall length. In the immediate postnationalisation period in Britain (approximately 1951 onwards) “national gauges” for rail passenger vehicles (known as C1) and freight vehicles (W5) were defined, based upon the vehicle gauges used by the majority of component railway companies absorbed into British Railways (BR). C1 and W5 gauges are geometric gauges, requiring knowledge of both vehicle parameters and curve geometry in order to calculate the clearance to a structure. A clearance of 150 mm (6 in.) was usually allowed, comprising 100 mm potential vehicle movement on its suspension and 50 mm for potential track positional and geometric errors.

In Great Britain, details of current and historic gauges, together with other useful information on current gauging practice may be found in a guidance note published by the Rail Safety and Standards Board.<sup>10</sup>

## B. SWEEP ENVELOPES

### 1. Kinematic Envelopes

In the late 1970s and 1980s cost engineering became prevalent in Britain, particularly in the area of track maintenance. A given ride quality can be achieved by maintaining high-quality track geometry or by providing softer vehicle suspensions. The former solution is particularly expensive since as track quality is raised the cost of providing this increases exponentially. The new generation of rolling stock then being commissioned could readily be given suspension capable of providing adequate ride comfort on poorer track. Air suspension provided this mechanism, but at the expense of having greater kinematic movement (movement associated with the speed of the vehicle). The methodologies described would have meant that the infrastructure would have required enlargement to maintain clearance. However, it was recognised that by relating kinematic movement to operating environment, the locations where enlarged infrastructure was required could be minimised. A publication known as “Design Guide BaSS 501”<sup>3</sup> provided a methodology whereby the kinematic envelope of a vehicle (the space required by a given vehicle, moving at speed) at a specific location could be manually calculated from a number of input parameters. The techniques used are quasistatic, equating dynamic conditions to stationary forces, and are generally conservative. Nevertheless, the techniques have been very successful in allowing larger trains to operate on restrictive infrastructure at minimal cost. In particular, a derivation of the technique has allowed tilting trains to be designed for Britain that would otherwise have been of a nonviable cross section if traditional gauging rules were applied.

### 2. Dynamic Envelopes

Increasingly, the conservatism of quasistatic gauging has challenged the development of larger vehicles. Furthermore, certain basic assumptions about vehicle behaviour are oversimplifications that are necessary to create a technique capable of manual calculation. With the advent of computerised gauging software (ClearRoute™) and vehicle dynamics simulation software (VAMPIRE™), the millions of calculations necessary for the calculation of the dynamic gauging performance of vehicles can be undertaken in a practical timescale. (ClearRoute is the registered trademark of Laser Rail Ltd. and VAMPIRE is the registered trademark of AEA Technology Rail.)

## C. HYBRID GAUGES

### 1. Pseudokinematic Gauges

A pseudokinematic gauge is where maximum kinematic movements are included in the gauge. It is common for light rail and metro systems to use a vehicle gauge that includes all suspension movement for particular vehicles (this is sometimes known as a *red-line* kinematic gauge). The system used across Europe is a further development of this, using a reference profile to define a notional boundary between train and infrastructure under certain, prescribed limits, beyond which both vehicle builder and infrastructure controller must make adjustment. Pseudokinematic gauges work well for new infrastructure, but lead to the restriction of vehicle size as softer suspension is introduced.

### 2. Kinematic Gauges

It should be noted that the swept envelope of a vehicle is really a series of swept envelopes, since some parts of a vehicle move more than others (depending on where the section of the vehicle is in relation to bogie or axle centres), and some sections may have projections. In particular, the cross

section of a vehicle at the bogie/axle positions will exhibit zero throw, whereas the section located in the centre of the vehicle will have maximum throw towards the inside of a curve. Similarly, the ends of a vehicle will have maximum throw towards the outside of a curve. A kinematic gauge (examples being those which define the GB/SNCF/SNCB Eurostar<sup>9</sup> vehicle and the new W12 freight gauge) is a union of different kinematic vehicle dimensions defining the largest envelope under given operating conditions. W12, for instance, consists of many thousands of gauge diagrams appropriate to curve radius, installed cant, and speed. The Eurostar gauge is computer generated from the same set of parameters.

Gauges refer not only to a cross-sectional profile, but also to a set of rules that must be applied. A basic understanding of the gauge definitions will show that the clearance required for safe operation is intrinsically linked to the derivation of the gauge and the parameters considered.

### D. INTERNATIONAL METHODS

#### 1. UIC

UIC gauging methods are defined in the 505 series of leaflets, and use reference gauges as the basis for gauging<sup>4-6</sup> (Figure 7.1). The method dates back to 1913 and has been developed as a hand-calculated technique that contains a number of simplifications. To ensure safety it is very conservative. No clearance is required — the conservatism ensures that contact is not physically possible. From a vehicle perspective, this reference profile defines a base gauge into which the vehicle must fit under certain defined conditions.

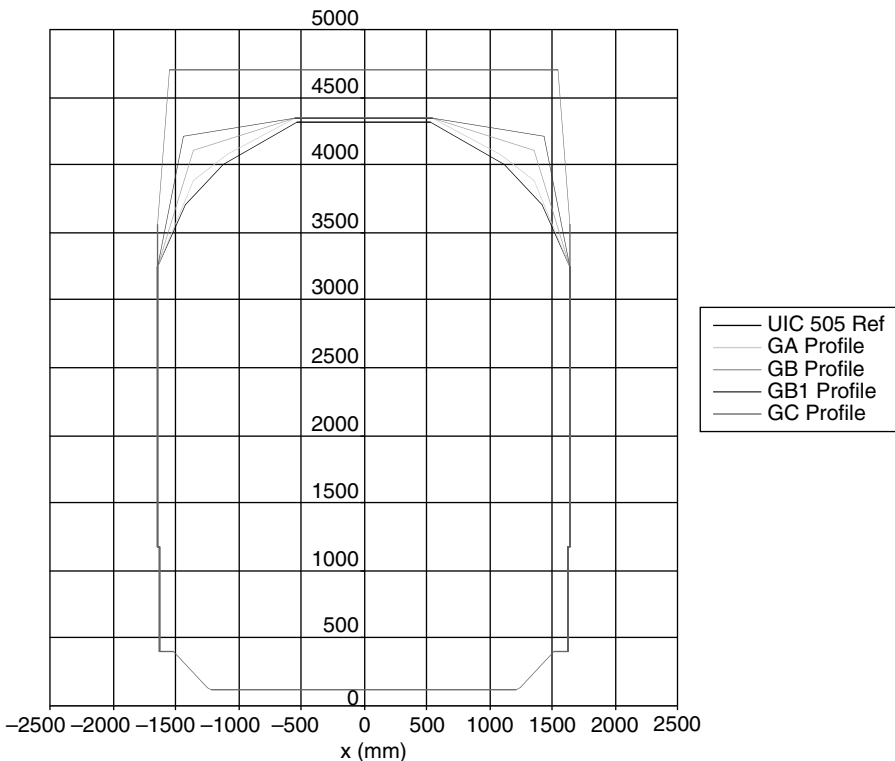


FIGURE 7.1 Basic UIC reference profiles.

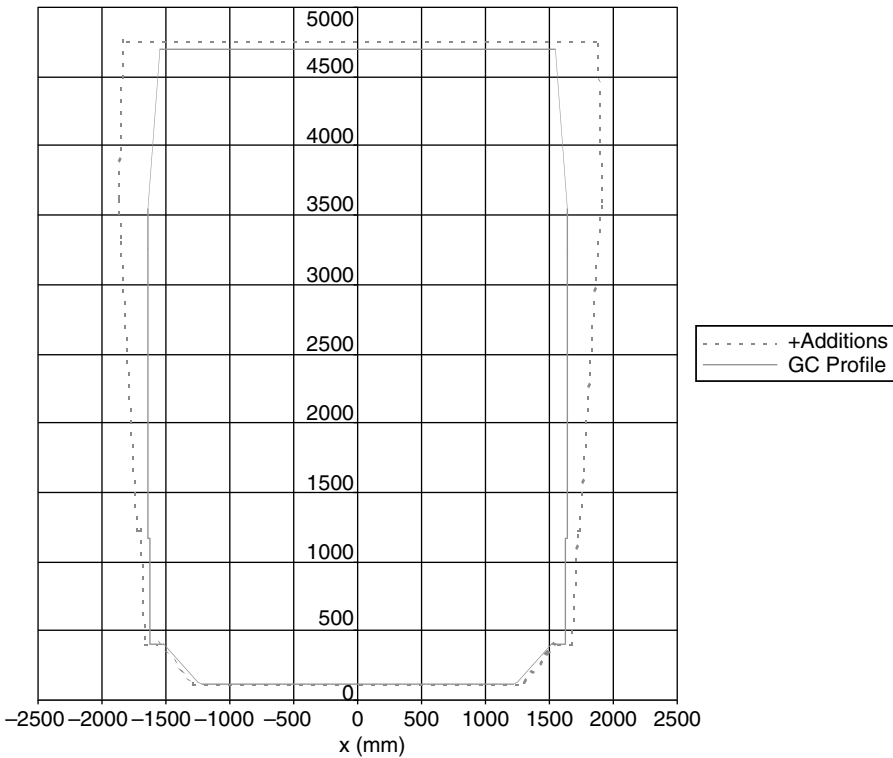
The conditions under which vehicles must be contained within the reference dimensions are (inclusively):

- Horizontally, on a 250 m radius curve
- Under 50 mm of installed cant and 50 mm of developed cant deficiency
- With full horizontal suspension travel
- Stationary

Note that this is not a vehicle gauge in the British sense. It represents a vehicle snapshot, which requires further analysis in applying this to the infrastructure. In particular, the following effects (known as infrastructure additions) must be taken into account:

- Authorised projections on curves and throws on curves between 150 and 250 m radius (note that UIC vehicles are not designed to run through curves of less than 150 m radius)
- The effect of track quality associated with speed
- Gauge widening
- Roll from cant excess or deficiency above 50 mm
- Track alignment tolerances
- Static and dynamic cross level error
- Possible vehicle loading asymmetry
- Vertical curvature

Figure 7.2 shows a developed vehicle gauge for a specific condition (1000 m horizontal curve, 150 mm installed cant, 160 kph developing a cant deficiency of 150 mm, good track, 1000 m



**FIGURE 7.2** UIC GC reference profile with infrastructure additions.

vertical curve). Under strict UIC rules, it is not necessary to provide any clearance to this envelope, there is sufficient conservatism in the calculation process to ensure that vehicle-structure contact can never occur. However, the UIC process by this very conservatism makes poor use of the space envelope that would normally be used in Britain.

## II. COMPONENTS OF GAUGING

A full gauging model requires the interactions between structure, track, and vehicle to be understood. Before examining these interactions an understanding of the behaviour of the individual components is required.

### A. STRUCTURE

#### 1. Shape

In Britain, clearance issues mainly relate to arch bridges and tunnels. Containers, particularly, provide an obvious “square peg in a round hole” challenge when trying to run the former through the latter. Overbridges generate height restrictions and platforms generate width restrictions. All obstacles in the vicinity of the train must be measured.<sup>11,12</sup>

#### 2. Accuracy of Measurement

Accuracy of structure measurement is becoming increasingly significant. As analysis methodologies improve, conservatism in infrastructure measurement that results from inaccurate measurement becomes less acceptable. In particular, while it may be possible to define the swept envelope of a vehicle to within a few millimetres, the accuracy of many structure measuring techniques may be poorer than 50 mm. In order to maximise infrastructure capacity, it is important that an opportunity is not lost through poor measurement accuracy.

Figure 7.3 shows various structure measuring systems on a graph of measurement accuracy (bottom axis) and relative cost (vertical axis).

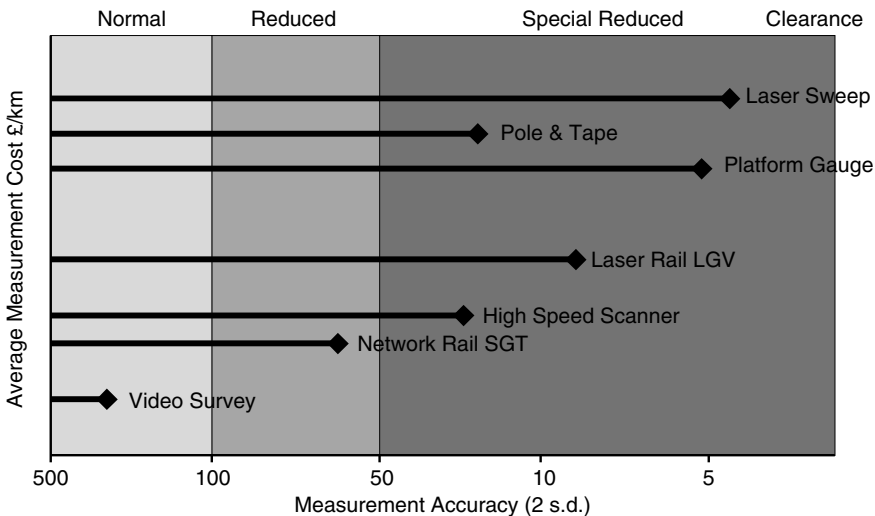


FIGURE 7.3 Structure survey systems showing relative accuracy and running cost.

The top axis shows bands relating to clearance regimes used on the British infrastructure. Normal clearances require minimal control measures, reduced clearances require better management of track position, and special reduced clearances require stringent controls.

Measurement cost relates to a cost per kilometre, assuming that the system is fully utilised for a measuring shift. Costs arise from a variety of sources, and include the following components:

- Capital cost
- Maintenance cost
- Operating cost
- Possession/protection cost
- Train path cost
- Data processing cost
- Effective working shift length

Measurement accuracy is the sum of inaccuracies in the measuring system. Quoted accuracies provide only an indication of this. How the measuring system relates measurement of the structure to the rails is just as important as its ability to measure the structure, but is often overlooked. Although a measuring system may be quoted as having a particular accuracy, the statistical accuracy should be considered in defining what accuracy of measurement *may be expected*. Quoted accuracies are often misleading, and if incorrectly used may result in unsafe calculations. Generally, an accuracy of two standard deviations would be used, representing a 95% confidence limit, based upon a normal distribution. System calibration over the full range of measurement provides the only reliable method of defining measuring system accuracy, particularly for more accurate systems.

From [Figure 7.3](#), it can be seen that some methods of measurement are more appropriate than others in terms of accuracy, cost, and the clearance regime that is being monitored. Video assessment, while cheapest, is a method of safely determining which structures are sufficiently far from the track to present no risk. It cannot be used to identify locations that may operate under



**FIGURE 7.4** The LaserSweep profiler.



**FIGURE 7.5** Road–Rail laser gauging vehicle, operated by Laser Rail Ltd.

special reduced clearances. LaserSweep™ (Figure 7.4) would not be economic to use for screening purposes, and is best restricted to locations where clearances may be tight. In considering an appropriate measuring method, it must be remembered that traffic type and future change of use may indicate that a more accurate measuring method should be used (Figure 7.5), if this proposed traffic were likely to reduce available clearance (for example, a tilting train). A structure gauged for containerised freight would require more accurate measurements in the area of the top corner of the container.

## **B. TRACK**

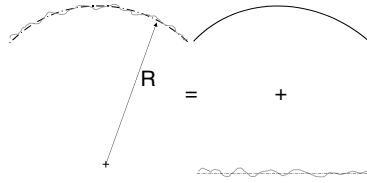
### **1. Track Position**

Knowledge of the position of the track, and the amount by which it may move, is vital in accurate gauging calculations. Track position controls vehicle position. Track position may vary as a result of traffic loading, effects of weather, and most importantly, the movement that is allowed for maintenance and alignment purposes. Some track is better restrained than others. Slab track is generally very stable. Ballast glueing can reduce lateral movement (although it would not affect vertical movement significantly). Strutting sleepers against platforms will generally reduce lateral movement towards the platform. Track fixity is discussed later in this chapter.

### **2. Track Geometry**

As we try to model vehicle behaviour more accurately, irregularities of track geometry become increasingly important. In basic gauging, the only geometric input comes from the curve radius used in the calculation of throws. In the more complex models, the full spectrum of track geometry must be considered.





**FIGURE 7.6** Superposition of track design alignment and irregularity.

Track geometry is essentially the variation of lateral and vertical track position in relation to the longitudinal position. On perfect, straight track, there is no track geometry. However, track is neither constantly straight nor perfect, and consists of straights, curves, and track irregularities. This is generally referred to as design geometry and roughness. These parameters are handled separately, in that they have different effects on vehicles. However, in practice, the boundary between geometry and defect is impossible to define. The key issue is one of wavelength. Generally, design geometry is of long wavelength and defects are of short wavelength. However, it is possible for deviations to be longer than design alignments. In particular, some transition curves are very short, and would, in other circumstances, be considered to be irregularities.

In Figure 7.6, the shape on the left would be considered to be normal curved track. It is a summation of the true curvature (shown top right) and irregularity (shown bottom right). If the exact (design) curvature is known, then the irregularity can easily be calculated. In practice, it is unusual (in Britain) to know the exact design curvature, and thus approximations must be performed to extract the two shapes. This process is known as filtering. The use of a high-pass filter (one which lets high frequency, short wavelength through) would produce the roughness from the measured, compound profile. This process is used to extract track roughness from track topography in order to determine when to maintain the track. The curvature of the track can be calculated by using a low pass filter. This process is sometimes referred to as regression, the regressed alignment being the underlying geometric shape of the track.

The filter used makes a significant difference in the separation of track roughness and track geometry. A commonly used filter is the Butterworth filter, originally developed for medical instrument technology in removing 50 Hz mains “hum” from body electrode signals. However, this filter, on its own, creates a spacial phase shift (where there is a longitudinal movement of peaks and troughs), which may be corrected by performing a reverse pass of the same filter.

It is important to understand how to use track geometry data in gauging calculations. It will become clear that the curvature data and the roughness data are used for different purposes. However, measuring track provides a single curvature reading, the nature of which is dependent on how it is measured and processed.

High-speed systems tend to use inertial geometry measurement, sampling at frequent intervals. Such systems are principally used for track quality recording, but have been adapted for use on gauging systems. Inertial systems are best at measuring high rates of change of curvature. Since sharper curves provide the greatest input to gauging calculations, these systems may be used provided the roughness (which they can also measure well) is removed. If unfiltered geometry is used, there is a risk of under- or overcalculating throws, and double-counting dynamic effects.

Manual systems tend to sample infrequently, usually every 10 to 20 m. As such they tend to pick up more general curvature without significant effect from track geometry errors. However, it should be noted that track faults will have an effect on the measured curvature. The true “design” curvature may need to be extracted using methods such as “Hallade,” a filtering technique using a combination of mathematics and human skill that is used to determine optimal lateral track alignment.

**C. VEHICLE**

**1. Geometric Considerations — Overthrow on Curves**

The axles of a railway vehicle form the end points of a chord placed on curved track. The body represents an extension of this chord. As the vehicle traverses the curve the centre of the vehicle is thrown towards the inside of the curve, and the end of the vehicle is thrown towards the outside of the curve. The overthrow effect increases with vehicle length and tighter curvature. A bogie is simply a vehicle with centre throw only. Vertical curvature is not generally an issue on main line railways, but is often considered on metro and light rail systems.

The equations for calculating throw are shown below. The simplified equations ignore some small angle effects leading to marginal inaccuracy, but are useful for quick calculations.

If we consider Figure 7.7, the overthrow at a point on a vehicle body is the difference between the radial distance from the track centreline to the point, and the lateral distance from the vehicle centreline to the point ( $W_o$  or  $W_i$ ). This is calculated with the vehicle stationary.

Consider a vehicle with bogie centres  $L$ , and a bogie axle semispacing of  $a_o$  (the actual axle spacing is  $2a_o$ ).

The inner overthrow of a point  $U_i$  from the centre of a the vehicle is:

$$R - W_i - \sqrt{[U_i^2 + (J - W_i)^2]}$$

The outer overthrow of a point  $U_o$  from the centre of the vehicle is:

$$\sqrt{[U_o^2 + (J + W_o)^2]} - R - W_o$$

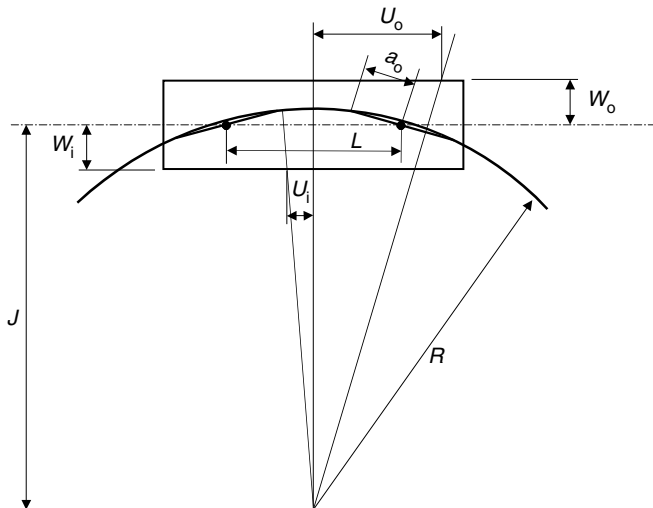
where  $J = \sqrt{[R^2 - a_o^2 - L^2/4]}$

The simplified equations are:

$$\text{Inner throw} = 125(L^2 - (L - 2x)^2)/R[\text{body}] + 500(a_o^2/R)[\text{bogie}]$$

$$\text{Outer throw} = 125((L + 2x)^2 - L^2)/R[\text{body}] - 500(a_o^2/R)[\text{bogie}]$$

where  $x = L/2 - U_o$ .



**FIGURE 7.7** Curve overthrow diagram.

## 2. Kinematic Considerations

The kinematic movement of a vehicle is the position adopted by the vehicle resulting from the forces applied, and allowances included. These can be summarised as:

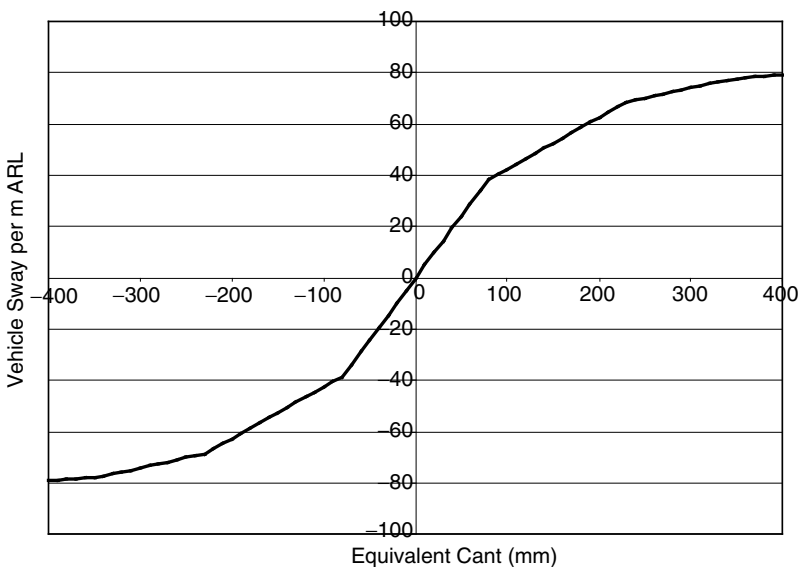
- Sway and roll drop due to curving forces
- Sway, lift, and drop due to motion
- Drops due to loading and suspension condition
- Vertical vehicle tolerances
- Lateral vehicle tolerances
- Lateral, vertical, and roll due to tilting suspension

### a. Movement from Curving Forces

A vehicle moves on its suspension in relation to applied lateral and vertical forces. In simple models, the total amount of lateral and vertical suspension travel (limited by bump stops) is used to determine the required clearance, and no relationship is assumed. As stated earlier, there is a low risk in this approach, but it makes poor use of infrastructure space.

BaSS 501 introduced the concept of relating vehicle movements to applied rolling force, expressed as an “equivalent cant.” This quasistatic method equates the movements of a vehicle operating in a dynamic environment to an applied static cant. The movement of the vehicle is defined simplistically as having components of vertical, lateral, and roll in a single plane (effects of yaw and pitch are ignored). A sample relationship between vehicle sway and applied cant is shown in Figure 7.8.

Figure 7.8 shows cant on the horizontal axis (cant excess to the left, cant deficiency to the right) and vehicle sway on the vertical axis (positive is sway to the outside of the curve). It can be seen that as cant increases (or decreases), the sway in the appropriate direction (inward or outward) increases (or decreases) also. The relationship is nonlinear, but is simplified into a series of straight-line relationships that relate to the increasing stiffness of the vehicle suspension as travel of the various stages is consumed. The “break points” occur, for instance at lateral bump stop contact, lateral



**FIGURE 7.8** Typical relationship between applied cant and sway.

metal bump stop contact, secondary vertical bump stop contact, and primary vertical bump stop contact. Similar graphs may be produced for roll drops. The general equations are:

$$\text{Sway} = K_{\text{sway}}D + E$$

$$\text{Drop} = K_{\text{drop}}D + J$$

where  $D$  is the (equivalent) cant,  $K_{\text{sway}}$ ,  $K_{\text{drop}}$ ,  $E$ , and  $J$  are the linear equation parameters for each segment.

These may be found by referring to a drawing for a given vehicle known as the “kinematic envelope.”

### *b. Movement due to Motion*

In normal curving, the force applied to a vehicle is related to the vehicle speed ( $V$ ) according to the following approximate formula (on conventional 1435-mm gauge track).

$$\text{Equilibrium Cant} = 11.82V^2/\text{Radius (m)}$$

If an equilibrium cant of 100 mm is required by a particular combination of radius and speed, then cant deficiency is the amount by which the applied cant is less than 100 mm, or cant excess is the amount by which the applied cant is greater than 100 mm. In an idealised world, this would be the only input of speed on the vehicle.

However, track is imperfect. As a result of unevenness of the rails, there will be local variations of the cant that the vehicle sees. Since this is a dynamic phenomenon, the effect on the vehicle suspension is likely to depend on a number of factors, in particular, the roughness of the track and the mass/inertia system of the train. BaSS 501 considers inputs due to track roughness as a component of equivalent cant (being the sum of actual cant experienced and other roll inducing effects expressed as cants), using a parameter known as  $K_{\text{speed}}$ . This parameter defines a linear relationship between the notional force applied to the vehicle (expressed as a cant) and the speed of a vehicle. Typically,  $K_{\text{speed}}$  is around 0.5, meaning that at 100 km/h, the rolling force seen by the vehicle, acting on the suspension, is equivalent to an additional  $\pm 50$  mm of cant above that caused by curving. It should be emphasised that equivalent cant is an *input* to the suspension relationship given in the previous section. Although the relationship between speed and equivalent cant is linear, the actual suspension movement is unlikely to be so.

Additionally, the vehicle responds to vertical track irregularities. There is no simple method to predict these. Accordingly, upward and downward movements of the vehicle calculated on the remaining suspension travel for given load cases are defined. This is known as “dynamic drop,” although the term “dynamic lift” is appropriate for upward movements. Both cases need to be considered simultaneously, since these define the “bounce” of the vehicle. A number of techniques have been used to limit these according to the true amount of suspension travel available once roll drop is considered (by relating it to equivalent cant) and to linearise the value with speed.

### *c. Critical Speeds*

Speed must also be considered in relation to the maximum sways (and drops) of a vehicle that it generates in service. A vehicle will usually be designed to run at a maximum line speed. This is generally limited by cant deficiency based upon passenger comfort. The faster a conventional vehicle travels around curves, the more it will sway towards the outside of the curve, limited only by suspension travel. However, we must consider the possibility that the vehicle may travel at reduced speed, or may even be stationary. The maximum static force on a vehicle to the inside of a curve occurs when stationary, due to an excess of cant. It is frequently assumed that this is the worst

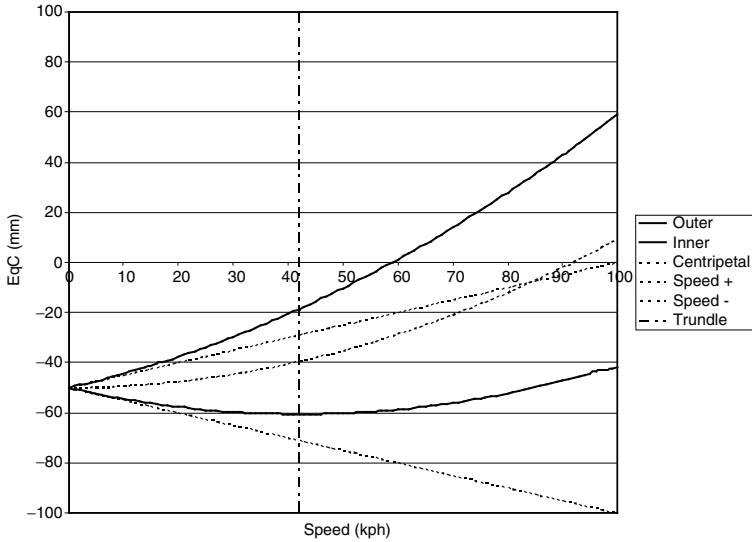


FIGURE 7.9 Components of trundle speed.

case for inside curve clearances. However, Figure 7.9 demonstrates the concept of “trundle speed,” a speed at which quasistatic force, and thus sway, to the inside of the curve is maximised.

Figure 7.9 shows a 2000 m radius curve with 50 mm installed cant. The vehicle has a  $K_{\text{speed}}$  of 0.5. There are three faint lines:

- The two linear lines radiating from  $-50$  mm equivalent cant represent the equivalent cant (leading to sway) as a result of the speed of the vehicle on imperfect track. Thus, at 100 km/h, the vehicle will experience an equivalent cant due to speed of  $-50 \pm 50$  mm. This is the same on both straight and curved track, although there would be no installed cant on the former.
- The parabolic line radiating from  $-50$  mm equivalent cant is the component of equivalent cant due to curving, which is towards the outside of the curve and increases with the square of speed. From this line, the balancing speed on this particular curve (considering radius and installed cant) is 92 km/h.

The solid lines show the summation of these forces, the upper being towards the outside of the curve, and the lower being towards the inside. It can be seen that the outside equivalent cant always increases as speed increases, confirming that maximum sway to the outside of the curve occurs at maximum speed. To the inside, it can be seen that the summation results in a minimum at 42 km/h, where the equivalent cant to the inside is a maximum. This speed is known as the trundle speed.

If we consider the equations used to develop this relationship, we can derive this value mathematically:

$$\text{Equivalent Cant } (D) \text{ due to curving} = 11.82V^2/R$$

$$\text{Equivalent Cant } (D) \text{ due to speed} = \pm K_{\text{speed}}V$$

$$D_{\text{outside}} = 11.82V^2/R + K_{\text{speed}}V - C(\text{Installed Cant})$$

$$D_{\text{inside}} = 11.82V^2/R - K_{\text{speed}}V - C(\text{Installed Cant})$$

Differentiating:

$$dD_{\text{inside}}/dV = 23.64V/R - K_{\text{speed}}$$

Resolving gives:

$$V_{\text{trundle}} = K_{\text{speed}}R/23.64$$

Trundle speed is thus directly related to radius and  $K_{\text{speed}}$ . Since the radius of straight track is infinite, this suggests that the trundle speed on straight track is also infinite. In practice, as can be seen from the graph, it means that above certain radii, the graph will not exhibit a minima below line speed and the maximum inward sway will occur at line speed. Trundle speed is lowest on tight curves.

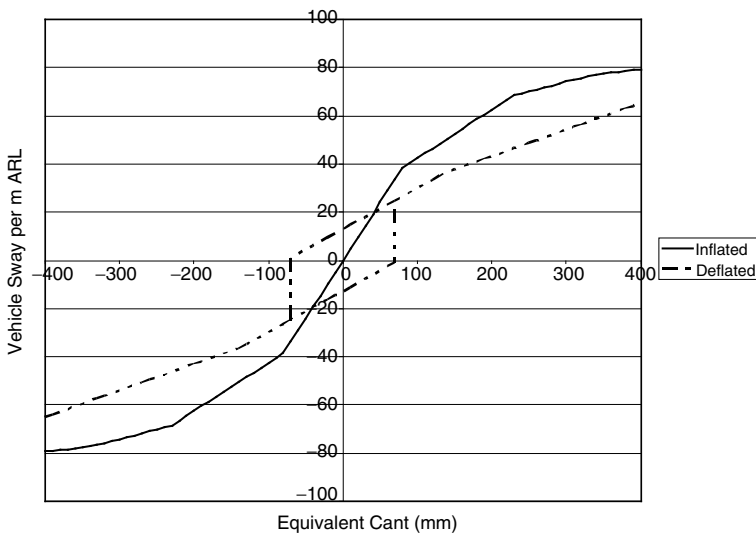
*d. Effect of Loading*

Different relationships occur for different suspension loading and failure conditions. In particular, Figure 7.10 shows relationships defined as inflated and deflated, referring to airbag condition. The possibility of airbag failure (or accidental isolation) must be considered in analysing clearances. As can be seen from the graph, deflation of airbags results in a stiffer vehicle, but which may have a “locked-in” suspension lateral movement (see also time-related effects later). At lower cants, and particularly on lower parts of the vehicle body, this lateral offset may be the worst defining movement of the vehicle. Note also that there is likely to be some hysteresis, since this locked-in movement requires a force in the opposite direction for it to break out. The value of this friction is not defined in BaSS 501 (this assumes a simple locked-in movement at zero equivalent cant) and analysis must consider that its effect can simultaneously exist to both the inside and outside of curves.

In the deflated case, the sway equation becomes:

$$\text{Sway} = K_{\text{sway}}D + E + C$$

where  $C$  is the value of locked-in suspension movement.



**FIGURE 7.10** Relationship between equivalent cant and sway for vehicles with air suspension.

Obviously, the suspension performance relates to many factors, the principal ones being the passenger (or freight) load, and the condition of the springs. The normal conditions analysed are:

- *Tare or tare inflated.* The condition where the vehicle is likely to be tallest, since the springs are least compressed.
- *Laden or crush inflated.* This condition results in the greatest sways, due to the mass of passengers or load and higher centre of gravity.
- *Crush deflated.* This condition has a larger locked-in lateral component, but smaller sways, leading to greater risk to low structures (such as platforms). It is also the lowest position in which a vehicle may operate.
- *Tare deflated.* This is an unusual condition, in that it is only likely to occur on depot routes, where there are no passengers. Its principal use is in clearing low structures where a tare inflated vehicle may not pass, frequently for delivery.

On coil sprung vehicles, failure of springs is considered unlikely, so that these conditions are not generally analysed.

The sway graphs can be calculated theoretically, either by formulae or by using VAMPIRE™ or a similar dynamics package. The results of this modelling can be confirmed by sway testing, where the vehicles are subjected to a range of cants in various conditions, and the sway and drop of different positions on the body measured.

#### e. Time Factors

Time has not generally been considered in terms of gauging, but increasingly, the effects are becoming significant in the context of accurate analysis. Two particular issues should be considered:

- Air suspension — these systems generally have self-levelling valves that compensate for vehicle asymmetry caused by loading and curving forces, tending to compensate for roll on curves. The time constants for such systems are generally long, and thus unlikely to have a noticeable effect on normal, at speed, analyses. However, where vehicles are moving very slowly, or are stationary, the compensation effect may need to be considered.
- Locked-in suspension movement — where air suspensions are run in deflated failure mode, hysteresis, as shown on the earlier suspension characteristics graph, is gradually shaken out by normal track oscillations over a short period of time. Such effects can only be considered by advanced dynamic prediction methods, as described later.

#### f. Vehicle Height

A vehicle has a nominal, static height. This would normally be the tare inflated condition. When loaded (passengers or freight) the suspension is compressed depending on the loading, resulting in a lowered static height for this condition. Loading is defined by strict rules, and there may be different operating conditions associated with different loadings. The static height is also reduced in the case of airbag deactivation or failure. Occasionally, overinflation of airbags is considered.

### 3. Vehicle Tolerances

Commonly considered tolerances are:

- Uncompensated wheel wear — this is the amount of wheel wear that can develop before being compensated by shimming of the suspension. A worn vehicle will be lower than

a new vehicle, and thus this parameter must be considered when analysing lower-body clearances (platforms, etc.).

- Suspension creep — with age, rubber suspension components compress (creep) and thus lower the body. This parameter must also be considered when assessing lower-body clearances.
- Body build tolerance (BOD) — this parameter represents tolerances in the building processes, and must be added to the static shape of the vehicle. Construction methods generally produce smaller tolerances at the vehicle solebar, and thus it is possible to have a varying BOD for different positions on the vehicle.
- Height setting tolerance — this value affects both the maximum and minimum height of the vehicle and depends on the accuracy with which the static height of the vehicle air suspension may be set.
- Air bag compensation — a self-levelling system on air suspensions means that over a period of time, the suspension gradually corrects for load imbalance. This is particularly noticeable on canted platforms, either as passengers embark or disembark, or as the cant excess is gradually compensated. Being a relatively slow process, this parameter, on its own, is unlikely to be an issue. Where a vehicle stops on a curve the effect is to reduce sway. However, there may be a tolerance of operation of the self-levelling valves, and these are occasionally considered.
- Vehicle yaw — this is a lateral movement of the end of the vehicle in relation to the centre. It is not strictly a tolerance, but is occasionally included as one. Vehicle yaw affects sections progressively the further they are from the centre of the vehicle. Usually this is only considered if their effect is of significance in relation to the clearance regime under which the vehicle is operating.
- Vehicle pitch — this is the vertical equivalent of yaw, and the same considerations apply.

In defining kinematic movements, it is usual to refer to specific points of significance on a vehicle. Typically, these would be:

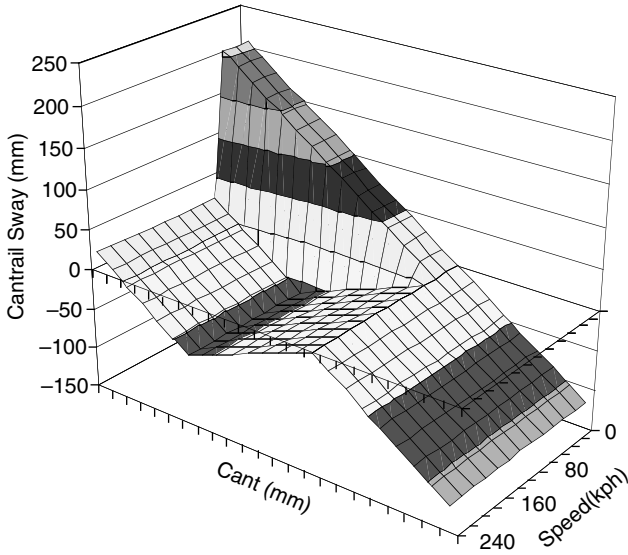
- *Cantrail*. A notional line drawn along the vehicle, which for passenger stock represents the upper limit of the body side and the start of the roof contour. This height represents a combination of semiwidth and sways likely to present the greatest risk of infringement to the arch bridges prevalent in the British infrastructure.
- *Waist*. The widest part of the vehicle (statically) which is likely to present the greatest risk of infringement to passing vehicles (with the possible exception of tilting trains).
- *Step*. Traditionally the part of the vehicle designed to come into close proximity to the infrastructure (platforms).

Using BaSS 501, the sways and drops of cantrail, waist, step and occasionally yaw damper may be calculated.

#### a. *Tilting Trains*

The basic relationships described are valid for most normal vehicles. However, tilting trains have a more complex relationship. In this case not only do the inputs from curving forces need to be considered but also the effect of the active suspension (which principally operates in roll). The relationship between tilt angle, cant deficiency, and speed varies between trains, and is nonlinear. Whereas the graphs of body sway for conventional vehicles relate simply to cant deficiency or cant excess, in the case of a tilting train the relationship is three-dimensional, having inputs of cant deficiency and speed to determine a series of sways. In general, tilting trains behave conventionally below a cut-in speed known as the tilt threshold speed or on cant excess. Vehicles





**FIGURE 7.11** Relationship between speed, cant, and sway for a tilting train at the cantrail.

with tilt locked out or in tilt failure behave as conventional vehicles, although the positional error of the body in the latter case will also need to be considered.

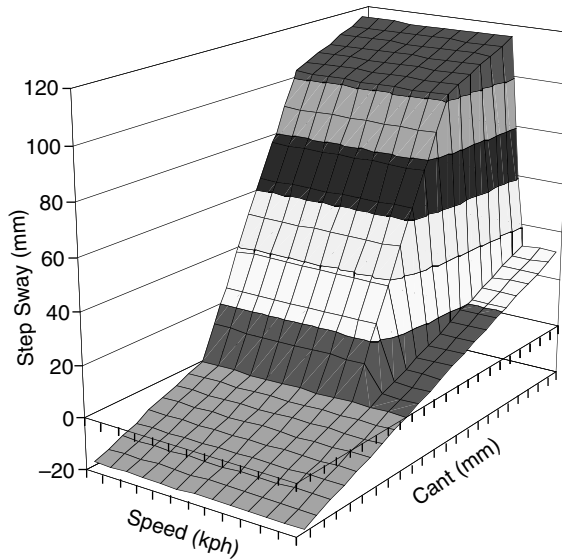
Figure 7.11 shows an idealised relationship for a tilting train suspension, showing the effects of speed and cant on the sway at the cantrail. The horizontal axis shows cant deficiency as positive, and cant excess as negative. A positive sway on the vertical axis is towards the outside of the curve. It can be seen that at low speeds ( $< 50$  km/h on this particular vehicle) and at cant excess (–ve cant), a similar relationship to that of a conventional train can be seen (i.e., a simple linear relationship, where increased cant deficiency results in a greater sway to the outside of the curve). In the tilt active area, it can be seen that the tilt system causes the vehicle to lean inwards where a normal vehicle would lean outwards (i.e., under cant deficiency). As the tilt movement is used up, the vehicle again begins to move outwards at high cant deficiencies. If the point being considered is above the tilt centre (normally the case with the cantrail), then the characteristic inward tilt can be seen.

If the point is below the tilt centre (usually the case with a vehicle step), then tilt angle is additive to the roll caused by cant deficiency. This is closer to the performance of a conventional train, but with additional roll. Vehicles whose tilt centre is high (such as Talgo, where the bodies are suspended from a high level suspension) behave similar to conventional trains, but with significantly greater sways. These vehicles sway more at the step than at the cantrail (Figure 7.12).

For tilting trains, the relative angles of primary and secondary suspensions must also be considered, since at high tilt angles the lateral force of curving may cause a significant compression of the secondary suspension, which is now not operating truly vertically. This component is known as compression drop.

The number of operational cases of tilting trains is higher than the four conventional cases, since the possibility of failure of the tilt system or other parts of the active suspension must be considered. Also, an effect known as tilt lag means that as the train travels onto and off transition curves the active suspension is slightly delayed in its response for electromechanical reasons. For some trains this can be as much as  $6^\circ$ . There is thus an entire range of situations that may occur individually or concurrently, including:

- Tare or laden (various loading factors)
- Inflated or deflated air suspension



**FIGURE 7.12** Relationship between speed, cant, and sway for a tilting train at the step.

- Active, active lagged, passive (locked), or failed tilt system
- Active or failed active suspension components

This may result in many potential situations requiring analysis.

*i. Effects of Speed*

In the analysis of tilting trains, the speed of the vehicle has been included into the basic relationship between applied force and suspension movement. However, this results in more complicated analyses than those of conventional vehicles, where sway is directly related to speed.

*ii. Critical Speeds*

For conventional vehicles, maximum speed and trundle speed are the only critical speeds that need to be considered. The nature of conventional suspension means that these cover not only worst sways to the outside and inside, respectively, but also worst drops on each side. With active suspensions, critical speeds are not as easy to calculate. Consider the graph of cantrail sway in relation to speed (Figure 7.11). If the vehicle is operating below the tilt threshold speed, then the worst sway to the outside occurs at maximum speed. If operating above tilt threshold speed, then the outward critical speed is dependent upon cant deficiency (also dependent on speed). At low cant deficiencies, the critical speed is likely to be the tilt threshold speed. However, above the limit of tilt compensation, critical speed can again be the maximum speed. There are a number of intermediate combinations of speed and cant deficiency leading to critical speeds. This is further complicated, for example:

- High-speed cases can lead to sways to the inside greater than low speed cases.
- Points above the tilt centre behave differently to points below it, where more conventional rules apply.
- Worst sway does not necessarily mean worst drop, since this depends on the tilt system geometry.

Crucial parts of the train may need to be assessed over an entire speed spectrum to ensure that all combinations of sway and drop over the speed range are considered.

BaSS 501 provided a simple methodology that allowed the sways and drops of various points of the vehicle to be calculated and related to the position of a structure or adjacent vehicle. Advanced gauging software (ClearRoute™) enables the entire vehicle shape to be modelled, providing a more complete calculation of clearance. The software enables large volumes of structure data to be analysed quickly, and is the only practical method of gauging tilting trains where a large number of load cases and speed combinations must be calculated. This system was used for the entire gauging analysis for the remodernisation of the British west coast main line (1998–2004) for the introduction of class 390 tilting trains.

iii. Time Factors

Tilting and active suspensions, by their nature, have delayed responses either by design or due to the time required to provide a measured response to inputs. The most common form of tilting suspensions measure cant deficiency and curvature on a leading bogie, and calculate the required tilt demand from this, which is applied to the leading vehicle and to trailing vehicles. In order to avoid false responses to track irregularities and ensure that there is only a response to true curving forces, a delay period (normally no more than one second) is provided, during which no tilt is applied to the lead vehicle. This is progressively less pronounced on trailing vehicles where the time lag is less. This effect is known as tilt precedence. A further effect is that the tilt system may not be able to respond at the same rate as transition curves develop. Figure 7.13 illustrates the tilt lag phenomenon.

In Figure 7.13, the horizontal axis shows the position of a train entering into a curve, which starts 100 m into the diagram. The solid line shows a linear cant transition for this curve, in degrees. On this particular curve the maximum cant is 6° (approximately 150 mm), and the transition is 100 m long. At 50 m/sec this represents a cant gradient of 75 mm/sec, typical of a tilting train at its enhanced speed. The dashed line shows the response of the tilt system. The system does not respond for the first 50 m of the curve (1 sec at 50 m/sec) and then responds at a rate of 2° per second. The dotted line shows the imbalance between tilt required and tilt achieved. A maximum tilt lag of 4° develops in this particular scenario.

Tilt lag means that in some cases the use of conventional gauging models will not provide adequate clearance assessment. In these cases it is necessary to use lead-lag models, where the kinematic envelope of the vehicle is expanded to include this error. Tilt lag refers to the error that develops as a vehicle moves onto a transition, and tilt lead (technically a misnomer) refers to the

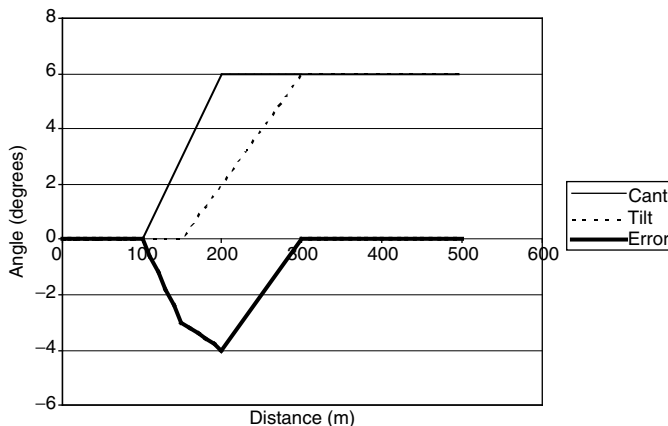


FIGURE 7.13 Tilt lag relationship.

opposite error that develops as a vehicle moves off a transition. In the latter case, it is important to note that the effect can occur on tangent track, where simple analyses would normally be performed.

### *b. Advanced Modelling*

The advent of ClearRoute™ and VAMPIRE™ computer simulation software meant that many of the simplifications inherent in BaSS 501 could be removed given that there was no longer a need to produce a hand-calculation process.

BaSS 501 has a number of inherent simplifications:

- It assumes that all lateral movement of the vehicle is from roll generated by cant forces. Pure lateral irregularities are not considered dynamically.
- It assumes a linear relationship between speed and equivalent cant leading to sway. In practice, this relationship is nonlinear as a result of harmonic responses of the spring/mass system.
- It assumes that all generated sway is upper sway, where the roll centre is low.
- The locked in movements predicted by quasistatic analysis are rapidly shaken-out by dynamic movement of the vehicle.

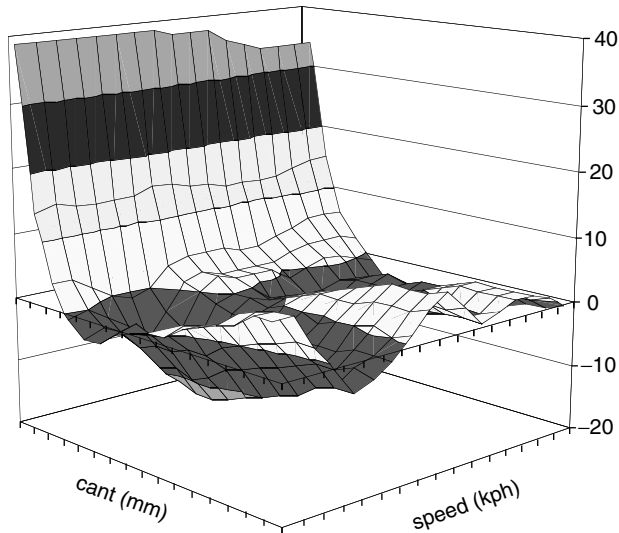
Using VAMPIRE™ it is possible to perform full dynamic simulations of the vehicle, and thus consider true vehicle behaviour within the bounds of a simulation that is now considered to be extremely accurate. The process is as follows:

- A range of real track data is assembled, according to the probable range of track roughness that will be experienced.
- The vehicle is run, in a variety of suspension conditions, over a full range of applied cants and speeds. (It is not necessary to consider radius, since the cant deficiency or excess drives the behaviour.)
- A series of lookup tables are produced defining lateral, vertical, and roll performance of the vehicle or suite of vehicles (different configurations of the same vehicle can behave differently, and there can be a different behaviour when running in different directions).
- The curving behaviour is calculated (as explained in vehicle–track interaction).

The process of defining the relationships between track inputs and vehicle body dynamic behaviour is statistically based. Track inputs are provided from a variety of typical track geometries appropriate to the speed of the vehicle. In general, lower speed track provides greater dynamic inputs to the vehicle. The resultant body movements generated at each combination of applied cant and speed are summarised statistically as a mean and standard deviation of lateral, vertical, and roll movements, onto which a 95% certainty limit is applied. The definition of appropriate track quality indices, combined with the maintenance regime of the railway are an important factor in determining the level of risk in the gauging calculations. In common with most railway administrations, British railways are seeking to provide an appropriate safety factor without considering all events to be concurrent. Techniques of uncertainty analysis are increasingly being used.

Figure 7.14 shows the sway predicted at the cantrail by the different models on the vertical axis. The horizontal axes show the inputs of cant and speed into the model. The following observations can be made:

- Sways predicted by both methods of analysis are similar, indicating a generally good correlation between the techniques.
- The nonlinearity associated with speed is clearly visible. In some cases BaSS 501 overpredicts and in others underpredicts.



**FIGURE 7.14** Comparison of cantail sways for the same vehicle modelled by BaSS 501 and VAMPIRE™.

- Generally, BaSS 501 overpredicts sways.
- The significant conservatism of BaSS 501 at low cants results from the shakeout of hysteresis due to the movement of the vehicle that is not considered in the quasistatic analysis.

It should be noted that while the above graph shows axes of speed and cant, these are not independent of each other, a factor that should be considered in understanding the effect of speed. On a given curve, increasing speed will have the effect of increasing cant deficiency, while decreasing speed will reduce this or generate cant excess. The dependency between cant and speed is a squared relationship, as described in Section II.C.2.b.

Such dynamic systems can provide vehicle movement information associated with particular track geometry, and in real time. However, real-time gauging is flawed in that it takes no account of the spectrum of track geometry that may develop as track deteriorates, or is maintained. In particular, a track defect that causes a vehicle to sway away from a structure (providing clearance) could cause a gauging infringement if removed.

### III. INTERACTION BETWEEN GAUGING COMPONENTS

#### A. VEHICLE-TRACK INTERACTION

##### 1. Wheelset Movement

The primary interface between vehicle and track occurs at the wheel-rail interface. The wheelset has freedom to move within the rails, limited by flange contact. The size of the gap depends upon the gauge of the track, the wheel flange wear, and the rail sidewear. The various gauging models handle this interface in various ways.

In simple analyses, this interface is ignored since its effect is small.

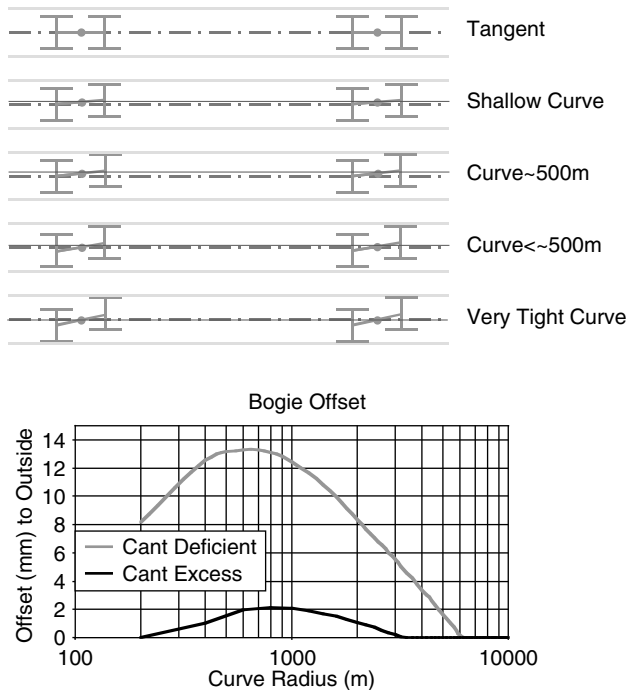
In more detailed analyses, such as BaSS501, it is assumed that all possible combinations of wide gauge, wheel flange wear, and rail sidewear occur simultaneously. Typical values would assume an 8-mm flange — rail gap, 3 mm of wheel flange wear, and 6 mm of rail sidewear.

Owing to the curving nature of bogies, this latter value is usually reduced to 3 mm, since it would be very unlikely for all wheelsets to be running to either the outside or inside rail (as demonstrated below). Thus, a global value of 14 mm is often used.

In complete analyses, the nature of curving is considered and is used to correct the centreline position of the vehicle for the curving behaviour of the bogies or wheelsets. It has been found that in cant-deficient situations (where there is insufficient installed cant to balance curving force) the bogies (and hence the vehicle) move towards the outside of the curve. Figure 7.15 shows the approximate bogie behaviour as curve radius varies. On tangent (straight) track, the wheelsets assume a mean position running centrally between the rails, and there is generally no offset, although asymmetric running on straight track has been observed on some flexible-frame bogies. As radius progressively tightens, the angle of attack increases as the leading wheelset moves towards the outside rail. The trailing wheelset continues to follow a path more centrally between the rails. At approximately 500 m, the leading wheelset will come into flange contact. This is the point of maximum outward bogie movement. As radius further decreases, the angle of attack increases further by the trailing wheelset moving towards the inside rail. An extremely sharp curve would cause the trailing wheelset to come into flange contact with the inner rail, giving a resultant zero offset. In practice, radii this severe will not be encountered. It must be emphasised that the exact relationship of offset to curve radius is complex and vehicle specific. It requires complex modelling software, such as VAMPIRE™ to generate the exact relationship.

It is necessary to consider a spectrum of operating conditions to determine the maximum and minimum wheelset movements at different radii. This will include modeling various conditions of worn wheel profile. In the gauging analysis, the maximum outward wheelset movement will be applied to cant-deficient cases, the minimum movement will be applied to cant-excess cases.

UIC rules consider the behaviour or wheel–rail interaction, and in particular, gauge widening and bogie alignment as part of the structure-vehicle relationship.



**FIGURE 7.15** Relationship between curve radius and outward wheelset movement.

## B. TRACK-STRUCTURE INTERACTION

### 1. Track Tolerances

The relationship between track position and structures is a significant factor to consider in gauging. The following components of track tolerance must be considered.

#### a. Lateral Track Positional Tolerance

Sometimes known as the “track alignment error,” this relates to the possible movement of the track over its maintenance cycle. Normal ballasted track is generally maintained by tamping, with slues being applied to correct geometric errors. In normal circumstances, track is maintained within a tolerance of  $\pm 25$  mm. Datum plates are used to provide guidance to machine operators on track position, and overhead contact wire registration (if present) normally requires the track to be maintained to this tolerance. Where the track is nonelectrified, and datum plates are not present, care must be taken to ensure (by more frequent measurement) that track remains within positional tolerance. Normal ballasted track is known as low fixity. High-fixity track, such as slabs, may be held to much tighter tolerances (even zero). Ballast glueing and strutting tracks against platforms are considered medium fixity, with a tolerance of  $\pm 15$  mm.

#### b. Vertical Track Positional Tolerance

Track level deteriorates under the effect of traffic and time. In general, the settlement of ballasted track is logarithmic in nature. Rapid settlement occurring immediately after maintenance becomes more linear as quality deteriorates towards the end of a maintenance cycle. Over a maintenance cycle, of 1 to 2 years, depending on track condition and quality, it would be expected that track would settle around 25 mm from its highest level. Unfortunately, it is difficult to apply this tolerance without knowing the position within the maintenance cycle. Recently maintained track will settle by up to 25 mm, but track which is just about to be maintained could be lifted by 25 mm. By assuming track to be at a position that could be lifted 15 mm and lowered by 10 mm provides a regime which statistically covers a large part of the maintenance cycle.

#### c. Cross-Level Error

On low-fixity track, it is generally assumed that cross level may vary by  $\pm 20$  mm in relation to that required, as a result of differential settlement or measuring errors. It is considered that half of this value ( $\pm 10$  mm) would be long wavelength and half ( $\pm 10$  mm) would be short wavelength. The long wavelength (static) component affects steady state curving forces and vehicle position. The short wavelength component affects dynamic performance of the vehicle. In applying this error, it is usual to consider the long wavelength component in relation to track fixity. High fixity (slab track) may be laid to such precision that there is no long wavelength error, and a zero value may be used. However, it is unusual to reduce the short wavelength cross-level error significantly below  $\pm 10$  mm. This latter value is usually included in the vehicle model (although it is a track parameter) and is implicitly related to  $K_{\text{speed}}$  in BaSS 501 calculations and to the track geometry files in dynamic simulations.

#### d. Sidewear

On tight curves, rail sidewear tends to occur. Its formation can be slowed by lubrication, and is generally a high-rail problem. However, it serves to widen gauge and affect the vehicle positioning on the rails. The amount of sidewear included in analyses depends on whether it can develop (unusual on straight track) and what the maintenance intervention level is: 6 to 9 mm is a normal

sidewear limit. However, as discussed earlier, the amount of movement that this can generate in the vehicle is generally less.

### C. STRUCTURE-VEHICLE INTERACTION

#### 1. Clearances

Clearance is required for a variety of reasons. Historically, clearance provided the safe boundary between vehicle and structure where there were significant unknowns in each, which has included suspension movements, tolerances, and inaccuracies in the measurement of structures. Clearance provides space to allow for aerodynamic effects and for safe walkways.

As vehicle behaviour and system tolerances are better understood, a differentiation between what is calculable and what remains unknown, is possible. Unfortunately, this has not always led to a relaxation of clearances as tolerances are extracted, which increasingly leads to conservatism — and smaller trains. Modern trains, with air suspension and about which the behaviour is well understood, tend to be smaller inside than their predecessors while occupying what appears to be a larger swept envelope.

Pressures on the infrastructure, especially in the face of an increasing need to move larger intermodal freight containers (notably 9 ft 6 in. high × 2500–2600 mm wide ISO boxes), require clearances to be specified frugally if rail is to survive in the increasingly competitive environment offered by road transport where larger paths routinely exist.

Clearance is about risk management. The larger the clearance provided, the smaller the risk, and thus the need for control measures is minimised. Modern standards specify clearance according to risk regime, where the available clearance dictates what control measures are required.<sup>7,8</sup> Typically, actual clearances greater than 100 mm are defined as normal, whereas below this reduced and special reduced clearances (the latter being clearances >0 mm) require increasingly rigorous control measures. Control measures involve processes to control track position such as slab track, glued ballast, etc. The regime of inspection is also important, ensuring that tight structures are inspected more regularly than those that are well clear of the track.

UIC rules require a reference profile to be enlarged for various effects. Clearance between the developed reference profiles is not specifically mandated.

It is important to consider risk in relation to the methodologies being used. Figure 7.16 presents a number of analytical methodologies and considers the risk associated with using them

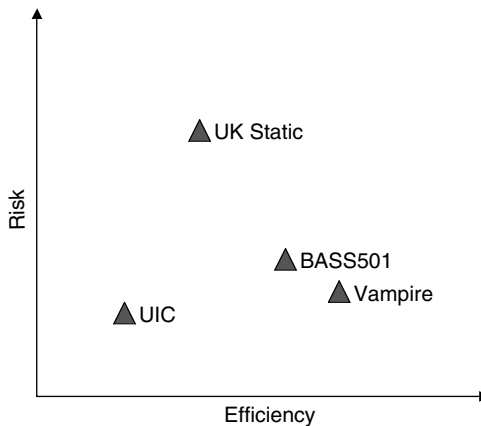


FIGURE 7.16 Risk vs. efficiency for various gauging methodologies.



(based upon the likelihood of there being situations where more clearance is required than is actually provided) and the efficiency of use of space over a typical length of route.

If we consider typical British static gauging, then a fixed clearance is provided, which provides adequate clearance in all but the worst combination of systems failure and extreme track geometry. In most situations, the vehicle is well clear of the infrastructure, and very poor use is made of the infrastructure when the vehicle is not running at these extreme limits (for instance, straight track). In contrast, using VAMPIRE™ analyses allows the actual clearance to be assessed in all situations. By ensuring that this is adequate, risk is low. Through accurate analysis, more use can be made of the infrastructure by only enlarging limiting structures, rather than increasing the gauge of a whole route.

## 2. Stepping Distances

While not strictly a clearance issue, stepping distances are an integral component of gauging analyses, and are generally the most difficult to resolve. To provide an adequate clearance to a moving vehicle, while still providing safe passenger access and egress to stationary vehicles, involves considering opposite, worst case scenarios.

Clearance analysis involves calculating the tightest reasonable clearance that may develop. Stepping calculation involves determining the maximum distance between a platform edge and vehicle step that may develop. In the latter case, it is customary to consider the static (thrown) position of a vehicle in relation to the platform edge. In Britain, the Health and Safety Executive's Railway Inspectorate<sup>2</sup> require maximum stepping distances of:

- Lateral: 275 mm
- Vertical: 250 mm
- Diagonal: 350 mm

This is known as the “stepping triangle,” although it does not conform to Pythagoras' rule.

The values required are theoretical, taking no account of many tolerances that affect the actual stepping distance. In particular, air suspension system performance (self-levelling valves), installed cant, and track tolerances can have a significant effect on the stepping distances measured in relation to those calculated. However, the HMRI *Guidance* values do provide a sensible benchmark for static values.

Improvement of stepping distances is likely to be a characteristic of increasing disability regulation.

## REFERENCES

1. Gauging Policy Consultation — Strategic Rail Authority, 2004.
2. Railway Safety Principles and Guidance — Her Majesty's Railway Inspectorate, Health and Safety Executive, 1996.
3. Kinematic Envelope and Curve Overthrow Calculations (Design Guide 501) — Bogie and Suspension Section (BaSS), Department of Mechanical and Electrical Engineering, British Railways Board, 1985.
4. UIC Code 505-1. *Rolling Stock Construction Gauge* — UIC, 8th ed., 1997.
5. UIC Code 505-4. *Effects of the Application of the Kinematic Gauges Defined in the 505 Series of Leaflets on the Positioning of Structures in Relation to the Tracks and of the Tracks in Relation to Each Other* — UIC, 3rd ed., 1977.
6. UIC Code 505-5. *Basic Conditions to Common Leaflets 505-1 to 505-4* — UIC, 2nd ed., 1977.
7. Railway Group Standard GC/RT5212, Issue 1 — Rail Safety and Standards Board, 2003.

8. Railway Group Standard GM/RT2149, Issue 3 — Rail Safety and Standards Board, 2003.
9. General Gauging Procedure for Class 373 (Eurostar) Trains — Mechanical Systems Engineer Report No. 42, European Passenger Services, British Rail, 1995.
10. Guidance on Gauging GE/GN8573, Issue 1 — Rail Safety and Standards Board, 2004.
11. Johnson, D. M., *Square Boxes into Round Holes*, Paper to the Railway Civil Engineers Association, Institution of Civil Engineers, September, 2003.
12. Johnson, D. M., *Technology for Structure Gauging*, Paper to the Rail Track Association of Australia, Darwin, June 2004.

---

# 8 Railway Vehicle Derailment and Prevention

*Huimin Wu and Nicholas Wilson*

## CONTENTS

I.	Introduction .....	210
II.	History and Statistics .....	210
III.	Railway Vehicle Derailment Mechanisms and Safety Criteria .....	212
A.	Flange Climb Derailment.....	212
1.	Wheel Climb Process .....	212
2.	Flange Climb Safety Criteria .....	214
a.	Nadal Single-Wheel L/V Limit Criterion .....	214
b.	Weinstock Criterion .....	216
c.	Duration Based Criterion .....	217
d.	AAR Wheel Climb Duration Limit (U.S.) .....	218
e.	FRA Wheel Climb Distance Limit (U.S.) .....	218
f.	Proposed TTCI Wheel Climb Distance Criterion .....	218
B.	Application of Flange Climb Derailment Criteria .....	220
1.	Flange Climb due to Low Flange Angle .....	220
2.	Increase of Flange Length Can Increase Flange Climb Distance Limit.....	221
3.	Flange Climb due to High Coefficient of Friction at Wheel–Rail Interface.....	221
4.	Flange Climb of Independently Rotating Wheels .....	221
C.	Derailments Caused by Gauge Widening and Rail Rollover .....	223
1.	The AAR Chapter XI Rail Roll Criterion .....	224
2.	The Gauge Widening Criterion.....	225
3.	Effect of Hollow-Worn Wheels on Gauge Widening and Rail Roll Derailment .....	226
D.	Derailment Caused by Track Panel Shift .....	228
1.	Causes of Track Panel Shift.....	229
2.	Panel Shift Criterion.....	229
E.	Derailment Caused by Vehicle Lateral Instability .....	230
IV.	Prediction of Derailment.....	231
A.	Assessment of Wheel/Rail Parameters .....	231
B.	Dynamic Simulation of Vehicle–Track Interaction .....	232
C.	Track Tests .....	232
V.	Prevention of Derailment.....	233
A.	Wheel/Rail Profiles .....	233
1.	Addressing Wheel Flange Angle .....	233
2.	Removing Hollow-Worn Wheels.....	233
B.	Independently Rotating Wheels.....	234

C. Installation of Guard Rail or Restraining Rail on Sharp Curves.....	234
D. Optimising Bogie Suspension and Bogie Yaw Stiffness .....	234
E. Lubrication .....	235
F. Track Geometry Inspection and Maintenance.....	235
G. System Monitoring.....	235
References.....	235

## I. INTRODUCTION

Railway vehicle derailment can cause significant casualties and property loss. Avoidance of derailment is vital to railways for both safety and economic reasons.

Railway vehicle derailments are the results of wheels running off the rails which provide the support and guidance. The reason for wheels running off rails can be very complicated. However, the final scenario of derailment can result in wheels climbing off the rail, rail gauge widening, or rail rollover that causes wheels to fall between rails. Therefore, any conditions that may reduce the lateral guidance provided by rail can increase the risk of derailment.

Note that the derailments discussed in this chapter relate only to the cause of losing lateral constraint at wheel and rail interface. Derailments due to other causes, such as component failure, are not covered in this chapter.

## II. HISTORY AND STATISTICS

Derailment has always been one of the major concerns for railway operations since the first day of wheels running on rails. The essential feature of wheels running on rails creates a unique challenge for railways to ensure that wheels stay on the rail. The high speed operation developed in the past century demands a more strict control of vehicle lateral guidance.

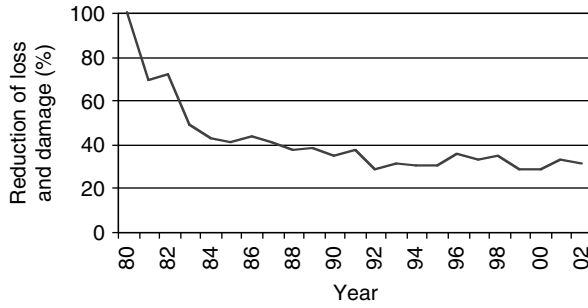
Railway technologies have advanced significantly in recent years and safety levels are high compared with the early days and also compared with other transport modes. Derailments however, unfortunately, still frequently occur. Typing “railway derailment” into an internet search, produces pages of derailment incidences around the world, due to different kinds of causes. A very old report can be found that describes the derailment on November 10, 1881 at Carnforth, U.K.<sup>1</sup> This derailment resulted in four passengers being injured, three carriages damaged, and some track damage, and was concluded to be due to the signalman having shifted the facing-points before the entire train had passed over them.

Recently, nearly 40 people were killed when a packed express train crashed in northwestern Turkey on its way from Istanbul to Ankara. It was not immediately clear what had caused the train to derail, but survivors described feeling the carriage violently shaking before the crash.<sup>2</sup>

Railroads in the United States began reporting accidents to the federal government following the passage of the Report Act of 1910.<sup>3</sup> A database of accident/incident has been established since then. Derived from this database, [Figure 8.1](#) displays the North American Freight Railways safety trends between 1980 and 2002.<sup>4</sup> Note that the values for 1980 are taken as a reference defined as 100%. Compared with 1980, a significant reduction in loss and damage has been achieved. However, the trend has been stable since 1992.

According to the FRA accident/incident database, the leading train accident causes are<sup>5</sup>:

- Rail, joint bar, and anchoring
- Track geometry defect
- General switching rules
- Wheels
- Axles and journal bearings



**FIGURE 8.1** Loss and damage trends between 1980 and 2000. U.S. (FRA, Railroad Safety Statistics Annual Report 2002).

- Switches
- Frogs, switches and track appliances
- Bogie components
- Train handling/train makeup
- Highway rail grading

In a study of derailments caused by hollow wheels,<sup>6</sup> this database revealed that 8862 reportable derailment incidents occurred between 1998 and 2000. From a review of over 300 derailment incidence cause codes defined by the Federal Railroad Administration (FRA), 53 cause codes were identified as being likely to be influenced by poor wheel–rail interactions.

By further searching the derailment database, 1796 derailments were found relevant to these 53 cause codes between 1998 and 2000. Table 8.1 shows the distribution of the 1796 incidents.

Wide gauge, track alignment, bogie hunting, and wheels with worn tread and flanges were given special attention as they relate directly to poor wheel–rail interaction and are likely to be influenced by hollow-worn wheels. As indicated in Table 8.1, these four cause types are responsible for about 50% of derailments related to the 53 incidence cause codes. Noticeably, wide gauge was reported as the cause of approximately 8% of total reported derailments and approximately 40% of derailments related to the 53 selected cause codes.

Among a large number of derailments, many were caused by mechanical failures of either tracks or vehicles. Some of these failures might be caused by poor designs and others were possibly the accumulative results of undesired performance or poor maintenance.

**TABLE 8.1**  
**Distribution of the Researched Incidents**

Year	Total Reportable Incidents	Incidents (53 codes)	Wide Gauge	Track Alignment	Hunting	Worn Tread and Flange	Others
2000	3193	673	295	46	7	13	312
1999	2924	612	234	52	6	18	302
1998	2745	511	189	54	3	6	259
Sum	8872	1796	718	152	16	37	873

### III. RAILWAY VEHICLE DERAILMENT MECHANISMS AND SAFETY CRITERIA

Railway derailments due to loss of the lateral guidance at the wheel and rail interface may be classified into four major causes: wheel flange climb, gauge widening, rail rollover, and track panel shift, based on the ways that wheel–rail lateral constraints are lost.

A derailment review by Blader discussed the mechanism of these types of derailment and some related test methods.<sup>7</sup>

#### A. FLANGE CLIMB DERAILMENT

Wheel flange climb derailments are caused by wheels climbing onto the top of the railhead then further running over the rail. Wheel climb derailments generally occur in situations where the wheel experiences a high lateral force combined with circumstances where the vertical force is reduced on the flanging wheel. The high lateral force is usually induced by a large wheelset angle-of-attack. The vertical force on the flanging wheel can be reduced significantly on bogies having poor vertical wheel load equalisation, such as when negotiating rough track, large track twist, or when the car is experiencing roll resonances. The forces between the wheel and the rail are explained in more detail in Chapter 4.

Flange climb derailments generally occur on curves. The wheels on the outer rail usually experience a base level of lateral force to vertical force ratio ( $L/V$ ) that is mainly related to:

- Curve radius
- Wheel–rail profiles
- Bogie suspension characteristics
- Vehicle speed

These factors combine to generate a base wheelset angle of attack, which in turn generates the base level of lateral curving force.

A significantly misaligned bogie is likely to induce higher wheelset angle of attack. Furthermore, any track irregularities and dynamic discontinuities may lead to an additional increase of the wheel  $L/V$  ratio. When this ratio exceeds the limit that the wheel can sustain, flange climb occurs.

Wheel climb derailments can also occur on tangent track when track irregularities and vehicle lateral dynamic motion are severe, such as during vehicle hunting and aggressive braking.

#### 1. Wheel Climb Process

The lateral velocity of a wheel due to its rotational velocity is given by

$$V_t = -\omega r \sin(\psi) \quad (8.1)$$

where  $V_t$  is the lateral velocity of a wheelset,  $r$  is rolling radius, and  $\psi$  is wheelset angle of attack.

Figure 8.2 shows a plan view of a wheelset with a yaw angle relative to the track. This angle, which is known more commonly as the angle of attack, contributes to the lateral creepage through a component of the wheelset's rotational velocity.

If the wheelset has a lateral velocity in addition to the component of lateral velocity due to its rotation, the net lateral velocity of the wheelset at the contact zone, assuming the angle of attack to be small, ( $\psi = \sin \psi$ ), is given by,

$$V_y = \dot{y} - \omega r \psi \quad (8.2)$$

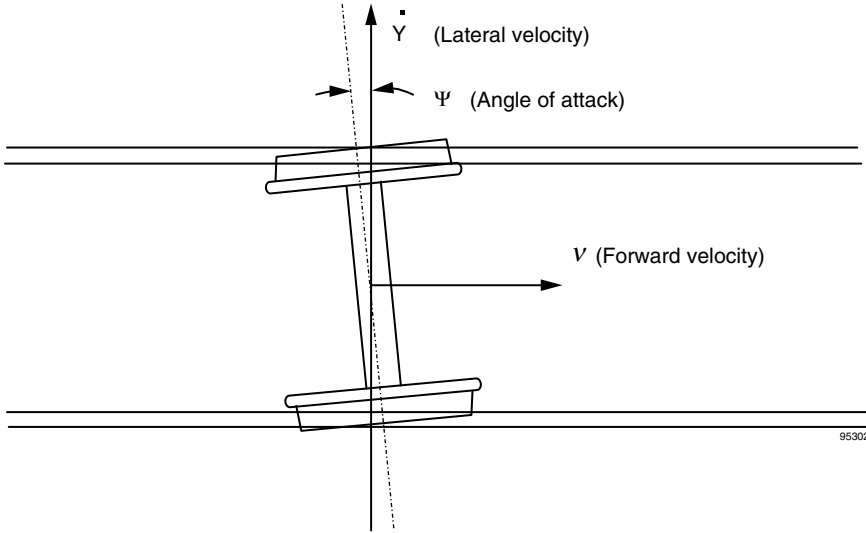


FIGURE 8.2 Wheelset angle of attack.

The lateral creepage is defined as the wheel–rail relative lateral velocity divided by the forward velocity.

$$\gamma_y = \left( \psi - \frac{\dot{y}}{V} \right) \sec(\delta) \tag{8.3}$$

The term  $(\psi - \dot{y}/V)$  is commonly known as the effective angle of attack and is a function of the wheelset lateral velocity. It is clear that if the wheelset is moving towards flange contact with a positive angle of attack, the lateral velocity tends to reduce the effective angle of attack.

Since the term  $\sec(\delta)$  always has a positive value during flange climb, the direction of the lateral creepage is dependent on the sign of the term  $(\psi - \dot{y}/V)$ . The lateral creepage equals zero when  $\psi$  equals  $\dot{y}/V$ . The lateral creepage changes direction when  $\psi < \dot{y}/V$ . The spin creepage also affects the lateral creep force. The direction of the lateral creep force depends on the resultant of the contribution of both the lateral and spin creepages.

The process of the wheel flange climbing up the gauge face and onto the head of the rail may be illustrated in three phases, as Figure 8.3 shows. A single point of contact is assumed in this description.

In phase 1, under the influence of a lateral force, the wheel moves to right towards flange contact. This produces a lateral creep force, acting on the wheel, which is opposing flange climb. In phase 2, as the flange contact angle increases, the wheelset lateral velocity decreases. As a result,

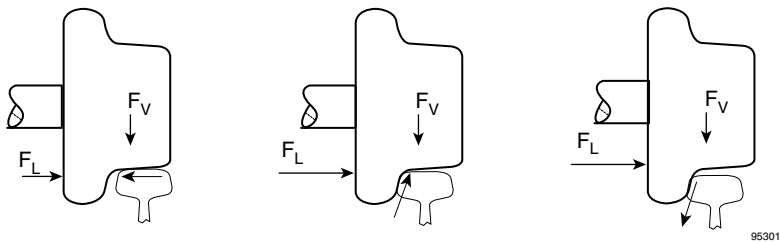


FIGURE 8.3 Process of derailment.

the lateral creepage and creep force reverse direction due to the change in sign of the effective angle of attack. During this phase, the lateral creep force is assisting the wheel to climb. After the maximum contact angle has been passed, the wheelset lateral velocity increases and the wheelset lateral displacement increases rapidly. As a result of the changing wheelset lateral velocity, the effective angle of attack approaches zero and then changes sign. Consequently, the lateral creepage and creep force also reverse direction and, once again, the lateral creep force opposes the climbing motion of the wheel, as shown in phase 3.

## 2. Flange Climb Safety Criteria

Wheel flange climb derailment phenomena have been investigated for more than 100 years. Several flange climb safety criteria have been proposed. These criteria have been used by railway engineers (globally or locally) as guidelines for safety certification testing of railway vehicles. The following are examples of published criteria:

1. Nadal single-wheel  $L/V$  limit criterion
2. Weinstock axle-sum  $L/V$  limit criterion
3. High speed passenger distance limit (5 ft) — FRA, U.S.
4. CHXI 50-millisecond time limit — Association of American Railroads, U.S.
5.  $L/V$  time duration criterion — proposed by Japanese National Railways (JNR)
6.  $L/V$  time duration criterion — proposed by Electromotive Division of General Motors (EMD)
7. Wheel climb distance criterion — proposed by Transportation Technology Centre, Inc. (TTCI)

Criteria 1 and 2 are related to the  $L/V$  ratio limits. The rest are related to the time or distance limits, which are applied to limit the exceeding duration of the  $L/V$  ratio limit, in either time or distance scale. The wheel climb would be very likely to occur as both the  $L/V$  ratio criterion and duration limit are exceeded. There follows a brief description for each criterion listed above.

### a. Nadal Single-Wheel $L/V$ Limit Criterion

The Nadal single-wheel  $L/V$  limit criterion, proposed by Nadal for the French railways, has been used throughout the railway community.<sup>8</sup> Nadal established the original formulation for limiting the  $L/V$  ratio in order to minimise the risk of derailment. Nadal assumed that the wheel was initially in two-point contact with the flange point leading the tread point. He concluded that the wheel material at flange contact point was moving downwards relative to the rail material, due to the wheel rolling about the tread contact. He further theorised that wheel climb occurs when the downward motion ceases with the friction saturated at the contact point. Based on Nadal's assumption and a simple equilibrium of the forces between a wheel and rail at the single point of flange contact, as illustrated in [Figure 8.4](#), Equation 8.4 can be derived.

$$F_3 = V \cos \delta + L \sin \delta = V \left( \cos \delta + \frac{L}{V} \sin \delta \right)$$

$$\left\{ \begin{array}{ll} F_2 = V \sin \delta - L \cos \delta = V \left( \sin \delta - \frac{L}{V} \cos \delta \right) & \text{when } (V \sin \delta - L \cos \delta) < \mu \times F_3 \\ F_2 = \mu \times F_3 & \text{when } (V \sin \delta - L \cos \delta) \geq \mu \times F_3 \end{array} \right\} \quad (8.4)$$



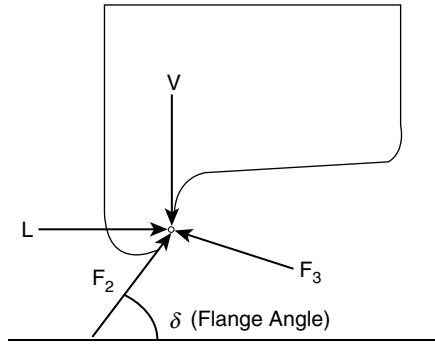


FIGURE 8.4 Forces at flange contact location.

From Equation 8.4, the L/V ratio can be expressed as

$$\frac{L}{V} = \frac{\tan\delta - \frac{F_2}{F_3}}{1 + \frac{F_2}{F_3} \tan\delta} \tag{8.5}$$

Nadal’s famous L/V ratio limiting criterion, given by Equation 8.5, was proposed for the saturated condition  $F_2/F_3 = \mu$ .

$$\frac{L}{V} = \frac{\tan\delta - \mu}{1 + \mu \tan\delta} \tag{8.6}$$

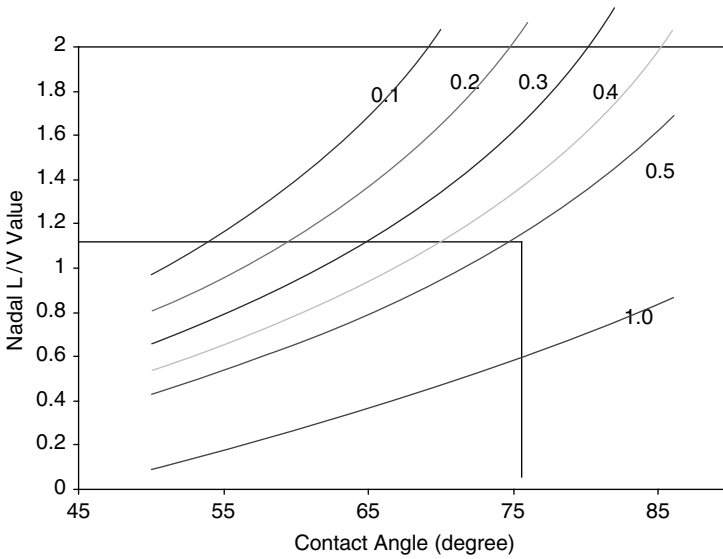
If the maximum contact angle is used, this equation gives the minimum wheel L/V ratio at which flange climb derailment may occur, for the given contact angle and friction coefficient  $\mu$ . In other words, below this L/V value, flange climb cannot occur.

Figure 8.5 plots Equation 8.6 for the coefficient of friction range between 0.1 and 1.0.

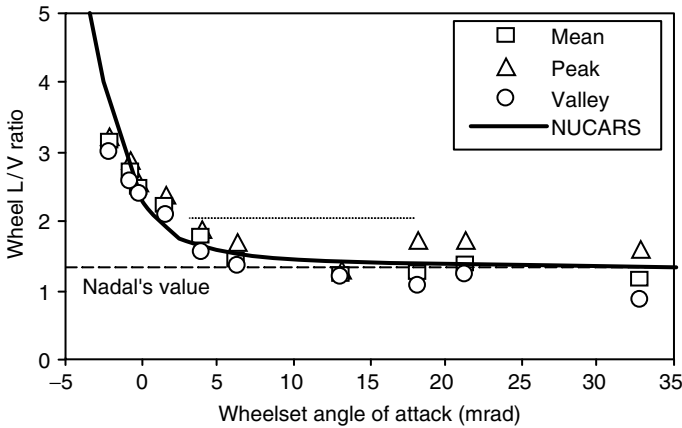
It indicates that the larger the maximum contact angle, the higher the L/V ratio limit required for flange climb. Figure 8.5 also indicates for the same contact angle, the lower of friction coefficient, the higher the L/V ratio limit required for flange climb.

To explain the effect of wheelset angle of attack (defined in Figure 8.2) on wheel L/V ratio limit, Figure 8.6 displays an example of a single-axle wheel climbing. These are the results from simulations using NUCARS™ (TTCI’s rail vehicle dynamic simulation software) and flange climb tests conducted using the Association of American Railroads Track-Loading Vehicle.<sup>9</sup> The wheels used in this example have a flange angle of 75°. In this example, wheel climb will not occur for an L/V ratio level below the solid line for a specified angle of attack. Figure 8.6 also indicates that for large wheelset angles of attack (about 10 mrad in Figure 8.6), derailments occurred at Nadal’s value. However, for smaller and negative angles of attack, the L/V ratio required for derailment increased considerably.

In summary, Nadal’s criterion agrees with situations when a large angle of attack is experienced, and is conservative for small angles of attack. It does not consider the effects of friction coefficient of the nonflanging wheel on the flanging wheel climbing, which will be discussed in a later section. It assumes flange-climbing derailment is instantaneous once the L/V limit has been exceeded. Both field tests and simulations have proved that wheel flange climb derailments would only occur when the L/V ratio limit has been exceeded for a certain distance limit or time duration limit.



**FIGURE 8.5** Relationship of limiting wheel L/V ratio and maximum wheel–rail contact angle. The friction coefficients showed a range of 0.1 to 1.0.



**FIGURE 8.6** Effect of wheelset angle of attack on wheel L/V ratio limit.

*b. Weinstock Criterion*

In 1984, Weinstock proposed a less conservative wheel flange climb criterion.<sup>10</sup> This criterion predicts incipient derailment by summing the absolute values of L/V on the two wheels on the same axle, known as the “axle sum L/V” ratio. The L/V ratio on the flanging wheel is evaluated using the Nadal equation, and the L/V ratio on the nonflanging wheel is generally used to represent the friction coefficient at wheel–rail interface since the nonflanging wheel contacts the rail at wheel tread-rail crown region with a low contact angle (0 to 3° in general).

Weinstock suggested that sum of Nadal value on the flanging wheel and the coefficient of friction on the non-flanging wheel might provide a more accurate criterion than Nadal criterion, especially at small or negative angle-of-attack because Nadal criterion only considered the flange wheel L/V ratio along.

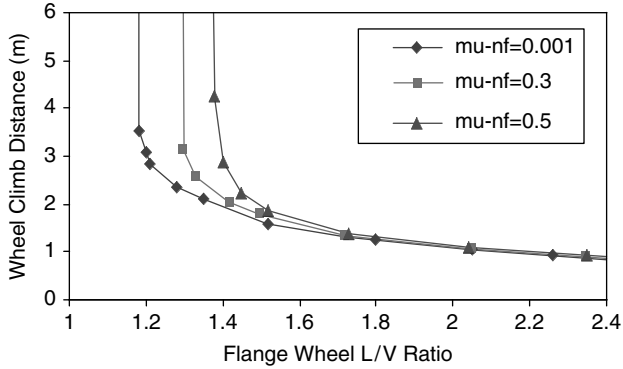


FIGURE 8.7 Effect of nonflanging wheel friction coefficient (5 mrad wheelset angle of attack, 75° flange angle).

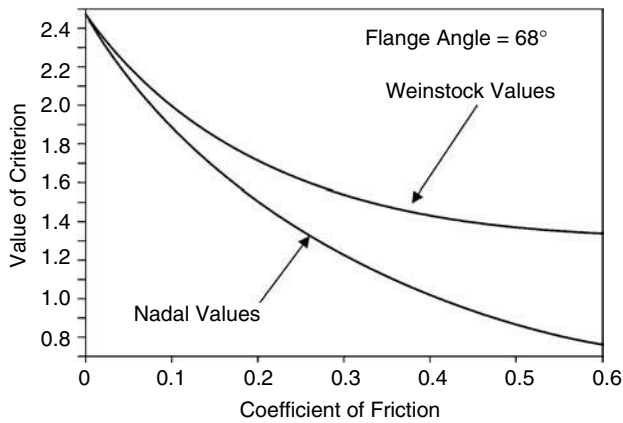


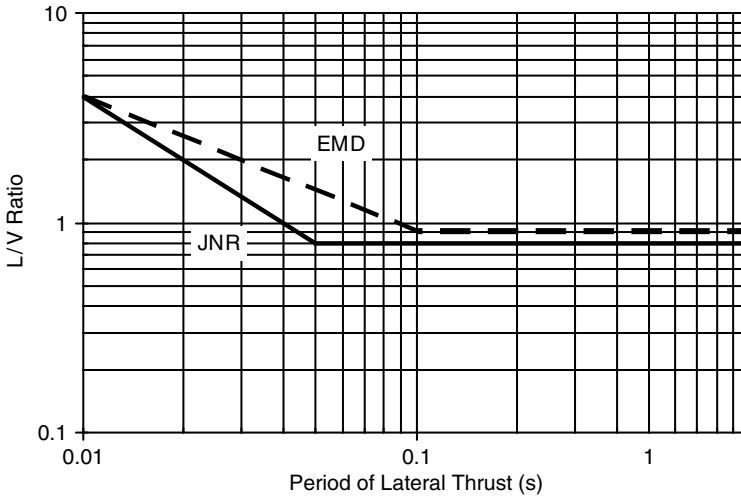
FIGURE 8.8 Comparison of Nadal and Weinstock criteria.<sup>7</sup>

Note that if the coefficient of friction on the nonflange wheel approaches zero, the Weinstock criterion converges with that of Nadal since there is no contribution of L/V value from the nonflange wheel. A flange climb derailment study conducted by Wu and Elkins<sup>11,12</sup> showed that the L/V ratio limit for the flanging wheel increases with the increase of friction coefficient on the nonflange wheel (Figure 8.7). In Figure 8.7, wheel climb will not occur for an L/V ratio level less than the asymptotic line for each friction coefficient level on the nonflanging wheel.

The Weinstock criterion retains the advantage of simplicity. It can be measured with an instrumented wheelset, which measures the values of L/V ratio on both wheels on an axle. It is not only more accurate than Nadal’s criterion, it also has the merit of being less sensitive to errors or variations in the coefficient of friction. Figure 8.8 compares the Nadal and Weinstock criteria variation with the coefficient of friction at the wheel–rail interface.

c. Duration Based Criterion

While investigating the duration of the single wheel L/V criterion necessary for derailment, researchers at the JNR proposed a modification to Nadal’s criterion.<sup>13</sup> They suggested that for a duration of lateral thrust (lateral force impulse) less than 50 msec, such as might be expected during



**FIGURE 8.9** JNR and EMD flange climb duration criteria.

flange impacts while hunting, the allowable value of the L/V criterion should be increased, as shown in Figure 8.9. The analytical expression for the JNR criterion is given as Equation 8.7. An even less conservative approach was proposed by the Electromotive Division of General Motors (EMD).<sup>14</sup> Its L/V criterion is also shown in Figure 8.9.

$$\frac{L}{V} = \pi \left\{ \frac{i_B}{G} \right\} \left\{ \frac{\tan(\delta) - \mu}{1 + \mu \tan(\delta)} \right\} \sqrt{\frac{hP_w}{gP}} \left\{ \frac{1}{T} \right\} \quad (8.7)$$

*d. AAR Wheel Climb Duration Limit (U.S.)*

Based on the JNR and EMD research, and considerable experience in on-track testing of freight vehicles, a 0.05 second (50 msec) time duration was adopted by the AAR for the Chapter XI certification testing of new freight vehicles.<sup>15</sup> The Chapter XI criterion states that: “The individual wheel L/V should not exceed 1.0 on any wheels measured. The instantaneous sum of absolute wheel L/V’s on any axle shall not exceed 1.5.” “(Those values) not to exceed indicated value for a period greater than 50 msec per exceedence.”

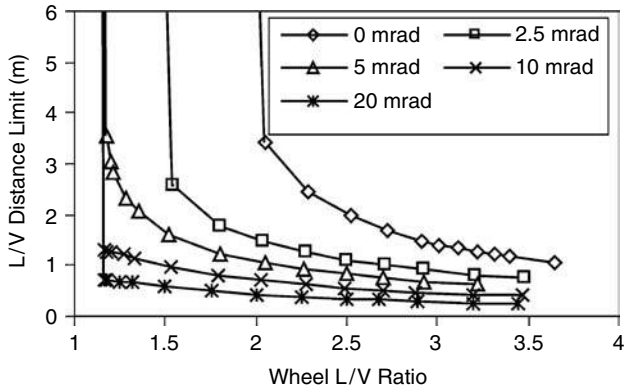
This time duration has since been widely adopted by test engineers throughout North America for both freight and passenger vehicles that have adopted wheel flange angle of 75°.

*e. FRA Wheel Climb Distance Limit (U.S.)*

A flange climb distance limit of 5 ft was adopted by the FRA, U.S. for the class 6 (and higher) high speed track standards.<sup>16</sup> This distance limit appears to have been based partly on the results of the joint AAR/FRA flange climb research conducted by TTCI<sup>11</sup> and also on experience gained during the testing of various commuter rail and passenger vehicles.

*f. Proposed TTCI Wheel Climb Distance Criterion*

The TTCI flange climb criterion was developed for North American freight cars using AAR1B wheel profile with a 75° flange angle at speeds below 80 km/h in curving.<sup>11,12</sup> This criterion encompasses two limits, the single wheel L/V limit and the L/V distance limit. The distance limit is



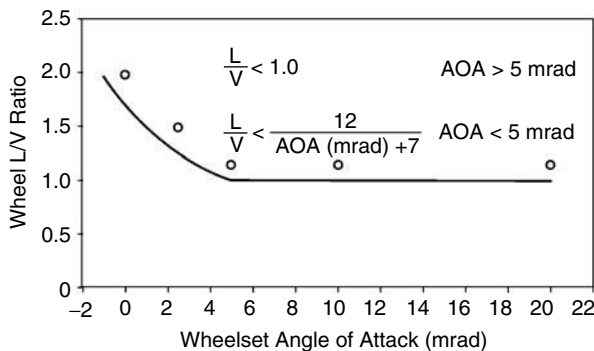
**FIGURE 8.10** Effect of wheelset angle of attack on L/V distance limit,  $\mu = 0.5$ , no longitudinal creepage.

the maximum distance that the single wheel L/V limit can be exceeded without risk of flange climb derailment. It is possibly the first time that the wheelset angle of attack has explicitly been included in the flange climb criterion. Figure 8.10 shows the simulation results of the L/V distance limits under different wheelset angle of attack. The test and simulation results showed that the distance limit is a function of wheelset angle of attack.

The following are the proposed criteria. Since measurements of angle of attack are usually difficult in track test, the criterion is given in two forms: one for the use of simulations in terms of wheelset angle of attack, and one for the use of track test in terms of track curvature (in degrees). Figure 8.11 and Figure 8.12 graphically display the criteria as a function of angle of attack.

Note that both the wheel L/V ratio limit and distance limit will converge to a constant value as the wheelset angle of attack reaches a certain level.

TTCI is currently conducting research to update the proposed criterion and further develop flange climb derailment L/V ratio and distance criteria for application to the North American Freight Railroads and North American Transit Operation. The new criteria will have a more general form for application to the variety of wheel profile designs used by different freight vehicles and transit systems.<sup>17,18</sup> Specific criteria will be specified for the AAR Chapter XI acceptance tests, and for selected transit systems.



**FIGURE 8.11** Proposed single wheel L/V criterion as a function of angle of attack.

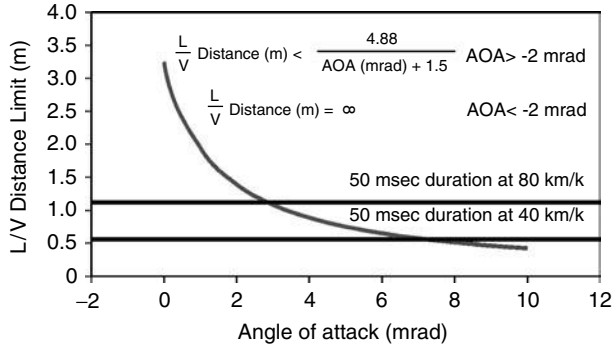


FIGURE 8.12 Proposed L/V distance limit as a function of angle of attack.

### B. APPLICATION OF FLANGE CLIMB DERAILMENT CRITERIA

#### 1. Flange Climb due to Low Flange Angle

Figure 8.13 provides two examples of wheel flange angles. One is a wheel profile with a flange angle of 75° and the other has a 63° flange angle. Referring to Figure 8.5, at a friction coefficient of 0.5, which represents the dry wheel–rail contact condition, the limiting L/V value is 1.13 for wheels with a 75° flange angle, according to the Nadal criterion, and 0.73 for wheels with a 63° flange angle. Clearly, wheels with low flange angles have a higher risk of flange climb derailment.

For historic reasons, some railway systems have adopted relatively low wheel flange angles in the range of 63 to 65°. New systems now generally start with a wheel profile having a flange angle of 72 to 75°.

A wheel profile with a higher flange angle can reduce the risk of flange climb derailment and can have much better compatibility with any new designs of vehicle/bogie that may be introduced compared to wheels with lower flange angles. Also, with a higher L/V ratio limit, high flange angles will tolerate greater levels of unexpected track irregularity.

In the *Track Design Hand Book for Light Rail Transit*,<sup>19</sup> a wheel flange angle of 70° was proposed, based on Professor Heumann’s design. *APTA Passenger Rail Safety Standard Task Force Technical Bulletin*<sup>20</sup> provided guidance on reducing the probability of wheel climb derailment by suggesting a minimum wheel flange angle of 72° (suggested tolerances are + 3 and – 2°).

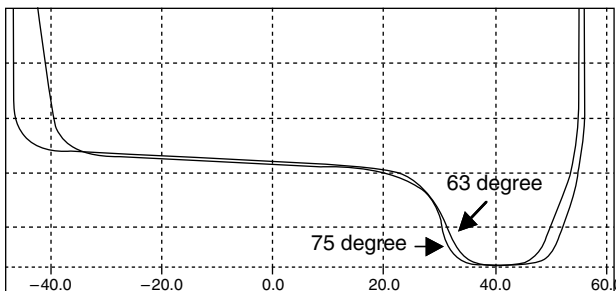


FIGURE 8.13 Wheels with 75 and 63° flange angle.

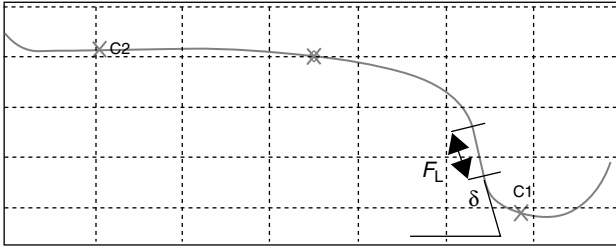


FIGURE 8.14 Definition of flange length.

## 2. Increase of Flange Length Can Increase Flange Climb Distance Limit

The flange length is defined as a measure of wheel flange contour with an angle above a given degree, explained in Figure 8.14. The flange angle may not be a constant in the length  $F_L$ , but must be above a specified value.

A concept of increasing flange length to increase the flange climb distance limit was proposed by Wu and Elkins<sup>11</sup> and further validated by Wilson et al.<sup>21</sup> They concluded that increasing flange length would increase flange climb distance appreciably at a lower angle of attack (approximately 5 mrad), and produce only a small increase in climb distance at a higher angle of attack.

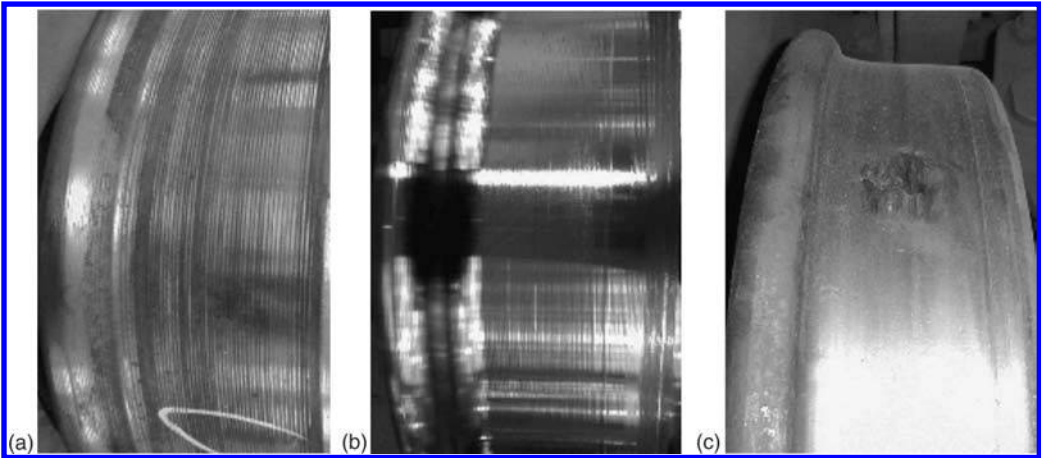
## 3. Flange Climb due to High Coefficient of Friction at Wheel–Rail Interface

Flange climb derailments have been reported to occur at curves or switches in maintenance yards when the cars were just out of the wheel truing machines. This type of derailment is probably caused by the wheel surface roughness after wheel truing. Figure 8.15 compares a wheel surface just after truing and the surface after many miles of running. The left wheel in Figure 8.15 was trued by the milling type machine with very clear cutting traces on the surface; the middle one was trued by the lathe type machine with shallower cutting traces; and the right wheel was back from operation with a smooth surface but a flat spot on the tread.

Generally, the coefficient of friction for dry and smooth steel-to-steel contact is about 0.5. The effective friction coefficient for the rough surface could be much higher. For example, if the coefficient reaches 1.0, the  $L/V$  limit, as shown in Figure 8.5, would be 0.5 for a  $75^\circ$  flange angle and 0.3 for a  $63^\circ$  flange angle. Therefore, the rough surface produced by wheel truing could significantly reduce the  $L/V$  limit for flange climb. A low flange angle would further increase the derailment risk. Addressing the final surface tuning and wheel–rail lubrication after reprofiling are two possible remedies to improve the surface condition after reprofiling.

## 4. Flange Climb of Independently Rotating Wheels

Wheels mounted on a solid axle must rotate at the same speed. To accommodate running in curves, a taper is usually provided on the wheel tread. The wheelset shifts sideways, as shown in Figure 8.16, to allow the outer wheel to run with a larger rolling radius than the inner wheel. The resulting longitudinal creep forces at the wheel–rail interfaces on wheels of the same axle form a moment that steers the bogie around curves (Figure 8.16). Previous flange climb studies have indicated that as the ratio of longitudinal force to vertical force increases, the wheel  $L/V$  ratio required for derailment also increases (Figure 8.17). Therefore, the Nadal flange climb criterion can be relaxed based on the level of longitudinal force. The flange climb would occur at an  $L/V$  ratio above the Nadal limiting value in the presence of longitudinal force.

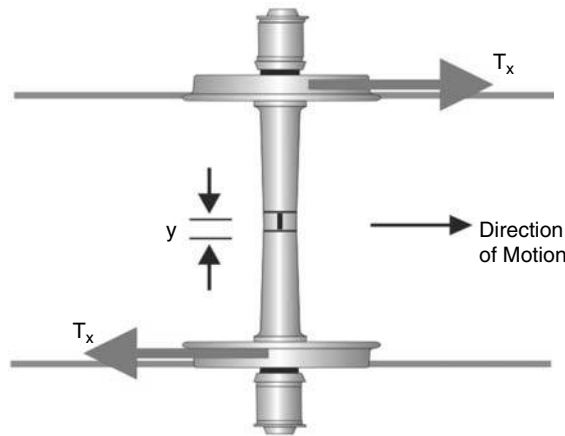


**FIGURE 8.15** Comparison of wheel surface roughness. (a) Surface after wheel truing from milling type machine, (b) surface after wheel truing from lathe type machine, (c) surface of wheel back from operation with a flat spot.

In simple terms, the longitudinal steering forces can be viewed as “using up” some of the available wheel–rail friction. This reduces the effective friction coefficient for flange climbing, increasing the  $L/V$  ratio required for flange climb.

Independently rotating wheels can rotate at different speeds and therefore produce no longitudinal forces to form a steering moment. This can lead to higher wheelset angles of attack, consequently higher lateral forces (before reaching to the saturation), higher  $L/V$  ratios, and increased wheel and rail wear. In addition, since there are no longitudinal forces (the line of  $f_{long} = 0$  in Figure 8.17), the wheel–rail friction acts entirely in the lateral direction, resulting in the shortest distance to climb and greater flange climb risk.

In Section III.A.2.a, the conservative nature of the Nadal criterion was discussed. However, for independently rotating wheels, any  $L/V$  values that exceed the Nadal limit would cause wheel flange climb because there is no relaxation from the effect of longitudinal force and friction coefficient level on the nonflanging wheel. Therefore, independently rotating wheels have less



**FIGURE 8.16** Steering moment formed by wheel longitudinal forces due to different rolling radius on two wheels.



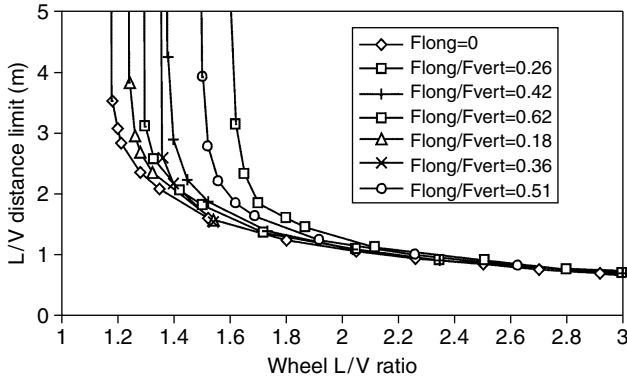


FIGURE 8.17 Effect of wheel L/V ratio on wheel climb distance (5 mrad wheelset angle of attack).<sup>11</sup>

tolerance to track irregularities that may suddenly increase wheel lateral forces or reduce vertical forces.

In summary, vehicles with independently rotating wheels need to be carefully designed to control flange climb and wheel wear. Additional control mechanisms, such as linkages or active control systems, can be used to steer the wheelset on curves and track perturbations. Without such control mechanisms, the wheel–rail profiles, vehicle–track maintenance and wheel–rail friction will need to be much more strictly controlled and monitored to prevent wheel flange climb.

### C. DERAILMENTS CAUSED BY GAUGE WIDENING AND RAIL ROLLOVER

Derailments caused by gauge widening usually involve a combination of wide gauges and large lateral rail deflections (rail roll), as shown in Figure 8.18. Large lateral forces from the wheels act to spread the rails in curves. Both rails may experience significant lateral translation and/or railhead roll, which often cause the nonflanging wheel to drop between rails. Figure 8.19 was produced based on a photo of an actual derailment caused by rail rollover, in which the nonflange wheel fell between rails and the outer rail was rolled over. Frequently, the inner rail will rollover due to contact with hollow-worn wheels, as discussed in Section III.C.3.

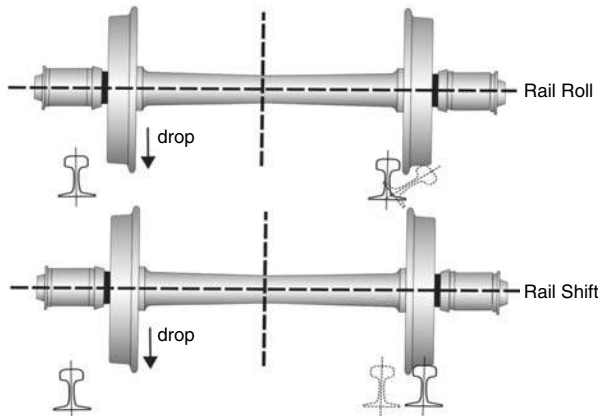
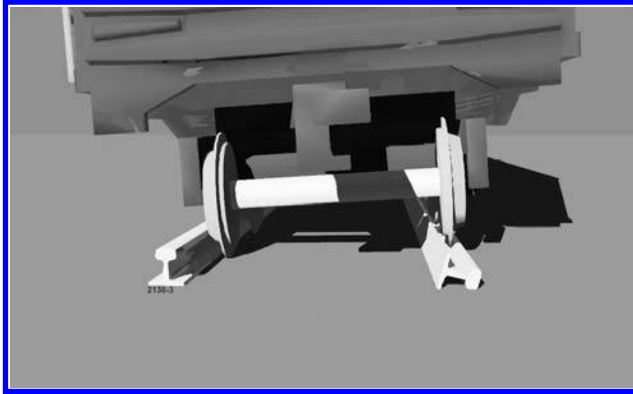


FIGURE 8.18 Gauge widening derailment.<sup>7</sup>



**FIGURE 8.19** Nonflange wheel falls between the rails, while the outer rail is rolled over.

When a bogie experiences poor steering, the wheelsets may experience high angles of attack in curves, resulting in large lateral forces exerted on the rails. The poor steering can be caused by inadequate suspensions (generally indicated by low warp or skew stiffness), high bogie turning resistance, misaligned axles, poor wheel and rail profile compatibilities,<sup>22</sup> and wheels having significant tread hollowing. The dynamic forces caused by track lateral perturbations and curve entry/exit spirals can intensify the lateral force level to deflect the rail further.

Locomotives produce high traction forces on rails. Since their bogies are long and can result in large angles of attack, three-axle locomotives (which are now common for heavy haul operation), have been considered as an important cause of gauge widening and rail rollover derailments.

Rail gauge wear is another cause for gauge widening, and is further discussed in Section III.C.2.

### 1. The AAR Chapter XI Rail Roll Criterion

The AAR Chapter XI rail roll criterion is established by using the  $L/V$  force ratio. The rail is assumed to rotate about the rail base corner under the load, as shown in Figure 8.20. The roll moment about the pivot point is given by,

$$M = Vd - Lh \quad (8.8)$$

Under an equilibrium condition, just before the rail starts to roll,  $M$  approaches to zero, then,

$$\frac{L}{V} = \frac{d}{h} \quad (8.9)$$

This  $L/V$  ratio is considered as the critical value to evaluate the risk of rail roll. When the  $L/V$  ratio is larger than the ratio of  $d/h$ , the risk of rail roll becomes high.

The critical  $L/V$  ratio for rail roll can vary from above 0.6 for contact at the gauge side to approximately 0.2 when the contact position is at the far-field side based on the dimension of rails. This is because the distance  $d$  (in Figure 8.20) is reduced. Note that this  $L/V$  ratio is calculated assuming that neither the rail fasteners nor the torsional stiffness of the rail section provide any restraint.

When considering the torsional rigidity of rail and the vertical force applied to the rail by the adjacent wheels, a criterion that only considers the forces due to a single wheel may be too conservative for predicting the stability of the rail. Therefore, the limiting criterion has counted the combined forces from all wheels on the same side of the bogie. Hence, a truck (bogie) side  $L/V$  ratio

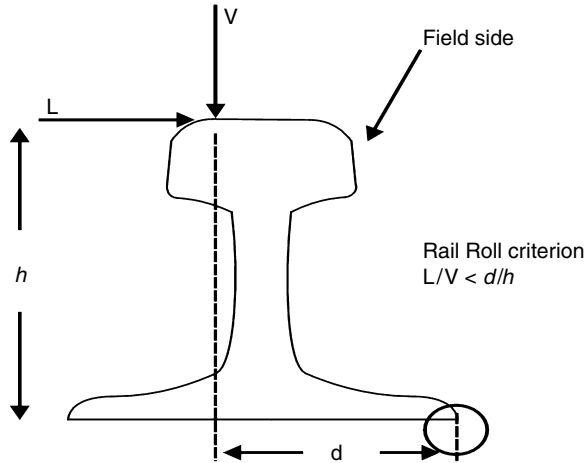


FIGURE 8.20 Illustration of rail roll criterion.

is defined by:

$$\frac{L}{V} = \frac{\text{sum of Lateral forces on truck side}}{\text{sum of Vertical forces on truck side}} \tag{8.10}$$

In Chapter XI of the *Manual of Standards and Recommended Practice* of AAR, the bogie side L/V ratio has been limited to below 0.6 in the vehicle yaw and sway tests.

**2. The Gauge Widening Criterion**

The gauge widening criterion is related to the wheel and rail geometries and their relative positions, as illustrated in Figure 8.21.

When the wheel drops between the rails, as in Figure 8.18, the geometry of wheel and rail must meet the following expression,

$$G \geq B + W + f_w \tag{8.11}$$

where  $G$ ,  $B$ ,  $W$ , and  $f_w$  are the rail gauge distance, wheel back-to-back space, wheel width, and flange thickness, respectively.

Therefore, a safety margin ( $S$ ), expressed in Equation 8.12, represents the minimum overlap of wheel and rail required on the nonflanging wheel, when the flanging wheel contacts the gauge face of the rail. In this circumstance, the instantaneous flangeway clearance on the flanging

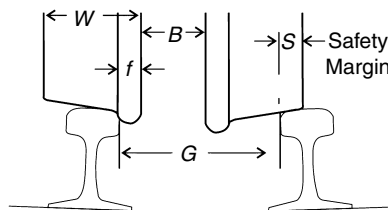


FIGURE 8.21 Wheel and rail geometry related to gauge widening derailment.

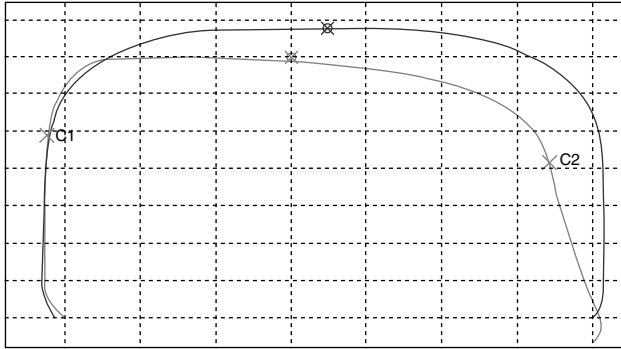


FIGURE 8.22 Gauge widen caused by rail gauge wear.

wheel is zero.

$$(B + W + f_w) - G > S \quad (8.12)$$

In general, the wheel back-to-back space ( $B$ ) is a constant for a solid axle, and so is the wheel width ( $W$ ). However, the flange thickness ( $f_w$ ) is gradually reduced as the wheel wears. The track gauge variations are influenced by multiple factors. As discussed in the previous section, rail roll and the lateral movement of rail due to weakened fasteners can widen the gauge. Rail gauge wear can also contribute to gauge widening, as shown in Figure 8.22, that gives a gauge wear of about 8 mm.

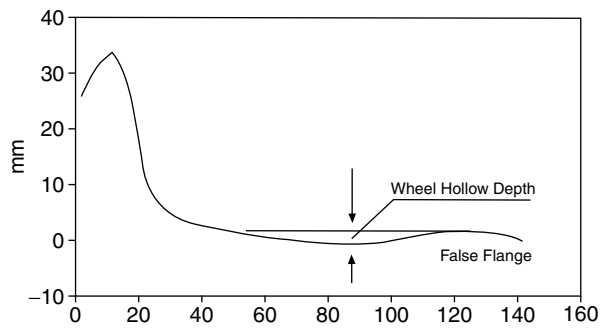
The North American interchange wheel of AAR1B is taken as an example. The back-to-back spacing of the AAR1B is 1350 mm, the wheel width is 145 mm, and the flange thickness is 35 mm. With a standard gauge of 1435 mm, the safety margin is 95 mm. A maximum 31.5-mm gauge widening (include the rail gauge wear measured under an unloaded condition) from the standard value is allowed for a freight vehicle operating in the speed range of 40 to 60 km/h, if a maximum of 15-mm wheel flange wear is allowed. Under this extreme condition, the overlapping is reduced to 48.5 mm. Therefore, any lateral shift and rotation of rail under the loaded condition can further reduce the overlapping to increase the risk of the wheel falling between the rails, especially on poorly maintained track.

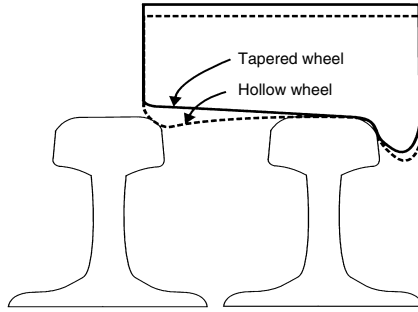
### 3. Effect of Hollow-Worn Wheels on Gauge Widening and Rail Roll Derailment

Wheel hollowing is defined as the vertical difference in rolling radius between the end of the tread and the minimum point around the middle of the tread. The value is found by placing a horizontal line through the highest point on the end of the tread. The wheel tread hollow tends to form a false flange at the end of the tread, as Figure 8.23 illustrates.

Figure 8.24 shows the probable contact condition of a measured hollow-worn wheel on a measured low rail. With wide gauge, or with combined wide rail gauge and thin wheel flange, the false flange of the hollow wheel is likely to contact the top of the low rail towards the field side on curves. Referring to the rail rollover criterion, stated in Equation 8.9, the value of  $d$  would be quite low under this condition, leading to a low ratio of  $d/h$ . Therefore, any truck (bogie) side  $L/V$  ratio larger than this  $d/h$  ratio would increase the risk of rail rollover, or put excessive forces on the fasteners. Figure 8.25 shows an example of loose spikes caused by repeated contact towards the field side of the rail.

The contact between the false flange of a hollow wheel and the field side of the inner rail in a curve, can lead to an adverse rolling radius difference condition between the inside wheel and





**FIGURE 8.26** False flange of a hollow-worn wheel applies a roll force to the stock rail on a riserless switch point.

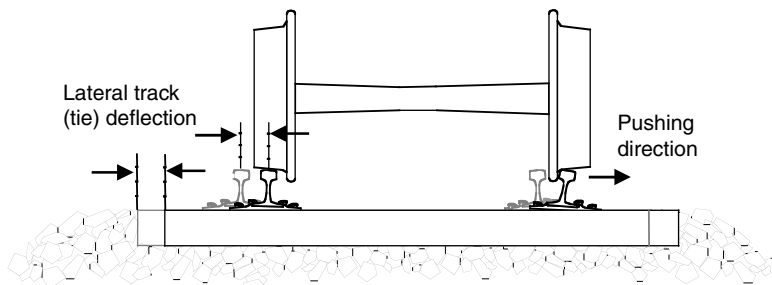
the outside wheel in curves. Under certain conditions of wheel–rail flange lubrication, this can cause the wheelset steering forces to reverse, increasing the wheelset angles of attack, and lateral forces, causing increased gauge widening and rail rollover.<sup>23</sup>

Switches are consistently among one of the most common track-related causes of derailments. Some incidents of rail rollover derailment in switches are expected to be directly related to hollow wheel profiles. Especially at riserless switch points, a car with hollow-worn wheels may cause rail rollover.<sup>24</sup>

For the vehicle trailing point moves, the wheel stays on the switch point with the false flange hanging down below the level of the stock-rail running surface. As the wheel approaches the point where the two railheads converge, the false flange can strike the side of the stock rail (see Figure 8.26). The high force produced from the false flange–stock rail interaction may either cause a stock rail to roll out, or, in severe cases, a wheel to climb.

#### D. DERAILMENT CAUSED BY TRACK PANEL SHIFT

Track panel shift is the cumulative lateral displacement of the track panel, including rails, tie plates and ties, over the ballast, as shown in Figure 8.27. A small shift of these components may not immediately cause the loss of guidance to bogies. However, as the situation gradually depreciates to a certain level, wheels could lose guidance and drop to the ground at some speed. The derailments caused by track panel shift usually result in one wheel falling between the rails and the other falling outside of the track.



**FIGURE 8.27** Lateral track panel shift.

### 1. Causes of Track Panel Shift

Track panel shift is a lateral misalignment phenomenon primarily caused by repeated lateral axle loads. Tracks that possess low resistance to lateral force, such as poorly laid track, newly laid track, and newly maintained track, often show separations between track panel and ballast. Track panels could shift under large lateral forces under such conditions. The capacity of a track panel to resist lateral movement is measured by lateral track strength and stiffness. Soft subgrade may also allow the panel to shift more freely.

Track panel shift has become increasingly important as both speed and load increase and more continuous welded rail is placed in use. The increase in speed may result in an increase in the unbalanced forces on curves or poorly aligned track by curving considerably above the balance speed. At high imbalance speeds, the wheelsets of many bogie designs stop generating gauge spreading forces, and instead, both wheels generate wheel–rail forces that act to force the rails outwards in a curve, leading to panel shift.

The escalation in load can increase the magnitude of lateral force at the wheel and rail interface. Continuous welded rail may buckle due to the longitudinal force caused by temperature change. Aggressive acceleration and braking can also induce large forces to cause panel shift on track with poor lateral resistance.

Knothe and Bohm<sup>25</sup> commented that how near the BB 9104 locomotive came to a catastrophe when it archived the world record (of 331 km/h in 1955). The test locomotive has caused a strong sinusoidal alignment fault of the track. There seem to have been two causes. The first is probably that the track was tamped just before the record test. By this maintenance operation the strength of track to resist lateral displacements was reduced. The second cause is that the locomotive was unstable, thus exerting high lateral forces on the dislodged ballast bed.

### 2. Panel Shift Criterion

Lateral track strength indicates the capability of track to resist track buckling and to retain lateral alignment under traffic. Track buckling is defined as the lateral deformation of track due to high compressive rail force in the longitudinal direction at a temperature above the rail neutral temperature. Track panel shift normally accumulates gradually. However, when the critical load level is exceeded, panel shift increases rapidly with the number of repeated load applications.

The definition of the critical lateral load can be based on either accumulated or incremental deformation after each load application. Note that with each load pass, the increment of total deformation (elastic plus residual) remains constant for a stable track. Below critical loads, the elastic deformation remains constant and the residual deformation tends to zero. With each load pass over a segment of unstable track, both elastic and residual deformation increment will grow.

Research by the French National Railways suggested that the limiting lateral axle load can be defined in a general expression (Equation 8.13) for preventing excessive track panel shift,<sup>7,26</sup>

$$L_c = aV + b \tag{8.13}$$

where  $L_c$  is the critical lateral load and  $V$  is the vertical axle load. Table 8.2 lists two groups of suggested values of  $a$  and  $b$ .

A further multiplying factor of 0.85 was specified for use in Canada<sup>7</sup> to allow for the quality of tie-down and ties-in in some tracks. Then, Equation 8.13 is reformatted as,

$$L_c = 0.85(aV + b) \tag{8.14}$$

**TABLE 8.2**  
**Suggested Values of  $a$  and  $b$**

Prud'homme	$a = 0.333$ $b = 9.96$ kN for uncompacted ballast	$a = 0.333$ $b = 14.99$ kN for well compacted ballast and concrete ties
Ahlbeck and Harrison	$a = 0.4$ $b = 9.96$ kN for uncompacted ballast on wood ties	$A = 0.7$ $b = 24.6$ kN for compacted ballast on wood ties

A multiplying factor proposed by Ahlbeck and Harrison considered the effect of track curvature and temperature in determining the lateral force limit.<sup>7</sup>

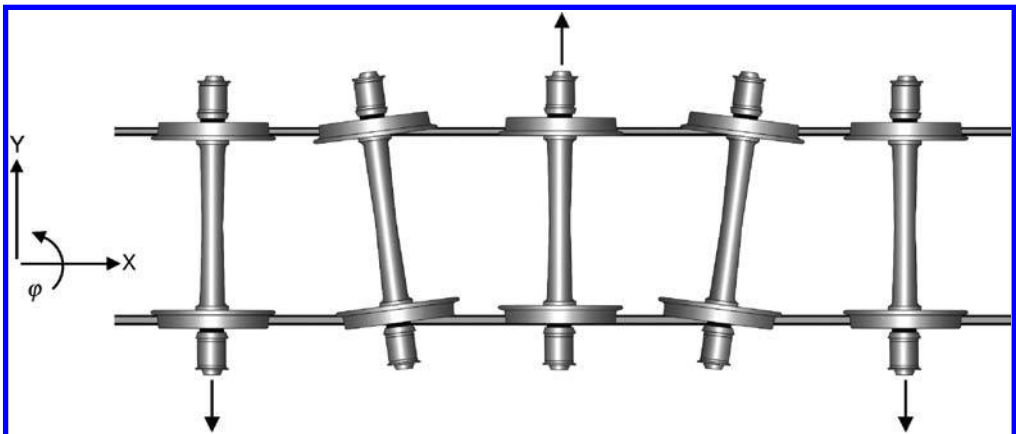
$$A = 1 - \frac{S\Delta}{22,300}(1 + 0.46D) \quad (8.15)$$

where  $S$  is the area of rail section (square inches),  $\Delta$  is temperature change (degree Fahrenheit) and  $D$  is degree of track curvature.

The lateral track strength and track panel shift study conducted by Li and Shust<sup>26</sup> concluded that vertical axle load has a major effect on the resistance of a track panel to lateral deflection. The lateral track strength and stiffness would only be valid for a given vertical load. In their in motion panel shift test, the critical lateral axle force for causing panel shift is approximately 15 to 30% higher for track with concrete ties than wood tie track.

## E. DERAILMENT CAUSED BY VEHICLE LATERAL INSTABILITY

On tangent track, the wheelset generally oscillates around the track centre due to any vehicle and track irregularities, as shown in Figure 8.28. This movement occurs because vehicle and track are never absolutely smooth and symmetric. This self-centring capability of a wheelset is induced by the coned shape of the wheel tread. However, as speed is increased, if the wheelset conicity is high, the lateral movement of wheelset, as well as the associated bogie and car body motion, can cause oscillations with large amplitude and a well-defined wavelength. The lateral movements are limited only by the contact of the wheel flanges with the rail. This vehicle dynamic response is also termed as vehicle hunting, and can produce high lateral forces to damage track and to cause derailments.



**FIGURE 8.28** Wheelset oscillates around the track centre.



Derailments caused by vehicle hunting can have derailment mechanisms of all the four types discussed in the previous sections. The high lateral force induced from hunting may cause wheel flange climbing on the rail, gauge widening, rail rollover, track panel shift, or combinations of these. The safety concerns for this type of derailment, usually occurring at higher speeds, make it an important area of study.

Hunting predominantly occurs in empty or lightweight vehicles. The critical hunting speed is highly dependent on the vehicle/track characteristics. When vehicle hunting is onset, the displacements of wheelset are generally large, alternatively flanging from one side of the rails to the other. Considering the wheel/rail geometry and the creep force saturation, the vehicle/track system under hunting conditions should be treated as nonlinear. Investigation of the critical speed for such a system with nonlinearities is to examine the vehicle dynamic response to a disturbance using a numerical solution of the equations of motion.<sup>27</sup>

Vehicle simulation computer models, which include the processes to solve these equations of motion, are often used to predict the hunting speed. Track tests are also generally required to either validate the hunting speed predicted by modelling or ensure the system operating speed is below the hunting onset speed.

The effective conicity of wheel–rail contact has considerable influence on the vehicle hunting speed. As wheelset conicity increases, the onset critical speed of hunting decreases. For this reason, it is important when designing wheel and rail profiles to ensure that, for a specific bogie/vehicle, the critical hunting speed is above the operating speed.

## IV. PREDICTION OF DERAILMENT

Three types of approach are usually used for predicting the risks of derailment or diagnosing the causes of derailments: assessment of wheel/rail parameters, dynamic simulations of vehicle–track interaction and vehicle performance track tests. In many cases, all three approaches are applied.

### A. ASSESSMENT OF WHEEL/RAIL PARAMETERS

Assessment of wheel/rail parameters may predict the risks of derailment that are the result of unfavourable wheel–rail contact. Since wheel and rail are operating as a mechanical system involving two-body contact and interaction, this assessment of wheel/rail parameters should include wheels operating on the line where rails are measured (or designed for the new line condition). The contact parameters that may affect derailment include:

- Maximum contact angle and length of flange (related to flange climb)
- Rolling radius difference on curves (related to flange climb)
- Effective conicity (related to vehicle lateral instability on tangent track)
- Rail gauge (related to gauge widening)
- Contact positions (related to rail rollover)
- Wheel–rail contact conformity (affects bogie steering and level of lateral forces)
- Rolling radius difference of two wheels on a same axle (affects bogie steering)
- Level of wheel tread hollowing (related to rail rollover, also affects wheel climb and vehicle instability under certain conditions, and rolling radius difference)
- Significant wear of wheel flange and rail gauge can increase wheelset lateral movement and may reduce the effectiveness of restraining rail on curves

A comprehensive view of wheel–rail contact at a system level is important to reveal the overall patterns of the contact. For example, thousands of wheels with different profiles (due to different levels of wear or resulting from different bogie performances) contact a section of rail at different positions and could produce different levels of contact stress. Therefore, the performances of the majority of wheel–rail pairs are of interest in system assessment.

However, derailments are often related to the behaviours of individual wheels and rails that possess undesired shapes. In diagnosing the causes of derailment, the wheel profiles on derailed cars and rail profiles at derailment sections must be measured and analysed to study the contributions from the wheel/rail shapes.

## B. DYNAMIC SIMULATION OF VEHICLE–TRACK INTERACTION

Since the 1980s, computer simulations have been extensively used to study vehicle–track interaction (see Chapter 12). They are also useful tools to diagnose the causes of derailments. For some derailments, the explicit causes cannot be simply identified. Some derailments cannot even be repeated by track tests that appear to have similar conditions as the derailed cars and tracks. The advantages of a computer modelling study are that parameters of vehicle/track can be conveniently (also cost effectively) varied to investigate the effects of either single parameters or combinations of multiple parameters on derailments. Consequently, if the simulation reveals derailment risk, modifications to vehicle and track for preventing derailments can also be defined by a parametric study.

The flange climb risk can be evaluated in the simulation using the flange climb criteria discussed in Section III.A including the L/V ratio limit and exceeding distance (or time) limit. Simulation of derailments caused by gauge widening and panel shift under dynamic load requires more advanced vehicle/track simulation models that include the capability of rail roll under the load, which results in changing wheel–rail contact condition, and a detailed track model to describe the structure below the rail.

The accuracy of the description of vehicle and track parameters is crucial for simulation to reflect the actual responses of vehicle and track. For example, a sudden change of wheel/rail forces due to large track lateral irregularity combined with wheels with low flange angle could lead to derailment. Therefore, only this irregularity and the shape of the wheels are accurately described in the model with other vehicle/track parameters, the derailment scenario will be reproduced by the simulation. When evaluating a system, it is important to examine the vehicles under the worst track conditions possibly allowed in the system in order to ensure safe operation under those conditions. When examining newly designed (or modified) vehicles, a certain limit of wear of vehicle elements and wheel profiles should also be considered in the simulations.

## C. TRACK TESTS

On-track tests are generally required for new designs or modifications of vehicles. Track tests are also often conducted for diagnosing performance problems, including derailment, caused by vehicle or track conditions, or a combination of both. These vehicles may have been examined by computer simulation. However, precisely describing every element and parameter in the vehicle/track system is very difficult, especially for those nonlinear elements, such as friction elements, damping elements, and gaps/stops. Modelling can reveal the trends or probable performances with regard to derailment. Track tests will demonstrate the actual performance under different test conditions for the test unit. Therefore, computer simulations and track tests are often combined efforts in derailment evaluations or diagnoses.

On-track tests of new vehicle designs is required for the North American Freight Service. Chapter XI of the *Manual of Standard and Recommended Practices*, approved by the Association of American Railroads, describes the regimes of vehicle performance to be examined and the required test conditions.<sup>15</sup> The test regimes in Chapter XI include:

- Hunting (vehicle lateral instability)
- Constant curving
- Spiral

- Twist, roll
- Pitch
- Yaw, sway
- Dynamic curving

Usually, the test unit is instrumented with a number of gauges to measure forces, accelerations, and displacements at critical locations on the test unit depending on the test objective. Instrumented wheelsets have been applied in recent years to determine more accurately wheel–rail interaction forces. The wheel L/V ratio limit (or axle L/V ratio limit) and the exceeding distance (or time) limit are used in Chapter XI tests to evaluate vehicle curving performance. Wheel unloading is also an important criterion used in the Chapter XI test. It limits the level of minimum vertical wheel force to no less than 10% of static load.

The FRA (U.S.A.) has also implemented a requirement for testing of new passenger trains that run at speeds above 90 mi/h (144 km/h) for passenger cars and 80 mi/h (128 km/h) for freight cars.<sup>16</sup>

The Railway Safety and Standard Board of the U.K. has issued railway groups standards relevant to engineering acceptance. These standards describe the permissible track forces for railway vehicles.<sup>28,29</sup>

## V. PREVENTION OF DERAILMENT

Most derailments may be prevented if the problems, related either to vehicles or tracks, that lead to derailment could be identified and adequate preventive actions could be taken at an early stage.

The four types of derailments discussed in the previous sections have a common cause of high lateral force at the wheel–rail interface. Therefore, any conditions that lead to high lateral forces or lead to lower the ability of the system to sustain the force should be corrected.

In this section, some general preventive methods are introduced. However, due to the wide range of vehicle types and track conditions, any methods that are adopted by a system to prevent derailment must be carefully assessed by considering the specific vehicle and track conditions in that system to ensure the effectiveness of the methods.

### A. WHEEL/RAIL PROFILES

#### 1. Addressing Wheel Flange Angle

To prevent flange climb derailment, the maximum wheel flange angle should be sufficiently high to increase the allowed L/V ratio limit. For a new wheel profile design, a higher flange angle should be emphasised. A flange angle above 70° is generally recommended. Should flange climb derailments be a concern to an existing system that has adopted wheels with a low flange angle, a transition to a higher flange angle might be considered. However, this transition needs to be carefully planned according to the capacity of wheel truing and rail grinding in the system.<sup>30</sup>

#### 2. Removing Hollow-Worn Wheels

Removing significantly hollowed worn wheels from the system may reduce the risk of gauge widening and rail rollover derailment, as described in Section III.C.3. Hollow wheels can also reduce rolling radius difference required in curving and increase lateral instability on tangent track.<sup>31,32</sup> A starting 4-mm hollow wheel removing limit has been recommended by TTCI for the North American interchange operation. The aim is to eventually remove the wheels with 3-mm hollow tread from the service.<sup>33</sup>

## B. INDEPENDENTLY ROTATING WHEELS

As discussed in Section III.B.3, independently rotating wheels tend to run with a larger angle of attack in curves than the conventional coupled wheelsets and can generate greater lateral forces that increase the risk of wheel climb. Therefore, independent rotating wheels require more carefully designed wheel profiles and control mechanisms for curving. Elkins<sup>34</sup> and Suda et al.<sup>35</sup> have proposed self-steering bogies with independent rotating wheels equipped on the trailing axle only. In recent years, the concept of active control has been studied.<sup>36,37</sup> As the yaw angle and lateral motion of wheels can be accurately controlled, more applications of independent rotating wheels can be expected.

## C. INSTALLATION OF GUARD RAIL OR RESTRAINING RAIL ON SHARP CURVES

Restraining rails and guard rails have been frequently applied in transit operations on sharp curves to prevent flange climb derailment (or to reduce gauge wear on the high rail). The restraining/guard rails are generally installed inside of the low rail, as shown in Figure 8.29. In extremely sharp curves restraining rails are sometimes installed on both the inside and outside rails.

The clearance between the low rail and the restraining rail is critical for the effectiveness of restraining rails. Too tight clearance may reduce wheelset rolling radius difference required for bogie curving by limiting the flange contact on the high rail. Overwide clearance may completely lose the restraining function.

Wear at the wheel flange back and the contact face of the restraining rail can vary the clearance between the low rail and the restraining rail. The wheel flange and high rail gauge wear can affect the amount of wheelset lateral shift on curves. Note that track lateral geometry irregularities, including alignment and gauge variations can also affect the performance of restraining rails.

## D. OPTIMISING BOGIE SUSPENSION AND BOGIE YAW STIFFNESS

The suspension design of a bogie affects its steering capability. Bogies with soft primary suspensions that allow the axles to steer in curves generally generate lower lateral forces than bogies with stiff primary suspensions. However, if the primary suspension is too soft, high-speed stability may be reduced. Therefore, optimising bogie suspension to meet the specific requirements of a system (mainly in terms of track condition and operating speed) is essential for the bogie design.

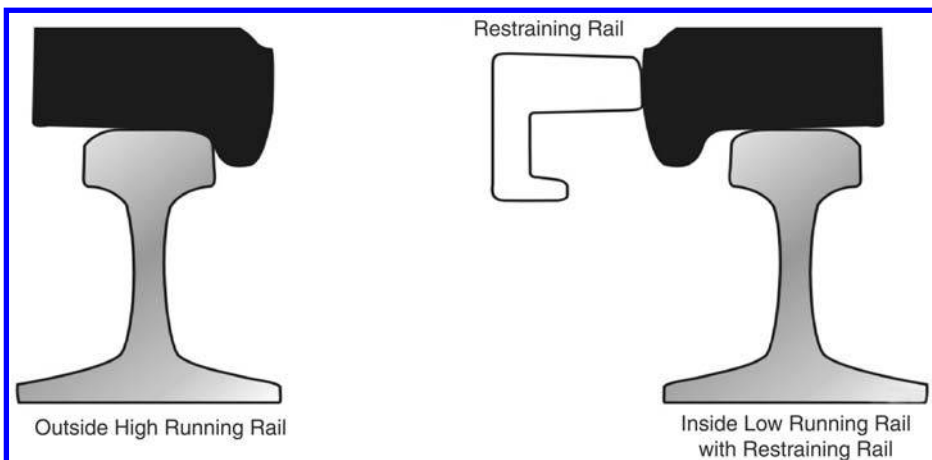


FIGURE 8.29 Restraining rail.

## E. LUBRICATION

Friction plays an important role in wheel-rail interface. It affects many wheel-rail interaction scenarios. With respect to derailment, proper lubrication at the wheel-rail interface can reduce wheel lateral forces because wheel lateral creep force ( $F_{\text{lat}}$ ) saturates at a lower level ( $F_{\text{lat}} = \mu N$ ,  $N$  is the contact normal force). Therefore, the potential of wheel climbing and gauge widening is reduced. As illustrated in Figure 8.4, the limiting wheel L/V ratio for wheel climb increases with the decreases in the friction coefficient at contact surface.

In recent years, top-of-rail friction management, with an intermediate range of coefficient of friction (from 0.2 to 0.4), has been tested.<sup>38,39</sup> Application of top-of-rail friction modifiers is intended to reduce rolling resistance, corrugation, and eliminate wheel squeal. Top-of-rail friction management can also be expected to reduce vehicle lateral instability on tangent track.

## F. TRACK GEOMETRY INSPECTION AND MAINTENANCE

Severe track lateral and vertical irregularities and gauge widening are significant causes of derailment. On many railways, track geometry recording cars are used regularly to survey the track. The frequency of survey varies greatly, between monthly and yearly, according to track category.<sup>16, 40</sup> Track is categorised based on a combination of maximum speed and annual tonnage. Regular visual inspection is often carried out at much shorter intervals (between once and twice a week). Regulations of track geometry for different regions and railways may be different. However, maintenance is normally required when the track geometry deviations exceed the specified limits.

Note that newly laid or newly maintained track requires special attention due to its low lateral strength. Speed restrictions may be required for a period to let the track settle.

## G. SYSTEM MONITORING

Derailments usually occur from a combination of unfavourable vehicle and track conditions. Normally, only a single vehicle with a particular problem derails in a section of track with an adverse condition. On the other hand, a bogie with a particular defect usually does not derail in every section of track. A program of system monitoring can be implemented to detect any derailment-related vehicle and track problems, including wheel/rail profiles, in their early stages. Then, corrective action can be promptly taken to prevent occurrences of derailment.

## REFERENCES

1. <http://home.clara.net/gw0hqd/bumps/101181/101181.htm> (Carnforth derailment)
2. [http://news.com.au/common/story\\_page/0,4057,10221240%255E1702,00.html](http://news.com.au/common/story_page/0,4057,10221240%255E1702,00.html) (Turkey derailment)
3. FRA Guide for Prepare Accident/Incident Report, U.S. Department of Transportation, Federal Railroad Administration, May 2003.
4. Railroad Safety Statistics, Annual Report 2002, U.S. Department of Transportation, Federal Railroad Administration.
5. *Safety Database Analysis*, Transportation Technology Center, Inc., 2002.
6. Harry, T., Wu, H., and Guins, T., *The Influence of Hollow-Worm Wheels on the Incidence and Costs of Derailments*, Association of American Railroads/Transportation Technology Center, Inc. report R-965, February 2004.
7. Blader, F.B., *A Review of Literature and Methodologies in the Study of Derailments Caused by Excessive Forces at the Wheel/Rail Interface*, Association of American Railroads Report R-717, December 1990.
8. Nadal, M.J., *Locomotives a Vapeur*, Collection Encyclopedie Scientifique, Biblioteque de Mecanique Appliquee et Genie, Vol. 186, Paris, 1908.

9. Shust, W.C., Elkins, J.A., Kalay, S., and El-Sibaie, M., *Wheel-Climb Derailment Tests using AAR's Track Loading Vehicle*, Association of American Railroads Report R-910, December 1997.
10. Weinstock, H., *Wheel Climb Derailment Criteria for Evaluation of Rail Vehicle Safety*, Proceedings of ASME Winter Annual Meeting, 84-WA/RT-1, New Orleans, LA, 1984.
11. Wu, H. and Elkins, J., *Investigation of Wheel Flange Climb Derailment Criteria*, Association of American Railroads Report R-931, July 1999.
12. Elkins, J. and Wu, H., *New Criteria for Flange Climb Derailment*, IEEE/ASME Joint Railroad Conference, Newark, NJ, April 4–6, 2000.
13. Matsudaira, T., *Dynamics of High Speed Rolling Stock*, Japanese National Railways RTRI Quarterly Reports, Special Issue, 1963.
14. Koci, H. H. and Swenson, C. A., *Locomotive Wheel-Loading — A System Approach*, General Motors Electromotive Division, LaGrange, IL, February, 1978.
15. M1001, AAR Mechanical Division, *Manual of Standards and Recommended Practices*, Section C — Part II, Volume 1, Chapter XI, Section 11.5.2 Track-Worthiness Criteria, Adopted 1987, Revised 1993.
16. Federal Railroad Administration. *Track Safety Standards, Part 213, Subpart G*, September, 1998.
17. Wu, H., Shu, X., and Wilson, N., *Flange Climb Derailment Criteria and Wheel/Rail Profile Management and Maintenance Guidelines for Transit Operations — Appendix B*, Transit Cooperate Research Program Report 71, Vol. 5, 2005.
18. Wilson, N., Shu, X., Wu, H., and Tunna, J., *Distance-Based Flange Climb LV Criteria*, Association of American Railroads/Transportation Technology Center, Institute of Technology Digest, TD-04-012, July 2004.
19. *Track Design Handbook for Light Rail Transit*, Transit Cooperate Research Program, Report 57, 2000.
20. *Passenger Rail Safety Standard Task Force Technical Bulletin*, 1998-2001, Part 1.
21. Wilson, N., Shu, X., and Kramp, K., *Effect of Independently Rolling Wheels on Flange Climb Derailment*, Proceedings of ASME International Mechanical Engineering Congress, November, 2004.
22. Mace, M., Dibrito, D., Blank, R., Keegan, S., and Michael, Allran, *Effect of Wheel Profiles on Gauge Widening Behaviour*, Proceeding of ASME/IEEE Joint Railroad Conference, March, 1994.
23. Mace, S., Pena, R., Wilson, N., and Dibrito, D., Effect of wheel-rail contact geometry on wheelset steering forces, *Wear*, 191(1–2), January, 1996.
24. Singh, S. and Davis, D.D., *Effect of Switch-Point Risers on Turnout Performance*, Association of American Railroads, Transportation Technology Center, Inc., Technology Digest 98-018, June 1998.
25. Knothe, K. and Bohm, F., History of stability of railway and road vehicles, *Vehicle Syst. Dyn.*, 31(5–6), June, 1999.
26. Li, D. and Shust, W., *Investigation of Lateral Track Strength and Track Panel Shift Using AAR's Track Loading Vehicle*, AAR/TTCI report R-917, December 1997.
27. Wickens, A., *Fundamentals of Rail Vehicle Dynamics — Guidance and Stability*, Swets & Zeitlinger, Lisse, 2003.
28. *Permissible Track Forces for Railway Vehicles*, British Railways Board, GM/TT0088, October 1993.
29. *Commentary on Permissible Track Forces for Railway Vehicles*, Railtrack, UK, 1995.
30. Wu, H., Shu, X., and Wilson, N., *Flange Climb Derailment Criteria and Wheel/Rail Profile Management and Maintenance Guidelines for Transit Operations — Appendix A*, Transit Cooperate Research Program Report 71, Vol. 5, 2005.
31. Sawly, K., Urban, C., and Walker, R., *The Effect of Hollow-Worn Wheels on Vehicle Stability in Straight Track*, Proceedings of Contact Mechanics and Wear of Rail/Wheel System, The Sixth International Conference, June, 2003.
32. Sawley, K. and Wu, H., *The Formation of Hollow-Worn Wheels and Their Effect on Wheel/Rail Interaction*, Proceedings of Contact Mechanics and Wear of Rail/Wheel System, The Sixth International Conference, June, 2003.
33. Sawley, K. and Clark, S., *The Economics of Removing Hollow Wheels from Service*, Association of American Railroads/Transportation Technology Center, Ins. Technology Digest, TD-99-034, December 1999.
34. Elkins, J. A., *Independently Rotating Wheels. A Simple Modification to Improve the Performance of the Conventional Three-Piece Truck*, Proceedings of the Ninth International Wheelset Congress, Montreal, Quebec, Canada, September, 1988.

35. Suda, Y., Nishimura, R., Kata, N., Matsumoto, A., Sata, Y., Ohno, H., Tanimoto, M., and Miyauchi, E., Self-steering tricks using unsymmetrical suspension with independently rotating wheels — Comparison between stand tests and calculations, *Vehicle Syst. Dyn.*, 33, 2000.
36. Mei, T.X. and Goodall, R.M., Wheelset control strategies for a two-axle railway vehicle, *Vehicle Syst. Dyn.*, 33, 2000.
37. Shen, S., Mei, T. X., Goodall, R. M., Pearson, J., and Himmelstein, G., *A Study of Active Steering Strategies for Railway Bogies*, 18th IAVSD Symposium, Dynamics of Vehicles on Road and Tracks, August, 2003.
38. Eadie, D. and Stantoro, M., *Railway Noise and the Effect of Top of Rail Liquid Friction Modifiers: Changes in Sound and Vibration Spectral Distributions*, Proceedings of Contact Mechanics and Wear of Rail/Wheel System, The Sixth International Conference, June, 2003.
39. *Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues*, International Heavy Haul Association, May 2001.
40. Federal Railroad Administration. *Track Safety Standards*, Part 213, Subpart A to F, September, 1998.

---

# 9 Longitudinal Train Dynamics

*Colin Cole*

## CONTENTS

I.	Introduction .....	239
A.	An Overview of Longitudinal Train Dynamics .....	240
II.	Modelling Longitudinal Train Dynamics .....	241
A.	Train Models .....	241
B.	Wagon Connection Models .....	243
1.	Conventional Autocouplers and Draft Gear Packages.....	244
2.	Slackless Packages.....	254
3.	Drawbars .....	254
C.	Locomotive Traction and Dynamic Braking.....	255
D.	Pneumatic Brake Models.....	259
E.	Gravitational Components .....	260
F.	Propulsion Resistance .....	261
G.	Curving Resistance .....	263
H.	Train Dynamics Model Development and Simulation.....	263
III.	Interaction of Longitudinal Train and Lateral/Vertical Wagon Dynamics .....	264
A.	Wheel Unloading on Curves due to Lateral Components of Coupler Forces .....	264
B.	Wagon Body Pitch due to Coupler Impact Forces .....	264
C.	Bogie Pitch due to Coupler Impact Forces .....	265
IV.	Longitudinal Train Crashworthiness .....	266
A.	Vertical Collision Posts .....	266
B.	End Car Crumple Zones .....	267
V.	Longitudinal Comfort .....	267
VI.	Train Management and Driving Practices.....	269
A.	Train Management and Driving Practices.....	269
1.	Negotiating Crests, Dips, and Undulations .....	270
2.	Pneumatic Braking.....	270
3.	Application of Traction and Dynamic Braking .....	271
4.	Energy Considerations .....	272
5.	Distributed Power Configurations .....	273
VII.	Conclusions .....	275
	Acknowledgments .....	276
	Nomenclature.....	276
	References.....	277

## I. INTRODUCTION

Longitudinal train dynamics is discussed from the background of the Australian Railway industry. The technology and systems used draw from both British and North American systems. Structure



and rollingstock gauges are clearly influenced by the British railway practice, as are braking systems. Wagon couplings on freight trains are predominately autocouplers with friction wedge type draft gear packages showing the North American influence. Privately owned railways on iron ore mines in the Australia's North West show even more North American influence with American style braking and larger structure and rollingstock gauges. Australia is also characterised by three track gauges, a legacy of colonial and state governments before federation. The presence of narrow gauges of 1067 mm results in a large fleet of rollingstock with a design differing from standard gauge rollingstock in North America, Britain, and the southern states of Australia.

This chapter is arranged to firstly give an overview of longitudinal train dynamics. The second section goes into considerable detail on approaches to modelling longitudinal train dynamics. The most space is given to the modelling of the wagon connection model. Subsections are also devoted to modelling traction and dynamic braking systems, rolling resistance, air resistance, curving resistance, the effect of grades, and pneumatic braking. The subsection on pneumatic braking only provides an explanation of the effect of pneumatic braking on train dynamics. Modelling pneumatic braking systems would require a chapter in itself. Further more brief chapter sections are included on the interaction of longitudinal train dynamics with lateral/vertical wagon dynamics, crash-worthiness, comfort and train management, and driving practices.

## A. AN OVERVIEW OF LONGITUDINAL TRAIN DYNAMICS

Longitudinal train dynamics is defined as the motions of rollingstock vehicles in the direction of the track. It therefore includes the motion of the train as a whole and any relative motions between vehicles allowed due to the looseness of the connections between vehicles. In the railway industry, the relative motion between vehicles is known as "slack action" due to the correct understanding that these motions are primarily allowed by the free slack in wagon connections, coupling free slack being defined as the free movement allowed by the sum of the clearances in the wagon connection. These clearances consist of clearances in the autocoupler knuckles and draft gear assembly pins. Cases of slack action are further classified in the Australian industry vernacular as run-ins and run-outs. The case of a run-in describes the situation where vehicles are progressively impacting each other as the train compresses. The case of a run-out describes the opposite situation where vehicles are reaching the extended extreme of connection free slack as the train stretches. Longitudinal train dynamics therefore has implications for passenger comfort, vehicle stability, rollingstock design, and rollingstock metal fatigue.

The study and understanding of longitudinal train dynamics was probably firstly motivated by the desire to reduce longitudinal oscillations in passenger trains and in so doing improve the general comfort of passengers. The practice of power braking, that being keeping power applied with minimum air braking, is still practiced widely in Australia on passenger trains. Power braking is also used on partly loaded mixed freight trains to keep the train stretched during braking and when operating on undulating track. In the Australian context, the study of longitudinal train dynamics is evidenced in technical papers coinciding with the development of heavy haul unit trains for the transport of coal and iron ore. Measurement and simulation of in-train forces on such trains in the Queensland coal haulage was reported by Duncan and Webb.<sup>1</sup> Moving to trains of double existing length was reported at the same time in New South Wales in a paper by Jolly and Sismey.<sup>2</sup> Interest was also evident in South Africa with the publication of a paper focused on train handling techniques on the Richards Bay Line.<sup>3</sup> The research was driven primarily by the occurrences of fatigue cracking and tensile failures in autocouplers. From these studies<sup>1-3</sup> an understanding of the force magnitudes and an awareness of the need to limit these forces with appropriate driving strategies was developed. During these developments, the first measurement of in-train forces in long trains utilising distributed locomotive placement were completed. An important outcome was that a third type of in-train force behaviour was identified. Prior to these studies in-forces were divided into two types, namely, steady forces and impact forces. Steady in-train forces are

associated with steady applications of power or braking from the locomotives or train air braking, combined with drag due to rolling resistance, air resistance, curve drag, and grades. Impact in-train forces are associated with run-in and run-out occurrences due to changes in locomotive power and braking settings, changes in grade and undulations. In trains with distributed power, a new force phenomena known as low frequency oscillations was identified. This new behaviour was further classified into two distinct modes, namely cyclic vibration and sustained longitudinal vibration.<sup>1</sup> Sustained longitudinal vibration occurred only when the entire train was in a single stress state, either tensile or compression. The oscillation was underdamped and approximated to a smooth sinusoid. Of interest was that the magnitude of the in-train force associated with this low frequency oscillation could approach the magnitude of the steady in-train force, representing a substantial increase in possible fatigue damage and the risk of vehicle instability. Cyclic vibrations were characterised by oscillations approximating a square wave and occur due to run-in/run-out behaviour. Cyclic vibration differed from impacts in that the vibrations could be sustained for several seconds. The need to control, and where possible reduce, in-train forces resulted in the development of longitudinal train simulators for both engineering analysis and driver training.

More recent research into longitudinal train dynamics was started in the early 1990s, motivated not this time by equipment failures and fatigue damage, but derailments. The direction of this research was concerned with the linkage of longitudinal train dynamics to increases in wheel unloading. It stands to reason that as trains get longer and heavier, in-train forces get larger. With larger in-train forces, lateral and vertical components of these forces resulting from coupler angles on horizontal and vertical curves are also larger. At some point these components will adversely affect wagon stability. The first known work published addressing this issue was that of El-Siabi,<sup>4</sup> which looked at the relationship between lateral coupler force components and wheel unloading. Further modes of interaction were reported and simulated by McClanachan et al.<sup>5</sup> in 1999, detailing wagon body and bogie pitch.

Concurrent with this emphasis on the relationship between longitudinal dynamics and wagon stability is the emphasis on train energy management. The operation of larger trains meant that the energy consequences for stopping a train become more significant. Train simulators were also applied to the task of training drivers to reduce energy consumption. Measurements and simulations of energy consumed by trains normalised per kilometre–tonne hauled have showed that different driving techniques can cause large variances in the energy consumed.<sup>6,7</sup>

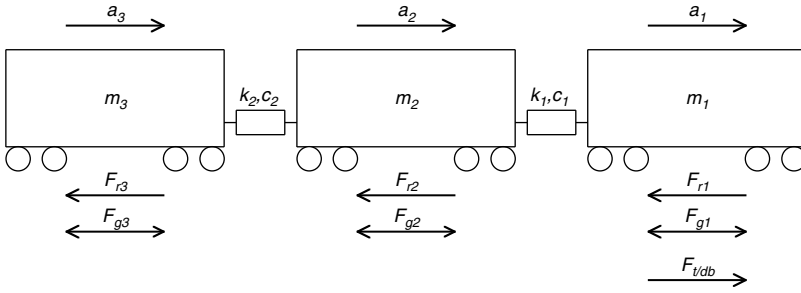
## II. MODELLING LONGITUDINAL TRAIN DYNAMICS

### A. TRAIN MODELS

The longitudinal behaviour of trains is a function of train control inputs from the locomotive, train brake inputs, track topography, track curvature, rollingstock and bogie characteristics, and wagon connection characteristics.

The longitudinal dynamic behaviour of a train can be described by a system of differential equations. For the purposes of setting up the equations, modelling, and simulation, it is usually assumed that there is no lateral or vertical movement of the wagons. This simplification of the system is employed by all known rail specific, commercial simulation packages and by texts such as Garg and Dukkipati.<sup>8</sup> The governing differential equations can be developed by considering the generalised three mass train in [Figure 9.1](#). It will be noticed that the in-train vehicle, whether locomotive or wagon, can be classified as one of only three connection configurations, lead (shown as  $m_1$ ), in-train, and tail. All vehicles are subject to retardation and grade forces. Traction and dynamic brake forces are added to powered vehicles.

It will be noted on the model in [Figure 9.1](#) that the grade force can be in either direction. The sum of the retardation forces,  $F_r$  is made up of rolling resistance, curving resistance or curve drag, air resistance and braking (excluding dynamic braking which is more



**FIGURE 9.1** Three mass train model, where:  $a$  is vehicle acceleration,  $m/sec^2$ ;  $c$  is damping constant,  $Nsec/m$ ;  $k$  is spring constant,  $N/m$ ;  $m$  is vehicle mass,  $kg$ ;  $v$  is vehicle velocity,  $m/sec$ ;  $x$  is vehicle displacement,  $m$ ;  $F_g$  is gravity force components due to track grade,  $N$ ;  $F_r$  is sum of retardation forces,  $N$ ; and  $F_{t/db}$  is traction and dynamic brake forces from a locomotive unit,  $N$ .

conveniently grouped with locomotive traction in the  $F_{t/db}$  term). Rolling and air resistances are usually grouped as a term known as propulsion resistance,  $F_{pr}$ , making the equation for  $F_r$  as follows:

$$F_r = F_{pr} + F_{cr} + F_b$$

where  $F_{pr}$  is the propulsion resistance;  $F_{cr}$  is the curving resistance; and  $F_b$  is the braking resistance due to pneumatic braking.

The three mass train allows the three different differential equations to be developed. With linear wagon connection models the equations can be written as:

$$m_1 a_1 + c_1(v_1 - v_2) + k_1(x_1 - x_2) = F_{t/db} - F_{r1} - F_{g1} \tag{9.1}$$

$$m_2 a_2 + c_1(v_2 - v_1) + c_2(v_2 - v_3) + k_1(x_2 - x_1) + k_2(x_2 - x_3) = -F_{r2} - F_{g2} \tag{9.2}$$

$$m_3 a_3 + c_2(v_3 - v_2) + k_2(x_3 - x_2) = -F_{r3} - F_{g3} \tag{9.3}$$

Note that a positive value of  $F_g$  is taken as an upward grade, i.e., a retarding force.

Allowing for locomotives to be placed at any train position and extending equation notation for a train of any number of vehicles, a more general set of equations can be written as:

For the lead vehicle:

$$m_1 a_1 + c_1(v_1 - v_2) + k_1(x_1 - x_2) = F_{t/db1} - F_{r1} - F_{g1} \tag{9.4}$$

For the  $i$ th vehicle:

$$m_i a_i + c_{i-1}(v_i - v_{i-1}) + c_i(v_i - v_{i+1}) + k_{i-1}(x_i - x_{i-1}) + k_i(x_i - x_{i+1}) = F_{t/dbi} - F_{ri} - F_{gi} \tag{9.5}$$

For the  $n$ th or last vehicle:

$$m_n a_n + c_{n-1}(v_n - v_{n-1}) + k_{n-1}(x_n - x_{n-1}) = F_{t/dbn} - F_{rn} - F_{gn} \tag{9.6}$$

By including the  $F_{t/db}$  in each equation, thus on every vehicle, the equations can be applied to any locomotive placement or system of distributed power. For unpowered vehicles  $F_{t/db}$  is set to zero.

For nonlinear modelling of the system, the stiffness and damping constants are replaced with functions. It is usual to express stiffness as a function of displacement and incorporate coupler slack and piece-wise-linear approximations of draft gear response. Damping is usually expressed as a function of velocity. More complex functions, incorporating a second independent variable, (i.e., displacement and velocity for a stiffness function), can also be used. The generalised nonlinear equations are therefore:

For the lead vehicle:

$$m_1 a_1 + f_{wc}(v_1, v_2, x_1, x_2) = F_{t/db1} - F_{r1} - F_{g1} \quad (9.7)$$

For the  $i$ th vehicle:

$$m_i a_i + f_{wc}(v_i, v_{i-1}, x_i, x_{i-1}) + f_{wc}(v_i, v_{i+1}, x_i, x_{i+1}) = F_{t/dbi} - F_{ri} - F_{gi} \quad (9.8)$$

For the  $n$ th or last vehicle:

$$m_n a_n + f_{wc}(v_n, v_{n-1}, x_n, x_{n-1}) = F_{t/dbn} - F_{rn} - F_{gn} \quad (9.9)$$

where  $f_{wc}$  is the nonlinear function describing the full characteristics of the wagon connection.

Solution and simulation of the above equation set is further complicated by the need to calculate the forcing inputs to the system, i.e.,  $F_{t/db}$ ,  $F_r$ , and  $F_g$ . The traction-dynamic brake force term  $F_{t/db}$  must be continually updated for driver control adjustments and any changes to locomotive speed. The retardation forces,  $F_r$ , are dependent on braking settings, velocity, curvature, and rollingstock design. Gravity force components,  $F_g$ , are dependent on track grade and, therefore, the position of the vehicle on the track. Approaches to the nonlinear modelling of the wagon connection and modelling of each of the forcing inputs are included and discussed in the following sections.

## B. WAGON CONNECTION MODELS

Perhaps the most important component in any longitudinal train simulation is the wagon connection element. The autocoupler with friction type draft gears is the most common wagon connection in the Australian and North American freight train systems. It also, perhaps, presents the most challenges for modelling and simulation due to the nonlinearities of air gap (or coupler slack), draft gear spring characteristic, (polymer or steel), and stick-slip friction provided by a wedge system. Due to these complexities, the common autocoupler-friction type draft gear wagon connection will be examined first. Other innovations such as slackless packages, drawbars, and shared bogies are then more easily considered.

## 1. Conventional Autocouplers and Draft Gear Packages

A conventional autocoupler and draft gear package is illustrated in the schematic in Figure 9.2. A schematic of the wedge arrangement of the draft gear unit is included in Figure 9.3. Variations on the arrangement shown in Figure 9.3 exist. Some designs include an additional taper in the housing or are provided by additional wedges as shown in Figure 9.4. The stick–slip nature of the friction wedges has also led to recent innovations such as those shown in Figure 9.5, which include a release spring. In a design of this type, the release spring is provided to unlock the outside wedge thereby releasing the friction wedges.

When considering a wagon connection, two autocoupler assemblies must be considered along with gap elements, and also stiffness elements describing flexure in the wagon body. A wagon connection model will therefore appear as something similar to the schematic in Figure 9.6. Modelling the coupler slack is straightforward, a simple dead zone. Modelling of the steel components including wagon body stiffness can be provided by a single linear stiffness. Work by Duncan and Webb<sup>1</sup> from test data measured on long unit trains identified cases where the draft gear wedges locked and slow sinusoidal vibration was observed. The behaviour was observed in distributed power trains when the train was in a single stress state. The train could be either in a tensile or compressed condition. The stiffness corresponding to the fundamental vibration mode observed was defined as the locked stiffness of the wagon connections. The locked stiffness value for the trains tested, (consisting of 102 coal hopper cars each of 80 tonne gross mass), was nominally in the order of 80 MN/m.<sup>1</sup> As the locked stiffness is the limiting stiffness of the system, it must be incorporated into the wagon connection model. The locked stiffness is the sum of all the stiffness' added in series, which includes the components such as the coupler shank, knuckle, yoke, locked draft gear, and wagon body. It also includes any pseudo-linear stiffness due to gravity and bogie steer force components, whereby a longitudinal force is resisted by gravity as a wagon is lifted or forced higher on a curve. The limiting stiffness of a long train may therefore vary for different wagon loadings and on-track placement.

Wagon connection modelling can be simplified to a combined draft gear package model equivalent to two draft gear units and includes one spring element representing locked or limiting stiffness, Figure 9.7.

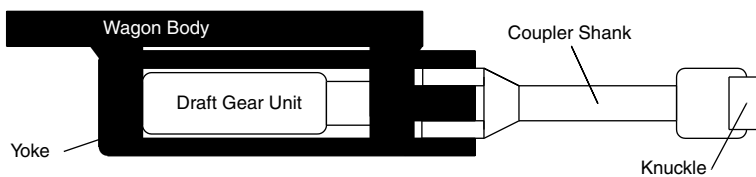


FIGURE 9.2 Conventional autocoupler assembly.

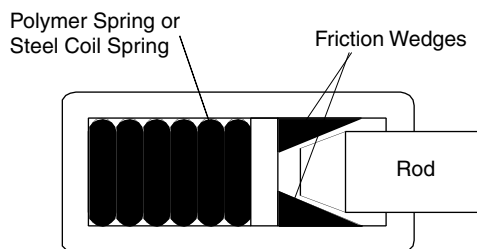


FIGURE 9.3 Friction type draft gear unit.

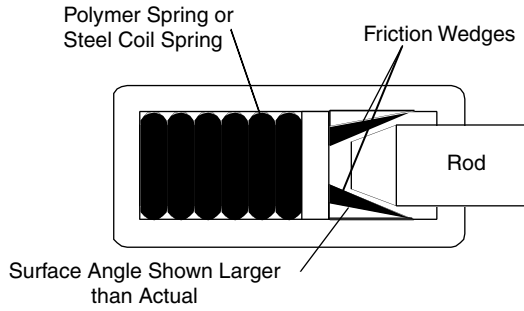


FIGURE 9.4 Friction type draft gear unit with angled surfaces.

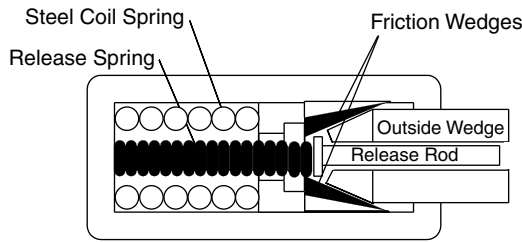


FIGURE 9.5 Friction type draft gear with release spring.

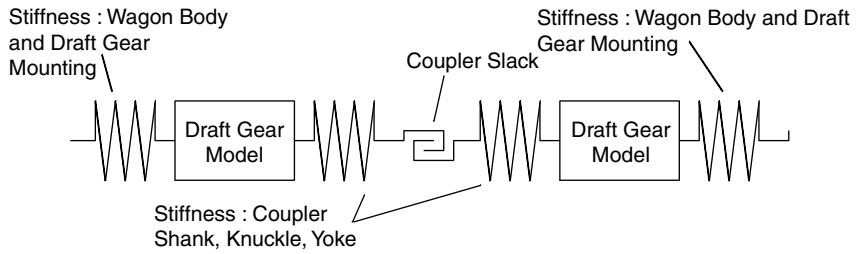


FIGURE 9.6 Components in a wagon connection model.

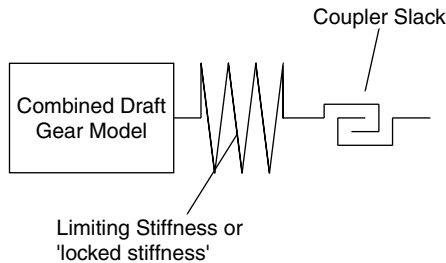


FIGURE 9.7 Simplified wagon connection model.

Determination of the mathematical model for the draft gear model has received considerable attention in technical papers. For the purposes of providing a model for train simulation, a piecewise linear model representing the hysteresis in the draft gear friction wedge (or clutch) mechanism is usually used.<sup>1,9</sup> The problem of modelling the draft gear package has been approached in several

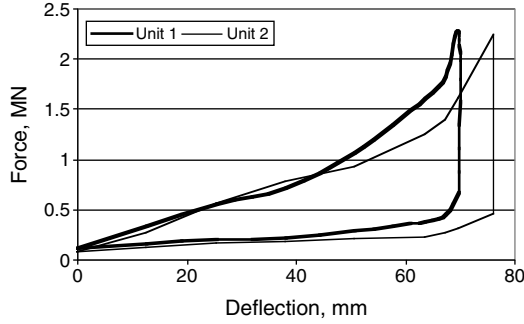


FIGURE 9.8 Typical manufacturer’s draft gear response data.

ways. In early driver training simulators when computing power was limited, it was common practice to further reduce the complexity of the dynamic system by lumping vehicle masses together and deriving equivalent connection models. As adequate computational capacities are now available it is normal practice to model each wagon in detail.<sup>9,10</sup> It would seem reasonable in the first instance to base models on the hysteresis published for the drop hammer tests of draft gear units. Typical draft gear response curves are shown in Figure 9.8.

The first thing to remember is that the published data, as shown in Figure 9.9, represents the extreme operating behaviour simulated by a drop hammer test. The drop hammer of 12.27 tonne (27,000 lb) impacts the draft gear at a velocity of 3.3 m/sec, this simulating an inter-wagon impact with a relative velocity between wagons of 6.6 m/sec, (23.8 km/h). In normal train operation it would be hoped that such conditions are quite rare. Data recording of in-train forces of unit trains in both iron ore and coal haulage systems in Australia revealed that draft gear stiffness in normal operation could be very different from that predicted by drop hammer test data.<sup>1,9</sup> The approach taken by Duncan and Webb<sup>1</sup> was to fit a model to the experimental model, as shown in Figure 9.10, using piecewise linear functions.

It will be noted that the model proposed by Duncan and Webb includes the locked stiffness, as discussed earlier. A significant outcome from the train test data reflected in the model in Figure 9.10 was that unloading and loading could occur along the locked curve whenever the draft gear unit was locked. This cyclic loading and unloading could occur at any extension. Data from this program,<sup>1</sup> and later by Cole,<sup>9</sup> confirmed that the draft gear unit would remain locked until the force level reduced to a point close to the relaxation or unloading line. Due to individual friction characteristics, there is considerable uncertainty about where unlocking occurs. In some cases unlocking was observed below the unloading curve.

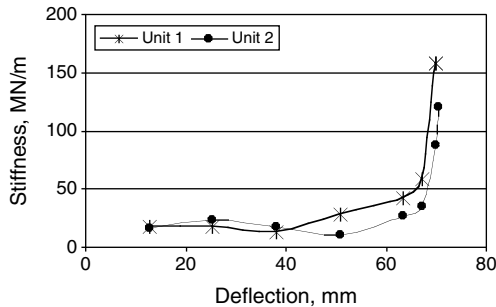
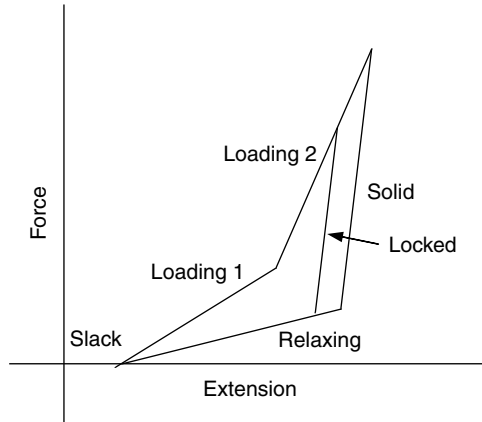


FIGURE 9.9 Draft gear package stiffness–drop hammer tests.



**FIGURE 9.10** Piecewise linear wagon connection model as proposed by Duncan and Webb.<sup>1</sup>

Further refinement of wagon connection modelling was proposed by Cole.<sup>9</sup> The difficulty presented in the work by Duncan and Webb<sup>1</sup> is that draft gear units, and the mathematical models used to represent them, differ depending on the regime of train operation expected. Clearly, if extreme impacts were expected in simulation due to shunting or hump yard operations, a draft gear model representing drop hammer test data would be appropriate. Conversely, if normal train operations were expected, a wagon connection model as proposed in Figure 9.10 would be appropriate. It was noted by Cole<sup>9</sup> that the stiffness of the draft gear units for small deflections varied, typically 5 to 7 times the stiffness indicated by the drop hammer test data, but could be up to 17 times stiffer. The stiffness levels indicated by the units in Figure 9.8 are shown in Figure 9.9. The multiplier of up to 17 times indicates that the stiffness for small deflections could exceed the locked and/or limiting stiffness indicated in experimental data. It is therefore evident that for mild inter-wagon dynamics (i.e., gradual loading of draft gear units) the static friction in the wedge assemblies is large enough to keep draft gears locked. A model incorporating the wedge angles and static and dynamic friction is therefore proposed.

The draft gear package can be considered as a single wedge spring system as shown in Figure 9.11. The rollers provided on one side of the compression rod can be justified in that the multiple wedges are arranged symmetrically around the outside of the rod in the actual unit. It will be realised that different equilibrium states are possible depending the direction of motion, wedge angles, and surface conditions. The free body diagram for increasing load (i.e., compressing) is shown in Figure 9.11. The state of the friction  $\mu_1 N_1$  on the sloping surface can be any value between  $\pm \mu_1 N_1$ . The fully saturated cases of  $\mu_1 N_1$  are drawn on the diagram. If there is sliding action in the direction for compression, then only the Case 1 friction component applies. Case 2 applies if a prejammed state exists. In this case, the rod is held in by the jamming action of the wedge. If the equations are examined, it can be seen that for certain wedge angles and coefficients of friction, wedges are self-locking.

*Examining the rod:*

$$\text{Case 1 : } F_c = N_1(\sin \phi + \mu_1 \cos \phi) \tag{9.10}$$

$$\text{Case 2 : } F_c = N_1(\sin \phi - \mu_1 \cos \phi) \tag{9.11}$$



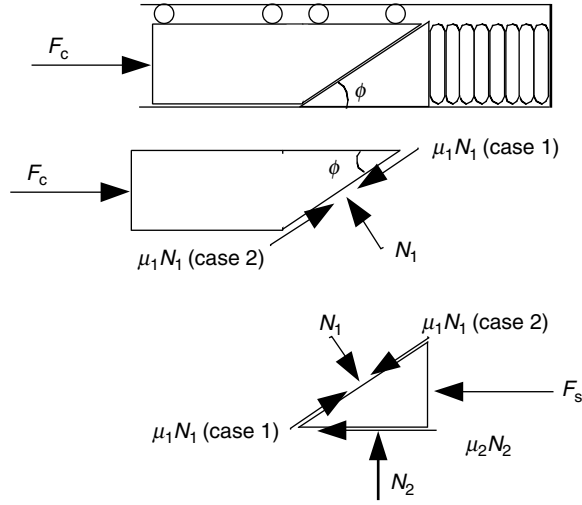


FIGURE 9.11 Free body diagram of a simplified draft gear rod–wedge–spring system.

For self locking,  $N_1$  remains nonzero when  $F_c$  is removed therefore:

$$\sin \phi = \mu_1 \cos \phi$$

i.e., if  $\sin \phi < \mu_1 \cos \phi$ , then a negative force  $F_c$  is required to extend the rod.

From this inequality it can be seen that for self locking:

$$\tan \phi < \mu_1$$

The relationship between wedge angle and friction coefficient can therefore be plotted as shown in Figure 9.12.

Further insight can be gained if the equations relating the wedge forces to the coupler force and polymer spring force are developed, again assuming saturated friction states and direction shown in Case 1 for  $\mu_1 N_1$  giving:

$$F_c = F_s(\mu_1 \cos \phi + \sin \phi) / [(\mu_1 - \mu_2)\cos \phi + (1 + \mu_1 \mu_2)\sin \phi] \tag{9.12}$$

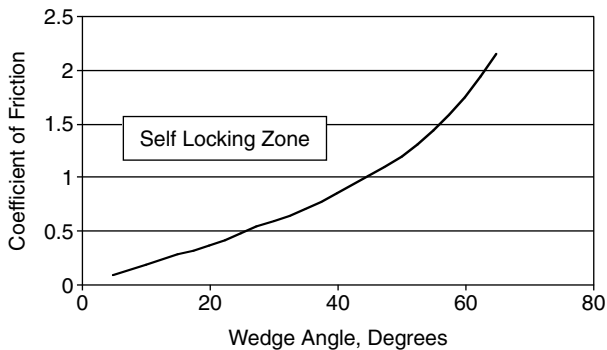


FIGURE 9.12 Friction wedge self locking zone.

If it is assumed that  $\mu_1 = \mu_2$ , and that both surfaces are saturated, then the equation reduces to:

$$F_c = F_s(\mu \cot \phi + 1)/(1 + \mu^2) \tag{9.13}$$

The other extreme of possibility is when there is no impending motion on the sloping surface due to the seating of the rod and wedge, the value assumed for  $\mu_1$  is zero, Equation 9.10 reducing to:

$$F_c = F_s \tan \phi / [\tan \phi - \mu_2] \tag{9.14}$$

If the same analysis is repeated for the unloading case, a similar equation results,

$$F_c = F_s \tan \phi / [\tan \phi + \mu_2] \tag{9.15}$$

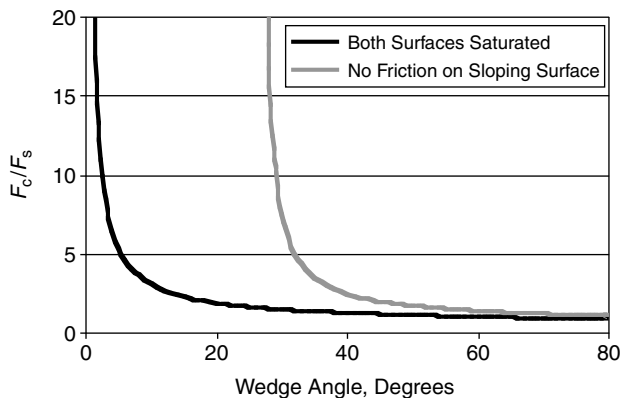
At this point, it is convenient to define a new parameter, namely friction wedge factor, as follows:

$$Q = F_c / F_s \tag{9.16}$$

Using the new parameter, the two relationships 9.13 and 9.14 are plotted for various values of  $\phi$  in Figure 9.13.

The above plots illustrate the significance of the sloping surface friction condition. Measured data indicated in Ref. 19 showed that the stiffness of draft gear packages can reach values  $\sim 7$  times the values obtained in drop hammer tests. Assuming the polymer spring force displacement characteristic is of median slope between loading and unloading curves, the friction wedge factor required will be  $Q \sim 15$ . From disassembled draft gear packages, wedge angles are known to be in the range of 30 to 50°.

The only aspect of the model that now remains to be completed is the behaviour of the friction coefficient. Estimation of these values will always be difficult due to the variable nature of the surfaces. Surface roughness and wear ensure that the actual coefficients of friction can vary, even on the same draft gear unit, resulting in different responses to drop hammer tests. It is also difficult to estimate the function that describes the transition zone between static and minimum kinetic friction conditions and the velocity at which minimum kinetic friction occurs. For simplicity and a first approximation, a piece-wise-linear function can be used as shown in Figure 9.14.



**FIGURE 9.13** Friction wedge factor for  $\mu = 0.5$ .

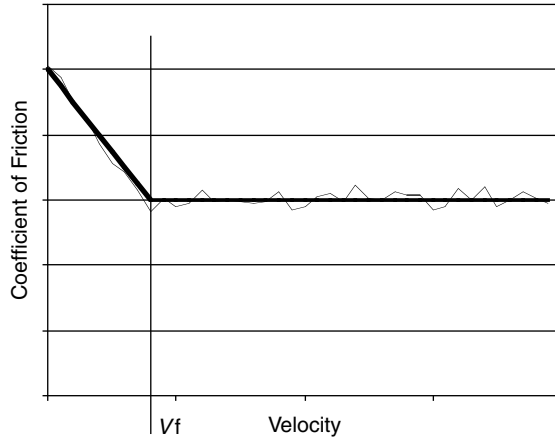


FIGURE 9.14 Piece-wise-linear approximation of wedge friction coefficient.

The friction coefficient  $\mu$  was therefore given by:

$$\begin{aligned} \mu &= \mu_s && \text{for } v = 0 \\ \mu &= \mu(v) && \text{for } 0 < v < V_f \\ \mu &= \mu_k && \text{for } v \geq V_f \end{aligned} \tag{9.17}$$

where  $\mu(v)$  can be any continuous function linking  $\mu_s$  and  $\mu_k$ . Key data for the model therefore becomes, wedge angle  $\phi$ , kinetic friction velocity  $V_f$ , static friction coefficient  $\mu_s$  and kinetic coefficient of friction  $\mu_k$ . and the spring force  $F_s$ . If the assumption is taken that there is no impending motion on the sloping wedge surface and that the  $\mu_1 N_1$  term is small, Equation 9.14 and Equation 9.15 can be used as a starting point for a draft gear model. Alternatively the more complex Equation 9.12, could be used, but it will be shown that sufficient model flexibility will be achieved using the simplified Equation 9.14 and Equation 9.15.

By tuning the various parameters, the model can be adjusted to match both the drop hammer test data and mild impact data from normal train operations. It will be noted in Figure 9.15 that

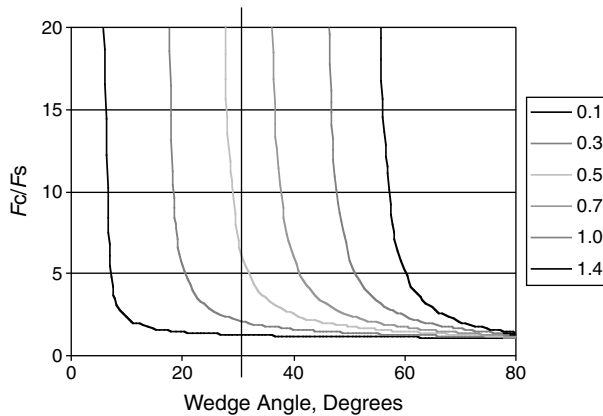


FIGURE 9.15  $F_c/F_s$  ratios for various coefficients of friction.

various coefficients for  $\mu_s$  and  $\mu_k$  can be selected to adjust the span of the model. The difference between loading and unloading curves is determined by the wedge geometry and friction coefficient, Equation 9.14 and Equation 9.15. The nonlinearity of the polymer or steel draft gear springs can be modelled by a piecewise linear for spring force,  $F_s$ . The difference in deflection noted between impact and gradual-loading conditions can be adjusted by selection of friction coefficient parameters, Figure 9.16. Values for wedge angle,  $\phi$ , can be manipulated to increase or decrease the size of the hysteresis, Figure 9.17. The friction parameters can be manipulated to obtain the trajectories of the upper curve that are desired to fit with measured data, Figure 9.18. Having reached this point, a comprehensive wagon connection model can be implemented either by combining two draft gear models, a slack element and a locked stiffness element, or by setting up

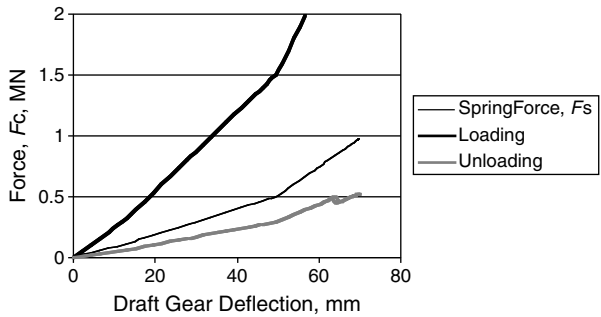


FIGURE 9.16 Sample draft gear wedge model output.

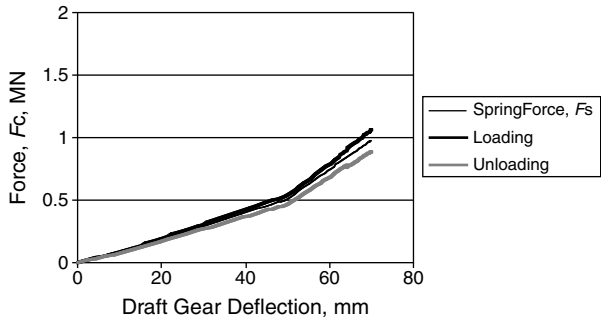


FIGURE 9.17 Effect of increased wedge angle.

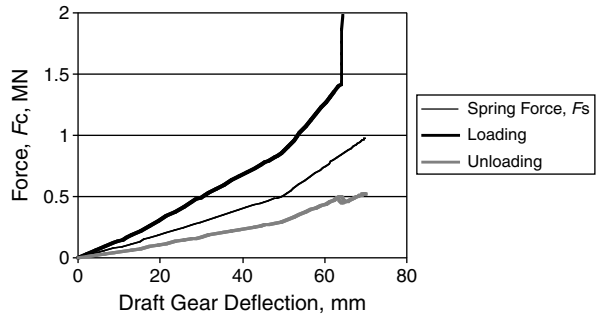
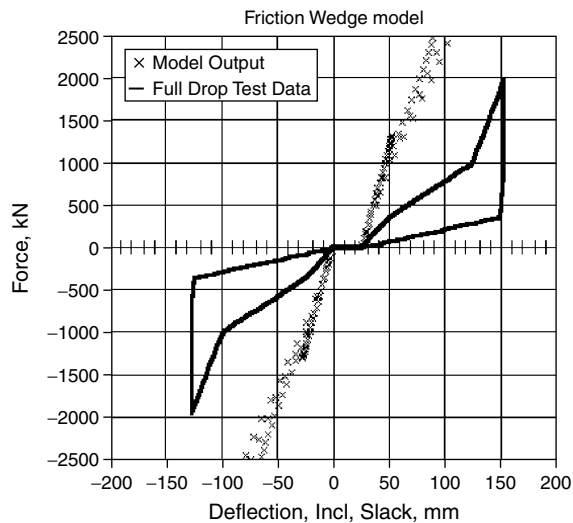


FIGURE 9.18 Effect of lowering kinetic friction coefficient.

a single look up table representing the two draft gear springs in series and then tuning draft gear model parameters to suit a draft gear pair. The slack element can be added either in the look up table or added in series. Different parameters can be chosen for loading and unloading curves. The small kick in the unloading curve is observed in some test data. The dynamicist can also implement slightly different values of,  $\phi$ ,  $V_f$ ,  $\mu_s$  and  $\mu_k$  for the unloading curve if required to obtain a good fit to the experimental data.

There is always room for debate as to whether a complex model as described here is justified when compared to the simpler yet detailed work by Duncan and Webb.<sup>1</sup> The user may decide on the complexity of the model according to the purpose and accuracy required for the simulation studies being completed. While the wedge friction model adjusts for different impact conditions, its use is really justified for simulations where these conditions are expected to vary. The use of the wedge model for the unloading curve is an area where a simple lookup table may suffice, as it is only the loading curve data that shows large variations in stiffness. The following figures show the response of the model with sine wave inputs with frequencies of 0.1, 1.0, and 10.0 Hz. This frequency range covers both normal train operation and loose shunt impact conditions, Figures 9.19 to 9.21, inclusive. These results were obtained by applying the friction wedge model only to the loading curve.<sup>9</sup>

Train simulations using this model are given in Figure 9.22 and Figure 9.23. Force and acceleration traces are plotted for train positions: first and last connected couplers and wagons positioned at intervals of 20% of train length. Both simulations are for a distributed power train for which a throttle disturbance is added at time = 38 sec. The first result showing sinusoidal locked behaviour has the train situated on a crest with the top of the crest situated in the first wagon group. The second result uses exactly the same control input on flat track. It will be noted in Figure 9.22 that while longitudinal forces behave sinusoidally between time = 40 sec and time = 100 sec, wagon accelerations are steady demonstrating locked draft gear behaviour. This is contrasted with the oscillatory nature of wagons longitudinal behaviour for the same period and control input on flat track, Figure 9.23. The difference between what Duncan and Webb referred to as sustained longitudinal vibration and cycle vibration can be identified in Figure 9.22 and Figure 9.23, respectively.



**FIGURE 9.19** Draft gear model response — slow loading (0.1 Hz).<sup>9</sup> Source: From Cole, C., Improvements to wagon connection modelling for longitudinal train simulation, *Conference on Railway Engineering*, Rockhampton, Institution of Engineers, Australia, pp. 187–194, 1998. With permission.

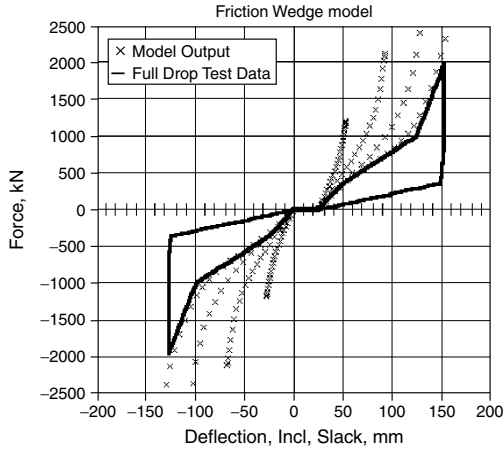


FIGURE 9.20 Draft gear model response — mild impact loading (1 Hz).<sup>9</sup> Source: From Cole, C., Improvements to wagon connection modelling for longitudinal train simulation, *Conference on Railway Engineering*, Rockhampton, Institution of Engineers, Australia, pp. 187–194, 1998. With permission.

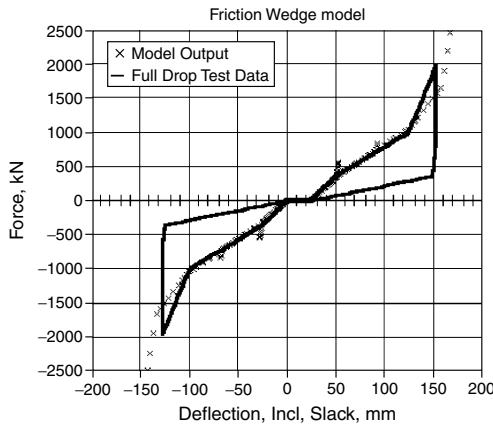


FIGURE 9.21 Draft gear model response — shunt impact (10 Hz).<sup>9</sup> Source: From Cole, C., Improvements to wagon connection modelling for longitudinal train simulation, *Conference on Railway Engineering*, Rockhampton, Institution of Engineers, Australia, pp. 187–194, 1998. With permission.

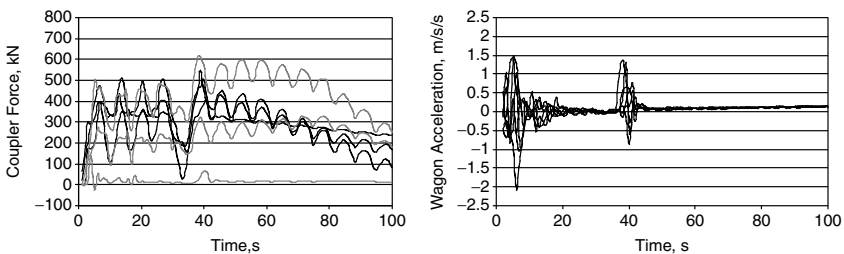


FIGURE 9.22 Simulation results showing “locked” draft gear behaviour — crest track.

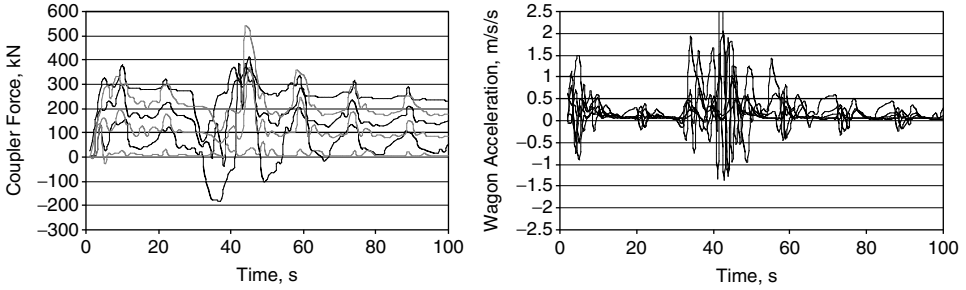


FIGURE 9.23 Simulation results showing “unlocked” draft gear behaviour — flat track.

### 2. Slackless Packages

Slackless draft gear packages are sometimes used in bar-coupled wagons or integrated into shared bogie designs. The design of slackless packages is that the components are arranged to continually compensate for wear to ensure that small connection clearances do not get larger as the draft gear components wear. Slackless packages have been deployed in North American train configurations such as the trough train<sup>11</sup> and bulk product unit trains.<sup>12</sup> The advantage of slackless systems is found in reductions in longitudinal accelerations and impact forces of up to 96 and 86%, respectively as reported in.<sup>11</sup> Disadvantages lie in the inflexibility of operating permanently coupled wagons and the reduced numbers of energy absorbing draft gear units in the train. When using slackless coupled wagon sets, it is usual that the autocouplers at each end are equipped with heavier duty energy absorbing draft gear units. The reduced capacity of these train configurations to absorb impacts can result in accelerated wagon body fatigue or even impact related failures during shunting impacts. Modelling slackless couplings is simply a linear spring limited to a maximum stiffness appropriate to the coupling type, wagon body type, and wagon loading. A linear damper of very small value should be added to approximate small levels of damping available in the connection from friction in pins, movement in bolted or riveted plates, etc. (Figure 9.24).

### 3. Drawbars

Drawbars refer to the use of a single link between draft gear packages in place of two auto couplers. Drawbars can be used with either slackless or energy absorbing draft gear packages. The most recent fleet of coal wagons commissioned in Queensland utilises drawbars with energy absorbing dry friction type draft gear packages. In this case, wagons are arranged in sets of two with

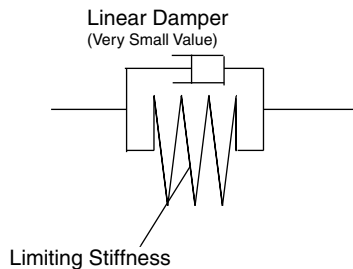
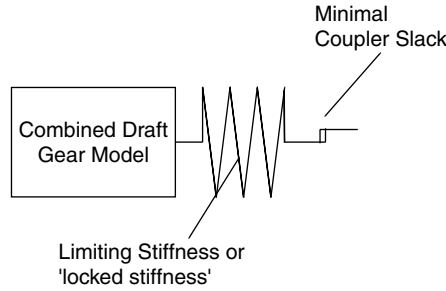


FIGURE 9.24 Wagon connection model — slackless connection.



**FIGURE 9.25** Wagon connection model — drawbar coupled wagon.

conventional autocouplers at either end. Drawbar connections, which connect to energy absorbing draft gear, have the advantage of retaining full capability to absorb impact energy. Modelling drawbars with energy absorbing draft gear units is simply a matter of removing most of the coupler slack from the model, Figure 9.25.

### C. LOCOMOTIVE TRACTION AND DYNAMIC BRAKING

When developing a train model it is logical to treat tractive effort and dynamics braking in the same mathematical model, as both introduce forces to the train via the locomotive–wagon connections. The modelling of locomotive traction/dynamic brake systems is a subject in itself. The complexity of the model required will depend on the particular aspect(s) of locomotive performance that are important for analysis and/or how complex the installed locomotive control systems are. Modern locomotive design has incorporated many performance improvement features. For a fully detailed model, the following may need consideration:

- Torque derating due to thermal effects.
- Limited power application control (pollution control).
- Adhesion limit.
- Traction slip controls.
- Steerable traction bogies.
- Extended range dynamic braking.

The traction control, known as throttle notch, is used to set a current reference. Typically, diesel electric locomotives have eight notches or levels of throttle adjustment. Fully electric haulage locomotives may have differing control systems, e.g., in Australia there are electric locomotives in service with 31 notches (i.e., 32 control positions, 2<sup>5</sup>). At low speeds the traction system is limited by current so tractive effort is applied proportionally to throttle notch levels. Tractive force delivered in this region may be independent of speed, or reduce with speed, depending on the locomotive characteristics and control. At higher speeds the system is limited by power so the tractive effort available decreases at increased speeds according to force velocity product  $P = F_{t/db}v$ . An example of a typical locomotive performance curve is given in Figure 9.26. It will also be noticed that because the control is a current reference, the power curve is proportional to the square of the throttle notch.

A typical equation set for modelling tractive effort would be:

$$\text{For } F_{t/db}v < (N^2/64)P_{\max} \quad F_{t/db} = (N/8)Te_{\max} - k_f v \tag{9.18}$$



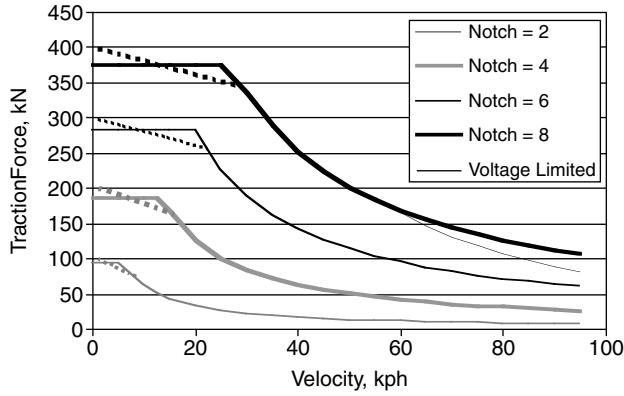


FIGURE 9.26 Typical tractive effort performance curves — diesel electric.

$$\text{Else } F_{t/db} = (N^2/64)P_{\max}/v \tag{9.19}$$

where  $N$  is the throttle setting in notches, 0 to 8;  $P_{\max}$  is the maximum locomotive traction horsepower,  $W$ ;  $Te_{\max}$  is the maximum locomotive traction force,  $N$ ; and  $k_f$  is the torque reduction,  $N/(m/sec)$ .

While a reasonable fit to the published power curves may be possible with a simple equation of the form  $P = F_{t/db}v$ , it may be necessary to modify this model to reflect further control features or reflect changes in efficiency or thermal effects at different train speeds. It is common for the traction performance characteristic to fall below the power curve  $P_{\max} = F_{t/db}v$  at higher speeds due to limits imposed by the generator maximum voltage. Enhanced performance closer to the power curve at higher speeds is achieved on some locomotives by adding a motor field weakening control.<sup>13</sup> It can be seen that accurate modelling of locomotives, even without the need to understand the electrical detail, can become quite complicated. In all cases the performance curves should be viewed and as much precise detail as possible should be obtained about the control features to ensure the development of a suitable model.

It is typical for locomotive manufacturers to publish both the maximum tractive effort and the maximum continuous tractive effort. The maximum continuous tractive effort is the traction force delivered at full throttle notch after the traction system has heated to maximum operating temperature. As the resistivity of the windings increase with temperature, motor torque, which is dependent on current, decreases. As traction motors have considerable mass, considerable time is needed for the locomotive motors to heat and performance levels drop to maximum continuous tractive effort. A typical thermal derating curve for a modern locomotive is shown in Figure 9.27.

Manufacturer’s data from which performance curves such as in Figure 9.26 are derived can usually be taken to be maximum rather than continuous values. If the longitudinal dynamics problem under study has severe grades, and locomotives are delivering large traction forces for long periods, it will be necessary to modify the simple model represented in Figure 9.26 with a further model adding these thermal effects.

A recent innovation in locomotive control is the inclusion of a power application rate limit. The effect of this control is that the power (or dynamic brake) can be applied no faster than a preset rate by the manufacturer irrespective of how fast the driver sweeps the control. Opinions differ as to whether this system was included as an innovation to reduce train dynamics or due to engine design considerations. Records in the Australian patent office identify the system as a pollution

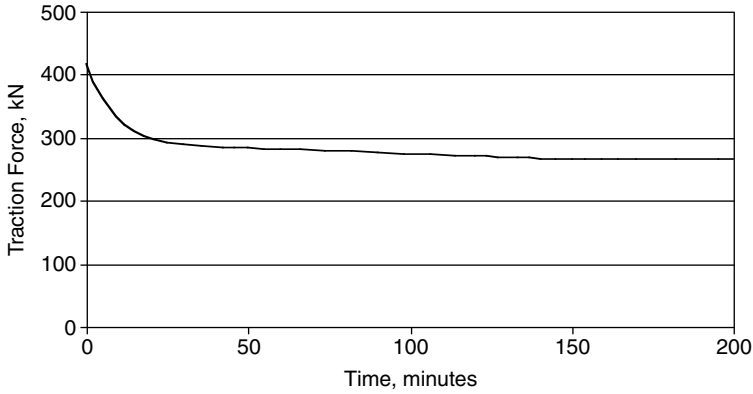


FIGURE 9.27 Tractive effort thermal derating curve.

control — slowing the rate at which the throttle can be applied reduces smoke emissions. At least one fleet of locomotives in Australian service have the application of full power limited to a period not shorter than 80 sec, rate limit being 1.25%/sec. The application of power limited to this rate has a significant effect on the train dynamics and the way trains like this are driven. It therefore must be superimposed on the traction force model.

A key parameter in any discussion about tractive effort is rail–wheel adhesion or the coefficient of friction. Prior to enhancement of motor torque control, a rail–wheel adhesion level of  $\sim 0.20$  could be expected. With modern locomotive traction control, higher values of adhesion reaching  $\sim 0.35$  are obtained with manufacturers claiming up to 0.46 in published performance curves. It needs to be remembered that a smooth control system can only deliver an adhesion level up to the maximum set by the coefficient of friction for the wheel–rail conditions. Wheel–rail conditions in frost and snow could reduce adhesion to as low as 0.1. Superimposing adhesion levels on Figure 9.26, as shown in Figure 9.28, shows how adhesion is significant as a locomotive performance parameter.

The use of dynamic brakes as a means of train deceleration has continued to increase as dynamic brake systems have been improved. Early systems, as shown in Figure 9.29, gave only a variable retardation force and were not well received by drivers. As the effectiveness was so

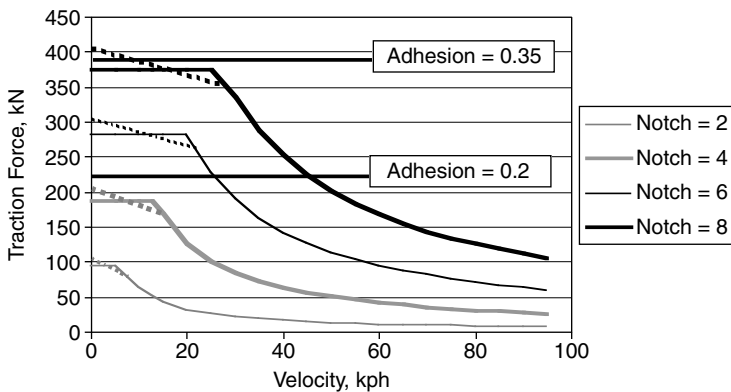


FIGURE 9.28 Tractive effort performance curves — showing effect of adhesion levels.

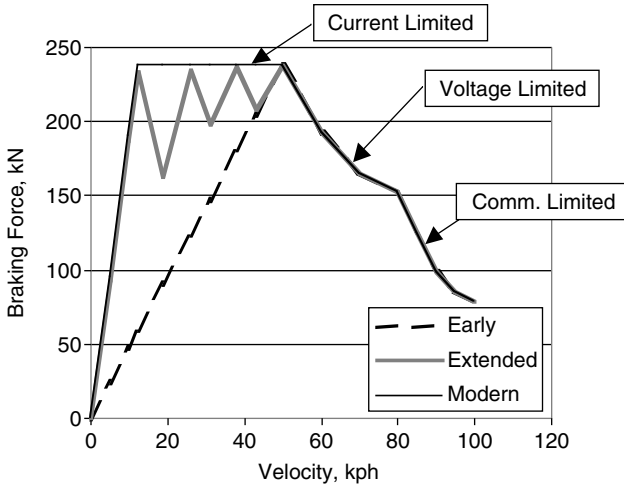


FIGURE 9.29 Dynamic brake characteristics.

dependent on velocity, the use of dynamic brakes gave unpredictable results unless a mental note was made of locomotive velocity and the driver was aware of what performance to expect. Extended range systems, which involved switching resistor banks, greatly improved dynamic brake usability on diesel electric locomotives. More recent locomotive packages have provided large regions of maximum retardation at steady force levels. The performance of the dynamic brake is limited at higher speeds by current, voltage, and commutator limits. Performance at low speeds is limited by the motor field. Designers now try to achieve full dynamic brake force at as low a velocity as possible. Recent designs have achieved the retention of maximum dynamic braking force down to 10 km/h. Dynamic braking is usually controlled as a continuous level rather than a notch, but again some locomotives may provide discrete control levels. The way in which the control level affects the braking effort differs for different locomotive traction packages. Four different dynamic brake characteristics have been identified, but further variations are not excluded, Figure 9.30 and Figure 9.31.

Later designs (shown on the left in Figure 9.30 and Figure 9.31) provide larger ranges of speed where a near constant braking effort can be applied. Modelling of the characteristic can be achieved by fitting a piecewise linear function to the curve, representing 100% dynamic braking force. The force applied to the simulation can then be scaled linearly in proportion to the control setting. In some configurations it will be necessary to truncate the calculated value by different amounts,

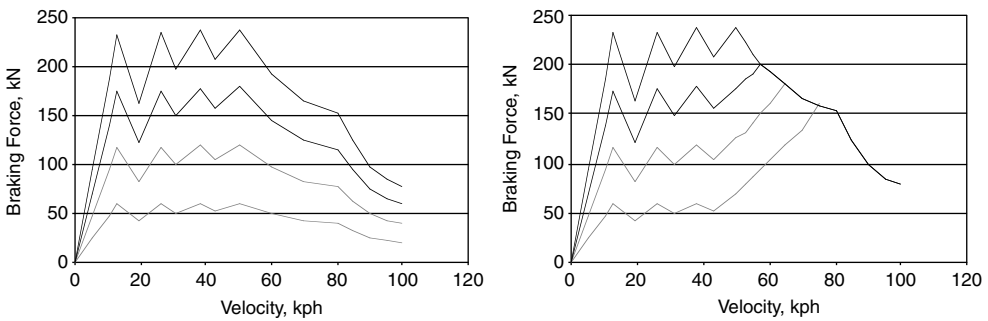


FIGURE 9.30 Dynamic brake characteristics — diesel electric locomotives.

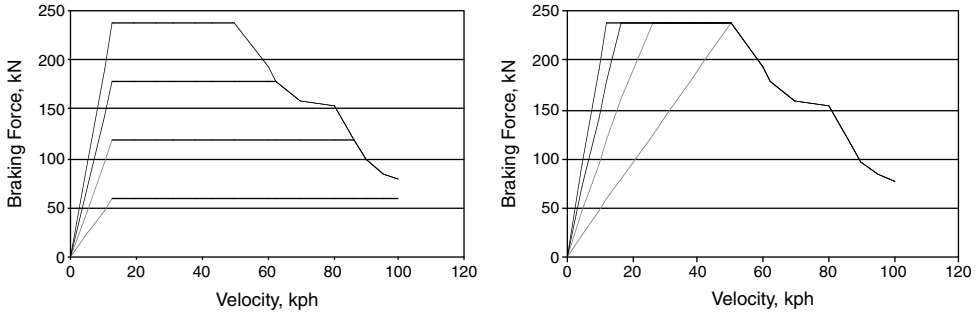


FIGURE 9.31 Dynamic brake characteristics — electric locomotives.

see characteristics on the right hand side of Figure 9.30 and both characteristics in Figure 9.31. In these cases a combination of look up tables and mathematical functions will be required.

### D. PNEUMATIC BRAKE MODELS

The modelling of the brake system requires the simulation of a fluid dynamic system that must run in parallel with the train simulation. The output from the brake pipe simulation is the brake cylinder force, which is converted by means of rigging factors and shoe friction coefficients into a retardation force that is one term of the sum of retardation forces  $F_r$ .

Modelling of the brake pipe and triple valve systems is a subject in itself and therefore will not be treated in this chapter beyond characterising the forces that can be expected and the effect of these forces on train dynamics. The majority of freight rollingstock still utilises brake pipe-based control of the brake system. The North American system differs in design from the Australian/U.K. systems, but both apply brakes sequentially starting from the point where the brake pipe is exhausted. Both systems depend on the fail-safe feature whereby the opening of the brake valve in the locomotive, or the failure of the brake pipe allowing loss of brake pipe pressure, results in the application of brakes in the train.

The implications for train dynamics is that the application of brakes can be accompanied by severe slack action as the brakes nearest the lead of the train, closest to the brake control, apply brakes first. For brakes applied at the lead of a group of wagons 700 m long, the initiation of braking at the last wagon typically lags the lead application by ~5 sec, Figure 9.32. The brake system shown in Figure 9.32 benefits from distributed locomotives allowing the release of air at the lead and mid train positions. The response at the mid point of the first wagon group (Vehicle 26) is faster

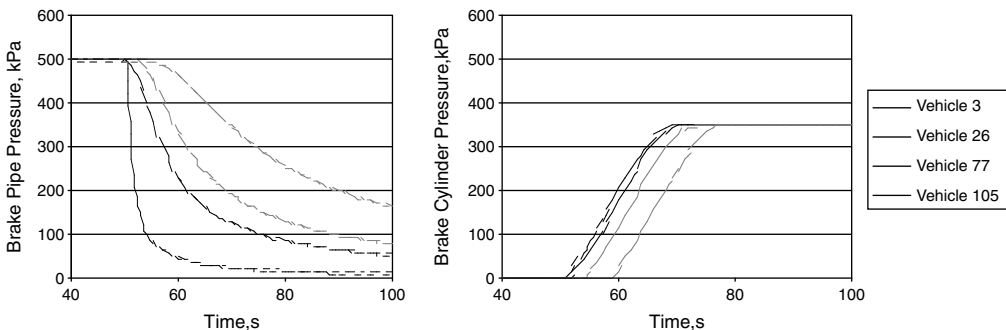


FIGURE 9.32 Brake pipe and cylinder responses — emergency application.

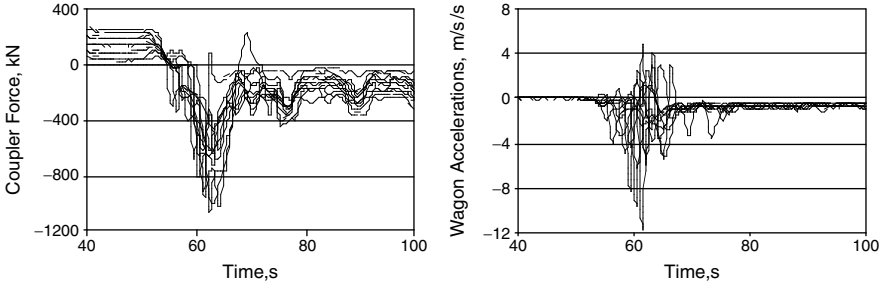


FIGURE 9.33 Coupler forces and wagon accelerations — emergency application.

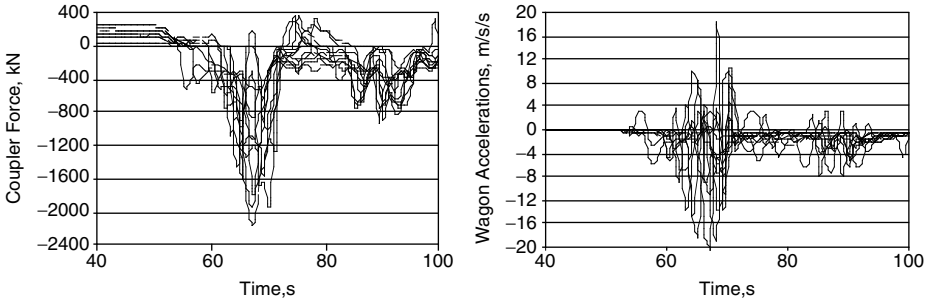


FIGURE 9.34 Coupler forces and wagon accelerations — emergency application — increased slack to 75 mm.

due to the brake pipe exhausting from both ends. The slower responses at positions 77 and 105 are typical of a head end train with only one pipe exhaust point. Coupler forces and associated wagon accelerations for first and last wagon connections and at vehicles at intervals of 10% of train length are shown in Figure 9.33. The same simulation is repeated to obtain the coupler forces in Figure 9.34 with the coupling slack increased from 25 to 75 mm, illustrative of the significance of slack action in brake applications.

### E. GRAVITATIONAL COMPONENTS

Gravitational components,  $F_g$ , are added to longitudinal train models by simply resolving the weight vector into components parallel and at right angles to the wagon body chassis. The parallel component of the vehicle weight becomes  $F_g$ . On a grade, a force will either be added to or subtracted from the longitudinal forces on the wagon, Figure 9.1 and Figure 9.35.

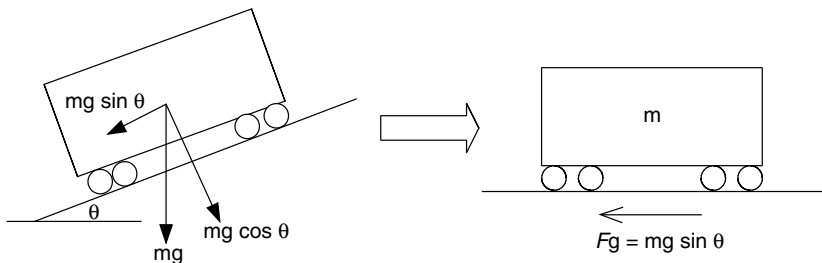
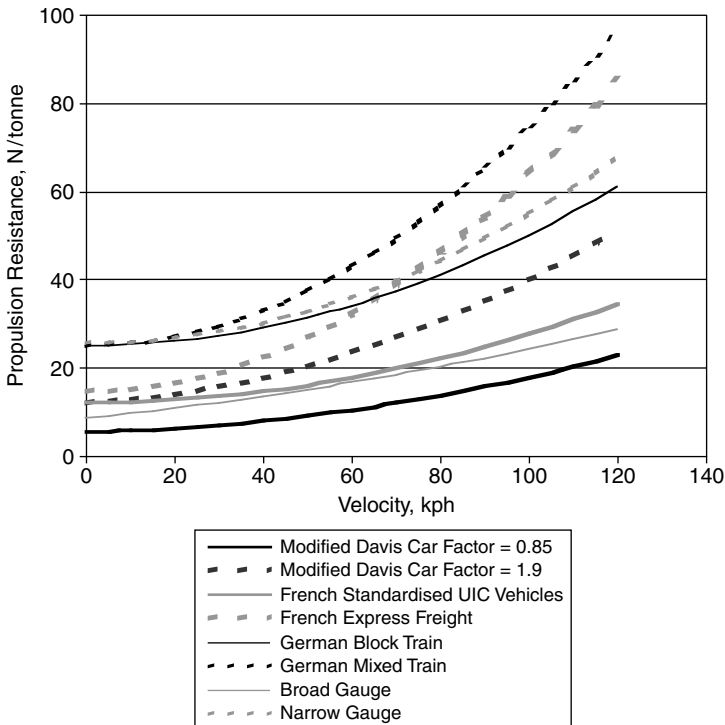


FIGURE 9.35 Modelling gravitational components.

The grade also reduces the sum of the reactions of the wagon downward on the track. This effect has implications for propulsion resistance equations that are dependent on vehicle weight. However, the effect is small and, due to the inherent uncertainty in propulsions resistance calculations, it can be safely ignored. Taking a 1 in 50 grade as an example, gives a grade angle of  $1.146^\circ$ . The cosine of this angle is 0.99979. The reduction in the sum of the normal reactions for a wagon on a 1 in 50 grade (or 2%) is therefore 0.02%. Grades are obtained from track plan and section data. The grade force component must be calculated for each vehicle in the train and updated each time step during simulation to account for train progression along the track section.

**F. PROPULSION RESISTANCE**

Propulsion resistance is usually defined as the sum of rolling resistance and air resistance. In most cases, increased vehicle drag due to track curvature is considered separately. The variable shapes and designs of rollingstock, and the complexity of aerodynamic drag, mean that the calculation of rolling resistance is still dependent on empirical formulae. Typically, propulsion resistance is expressed in an equation of the form of  $R = A + BV + CV^2$ . Hay presents the work of Davis which identifies the term  $A$  as journal resistance dependent on both wagon mass and the number of axles, an equation of the form  $R = ax + b$ , giving in imperial units  $1.3wn + 29n$ , where  $w$  is weight per axle and  $n$  is the number of axles, is quoted in Ref. 14. The second term is mainly dependent on flanging friction and therefore the coefficient  $B$  is usually small (nonexistent in some empirical formulae) and the third term is dependent on air resistance. The forms of propulsion resistance equations used and the empirical factors selected vary between railway systems reflecting the use of equations that more closely match the different types of rollingstock and running speeds (Figure 9.36). An instructive collection of propulsion resistance formulae has been assembled from



**FIGURE 9.36** Propulsion resistance equations compared — freight rollingstock.

**TABLE 9.1**  
**Empirical Formulas for Propulsion Resistance–Freight Rollingstock**

Description	Equation 9.20
Modified Davis equation (U.S.A.)	$K_a[2.943 + 89.2/m_a + 0.0306V + 1.741k_{ad}V^2/(m_a n)]$ $K_a = 1.0$ for pre 1950, 0.85 for post 1950, 0.95 container on flat car, 1.05 trailer on flat car, 1.05 hopper cars, 1.2 empty covered auto racks, 1.3 for loaded covered auto racks, 1.9 empty, uncovered auto racks $k_{ad} = 0.07$ for conventional equipment, 0.0935 of containers and 0.16 for trailers on flatcars
French Locomotives	$0.65m_a n + 13n + 0.01m_a nV + 0.03V^2$
French Standard UIC vehicles	$9.81(1.25 + V^2/6300)$
French Express Freight	$9.81(1.5 + V^2/(2000\dots2400))$
French 10 tonne/axle	$9.81(1.5 + V^2/1600)$
French 18 tonne/axle	$9.81(1.2 + V^2/4000)$
German Strahl formula	$25 + k(V + \Delta V)/10$ $k = 0.05$ for mixed freight trains, 0.025 for block trains
Broad gauge (i.e., 1.676 m)	$9.81[0.87 + 0.0103V + 0.000056V^2]$
Broad gauge (i.e., $\sim 1.0$ m)	$9.81[2.6 + 0.0003V^2]$

$K_a$  is an adjustment factor depending on rollingstock type;  $k_{ad}$  is an air drag constant depending on car type;  $m_a$  is mass supported per axle in tonnes;  $n$  is the number of axles;  $V$  is the velocity in kilometres per hour; and  $\Delta V$  is the head wind speed, usually taken as 15 km/h.

Ref. 14 and work by Profillides.<sup>15</sup> All equations are converted to SI units and expressed as Newtons per tonne mass (see Table 9.1 and Table 9.2).

Even with the number of factors described in Table 9.1 and Table 9.2, the effects of many factors are not, and usually cannot be, meaningfully considered. If the rollingstock design area is considered, how are the instances of poor bogie steer causing wheel squeal quantified? The equations do not include centre bowl friction, warp stiffness or wheel–rail profile information. In the area of air resistance, wagon body design is more variable than suggested by the few adjustment factors presented here. The dynamicist should therefore be aware that considerable differences between calculations and field measurements are probable.

**TABLE 9.2**  
**Empirical Formulas for Propulsion Resistance–Passenger Rollingstock**

Description	Equation 9.21
French passenger on bogies	$9.81(1.5 + V^2/4500)$
French passenger on axles	$9.81(1.5 + V^2/(2000\dots2400))$
French TGV	$2500 + 33V + 0.543V^2$
German Sauthoff Formula Freight (Intercity Express, ICE)	$9.81[1 + 0.0025V + 0.0055((V + \Delta V)/10)^2]$
Broad gauge (i.e., 1.676 m)	$9.81[0.6855 + 0.02112V + 0.000082V^2]$
Narrow gauge (i.e., $\sim 1.0$ m)	$9.81[1.56 + 0.0075V + 0.0003V^2]$

$K_a$  is an adjustment factor depending on rollingstock type;  $k_{ad}$  is an air drag constant depending on car type;  $m_a$  is mass supported per axle in tonnes;  $n$  is the number of axles;  $V$  is the velocity in kilometres per hour; and  $V$  is the head wind speed, usually taken as 15 km/h.

## G. CURVING RESISTANCE

Curving resistance calculations are similar to propulsion resistance calculations in that empirical formulae must be used. Rollingstock design and condition, cant deficiency, rail profile, rail lubrication, and curve radius will all affect the resistance imposed on a vehicle on the curve. As rollingstock design and condition, rail profile, and cant deficiency can vary, it is usual to estimate curving resistance by a function relating only to curve radius. The equation commonly used is<sup>14</sup>:

$$F_{cr} = 6116/R \quad (9.22)$$

where  $F_{cr}$  is in Newtons per tonne of wagon mass and  $R$  is curve radius in metres.

Rail flange lubrication is thought to be capable of reducing curving resistance by 50%. The curving resistance of a wagon that is stationary on a curve is thought to be approximately double, i.e., 200% of the value given by Equation 9.22.

## H. TRAIN DYNAMICS MODEL DEVELOPMENT AND SIMULATION

As can be seen from the preceding sections, the modelling of the train as a longitudinal system involves a range of modelling challenges for the dynamicist. The basic interconnected mass–damper–spring type model, representing the train vehicle masses and wagon connections, is complicated by nonlinear gap, nonlinear spring, and stick slip friction elements. The complexity and detail, which is chosen for models such as the wagon connection element, may limit the choices available in the modelling and simulation software used. Software packages with predefined model blocks and look up tables etc. can usually be used, sometimes with difficulty, to model systems of this complexity. The wagon connection models used by Duncan<sup>1</sup> and Cole<sup>9</sup> were both implemented only as code subroutines. In some cases, a subroutine or function written in a programming language will be easier to develop than a complex combination of re-existing stiffness' and dampers from a software library.

Having developed a suitable connector for the mass–damper–spring, i.e.,  $f_{wc}(v_i, v_{i+1}, x_i, x_{i+1})$ , the remaining subsystems for traction, braking, resistance forces, and control inputs require modelling and data bases must be provided. Again, software packages with predefined modelling features can be used, but code scripts will also usually be required to work with track databases or for more complex models. The pneumatic braking system, not treated in detail in this chapter, will require a complete time stepping simulation of its fluid flow dynamics. The pneumatic braking model must interface with the train simulation model at the locomotive control input subsystem to receive brake control inputs. The output from the brake model, cylinder pressures, must be scaled by cylinder sizes, brake rigging, and brake shoe friction coefficients to give retardation forces which are applied to the vehicle masses. If the brake model is a fully detailed gas dynamics model, it will usually require a much smaller time step than the train mass–damper–spring model. It is not unusual for this problem to be solved by completing several integration steps of brake pipe simulation for every one integration step of the train mass–damper–spring model. Such models are computationally expensive and until recently would only be found in engineering analysis simulators. Many existing rail industry specific train simulation software packages, because of the era in which they were developed, utilise some simplification of the brake model to allow reasonable run times for simulation studies. This is particularly the case for driver training simulators where a design criteria is that the graphics and experience of the simulator must be at real time speed.

Train simulators with highly nonlinear and hard limited connections, as described in this section, can be simulated successfully with explicit schemes such as Fourth Order Runge Kutta. The simulation examples presented in this chapter utilise this solver with a 10 m/sec time step. Some simulations and variations of the wagon connection model have been found to require a slightly smaller step. A discussion of numerical methods is given in Ref. 8. The advantage of the



Runge Kutta scheme is that it is self starting and forward solving. This simplifies the starting of the simulation and reduces number of initial conditions that need to be set.

### III. INTERACTION OF LONGITUDINAL TRAIN AND LATERAL/VERTICAL WAGON DYNAMICS

Traditionally, the study of longitudinal train dynamics has considered wagons as single degree of freedom masses connected with either spring–damper units or nonlinear wagon connection models. The inputs to the system are arranged and applied to the model in the longitudinal direction, i.e., locomotive forces, grade forces, and resistance forces. Similarly, the study of wagon dynamics has focused on lateral and vertical wagon dynamics with the inputs being from the track geometry. Depending on the way a rail system develops, there could be cases where the interaction of train and wagon dynamics should be considered. Interaction has the potential to become a problem when freight train lengths become large, giving rise to both larger steady and impact forces. Impacts can be reduced by reducing coupling slack. Larger steady forces can be reduced by adopting distributed power configurations and appropriate control techniques. It should not be assumed that adopting distributed power alone will reduce all in-train forces. The discussion in Section VI shows that inappropriate use of distributed power can lead to very high in-train forces and pull-aparts. Another key factor is the rate at which rail infrastructure development matches the rollingstock development. When the rail infrastructure is driven by high-speed passenger train requirements, mild curvatures will tend to ensure that lateral components of in-train forces are minimal. Three modes of train–wagon interaction can be considered:

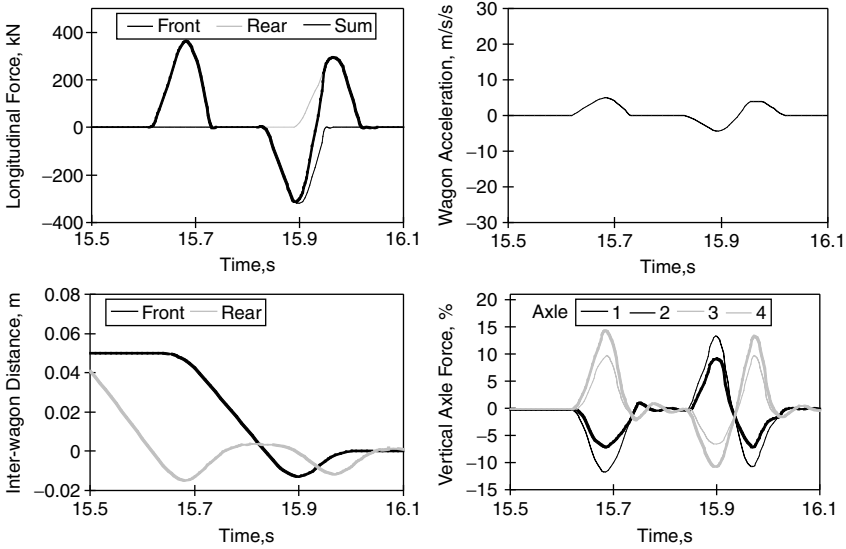
- Wheel unloading on curves due to lateral components of coupler forces.
- Wagon body pitch due to coupler impact forces.
- Bogie pitch due to coupler impact forces.

#### A. WHEEL UNLOADING ON CURVES DUE TO LATERAL COMPONENTS OF COUPLER FORCES

The published work usually referred to in this area is that of El-Sibaie,<sup>4</sup> which presented experimental data and simulation of wheel unloading due to lateral force components in curves. Wheel unloading was shown to increase for: increased in-train forces, decreased curve radius and long–short wagon combinations (i.e., the effect of differing wagon body lengths and bogie overhang distances). The usual method of analysis is to complete a longitudinal train simulation to obtain coupler force data. Coupler angles are then calculated allowing lateral components of coupler forces to be calculated. The lateral force components are then applied to a fully detailed wagon dynamics model to study the resulting wheel unloading and lateral/vertical wheel force ratios.

#### B. WAGON BODY PITCH DUE TO COUPLER IMPACT FORCES

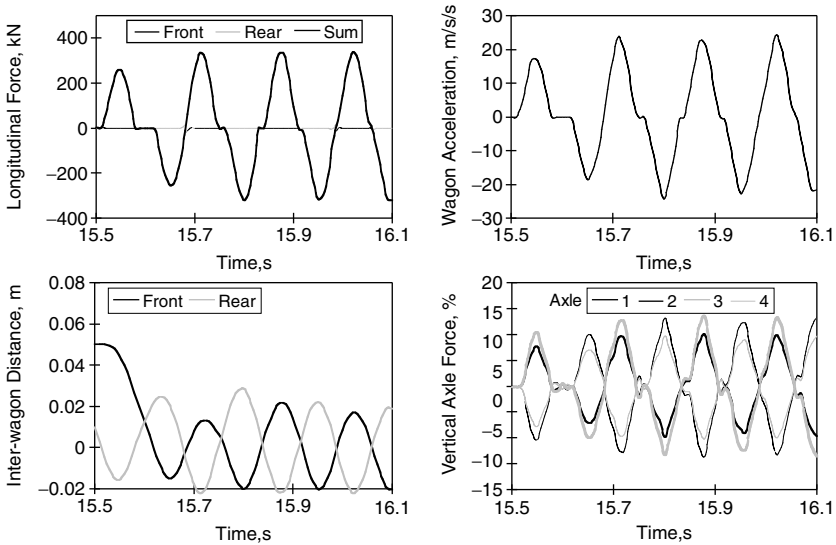
Wagon body pitch can occur in response to a longitudinal impact force and is due to the centre of mass of the wagon being higher than the line of action of the coupler. As this is the mechanism, body pitch is most likely in loaded wagons. Body pitch is unlikely in empty wagons as the centre of mass is usually close to coupler level. Simulation studies using various models and packages were published by McClanachan et al.<sup>5</sup> A sample result from this paper is shown in [Figure 9.37](#). The longitudinal force of  $\sim 380$  kN results in body pitch and at least 10% wheel unloading.



**FIGURE 9.37** Wagon body pitch — loaded in a loaded unit train.<sup>5</sup> Source: From McClanachan M., Cole C., Roach D., and Scown B., *The Dynamics of Vehicles on Roads and on Tracks-Vehicle Systems Dynamics Supplement 33*, Swets & Zeitlinger, Amsterdam, pp. 374–385, 1999. With permission.

### C. BOGIE PITCH DUE TO COUPLER IMPACT FORCES

Similarly, coupler impacts can be sufficient to accelerate or decelerate the wagon so rapidly that the bogies will pitch. This behaviour is most likely to occur in empty wagons. When a wagon is empty, the line of action of the coupling force is close to the same level as the wagon body centre of mass.



**FIGURE 9.38** Bogie pitch — empty wagon in empty unit train.<sup>5</sup> Source: From McClanachan M., Cole C., Roach D., and Scown B., *The Dynamics of Vehicles on Roads and on Tracks-Vehicle Systems Dynamics Supplement 33*, Swets & Zeitlinger, Amsterdam, pp. 374–385, 1999. With permission.

The bogie mass is a significant percentage of empty wagon mass, typically,  $\sim 20\%$  (per bogie) and therefore having significant inertia. Acceleration and deceleration is applied to the bogie at the centre bowl connection, some distance above the bogie centre of mass. The result being that significant wheel unloading, 50%, due to bogie pitch can be both measured and simulated,<sup>5</sup> Figure 9.38. Even worse wheel unloading could be expected for an empty wagon placed in a loaded train where impact conditions can be more severe. The case of an empty wagon in a loaded train combines low wagon mass with larger in-train forces — more severe longitudinal wagon accelerations.

#### IV. LONGITUDINAL TRAIN CRASHWORTHINESS

Crashworthiness is a longitudinal dynamics issue associated with passenger trains. Design requirements of crashworthiness are focused on improving the chances of survival of car occupants. There are two areas of car design related to longitudinal dynamics that require attention and will be mandated by safety authorities in most countries. Passenger cars require:

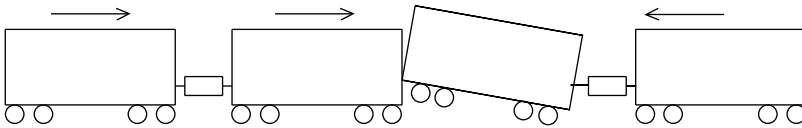
- Vertical collision posts.
- End car crumple zones.

##### A. VERTICAL COLLISION POSTS

The requirement is based on the scenario of a wagon becoming uncoupled or broken away and then climbing the next car. The chassis of the raised wagon, being much stronger than the passenger car upper structure, can easily slice through the car causing fatalities and horrific injuries, Figure 9.39. Design requirements to improve occupant survival include the provision of vertical collision posts that must extend from the chassis or underframe to the passenger car roof, Figure 9.40. Standards will differ depending on the expected running speeds and country of operation. The specification in Australia for operation on the Defined Interstate Rail Network<sup>16</sup> requires the following forces to be withstood without the ultimate material strength being exceeded.

At total longitudinal force of 1100 kN distributed evenly across the collision posts. The force applied 1.65 m above the rail level.

A horizontal shear force of 1300 kN applied to each individual post fitted at a level just above the chassis or underframe.



**FIGURE 9.39** Collision illustrating wagon climb. *Source:* From McClanachan M., Cole C., Roach D., and Scown B., *The Dynamics of Vehicles on Roads and on Tracks-Vehicle Systems Dynamics Supplement 33*, Swets & Zeitlinger, Amsterdam, pp. 374–385, 1999. With permission.



**FIGURE 9.40** Passenger car showing placement of vertical collision posts.

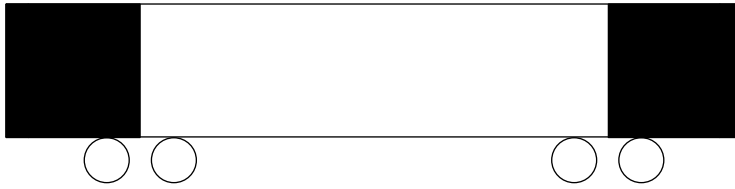


FIGURE 9.41 Passenger car showing crumple zones.

## B. END CAR CRUMPLE ZONES

A further requirement for crashworthiness is energy absorption. Again, standards and specifications will differ depending on the expected running speeds and country of operation. In Australia it is a requirement that energy absorption elements within draft gears will minimise effects of minor impacts. The minimum performance of draft gears is the requirement to accommodate an impact at 15 km/h.<sup>18</sup> The code of practice also requires that cars include unoccupied crumple zones between the headstock and bogie centres to absorb larger impacts by plastic deformation, Figure 9.41.

## V. LONGITUDINAL COMFORT

Ride comfort measurement and evaluation is often focused on accelerations in the vertical and lateral directions. The nature of longitudinal dynamics is that trains are only capable of quite low steady accelerations and decelerations due to the limits imposed by adhesion at the wheel–rail interface. Clearly, the highest acceleration achievable will be that of a single locomotive giving a possible  $\sim 0.3$  g assuming 30% wheel–rail adhesion and driving all wheels. Typical train accelerations are of course much lower, of the order 0.1 to 1.0 m/sec<sup>2</sup>.<sup>15</sup> Braking also has the same adhesion limit but rates are limited to values much lower to prevent wheel locking and wheel flats. Typical train deceleration rates are of the order 0.1 to 0.6 m/sec.<sup>15</sup> The higher values of acceleration and deceleration in the ranges quoted correspond to passenger and suburban trains. The only accelerations that contribute to passenger discomfort or freight damage arise from coupler impact transients. The nature of these events are irregular so frequency spectral analysis and the development of ride indexes are inappropriate in many instances. It is more appropriate to examine maximum magnitudes of single impact events.

For comparison, the maximum acceleration limits specified by various standards that can be applied to longitudinal comfort are plotted in Figure 9.42.

The levels permitted for 1 min exposure are plotted for the fatigue-decreased proficiency boundary (FDPB) and for the reduced comfort boundary (RCB), as per AS 2670 and are plotted in Figure 9.42 to compare with peak or maximum criteria found in other standards. Further insight is gained if the longitudinal oscillations are assumed to be sinusoidal and displacement levels associated with these acceleration levels are also plotted. The displacement amplitudes permitted for various frequencies are plotted in Figure 9.43.

In Australia, the now outdated Railways of Australia (ROA) Manual of Engineering Standards and Practices<sup>17</sup> included calculations of ride index only for vertical and lateral directions. The only reference to longitudinal comfort was a peak limit of 0.3 g (2.943 m/sec) applying to accelerations for all three directions. The 0.3 g limit applied over a bandwidth of 0 to 20 Hz, thereby describing maximum longitudinal oscillation accelerations and displacements in the range of 75 to 0.2 mm in the range of vibration frequencies from 1 to 20 Hz, as shown by Figure 9.42 and Figure 9.43. The newer standard, Code of Practice for the Defined Interstate Rail Network,<sup>18</sup> more specifically excludes the evaluation of longitudinal comfort with peak accelerations specified only for vertical

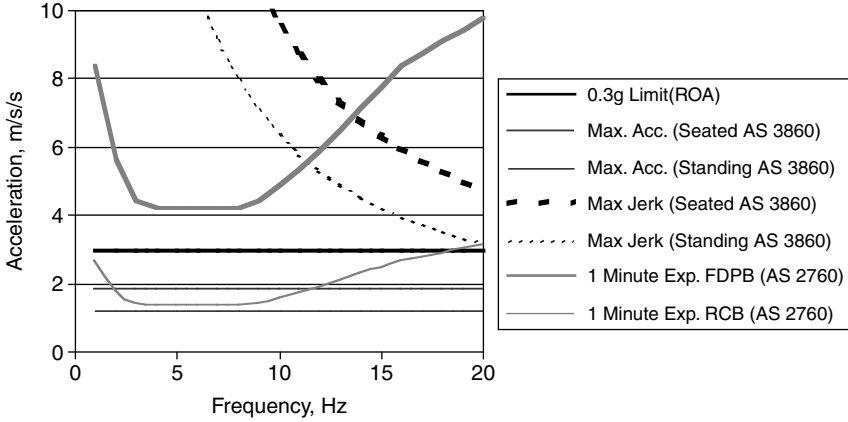


FIGURE 9.42 Passenger comfort acceleration limits.

and lateral dynamics. Both standards refer to Australian Standard AS2670<sup>19</sup> stating that vibration in any passenger seat shall not exceed either the “reduced comfort boundary” (RCB), or the “fatigue decreased proficiency boundary” (FDPB), of AS2670 in any axis. So, while not specifying calculations in the railway standards, a criteria for longitudinal comfort can be drawn from the general Australian Standard, AS 2670. Another Australian standard which is useful when considering longitudinal comfort issues is AS3860 (Fixed Guideway People Movers).<sup>20</sup> This standard gives maximum acceleration limits for sitting and standing passengers. It also gives maximum values for “jerk”, the time derivative of acceleration. A jerk limit slightly lower than the Australian standard AS 3860 of 1.5 m/sec<sup>3</sup> is also quoted by Profillidis.<sup>15</sup>

More elaborate treatment of longitudinal comfort was located in the UIC Leaflet 513, Guidelines for evaluating passenger comfort in relation to vibration in railway vehicles, issued 1/11/2003.<sup>21</sup> The UIC approach integrates longitudinal accelerations into a single parameter. Vertical, lateral, and longitudinal accelerations are measured and weighted with appropriate filters. Root mean square values of accelerations taken over 5 sec time blocks are calculated. The test data sample is of 5 min duration. The 95th percentile point in each event distribution is then used to calculate a single parameter. The equation for the simplified method (where measurements are

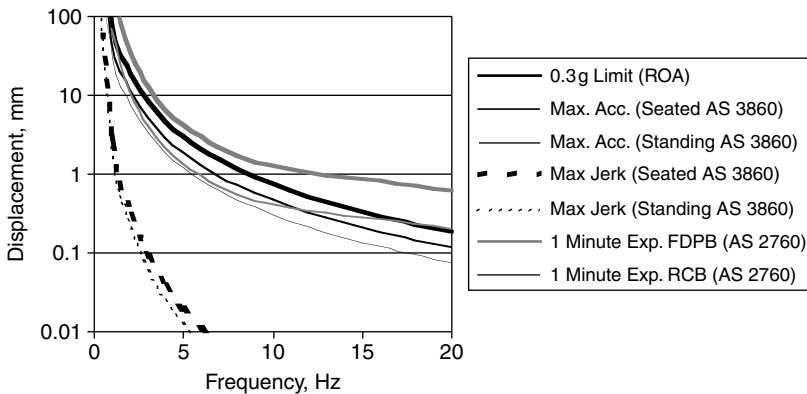


FIGURE 9.43 Passenger comfort displacement limits.

taken on the vehicle floor) is quoted below.

$$N_{MV} = 6\sqrt{(a_{XP95})^2 + (a_{YP95})^2 + (a_{ZP95})^2} \quad (9.23)$$

where  $a_{XP}$  is acceleration in the longitudinal direction;  $a_{YP}$  is acceleration in the lateral direction; and  $a_{ZP}$  is acceleration in the vertical direction.

A further equation is available for standing passengers, this time using 50 percentile points from event distributions.

$$N_{VD} = 3\sqrt{16(a_{XP50})^2 + 4(a_{YP50})^2 + (a_{ZP50})^2 + 5(a_{ZP95})^2} \quad (9.24)$$

Ride criteria using the above index parameters is:  $N < 1$ : very comfortable;  $1 < N < 2$ : comfortable;  $2 < N < 4$ : medium;  $4 < N < 5$ : uncomfortable; and  $N > 5$ : very uncomfortable.

## VI. TRAIN MANAGEMENT AND DRIVING PRACTICES

### A. TRAIN MANAGEMENT AND DRIVING PRACTICES

Train management and driving practices has received considerable attention in literature dating back several decades. Technology developments, such as the transitions from steam to diesel-electric locomotives, improved locomotive traction control systems, remote control locomotives, operation of very long heavy haul units trains, and the operation of high speed passenger services, have ensured that this area continues to evolve. Train management and driving practices will differ for different rail operations. Suburban train drivers will be motivated primarily by the need to run on time. A secondary consideration may be energy consumption. Longitudinal dynamics will have minimal consideration as cars are connected with minimal slack and usually have distributed traction and slip controls for both traction and braking. Passenger express services will be similarly motivated. Slow passenger services with locomotive hauled passenger cars will share the concerns of running on time with the next priority being the smoothness of passenger ride. Passenger train driver practice often includes energy consumptive power braking to minimise slack action. Where locomotives have excess power, train drivers have been known to operate with a minimum brake application on for several kilometres to reduce slack action over undulations. Mixed freight train practice, while not motivated by passenger comfort, will share some similar driving practices to ensure train stability. This is particularly the case when trains are operated with mixes of empty and loaded wagons. Running on time will be an emphasis on some systems depending on the type of freight. Differing from passenger systems, energy consumption is a significant freight cost factor and is emphasised in freight operations. The operation of bulk product/uniform module type freight trains (unit trains or block trains), e.g., carrying minerals, grain, containers etc, can be optimised to the specific source/destination requirements. In some cases, timeliness is a secondary concern while tonnage per week targets must be achieved.

Despite the differences in operation, a common thread to train management is the issue of speed control and hence management of train momentum. For suburban passenger trains, speed must be managed to ensure timeliness and adequate stopping distances for signals and for positioning at platforms. For longer locomotive hauled passenger, freight, and heavy haul unit trains, the problem of momentum control becomes even more significant due to the larger masses involved. In general, it is desirable to apply power as gradually as possible until in-train slack is taken up. During running it is desirable to minimise braking and energy wastage utilising coasting where possible. Route running schedules will limit the amount of time that the train can coast. Longer trains can coast over undulating track more easily than shorter trains due to grade forces being partially balanced within the train length. Stopping is achieved at several different rates.

Speed can be reduced by removing power and utilising rolling resistance (slowest), application of dynamic braking, application of minimum pneumatic braking, service application of pneumatic braking, and emergency application of pneumatic braking (fastest). The listed braking methods are also in order of increasing energy wastage and increased maintenance costs. The finer detail of braking practice will also depend on the usability and performance of the dynamic brake, Figure 9.29.

As suburban and high speed passenger trains could be classed as single vehicles due to the minimal slack in couplings and “viewable length”, the following discussion will be limited to slow passenger and longer freight and unit trains where the interplay of timeliness, energy conservation, and train dynamics must be considered.

## 1. Negotiating Crests, Dips, and Undulations

In negotiating crests and dips, the driver has the objectives of minimising the power loss in braking and managing in-train forces. In approaching the top of a crest, at some point close to the top (depending on grades, train size, etc.) power should be reduced to allow the upgrade to reduce train speed. The objective being that excess speed requiring severe braking will not occur as the train travels down the next grade. Similarly, when negotiating dips, power should be reduced at some point approaching the dip to allow the grade to bring the train to track speed as it travels through the dip. It can be seen that there is considerable room for variations in judgment and hence variation in energy usage. Work published in Ref. 7 indicated variations in fuel usage of up to 42% due primarily to differences in the way drivers manage the momentum of trains.

The handling of undulations presents several difficulties for train dynamics management due to slack action in the trains. The presence of undulations in track mean that slack action can occur within the train even while under steady power. Using techniques typical of passenger train operation it is often the case that power braking is used to keep the train stretched. The practice of power braking is the application of a minimum level of the pneumatic brake to all the wagons but not the locomotive. Locomotive tractive power is still applied. Simulation studies in Ref. 6 showed that power braking on the specified undulating track section succeeded in improving train dynamics only in the lead section of the train. The results of the paper should be utilised with some caution as the freight train under study appears to consist of uniformly loaded wagons and the assessment is based on coupler force data. The implication of mixed freight operations, with some empty or lightly loaded wagons, or hopper wagon unit train operations where a wagon is left unloaded, is not discussed in the paper. The risk of increased wheel unloading due to lateral coupler force components or due to bogie pitch due to force impacts, as discussed in Section III, is increased by the combination of larger in-train forces (as experienced in a loaded train), with a lightly loaded wagon. The use of power braking, while not reducing forces significantly, may still provide useful damping of longitudinal accelerations of lightly loaded wagons.

## 2. Pneumatic Braking

Braking techniques and practices are in part dictated by the specific requirements of the brake system. The Australian triple valve, North American AB valve, and European Distributor systems all utilise pressure differences between pipe pressure and on-wagon reservoirs to effect control. Brake pipe pressure is dropped by exhausting air via a valve in the drivers cabin. Due to the design, the minimum brake pipe application is usually of the order of 50 kPa reduction in brake pipe pressure. This will deliver 30% of the maximum brake pressure to the brake cylinders. This application is called a “minimum”. Drivers can also apply brakes using brake pipe pressure reductions of up to 150 kPa. These applications are called “service” applications. Full service brake cylinder pressures are reached in cylinders when the full 150 kPa application is applied. The brake pipe pressure can also be completely exhausted and this type of application is called an

“emergency” application. Emergency applications result in the maximum pressures in brake cylinders being applied. In Australia this is slightly greater than full service pressure due to valving design. In the North American system, a second reservoir of air is released during an emergency application giving a significantly higher cylinder pressure for emergency brake applications. Due to the slightly differing designs of the brake system and the policies of rail operators, driver braking practices will vary between countries and rail systems. The following practices are noted:

- Minimum applications without application of locomotive brakes.
- Minimum applications with application of locomotive brakes.
- Minimum applications with locomotive power applied application (power braking).
- Service applications without application of locomotive brakes.
- Service applications with application of locomotive brakes.
- Emergency applications.
- Penalty applications (automatic emergency in response to vigilance systems).
- Requirement to make a large service reduction after several minimum applications to ensure on-wagon valves are all operating correctly.
- Requirement to maintain any reduction for a time period.

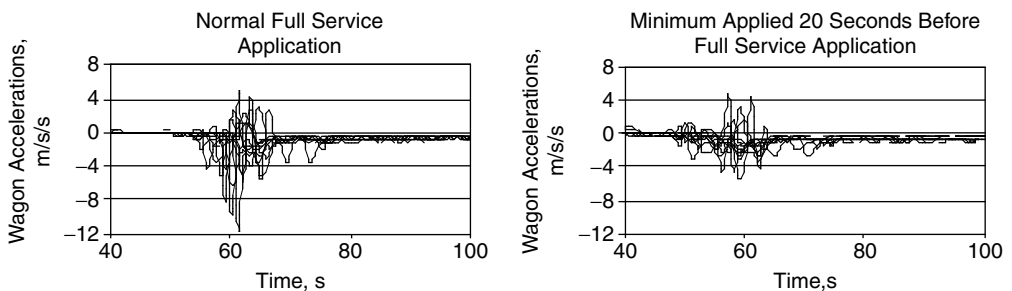
The use of minimum applications to either stretch the train or the use of minimum with power applied as a first stage of braking can reduce wagon accelerations and therefore improve in-train stability, Figure 9.44.

The most recent innovation in train braking is the development of electro-pneumatic (ECP) braking, although take up by freight operators has been slow.

The capability of the system to apply all brake cylinders simultaneously will reduce coupler impacts during brake applications and improve vehicle stability. Driving practices required to ensure the correct operation of the triple valves would also be expected to disappear.

### 3. Application of Traction and Dynamic Braking

The improved control systems for both tractive effort and dynamic braking has greatly improved locomotive performance in recent years with higher adhesion levels and greater ranges of speed where dynamic braking is effective. Significant improvement to traction systems can be found in slip controls and steering bogies. In practice, ground radar based slip controls give slightly better results than systems based on minimum locomotive drive axle speed. For train systems where the majority of running speeds fall within the flat region of dynamic brake response, driving strategies have been developed to predominantly use dynamic braking. An important practice is to ensure that drivers allow a period of time between the end of a throttle application and the beginning of a dynamic brake application or *vice versa*. This time period allows inter-wagon states to slowly move



**FIGURE 9.44** Wagon accelerations compared — different braking strategies.



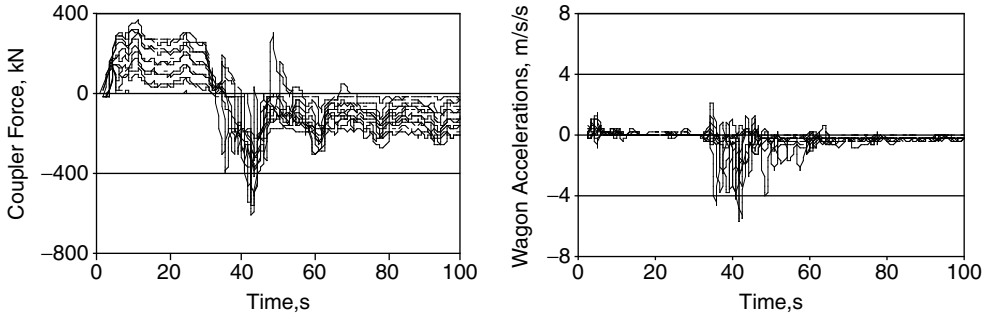


FIGURE 9.45 Power to dynamic brake transition without pausing.

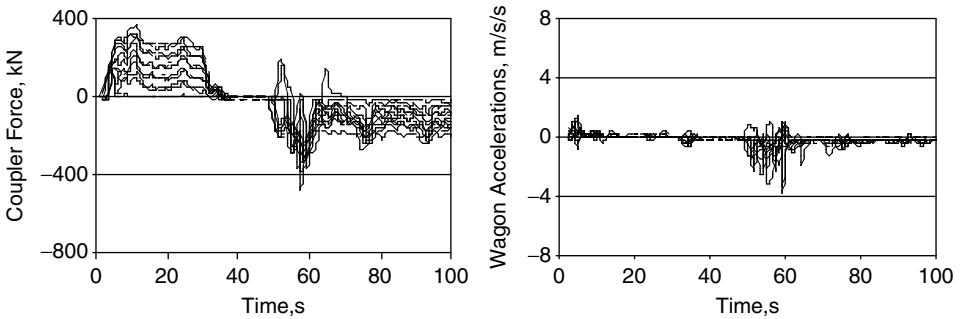


FIGURE 9.46 Power to dynamic brake transition with 20 second pause.

from stretched to bunched or *vice versa*, preventing large impact forces. Examples are simulated in Figure 9.45 and Figure 9.46.

**4. Energy Considerations**

Minimisation of energy usage is often a popular emphasis in train management. It is helpful to examine the way energy is utilised before innovations or changes to practice are adopted. Air resistance, for example, is often over-stated. A breakdown of the Davis equation<sup>14</sup> shows the significance of air resistance compared to curving resistance and rolling resistance factors and grades, Figure 9.47. It will be noticed that on a 1 in 400 grade, 0.25% is approximately equal to the propulsion resistance at 80 km/h.

The minimum energy required for a trip can be estimated by assuming an average train speed and computing the sum of the resistances to motion, not forgetting the potential energy effects of changes in altitude. The work carried out to get the train up to running speed once must also be added. As the train must stop at least once, this energy is lost at least once. Any further energy consumed will be due to signalling conditions, braking, stop-starts, and the design of grades. Minimum trip energy can be estimated as:

$$E_{\min} = \frac{1}{2} m_t v^2 + m_t g h + \sum_{i=1}^q \left( m_i \sum_{j=1}^r \left[ \int_0^{x=l_{c_j}} F_{crj} dx \right] \right) + \sum_{i=1}^q \left( m_i \int_0^{x=L} F_{prj} dx \right) \quad (9.25)$$

where  $E_{\min}$  is the minimum energy consumed, J;  $g$  is gravitational acceleration in  $m/sec^2$ ;  $h$  is the net altitude change, m;  $L$  is the track route length, m;  $l_{c_j}$  is the track length of curve  $j$ , m;  $m_i$  is

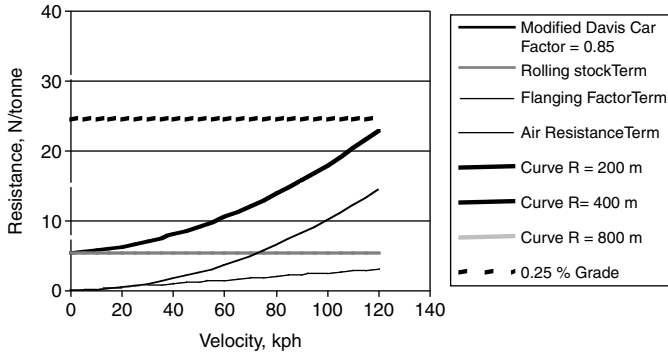


FIGURE 9.47 Comparative effects of resistances to motion.

individual vehicle mass  $i$ , kg;  $m_t$  is the total train mass, kg;  $F_{crj}$  is the curving resistance for curve  $j$  in Newtons;  $F_{pri}$  is the propulsion resistance for vehicle  $i$  in Newtons;  $q$  is the number of vehicles; and  $r$  is the number of curves.

Unless the track is extremely flat and signalling conditions particularly favourable, the energy used will be much larger than given by the above equation. However, it is a useful equation in determining how much scope exists for improved system design and practice. It is illustrative to consider a simple example of a 2000 tonne freight train with a running speed of 80 km/h. The work carried out to bring the train to speed, represented in Equation 9.25, by the kinetic energy term, is lost every time the train must be stopped and partly lost by any brake application. The energy loss per train stop in terms of other parameters in the equation are given in Table 9.3.

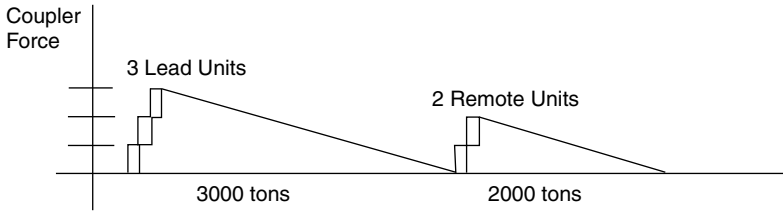
What can be seen at a glance from Table 9.3 is the very high cost of stop starts compared to other parameters. Air resistance becomes more significant for higher running speeds. High densities of tight curves can also add considerable costs. It should be noted that this analysis does not include the additional costs in rail wear or speed restriction also added by curves.

**5. Distributed Power Configurations**

Perhaps a landmark paper describing the operation of remote controlled locomotives was that of Parker,<sup>22</sup> referred to by Van Der Meulen.<sup>3</sup> The paper details the introduction of remote controlled locomotives to Canadian Pacific. The paper is comprehensive in its description of the equipment used, but, most importantly, it examines the issues concerning remote locomotive placement and includes operational case studies. Parker notes that the usual placement of the remote locomotives is at the position two thirds along the train. For operation on severe grades it was recommended that

**TABLE 9.3**  
Energy Losses Equivalent to One Train Stop for a Train Running at 80 km/h

Energy Parameter	Equivalent Loss	Units
Gravitational potential energy (second term Equation 9.25)	~25	Metres of altitude
Curving resistance (third term Equation 9.25)	~16	Kilometres of resistance due to curvature of 400 m radius
Propulsion resistance (fourth term Equation 9.25)	~18	Kilometres of propulsion resistance
Air resistance (Part of propulsion resistance)	~38	Kilometres of air resistance

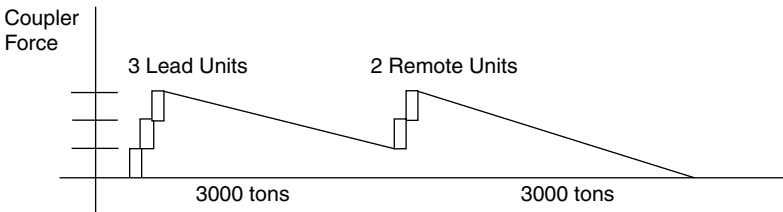


**FIGURE 9.48** Two trains connected configuration.<sup>22</sup> Source: From Parker, C. W., *Rail. Eng. J.*, January, 1974. With permission.

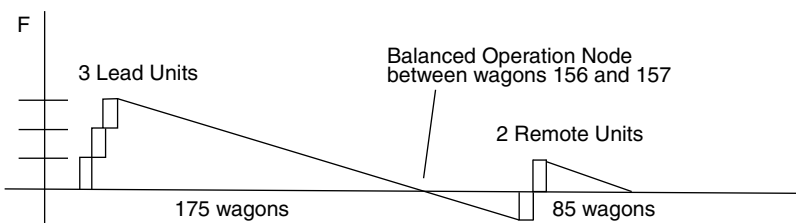
locomotives be positioned in proportion to wagon tonnages, i.e., “two trains connected”. Parker’s diagrams are redrawn in Figure 9.48 to Figure 9.50. The position and movement during operation of the point of zero coupler force, or “node,” was discussed in the paper at some length. Particular problems were noted with the two trains connected configuration in that if the lead locomotive units slow down relative to the remote units, the node moves forward. Under the resulting increased load the remote units will then slow down allowing the node to travel backward. Of interest was the author’s note that the relative speeds of lead and remote locomotive groups could differ by as much as 8 km/h. If the dynamic action in the train is severe enough, the front half of the train will attempt to accelerate the remote locomotive. This can result in large coupler force peaks or coupler failure. The movement of the units either forward or backward, Figure 9.49 and Figure 9.50, is recommended. It will be noticed that the units on the abscissa in Figure 9.50 are inconsistent with previous figures. This is consistent with Parker’s paper.<sup>22</sup>

Parker<sup>22</sup> also details several incidents of coupler failure due to the combination of track grade conditions and incorrect train control techniques. These are summarised briefly as follows:

1. *Starting a train on a crest:* in this case the coupler behind the remote locomotive failed. The locomotives, both lead and remote, were powered equally using multiple unit



**FIGURE 9.49** Remote locomotives placed ahead of a balanced operation node.<sup>22</sup> Source: Parker, C. W., *Rail. Eng. J.*, January, 1974. With permission.



**FIGURE 9.50** Remote locomotives placed behind a balanced operation node.<sup>22</sup> Source: Parker, C. W., *Rail. Eng. J.*, January, 1974. With permission.

operation. The failure occurred because the force applied by the lead locomotives was not required to haul the lead portion of wagons as they were located on a down hill grade. The force was therefore transferred through to the rear wagon group. The highest force was therefore generated behind the remote locomotives. The train could have been started successfully by powering the remote locomotives alone.

2. *Wheel slip on a heavy grade*: in this case the train was on a steep grade of greater than 2% and a train separation occurred near the front of the train. The failure occurred due to slippage or a momentary power loss that developed in the lead. The remote units then took up the slack and pushed the wagons in the front wagon rake. The lead locomotive then regained adhesion and accelerated forward causing severe coupler impacts and the failure of a coupler in a wagon near the front of the train. The problem can be solved by a slight reduction in the lead locomotive power setting. This requirement led to the incorporation of a device termed the “lead unit power reduction feature” into the multiple unit system.
3. *Effect of changes of grade*: the author noted several cases of short descents encountered after ascending a grade. In such cases the problem was the same as when starting a train on a crest. Excess traction force was transferred to the remote wagon rake resulting in separation behind the remote locomotives. Again, the reduction of lead locomotive power using independent control could have been used to prevent the problem. However, control strategies can become quite complicated if descents are short and followed by another grade. Power reductions that are too large could result in the remote units being stalled or slipping. The driver must therefore attempt to balance power settings and keep locomotive speeds the same.
4. *Braking under power*: the author<sup>22</sup> notes a case where separation occurred near the lead locomotives due to braking under power with the tail of the train on a slight grade. The traction force of the remote locomotive, combined with the grade, bunched the wagons in the front half of the train to compress the draft gears in that region. The compressed draft gears then forced the front locomotives forward causing a separation. The problem could have been prevented by reducing the remote locomotives to idle before the brakes were applied.

While the case studies by Parker<sup>22</sup> are not exhaustive, they are illustrative of the types of issues that arise in long train distributed power operation. It can be seen that driving strategies appropriate to the track topography are required. Simulation of train dynamics is a key tool in gaining an understanding of the train dynamics that could occur on a particular route. In such cases, attention to representative modelling of wagon connection elements, locomotive traction, adhesion, and braking characteristics are of utmost importance.

## VII. CONCLUSIONS

Detailed nonlinear models with stick–slip features for the simulation and study of longitudinal train dynamics have been developed, allowing increased understanding of long train dynamics. There still remains scope for further modelling and validation of models of existing draft gear packages. Further research should also be directed to new draft gear package designs.

The area of the interaction of train–wagon dynamics is an emerging area of research where train operators are operating longer trains on infrastructures with tighter curves.

Advances in locomotive controls and bogie design in recent years warrant the development of improved and more detailed traction and dynamic braking models.

The adoption of electro-pneumatic controlled brakes in freight train systems will improve wagon stability during braking and add a new variation to train management and driving practice.

## ACKNOWLEDGMENTS

The first acknowledgement is directed to Queensland Rail, Australia, who through sponsorship of collaborative research provided funding and technical support for the research into the longitudinal train dynamics simulation and train dynamics management for both the author's Doctoral studies and subsequent industry projects. The second acknowledgement is to the Center of Railway Engineering, Rockhampton, Australia and the team of researchers, programmers, and technicians with whom the author works and whose efforts ensure robust research and supporting field measurements. Finally, the author acknowledges and thanks his employer, Central Queensland University, Rockhampton, Australia for releasing work time for him to write this chapter.

## NOMENCLATURE

$a$ :	Vehicle acceleration, $\text{m/sec}^2$
$a_{XP}$ :	Acceleration in the longitudinal direction, $\text{m/sec}^2$
$a_{YP}$ :	Acceleration in the lateral direction, $\text{m/sec}^2$
$a_{ZP}$ :	Acceleration in the vertical direction, $\text{m/sec}^2$
$c$ :	Damping constant, $\text{Ns/m}$
$f_{wc}$ :	Nonlinear wagon connection function or subroutine.
$g$ :	Gravitational acceleration in $\text{m/sec}^2$
$h$ :	Net altitude change, $\text{m}$
$k$ :	Spring stiffness, $\text{N/m}$
$k_{ad}$ :	Air drag constant depending on car type
$k_f$ :	Locomotive torque reduction factor, Newton per metre per second $\text{N}/(\text{m/sec})$
$m$ :	Vehicle mass, $\text{kg}$
$m_a$ :	Mass supported per axle in tonnes
$m_i$ :	Vehicle mass $i$ , $\text{kg}$
$m_t$ :	train mass in $\text{kg}$
$n$ :	Number of axles
$v$ :	Vehicle velocity, $\text{m/sec}$
$x$ :	Vehicle displacement, $\text{m}$
$E_{\min}$ :	Minimum energy consumed, $\text{J}$
$F_b$ :	Braking resistance due to pneumatic braking, $\text{N}$
$F_c$ :	Coupler Force, $\text{N}$
$F_{cr}$ :	Curving resistance, $\text{N}$
$F_g$ :	Gravity force components due to track grade, $\text{N}$
$F_{pr}$ :	Propulsion resistance, $\text{N}$
$F_r$ :	Sum of retardation forces, $\text{N}$
$F_s$ :	Draft gear spring force, $\text{N}$
$F_{t/db}$ :	Traction and dynamic brake forces from a locomotive unit, $\text{N}$
$K_a$ :	Adjustment factor depending on rollingstock type
$L$ :	Track route length, $\text{m}$
$N$ :	Throttle setting in notches, 0 to 8
$P$ :	Locomotive power, Watts
$P_{\max}$ :	Maximum locomotive traction horsepower, Watts
$Q$ :	Friction wedge factor
$R$ :	Curve radius, $\text{m}$
$Te_{\max}$ :	Maximum locomotive traction force, Newtons
$V$ :	Velocity in kilometres per hour
$\Delta V$ :	Head wind speed, usually taken as $15 \text{ km/h}$
$\phi$ :	Wedge angle

$V_f$ :	Kinetic friction velocity, m/sec
$\mu_s$ :	Static friction coefficient
$\mu_k$ :	Kinetic coefficient of friction

## REFERENCES

1. Duncan, I. B. and Webb, P. A., The Longitudinal Behaviour of Heavy Haul Trains Using Remote Locomotives, *Fourth International Heavy Haul Conference*, Brisbane, pp. 587–590, 1989.
2. Jolly, B. J. and Sismey, B. G., Doubling the Length of Coals Trains in the Hunter Valley, *Fourth International Heavy Haul Conference*, Brisbane, pp. 579–583, 1989.
3. Van Der Meulen, R. D., Development of Train Handling Techniques for 200 Car Trains on the Ermelo-Richards Bay Line, *Fourth International Heavy Haul Conference*, Brisbane, pp. 574–578, 1989.
4. El-Sibaie, M., Recent Advancements in Buff and Draft Testing Techniques, *Fifth International Heavy Haul Conference*, Beijing, 1993.
5. McClanachan, M., Cole, C., Roach, D., and Scown, B., An Investigation of the Effect of Bogie and Wagon Pitch Associated with Longitudinal Train Dynamics, *The Dynamics of Vehicles on Roads and on Tracks-Vehicle System Dynamics Supplement 33*, Swets & Zeitlinger, Amsterdam, pp. 374–385, 1999.
6. Scown, B., Roach, D., and Wilson, P., Freight Train Driving Strategies Developed for undulating Track Through Train Dynamics Research, *Conference on Railway Engineering, Adelaide*, Rail Technical Society of Australia, Australia, pp. 27.1–27.12, 2000.
7. Simson, S., Cole, C., and Wilson, P., Evaluation and Training of Train Drivers during Normal Train Operations, *Conference on Railway Engineering, Wollongong*, Rail Technical Society of Australia, Australia, pp. 329–336, 2002.
8. Garg, V. K. and Dukkipati, R. V., *Dynamics of Railway Vehicle Systems*, Academic Press, New York, 1984.
9. Cole, C., Improvements to Wagon Connection Modelling for Longitudinal Train Simulation, *Conference on Railway Engineering, Rockhampton*, Institution of Engineers, Australia, pp. 187–194, 1998.
10. Muller, L. and Witt, T., TRAIN — A Computer Model for the Simulation of Longitudinal Dynamics in Trains, *Conference on Railway Engineering, Rockhampton*, Institution of Engineers, Australia, pp. 181–186, 1998.
11. Wolf, G. P. and Kieres, K. C., Innovative Engineering Concepts for Unit Train Service: The Slackless Drawbar Train and Continuous Center Sill Trough Train, *The Fourth International Heavy Haul Railway Conference*, Brisbane, pp. 124–128, 1989.
12. Bartley, G. W. and Cavanaugh, S. D., *The Second Generation Unit Train*, *The Fourth International Heavy Haul Railway Conference*, Brisbane, pp. 129–133, 1989.
13. Andrews, H. I., *Railway Traction*, Elsevier Science Publishers, The Netherlands, 1986.
14. Hay, W. W., *Railroad Engineering*, 2nd ed., Wiley, New York, pp. 69–82, 1982.
15. Profillidis, V. A., *Railway Engineering*, 2nd ed., Ashgate, Aldershot, 2000.
16. RCP-6102: Locomotive Hauled Cars, Part 6 Passenger Cars, Code of Practice for the Defined Interstate Rail Network — Vol. 5: Rollingstock, pp. 6–7, 2003.
17. Section 12, Railways of Australia Manual of Engineering Standards and Practices Adopted 24/10/91, pp. 12.37–12.40.
18. RCP-6103: Locomotive Hauled Cars, Part 6 Passenger Cars, Code of Practice for the Defined Interstate Rail Network — Vol. 5: Rollingstock, pp. 11–12, 2003.
19. AS2760 Evaluation of Human Exposure to Whole Body Vibration, Standards Association of Australia 1990.
20. AS3860 Fixed Guideway People Movers, Standards Association of Australia 1991.
21. UIC Leaflet 513, Guidelines for evaluating passenger comfort in relation to vibration in railway vehicles, issued 1/11/2003, UIC International Union of Railways, Paris.
22. Parker, C. W., Design and Operation of Remote Controlled Locomotives in Freight Trains, *Rail. Eng. J.*, Jan 1974.

---

# 10 Noise and Vibration from Railway Vehicles

*David Thompson and Chris Jones*

## CONTENTS

I.	Introduction .....	280
A.	The Importance of Noise and Vibration.....	280
B.	Basics of Acoustics .....	280
C.	Sources of Railway Noise and Vibration .....	282
II.	Rolling Noise .....	282
A.	Mechanism of Rolling Noise Generation.....	282
B.	Surface Roughness .....	283
C.	Wheel Dynamics .....	285
D.	Track Dynamics .....	287
E.	Wheel–Rail Interaction .....	288
F.	Noise Radiation .....	289
G.	Computer Packages.....	290
III.	Reducing Rolling Noise.....	290
A.	Controlling Surface Roughness .....	291
B.	Wheel-Based Solutions .....	291
1.	Wheel Damping.....	291
2.	Wheel Design .....	292
C.	Track-Based Solutions .....	293
1.	Low Noise Track.....	293
2.	Slab Tracks .....	294
D.	Local Shielding and Barriers .....	295
IV.	Impact Noise .....	296
A.	Introduction .....	296
B.	Wheel Flats .....	296
C.	Predicting Impact Noise from Wheel Flats.....	298
D.	Rail Joints.....	299
E.	Reducing Impact Noise.....	301
V.	Curve Squeal .....	301
A.	Mechanism of Squeal Noise Generation.....	301
B.	Reducing Squeal Noise .....	303
VI.	Other Sources of Noise.....	304
A.	Aerodynamic Noise.....	304
B.	Power Unit Noise.....	304
VII.	Vehicle Interior Noise.....	304
A.	Vehicle Interior Noise Levels.....	304
B.	Measurement Quantities for Interior Noise.....	305

C.	Airborne Transmission.....	307
D.	Structure-Borne Transmission .....	308
E.	Prediction of Interior Noise .....	310
VIII.	Ground-Borne Vibration and Noise.....	310
A.	Overview of Vibration Phenomena .....	310
B.	Surface Vibration Propagation.....	311
C.	Tunnel Vibration .....	315
D.	Vibration Isolating Tracks .....	317
E.	Summary .....	319
IX.	Vibration Comfort on Trains .....	320
A.	Introduction .....	320
B.	Assessment of Vibration Comfort in Trains .....	320
C.	Effects of Vehicle Design .....	321
	References.....	322

## I. INTRODUCTION

### A. THE IMPORTANCE OF NOISE AND VIBRATION

Environmental noise is an issue that has seen increased awareness in recent years. Within the European Union (E.U.) it is estimated<sup>1</sup> that 20% of the population live in areas with unacceptable noise levels.<sup>a</sup> Noise is often cited as a major factor contributing to people's dissatisfaction with their environment. While this noise exposure is usually due mainly to road traffic, trains also contribute significantly in the vicinity of railway lines. Road vehicles and aircraft have long been the subject of legislation that limits their noise emissions. The E.U. has therefore recently introduced noise limits for new rail vehicles. These have been implemented as part of the Technical Specifications for Interoperability (TSIs), which initially cover high speed trains<sup>2</sup> and are being extended to include conventional trains. They state noise limits for new trains under both static and running conditions.

By contrast with exterior noise, the noise inside a vehicle (road or rail) is not generally the subject of legislation, apart from the noise inside the driver's cab. For road vehicles, noise is actually used as a major factor to distinguish vehicles from their competitors and to attract people to buy a particular vehicle. As rail vehicles are for mass use, interior noise is subject instead to specifications from the purchasing organisation. These are usually limited to ensuring that problems are eliminated and that the vehicles are fit for their purpose.

Railway operations also generate vibrations that are transmitted through the ground into neighbouring properties. These can lead either to feelable vibration (in the range 4 to 80 Hz) or to low frequency rumbling noise (30 to 250 Hz). Vibrations are also transmitted into the vehicle itself, affecting passenger comfort.

### B. BASICS OF ACOUSTICS

The field of acoustics is too large to cover in detail here. This chapter therefore gives only a very brief overview of some basic quantities. The interested reader is referred to textbooks on the subject for further details.<sup>3,4</sup>

Sound consists of audible fluctuations in pressure, usually of the air. These propagate through the air as waves with a wave speed, denoted by  $c_0$ , of about 340 m/sec in air at 20°C.

<sup>a</sup> This is expressed as levels above 65 dB,  $L_{Aeq}$ . The  $L_{Aeq}$  is the A-weighted equivalent noise level averaged over a period of, for example, a day (or night).



Simultaneously, fluctuations in air density and particle motion also occur. To express the magnitude of a sound, the root mean square (rms) sound pressure is usually used:

$$p_{\text{rms}} = \left( \frac{1}{T} \int_{t_1}^{t_1+T} p^2(t) dt \right)^{1/2} \quad (10.1)$$

where  $p(t)$  is the instantaneous sound pressure and  $T$  is the averaging time. Much use is made of frequency analysis, whereby sound signals are decomposed into their frequency content (e.g., using Fourier analysis). The normal ear is sensitive to sound in the frequency range 20 to 20,000 Hz (the upper limit reduces with age and with noise exposure) and to a large range of amplitudes (around six orders of magnitude). Owing to these large ranges, and to mimic the way the ear responds to sound, logarithmic scales are generally used to present acoustic data. Thus amplitudes are expressed in decibels. The *sound pressure level* (NB level implies decibels) is defined as:

$$L_p = 10 \log_{10} \left( \frac{p_{\text{rms}}^2}{p_{\text{ref}}^2} \right) = 20 \log_{10} \left( \frac{p_{\text{rms}}}{p_{\text{ref}}} \right) \quad (10.2)$$

where the reference pressure  $p_{\text{ref}}$  is usually  $2 \times 10^{-5}$  Pa. Frequencies (expressed in Hz) are also generally plotted on logarithmic scales, with *one-third octave bands* being a common form of presentation. The frequency range is divided into bands that are of equal width on a logarithmic scale. The centre frequencies of each band can be given by  $10^{(N/10)}$  where  $N$  is the band number, although by convention they are rounded to particular values. Bands 13 to 43 cover the audible range.

The total sound emitted by a source is given by its power,  $W$ , which in decibel form is given as the *sound power level*:

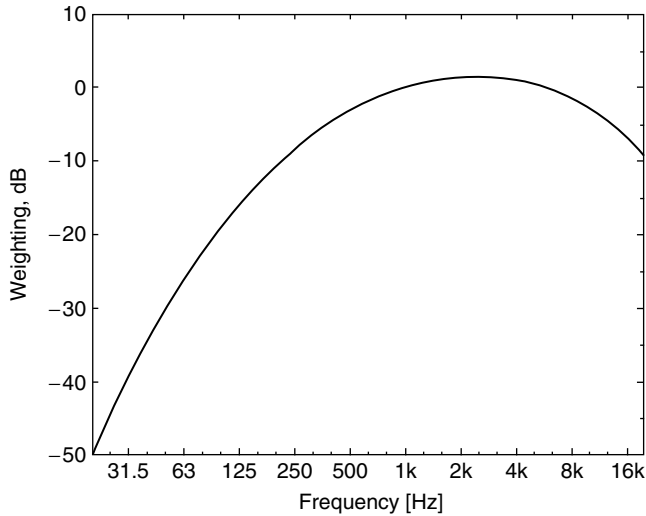
$$L_W = 10 \log_{10} \left( \frac{W}{W_{\text{ref}}} \right) \quad (10.3)$$

where the reference power,  $W_{\text{ref}}$ , is usually  $10^{-12}$  W. The power is generally proportional to the square of the sound pressure, so that a 1-dB increase in sound power level leads to a 1-dB increase in sound pressure level at a given location. However, sound pressure also depends on the location, usually reducing as the receiver becomes further from the source. For a compact point source this reduction is 6 dB per doubling of distance, while for a line source it is 3 dB per doubling. Other quantities can also be expressed in decibels following the pattern of Equation 10.2 and Equation 10.3.

It should be realised that sound generation is often a very inefficient process. The *proportion* of the mechanical power of a typical machine that is converted into sound is often in the range  $10^{-7}$  to  $10^{-5}$ . Sound is generated by various mechanisms, but the two main ones are:

- *Structural vibrations* — the vibration of a structure causes the air around it to vibrate and transmit sound, e.g., a drum, a loudspeaker, wheels and rails.
- *Aerodynamic fluctuations* — wind, particularly turbulence and flow over solid objects, also produces sound, e.g., jet noise, turbulent boundary layer noise, exhaust noise, fan noise.

It can be pointed out that *noise* differs from *sound* in that noise is unwanted sound. While the acceptability of sound levels and signal content varies greatly between individuals, it is important to include some approximation to the way the ear weights different sounds. Several weighting curves have been devised, but the A-weighting (Figure 10.1) is the most commonly used. This approximates the inverse of the equal loudness curve at about 40 dB. As the ear is most



**FIGURE 10.1** The A-weighting curve.

sensitive around 1 to 5 kHz and much less sensitive at low and high frequencies, more prominence is given to this central part of the spectrum. The overall sound level is often quoted as an A-weighted value, meaning that this weighting curve is applied to the spectrum before calculating the total.

Another overall measure of the magnitude of a sound is the *loudness*. Strictly, this is a subjective quantity, but there are ways of calculating a loudness value from a one-third octave band spectrum.<sup>5</sup> However, this is less commonly used than the A-weighted decibel. It should be borne in mind that an increase of 10 dB is perceived as a doubling of loudness, while a change of less than 3 dB is normally imperceptible.

## C. SOURCES OF RAILWAY NOISE AND VIBRATION

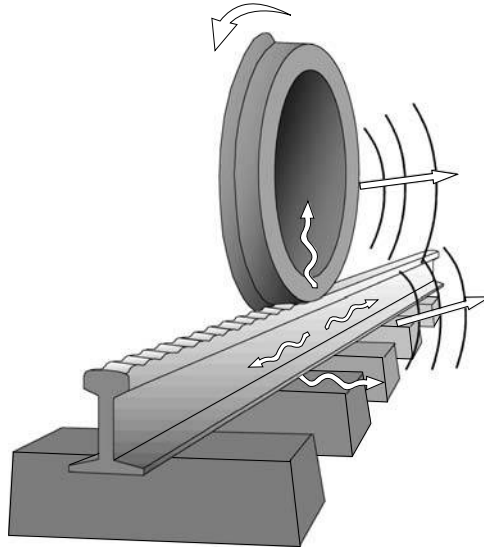
In the case of railway noise, both of the above types of mechanism apply. Aerodynamic noise is important for high-speed operation and is generated by unsteady airflow, particularly over the nose, intercarriage joints, bogie regions, louvres, and roof-mounted equipment such as pantographs. However, mechanical sources of noise are also present on a train and these dominate the overall noise for speeds up to about 300 km/h.

The most important mechanical noise source from a train is generated at the wheel–rail contact. Rolling noise is caused by vibrations of the wheel and track structures, induced at the wheel–rail contact point by vertical irregularities in the wheel and rail surfaces. A similar mechanism leads to noise due to discontinuities in the wheel or rail surface (impact noise). Squeal noise occurs in sharp curves and is induced by unsteady friction forces at the wheel–rail contact. Finally, ground-borne vibration and noise are caused by track and wheel irregularities and by the movement of the set of axle loads along the track. Each of these sources of noise and vibration are discussed in turn in the following sections.

## II. ROLLING NOISE

### A. MECHANISM OF ROLLING NOISE GENERATION

As indicated above, rolling noise is usually the dominant source of noise from moving trains at speeds below about 300 km/h. It can be attributed to components radiated by vibration of both



**FIGURE 10.2** Schematic view of how rolling noise is generated at the wheel–rail interface.

the wheels and track. This vibration is caused by the combined surface roughness at their interface, as shown in Figure 10.2.

The relative importance of the components of sound radiation from the wheel and track depends on their respective designs as well as on the train speed and the wavelength content of the surface roughness. In most cases both sources (wheel and track) are significant. As the noise radiation depends on the roughness of both the wheel and track, it is possible that a rough wheel causes a high noise level that is mainly radiated by the track vibration or vice versa. It is therefore difficult to assign noise contributions solely to the vehicle or infrastructure.

## B. SURFACE ROUGHNESS

Irregularities with wavelengths between about 5 and 500 mm cause the vibrations of relevance to noise. When a wavelength  $\lambda$ , in m, is traversed at a speed  $\nu$ , in m/s, the associated frequency generated (in Hz) is given by:

$$f = \frac{\nu}{\lambda} \quad (10.4)$$

The corresponding amplitudes range from over 50  $\mu\text{m}$  at long wavelengths to much less than 1  $\mu\text{m}$  at short wavelengths. Typical wheel roughness spectra are shown in Figure 10.3. These are given in decibels relative to 1  $\mu\text{m}$  (using a definition equivalent to Equation 10.2), expressed in one-third octave bands over wavelength.

In the TSIs,<sup>2</sup> a standard is included for the roughness of a test track that is used to measure vehicles. The roughness should be less than a specified spectrum, shown in Figure 10.4. This represents good quality track. The purpose of this is to ensure that variations in rail roughness from one site to another do not significantly affect the measurement, as the wheel roughness will usually be at least as large as the rail roughness (see Figure 10.3).

The wheel–rail contact does not occur at a point but over a small area. The *contact patch* is typically 10- to 15-mm long and a similar width. When roughness wavelengths are short compared with the contact patch length, their effect on the wheel–rail system is attenuated. This effect is

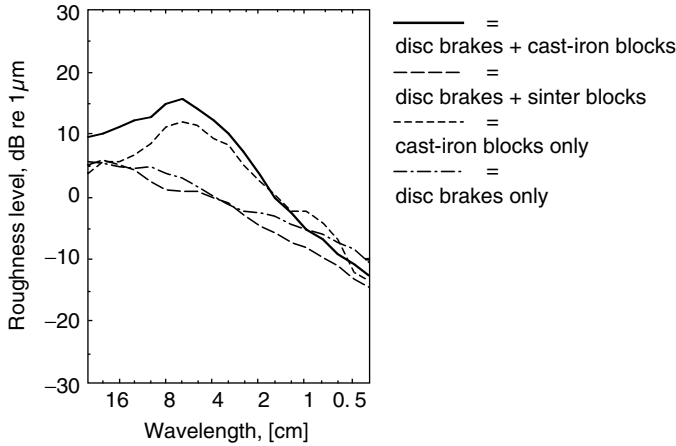


FIGURE 10.3 Typical wheel roughness spectra.<sup>6</sup>

known as the contact filter. This is significant from about 1 to 1.5 kHz for a speed of 160 km/h, and at lower frequencies for lower speeds.

In early analytical models for this effect,<sup>7</sup> the extent of the correlation of the roughness across the width of the contact had to be assumed since very detailed roughness data were not available. Figure 10.5a shows results from this model for a contact patch length of 11 mm. The parameter  $\alpha$  determines the extent of correlation across the width that is assumed. More recently, Remington has developed a numerical discrete point reacting spring (DPRS) model.<sup>8</sup> This model is intended to be used with roughness measurements obtained on multiple parallel lines a few millimetres apart. Figure 10.5b from Ref. 9, shows results obtained using a series of such measurements in combination with the DPRS model. This confirms the validity of the analytical model at low frequencies but indicates that the filtering effect is less severe at high frequencies than the analytical model suggests.

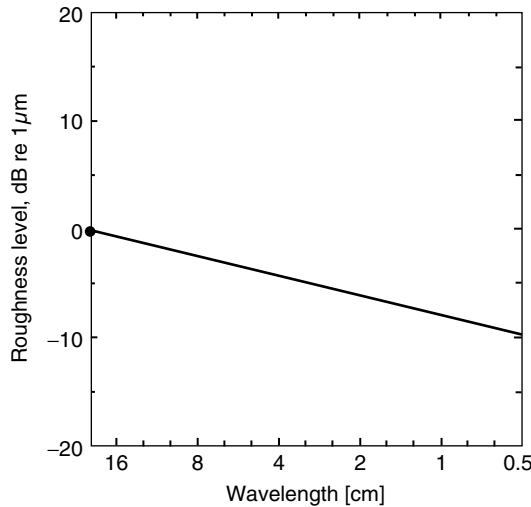
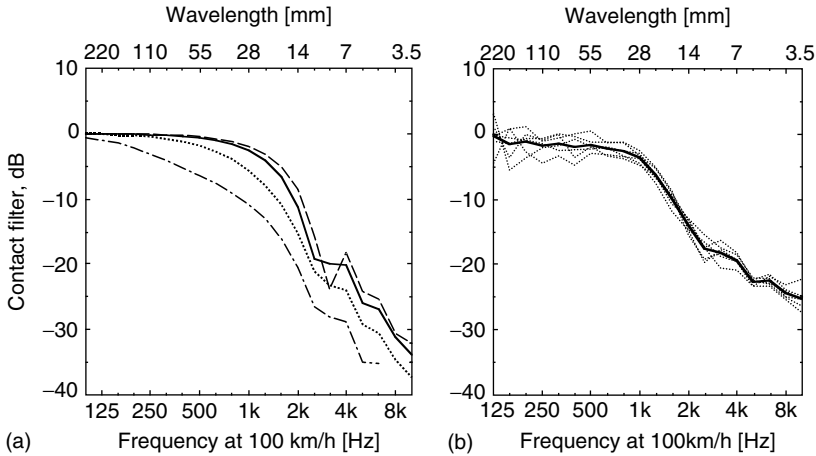


FIGURE 10.4 Maximum roughness allowed for vehicle noise measurements according to high-speed train TSI.<sup>2</sup>



**FIGURE 10.5** Contact filter due to contact patch of semiaxis length 5.7 mm. (a) Analytical model: —  $\alpha = 1$ , - -  $\alpha = 0.1$ , ...  $\alpha = 3$ , - . -  $\alpha = 10$ . (b) Numerical DPRS model for data from three cast-iron block-braked wheels, one disc-braked wheel, and two sinter block braked wheels. — Mean of six wheels.<sup>9</sup> *Source:* From Thompson, D. J., *J. Sound Vib.*, 267, 523–535, 2003, Elsevier. With permission.

### C. WHEEL DYNAMICS

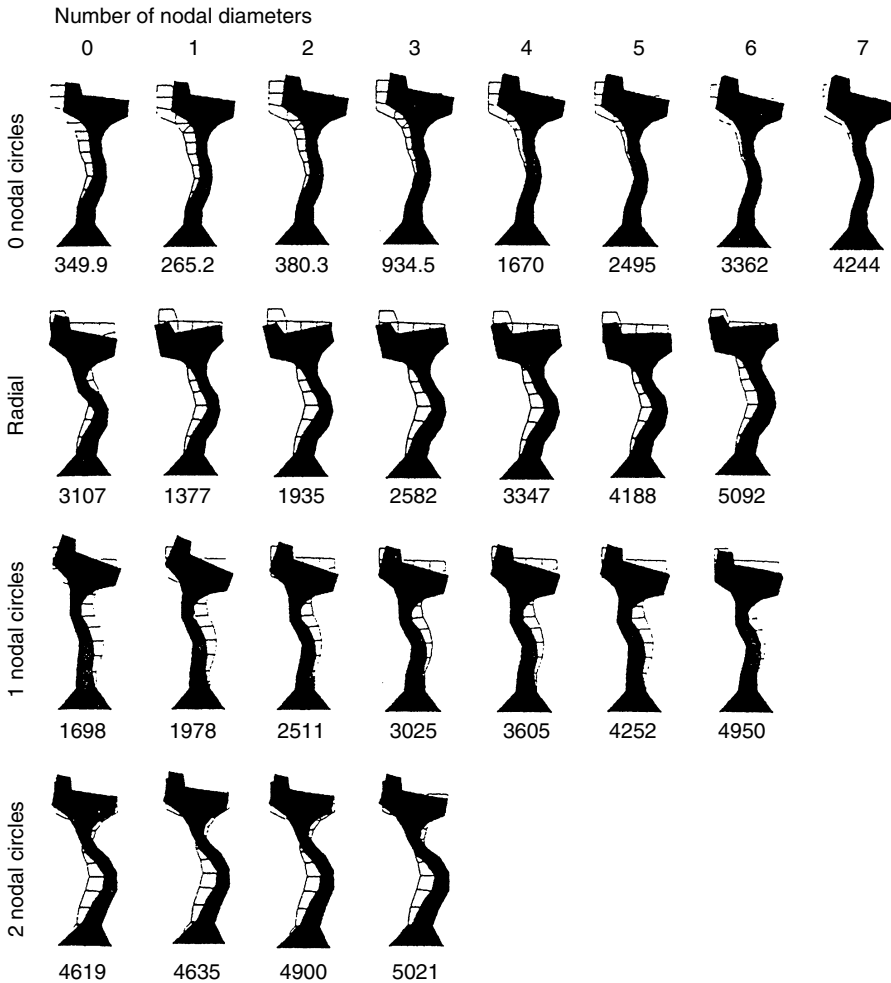
A railway wheel is a lightly damped resonant structure, which when struck rings like a bell, a structure which it strongly resembles. As with any structure, the frequencies at which it vibrates freely are called its “resonance” or “natural” frequencies and the associated vibration pattern is called the mode shape.

Wheels are usually axisymmetric (although the web is sometimes not). Their normal modes of vibration can therefore be described in terms of the number of diametral node lines — lines at which the vibration pattern has a zero. A flat disc, to which a wheel can be approximated, has out-of-plane modes that can be described by the number of nodal diameters,  $n$ , and the number of nodal circles,  $m$ . A perfectly flat disc also has in-plane radial modes with  $n$  nodal diameters and circumferential modes with  $n$  nodal diameters. In-plane modes with nodal circles occur for railway wheels above 6 kHz.

A railway wheel differs from a flat disc, having a thick tyre region at the perimeter and a thick hub at the centre connecting the wheel to the axle. A railway wheel is also not symmetric about a plane perpendicular to its axis. The tyre region is asymmetric due to the flange, and the web is usually also asymmetric, at least on wheels designed for tread braking, the curved web being designed to allow for thermal expansion. An important consequence of this asymmetry is that radial and out-of-plane (axial) modes are coupled.

The finite element method can be used quite effectively to calculate the natural frequencies and mode shapes of a railway wheel. Figure 10.6 shows an example of results for a UIC 920-mm freight wheel.<sup>10</sup>

The cross-section through the wheel is shown, along with an exaggerated form of the deformed shape in each mode of vibration. Each column contains modes of a particular number of nodal diameters,  $n$ . The first row contains axial modes with no nodal circle. These have their largest out-of-plane vibration at the running surface of the wheel. These modes are usually excited in curve squeal (see Section V below) but are not excited significantly in rolling noise. The second and third rows contain one-nodal-circle axial modes and radial modes. Owing to the asymmetry of the wheel cross-section, and their proximity in frequency, these two sets of modes are strongly coupled, that is, both contain axial *and* radial motion. It is these modes that are most strongly excited by roughness during rolling on straight track, due to their radial component at the wheel–rail contact point.

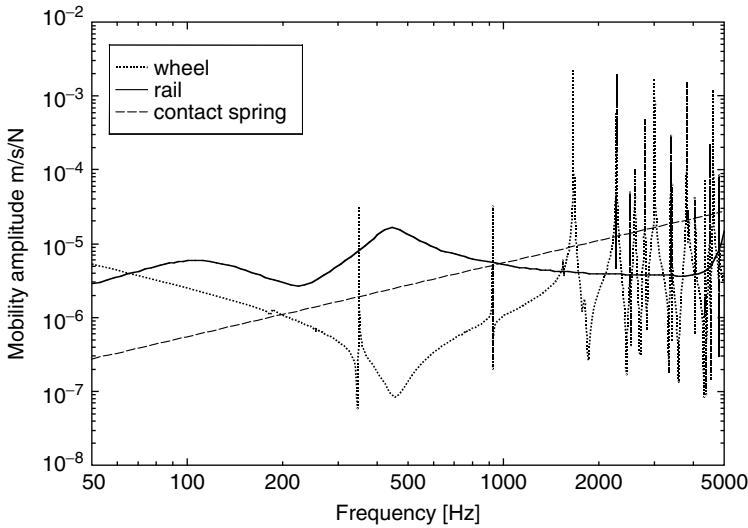


**FIGURE 10.6** Modes of vibration and natural frequencies (in Hz) of UIC 920-mm freight wheel calculated using finite elements.<sup>10</sup> Source: From Thompson, D. J., *J. Sound Vib.*, 231, 519–536, 2000, Elsevier. With permission.

The modes shown in Figure 10.6 are of the wheel alone, constrained rigidly at the inner edge of its hub. The first column of modes,  $n = 0$ , are in practice coupled to extensional motion in the axle, and the second set,  $n = 1$ , are coupled to bending motion in the axle. As a result of this coupling with the axle, which is constrained by the roller bearings within the axle boxes, these sets of modes experience greater damping than the modes with  $n \geq 2$ . The latter do not involve deformation of the axle and therefore are damped only by material losses; their modal damping ratios are typically about  $10^{-4}$ .

In order to couple the wheel to the track in a theoretical model, the frequency response functions of the wheel at the interface point are required. These may be expressed in terms of receptance, the vibration displacement due to a unit force as a function of frequency. Alternatively, mobility, the velocity divided by force, or accelerance, the acceleration divided by force, can be used.

Such frequency response functions of a structure can be constructed from a modal summation. For each mode, the natural frequency  $f_{mn}$  is written as a circular frequency  $\omega_{mn} = 2\pi f_{mn}$ .



**FIGURE 10.7** Vertical mobilities of the wheel–rail system. Radial mobility of UIC 920-mm freight wheel, vertical mobility of track with moderately soft pads and contact spring mobility.

Then the response at circular frequency  $\omega$ , in the form of a receptance  $\alpha_{jk}^W$  is:

$$\alpha_{jk}^W = \sum_{n,m} \frac{\psi_{mnj}\psi_{mnk}}{m_{mn}(\omega_{mn}^2 - \omega^2 + 2i\zeta_{mn}\omega\omega_{mn})} \tag{10.5}$$

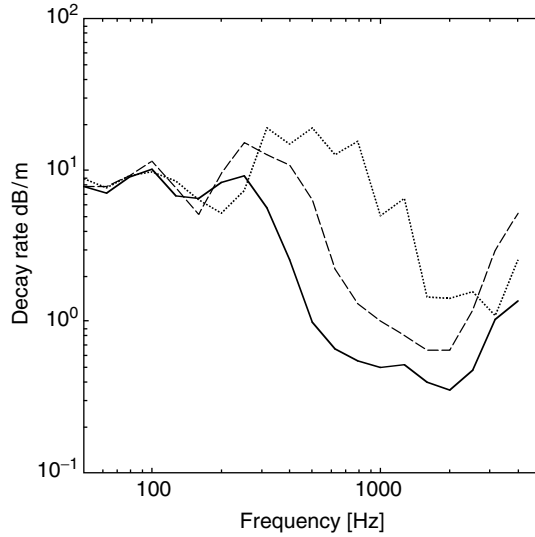
where  $\psi_{mnj}$  is the mode shape amplitude of mode  $m,n$  at the response position,  
 $\psi_{mnk}$  is the mode shape amplitude of mode  $m,n$  at the force position,  
 $m_{mn}$  is the modal mass of mode  $m,n$ , a normalisation factor for the mode shape amplitude,  
 $\zeta_{mn}$  is the modal damping ratio of mode  $m,n$ ,  
 $i$  is the square root of  $-1$ .

Figure 10.7 shows the radial point mobility of a wheelset calculated using the normal modes from a finite element model as shown in Figure 10.6. This is based on Equation 10.5 multiplied by  $i\omega$  to convert from receptance to mobility. At low frequencies the mobility is inversely proportional to frequency, corresponding to mass-like behaviour. Around 500 Hz an antiresonance trough appears and above this frequency the curve rises in stiffness-like behaviour until a series of sharp resonance peaks are reached at approximately 2 kHz. These peaks are the axial one-nodal-circle and radial sets of modes, identified in Figure 10.6.

### D. TRACK DYNAMICS

The dynamic behaviour of track is described in detail in Chapter 6. A typical track mobility is also shown in Figure 10.7. This is predicted using a model based on a continuously supported rail, which neglects the effects of the periodic support. A broad peak at around 100 Hz corresponds to the whole track vibrating on the ballast. At the second peak, at approximately 500 Hz, the rail vibrates on the rail pad stiffness. The frequency of this peak depends on the rail pad stiffness. Above this frequency, bending waves propagate in the rail and can be transmitted over quite large distances.

The degree to which these waves are attenuated, mainly due to the damping effect of the pads and fasteners, affects the noise radiation from the rail. Figure 10.8 shows measured decay rates of vertical vibration for three different rail pads installed in the same track. The results for the middle



**FIGURE 10.8** Measured decay rate of vertical vibration along the track for three different rail pads: — 140, – – – 300, and ... 1000 MN/m.

value of pad stiffness corresponds to the mobility in Figure 10.7. The vertical bending waves are strongly attenuated in a region between 300 and 800 Hz which depends on the pad stiffness. This peak in the decay rate corresponds to the region between the two resonance peaks in Figure 10.7. Here, the sleeper mass vibrates between the pad and ballast springs and acts as a dynamic absorber to attenuate the propagation of waves in the track. The attenuation of lateral waves is generally smaller than for the vertical direction.

**E. WHEEL–RAIL INTERACTION**

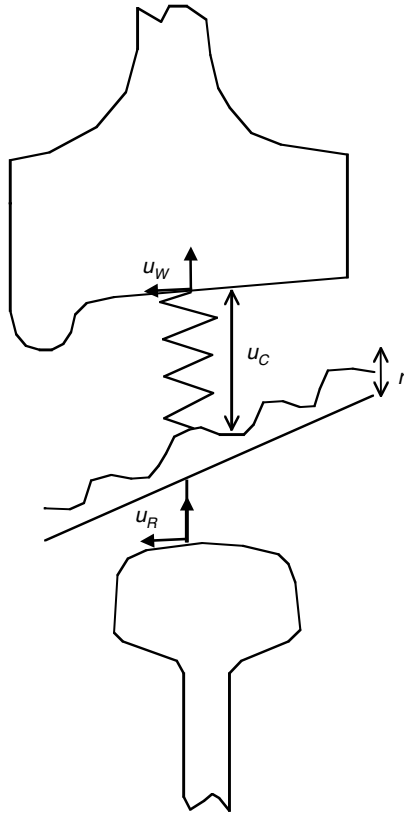
The wheel and rail are coupled dynamically at their point of contact. Between them local elastic deflection occurs to form the contact patch, which can be represented as a contact spring. Although this spring is nonlinear (see Chapter 4), for small dynamic deflections it can be approximated by a linearised stiffness,  $k_H$ .<sup>11</sup> This is shown as a mobility ( $= i\omega/k_H$ ) in Figure 10.7.

The coupled wheel–rail system is excited by the roughness, which forms a relative displacement input (see Figure 10.9). Here, the motion of the wheel is ignored and the system is replaced by one in which the wheel is static and the roughness is pulled between the wheel and rail (moving irregularity model). Considering only coupling in the vertical direction, from equilibrium of forces and compatibility of displacements, the vibration amplitude of the wheel ( $u_W$ ) and rail ( $u_R$ ) at a particular frequency can be written as:

$$u_W = \frac{\alpha_W r}{\alpha_W + \alpha_R + \alpha_C}; \quad u_R = \frac{-\alpha_R r}{\alpha_W + \alpha_R + \alpha_C} \tag{10.6}$$

where  $r$  is the roughness amplitude and  $\alpha_W$ ,  $\alpha_R$ ,  $\alpha_C$  are the vertical receptances of the wheel, rail, and contact spring, respectively. Clearly, where the rail receptance has a much larger magnitude than that of the wheel or contact spring,  $u_R \approx -r$ , that is, the rail is pushed down at the amplitude of the roughness. From Figure 10.7, this can be expected between approximately 100 and 1000 Hz. Changing the rail receptance in this frequency region has little effect on the rail vibration at the contact point (although the changes may affect the decay rates).





**FIGURE 10.9** The wheel–rail contact showing excitation by roughness of amplitude  $r$ .

In practice, coupling also exists in other directions as well as the vertical, notably the lateral direction. This modifies Equation 10.6 to yield a matrix equation, but the principle remains the same.

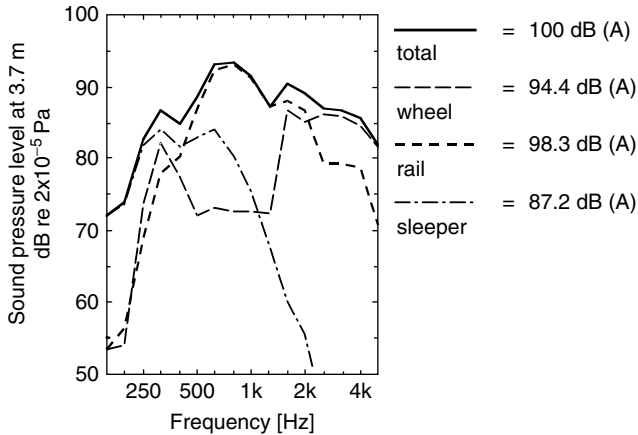
**F. NOISE RADIATION**

The vibrations of the wheel, rail, and sleepers all produce noise. In general, the sound power  $W_{\text{rad}}$  radiated by a vibrating surface of area  $S$  can be expressed as:

$$W_{\text{rad}} = \rho_0 c_0 S \sigma \langle \overline{v^2} \rangle \tag{10.7}$$

where  $\langle \overline{v^2} \rangle$  is the spatially averaged mean-square velocity normal to the vibrating surface,  $\rho_0$  is the density of air,  $c_0$  is the speed of sound in air, and  $\sigma$  is a factor called the radiation efficiency. Thus components radiate large amounts of noise if their vibration is large and/or their surface area is large and/or their radiation efficiency is high. The radiation efficiency is usually close to unity at higher frequencies and smaller than unity at low frequencies (where the radiating object is small compared with the wavelength of sound). Predictions of this factor can be obtained using numerical methods such as the boundary element method or, for simple cases, analytical models.

Figure 10.10 shows predictions of the noise from wheels, rails, and sleepers during the passage of a pair of similar bogies. This is shown in the form of the average sound pressure level at a location close to the track (3 m from the nearest rail). The wheel is the most important source of noise at high frequencies, above about 1.6 kHz. From Figure 10.7 it can be seen that this corresponds to the region in which many resonances are excited in the radial direction. Between approximately 400



**FIGURE 10.10** Predicted noise components from the wheels, rails, and sleepers for 920-mm diameter freight wheels at 100 km/h on a track with moderately soft rail pads.

and 1600 Hz the rail is the dominant source of noise. Here, the rail vibrates at the amplitude of the roughness. The support structure affects the decay with distance and hence the spatially averaged velocity. At low frequencies the sleeper radiates the largest component of noise. Here, the rail and sleeper are well-coupled and have similar vibration amplitudes, but the sleeper has a larger area and a radiation efficiency close to unity, whereas that of the rail reduces below 1 kHz.

Although the details of Figure 10.10 are specific to this combination of wheel and track design, train speed, and roughness spectrum, it is generally the case that the most important source is formed by the sleepers at low frequencies, the rails in the mid frequencies and the wheels at high frequencies. As speed increases the noise spectrum shifts towards higher frequencies, leading to a greater importance of the wheel.

## G. COMPUTER PACKAGES

The complete model for rolling noise that has been described in Section II.A to Section II.F has been implemented in a software package, TWINS<sup>12</sup> that is widely used in the industry. This is a frequency-domain model based on the moving irregularity formulation. It produces estimates of sound power and sound pressure spectra in one-third octave bands and allows the user to study the effect of different wheel and track designs on noise.

This model has also been the subject of extensive validation.<sup>13,14</sup> Comparisons between predictions and measurements for three track types, three wheel types, and four speeds gave overall sound levels that agreed within about  $\pm 2$  dB.<sup>13</sup> These predictions were updated in Ref. 14, along with new measurements for a range of novel constructions. Revisions to the software have improved agreement slightly. Agreement in one-third octave bands had a larger spread of around  $\pm 4$  dB, but this was at least partly due to uncertainties in the measured roughness inputs used in Ref. 13.

## III. REDUCING ROLLING NOISE

From the theoretical understanding it is clear that rolling noise can be reduced by:

1. Controlling the surface roughness.
2. Minimising the vibration response of wheels and tracks by adding damping treatments, by shape optimisation of wheels or rails, or by introducing vibration isolation.
3. Preventing sound radiation, for example, by using local shielding measures.

In each case attention must be given to the presence of multiple sources. If more than one source is important, overall reductions will be limited unless all sources are controlled. For example, if there are initially two sources (wheel and track) that contribute equally and one of them is reduced by 10 dB without affecting the other, the overall reduction will be limited to 2.5 dB.

## A. CONTROLLING SURFACE ROUGHNESS

From the vehicle designer's point of view, the main feature affecting the wheel roughness is the braking system. Traditional tread brakes, in which cast-iron brake blocks act on the wheel tread, lead to the development of high levels of roughness on the wheel running surfaces due to the formation of local hot spots. This can be seen from [Figure 10.3](#), the greatest differences in roughness being at the peak at approximately 6-cm wavelength. This high roughness in turn leads to higher levels of rolling noise. With the introduction of disc-braked vehicles, for example, the Mk III coach in the U.K. in the mid-1970s, it became apparent that disc braking can lead to quieter rolling stock. The difference in rolling noise between the Mk III and its tread-braked predecessor, the Mk II, was about 10 dB, mainly due to the difference in roughness. Modern passenger rolling stock is mostly disc braked for reasons of braking performance and this brings with it lower noise levels than older stock.

However, environmental noise is usually dominated by freight traffic. Freight vehicles are generally noisier and often run at night when environmental noise limits are tighter. For freight traffic in Europe a number of factors have meant that cast-iron brake blocks have remained the standard. These include cost, the longevity of wagons (typically 50 years), and most importantly, the UIC standards for international operation that have required the use of such brakes. However, since 1999 the UIC has been pursuing an initiative to replace cast-iron blocks with so-called K-blocks.<sup>15</sup> The idea is to introduce blocks made of a composite material that do not produce hot spots and therefore leave the wheel relatively smooth. If possible, these should be available as a "retro-fit" with no further modifications to the vehicles. In practice, the implementation of these blocks has a number of side effects including potentially higher wheel tread temperatures and so the development is ongoing with widespread introduction still a number of years away.

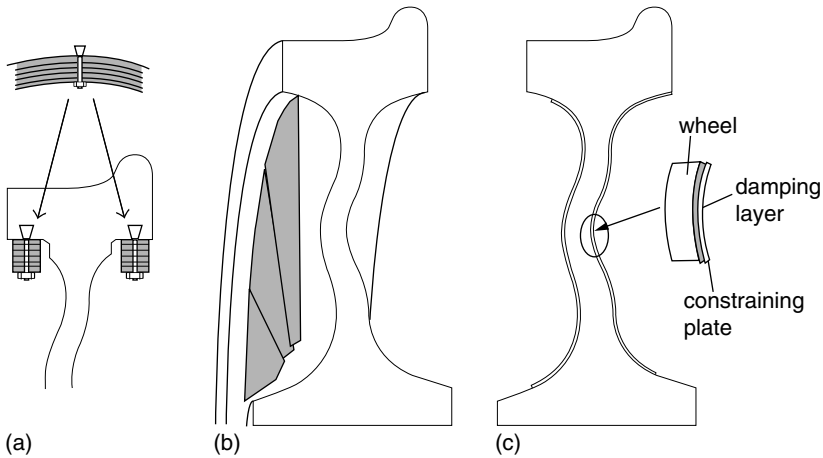
Rail corrugations are also a source of increased noise. A corrugated track can be 10 dB noisier than a smooth one for tread-braked wheels. For disc-braked wheels the difference can be up to 20 dB. Grinding of the rail for acoustic purposes is carried out, for example, in Germany, to maintain special low noise sections of track.

## B. WHEEL-BASED SOLUTIONS

### 1. Wheel Damping

One means of reducing the amount of noise radiated by the wheels is to increase their damping. Impressive reductions in the reverberation of wheels can be achieved by simple damping measures. However, a wheel in rolling contact with the rail is, in effect, already considerably more damped than a free wheel since vibration energy flows from the wheel into the track. To improve the rolling noise performance, the added damping must exceed this effective level of damping, which is one to two orders of magnitude higher than that of the free wheel.

Various devices have been developed to increase the damping of railway wheels by absorbing energy from their vibrations and thereby reducing the noise produced. Examples are shown schematically in [Figure 10.11](#). These include multiresonant absorbers ([Figure 10.11a](#)) which have been used in Germany since the early 1980s and are fitted to many trains including the ICE-1. Noise reductions of 5 to 8 dB are claimed for speeds of 200 km/h.<sup>16</sup> Another commercial form of damper involves multiple layers of overlapping plates known as the shark's fin damper ([Figure 10.11b](#)). Färm<sup>17</sup> found reductions of 1 to 3 dB(A) overall, associated with wheel noise reductions of 3 to 5 dB(A). Constrained layer damping treatments ([Figure 10.11c](#)) consist of a thin layer of



**FIGURE 10.11** Various wheel damping devices used on railway wheels: (a) tuned resonance devices, (b) shark's fin dampers, (c) constrained layer damping.

viscoelastic material applied to the wheel and backed by a thin stiff constraining layer (usually metal). Such a treatment was used on the class 150 DMU in the U.K. in the late 1980s and was applied to the whole vehicle fleet to combat a particularly severe curve squeal problem excited by contact between the wheel flange and the check rail. By careful design, sufficient damping can be achieved using constrained layer damping to also make significant reductions in rolling noise.<sup>18,19</sup>

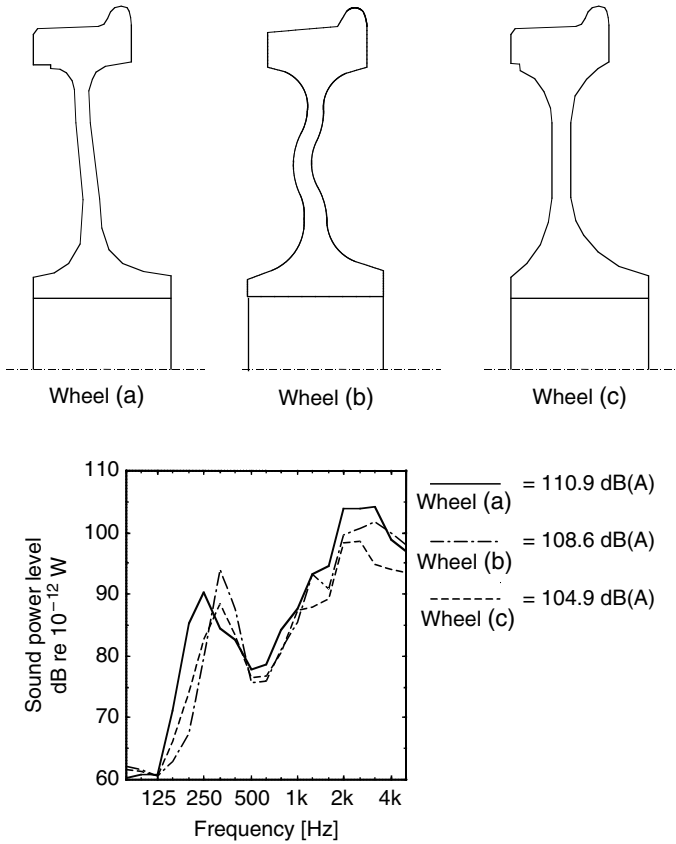
## 2. Wheel Design

Reductions in the wheel component of radiated noise can also be achieved by careful attention to the wheel cross-sectional shape. In recent years, manufacturers have used theoretical models such as TWINS<sup>12</sup> to assist in designing wheels for low noise.

As an example of the difference that the cross-sectional shape can have, three wheels are shown in [Figure 10.12](#). Wheel (a) is a German Intercity wheel, (b) is a UIC standard freight wheel, and wheel (c) was designed several years ago by the Technical University of Berlin on the basis of scale model testing.<sup>16</sup> [Figure 10.12](#) also shows the predicted noise components from the wheel in each case. The track component of noise (not shown) is not affected by these changes and remains the dominant source up to 1 kHz.

These results show that a straight web (wheel c) is beneficial compared with a curved web (wheel b). This is because the radial and axial motions are decoupled for a straight web. However, it is not always possible to use straight webs if tread brakes are used as the curve is included in the web to allow thermal expansion. Wheel (a) is particularly noisy, the main difference between this and wheel (c) being the transition between the inside of the tyre and the web and the web thickness. Its web is also slightly angled. Increasing the web thickness, and particularly the transition between the tyre and web, are effective means of reducing noise but also lead to increased unsprung mass. It is the case that wheels with profiles similar to (a) have shown appreciable rolling noise reductions by the addition of absorbers, whereas wheels such as (b) have shown much smaller reductions.

Another aspect of wheel design that can be used to reduce noise is the diameter. Smaller wheels have higher resonance frequencies, so it is possible by reducing the diameter to move most of the resonances out of the range of excitation (i.e., above approximately 5 kHz).<sup>20,21</sup> The upper frequency itself is somewhat increased for a smaller wheel due to a shift in the contact patch filter, but this effect is much less significant than the shift in resonance frequencies. The trend in recent years towards smaller wheels for other reasons is therefore advantageous for noise. This also



**FIGURE 10.12** TWINS predictions of wheel sound power from three types of wheel at a train speed of 160 km/h for the same roughness spectrum in each case, typical of a disc-braked wheel.

negates the increase in unsprung mass caused by increases in thickness. However, if the wheel size is reduced too much, the track noise will increase due to the reduction in contact filter effect.<sup>22</sup>

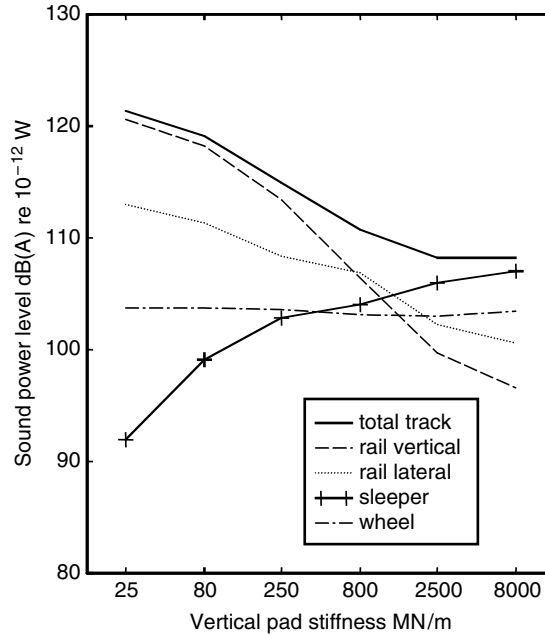
In the SILENT FREIGHT EU project a shape-optimised wheel was designed that allowed for tread braking.<sup>23</sup> It had a diameter of 860 mm, compared with the reference wheel, which was 920 mm, and it had a somewhat thicker web. The reduction in diameter was limited by the desire to allow retro-fitting to bogies intended for 920 mm wheels. Reductions of 3 dB in wheel component of noise were obtained from these changes.

## C. TRACK-BASED SOLUTIONS

### 1. Low Noise Track

To achieve significant reductions in overall noise, it is usually not sufficient to deal only with the wheel noise. There must be a corresponding reduction in noise from the track vibration. Two very important parameters of the track, that affect its noise emission and that are related to one another, are the stiffness of the rail pad and the decay rate of vibrations along the rail. A stiff rail pad causes the rail and sleeper to be coupled together over a wide frequency range. Conversely, a soft pad isolates the sleeper for frequencies above a certain threshold.

The lower the stiffness of the rail pad, the lower this threshold frequency. Soft rail pads therefore effectively isolate the sleepers and the foundation from the vibration of the rail, reducing the component of noise radiated by vibration of the sleepers. Part of the designed role of the rail pad



**FIGURE 10.13** Effect of rail pad stiffness on predicted components of rolling noise.

is to protect the sleeper and ballast from high impact forces. For this reason softer rail pads have become more commonplace in recent years. Unfortunately, softer rail pads also cause the vibration of the rail to propagate with less attenuation. As a greater length of rail vibrates with each wheel, this means more noise is generated by the rail, as shown in Figure 10.13. There is thus a compromise to be sought between the isolating and attenuating properties of the rail pad.<sup>24</sup>

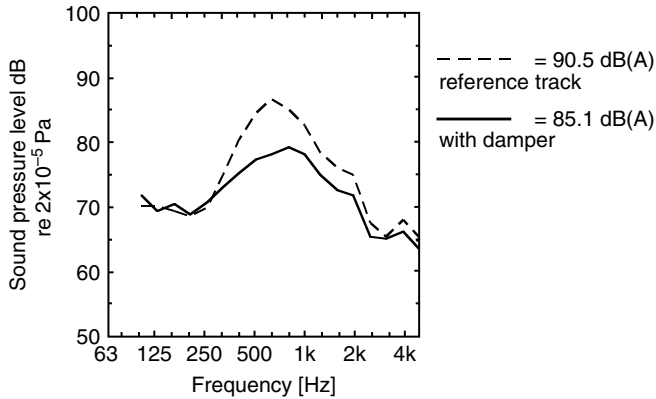
The recent E.U.-funded research project SILENT TRACK successfully developed and demonstrated low noise technology for the track. The most successful element was a rail damper. Multiple blocks of steel are fixed to the sides of the rail by an elastomer and tuned to give a high damping effect in the region of 1 kHz. This allows a soft rail pad to be used, to give isolation of the sleepers, while minimising the propagation of vibration along the rail.<sup>25</sup>

Figure 10.14 shows the noise reduction achieved in the field tests. In this case, a low noise wheel was used for the comparison to minimise the effect of the wheel on the total noise, but, even so, some wheel noise was present at high frequencies. The overall reduction in *track* noise is approximately 6 dB.

Tests with an optimum pad stiffness have been less successful. Although the effect of pad stiffness has been clearly demonstrated in field tests, the optimum for noise radiation is too stiff to be acceptable for other reasons, particularly track damage protection. Stiff pads are also believed to lead to a higher likelihood of corrugation growth, which in the long term, has a negative effect on the noise. The analysis of the acoustic performance of pads with different stiffnesses is further complicated by their load-dependent characteristics and other factors such as temperature variation.

## 2. Slab Tracks

Tracks mounted on concrete slabs have become more commonplace in the last few years, notably in Germany on the high speed lines. Such tracks are generally found to be noisier than conventional ballasted track, typically by 3 to 5 dB. This can be attributed to two features of such tracks. First, they tend to be fitted with softer rail fasteners in order to introduce the resilience normally



**FIGURE 10.14** Measured noise reduction from SILENT TRACK rail damper during the passage of a low noise wheel at 100 km/h.

given by the ballast. Second, they have a hard sound-reflecting surface, whereas ballast has an absorptive effect. The latter affects the overall noise by 1 to 2 dB.

A number of mitigation measures have been studied in Germany, in which absorbent material is added to the upper surface of the slab. This has the effect of reducing the reflections of sound from the slab surface. Where it is also possible to introduce some shielding of the rail noise, for example, by an integrated minibarrier, additional attenuation is possible. Such treatments have been found to reduce noise levels from slab track back to those of ballasted track.

For street-running trams a number of embedded rail systems are used, for example in Manchester, Sheffield, and Birmingham in the U.K. At first sight an embedded rail might be expected to be silent, as the rail is mostly hidden and therefore should not produce sound. In practice, the rail head is visible, and both it and the embedding material around it vibrate and produce sound.

Embedded rail systems offer the possibility of good rail attenuation rates, owing to the damping effect of the embedding material around the rail. They can also be constructed with relatively soft supports and therefore offer the potential to produce good vibration isolation.

**D. LOCAL SHIELDING AND BARRIERS**

Another means of reducing the sound radiation is to use a barrier. The efficiency is improved by placing the barrier as close as possible to the source. In the SILENT FREIGHT project, it was demonstrated that, at least for certain types of wheel, a shield mounted on the wheel covering the web can reduce the noise. A more general solution is to place an enclosure around the bogie. If used in combination with low barriers very close to the rail reductions of up to 10 dB can be achieved.<sup>26</sup>

Bogie shrouds and low barriers were also tested in SILENT FREIGHT and SILENT TRACK, but in these cases the objective was to find a combination that satisfied international gauging constraints. Unfortunately, this meant that the overall reduction was limited to less than 3 dB due to the inevitable gap between the top of the barrier and the bottom of the shroud.<sup>27</sup>

There are many other practical difficulties in enclosing the bogies, such as ventilation for the brakes and access for maintenance. Nevertheless, such vehicle-mounted screens are common on trams. Bogie fairings have also been tested on high-speed trains but in this case the objective was to reduce aerodynamic noise.

Conventional noise barriers at the trackside are used widely in some European countries and in Japan. Reductions of 10 to 20 dB are achievable, depending on the height of the barriers, but they

are expensive and visually intrusive especially if taller than about 2 m. Cost-benefit studies have shown that noise reduction at source can be cost-effective compared with barriers or, in combination, can allow the use of lower barriers for the same overall effect.<sup>28</sup>

## IV. IMPACT NOISE

### A. INTRODUCTION

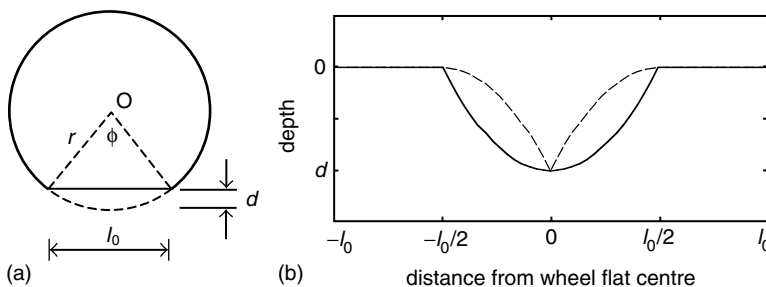
In the previous sections, noise due to random irregularities on the railhead and wheel tread has been considered. As well as this, larger discrete features occur on the running surfaces such as rail joints, gaps at points and crossings, dipped welds, and wheel flats. These cause high interaction forces, and consequently, noise. In some cases loss of contact can occur between the wheel and rail followed by large impact forces. Noise from such discrete features is often referred to as impact noise. Whereas rolling noise can be predicted using a linearised contact spring, in order to predict impact forces and noise the nonlinear contact stiffness must be included, for example, using a Hertzian deflection model (Chapter 4).

Early models for impact noise were essentially empirical.<sup>29</sup> To predict impact forces, time-domain models incorporating the nonlinearities in the contact zone, have been used, for example, by Clark et al.<sup>30</sup> and Nielsen and Igeland.<sup>31</sup> These models contain large numbers of degrees of freedom to represent the track. Nevertheless, they are limited to a maximum frequency of around 1500 Hz. In order to model impact noise up to approximately 5 kHz, simplified models of the wheel and rail have been used in a time-stepping model in order to determine the effects of the nonlinearities.<sup>32</sup> These are then used together with the TWINS model to predict the noise radiation.

### B. WHEEL FLATS

A wheel flat is an area of the wheel tread that has been worn flat, as shown schematically in Figure 10.15a. This usually occurs because the brakes have locked up under poor adhesion conditions at the wheel–rail contact, for example, due to leaves on the railhead during the autumn. Wheels with flats produce high levels of noise and impact loading of the track that can lead to damage of track components. Typically flats can be approximately 50 mm long, in extreme cases up to 100 mm. After their initial formation, flats become worn, i.e., rounded at their ends due to the high load concentration on the corners. A worn flat of a given depth is longer than the corresponding new flat.

Wheel flats introduce a relative displacement input to the wheel–rail system in the same way as roughness. The profile shape can be seen to correspond to a circular arc dip in the railhead.



**FIGURE 10.15** (a) An idealised flat of length  $l_0$  and depth  $d$ ; (b) — profile shape, - - - after geometric filtering. Source: From Wu, T. X. et al., *J. Sound Vib.*, 251, 115–139, 2002, Elsevier. With permission.



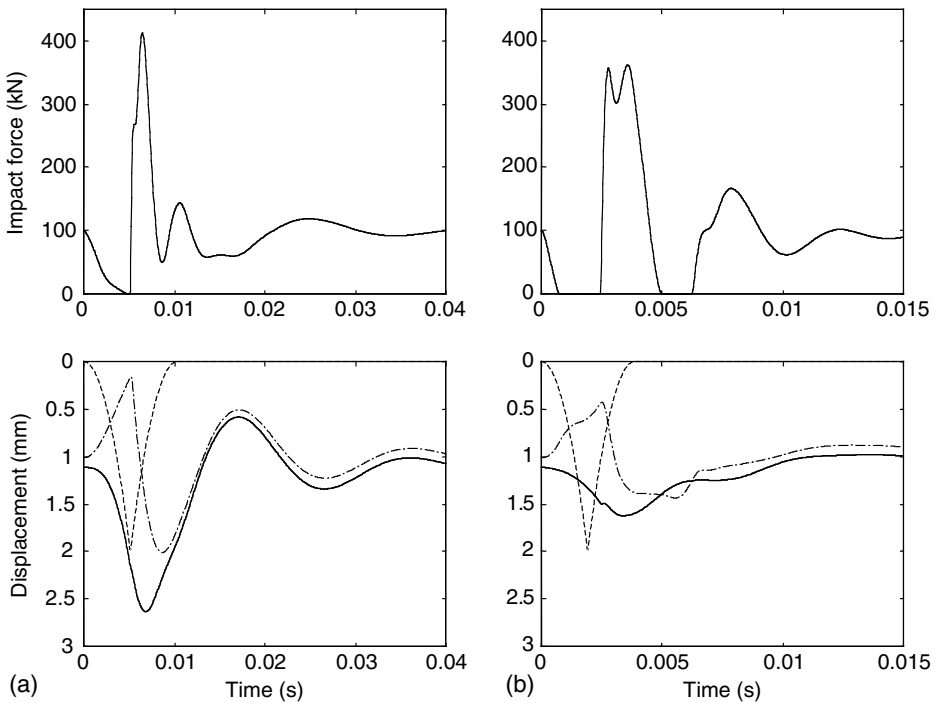
However, owing to the geometry of the wheel and rail surfaces, the actual displacement input is modified by the wheel curvature. For the idealised flat, shown in Figure 10.15a, the wheel first pivots downwards on the front corner of the flat, then pivots upwards again on the rear corner.<sup>32</sup> The resulting relative displacement input experienced by the wheel–rail system is shown in Figure 10.15b. In Ref. 32 it is shown that a worn wheel flat can be represented by a curve of a similar shape to that in Figure 10.15b but elongated.

Figure 10.16 shows examples of the calculated response of the wheel–rail system to a new wheel flat of depth 2 mm (length 86 mm) for a nominal contact force of 100 kN. The model used here represents the wheel as a mass and spring and the track by a simple state-space model fitted to the track mobility.<sup>32</sup>

When the indentation (relative displacement input due to the wheel flat) appears between the wheel and rail, the wheel falls and the rail rises. Since the wheel and rail cannot immediately follow the indentation owing to their inertia, the contact force is partly unloaded. At a train speed of 30 km/h (Figure 10.16a), full unloading first occurs.

After the relative displacement input reaches its maximum, the contact force increases rapidly until it reaches its peak (the wheel is now pivoting about the trailing edge of the flat). The peak force is here about four times as large as the static load. As the speed increases, contact is lost for longer periods during the unloading phase. At 80 km/h (Figure 10.16b), a second loss of contact can be seen to occur. However, the second impact is much smaller than the first one.

Comparisons with measured impact forces<sup>33</sup> suggest that the simplified geometry indicated in Figure 10.15 leads to overestimates of the contact force. Measured wheel flat profiles are required to give more accurate predictions.



**FIGURE 10.16** Predicted wheel–rail interaction and displacements of wheel and rail due to 2 mm newly formed wheel flat. (a) At train speed 30 km/h; (b) at 80 km/h: — wheel displacement, - - - rail displacement, ... relative displacement excitation. *Source:* From Wu, T. X. et al., *J. Sound Vib.*, 251, 115–139, 2002, Elsevier. With permission.

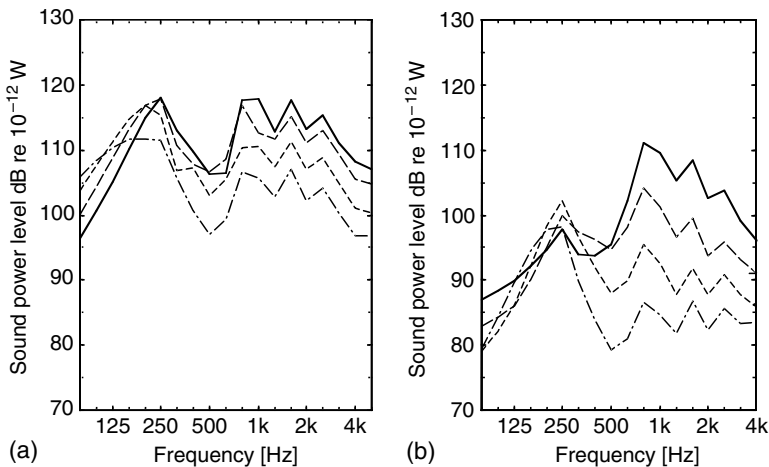
### C. PREDICTING IMPACT NOISE FROM WHEEL FLATS

It is not possible to use the contact force obtained from the impact model and apply it directly within the TWINS model, because the predicted interaction force is very sensitive to details of the wheel and track dynamics used in its prediction. With a modal wheel model, the force will have strong dips at the wheel resonance frequencies. The wheel response has only shallow peaks, just above the resonance frequencies. The interaction with the track thereby introduces apparent damping to the wheel.

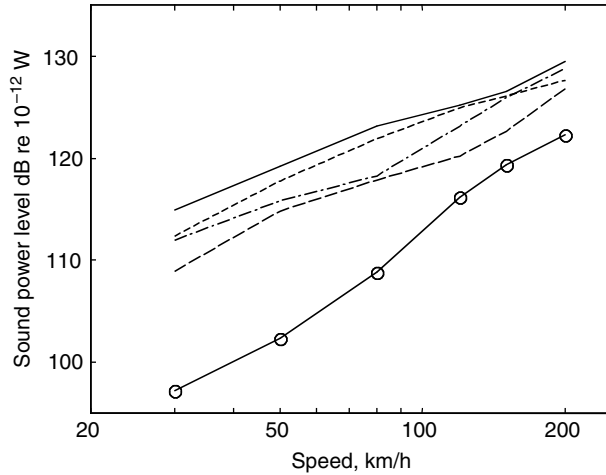
A hybrid approach has therefore been developed,<sup>32</sup> whereby an equivalent roughness spectrum is derived. This is defined such that the contact force spectrum obtained using the above *nonlinear* model is identical to that obtained using a *linear* model excited by the equivalent roughness spectrum. At this stage the wheel and track are represented by the same simple elements as above in both cases. The equivalent roughness spectrum can then be used as the input to a more detailed linear frequency-domain model, such as the TWINS model, to predict the noise due to the impact.

Example results are given in Figure 10.17a. This shows the sound power due to one wheel and the associated track vibration for a 2-mm deep new wheel flat at different speeds for 100 kN wheel load. Results correspond to the average over a whole wheel revolution. Figure 10.17b shows, for comparison, corresponding results for roughness excitation due to a moderate roughness (tread-braked wheel roughness). As the speed increases, the noise at frequencies above approximately 200 to 400 Hz increases in both cases. The increase in rolling noise with increasing speed is greater than that due to the flat. For the wheel flats considered here, the noise generated exceeds that due to the tread-braked wheel roughness at all speeds and in all frequency bands, although the noise due to roughness increases more rapidly with speed so that at sufficiently higher speeds it can be expected to dominate. For corrugated track, the noise due to roughness exceeds that due to wheel flats at 120 km/h.

Figure 10.18 provides a summary of the variation of the overall A-weighted sound power level with train speed. The predicted noise level due to conventional roughness excitation increases at a rate of approximately  $30 \log_{10} V$ , where  $V$  is the train speed, whereas the noise due to flats increases at an average of around  $20 \log_{10} V$  once loss of contact occurs. For example, loss of contact was found to occur for the newly formed 2-mm deep flat at speeds above 30 km/h and for a rounded 2-mm flat above 50 km/h. This variation with speed indicates that the radiated sound due to wheel flats continues to increase with increasing speed, even though loss of contact is occurring.



**FIGURE 10.17** Sound power level due to wheel and track: (a) 2 mm new wheel flat; (b) rolling noise from moderate roughness. — · — · — 30 km/h, ... 50 km/h, - - - 80 km/h, — 120 km/h. *Source:* From Wu, T. X. et al., *J. Sound Vib.*, 251, 115–139, 2002, Elsevier. With permission.



**FIGURE 10.18** Sound power radiated by one wheel and the associated track vibration: - - - 1 mm rounded flat, ... 2 mm rounded flat, - . - . 1 mm new flat, — 2 mm new flat, o—o rolling noise due to tread-braked wheel roughness. *Source:* From Wu, T. X. et al., *J. Sound Vib.*, 251, 115–139, 2002, Elsevier. With permission.

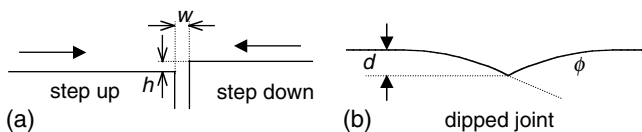
Impact noise from wheel flats is found to depend on the wheel load. The increase in noise between a load of 50 and 100 kN is about 3 dB. In contrast, the rolling noise due to roughness is relatively insensitive to wheel load.

### D. RAIL JOINTS

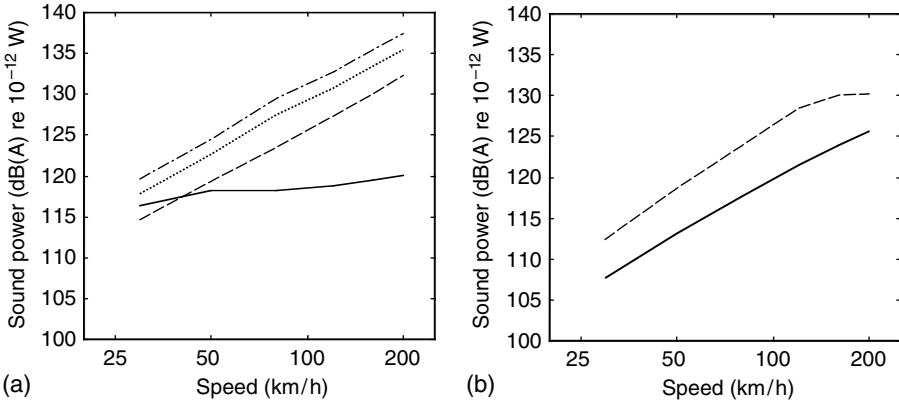
In a similar way to wheel flats, rail joints provide discrete inputs to the wheel–rail system that induce quite large contact force variations. Rail joints can be characterised by a gap width and a step height (either up or down) (see Figure 10.19a). Moreover, the rail often dips down to a joint on both sides (Figure 10.19b). Such dips are also present at welds, and are usually characterised in terms of the angle at the joint.

A similar approach has been used as that above to study the effects of rail joints.<sup>34,35</sup> The sound radiation was calculated using the same hybrid method as for the wheel flats. It was found, for realistic parameter values, that the gap width is insignificant compared with the step height and dip angle.

Results are shown in Figure 10.20a for undipped rail joints in the form of the total A-weighted sound power emitted by the wheel and rail during 1/8 sec. The results for a step-down joint are found to be virtually independent of the step height (only results for one value are shown) and also change very little with train speed. However, for step-up joints both the peak contact force and the sound power level increase with step height and with train speed. The sound power level from a single joint has a speed dependence of around  $20 \log_{10} V$ .



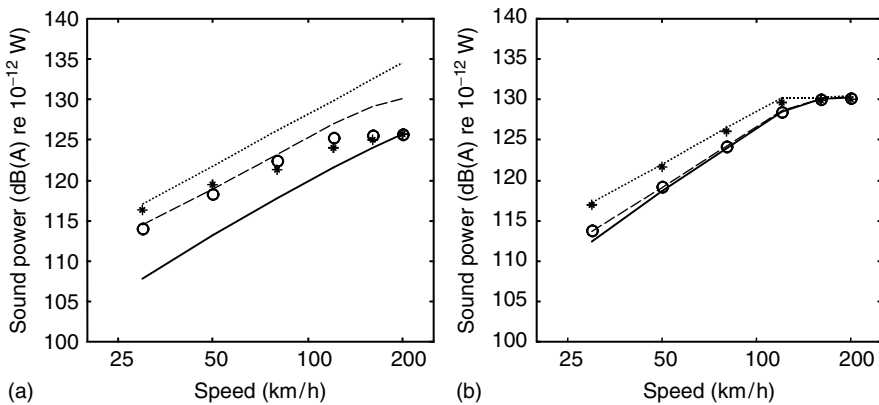
**FIGURE 10.19** Idealised rail joints showing (a) step height  $h$  and gap width  $w$ , (b) dip height  $d$  and angle  $\phi$ .



**FIGURE 10.20** A-weighted sound radiated by one wheel and the associated track vibration during 0.125 sec due to a wheel passing over: (a) flat rail joints, - - - 1 mm step-up, ... 2 mm step-up, - . - . 3 mm step-up, — 2 mm step-down; (b) dipped rail joints with no height difference, — 5 mm dip, - . - . 10 mm dip (all with 7 mm gap).

In Figure 10.20b, results are given for dipped joints with no height difference. Here a dip of 5 or 10 mm is considered as a quadratic function over a length of 0.5 m either side of the joint. A dip of 5 mm corresponds to a joint angle of 0.04 rad which is large although within a typical range, a dip of 10 mm corresponds to 0.08 rad which is severe. The 10-mm dip produces a similar noise level to a 1-mm step-up undipped joint, although for speeds above 120 km/h the noise level from the dip joint becomes independent of train speed.

Figure 10.21 shows the predicted noise for joints with both dipped rails and steps. The noise radiation generally increases with speed, regardless of whether loss of contact occurs. For the 5-mm dip, the noise level increases by 8 dB when the step height increases from 0 to 2 mm. For the step-down joints, the noise level is higher than without a step, although at higher speeds the dip has more effect than the step. The results for the 10-mm dip are similar for both step-up and step-down joints, indicating the dominance of the dip in this case.



**FIGURE 10.21** A-weighted sound power radiated by one wheel and the associated track vibration during 0.125 sec due to a wheel passing over different rail joints with 7 mm gap and 5 or 10 mm dip: (a) for 5 mm dip; (b) for 10 mm dip. ... 2 mm step-up, - - - 1 mm step-up, — no height difference, \* 2 mm step-down, o 1 mm step-down.

To compare these results with typical rolling noise results, the time base of the joint noise should be adjusted to the average time between joints. This shows<sup>34</sup> that rolling noise due to the tread-braked roughness considered above is similar to the average noise due to 5-mm dipped joints with no height difference (Figure 10.20b). With a height difference of 2 mm the average noise predicted from the joints increases to almost 10 dB greater than the rolling noise. Moreover, since the time between rail joints decreases as train speed increases, it is also found that the average noise level from joints increases at about  $30 \log_{10} V$ , similar to rolling noise.

## E. REDUCING IMPACT NOISE

In order to reduce impact noise it is clearly desirable to remove the cause, if this is possible. Wheel flats can be largely prevented by installation of wheel-slide protection equipment. Monitoring equipment is now widely used to identify wheels with flats, to allow them to be removed from service as quickly as possible for reprofiling. On main lines, jointed track has been mostly replaced by continuously welded rail in the last 30 years, although inevitably, joints such as expansion joints, track-circuit insulating joints, and points and crossings remain. Even so, measures such as swing-nose crossings allow the impact forces, and thus noise, to be minimised. Attention should also be given to ensuring that welded rail joints are as level as possible by using rail straightening equipment.

As the mechanism of impact noise is a vertical relative displacement excitation, counter-measures that are effective for rolling noise, such as are discussed in Section III, can be expected to work equally well for impact noise. This includes, for example, wheel damping, wheel shape optimisation, rail damping, and local shielding.

## V. CURVE SQUEAL

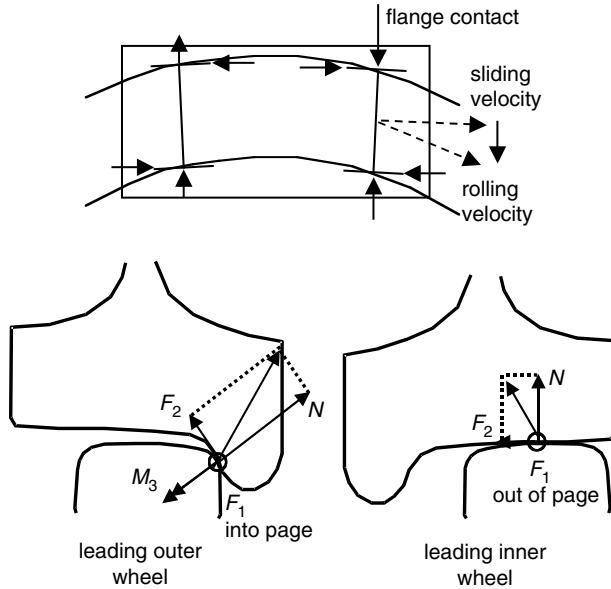
### A. MECHANISM OF SQUEAL NOISE GENERATION

Railway vehicles travelling around tight curves can produce an intense squealing noise. This is a particular problem where curved track exists in urban areas and it has been found to be annoying to both residents and railway passengers.

When a railway wheelset in a bogie traverses a curve it is unable to align its rolling direction tangentially to the rail (Figure 10.22). In sharp curves, this misalignment leads to large creep forces at the wheel–rail interface. The leading inner wheel of a bogie has its contact point with the rail towards the field side of the tread and experiences high lateral creepage. The leading outer wheel tends to be in flange contact, with the resultant lateral force acting inwards to ensure that the wheelset remains on the track. Longitudinal and spin creep forces also act as shown in Figure 10.22.

Figure 10.23a shows a typical “creep curve” relating creep force to creepage. At low values of creepage the magnitude of the creep force increases linearly. At high values of creepage the force becomes saturated, with a maximum value of  $\mu_0 N$ , where  $\mu_0$  is the friction coefficient and  $N$  is the normal load. In practice, however, the friction coefficient  $\mu$  is not a constant. It is usually recognised that dynamic or sliding friction coefficients are smaller than static ones. In fact, the friction coefficient depends on the sliding velocity, decreasing as the velocity increases. Thus, as creepage increases beyond the saturation point, the creep force once more reduces in amplitude (see Figure 10.23b). It is this falling amplitude at high creepage that is believed to be the main reason for the unstable dynamic behaviour leading to squeal noise.

By analogy with a damper, which gives a reaction force that is proportional to the relative velocity, the falling creep curve can be considered as a negative damping. Thus, the reaction force decreases as the relative velocity increases. Since wheel modes have very low levels of damping (see Section II), if this negative damping exceeds a certain level, it causes instability of the wheel modes making them prone to squeal.

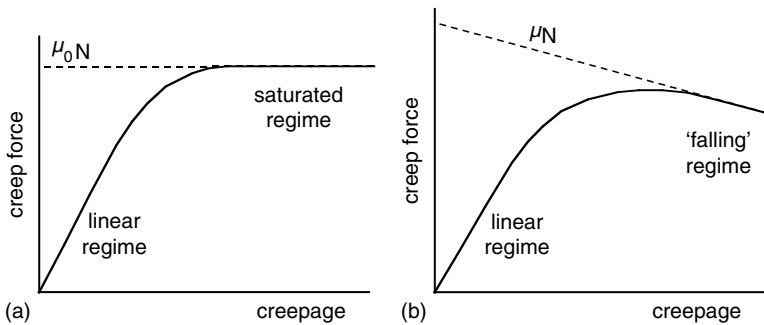


**FIGURE 10.22** Schematic view of forces acting on wheels of a bogie in a curve.  $N$  is normal load,  $F_2$  is lateral creep force,  $F_1$  longitudinal creep force, and  $M_3$  is spin moment.

There are two main types of curve squeal, characterised by the mechanisms of excitation:

1. Stick–slip excitation owing to lateral slip of the wheel due to its alignment.
2. Squeal due to wheel flange contact with the rail.

Observations indicate that the highest squeal noise amplitude is usually generated by the *leading inner* wheel of a four-wheeled bogie or two-axle vehicle. This noise is associated with stick–slip lateral motion at the contact between the wheel and the rail. The fundamental frequency of such squeal noise corresponds to a natural frequency of the wheel and is often in the range 200 to 2000 Hz. The wheel modes excited in this case are axial modes with no nodal circle and their maximum amplitude at the wheel tread (see first row of [Figure 10.6](#)).



**FIGURE 10.23** A typical creep force–creepage relationship: (a) for constant friction coefficient; (b) for velocity-dependent friction coefficient.

Contact between the wheel flange and the rail, which occurs at the *leading outer* wheel (and possibly the *trailing inner* wheel) in sharp curves, has generally been found to reduce the likelihood of stick–slip squeal due to lateral slip at this wheel. For example, Remington<sup>36</sup> concluded from laboratory experiments that flange contact reduces the level of squeal noise. However, it is thought that flange contact may generate a different form of squeal noise. Compared with squeal due to lateral slip, this generally has a considerably higher fundamental frequency, may have a lower level, and is often more intermittent in nature. Nevertheless, it can be a source of considerable annoyance. It is usually associated with flange contact; either with the outer running rail, with check rails in sharp curves, or wing rails in points and crossings. Compared with squeal due to lateral slip, flange noise has received much less attention.

Theoretical models for curve squeal have been developed by various authors. Rudd<sup>37</sup> (see also Remington<sup>36</sup>) indicated that instability of the lateral friction force was the most likely cause of squeal and gave a simple model. Fingberg<sup>38</sup> and Périard<sup>39</sup> have extended this basic model by including better models of the wheel dynamics, the friction characteristic, and the sound radiation from the wheel. Time-domain calculations allowed the squeal magnitude to be predicted as well as the likelihood of squeal to be determined. Heckl<sup>40</sup> also studied squeal using a simplified model and provided experimental validation using a small-scale model wheel.

De Beer et al.<sup>41</sup> extended these models, based on excitation by unstable lateral creepage, to include feedback through the vertical force as well as through the lateral velocity. Their model consists of two parts: a first part, in the frequency domain, can be used to determine instability and to predict which mode is most likely to be excited, and a second part, in the time domain, calculates the amplitude of the squeal noise.

This model has been extended further to allow for an arbitrary contact angle and to include lateral, longitudinal, and spin creepage.<sup>42</sup> This allows it to be applied to flange squeal as well as squeal due to lateral creepage.

## B. REDUCING SQUEAL NOISE

In discussing solutions for curve squeal it is of little value to quote decibel reductions. The nature of the instability is such that effective measures usually eliminate the squeal rather than reduce it. Curve squeal tests are also extremely unreproducible due to a high sensitivity to parameters such as temperature, humidity, train speed, track geometry, and wheel and rail wear.

Known solutions for curve squeal include lubrication using either grease or water or the application of friction modifiers that reduce the difference between static and sliding friction coefficients. If lubricants are used, it must be ensured that they do not lead to loss of adhesion as this could compromise safety. Grease is therefore only applied to the rail gauge corner or wheel flange. Although this may not be the primary cause of squeal noise, by modifying the curving behaviour this can nevertheless reduce the occurrence of squeal. Water sprays have also been used effectively in a number of locations.

Friction modifiers act by reducing or eliminating the falling friction characteristic without reducing the level of friction. These can be applied either to the track at the entrance to a curve, or on the vehicle. They have been shown to be very effective in eliminating squeal and can be applied to the top of the railhead without compromising traction or braking.<sup>43</sup>

Wheel damping treatments are also known to reduce the occurrence of squeal. In this case a small increase in the level of damping can be effective in eliminating squeal. In addition to the forms of damping discussed in Section III.B, ring dampers have been used as a simple means of increasing the damping of a wheel.<sup>44,45</sup>

Effective solutions can also be sought in the design of vehicles for curving in order to reduce the creepages. Unfortunately, this is often in conflict with the design of bogies for stability at high speed.

## VI. OTHER SOURCES OF NOISE

### A. AERODYNAMIC NOISE

Aerodynamic sound sources increase in sound power more rapidly with speed than mechanical sources. For an aeroacoustic monopole source, such as the pulsating flow from an exhaust pipe, the sound power increases with flow speed according to the fourth power of the speed. This means that the sound power level increases at a rate of  $40 \log_{10} V$ . For a dipole-type source, such as the tones generated by vortex shedding from a cylinder or turbulence acting on a rigid surface, the rate is  $60 \log_{10} V$ , whereas for a quadrupole source such as free turbulent flow the rate is  $80 \log_{10} V$ .

Aerodynamic sources become dominant for exterior noise of trains above a speed of approximately 300 km/h. Below this speed noise levels increase at about  $30 \log_{10} V$  (typical of rolling noise) whereas above this speed they increase at a rate of  $60 \log_{10} V$  or more.<sup>46</sup>

Where noise barriers are placed alongside the track, the wheel–rail noise may be attenuated by 10 to 15 dB, while leaving the aerodynamic sources from the upper part of the train and pantograph exposed. This causes aerodynamic noise to become important at lower speeds. Aerodynamic sources are also important for interior noise in high-speed trains, particularly the upper deck of double-deck trains where rolling noise is less noticeable.

Turbulent airflow, which can be caused by many different parts of a rail vehicle, is an important source of aerodynamic noise.<sup>47</sup> The locations of a number of sources have been identified and their strengths quantified in studies using specialised microphone arrays.<sup>48</sup> Important sources are found to fall into two main categories.<sup>49</sup> The first category, which is dipole in nature, is generated by airflow over structural elements: the bogies, the recess at the intercoach connections, the pantograph and electrical isolators on the roof, and the recess in the roof in which the pantograph is mounted. In addition, the flow over the succession of cavities presented by louvred openings in the side of locomotives is a source of aerodynamic noise, the form of which depends on the length and depth of the cavity. In the second category, which may have a dipole or quadrupole nature, noise is created due to the turbulent boundary layer.

Empirically based models for each source of aerodynamic noise from trains can be derived if the locations and source strengths are experimentally determined. Measurements may be complemented by the use of computational fluid dynamics (CFD) models. While a working theoretical model for the aerodynamic sources from trains is not yet available, it is the objective of current research.<sup>49</sup>

### B. POWER UNIT NOISE

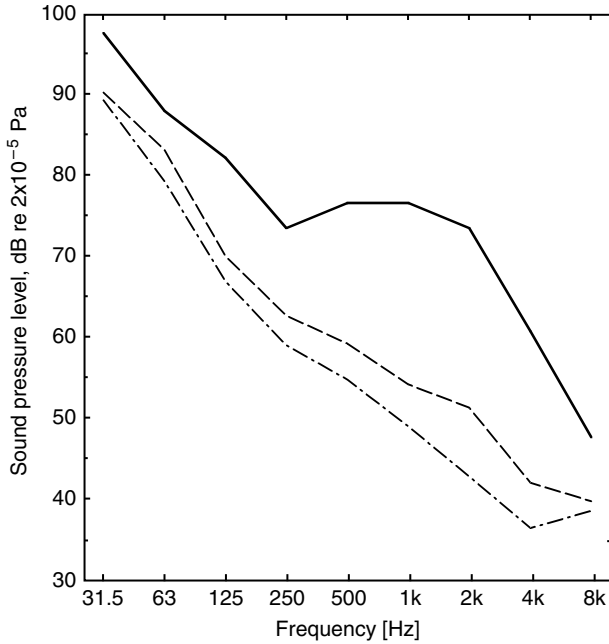
Power units on trains are generally either electric or diesel. Noise from diesel locomotives is mostly dominated by the engine and its intake and exhaust. Space restrictions often limit the ability to silence the exhaust adequately, although in modern locomotives this has been given serious attention. On electrically powered stock, and on diesels with electric transmission, the electric traction motors and their associated cooling fans are a major source of noise. Most sources of noise from the power unit are largely independent of vehicle speed, depending rather on the tractive effort required. The whine due to traction motors is an exception to this.

## VII. VEHICLE INTERIOR NOISE

### A. VEHICLE INTERIOR NOISE LEVELS

All the noise sources discussed above are also of relevance to interior noise in trains.<sup>50</sup> Noise is transmitted from each of these sources to the interior by both airborne and structure-borne paths, with structure-borne transmission often dominant at low frequencies and airborne transmission at high frequencies. The noise from the wheel–rail region is often the major source. Additionally, on





**FIGURE 10.24** Octave band spectra measured inside British rolling stock at 145 km/h (from Ref. 51). — Mk 1 vehicle (81 dB(A)), - - - Mk 2d vehicle (63 dB(A)), - . - . - Mk 3 vehicle (59 dB(A)).

vehicles with underfloor diesel engines, noise from the engine can be significant. Noise from the air-conditioning system, where this is present, can also require consideration in rolling stock. There is often very limited space in which to package the air-conditioning unit and ducts.

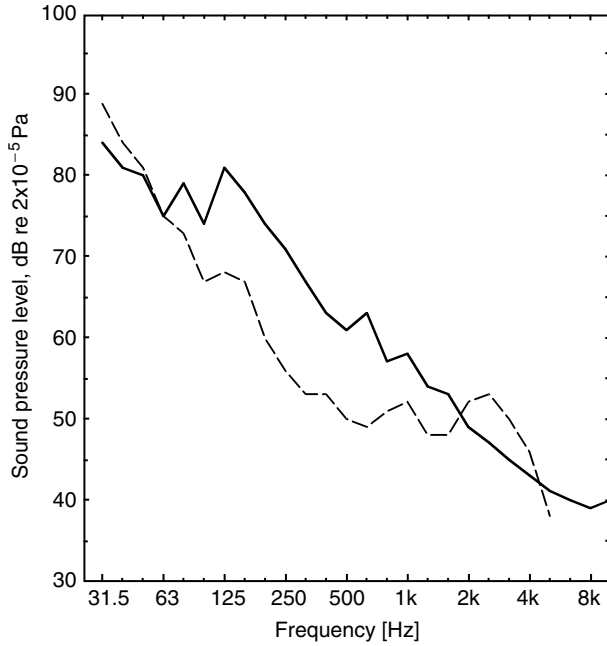
Examples of noise spectra inside British vehicles are shown in Figure 10.24 (taken from Eade and Hardy<sup>51</sup>). The Mk 1 vehicle was constructed in 1960 and had opening windows; the Mk 2d coach was introduced in 1970 and had sealed windows and air-conditioning; the Mk 3 coach was introduced in 1975 and has a similar interior to the Mk 2d but has disc brakes. The more modern stock can be seen to produce a considerable improvement at higher frequencies, but the differences at low frequencies are much more modest so that the low frequency noise now accounts for a much higher proportion of the total.

Other results are given in Figure 10.25 and Figure 10.26. The results in Figure 10.25 show that modern high-speed trains are quieter at 300 km/h than a conventional “rail car” at a lower speed.<sup>52</sup> Figure 10.26 shows measured spectra in an open saloon coach both in the open and in a tunnel.<sup>53</sup> The noise levels can be seen to increase considerably in the mid-frequency range in a tunnel owing to a greater contribution from the walls, windows, and roof.

**B. MEASUREMENT QUANTITIES FOR INTERIOR NOISE**

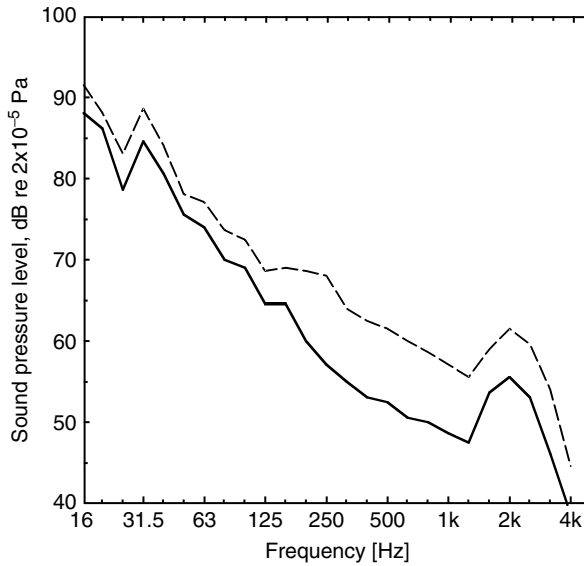
Conventionally, the A-weighted sound pressure level has been used to specify acceptable levels inside vehicles. Internationally agreed limits are 68 dB(A) in second class and 65 dB(A) in first class.<sup>53</sup> However, as seen above, the spectrum of noise inside trains contains considerable energy at low frequency. This low frequency sound energy can be a source of human fatigue, but is not effective in masking speech, for which noise in the range 200 to 6000 Hz is most effective.

Passenger requirements for noise inside a train vary from one person to another.<sup>54</sup> Clearly, it is desirable that the noise should not interfere with conversation held between neighbours. However, particularly for a modern open saloon-type vehicle, silence would also not be the ideal.



**FIGURE 10.25** One-third octave band spectra measured inside French rolling stock (from Ref. 52). — railcar at 160 km/h (72 dB(A)), - - - TGV at 300 km/h (64 dB(A)).

There should be sufficient background noise so that passengers talking do not disturb other passengers further along the vehicle (people talking loudly into mobile phones are a particular source of annoyance). According to Ref. 55, for example, the interior noise level should be at least 60 dB(A) to avoid disturbance by other passengers.



**FIGURE 10.26** One-third octave band spectra measured inside German rolling stock at 200 km/h on ballasted track (from Ref. 53). — in the open (64 dB(A)), - - - in tunnel (71 dB(A)).

Various alternative quantities exist that can be used to define acceptable environments. These include the B-weighted level, preferred speech interference level (PSIL), loudness level, alternative noise criteria (NCA), noise ratings (NR), and room criteria (RC).<sup>51,56</sup>

The interior sound level varies considerably within a vehicle. [Figure 10.27](#) shows some examples of measured results where a loudspeaker has been placed at one end of an open saloon vehicle. This was a Mk 2 coach dating from the 1960s, although the interior dated from the 1990s. The solid line shows the relative sound pressure along a line down the central gangway at the height of the headrests. Results are shown in three example one-third octave bands. At low frequencies, strong modal patterns are observed due to the long acoustic wavelength. At higher frequencies considerable decay in the sound level is observed along the coach due to the absorptive properties of the seats, carpets, etc. Additional attenuation is seen at the middle of the coach where two glass partial screens were present either side of the door.

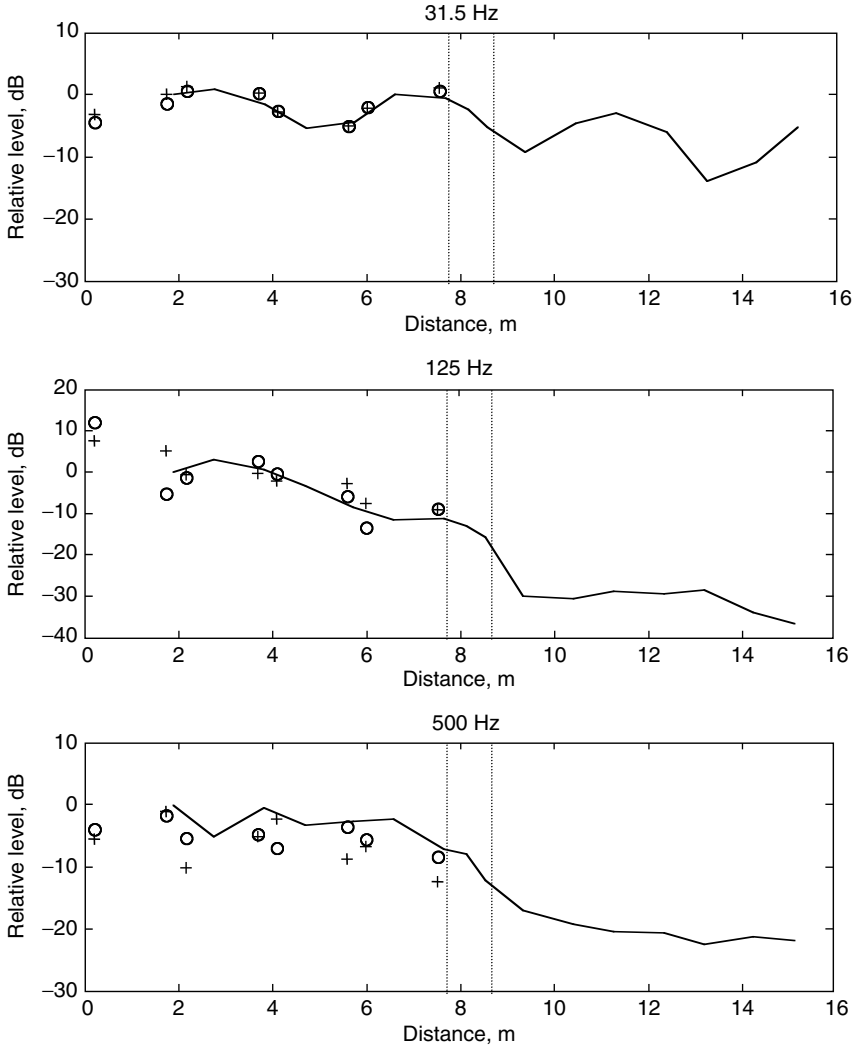
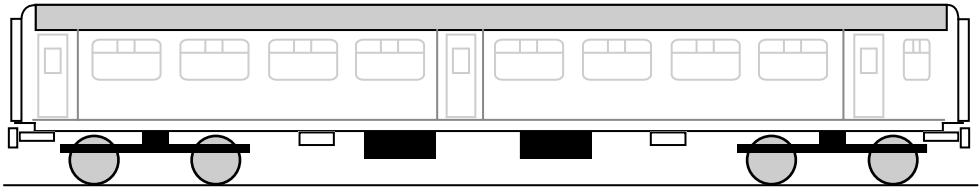
Also shown are measured results at positions in front of each seat headrest. The seats were arranged in groups of four with tables in between them. At low frequencies these measurements follow the same pattern as the gangway measurements, but at higher frequencies considerable differences can be seen between adjacent seated positions. These spatial variations may be experienced by passengers in the vehicles — the 500 Hz frequency band, for example, is quite important for speech interference. It can also be expected that differences will occur between left and right ear positions at an individual seat, leading to binaural effects. Clearly, in a running vehicle other source positions will apply, but these results serve to illustrate the general trends that can be expected.

### C. AIRBORNE TRANSMISSION

Airborne sound transmission into the vehicle occurs due to acoustic excitation of the vehicle floor, walls, windows, doors, and roof. The acoustic performance of a panel can be measured by placing it between two reverberant rooms and measuring the difference in sound pressure level between the two rooms.<sup>3</sup> The sound reduction index (or transmission loss) is the difference between the incident intensity level and the transmitted intensity level that can be derived from such a measurement after allowing for the size of the panel and the absorption in the receiver room.

A typical sound reduction index of a homogeneous panel is shown in [Figure 10.28](#). Generally, the sound reduction index of panels is dominated by the “mass law” behaviour in a wide frequency range. At high frequencies, the coincidence region occurs where the wavelengths in the structure and in air are similar. Here, a dip in the sound reduction index occurs, the extent of which depends on the damping. The mass law behaviour extends from the first resonance of the panel up to just below the critical (or coincidence) frequency. In this region the bending stiffness of the panel and its damping have no effect on the sound transmission (see [Ref. 3](#) for more details).

The use of light-weight constructions such as extruded aluminum or corrugated steel leads to a low sound reduction index. This follows from the mass law, which states that the sound reduction index reduces by 6 dB for a halving of panel mass. However, such structures tend to have a performance that is even worse than the mass law would suggest due to the presence of an extended frequency region over which coincidence effects occur. For example, [Figure 10.29](#) shows measurements of the sound reduction index from a 60-mm thick extruded aluminum floor of a railway vehicle with a 3-mm wall thickness taken from [Ref. 57](#) (similar results are also found in [Ref. 58](#)). Also shown is the “field incidence” mass law estimated for a homogeneous panel of similar mass.<sup>3</sup> Clearly, the extruded panel exhibits a much lower sound reduction index than this. It can be brought closer to the mass law behaviour by the use of a suspended inner floor and by adding a damping treatment to the extruded section.



**FIGURE 10.27** Relative internal sound levels in selected one-third octave bands in an ex-BR Mk 2 coach due to a sound source located at the left-hand end, 1.05 m above the floor. — measured at 1.05 m from floor along central gangway, o measured 0.05 m from head-rests of aisle seats (height 1.05 m from floor), + measured 0.05 m from headrests of window seats (height 1.05 m from floor). Positions of partial screens indicated by dotted lines.

### D. STRUCTURE-BORNE TRANSMISSION

As well as the airborne path, considerable sound power is transmitted to the vehicle interior through structural paths. This originates from the wheel–rail region as well as from underfloor diesel

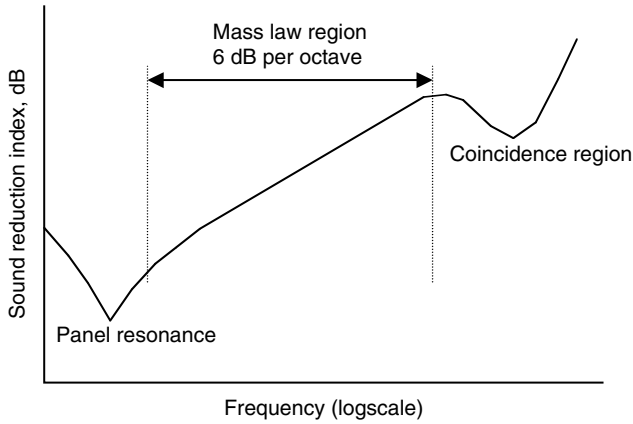


FIGURE 10.28 Typical sound reduction index of a homogeneous panel due to a diffuse incident field.

engines where these are present. Structure-borne engine noise could be reduced significantly in many cases by applying good mounting practice.<sup>50</sup> The mount stiffness must be chosen taking into account the frequency characteristics of the engine. An incorrect choice of stiffness can lead to amplification rather than attenuation of transmitted vibration. Flanking paths via pipes and hoses should also be avoided.

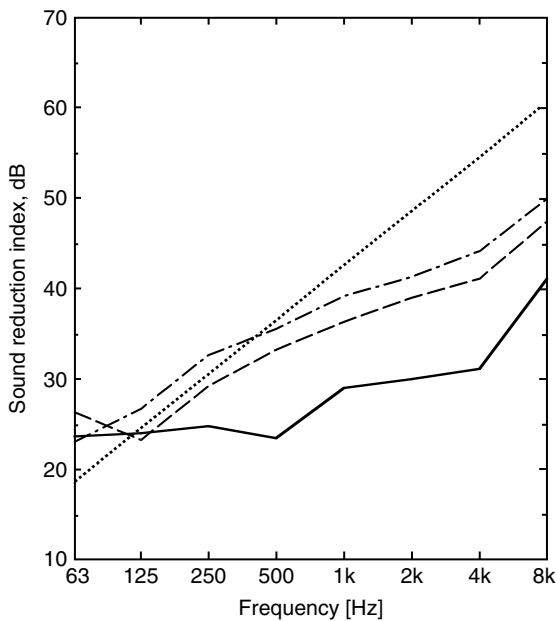


FIGURE 10.29 Octave band sound reduction index of extruded aluminum floor. — measured on bare floor panel, ... field incidence mass law for  $30 \text{ kg/m}^2$ , - - - measured for bare floor panel plus 12 mm suspended wooden deck, - . - measured for damped floor panel plus 12 mm suspended wooden deck (data taken from Ref. 57).

## E. PREDICTION OF INTERIOR NOISE

Deterministic methods, such as finite elements (FEM), may be applied at low frequencies to predict the vehicle interior noise.<sup>59</sup> Owing to the regular geometry an analytical model of the interior may also be used to construct the interior acoustic field on the basis of simple room modes.<sup>60</sup>

However, at high frequencies the number of modes becomes prohibitive for such approaches. The preferred analysis method for frequencies above approximately 250 Hz is therefore statistical energy analysis (SEA). This can be used in both predictive<sup>57,61</sup> and experimental modes.<sup>62</sup> However, in predictive mode it is not straightforward to define the coupling loss factors between the various subsystems, especially where use is made of aluminum extrusions<sup>58,63,64</sup> or other inhomogeneous constructions.<sup>65</sup> This is an area of continuing research. Moreover, as SEA is a statistical method, it provides an average result and cannot account easily for the spatial variations in sound field such as seen in [Figure 10.27](#).

## VIII. GROUND-BORNE VIBRATION AND NOISE

### A. OVERVIEW OF VIBRATION PHENOMENA

Just as environmental noise is receiving increased attention, so is the related environmental issue of ground vibration from rail traffic. Three distinct effects of ground vibration may be identified that arise predominantly from different types of railway.

Heavy axle-load freight traffic, travelling at relatively low speeds, causes vibration at the track of high amplitude that excites surface-propagating waves in the ground. This type of vibration often has significant components at very low frequency (below 10 Hz) and may interact with the frequencies of buildings rocking or bouncing on the stiffness of their foundations in the soil. This phenomenon is especially associated with soft soil conditions where it is found that significant levels of vibration may be propagated up to distances of the order of 100 m from the track. At these frequencies the vibration is perceived in the building as “whole body” vibration, which can be felt. This is usually assessed under the principles of ISO 2631.<sup>66</sup> High levels of vibration cause annoyance and, possibly, sleep disturbance. Complaints are often expressed in terms of concern over possible damage to property, although, for the levels of vibration normally encountered from trains, such concern is unlikely to be borne out when assessed against the criteria for building damage, e.g., BS 7385<sup>67</sup> or DIN 4150.<sup>68</sup>

Passenger trains also may cause significant levels of vibration, particularly electric multiple units with high unsprung masses. However, at sites of mixed traffic it is usually the case that a few freight trains, perhaps running at night, are identified as the worst cases and it is these which dominate the assessment of potential annoyance.

High-speed passenger trains sometimes travel at speeds in excess of the speed of propagation of vibration in the ground and embankments. This has been the concern of track engineers for some years because of the large displacements that can be caused in the track support structure and in electrification masts, etc. For its effect in causing ground vibration that propagates away from the track, this phenomenon may be compared with the bow wave from a ship or, more sensationally, the “boom” from a supersonic aircraft. Although the occurrences of this are comparatively rare, the topic has attracted considerable attention among researchers recently because of the expansion of the network of high-speed railways.<sup>69–71</sup> Hence, high-speed railways may also cause significant levels of surface vibration propagating to comparatively large distances from the track.

The third effect arises where trains run in tunnels and vibration is transmitted to buildings above. This has higher frequency content than vibration from surface tracks/trains. Although no direct airborne noise can be heard, vibration at the low end of the audible frequency range, from approximately 30 to 200 Hz, may excite bending in the floors and walls of a building which then radiate noise directly into the rooms. This rumbling noise may be found to be all the more annoying

because the source cannot be seen and no screening remedy is possible. This ground-borne noise is dealt with using soft baseplates, floating slab track, or other types of vibration-isolating track designs. While the design aims of new railway projects are, typically, to keep levels of structure-borne noise in properties to below maximum levels of 40 dB(A), it has been estimated, for instance, that around 56,000 households in London are subjected to this level and a very small number experience maximum levels of above 60 dB(A).<sup>72</sup>

## B. SURFACE VIBRATION PROPAGATION

In a layered ground, vibration propagates parallel to the surface via a number of wave types or modes. These are often called Rayleigh waves of different order (R-waves) and Love waves. The Rayleigh waves are also called P-SV waves since they involve coupled components of compressive deformation and vertically polarised shear deformation.<sup>b</sup> Here, the name P-SV wave is preferred and the term Rayleigh wave is reserved for the single such wave that exists in a homogeneous half-space. Love waves are decoupled from these and only involve horizontally polarised shear deformation and so are also known as SH waves. Since the vertical forces in the track dominate the excitation of vibration in the ground the SH waves are not strongly excited. They are not considered further in the present discussion.

To illustrate, examples of the wave mode shapes of the P-SV waves are shown in [Figure 10.30](#). These are the waves that propagate at 40 Hz in a ground modelled as a soft layer of weathered soil overlying a stiffer substratum of material. For the calculated results presented, the soft soil is a 2-m deep layer with a shear wave (S-wave) speed of 118 m/sec and dilatational wave (P-wave) speed 360 m/sec, and the substratum is a half-space with shear wave speed of 245 m/sec and dilatational wave speed of 1760 m/sec. Damping is included in both materials as a loss factor of 0.1.

If the wave number ( $= 2\pi/\lambda$  with  $\lambda$  the wavelength) is plotted as a function of frequency, the *dispersion curve* for the wave type is generated. [Figure 10.31](#) presents the dispersion diagram for the example soil structure (only the propagating P-SV modes are shown). Each line of the diagram represents a wave type associated with a cross-sectional mode of the layered soil. The inverse slope of a line from the origin to a point on a curve is equal to the phase velocity of that wave type at a particular frequency. The inverse slope of the curve itself gives the group velocity of the wave type. This is the speed at which energy is transported.

For this example set of soil parameters, at very low frequency, only a single mode exists and this has a wave speed close to that of the shear waves in the substratum. Around 15 Hz, the depth of the weathered material sustains a quarter wavelength of the shear wave. Above this frequency there is a “cut-on” of a wave that involves mostly deformation of the weathered layer material. With the onset of this mode, i.e., propagation via the layer material, a rise in the transmitted level of vibration is observed. It can be seen in [Figure 10.30](#) that the mode with the lower phase speed (left-hand picture) involves mainly deformation of the softer layer material whereas the second mode involves a stronger component of deformation in the half-space.

[Figure 10.31](#) shows that, as the frequency increases, higher order propagating wave types “cut-on” at frequencies of 23, 47, and 85 Hz. At high frequency, as the wavelengths of shear and compression become small compared with the depth of the weathered material layer, the wave number of the slowest wave converges towards that of the Rayleigh wave in a half-space of the layer material.

[Figure 10.31](#) also presents the dispersion curve of the wave propagating along a ballasted track structure unconstrained at its lower surface (parameters as in [Table 10.1](#)). This curve is superimposed on the ground wave curves. At the intersection of the track wave curve and the first

<sup>b</sup> The compressive or dilatational wave is often referred to as the P-wave (P stands for primary) and the shear wave as the S-wave (S stands for secondary). In a seismic survey the faster P-wave arrives at a detector first and the S-wave second.

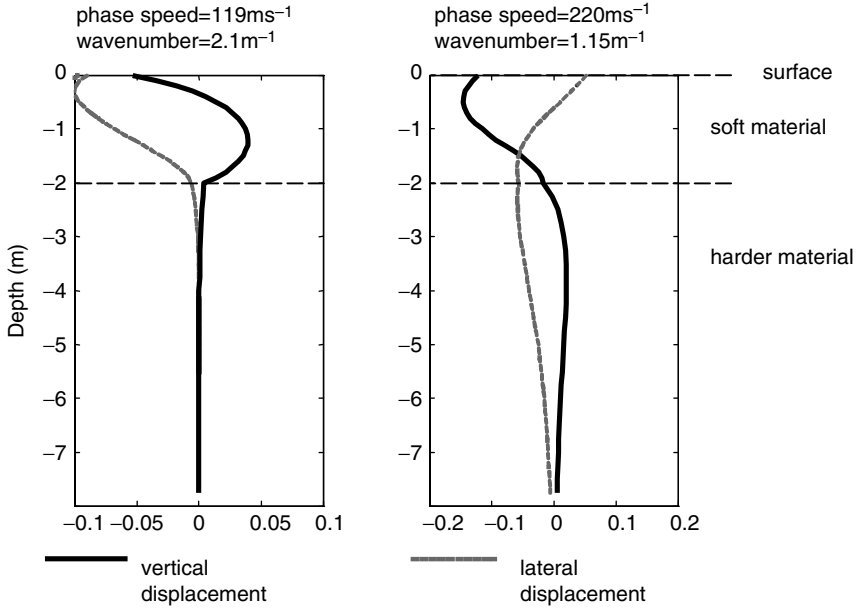


FIGURE 10.30 P-SV modes of the example layered ground structure at 40 Hz.

ground wave curve, around 35 Hz, waves may freely propagate at the same speed in both the track and the ground. The inverse of the phase speed at this intersection is represented by the slope of the straight line through this point and the origin indicated on the graph. If a moving but nonscillating load is applied to the track at this speed, it will excite a maximum level of ground vibration. Thus the track-ground system may exhibit a critical train speed that causes high levels of vibration. If a nonscillating load travels at a lower speed it will not directly excite propagating ground waves. The plot also shows that the response to a dynamic load not moving along the track would have a maximum at a frequency of 35 Hz. A more detailed discussion of the critical speed, and the influence of the track parameters on this, is given in Ref. 73.

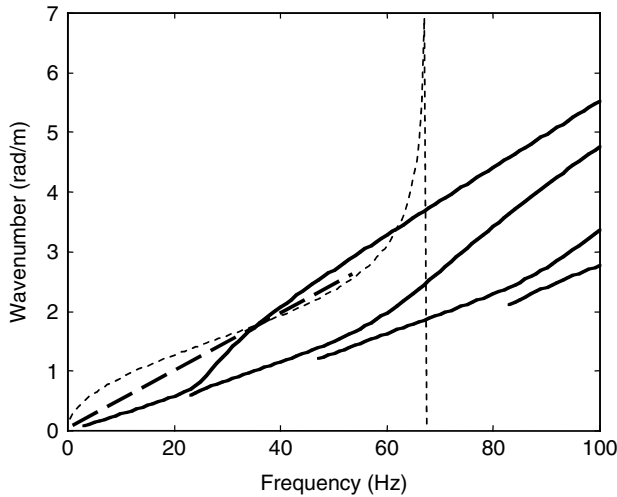


FIGURE 10.31 The dispersion diagram for propagating P-SV waves of the example ground (—), the dispersion diagram for a track (---), and a line representing the critical train speed (- -).



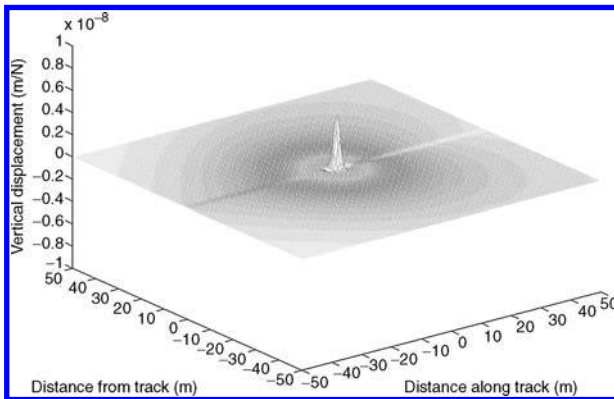
**TABLE 10.1**  
**Parameters Used for the Ballasted Track in Figure 10.31**

Mass of rail beam per unit length of track	120 kg/m
Bending stiffness of rail beam	$1.26 \times 10^7 \text{ Nm}^2$
Loss factor of the rail	0.01
Rail pad stiffness	$3.5 \times 10^8 \text{ N/m}^2$
Rail pad loss factor	0.15
Mass of sleepers per unit length of track	490 kg/m
Mass of ballast per unit length of track	3300 kg/m
Ballast stiffness per unit length of track	$1.775 \times 10^8 \text{ N/m}^2$
Loss factor of ballast	1.0
Effective contact width of railway and ground	2.7 m

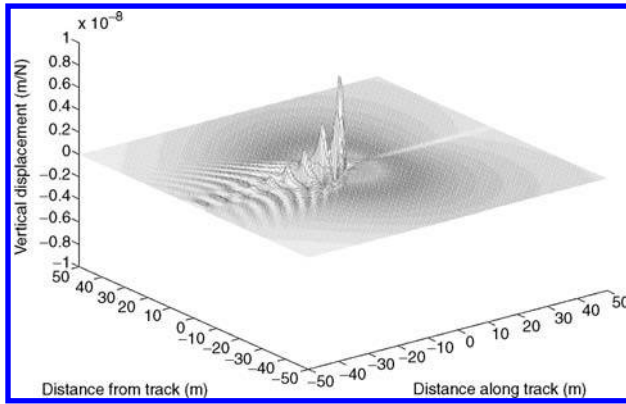
The wave field can be derived from the calculation of the excitation of waves in the ground due to excitation on the track.<sup>74,75</sup> The wave field for a single load moving along the track with a speed below that of any of the waves in the ground is shown in Figure 10.32. The displacement “dip” under the single load is indicated by the positive (upward) displacement under the track. Little effect is observed only a few metres away, although close to the track the passage of the quasistatic displacement pattern may be observed. Figure 10.33 shows what happens when the load travels at a speed near to the critical speed. Propagating waves may be seen travelling with significant amplitude away from the track. These exhibit the form of a “bow wave” because the load speed is greater than the speed of waves in the ground.

If vehicle models representing each of the vehicles of a train are coupled to the model for the track ground system, a theoretical model that predicts the complete vibration field can be produced.<sup>76</sup> A model in which the vibration excited by the moving axle loads of the whole train and that excited by the irregular vertical profile of the track for all the axles of a train has been validated by comparison with measured vibration for a number of sites.<sup>77</sup> Two illustrative results are presented here for vibration from an X2000 train running over very soft soil at a site called Ledsgård near Gothenburg in Sweden.

Figure 10.34 shows the measured vibration spectrum and that predicted for the whole X2000 train at 70 km/h (19.4 m/sec) for a point 7.5 m from the centre-line of the track. The measurements



**FIGURE 10.32** Displacement pattern in the moving frame of reference for a single nonoscillating axle load on the track moving at 83 m/sec, below the wavespeeds in the ground.

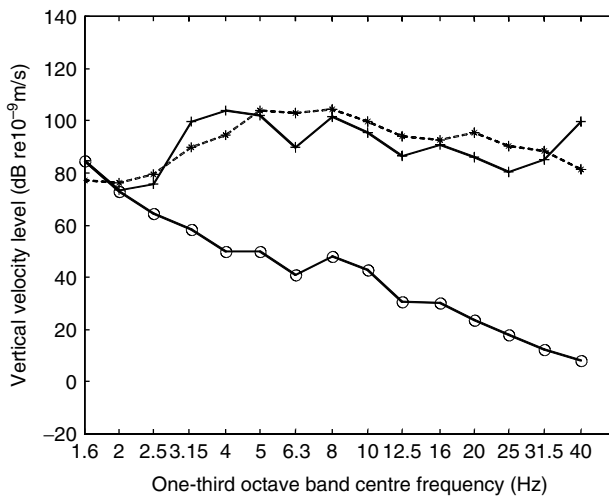


**FIGURE 10.33** Displacement pattern in the moving frame of reference for a single nonoscillating axle load on the track moving at 150 m/sec, close to the critical speed for this track-ground system.

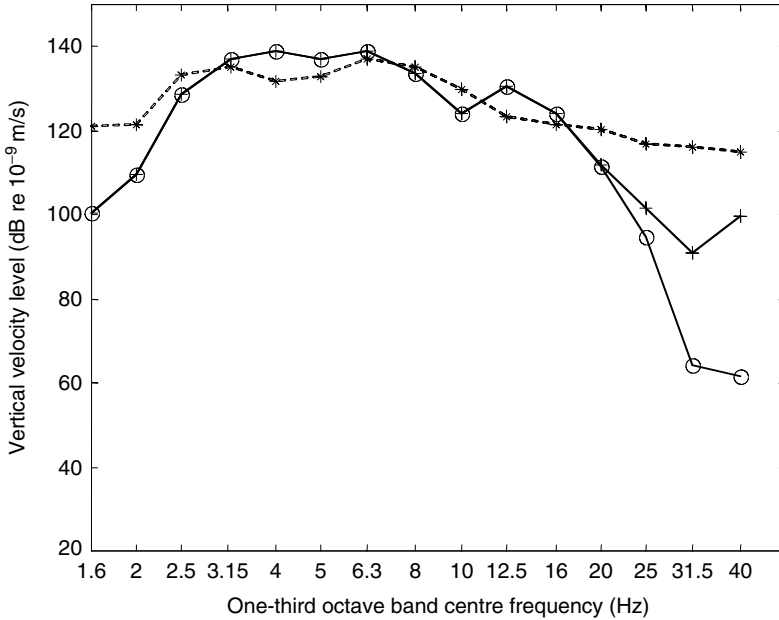
cover the frequency range from 1.6 to 40 Hz in one-third octave bands. Given that no specific track profile data were available and that therefore typical data were used, the figure shows that the predicted dynamically induced vibration level accounts well for the vibration measured over most of the frequency range. The quasistatically induced vibration is only important below 2.5 Hz.

Figure 10.35 compares predicted and measured vibration when the X2000 runs at 200 km/h (56 m/sec). In this case, the train speed is close to the critical value for this site at which wave numbers coincide approximately as illustrated, for a different ground, in Figure 10.31.

However, the coincidence of the ground wave and speed line occurs here over an extended frequency range between 3 and 10 Hz because the ground-wave dispersion curve has an almost constant slope in this range. The results show that, now, the observed level of vibration is about 35 dB higher and that it is due to the direct excitation of the propagating wave by the moving quasistatic axle loads. In summary, therefore, the effect of a ground vibration “boom” is



**FIGURE 10.34** Vertical velocity level at Ledsgård for train speed of 70 km/h (19.4 m/sec): o predicted level due to quasistatic loads; + predicted total level; \* measured level at 7.5 m. Source: From Sheng, X. et al., *J. Sound Vib.*, 267, 621–635, 2003, Elsevier. With permission.



**FIGURE 10.35** Vertical velocity level at Ledsgård for train speed of 200 km/h (56 m/s): o, predicted level due to quasistatic loads; + predicted total level; \* measured level at 7.5 m. *Source:* From Sheng, X. et al., *J. Sound Vib.*, 267, 621–635, 2003, Elsevier. With permission.

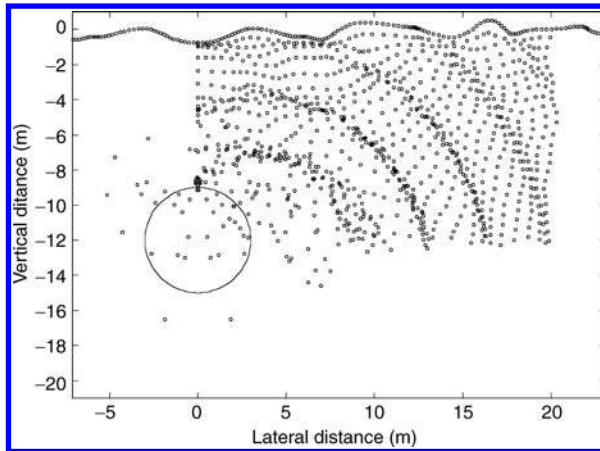
encountered at around 200 km/h at this site, whereas at lower speed, the observed vibration level is due to the dynamically induced component of vibration.

Ledsgård is unusual in having such a low ground-wave speed and the “boom” phenomenon is therefore not common. Nevertheless, high-speed lines are now being designed for speeds in excess of 300 km/h and the critical speed must be taken into account where these pass over soft ground. For conventional trains, and for the majority of sites where vibration problems occur, however, it may be said that the most important mechanism of vibration excitation is the irregular vertical profile of the track, possibly combined with out-of-round wheels. There also remains the possibility of the succession of load-displacement dips under each axle causing vibration at the axle passing frequencies for very low frequency vibration of buildings within a few metres of the track but this does not propagate far.

### C. TUNNEL VIBRATION

Attention is now directed towards vibration propagation from tunnels. For this discussion, the results of a coupled finite element/boundary element model are presented. These are for a typical 3.5-m outer radius, circular bore tunnel (20 m deep at the rail) with and without a concrete lining.<sup>78</sup> Ground properties typical of a deep clay formation have been used, namely, an S-wave speed of 610 m/sec and a P-wave speed of 1500 m/sec (implying a Poisson’s ratio of 0.4); the density of the material has been assumed to be 1700 kg/m<sup>3</sup> and the damping loss factor is 0.15. Boundary elements are used to represent the ground-tunnel interface and the ground surface from +50 to –20 m relative to the vertical centreline of the tunnel. The boundary elements model a ground of infinite extent.<sup>78</sup> The tunnel and invert structure are modelled using finite elements.

Figure 10.36 shows the exaggerated instantaneous particle displacement at a number of points in the ground to illustrate the wave pattern radiating away from an oscillating load at the base of the



**FIGURE 10.36** The vibration field around the unlined tunnel at 100 Hz. *Source:* From Jones, C. J. C., Thompson, D. J. and Petyt, M., *Transport J.*, 153(2), 121–129, 2002.

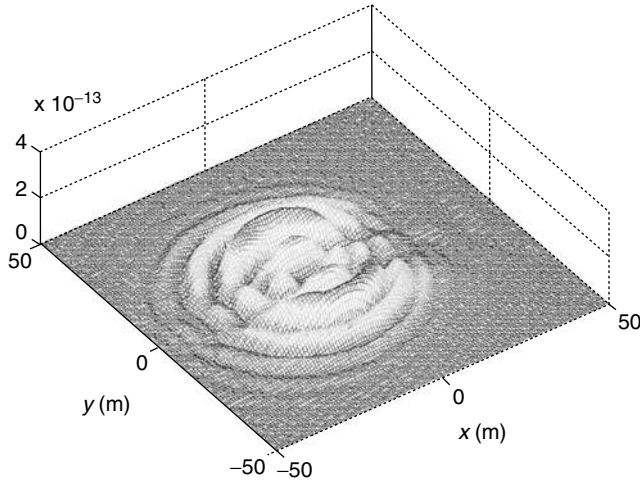
tunnel at high frequency. It shows that a relatively simple pattern of cylindrical wave fronts radiate towards the surface at greater distances from the tunnel. The strongest component of deformation in these waves is shear. At this frequency (100 Hz), the wavelengths of vibration are shorter than the diameter of the tunnel and therefore do not diffract around it. A “shadow zone” therefore exists in the region immediately above the tunnel. For this reason, the greatest amplitudes of response on the ground surface are at a distance of about 15 to 20 m from the tunnel alignment rather than directly above it.

Figure 10.37 shows the displacement of an unlined tunnel and one with a concrete lining at 100 Hz. Waves can be seen propagating from the invert slab with a high rate of decay around the ring of the unlined tunnel. These have the form of Rayleigh surface waves. Compared with the unlined tunnel, the amplitude of the response at the crown of the lined tunnel is much greater. The result is that, in the shadow zone, the tunnel structure design has a strong influence on the level of vibration at the surface.

It is clear from Figure 10.37 that the structure of the tunnel has an influence on the excitation of waves in the ground. An analysis of the waves that propagate along the tunnel structure is also possible. Figure 10.38 is similar to Figure 10.32 and Figure 10.33 in that it shows the calculated response on the surface of the ground to a moving oscillating load but, in this case, the load



**FIGURE 10.37** Amplified representation of the response showing waves round the unlined (left) and lined (right) tunnel rings at 100 Hz. *Source:* From Jones, C. J. C., Thompson, D. J. and Petyt, M., *Transport J.*, 153(2), 121–129, 2002.



**FIGURE 10.38** The vertical response amplitude of the surface of the ground to a load oscillating at 200 Hz and moving at 100 m/sec in the lined tunnel. The tunnel lies in the line  $y = 0$ , the invert is at a depth of 20 m.

is in a tunnel. The response can be seen to be asymmetrical because of the load speed. However, since the ground at tunnel depth is much stiffer than the soft surface conditions of [Figure 10.30](#) to [Figure 10.35](#), the effect of the moving load is very small even though a very high speed of 100 m/sec has been used.

As in [Figure 10.36](#), the highest levels of vibration in [Figure 10.38](#) can be seen to be about 15 m to the side of the centreline of the tunnel, with the propagation pattern beyond showing circular wave fronts with monotonic decay, while the vibration field above the tunnel is more complicated.

#### D. VIBRATION ISOLATING TRACKS

The main way in which the vibration from underground railways is controlled is by the use of soft or resilient elements in the vertical support of the track in order to provide some degree of *vibration isolation*. The principle of vibration isolation is illustrated in [Figure 10.39](#), using a simple single degree of freedom oscillator. The ratio of the amplitude of the force transmitted to the foundation to that of the oscillatory force applied to the mass is called the *transmissibility*. At very low frequency this ratio is unity; the whole force is transmitted as it would be in the static case. At the natural frequency of the system  $f_n$ , the force is increased. Above  $\sqrt{2}$  times the natural frequency, the transmissibility reduces to below unity and continues to decrease with increasing frequency. The effect of the damping in the support is also shown in [Figure 10.39](#). Here, a hysteretic damping model has been used (constant damping loss factor  $\eta$ ) that reflects the behaviour of elastomeric materials. The amplitude of the resonance is dependent on the damping in the support but the degree of vibration isolation at higher frequencies is not.

Vibration isolating tracks are commonplace in modern underground railway systems to reduce ground-borne noise and the subject is an important part of track design. They work on a principle similar to that shown in [Figure 10.39](#). The lower the stiffness of the support, the lower the natural frequency of the system will be and the greater the degree of vibration isolation at higher frequencies. The choice of support stiffness is, however, limited by the allowable vertical and lateral static displacements under the axle load of the train.

[Figure 10.40](#) shows some of the basic design concepts for vibration isolating track designs. The rail pad is not shown; it has a stiffness higher than that of the resilient element in each case but possibly still significant in the behaviour of the track design for the relevant frequency range.

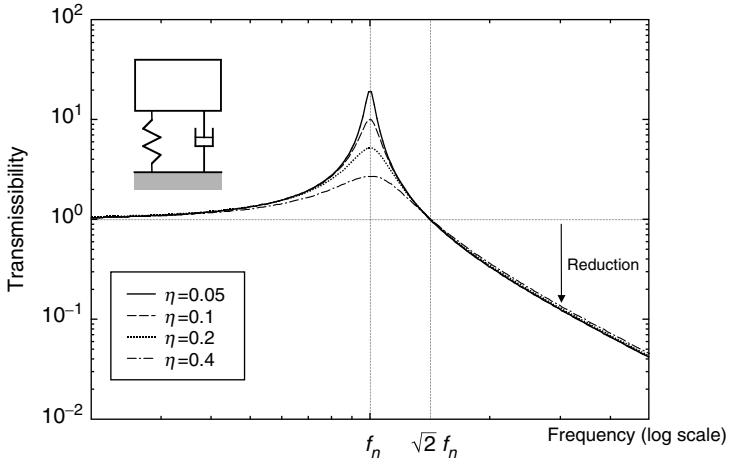


FIGURE 10.39 The force transmissibility of a hysteretically damped single degree of freedom system.

The ballast layer forms the resilient component of a conventional ballasted track. For this reason, slab tracks with normal pad stiffness give rise to increased vibration transmission compared with ballasted track. Soft baseplates are used to rectify this. For these, the lateral rail displacements are generally the limiting factor. Baseplate designs are therefore wide, or support the rail under the head, to avoid rail rotation and consequent gauge widening. Alternatively, gauge widening may be avoided by using resiliently mounted sleepers or floating slab track. These also increase the mass above the resilient element to decrease the natural frequency further.

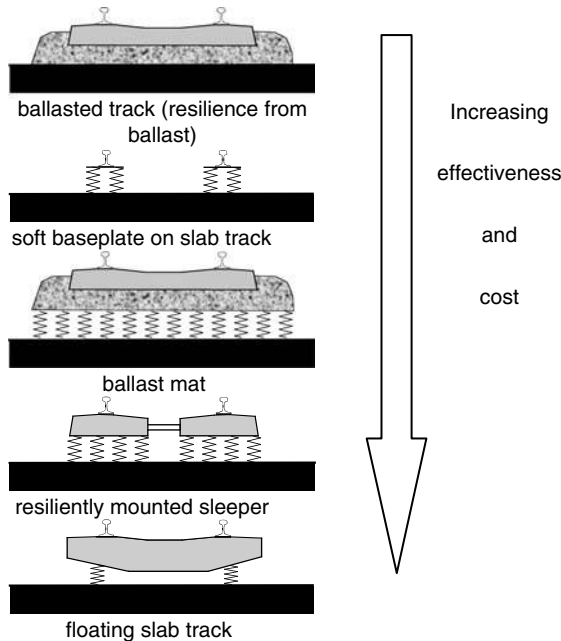
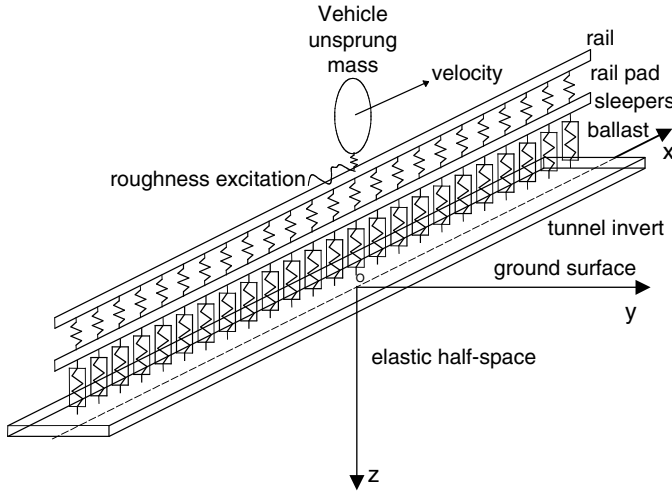


FIGURE 10.40 Design concepts for vibration isolating tracks.



**FIGURE 10.41** A model used to calculate the relative vibration isolation performance of different track designs.

Clearly, it is necessary to assess the vibration isolation performance of different track designs to obtain a required reduction in the vibration spectrum. Figure 10.41 presents a calculation model that has been extensively used for this purpose.<sup>79</sup> The figure shows the case for a ballasted track but different track models can be used. The model is used to predict the change in vibration response at a point of the surface of a half-space some distance from the track for a change in the parameters of the track or of the unsprung mass of the vehicle. The track is adequately modelled, for this frequency range, as an infinite layered beam structure with vertical stiffness for the ballast, pads, or baseplates distributed continuously along its length. The track-top irregularity excites a vertical dynamic force at the rail head. The model is solved in the frequency domain and hysteretic (loss factor) damping is included in all components. The elastic half-space model of the ground represents the frequency-dependent support stiffness under the track and provides a suitable means to sum the contributions of vibration from the waves propagating along the track. In this way, some geometrical and damping effects are taken into account in the propagation of vibration through the soil but the half-space does not represent a tunnel situation. The assumption is made that a vibration *reduction* due to the change in track design would be the same for a tunnel as for a half-space of realistic soil parameters. This is valid as long as the dynamic support stiffness of the ground is higher than that of the resilient element of the track.

With models similar to that shown in Figure 10.41 the performance of different track types can be evaluated for the specific vehicles and the other track parameters so that the track design can be chosen to obtain the required isolation performance at the minimum cost. The risk of the track-form leading to corrugation of the rail or high airborne noise levels in the tunnel must also be taken into account in the choice.

**E. SUMMARY**

Both low frequency vibration from trains running on tracks at grade and higher frequency vibration leading to ground-borne noise in buildings are major concerns for railways. It has been shown that low-frequency surface vibration from railways may be excited either by the movement of steady axle loads or dynamically as the axles run over the irregular profile of the track. While the latter is the more dominant mechanism for conventional train speeds and frequencies above a few hertz,

the movement of the axles may be the dominant mechanism for very low frequencies at some sites close to the track. Very high levels of vibration may result from high-speed trains which can travel faster than the wave speeds in the ground at sites with very soft soil.

Ground-borne noise from trains in tunnels cannot be predicted using simple decay with distance laws, especially for locations directly over the tunnel itself. In this case the vibration response is dependent on the tunnel structure. To reduce ground-borne noise, various vibration isolating track forms are used. These should be chosen with respect to the vehicles and the vibration reduction required at particular sites.

## IX. VIBRATION COMFORT ON TRAINS

### A. INTRODUCTION

The level of vibration in vehicles is a major influence on the perception of the quality of rail travel in comparison with other forms of transport. Vibration in the frequency range from approximately 0.5 to 80 Hz causes discomfort as “whole body” vibration and frequencies below this may cause nausea. The wavelengths in the vertical and lateral profiles of the track that give rise to this vibration are between approximately 1 and 70 m depending on the train speed. Of course, the comfort of passengers is a primary reason for the routine monitoring and maintenance that is central to track management for all railways.

### B. ASSESSMENT OF VIBRATION COMFORT IN TRAINS

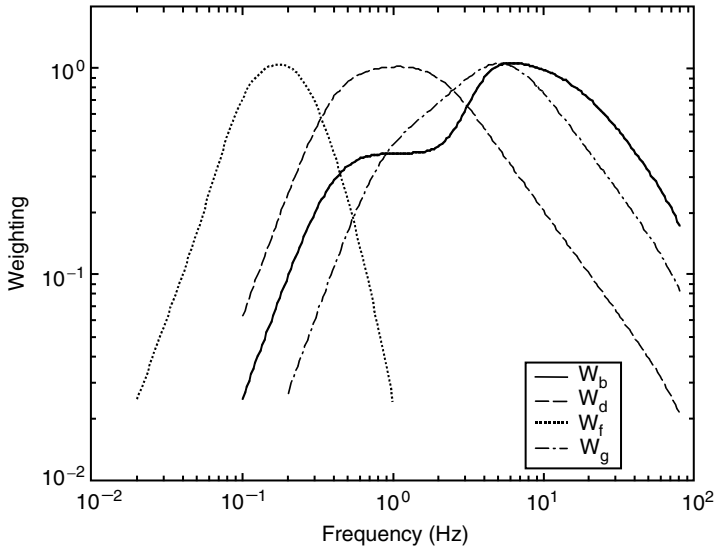
It is important to understand how measured vibration levels in vehicles are used to assess the likely reaction of passengers. A comprehensive background on this subject is given in Ref. 80; here, only an indicative overview is given.

The most commonly accepted principles of vibration perception assessment are laid out in the international standard ISO 2631-1 (1997), “Guide to the evaluation of human exposure to whole-body vibration,”<sup>66</sup> and also in BS 6841 (1987), “Measurement and evaluation of human exposure to whole-body mechanical vibration and repeated shock.”<sup>81</sup> These set out terms for consideration of health, comfort, incidence of motion sickness, and effects on human activities. Frequency weightings or “filters” are defined that reflect human sensitivity to vibration in a similar way to the A-weighting (Figure 10.1) is used for sound. Some of these are shown in Figure 10.42. In the assessment of ride comfort, the filter  $W_b$  is used in BS 6841 to weight rms vibration in the vertical (spinal) direction for both seated and standing passengers and filter  $W_d$  for the two components of lateral vibration. (There is a difference between ISO 2631 and BS 6841 in that the ISO standard uses a slightly different weighting for vertical vibration,  $W_k$ . However,  $W_b$  is used more in the railway industry as is recognized in a later *draft* standard ISO 2631-4, specifically for the railway industry.) Vibration in the frequency range 0.5 to 80 Hz is considered. It is measured, as appropriate, on the seat surface between the cushion and a subject, or on the carriage floor. Since measurements on the seat are dependent on the seated person, measurements should be carried out for a sample of subjects. Vibration on the seat back can also be important and is evaluated using other frequency weightings.

When considering the effects of vibration on human activity, weighting  $W_g$  is used for the vertical direction rather than  $W_b$ . For assessing the likelihood of vibration to cause motion sickness, weighting  $W_f$  is used for vertical vibration and the lower frequency range of 0.1 to 0.5 Hz is considered. No guidance is given in the standard on the influence of other components of vibration on motion sickness.

Meters and vibration analysis equipment are available that implement the frequency weighting filters and thereby evaluate the overall weighted levels of vibration. To combine the effects of





**FIGURE 10.42** Some of the frequency weightings for whole body vibration defined in BS 6841.

vibration entering the body at the seat, seat back, and the floor in different directions, the root sum of squares of these overall levels can be used.

It is for the rolling stock purchaser to set acceptable limits for vibration measured in this way according to the type of rolling stock taking into account factors such as the duration of journeys, number of standing passengers, line geometry standard, vehicle speed, etc. In practice, the standards that are set vary from one railway to another.

There are a small number of single-value indicators of ride quality. One is defined by BS 6841 which allows the measurement of weighted accelerations in 12 components on the seat, seat back, and floor. These overall levels are then multiplied by “axis multiplying factors” to give “component ride values” and these may be combined to give an overall ride index.

Another ride quality indicator for “average comfort”,  $N_{MV}$ , is defined by prENV 12229 (1996).<sup>82</sup> This uses overall accelerations in the vertical and the two horizontal component directions, weighted as in BS 6841 and ISO 2631, but the 95th percentile values of 60 separate 5-sec measurements are taken. The measure therefore becomes sensitive to rare events of high acceleration.  $N_{MV}$  is evaluated as six times the root sum of squares of these values. Values of  $N_{MV}$  are then rated in five bands from “very comfortable” ( $N_{MV} < 1$ ) to “very uncomfortable” ( $N_{MV} > 5$ ). Although it is suggested that all European railways should adopt this measure of “average comfort,” its complexity is a barrier to its acceptance in practice.

**C. EFFECTS OF VEHICLE DESIGN**

Passenger vehicle suspensions are designed to isolate the coach body for frequencies above approximately 2 Hz. The bogie ride dynamics and the design of the suspension are of primary importance through much of the lower frequency range to which humans are sensitive, especially for motion sickness. In the latter case, the human effects of tilt and cant deficiency are an important ongoing area for research.

The vibration level inside the coach body is also affected by its low order structural resonances. Typically, both vertical and lateral first-order modes of bending along the coach arise at frequencies around 10 Hz and resonances of the floor occur at frequencies just above this. The excited amplitude of these should be kept to a minimum by structural design of the coach avoiding the

coincidence of important modes and using damping treatments. The coach body modes should be kept above the frequency range to which humans are most sensitive (approximately 2 to 10 Hz). This is one reason for the trend towards light, stiff materials such as aluminum extrusions in the manufacture of rolling stock.

An additional trend is the introduction of vibration-isolated “walking” floors in passenger coaches. This is primarily aimed at reducing vibration in the audible frequency range that is important for the interior noise environment (Section VII) but can be effective in reducing vibration above approximately 20 Hz.<sup>83</sup>

The coach body vibration, both vertical and in the two lateral directions, is felt by passengers through the seat and the seat back. Seat dynamics must also therefore be considered. The coupled system of seat and human body exhibits a resonance of vertical vibration typically between 4 and 6 Hz and at a similar frequency in fore-aft vibration due to the stiffness of the backrest. With the very soft seats used on some old rolling stock, these resonances can cause the vibration at the floor level to be made worse for the passenger, rather than better, by the seat. For this reason new rolling stock often has much firmer seats than the stock it replaces.

## REFERENCES

1. Future Noise Policy, European Commission Green Paper, COM(96) 540 final, Brussels, 4 November, CB-CO-96-548-EN-C, 1996.
2. Commission Decision 2002/735/EC concerning the Technical Specification for Interoperability (TSI) relating to the rolling stock subsystem of the trans-European high-speed rail system. *Off. J. Eur. Communities* 12.9.2002 L245/402–506.
3. Fahy, F. and Walker, J., Eds., *Fundamentals of Noise and Vibration*, E&FN Spon, London, 1998.
4. Kinsler, L. E., Frey, A. R., Coppens, A., and Sanders, J. V., *Fundamentals of Acoustics*, Wiley, New York, 1982.
5. ISO 532-1975 Acoustics — method for calculating loudness level, International Organization for Standardization.
6. Dings, P. C. and Dittrich, M. G., Roughness on Dutch wheels and rails, *J. Sound Vib.*, 193, 103–112, 1996.
7. Remington, P. J., Wheel/rail noise, part IV: rolling noise, *J. Sound Vib.*, 46, 419–436, 1976.
8. Remington, P. J. and Webb, J., Estimation of wheel/rail interaction forces in the contact area due to roughness, *J. Sound Vib.*, 193, 83–102, 1996.
9. Thompson, D. J., The influence of the contact zone on the excitation of wheel/rail noise, *J. Sound Vib.*, 267, 523–535, 2003.
10. Thompson, D. J. and Jones, C. J. C., A review of the modelling of wheel/rail noise generation, *J. Sound Vib.*, 231, 519–536, 2000.
11. Wu, T. X. and Thompson, D. J., Theoretical investigation of wheel/rail non-linear interaction due to roughness excitation, *Veh. Syst. Dyn.*, 34, 261–282, 2000.
12. Thompson, D. J., Hemsworth, B., and Vincent, N., Experimental validation of the TWINS prediction program for rolling noise, part 1: description of the model and method, *J. Sound Vib.*, 193, 123–135, 1996.
13. Thompson, D. J., Fodiman, P., and Mahé, H., Experimental validation of the TWINS prediction program for rolling noise, part 2: results, *J. Sound Vib.*, 193, 137–147, 1996.
14. Jones, C. J. C. and Thompson, D. J., Extended validation of a theoretical model for railway rolling noise using novel wheel and track designs, *J. Sound Vib.*, 267, 509–522, 2003.
15. Hübner, P., The action programme of UIC, CER, and UIP ‘abatement of railway noise emissions on goods trains’, *J. Sound Vib.*, 231, 511–517, 2000.
16. Hölzl, G., A quiet railway by noise optimised wheels (in German), *ZEV + DET Glas. Ann.*, 188, 20–23, 1994.
17. Färm, J., Evaluation of wheel dampers on an intercity train, *J. Sound Vib.*, 267, 739–747, 2003.

18. Jones, C. J. C. and Thompson, D. J., Rolling noise generated by wheels with visco-elastic layers, *J. Sound Vib.*, 231, 779–790, 2000.
19. Cervello, S., Donzella, G., Pola, A., and Scepi, M., Analysis and design of a low-noise railway wheel, *Proc. Inst. Mech. Eng., J. Rail Rapid Transit*, 215F, 179–192, 2001.
20. Thompson, D. J. and Jones, C. J. C., A study of the use of vehicles with small wheels for determining the component of noise from the track, Proceedings of IOA Spring Conference, Salford, England, 2002.
21. Thompson, D. J., Wheel/rail noise generation, part II: wheel vibration, *J. Sound Vib.*, 161, 401–419, 1993.
22. Jones, C. J. C. and Edwards, J. W., Developing and testing of wheels and track components for reduced rolling noise from freight trains, Proceedings of Internoise '96, pp. 403–408, Liverpool, 1996.
23. Jones, C. J. C., Thompson, D. J., Frid, A., and Wallentin, M. O., *Design of a Railway Wheel with Acoustically Improved Cross-Section and Constrained Layer Damping*, Internoise 2000, Nice, France, August, pp. 673–678, 2000.
24. Vincent, N., Bouvet, P., Thompson, D. J., and Gautier, P. E., Theoretical optimization of track components to reduce rolling noise, *J. Sound Vib.*, 193, 161–171, 1996.
25. Thompson, D. J., Jones, C. J. C., and Farrington, D., *The Development of a Rail Damping Device for Reducing Noise from Railway Track*, Internoise 2000, Nice, France, August, pp. 685–690, 2000.
26. Jones, R. R. K., Bogie shrouds and low barriers could significantly reduce wheel/rail noise, *Rail. Gaz. Int.*, July, 459–462, 1994.
27. Jones, R., Beier, M., Diehl, R. J., Jones, C., Maderboeck, M., Middleton, C., and Verheij, J., *Vehicle-Mounted Shields and Low Trackside Barriers for Railway Noise Control in a European Context*, Internoise 2000, Nice, France, August, 2000.
28. Oertli, J., The STAIRRS project, work package 1: a cost-effectiveness analysis of railway noise reduction on a European scale, *J. Sound Vib.*, 267, 431–437, 2003.
29. Vér, I. L., Ventres, C. S., and Myles, M. M., Wheel/rail noise, part II: impact noise generation by wheel and rail discontinuities, *J. Sound Vib.*, 46, 395–417, 1976.
30. Clark, R. A., Dean, P. A., Elkins, J. A., and Newton, S. G., An investigation into the dynamic effects of railway vehicles running on corrugated rails, *J. Mech. Eng. Sci.*, 24, 65–76, 1982.
31. Nielsen, J. C. O. and Igeland, A., Vertical dynamic interaction between train and track — influence of wheel and track imperfections, *J. Sound Vib.*, 187, 825–839, 1995.
32. Wu, T. X. and Thompson, D. J., A hybrid model for the noise generation due to railway wheel flats, *J. Sound Vib.*, 251, 115–139, 2002.
33. Johansson, A. and Nielsen, J., *Railway Wheel Out-of-Roundness — Influence on Wheel–Rail Contact Forces and Track Response*, Proceedings of Wheelset Congress, Rome, 2001.
34. Wu, T. X. and Thompson, D. J., On the impact noise generation due to a wheel passing over rail joints, *J. Sound Vib.*, 267, 485–496, 2003.
35. Wu, T. X. and Thompson, D. J., *A Model for Impact Forces and Noise Generation Due to Wheel and Rail Discontinuities*, Eighth International Congress Sound and Vibration, Hong Kong, China, pp. 2905–2912, 2001.
36. Remington, P. J., Wheel/rail squeal and impact noise: what do we know? What don't we know? Where do we go from here?, *J. Sound Vib.*, 116, 339–353, 1987.
37. Rudd, M. J., Wheel/rail noise — part II: wheel squeal, *J. Sound Vib.*, 46, 381–394, 1976.
38. Fingberg, U., A model of wheel–rail squealing noise, *J. Sound Vib.*, 143, 365–377, 1990.
39. Périard, F., Wheel–rail noise generation: curve squealing by trams. Ph.D. thesis, Technische Universiteit Delft, 1998.
40. Heckl, M. A., Curve squeal of train wheels, part 2: which wheel modes are prone to squeal?, *J. Sound Vib.*, 229, 695–707, 2000.
41. De Beer, F. G., Janssens, M. H. A., and Kooijman, P. P., Squeal of rail bound vehicles influenced by lateral contact position, *J. Sound Vib.*, 267, 497–507, 2003.
42. Monk-Steel, A. and Thompson, D., *Models for Railway Curve Squeal Noise*, Seventh International Conference of Recent Advances in Structural Dynamics, Southampton, England, July, 2003.

43. Eadie, D. T., Santoro, M., and Powell, W., Local control of noise and vibration with Keltrack friction modifier and Protector trackside application: an integrated solution, *J. Sound Vib.*, 267, 761–772, 2003.
44. Wetta, P. and Demilly, F., Reduction of wheel squeal noise generated on curves or during braking, *11th International of Wheelset Congress*, Paris, June, pp. 301–306, 1995.
45. Kirschner, F., Koch, J. E., and Cohen, H. C., *Light Weight Vibration Damping Treatments for Railroad Wheels*, *11th International of Congress Acoustics*, Paris, 5, 129–132, 1983.
46. Mauclaira, B., *Noise Generated by High Speed Trains. New Information Acquired by SNCF in the Field of Acoustics Owing to the High Speed Test Programme*, *Internoise 1990*, Gothenburg, pp. 371–374, 1990.
47. King, W. F. III, A precis of developments in the aeroacoustics of fast trains, *J. Sound Vib.*, 193, 349–358, 1996.
48. Barsikow, B., Experiences with various configurations of microphone arrays used to locate sound sources on railway trains operated by DB-AG, *J. Sound Vib.*, 193, 283–293, 1996.
49. Talotte, C., Aerodynamic noise, a critical survey, *J. Sound Vib.*, 231, 549–562, 2000.
50. Hardy, A. E. J. and Jones, R. R. K., Control of the noise environment for passengers in railway vehicles, *Proc. Inst. Mech. Eng.*, 203F, 79–85, 1989.
51. Eade, P. W. and Hardy, A. E. J., Railway vehicle internal noise, *J. Sound Vib.*, 51, 403–415, 1977.
52. Guccia, L., Passenger comfort, general issues. Presented at Sixth International Workshop on Railway Noise, Ile des Embiez, France, 1998.
53. Wetschurck, R. and Hauck, G., Geräusche und Erschütterungen aus dem Schienenverkehr (Noise and ground vibration from rail traffic), In *Taschenbuch der Technischen Akusti*, 2nd ed., Heckl, M. and Müller, H. A., Eds., Springer, Berlin, 1995.
54. Hardy, A. E. J., Railway passengers and noise, *Proc. Inst. Mech. Eng.*, 213F, 173–180, 1999.
55. Willenbrink, L., *Noise Inside and Outside Vehicles and from Railway Lines*, *Proceedings of Internoise 73*, Copenhagen, 1973, pp. 362–371.
56. Hardy, A. E. J., Measurement and assessment of noise within passenger trains, *J. Sound Vib.*, 231, 819–829, 2000.
57. Shaw, N.J., The prediction of railway vehicle internal noise using statistical energy analysis techniques. M.Sc. thesis, Heriot-Watt University, 1990.
58. Kohrs, T., Structural acoustic investigation of orthotropic plates, Diploma Thesis, TU Berlin, 2002.
59. Bracciala, A. and Pellegrini, C., FEM Analysis of the internal acoustics of a railway vehicle and its improvements. WCRR 97, Florence, 1997.
60. Létourneux, F., Guerrand, S., and Poisson, F., Low-frequency acoustic transmission of high-speed trains: simplified vibroacoustic model, *J. Sound Vib.*, 231, 847–851, 2000.
61. Stegeman, B., Development and validation of a vibroacoustic model of a metro rail car using Statistical Energy Analysis (SEA). M.Sc. dissertation, Chalmers University, Gothenburg, Sweden, 2002.
62. de Meester, K., Hermans, L., Wyckaert, K., and Cuny, N., *Experimental SEA on a Highspeed Train Carriage*, *Proceedings of ISMA21*, Leuven, Belgium, pp. 151–161, 1996.
63. Geissler, P. and Neumann, D., *SEA Modelling for Extruded Profiles for Railway Passenger Coaches*, *Euro Noise 98, München, Germany*, 1998, pp. 189–194.
64. Xie, G., Thompson, D. J., and Jones, C. J. C., *A Modelling Approach for Extruded Plates*, *Proceedings of 10th International Congress on Sound and Vibration*, Stockholm, 2003.
65. Backström, D., Analysis of the sound transmission loss of train partitions. M.Sc. thesis, KTH Stockholm, 2001.
66. ISO 2631. *Evaluation of human exposure to whole body vibration, part 1 general requirements*, International Standards Organisation, Geneva, 1985.
67. BS 7385. *Evaluation and measurement for vibration in buildings, part 2: guide to damage levels from ground-borne vibration*, British Standards Institution, London, 1993.
68. DIN 4150-3. *Vibration in buildings — part 3: effects on structures*, Deutsches Institut für Normung e.V., Berlin, 1999, February.
69. Krylov, V. V., Generation of ground vibration by superfast trains, *Appl. Acoust.*, 44, 149–164, 1995.

70. Madshus, C. and Kaynia, A. M., High speed railway lines on soft ground: dynamic behaviour at critical train speed, *J. Sound Vib.*, 231, 689–701, 2000.
71. Sheng, X., Jones, C. J. C., and Petyt, M., Ground vibration generated by a load moving along a railway track, *J. Sound Vib.*, 228, 129–156, 1999.
72. Edwards, J. W., *Survey of Environmental Noise and Vibration from London Underground Trains, Proceedings of Internoise '96*, Liverpool, pp. 2029–2032, 1996.
73. Sheng, X., Jones, C. J. C., and Thompson, D. J., A theoretical study on the influence of the track on train-induced ground vibration, *J. Sound Vib.*, 272, 909–936, 2004.
74. Sheng, X., Jones, C. J. C., and Petyt, M., Ground vibration generated by a harmonic load acting on a railway track, *J. Sound Vib.*, 225, 3–28, 1999.
75. Jones, C. J. C., Sheng, X., and Petyt, M., Simulations of ground vibration from a moving harmonic load on a railway track, *J. Sound Vib.*, 231, 739–751, 2000.
76. Sheng, X., Jones, C. J. C., and Thompson, D. J., A theoretical model for ground vibration from trains generated by vertical track irregularities, *J. Sound Vib.*, 272, 937–965, 2004.
77. Sheng, X., Jones, C. J. C., and Thompson, D. J., A comparison of a theoretical model for vibration from trains with measurements, *J. Sound Vib.*, 267, 621–635, 2003.
78. Jones, C. J. C., Thompson, D. J., and Petyt, M., A model for ground vibration from railway tunnels, *Proc. Inst. Civ. Eng., Transport.*, 153, 121–129, 2002.
79. Jones, C. J. C., *Groundborne Noise from New Railway Tunnels, Proceedings of Internoise '96*, Liverpool, 1996, pp. 421–426.
80. Griffin, M. J., *Handbook of Human Vibration*, Academic Press, London, 1990.
81. BS 6841, *Measurement and Evaluation of Human Exposure to Whole-Body Mechanical Vibration and Repeated Shock*, 1987.
82. CEN, *Railway Applications -Ride Comfort of Passengers — Measurement and Evaluation*, prENV 12299, 1996.
83. Wollström, M., Internal noise and vibrations in railway vehicles — a pilot study. Report TRITA — FKT 1998:44 of the Railway Technology Group, Department of Vehicle Engineering, Royal Institute of Technology (KTH), Stockholm, 1988.

---

# 11 Active Suspensions

*R. M. Goodall and T. X. Mei*

## CONTENTS

I.	Introduction .....	328
II.	Basics of Active Suspensions .....	328
	A. Concepts .....	328
	B. Active and Semi-Active .....	328
	C. Design Considerations .....	330
III.	Tilting Trains.....	331
	A. Concept and Equations .....	331
	B. Mechanical Configurations .....	333
	C. Control: Strategies and Assessment.....	334
	1. Control Approaches.....	334
	2. Assessment of Controller Performance .....	336
	D. Summary of Tilting.....	338
IV.	Active Secondary Suspensions .....	338
	A. Concepts and Requirements.....	338
	B. Configurations .....	339
	C. Control Strategies.....	339
	1. Sky-Hook Damping.....	339
	2. Softening of Suspension Stiffness.....	342
	3. Low-Bandwidth Controls .....	342
	4. Modal Control Approach .....	343
	5. Model-Based Control Approaches .....	344
	6. Actuator Response.....	344
	7. Semi-Active Control .....	344
	D. Examples .....	345
	1. Servo-Hydraulic Active Lateral Suspension.....	345
	2. Shinkansen/Sumitomo Active Suspension.....	346
V.	Active Primary Suspensions .....	347
	A. Concepts and Requirements.....	347
	B. Configurations .....	348
	C. Control Strategies.....	349
	1. Stability Control — Solid-Axle Wheelset .....	349
	2. Stability Control — Independently Rotating Wheelset.....	349
	3. Steering Control — Solid-Axle Wheelset .....	350
	4. Guidance Control — Independently Rotating Wheelset .....	350
	5. Integrated Control Design .....	351
	6. Assessment of Control Performance.....	351
	D. Examples .....	352
VI.	Technology .....	353

A. Sensing and Estimation Techniques ..... 353  
 B. Actuators ..... 354  
 C. Controllers and Fault Tolerance ..... 355  
 VII. Long Term Trends ..... 355  
 Nomenclature..... 355  
 References..... 356

**I. INTRODUCTION**

It is clear from the preceding chapters that the subject of railway vehicle dynamics has developed principally as a mechanical engineering discipline, but an important technological change is starting to occur through the use of active suspension concepts. The use of advanced control has been common for many decades in the power electronic control of traction systems, and it is now firmly established as the standard technology which has yielded substantial benefits, but its application to suspensions is much more recent. Although the term “active suspension” is commonly taken to relate to providing improved ride quality in fact, it is a generic term which defines the use of actuators, sensors, and electronic controllers to enhance and/or replace the springs and dampers that are the key constituents of a conventional, purely mechanical, “passive” suspension; as such it can be applied to any aspect of the vehicle’s dynamic system.

**II. BASICS OF ACTIVE SUSPENSIONS**

Vehicle dynamicists have been aware of active suspensions for some time, with major reviews having been undertaken in 1975, 1983, and 1997,<sup>1-3</sup> but so far they have only found substantial application in tilting trains — which can now be thought of an established suspension technology. However, there are two other major categories: active secondary suspensions for improved ride quality, and active primary suspensions for improved running stability and curving performance. The sections which follow in this chapter deal with these three categories in turn: tilting, active secondary, and active primary suspension, but first there are a number of general principles and considerations which need to be explained.

**A. CONCEPTS**

The general scheme of an active suspension is shown in diagrammatic form in [Figure 11.1](#). The input/output relationship provided by the suspension, which in the passive case is determined solely by the values of masses, springs, dampers, and the geometrical arrangement, is now dependent upon the configuration of sensors and actuators, and upon the control strategy in the electronics (almost invariably now involving some form of software processing). For all the three categories it will be seen that the introduction of active control enables things to be achieved that are either not possible or extremely difficult with a passive suspension.

**B. ACTIVE AND SEMI-ACTIVE**

The greatest benefits can be achieved by using fully-controllable actuators with their own power supply, such that the desired control action (usually a force) can be achieved irrespective of the movement of the actuator. Energy can flow from or to the power supply as required to implement the particular control law. This is known as a “full-active” suspension, but it is also possible to use a “semi-active” approach in which the characteristic of an otherwise passive suspension component can be rapidly varied under electronic control — see [Figure 11.2](#). Semi-active

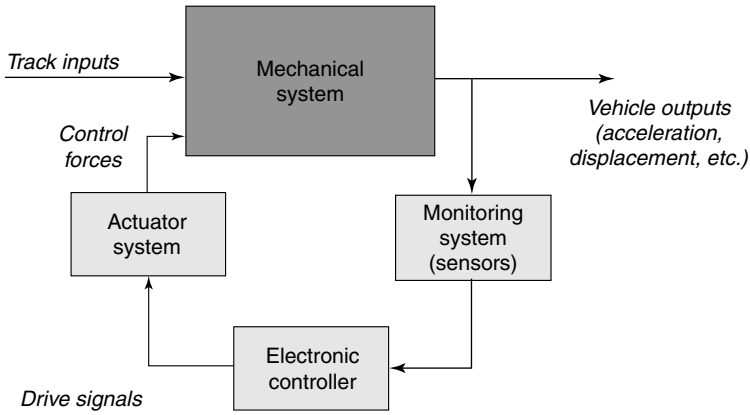


FIGURE 11.1 Generalised active suspension scheme.

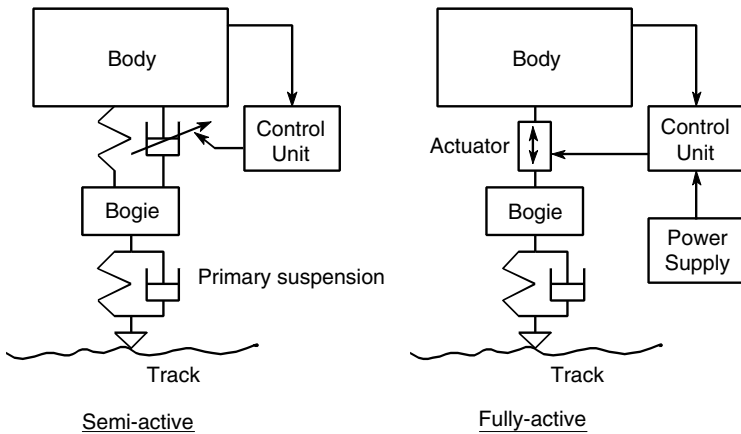


FIGURE 11.2 Semi- and full-active control.

suspensions usually use controllable dampers of some kind, although the concept is not restricted to dampers.<sup>a</sup>

The benefit of the semi-active approach compared with full-active is one of simplicity, because a separate power supply for the actuator is not needed. The disadvantage of a semi-active damper is that the force remains dependent upon the speed of damper movement, which means that large forces cannot be produced when its speed is low, and, in particular, it cannot develop a positive force when the speed reverses because it is only possible to dissipate energy, not inject it. Figure 11.3 clarifies the limitation by showing areas on the force–velocity diagram that are available for a semi-active damper based upon its minimum and maximum levels, whereas an actuator in a full-active system can cover all four quadrants. This limitation upon controllability restricts the performance of a semi-active suspension to a significant degree.<sup>4</sup>

Closely related is an option known variously as semi-passive, adjustable passive or adaptive passive, in which the characteristics are varied on the basis of a variable which is not influenced by the dynamic system being controlled (e.g., as a function of vehicle speed).

<sup>a</sup> An interesting option would be the use of an electronically-controllable spring to provide a semi-active suspension, but as far as the authors are aware, no such device has been invented.



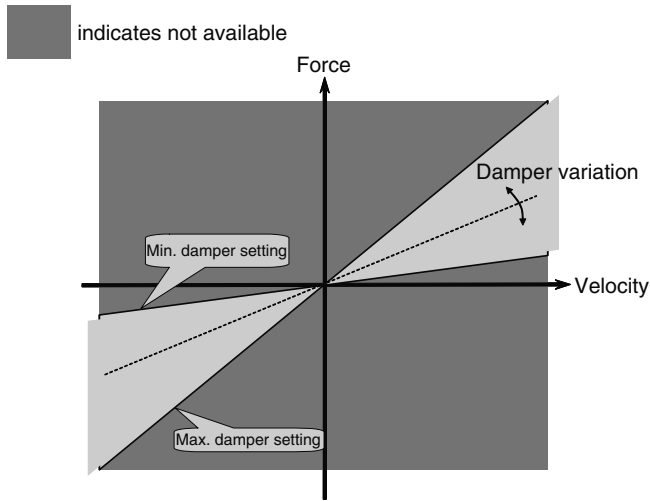


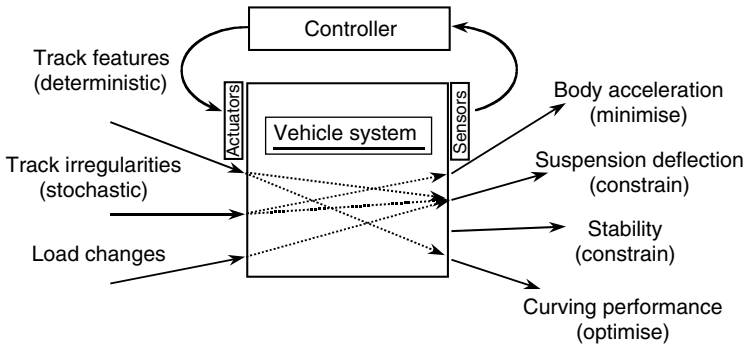
FIGURE 11.3 Force–velocity diagram for semi-active damper.

### C. DESIGN CONSIDERATIONS

For designing active suspension systems such as these, an important difference arises compared with passive suspensions. A conventional suspension is designed with as accurate a model as possible so that the computer simulation can predict the on-track performance effectively. The designer then adjusts the values of the suspension components based upon well-understood expectations for the particular vehicle configuration until the required performance is achieved. However, for an active suspension, it is important to distinguish between the design model and the simulation model: the former is a simplified model used for synthesis of the control strategy and algorithm, whereas the latter is a full-complexity model to test the system performance, i.e., as used for conventional suspensions. The importance of having an appropriately simplified design model is less profound when “classical” control design techniques are being used, although even here key insights arise with simplified models; the real issue arises when modern model-based design approaches are being used, either for the controller itself or for estimators to access difficult or impossible to measure variables, in which case the controller and/or estimator assumes a dynamic complexity equal to or greater than that of the design model. Since a good simulation model of a railway vehicle will usually have more than a hundred states, a controller based upon this model would at best be overly complex to implement, at worst impossible because some of the states may be uncontrollable or unobservable.

There are formal methods for reducing the model complexity, but often engineering experience will provide a suitable abstraction. For example, there is a relatively weak coupling between the vertical and lateral motions of rail vehicles and, depending on the objectives, only selected degrees of freedom need to be included in the design model. Common simplifications are based around a vehicle model that is partitioned into side-view, plan-view, and end-view models: the side-view model is concerned with the bounce and pitch degrees of freedom, and can be used for active vertical suspensions; the plan-view model deals with the lateral and yaw motions, and can be used for active lateral suspensions and active steering/stability control; the end-view model covers the bounce, lateral, and roll motions, and can be used for the design of tilting controllers.

It is, of course, essential that such modelling software can support the integration of the controller into the mechanical system. This can be achieved within a single package, but, there is



**FIGURE 11.4** Design process.

a strong argument for distinct but well-integrated software, i.e., one of the many MBS dynamics packages in combination with a control design package such as Matlab/Simulink<sup>®</sup>. Ideally, there should be a number of interface possibilities: controllers designed using the simplified design model need to be exported into the MBS package for simulation purposes; equally it is often valuable to be able to export a complex but linearised model from the MBS package for further controller evaluation using the targeted analytical tools provided for controller design; and finally, running the two packages simultaneously in a co-simulation mode is also important because this avoids the need for conversion and export, although the data transfer process must be robust.

A final point is illustrated by Figure 11.4, which emphasises the multi-objective nature of the design process. There are a variety of input types and output variables that must be considered, and each output will be affected by different combinations of inputs. The design will require an optimisation involving constraints. For example; an active secondary suspension design must minimise the frequency-weighted accelerations on the vehicle body without exceeding the maximum suspension deflection; an active primary suspension must optimise the curving performance whilst maintaining adequate levels of running stability on straight track; etc.

### III. TILTING TRAINS

The earliest proposals for tilting trains go back into the first half of the 20th century, but it was not until the 1960s and 1970s that experimental developments were aimed towards producing operational trains for prestigious high-speed routes. These emerged as the Talgo Pendular in Spain (1980), the APT in the UK and the LRC in Canada (1982), the first ETR 450 Pendolino trains in Italy (1988), and the X2000 in Sweden (1990). A similar pattern occurred in Japan, although the developments there were aimed at the regional/narrow-gauge railways rather than the high-speed Shinkansen. The 1990s saw tilting mature into a standard railway technology, with applications extending throughout most of Europe and Japan, and all the major rail vehicle manufacturers now offer and supply tilting trains for regional and high-speed applications.

#### A. CONCEPT AND EQUATIONS

Tilting trains take advantage of the fact that the speed through curves is principally limited by passenger comfort, and not by either the lateral forces on the track or the risk of overturning, although these are constraints that cannot be ignored. Tilting the vehicle bodies on curves reduces the acceleration experienced by the passenger, which permits higher speeds and provides a variety of operational benefits. The principles and basic equations related to tilting are relatively straightforward and are explained here in a manner that focuses upon the operational advantages.

There are two primary decisions that need to be made. The first is what maximum tilt angle is to be provided ( $\theta_{\text{tilt}}$ ), a decision based upon mechanical design of the vehicle, especially taking gauging issues into account. The second decision is what cant deficiency the passengers should experience on a steady curve ( $\theta_{\text{active}}$ ), which clearly is of primary importance to comfort. Given these two decisions, and the cant deficiency that applies for the passive (nontilting) case ( $\theta_{\text{passive}}$ ), it is possible to derive an equation for the increase in speed offered by tilt. Note that, although the curve radius and the acceleration due to gravity appear in the basic acceleration equations, they disappear when the equation is dealing with the fractional or percentage speed increase:

$$\text{speed increase} = \frac{V_{\text{active}} - V_{\text{passive}}}{V_{\text{passive}}} = \left\{ \sqrt{\frac{\sin(\theta_{\text{cant}} + \theta_{\text{tilt}} + \theta_{\text{active}})}{\sin(\theta_{\text{cant}} + \theta_{\text{passive}})}} - 1 \right\} \times 100\% \quad (11.1)$$

Although in principle the cant deficiency could be fully compensated by the tilting action, i.e., to make  $\theta_{\text{active}} = 0$ , in practice this is not sensible either from the operational or the ride comfort viewpoint. It is possible to recognise this by introducing a “cant deficiency compensation factor” ( $K_{\text{CD}}$ ), an important design parameter in the tilt controller, the choice of which will be discussed later.

$$K_{\text{CD}} = 1 - \frac{\theta_{\text{active}}}{(\theta_{\text{active}} + \theta_{\text{tilt}})}, \text{ i.e., } \frac{\theta_{\text{tilt}}}{\theta_{\text{active}} + \theta_{\text{tilt}}} \quad (11.2)$$

Consider some examples: track cant is usually  $6^\circ$ , and typically  $6^\circ$  of cant deficiency is applied for a nontilting train. Applying  $9^\circ$  of tilt and a cant deficiency of  $6^\circ$  for the tilting train, the calculation indicates a speed-up of 32% with a compensation factor of 60%. In this particular case, the passengers nominally experience the same comfort level on curves (although the passive vehicle will usually roll out by a small angle, typically less than  $1^\circ$ , so in practice tilting will give a small reduction in the curving acceleration). Another example might be where the tilting cant deficiency is reduced to  $4.5^\circ$ , perhaps to offer an improved ride comfort; using a slightly smaller tilt angle of  $8^\circ$ , the speed-up falls to 24% with a compensation factor of 64%.

Speeds on curves may, therefore, be theoretically increased by around 30% or more with tilting trains. However, the performance on curve transitions as well as the steady curves is important from a comfort point of view, and the comfort level can be predicted using a method described by a European standard.<sup>5</sup> It is based on an empirically-based method in which the percentage of passengers ( $P_{\text{CT}}$ ) that are likely to feel uncomfortable during the curve is determined from the lateral acceleration ( $\ddot{y}$ ), the lateral jerk ( $\dot{\ddot{y}}$ ), and the body roll velocity ( $\dot{\theta}$ ) experienced during the transition. Details of the method are given in the quoted reference, including the way in which the three measurements should be made. Equation 11.3a gives the appropriate empirically-derived equations, and the constants which must be used to calculate the  $P_{\text{CT}}$  factor, a separate calculation for seated and standing passengers derived from either simulated or measured performances of the vehicle at the entry to a curve Table 11.1 lists.

$$P_{\text{CT}} = |(A\ddot{y} + B\dot{\ddot{y}} - C)|_{\leq 0} + D\dot{\theta}^E \quad (11.3a)$$

There is also the issue of motion sickness. In contrast to the curve transition comfort level, which may be considered on a curve-by-curve basis, motion sickness is a cumulative effect, which comes as a consequence of a number of human factors, the exact nature of which is not fully understood. Again, the effect is aggravated on highly curvaceous routes with rapid transitions.<sup>6</sup> The degree to which the curving acceleration is compensated for by the tilting action is an important factor, but once this has been optimised, the only other mitigation measure is operating at lower speed.

**TABLE 11.1**  
**Constants for  $P_{CT}$  Equation**

Condition	A	B	C	D	E
Standing	28.54	20.69	11.1	0.185	2.283
Seated	8.97	9.68	5.9	0.120	1.626

## B. MECHANICAL CONFIGURATIONS

Broadly speaking, there are four mechanical arrangements which are possible to provide the tilting action.

The first is passive or pendular tilt, in which the secondary suspension is raised to around roof level in the vehicle: the vehicle centre of gravity is then substantially below the suspension and the body naturally swings outwards, reducing the lateral curving acceleration experienced by the passengers. This is a technique pioneered in the Talgo trains — the air springs are raised by means of vertical pillars at the vehicle ends, an arrangement made much easier by the articulated configuration of the trains.

A second approach is to achieve tilt directly by applying active control to the secondary roll suspension. One method which has been tried in both Europe and Japan is to apply differential control to the air springs, but this may cause a dramatic increase in air consumption and generally has not found favour, although one Japanese development has achieved it by transferring air between the air springs using a hydraulically-actuated pneumatic cylinder.<sup>7</sup> The alternative method of direct control of the roll suspension is by means of an active anti-roll bar (stabiliser), and this is applied in Bombardier's regional Talent trains. This uses the traditional arrangement consisting of a transversely-mounted torsion tube on the bogie with vertical links to the vehicle body, except that one of the links is replaced by a hydraulic actuator, and thereby applies tilt via the torsion tube.

The previous two arrangements are very much minority solutions, because most implementations use a tilting bolster to provide the tilt action. An important distinction is where this bolster is fitted compared with the secondary suspension, which leads to the third and fourth of the arrangements. With the tilting bolster above the secondary suspension, the increased curving forces need to be reacted by the secondary lateral suspension; since a stiffer lateral suspension is not consistent with the higher operating speed of a tilting train, in practice, either an increased lateral suspension movement or some form of active centring method is needed to avoid reaching the limits of travel.

The final arrangement has the tilting bolster below the secondary suspension, thereby avoiding the increased curving forces on the lateral suspension, and this is probably the most common of all schemes, the necessary rotation being achieved using either a pair of inclined swing links, or a circular roller beam. Typical schemes with inclined swing links and with a roller beam are shown in [Figure 11.5](#).

Actuators to provide tilt action have seen significant development since the early days of tilt. Some early systems were based upon controlling the air springs (i.e., intrinsically pneumatic actuation), but it was more normal to use hydraulic actuators because these tend to be the natural choice for mechanical engineers. However, experiments with electro-mechanical actuators in the UK in the 1970s, in Switzerland in the 1980s, and in Germany in the 1990s, paved the way for a progressive change away from the hydraulic solution. Electric motors controlled by solid-state power amplifiers drive screws fitted with high-efficiency ball or roller nuts to convert rotary to linear motion. They are less compact than hydraulic actuators at the point of application, but, overall, they provide significant system benefits and they are now employed in the majority of new

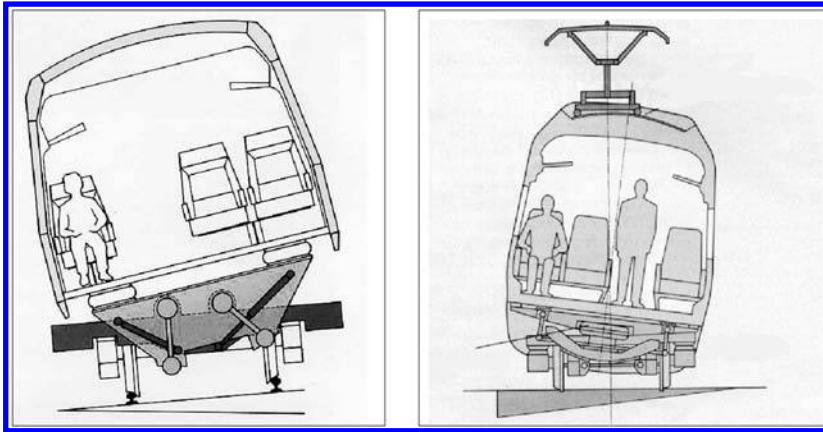


FIGURE 11.5 Tilt below secondary schemes.

European tilting trains. Interestingly, Japanese tilting technology has tended to use pneumatic actuators.

**C. CONTROL: STRATEGIES AND ASSESSMENT**

This section explains some of the essential control approaches that are possible to achieve effective tilting action, and then discusses how the performance of particular controllers can be assessed.

**1. Control Approaches**

The most intuitive control approach is to put an accelerometer on the vehicle body to measure the lateral acceleration that the tilt action is required to reduce, yielding the “nulling” controller shown in Figure 11.6(a). The accelerometer signal is used to drive the actuator in a direction that will bring it towards zero, i.e., a classical application of negative feedback. Implementation of the required value of  $K_{CD}$  can be achieved with a modification of the basic nulling controller to give a partial tilt action by including a measure of the tilt angle in the controller, as shown by the dotted arrow on the figure. However, there is a difficulty with this scheme due to interaction with the lateral

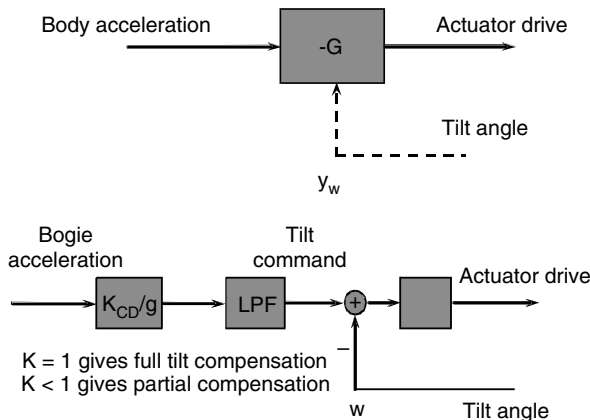


FIGURE 11.6 Tilt control methods.

suspension: the roll and lateral modes of the vehicle body are strongly coupled in a dynamic sense, and it can be shown that if the loop bandwidth is low enough not to interfere with the lateral suspension, it is then too slow-acting on the curve transition.

Figure 11.6(b) shows the next solution: the dynamic interaction problem can be avoided by putting the accelerometer on a nontilting part, in other words the bogie. This will then tell how much tilt is needed to reduce the lateral acceleration on the vehicle body, and can be multiplied by the factor  $K_{CD}$  which determines what proportion of the lateral acceleration is to be compensated;  $K_{CD} = 1$  gives 100% compensation, not a good idea for motion sickness reasons, and typically 60 or 70% compensation is used (as mentioned above). This “tilt angle command signal” then provides the input to a feedback loop which uses a measurement of the tilt angle.

Unfortunately there is still a problem, because the accelerometer on the bogie is not only measuring the curving acceleration, but also the pure lateral accelerations due to track irregularities. With the accelerometer on the vehicle body, these accelerations are reduced by the secondary suspension, but they are much larger when the accelerometer is on the bogie. Consequently, it is necessary to add a low-pass filter (LPF) to reduce the acceleration signals caused by the track irregularities, otherwise there is too much tilt action on straight track resulting in a worse ride quality. However, to apply sufficient filtering, there is also too much delay introduced at the start of the curve, so the full lateral curving acceleration is felt for a short time, even though it reduces to an acceptable level once properly on the curve.

Figure 11.7 shows the next step: the signal from the vehicle in front is used to provide precedence, carefully designed so that the delay introduced by the filter compensates for the precedence time corresponding to a vehicle length. In effect, this scheme is what most European tilting trains now use; sometimes roll and/or yaw gyros are used to improve the response, and normally a single command signal is generated from the first vehicle and transmitted digitally with appropriate time delays down the train.

The signal from the bogie-mounted accelerometer is essentially being used to generate an estimate of the true cant deficiency of the track’s design alignment, the difficulty being to exclude the effects of the track irregularities. An obvious development is to feed the vehicle controllers with signals from a database which defines the track, instead of from the accelerometer. Both the position of the vehicle along the track and the curve data contained in the database need to be known accurately for this approach to work effectively, but it is likely that such systems will become the norm in the future.

Japanese tilting trains often use a balise on the track ahead of the curve to initiate the tilting action, a technique which helps to mitigate the relatively slow response of the pneumatic tilt actuators.

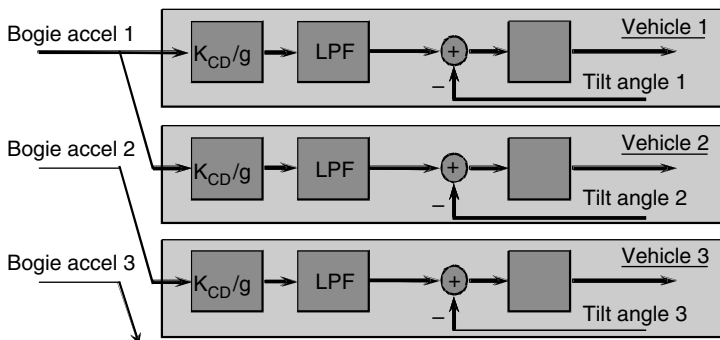


FIGURE 11.7 Precedence tilt control scheme.

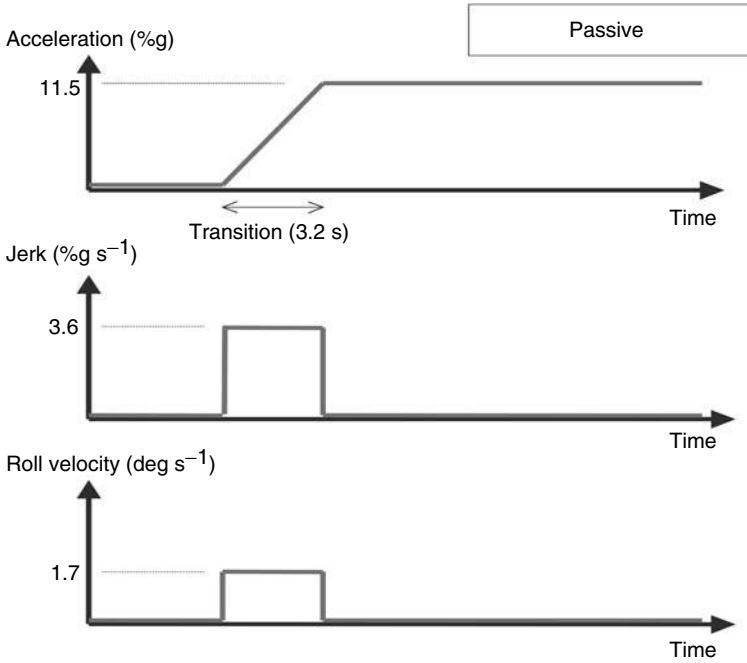


FIGURE 11.8 “Ideal” passive transition responses.

### 2. Assessment of Controller Performance

It is clear that what happens in the steady curve is important, however, the dynamic response during the transition must also be considered. In an ideal tilt control strategy, the tilt angle of the body should rise progressively, perfectly aligned both with the onset of curving acceleration and the rising cant angle, and the difficulties in achieving this kind of response have been explained above. Since the principal benefit of tilt is to be able to operate at higher speeds without degradation in

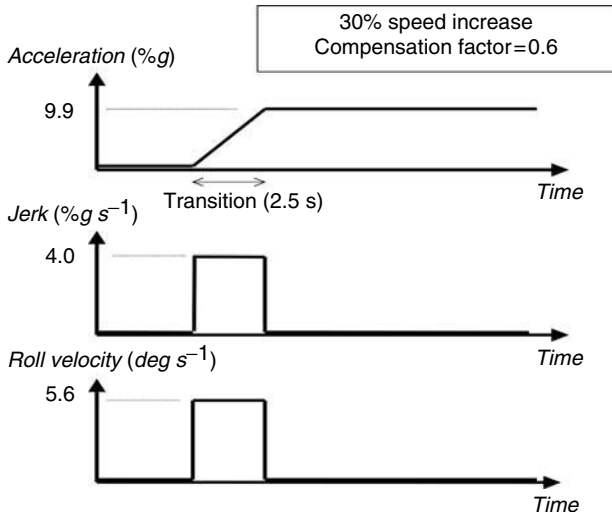


FIGURE 11.9 “Ideal” tilting responses.

passenger comfort, from a design point of view there are two issues: how well does the tilting vehicle perform on straight track, and how well does it perform on curve transitions?

The accelerometer-based control strategies means these two issues must, in practice, be traded off against each other — if the tilt action is fast to give good transition performance, in general, the straight track ride quality may be degraded. Qualitatively, a good tilt controller responds principally to the deterministic track inputs, and as much as possible ignores the random track irregularities. In order to assess different tilt control strategies in an objective manner, it is necessary to define appropriate criteria and conditions.

The straight track performance can be dealt with using a criterion of degrading the lateral ride quality by no more than a specified margin compared with the nontilting response, a typical value being 7.5%. Note that for assessing the tilt controller performance, this comparison must be made at the higher speed. Of course, a comparison of ride quality with a lower speed vehicle is also needed,

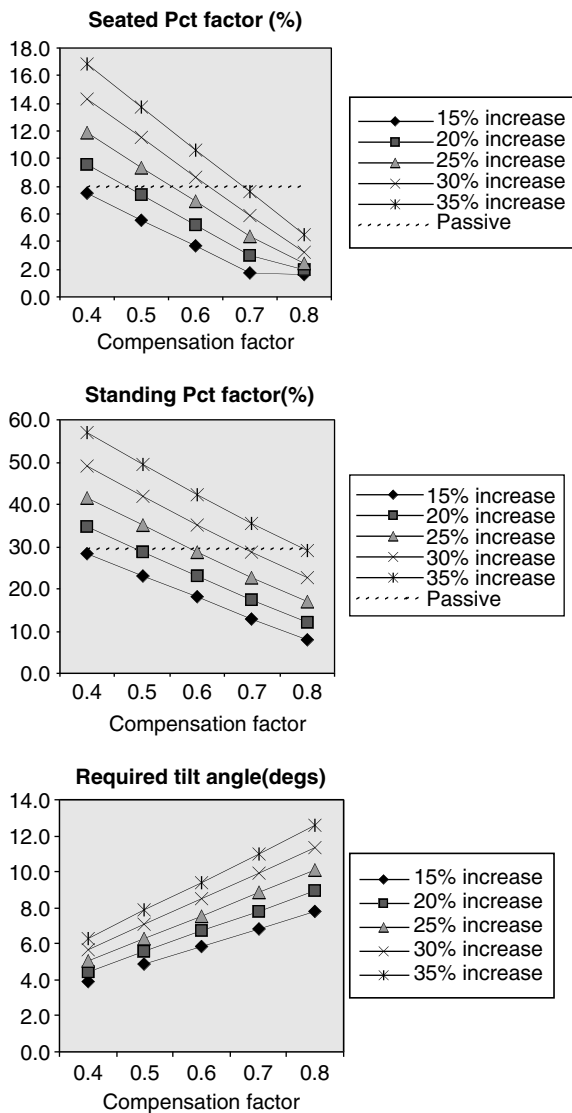


FIGURE 11.10 Comfort factors and tilt angle results.



but achieving a satisfactory ride quality at elevated speeds will require either an improved suspension or a better quality track, i.e., not a function of the tilt controller.

The curve transition response has to be separated into two aspects. Firstly, the fundamental tilting response, measured by the  $P_{CT}$  factors as described previously, must be as good as a passive vehicle at lower (nontilting) speed, otherwise the passenger comfort will inevitably be diminished, no matter how effective the tilt control is. It is possible, therefore, to introduce the idea of “ideal tilting” where the tilt action follows the specified tilt compensation perfectly, defined on the basis of the fundamental tilt system parameters — the operating speed (increase), maximum tilt angle, and the cant deficiency compensation factor. This combination of parameters can be optimised using the  $P_{CT}$  factor approach for deterministic inputs in order to choose a basic operating condition, and this will give “ideal”  $P_{CT}$  values (one for standing, one for sitting).

Consider, for example, the ideal transition responses for passive and tilting trains shown in [Figure 11.8](#) and [Figure 11.9](#), where the transition length gives a time of 3.2 sec for the passive vehicle and both cant and cant deficiency are  $6^\circ$ . (The passive response also includes the effect of a “passive roll-out” of  $1^\circ$ , but this is obviously vehicle-dependent.) [Figure 11.9](#) shows the corresponding acceleration, jerk, and roll velocity graphs for a particular tilting condition, i.e., 30% higher speed with a compensation factor of 0.6, but of course similar diagrams can be developed for other conditions.

The three graphs in [Figure 11.10](#) show the results of  $P_{CT}$  calculations undertaken with speed-up factors of between 15 and 35% and compensation factors from 40 to 80%, where the dotted horizontal lines show the values for the slower nontilting train, plus the corresponding tilt angle requirement. In this case, with a relatively slow transition, increasing the compensation factor improves the comfort level, although this is not necessarily the case with faster transitions; however, it can be seen that a larger tilt angle is required.

The other consideration is that it is necessary to quantify the additional dynamic effects which are caused by the suspension/controller dynamics as the transitions to and from the curves are encountered, which can be quantified as the deviations from the “ideal” response mentioned in the previous paragraph. These deviations relate to both the lateral acceleration and roll velocity, although the former is likely to be the main consideration. The performance in this respect will depend upon detailed characteristics of the controller, such as the filter in the command-driven scheme and the tuning parameters in the tilt angle feedback loop. It is clear that the deviations need to be minimised, but at present there is no information regarding their acceptable size, although the values derived for a normal passive suspension can be used as a guide.

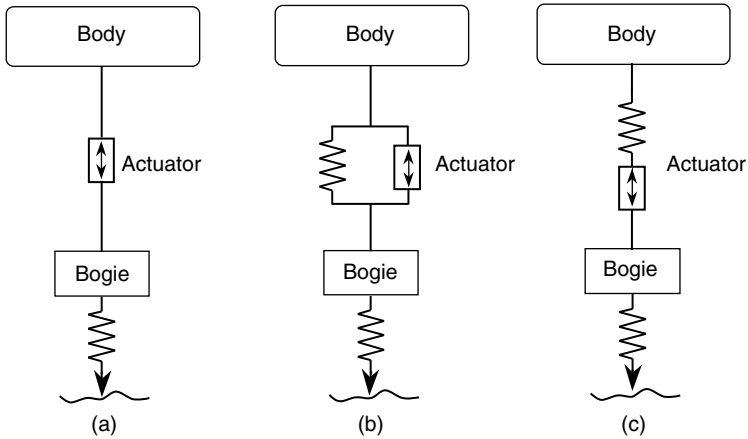
## D. SUMMARY OF TILTING

It should be emphasised that, although tilting seems in many ways to be a rather simple concept, it requires considerable care in practice and has taken a number of years to introduce reliable operational performance, and tilting controllers still need adjustment for specific route characteristics. It is likely that the state-of-the-art will continue to be developed in the years to come.

## IV. ACTIVE SECONDARY SUSPENSIONS

### A. CONCEPTS AND REQUIREMENTS

For the secondary suspensions, active controls improve the vehicle dynamic response and provide a better isolation of the vehicle body to the track irregularities than the use of only passive springs and dampers. Active control can be applied to any or all of the suspension degrees-of-freedom, but, when applied in the lateral direction, will implicitly include the yaw mode, and in the vertical direction will include the pitching mode. (Controlling in the roll direction is of course equivalent to tilting, which is essentially a particular form of active secondary suspension, but of sufficient



**FIGURE 11.11** Active secondary suspension actuator configurations.

importance to have its own section.) The improved performance can be used to deliver a better ride quality, but this is not directly cost-beneficial and so normally will be used to enable higher train speed whilst maintaining the same level of passenger comfort. The other possibility is to provide the same ride quality on less well aligned track, in which case the cost-benefit analysis needs to take account of the reduced track maintenance cost.

## B. CONFIGURATIONS

Active secondary suspensions can be used in the lateral and/or vertical directions and a number of actuator configurations are possible as illustrated in Figure 11.11.

Actuators can be used to replace the passive suspensions as shown in Figure 11.11(a) and the suspension behaviour will be completely controlled via active means. In practice, however, it is more beneficial that actuators are used in conjunction with passive components. When connected in parallel, as illustrated in Figure 11.11(b), the size of an actuator can be significantly reduced as the passive component will be largely responsible for providing a constant force to support the body mass of a vehicle in the vertical direction or quasi-static curving forces in the lateral direction. On the other hand, fitting a spring in series with the actuator, as shown in Figure 11.11(c), helps with the high frequency problem caused by the lack of response in the actuator movement and control output at high frequencies (see [Section IV.C.6](#), Actuator response), and in practice a combination of a parallel spring for load-carrying and a series spring to help with the high frequency response is the most appropriate arrangement. The stiffness of the series spring depends upon the actuator technology: a relatively high value can be used for technologies such as hydraulics that have good high frequency performance, and a softer value for other technologies which means that achieving a high bandwidth is more problematic.

The other option is to use actuators mounted between adjacent vehicles, although the improvement of ride quality is less significant and, in general, the design problem is more difficult because the complete train becomes strongly coupled in a dynamic sense via the actuators.

## C. CONTROL STRATEGIES

### 1. Sky-Hook Damping

There are different control approaches possible for active suspensions. A high bandwidth system, which deals with the random track inputs caused by irregularities, can be used to improve suspension performance largely through the provision of damping to an absolute datum.

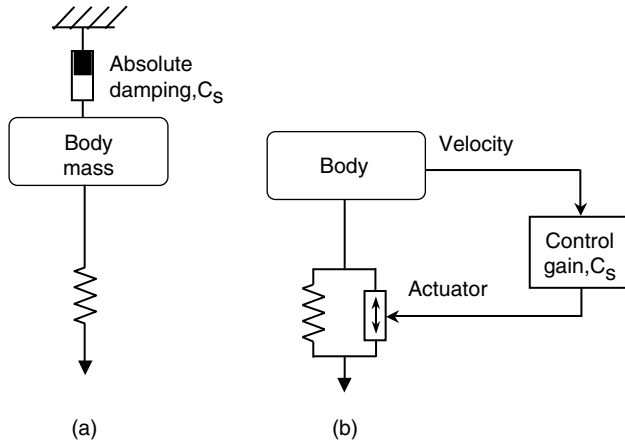


FIGURE 11.12 Sky-hook damping.

The principle of absolute damping is depicted in Figure 11.12(a), where a damper is connected from the mass to the sky, hence the term “sky-hook” damping. For practical implementations, the principle of the sky-hook damping can be realised by the arrangement shown in Figure 11.12(b). The feedback measurement is provided from a sensor mounted above the suspension on the body and the control demand is fed to the actuator which is placed between the vehicle body and the bogie.

A comparison between the passive and the sky-hook damping of a simple (one-mass) system illustrates the potential advantages of the active concept very well. For a passive damper, a higher level of modal damping can only be achieved at the expense of increased suspension transmissibility at high frequencies, as shown in Figure 11.13. For the sky-hook damper, however, the high frequency responses are independent of the damping ratio, and the transmissibility is significantly lower than that of the passive damping at all frequencies concerned. This is also the consequence of applying optimal control, as described in Ref. 8.

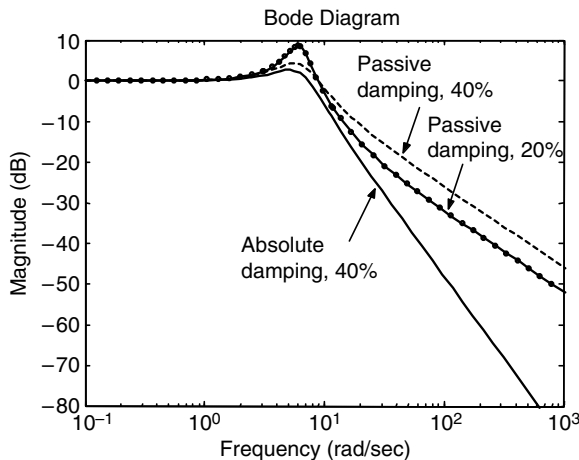


FIGURE 11.13 Comparison of passive and absolute damping.

The equation which implements the skyhook control law is simple, i.e.,

$$F_a = -C_s \frac{dz}{dt} \quad (11.3b)$$

where  $C_s$  is skyhook damping coefficient and  $F_a$  is the actuator force.

This yields the transfer function for a simple single-mass suspension as follows:

$$\frac{z}{z_t} = \frac{K}{K + sC_s + s^2M} \quad (11.4)$$

where  $K$  [N/m] and  $M$  [kg] are the spring constant and mass.

The equivalent transfer function for the passive suspension with a conventional damper having a coefficient  $C$  (Ns/m) is

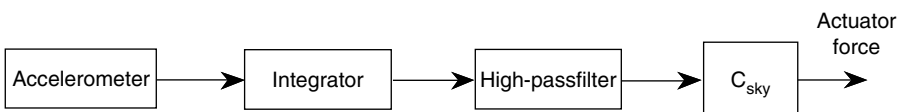
$$\frac{z}{z_t} = \frac{K + sC}{K + sC + s^2M} \quad (11.5)$$

from which it can be seen that the high frequency response is  $\propto 1/f$  for the passive suspension, compared with  $\propto 1/f^2$  for the active skyhook suspension, the overall effect of which was seen in [Figure 11.13](#).

Skyhook damping gives a profound improvement to the ride quality for straight track operation, however, it creates large deflections at deterministic features such as curves and gradients. Although this can be accommodated in the control design, e.g., by filtering out the low frequency components from the measurements which is largely caused by track deterministic features,<sup>9</sup> it is recognised that reducing the deterministic deflections to an acceptable level will compromise the performance achievable with “pure” skyhook damping. In fact, the absolute velocity signal that is required for skyhook damping will usually be produced by integrating the signal from an accelerometer, and so, in practice, it will also be necessary to filter out the low frequency components in order to avoid problems with thermal drift in the accelerometer — a typical scheme is shown in [Figure 11.14](#). In practice, the integrator and high-pass filter will normally be combined to provide a “self-zeroing” integration effect.

Whilst the use of a high-pass filter can eliminate the quasi-static suspension deflections due to the large quasi-state force of the skyhook damping on gradients or curves, it is less effective in reducing the transient suspension travel on track transitions, and in the selection of the filter cut-off frequency there is a difficult trade-off between the ride quality improvement of the vehicle body and the maximum movement of the suspension.

There are a number of possible solutions proposed to overcome the problem. The complementary filtering approach, as shown in [Figure 11.15](#), uses a relative damping force at the low frequency range in addition to the sky-hook damping at high frequencies, which results in a much improved trade-off. There are also Kalman filter based strategies where the effect of the track deterministic input can be minimised or the track features are directly estimated.<sup>10</sup> A typical trade-off comparison between different control approaches is given in [Figure 11.16](#), in this case for the vertical suspension of a vehicle running onto a gradient.<sup>9</sup>



**FIGURE 11.14** Practical implementation of skyhook damping.

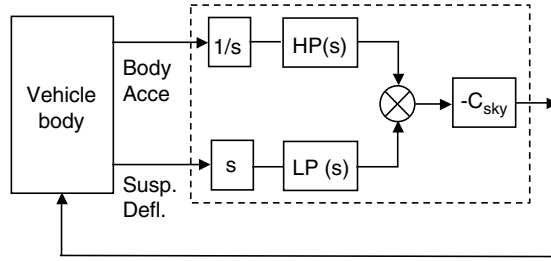


FIGURE 11.15 Complementary filters.

### 2. Softening of Suspension Stiffness

Another strategy is to create a softer suspension by controlling the actuator to cancel part of the suspension force produced by the passive stiffness. The control equation is of a simple form as shown in Equation 11.6, but note that positive feedback is used to reduce the overall stiffness to a value of  $(K - K_s)$ . The corresponding transfer function is not given because it is a trivial change to what was given for the passive suspension.

$$F_a = +K_s(z - z_t) \tag{11.6}$$

### 3. Low-Bandwidth Controls

Active secondary suspensions can also be used to provide a low bandwidth control, which is similar to tilting controls in that the action is intended to respond principally to the low frequency deterministic track inputs. In low bandwidth systems, there will be passive elements which dictate the fundamental dynamic response, and the function of the active element is associated with some low frequency activity. A particular use of the concept is for maintaining the average position of the suspension in the centre of its working space, thereby minimising contact with the mechanical

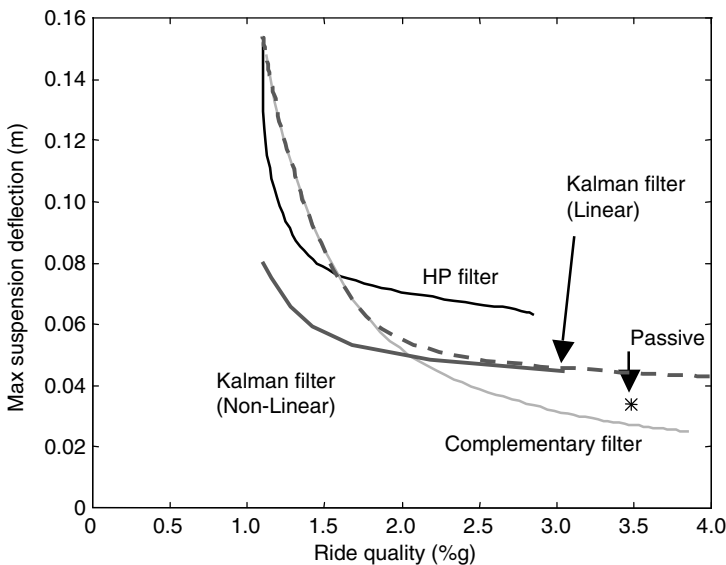


FIGURE 11.16 Trade-off between ride quality and suspension deflection.

limits of travel, and enabling the possibility of a softer spring to be used.<sup>11,12</sup> This is a powerful technique for the lateral suspensions because curving forces are large, and without centring action there may sometimes be significant reductions in ride quality whilst curving.

The idea of active levelling (or centring for a lateral suspension) can be achieved using the equation.

$$F_a = -K_L \int (z - z_t) dt$$

The suspension transfer function becomes

$$\frac{z}{z_t} = \frac{K_L + Ks + Cs^2}{K_L + Ks + Cs^2 + Ms^3} \tag{11.7}$$

The integral action changes it from second to third order, the effect of which is less obvious, but it can readily be shown that the suspension deflection  $(z - z_t)$  is zero in response to an acceleration input from the track, and it is this characteristic that corresponds to the self-levelling effect.

### 4. Modal Control Approach

For a conventional railway vehicle with two secondary suspensions between the body frame and the two bogies, it is possible to use local control for each suspension, i.e., the measurement from the sensor(s) mounted above either of the bogies is fed to the controller which controls the actuator on the same bogie. However, the tuning of control parameters may be problematic, as interactions between the two controllers via the vehicle body will be inevitable. To overcome this difficulty, a centralised controller for both suspensions may be used to enable independent control of the body modes.

Figure 11.17 shows how the lateral and yaw modes of a vehicle body can be separately controlled by using active suspensions in the lateral direction, and a similar scheme can be applied to actuators in the vertical direction to control the bounce and pitch modes. The output measurements from the two bogies are decomposed to give feedback signals required by the lateral and yaw controllers, respectively, and the output signals from the two controllers are then recombined to control two actuators at the two bogies accordingly. In this way, it is possible to apply different levels of control, in particular to reduce the suspension frequency and add more damping to the yaw (or pitch) mode, which is less susceptible to the low frequency deterministic inputs.

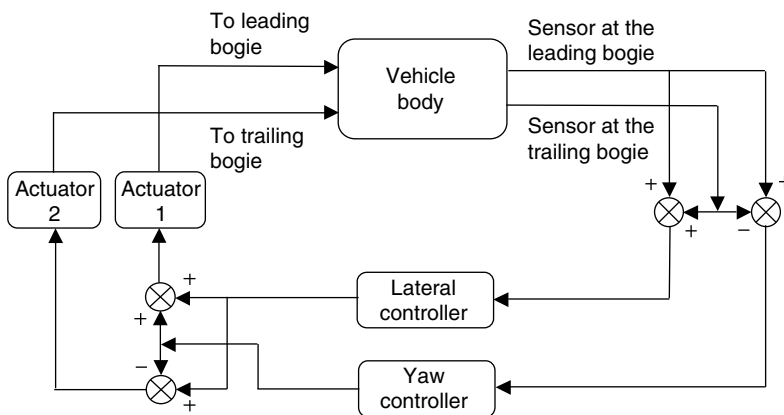


FIGURE 11.17 Modal control diagram.

### 5. Model-Based Control Approaches

Increased system complexity also encourages the use of mathematically rigorous design approaches such as optimal control, which enables a trade-off between ride quality and suspension deflection to be formally defined and optimised.<sup>13</sup> Equation 11.8 gives a typical cost function which is minimised in the design of an optimal controller to reflect the suspension design problem. Suitable choices of the weighting factors  $q_1$ ,  $q_2$ , and  $r$  (on the body acceleration  $a_b$ , suspension deflection  $x_b$  and actuator force  $F_a$ ) enable an appropriate design trade-off to be achieved.

$$J = \int (q_1 a_b^2 + q_2 x_d^2 + r F_a^2) dt \tag{11.8}$$

### 6. Actuator Response

In order to implement the control laws, for example, those listed in the previous subsection, it is necessary to have force control. However, very few actuator types inherently provide a force and so an inner force feedback loop is required, but it is important to appreciate that dynamics of this actuator force loop need to be significantly faster than is immediately obvious. The physical explanation can be seen from Figure 11.18, which is a generalised scheme of a force-controlled actuator.

The force command to the actuator would be generated by an active suspension controller, not shown here because it is useful to consider what happens even with a zero force command, which should in principle leave the suspension response unchanged compared with the passive suspension. The track input will impact upon the dynamic system, and this will cause actuator movement which the force control loop must counteract in order to keep its force as close as possible to zero. Remembering that the actuator will be connected across the secondary suspension, its movements at low frequencies will be small as the vehicle follows the intended features of the track, but relatively large at high frequencies as the suspension provides isolation by absorbing the track irregularities. How well the actuator generates the force required of it in the presence of the high frequency movement depends upon the characteristics of the actuator, and it is not possible to generalise. A more detailed analysis reveals that a force loop bandwidth in the region of 20 Hz will still yield noticeable degradations in the acceleration p.s.d. on the suspended mass at around 4 Hz, but this analysis is beyond the scope of this handbook because it is a detailed control engineering issue. However, studies of this problem can be found in Ref. 14.

### 7. Semi-Active Control

The basis of controlling a semi-active system is to replicate, as far as possible, the action of sky-hook damping.<sup>4</sup> Most semi-active control strategies are based upon achieving the demanded force

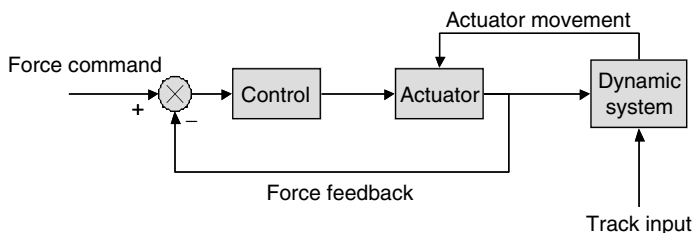


FIGURE 11.18 Actuator force control.

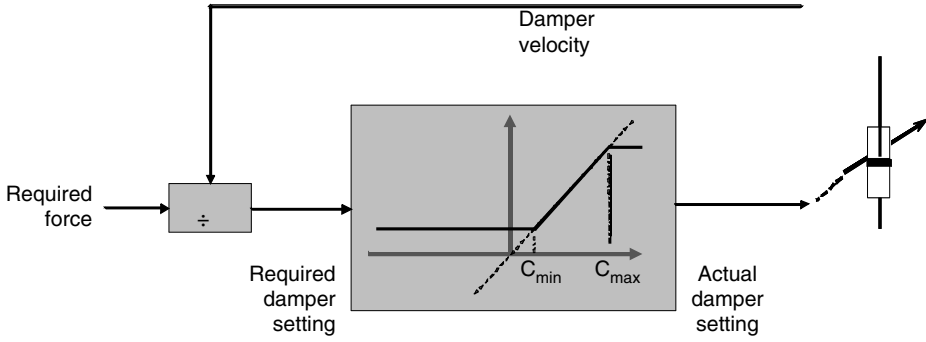


FIGURE 11.19 Controller for semi-active damper.

as closely as possible, but the actual damper setting is constrained to be between  $C_{min}$  and  $C_{max}$  — Figure 11.19 shows the control concept. To achieve operation in the upper left and lower right quadrants of the force–velocity diagram of Figure 11.3, for example, which would require a negative damper setting, the semi-active controller will simply apply  $C_{min}$ . As with full-active skyhook damping, this would potentially create large deflections in response to deterministic features; of course a semi-active damper cannot create the necessary forces, but prefiltering, as shown in Figure 11.14, is still required to ensure an effective control law.

Extra performance benefits are realised by adopting a modal approach, similar to that shown in Figure 11.17, but achievable improvements in ride quality depend upon both the minimum damper setting and the speed of response of the control action — valve switching speeds significantly less than 10 msec are needed to ensure effective implementation.

**D. EXAMPLES**

**1. Servo-Hydraulic Active Lateral Suspension**

The first full-scale demonstration of an active railway suspension was an active lateral secondary suspension using hydraulic actuators.<sup>15</sup> An actuator was fitted in parallel with the lateral secondary air suspension at each end of the vehicle, as can be seen in the left hand side of Figure 11.20. The performance obtained from a comprehensive series of tests is shown on the right, from which it can be seen that a large improvement in ride quality was obtained — a 50% reduction compared with the passive suspension.

The controller used a modal structure, shown in Figure 11.21, that provided independent control of the vehicle’s lateral and yaw suspension modes using the complementary filter technique.

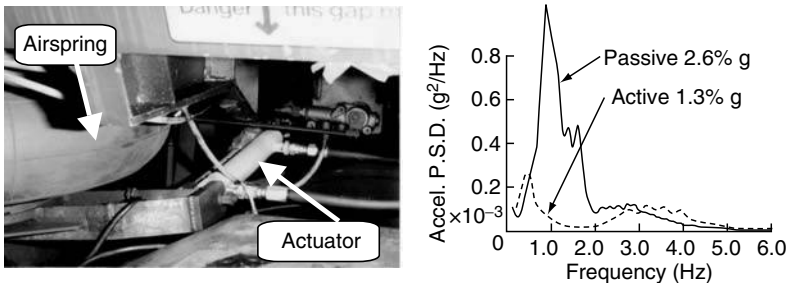


FIGURE 11.20 Servo-hydraulic actuator and experimental results for active lateral suspension.



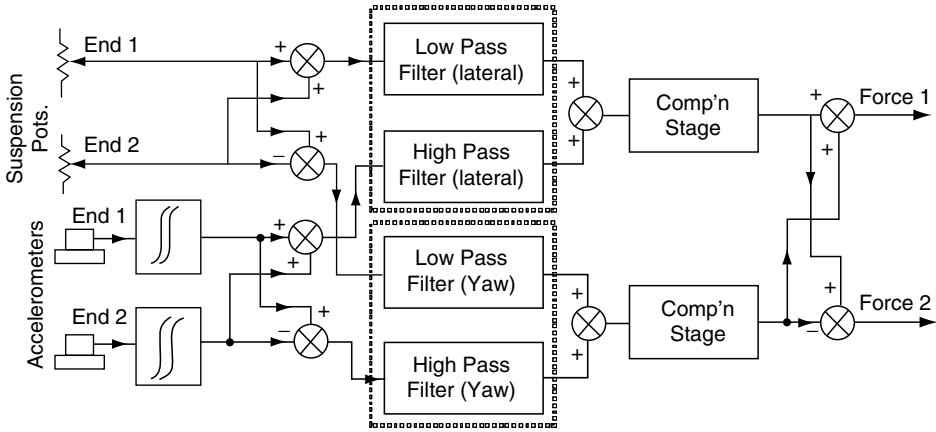


FIGURE 11.21 Controller for servo-hydraulic active lateral suspension.

Although hydraulic actuators provide a high bandwidth when used in normal applications, fast-acting force control loops (not shown in the diagram) were included to overcome the difficulty outlined above in the “actuator response” subsection, and to ensure adequate high frequency performance. Even with these inner loops, it can be seen that there is a small degradation above 3 Hz compared with the passive response.

### 2. Shinkansen/Sumitomo Active Suspension

The first commercial use of an active suspension was developed by Sumitomo for the East Japan Railway Company on their series E2-1000 and E3 Shinkansen vehicles, introduced in 2002.<sup>16</sup>

The main object of the control was the lateral vibration, i.e., closely related with riding comfort, the aim being to reduce by more than half the lateral vibration in the frequency range from 1 to 3 Hz. A pneumatic actuator system was adopted which has the advantage of easy maintenance and low cost, and is installed in parallel with a secondary suspension damper (see Figure 11.22). The damper is electronically-switched from a soft setting when active control is enabled, to the normal harder setting for passive operation.

An H-infinity controller was designed to provide robust vibration control using measurements from body-mounted accelerometers. It provides independent control of the yaw and lateral/roll

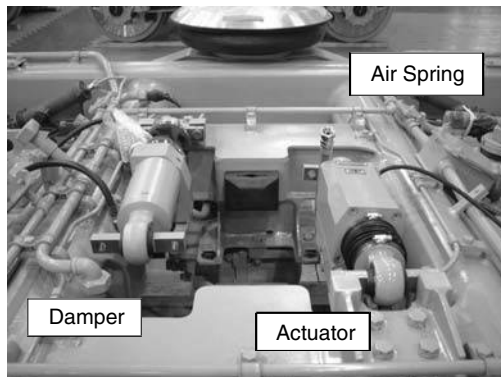


FIGURE 11.22 Actuator installation in bogie.

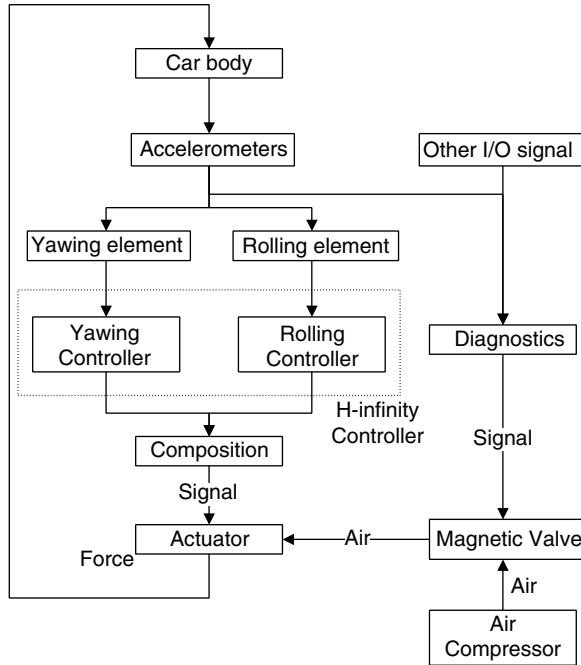


FIGURE 11.23 Overall scheme of control algorithm.

modes, with the yaw controller driving the two actuators in opposition, and the lateral/roll controller driving them in the same direction. Figure 11.23 is a diagram of the overall control scheme.

It was shown that improvements of between 5 and 9 dB in acceleration level were achievable (44–64% reduction); initially, it was a problem to achieve this kind of improvement in tunnel sections, and it was necessary to design a special controller that was switched in for use in tunnels.

## V. ACTIVE PRIMARY SUSPENSIONS

### A. CONCEPTS AND REQUIREMENTS

Although active control could be applied to vertical primary suspensions, in fact, there seems little to be gained from such an application. The main area of interest therefore relates to controlling the wheelset kinematics through the active primary suspensions. The important issue here is the trade-off between running stability (critical speed) and curving performance, which with a passive suspension is difficult, as has been outlined in earlier chapters. Various methods of passive mechanical steering to create radial alignment of the wheelsets on curves have been attempted with some improvement. However, the idea of using active control for the wheelset steering is relatively new and, therefore, mainly theoretical studies are described in this section.

There are two types of railway wheelset. As has been explained, a solid-axle wheelset consists of two coned or otherwise profiled wheels joined rigidly together by a solid-axle, which has the advantage of natural curving and self-centring, but when unconstrained exhibits a sustained oscillation in the lateral plane, often referred to as “wheelset hunting.” The structure of an independently-rotating wheelset is very similar to that of solid-axle wheelset except that two wheels on the same axle are allowed to rotate freely. The release of the rotational constraint between

the two wheels significantly reduces the longitudinal creepage on curves, but it loses the ability of natural curving and centring.

The control objectives for active primary suspensions are largely related to the wheelset configurations. For the solid-axle wheelset, the controller must produce a stabilisation effort for the kinematic mode and it must also ensure desirable performance on curves. For the independently-rotating wheelset, there is a weak instability mode which needs to be stabilised. However, more critically, a guidance control must be provided to avoid the wheelset running on flanges.

### B. CONFIGURATIONS

A number of actuation schemes are possible for implementing active steering. One of the obvious options is to apply a controlled torque to the wheelset in the yaw direction. This can be achieved via yaw actuators, as shown Figure 11.24(a), or, in practice, very likely by means of pairs of longitudinal actuators. Alternatively, actuators may be installed onto a wheelset in the lateral direction, as shown in Figure 11.24(b), but a drawback of the configuration is that the stabilisation forces also cause the ride quality on the vehicle to deteriorate. For the independently-rotating wheelset, there is a possibility of controlling the wheelset via an active torsional coupling between the two wheels, as illustrated in Figure 11.24(c). A more radical approach proposed is to remove the axle from the wheelset and to have two wheels mounted onto a wheel frame, as shown in

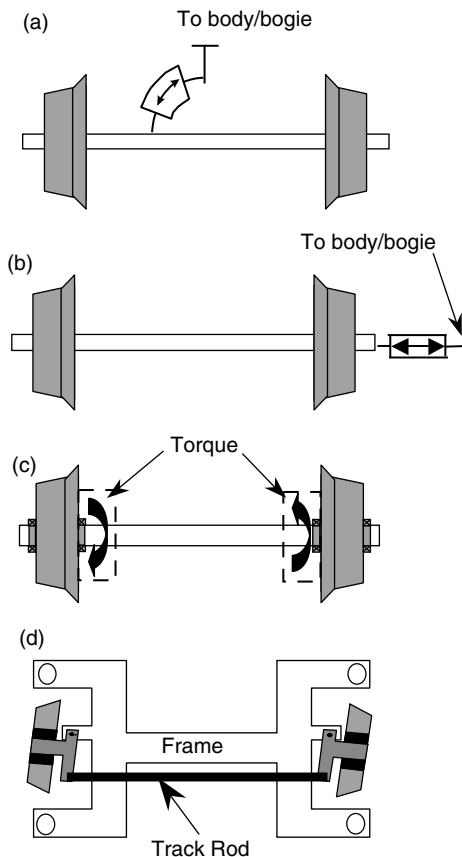


FIGURE 11.24 Actuation configurations for active steering.

Figure 11.24(d). It is then possible to apply a lateral force between the frame and the wheels to steer the wheel angle directly via a track rod, much like the steering of a car.

Similar to active secondary suspensions, the actuators for the primary suspensions can be used in combination with passive components. The passive stiffness can then be used primarily to provide the stabilisation function, whereas the actuator is used to produce an appropriate steering action on curves.

### C. CONTROL STRATEGIES

The control development for active primary suspensions ranges from separate design for stability and steering to integrated design approaches, as presented below.

#### 1. Stability Control — Solid-Axle Wheelset

The focus is on the stabilisation of the kinematic oscillation associated with the railway wheelset, but the control is ideally achieved in a way that it does not interfere with the natural curving and centring of the wheelset. One effective control technique is so-called active yaw damping, where a yaw torque from an actuator, as shown in Figure 11.24(a), is proportional to the lateral velocity of the wheelset.<sup>17</sup> The stabilising effect of the control technique can be shown using a linearised wheelset model given in Figure 11.25. It is clear from the figure that an unstable mode exists and that the inclusion of the active control loop produces positive damping to the mode. It can also be shown, using the figure, that an alternative and equally effective control method is to apply a lateral force proportional to the yaw velocity of the wheelset, a technique known as active lateral damping.<sup>17</sup> Both control techniques are difficult to realise using conventional passive components, but are relatively straightforward to implement with active means using sensors, controllers, and actuators.

#### 2. Stability Control — Independently Rotating Wheelset

An independently-rotating wheelset can still be unstable, even though the torsional constraint between the two wheels on the same axle is removed — a very effective measure that significantly reduces the longitudinal creep forces at the wheel–rail interface. The instability of an independently-rotating wheelset has been reported in Refs 17,18 and it is caused by the need of a longitudinal creep (albeit small) to rotate the wheels. However, the instability is much weaker compared to the kinematic oscillation of a solid-axle wheelset, and a high level of damping can be

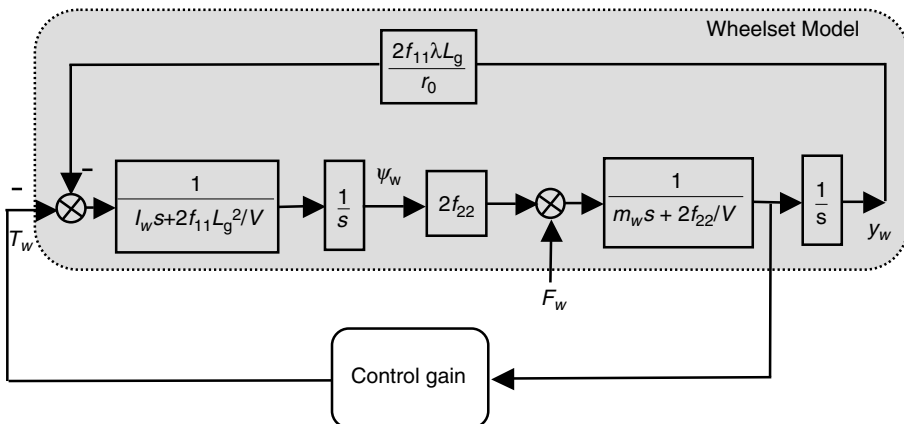


FIGURE 11.25 Active yaw damping.

attained with either a passive yaw damper or an active yaw moment control.<sup>19,20</sup> The latter is achieved by applying a yaw torque proportional to the lateral acceleration of the wheelset.

### 3. Steering Control — Solid-Axle Wheelset

When the stabilisation is obtained passively, or there are (passive) elements in the system that interfere with the natural curving action of the solid-axle wheelset, a steering action may be actively applied to provide a low bandwidth control that will eliminate, or at least reduce, the adverse effect on curves. Ideally, an active steering is required to achieve equal longitudinal creep between the wheels on the same axle (or zero force if no traction/braking) and equal creep forces in the lateral direction between all wheelsets of a vehicle. The first requirement is obviously to eliminate unnecessary wear and damage to the wheel–rail contact surfaces. The second requirement is concerned with producing and equally sharing the necessary lateral force to balance the centrifugal forces caused by the cant-deficiency.

A number of steering strategies are possible.<sup>21</sup> It can be readily shown that the perfect steering can be achieved if the angle of attack for two wheelsets (in addition to the radial angular position) can be controlled to be equal, and the bogie to be in line with the track on curves. This idea can be implemented by controlling the position of each actuator, such that the wheelset forms an appropriate yaw angle with respect to the bogie. As indicated in Equation 11.9 and Equation 11.10, the required yaw angle is determined by the track curve radius ( $R$ ), cant-deficiency (defining the necessary lateral force  $F_c$  for each wheelset), the creep coefficient ( $f_{22}$ ), and semi-wheelbase ( $l_x$ ).

$$\varphi_{\text{leading}} = \sin^{-1}\left(\frac{F_c}{2f_{22}}\right) - \sin^{-1}\left(\frac{l_x}{R}\right) \approx \frac{F_c}{2f_{22}} - \frac{l_x}{R} \quad (11.9)$$

$$\varphi_{\text{trailing}} = \sin^{-1}\left(\frac{F_c}{2f_{22}}\right) + \sin^{-1}\left(\frac{l_x}{R}\right) \approx \frac{F_c}{2f_{22}} + \frac{l_x}{R} \quad (11.10)$$

Alternately, a yaw torque can be applied such that it cancels out the effect of the longitudinal stiffness of the primary suspension which forces the wheelsets away from the pure rolling. As long as the cancellation occurs at frequencies significantly lower than that of the kinematic mode, the steering strategy will not compromise the stability. This can be realised by either measuring the relative yaw angle between the individual wheelset and the bogie and compensating for the primary forces, or by controlling the forces and/or moments of the primary suspension.<sup>22,23</sup>

### 4. Guidance Control — Independently Rotating Wheelset

For the independently-rotating wheelset, a different kind of steering action is required. The longitudinal creep is no longer an issue, which is solved by the introduction of the extra degree of freedom in the relative rotation between the two wheels. However, a guidance control becomes necessary to ensure that the wheelset will follow the track without running on flanges. To provide the necessary guidance action, it is obvious that the relative displacement between the wheelset and the track (i.e., the wheel–rail deflection) is the natural choice of feedback and the control design should then be straightforward. Sensing possibilities for the measurement vary from electromagnetic, eddy current to video imaging or optical techniques, but the potentially high cost and low reliability are the main obstacles for practical applications. Instead, angles between adjacent vehicles, as well as the vehicle body yaw rate, have been used as an indirect measurement of the track curvature and a steering action is applied to control the wheelset yaw motion.<sup>24,25</sup>

Another guidance method is to control the relative rotational speed between the two wheels.<sup>19</sup> Although there is no “hard” connection between the two wheels on an IRW axle, a control action can be formulated such that the actuator will steer the axle to achieve the zero-speed difference or a speed bias defined by the track curvature. This approach adds a damping effect between the two wheels via

the active means, however, it does not result in the stiff connection of the solid axle wheelset which forces the two wheels to be at the same angular position (rather than velocity) at all times.

### 5. Integrated Control Design

The approach to design separate controllers for the stability and steering/guidance is a pragmatic solution, and the integration of the two parts is, in general, not a problem as the two functions can be separated in the frequency domain. On the other hand, modern model based control techniques provide a more effective means to deal with the multi-objective nature of a complex control problem, although the control structures tend to be more dynamically complex.  $H_2$  optimal controls have been proposed to either maintain the natural curving of solid-axle wheelset or to provide the missing curving action for independently-rotating wheels.<sup>26</sup> Also, robust  $H_\infty$  controls have been studied to tackle the problem of parameter variations, such as the conicity and creepage deviating from their nominal values during operation.<sup>27</sup> The stability can be guaranteed in the design process and the focus is then on the other key issues such as curving performance, uncertainty, sensing, and actuation requirements.<sup>27</sup>

### 6. Assessment of Control Performance

At low speeds, the performance of active primary suspensions is measured by the reduction of creep forces and wear at the wheel–rail interface, and the focus is primarily on curved tracks where severe wear/noise may occur in passive vehicles. Many proposed active steering schemes deliver similar performances on constant curves, although the responses in transitions will be somewhat affected by different control design, which is less critical as track transitions are generally short. Compared with passive suspensions or even the radial steering (where wheelsets are mechanically forced to take a radial angle on curves), actively steered wheelsets provide significant performance improvements, as shown in Figure 11.26. The data has been obtained from a railway bogie with conventional solid-axle wheelset(s) and with much softer passive suspensions.<sup>21</sup> The creep forces produced in non-active cases would be much worse for vehicles with stiffer suspensions.  $F_x(w_1)$  is the longitudinal creep force of the leading wheelset of the bogie;  $F_x(w_2)$  is that of the trailing wheelset; and  $F_y(w_2 - w_1)$  is the difference in lateral creep forces between two wheelsets. Note that, while the longitudinal creep is undesirable except for traction purposes and should be reduced as much as possible, a certain level of the creep in the lateral direction will be inevitable in order to

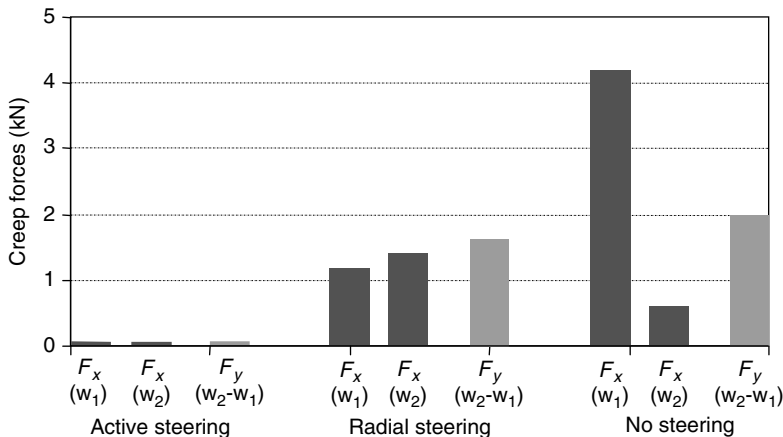


FIGURE 11.26 Steering performance comparison.

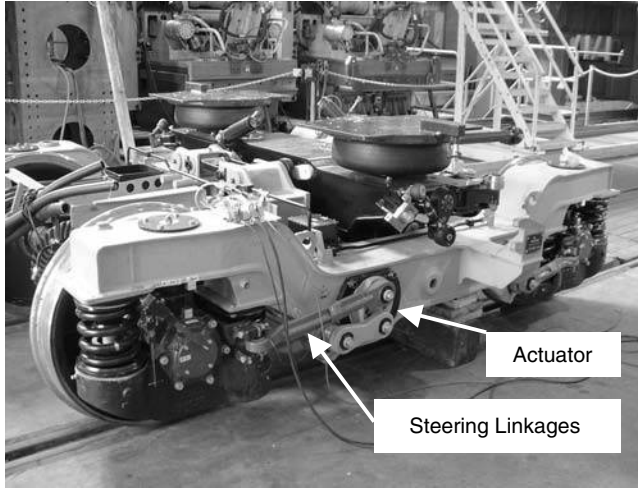


FIGURE 11.27 Actively controlled bogie.

produce a force to balance the cant-deficiency on curves. Therefore, the steering performance in the lateral direction is best assessed by examining the difference in the lateral forces, and a zero difference will indicate that track shifting forces at the two wheelsets are well balanced.

The performance of the active primary suspensions at high speeds is concerned with the running stability, and the level and speed of control effort required to control the wheelset kinematic mode and to cope with high frequency track irregularities. Those are affected more by wheelset and actuator configurations than by specific control strategies. In general, the solid-axle wheelset is much more demanding than the independently-rotating wheels as the latter arrangement allows the free rotating of the two wheels and is hence more readily adaptable to track positions.

D. EXAMPLES

This example presents an implementation and full size experiment of active control for railway wheelsets, the first example of its kind in the world.<sup>28</sup> Figure 11.27 shows a photograph of the

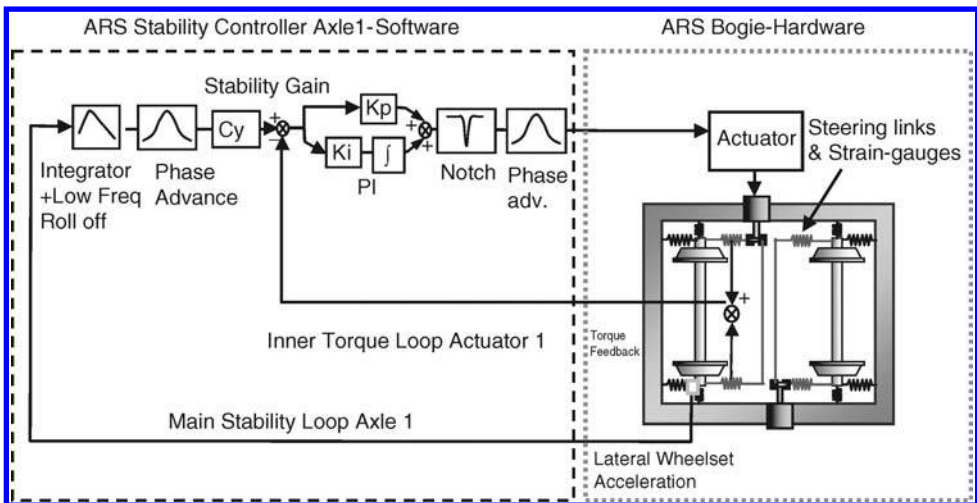
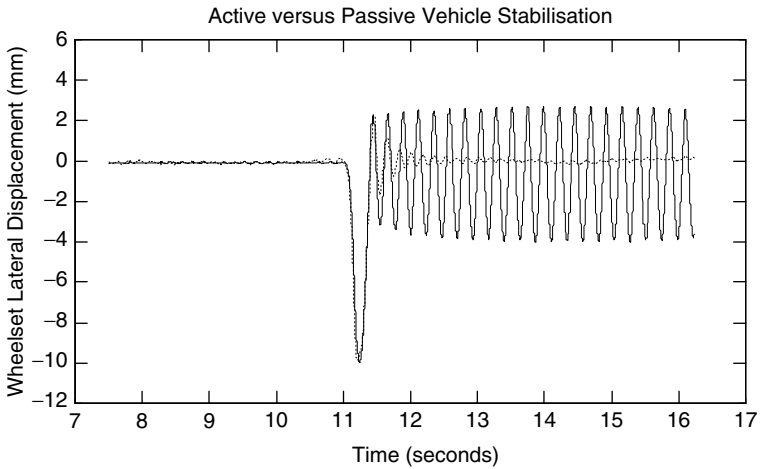


FIGURE 11.28 Stability control loop.



**FIGURE 11.29** Rig stability test result — active and passive control.

actively controlled bogie, which is a modified version of a Bombardier VT612 bogie. The bogie has a soft primary suspension and no secondary yaw dampers. In fact, the only stiffness in the longitudinal direction is due to the shear stiffness of the vertical suspensions. Removing the secondary yaw dampers offers significant advantages in terms of the vehicle's weight and comfort, however, once removed, stability and consequently high-speed operation are significantly compromised. Without active control, the modified bogie can reach a critical speed of around 90–100 km/h.

Active control is applied by means of two electrically-driven actuation mechanisms that apply independent yawing actions to each wheelset. Two a.c. servo-motors act through gearboxes, from which steering linkage mechanisms transfer the control action to the wheelsets. Control strategies for the stability and steering are designed separately,<sup>21,28</sup> but the two are brought together through an integration process to ensure there are no adverse interactions. Additional measures in the control loops are needed for reasons of practicality such as sensing and actuation, which is particularly important for the stability control due to the requirement of a high bandwidth control. Figure 11.28 shows the stability controller for one of the wheelsets, where an inner loop is added to ensure a fast dynamic response of the actuator to the torque demand from the stability control loop. The controller for the second wheelset is the same.

A fully actively controlled bogie was tested on a full size roller rig in Munich, Germany. Extensive stability tests and track file tests were performed and the controller successfully operated at speeds in excess of 300 km/h. Figure 11.29 shows results for both active control and the passive vehicle, and illustrates clearly the effectiveness of active stabilisation.

## VI. TECHNOLOGY

Technology of control concerned with the practicalities of implementation, the controller, sensors, and actuators, is an important issue. Satisfactory performances and costs are obviously essential, but, more critically, the safety and reliability requirements must be met before any applications can be considered.

### A. SENSING AND ESTIMATION TECHNIQUES

A large variety of suitable sensors is available, and the key aspects here relate to the conflict between the control requirements and practical issues such as the reliability and cost. In general, the



sensing for the active control of primary suspensions is more problematic than that for the secondary suspensions. The measurement of the wheelset movements, in particular those relative to the track, is highly desirable in order to control the wheelset effectively, but mounting effective sensors on the wheelset is extremely difficult and costly because of the harsh vibration environment.

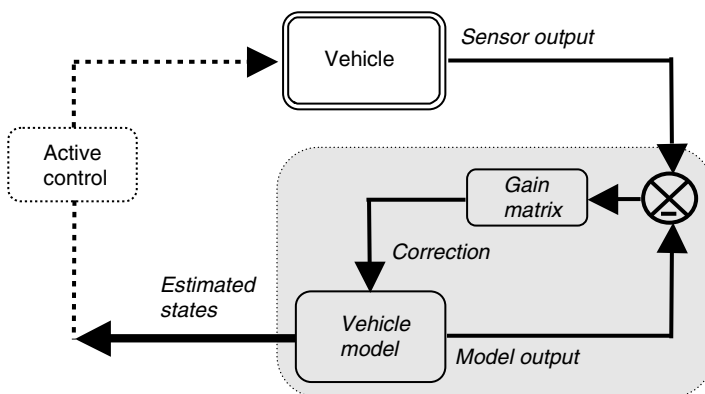
To enable a practical and cost-effective implementation of the active control schemes, model-based estimation techniques such as Kalman filters provide a very valuable alternative to the direct measurement. Figure 11.30 shows the principle of a model-based estimator. The measured output from the sensors is compared with the output from a mathematical model of the vehicle, and any deviations will produce a corrective action via the gain matrix to compensate for inaccuracies in the model and/or sensors. Estimated state variables, or some of the variables, are then used as the feedback signals for the controller as shown by the dotted line in the figure. The use of only inertial sensors on the wheelsets and bogies/body was first proposed and proved to give excellent results,<sup>29</sup> but it is also possible to remove the sensors from the wheelsets and replace them with bogie-based displacement sensors to provide the primary suspensions deflection.<sup>30</sup>

The use of more sophisticated equipment may become economically feasible in the future to measure directly some essential feedback signals and/or track features, e.g., by using track database and Global Positioning System (GPS), whereby the estimation may be simplified and its robustness improved.

**B. ACTUATORS**

The provision of high reliability actuation of sufficient performance is one of the main challenges in active suspensions. Capital cost of the total system is certainly important, but ease of installation, maintainability and maintenance cost, reliability and failure modes must all have essential inputs into the process of choosing and procuring the actuator system.

Actuator technologies, which are possible for active suspensions, are servo-hydraulic, servo-pneumatic, electro-mechanical, and electro-magnetic. Servo-hydraulic actuators themselves are compact and easy to fit, but when the power supply is included, the whole system tends to be bulky and inefficient, and there are important questions relating to maintainability. Pneumatic actuators are a possibility, particularly since the air-springs fitted to many railway vehicles can form the basic actuator, but the compressibility of air leads to inefficiency and limited controllability. Electro-mechanical actuators offer a technology with which the railway is generally familiar, and the availability of high performance servo-motors and high efficiency power electronics are favourable indicators. However, they tend to be less compact, and the reliability and life of the mechanical components needs careful consideration. Electro-magnetic actuators potentially offer an extremely



**FIGURE 11.30** Block diagram of a model-based estimator.

high reliability and high performance solution, but they tend to be very bulky and have somewhat limited travel.

### C. CONTROLLERS AND FAULT TOLERANCE

The availability of remarkable quantities of computing power means that the controller is unlikely to be a limiting factor in the implementation, although issues such as reliability, ruggedness etc. cannot be ignored.

More importantly, the issues of safety and reliability will have to be addressed satisfactorily. Any new technology must demonstrate that it can cope with component fault(s) without compromising passenger safety, and satisfy that any component fault would not lead to the system failure. On the other hand, the reliability and availability are of great importance to rail operators in order to maintain an effective operation of a rail system. Therefore, any active steering scheme must also meet the necessary standards of reliability.

Traditionally, mechanical components are used for the wheelset stabilisation and they are generally accepted as "safe." Safety is assured by having all safety-critical mechanical components designed as far as is practicable not to fail. This is achieved through a combination of conservative design, careful quality control during manufacture, and rigorous maintenance procedures during operation. However, failure modes in sensors and electronics are less definable, so it becomes necessary to reconsider the approach. Having a proven mechanical back-up for an active system, which takes over in the case of an electronic system failure, is one solution, but in the longer term this is not appropriate because it will detract from benefits. The alternative is a fault-tolerant active system based upon functional and/or analytical redundancy.

There are many fault tolerant studies carried out in many other applications, especially in the aerospace and process industry, but so far, very little is reported for railway vehicles except for a couple of studies on fault-tolerant sensing and fault detection for the provision of measurement data, and so this is an important area for development of both standards and technological solutions.

## VII. LONG TERM TRENDS

This chapter has covered the range of possible active railway vehicle suspension systems, from present-day tilting trains through to more speculative options that are, at the moment, little more than theoretical possibilities. Whether the active (secondary) suspensions are being used for improved passenger comfort, or active primary suspensions are being used to control the wheels and wheelsets, it should be clear to the reader that the use of active elements enables substantial performance improvements, improvements which are not possible with purely mechanical or passive solutions.

Active railway suspensions therefore represent an emerging and important technology, offering the railway industry a large variety of commercial and operational opportunities, although there are of course a number of major technical challenges. It is almost inevitable that the concepts will become progressively incorporated into railway vehicles, although it is less clear how quickly this will happen. However, many other industrial devices and systems which have started to replace or enhance mechanically-based products using electronic control concepts have never looked back, and so, almost certainly, the already-established tilting technology is just the starting point for active railway suspensions.

### NOMENCLATURE

$a_b$ :	body acceleration ( $\text{m/sec}^2$ )
$C$ :	coefficient of a passive damper ( $\text{Nsec/m}$ )
$C_s$ :	sky-hook damping coefficient ( $\text{Nsec/m}$ )

$F_w$ :	Lateral control force
$F_x$ :	Longitudinal creep force
$F_y$ :	Lateral creep force
$T_w$ :	Yaw control torque
$y_w$ :	Wheelset lateral displacement
$\psi_w$ :	Wheelset yaw angle
$F_a$ :	actuator force (N)
$F_c$ :	lateral force on curved track (N)
$f_{22}$ :	lateral creep coefficient (N)
$J$ :	optimisation index
$K$ :	spring constant (N/m)
$K_{CD}$ :	cant deficiency compensation factor
$K_L$ :	control gain of active leveling
$K_s$ :	control gain of suspension stiffness softening
$l_x$ :	semi wheelbase (m)
$M$ :	body mass (kg)
$q_1$ :	weighting factor for optimisation
$q_2$ :	weighting factor for optimisation
$R$ :	Curve radius (m)
$r$ :	weighting factor for optimisation
$V_{\text{active}}$ :	Vehicle speed in active case (m/s)
$V_{\text{passive}}$ :	Vehicle speed in passive case (m/s)
$x_d$ :	suspension deflection (m)
$z$ :	body displacement (m)
$z_t$ :	track displacement (m)
$\theta_{\text{active}}$ :	cant deficiency in active case
$\theta_{\text{cant}}$ :	cant angle of the track
$\theta_{\text{passive}}$ :	cant deficiency in passive case
$\theta_{\text{tilt}}$ :	tilt angle
$\varphi_{\text{leading}}$ :	required yaw angle at the leading wheelset (rads)
$\varphi_{\text{trailing}}$ :	required yaw angle at the trailing wheelset (rads)

## REFERENCES

1. Hedrick, J. K. and Wormley, D. N., Active suspensions for ground transport vehicles — a state of the art review, *Mechanics of Transportation Suspension Systems*, Vol. 15, Paul, B. et al., Eds., ASME AMD, New York, pp. 21–40, 1975.
2. Goodall, R. M. and Kortum, W., Active controls in ground transportation — a review of the state-of-the-art and future potential, *Vehicle Syst. Dyn.*, 12, 225–257, 1983.
3. Goodall, R. M., Active railway suspensions: implementation status and technological trends, *Vehicle Syst. Dyn.*, 28, 87–117, 1997.
4. O'Neill, H. R. and Wale, G. D., Semi-active suspension improves rail vehicle ride, *Comput. Control Eng. J.*, 5(4), 183–188, 1994, August.
5. Anon, *Railway applications ride comfort for passengers — measurement and evaluation*, DK CEN TC/256 WG/7, 4th draft rev, 1993.
6. Förstberg, J., Ride comfort and motion sickness in tilting trains — Human responses to motion environments in train and simulator experiments, *Doctoral thesis*, KTH, Stockholm, 2000.
7. Nishioka, Y., Tamaki, A., Isshihara, K., Okabe, N., and Torii, O., Tilt control system for railway vehicles using airsprings, *Sumitomo Search*, 56, 44–48, 1994, October.

8. Karnopp, D. C., Active and passive isolation of random vibration. Isolation of mechanical vibration impact and noise, *ASME Monograph*, Vol. 1, Snowdon, J. C. and Ungar, E. Em., Eds., p. 64–86, AMD No. 1, 1973.
9. Li, H. and Goodall, R. M., Linear and non-linear skyhook damping control laws for active railway suspension, *Control Eng. Practice*, 7(7), 843–850, 1999.
10. Mei, T. X., Li, H., and Goodall, R. M., Kalman filtering applied to actively controlled railway vehicle suspensions, *Trans. Inst. Meas. Control*, 23(3), 163–181, 2001.
11. Casini, C., Piro, G., and Mancini, G., The Italian tilting train ETR460, *Proceedings of the IMechE Conference STech 96*, paper C514/058/96, pp. 297–305, 1996.
12. Allen, D.H., Active bumpstop hold-off device, *Proceedings of the IMechE Conference Railtech 94*, paper C478/55/013, 1994.
13. Williams, R. A., Active suspensions — classical or optimal, *Vehicle Syst. Dyn.*, 114(1–3), 127–132, 1985.
14. Goodall, R.M., Pearson, J.T., and Pratt, I., Actuator technologies for active secondary suspensions, *Proceedings of the International Conference On Railway Speed-up Technology (JASM)*, Vol. 2, Japan, pp. 377–382, 1994.
15. Goodall, R.M., Williams, R.A., Lawton, A., and Harborough, P.R., Railway vehicle active suspensions in theory and practice, *Proceedings of the Seventh IAVSD Symposium*, Cambridge, UK, pp 301–316, 1981, September.
16. Tahara, M., Watanabe, K., Endo, T., Goto, O., Negoro, S., and Koizumi, S., Practical use of an active suspension system for railway vehicles, *International Symposium on Speed-up Technology for Railway and Maglev Systems 2003 — STECH'03*, JSME, 2003.
17. Goodall, R. M. and Li, H., Solid axle and independently-rotating Railway Wheelsets — a control engineering assessment of stability, *J. Vehicle Syst. Dyn.*, 33, 57–67, 2000.
18. Eickhoff, B. M., The application of independently rotating wheels to railway vehicles, *IMechE Proc.*, 205(Part F), 43–54, 1991.
19. Mei, T. X. and Goodall, R. M., Practical strategies for controlling railway wheelsets with independently rotating wheels, *J. Dyn. Syst., Meas. Control*, 125/3, 354–360, 2003, ASME.
20. Mei, T. X. and Lu, J. W., On the interaction and integration of wheelset control and traction system, *Vehicle Syst. Dyn.*, 41(Suppl.), 123–132, 2004.
21. Shen, S., Mei, T. X., Goodall, R. M., Pearson, J. T., and Himmelstein, G., A study of active steering strategies for a railway bogie, *Vehicle Syst. Dyn.*, 41(Suppl.), 282–291, 2004.
22. Shen, G. and Goodall, R. M., Active yaw relaxation for improved bogie performance, *Vehicle Syst. Dyn.*, 28, 273–289, 1997.
23. Perez, J., Busturia, J. M., Mei, T. X., and Vinolas, J., Combined active steering and traction for mechatronic bogie vehicles with independently rotating wheels, *Annu. Rev. Control*, 28(2), 207–217, 2004.
24. Hondius, H., Microprocessors harnessed to optimise radial steering, *Rail. Gaz. Int.*, 315–316, 1995, May.
25. Pederson, J., New S-trains for copenhagen, *Eur. Rail. Rev.*, 29–33, 1995, Nov.
26. Mei, T. X. and Goodall, R. M., Wheelset control strategies for a 2-Axle railway vehicle, *Vehicle Syst. Dyn.*, 33(Suppl.), 653–664, 2000.
27. Mei, T. X. and Goodall, R. M., Robust control for independently-rotating wheelsets on a railway vehicle using practical sensors, *IEEE Trans. Control Syst. Technol.*, 9(4), 599–607, 2001.
28. Pearson, J. T., Goodall, R. M., Mei, T. X., Shen, S., Kossmann, C., Polach, O., and Himmelstein, G., Design and experimental implementation of an active stability system for a high speed bogie, *Vehicle Syst. Dyn.*, 41(Suppl.), 43–52, 2004.
29. Mei, T. X. and Goodall, R. M., LQG solution for active steering of solid axle railway vehicles, *IEE Proc. — Control Theory Appl.*, 147, 111–117, 2000.
30. Pearson, J. T., Goodall, R. M., Mei, T. X., and Shen, S., *Kalman Filter Design for a High Speed Bogie Active Stability System*, UKACC Control, Bath, UK, 2004.

---

# 12 Simulation

*Oldrich Polach, Mats Berg, and Simon Iwnicki*

## CONTENTS

I.	Introduction .....	360
II.	Modelling Vehicle–Track Interaction .....	361
	A. Vehicle Models .....	361
	B. Vehicle Models — Body Components .....	363
	C. Vehicle Models — Suspension Components .....	366
	D. Track Models .....	369
	E. Wheel–Rail Contact Models .....	370
III.	Simulation Methods .....	371
	A. Multibody Systems and Equations of Motion .....	371
	B. Solution Methods .....	371
	C. Eigenvalue Analysis .....	372
	D. Stochastic Analysis .....	372
	E. Time-Stepping Integration .....	372
	F. Quasistatic Solution Method .....	373
IV.	Computer Simulation .....	373
	A. Historical Development .....	373
	B. Multibody Simulation Tools .....	376
	C. Flexible Bodies .....	378
	D. Benchmarking .....	378
V.	Dynamics in Railway Vehicle Engineering .....	379
	A. Introduction .....	379
	B. Railway Vehicle Engineering Processes .....	380
	C. Tasks and Methods in Dynamic Analysis .....	381
	D. Eigenbehaviour .....	382
	1. Eigenvalue Analysis .....	382
	2. Simulation of Eigenbehaviour .....	384
	E. Stability Analysis .....	385
	1. Introduction .....	385
	2. Linearised Stability Analysis .....	385
	3. Nonlinear Stability Analysis .....	390
	F. Run on Track with Irregularities .....	395
	1. Definition of Running Behaviour, Ride Characteristics and Comfort .....	395
	2. Ride Characteristics .....	396
	3. Ride Comfort .....	398
	G. Curving .....	401
	1. Assessment of Curving Properties .....	401
	2. Running Safety .....	403
	3. Track Loading and Wear .....	404

4. Curving Optimisation Using Self-Steering and Interconnected Wheelsets ..... 406

H. Running Dynamics under the Influence of External Loads ..... 409

    1. Influence of Crosswind..... 409

    2. Influence of Coupler Forces..... 410

    3. Interaction between Vehicle and Traction Dynamics ..... 412

VI. Conclusions ..... 415

Acknowledgments ..... 416

Nomenclature..... 416

References..... 417

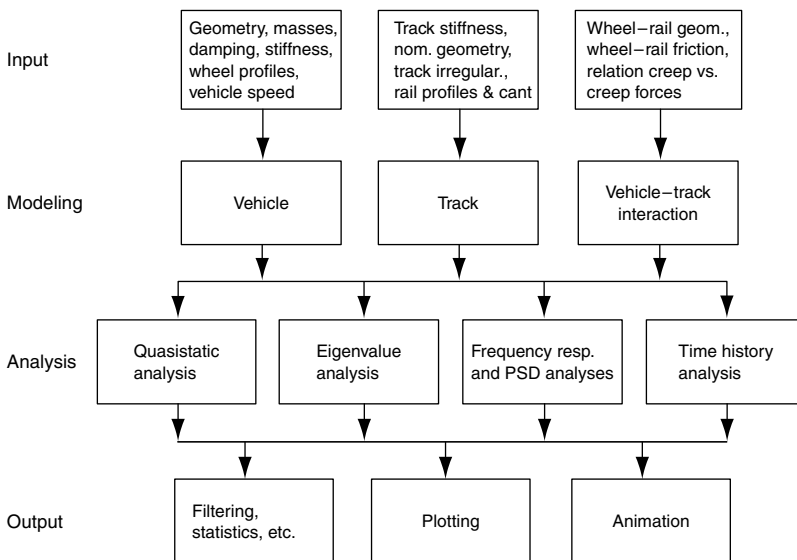
### I. INTRODUCTION

With the advent of powerful computers simulation of complex mechanical systems has become a real possibility. A computer model of a railway vehicle can be constructed and run on typical or measured track in a virtual environment, and a wide range of possible designs or parameter changes can be investigated. Outputs from the model can be set up to provide accurate predictions of the dynamic behaviour of the vehicle and its interaction with the track. Optimisation of suspension or other parts of the system can be carried out, and levels of forces and accelerations can be checked against standards to ensure safe operation.

Inputs to the model are usually made at each wheelset. Typical inputs are vertical and lateral track irregularities and deviations in gauge and cross level. These can be idealised discrete events, such as dipped joints or switches, or can be measured values from a track recording coach. Additional forces may be specified such as wind loading or from powered actuators.

Figure 12.1 shows a schematic summary of the main aspects of the simulation process.

The computer tools have developed from routines and programs used by researchers and engineers to solve specific problems, and the theoretical basis of the mathematical modelling used is now mature and reliable and programs originally written by research institutes have been developed into powerful, validated and user-friendly packages.



**FIGURE 12.1** A flow chart showing the process of computer simulation of the railway system (from Ref. 1).

This chapter covers the basic methods used in setting up a model of a typical railway vehicle and the types of analysis that can be carried out once the equations of motion have been produced. The historical development of the major simulation packages now used is also briefly covered. Typical analysis tasks such as modal analysis and stability, analysis of curving behaviour and the influence of external forces are then covered in detail with examples of typical applications. The main methods used for the assessment of the simulation results are also covered under each of these headings.

## II. MODELLING VEHICLE–TRACK INTERACTION

The potential and success of vehicle–track dynamic simulations very much depends on how well the system is mathematically modelled and fed with pertinent input data. The system modelling involves several fields of mechanics and, owing to its many complexities, is an engineering challenge. The choice of models for the system and its components depends on several aspects, mainly:

- Purpose of simulations, including requested output quantities and their accuracy
- Frequency range of interest
- Access to appropriate simulation packages
- Access to relevant model data
- Time and funding available

The first aspect above should reflect that quite coarse models may be sufficient for preliminary studies and that, for instance, the secondary suspension may not need to be that carefully modelled if we only are interested in wheel–rail forces, etc.

Regarding frequencies, the traditional frequency range of interest is 0 to 20 Hz. This low-frequency range covers the fundamental dynamics of the vehicle–track system. Obviously, this frequency range, and modelling, must be extended if we want to study noise issues (see Chapter 10). However, higher frequencies are also of interest for evaluation of vibrations (ride comfort, etc.) and wheel–rail forces (fatigue, etc.).

Today's MBD packages offer many modelling options but often the engineering guidelines on when to use the various options are limited. Depending on the experience, different engineers may therefore make different model choices. Also, the possibility to add user-defined models will affect this choice.

Some component models may be very advanced, and also accurate in themselves, but lack of appropriate input data strongly reduces their applicability. Such models often require measurements on the component in question or on similar ones.

Finally, there are also restrictions on available staff-time, calendar-time, and economical resources, which in practice, constrain the modelling possibilities.

In conclusion, the art of modelling requires engineering experience and judgment. It also requires a significant amount of relevant and reliable technical information on the vehicle–track system at hand.

In this section we try to provide some basic guidelines on modelling vehicle–track systems and their components. Below, we focus on vehicle modelling, but issues concerning track and wheel–rail modelling are also raised.

### A. VEHICLE MODELS

Railway vehicles consist of many components and for vehicle–track dynamic simulations we are interested in representing the mechanical properties of the main components. A main subdivision of the vehicle components can be made into body components and suspension components.

The dominating body components are the carbody, bogie frames and wheelsets and they essentially hold the vehicle mass (weight). Thus the inertia properties of the bodies are of primary interest. However, in many railway applications the body structural flexibility also needs to be considered. This especially holds for carbodies.

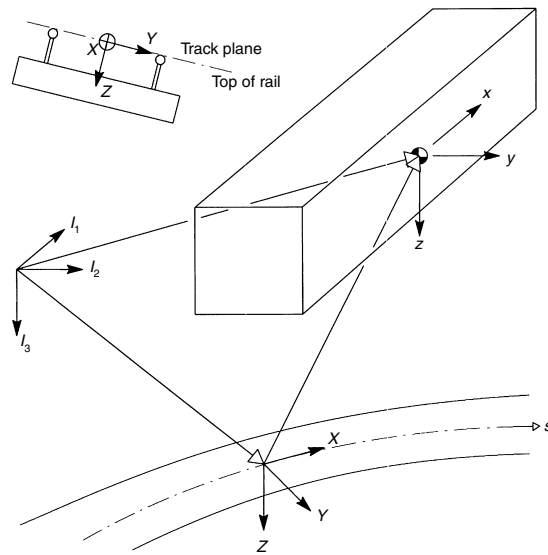
The main suspension components are various physical springs and dampers whose forces essentially are related to the displacements and velocities at the components. Traction rods, bump stops, antiroll bars, trailing arms, linkages, etc. also belong to this group of components.

This subdivision of vehicle components relates the vehicle modelling to the mechanics field of multibody dynamics (MBD) or multibody systems (MBS). This also implies that most of the vehicle degrees of freedom, or equations of motion, are assigned to the motions of the vehicle bodies.

In dealing with railway vehicles the body motions are often divided into large desirable motions, making the vehicle travel “from A to B,” and small undesirable motions. The large motions are known for a given track design geometry (curves, etc.) and a given vehicle speed history. For the large motions of the different vehicle bodies, the nominal positions of the bodies within the vehicle are also known.

For vehicles running on tangent track this subdivision of motions is simple since the large motions are given by the speed history alone. However, for curve negotiation the kinematics can be quite complicated. Figure 12.2 illustrates the principle of how the kinematics of vehicle bodies can be handled.

In Figure 12.2 an inertial or earth-fixed reference system (coordinate system)  $I_1-I_2-I_3$  is first introduced. Then, the large vehicle body motions are mainly represented by a track-following reference system,  $X-Y-Z$ , following the nominal track centreline with the speed of the vehicle. Possible track cant is also considered through the orientation of the track-following system. The small body motions can then be related to a body-following reference system,  $x-y-z$ , for instance, located at the nominal position of the body centre of gravity and with the same orientation as the system  $X-Y-Z$ . Provided the unknown body translations and rotations are small, linear kinematics is sufficient in formulating the system equations of motion.



**FIGURE 12.2** Reference systems for the kinematics of a railway vehicle body: inertial system  $I_1-I_2-I_3$ , track-following system  $X-Y-Z$ , and body-following system  $x-y-z$ .

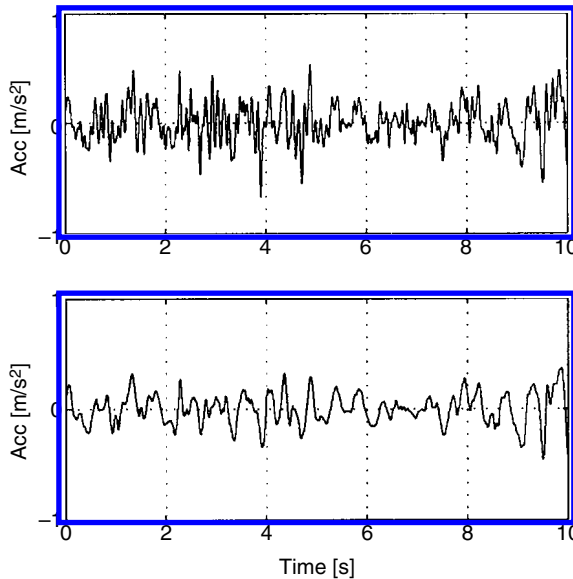


## B. VEHICLE MODELS — BODY COMPONENTS

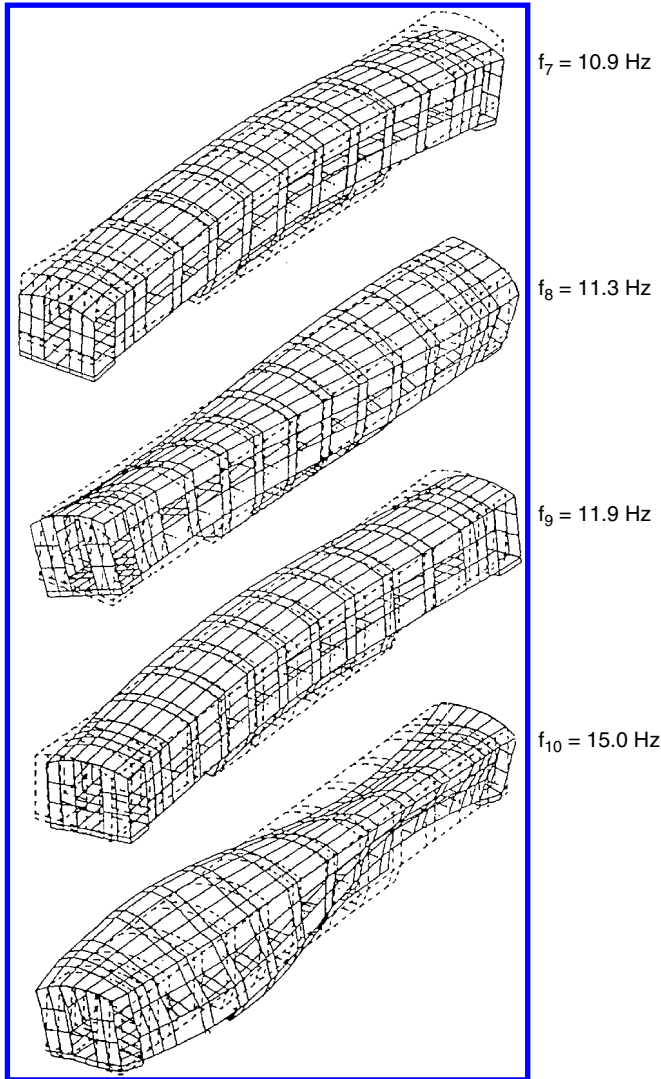
The carbody usually holds the main part of the vehicle mass. The mass properties of the carbody steel/aluminium structure are not difficult to calculate: a CAD or finite element (FE) analysis can give the mass, centre of gravity position, and mass moments of inertia of the carbody structure quite easily. However, in most applications the mass of the carbody structure is less than half the carbody mass. For a coach the additional mass is a result of the interior layout and interior/exterior equipment. For powered vehicles, especially locomotives, the carbody also carries some of the traction equipment. Often, it is quite hard to keep accurate records of all the pertinent masses and positions. Usually, the mass of this equipment is merged to the metal structure model. In many vehicle–track dynamics simulations the payload, i.e., passengers or goods, also needs to be considered. The corresponding mass might also be combined with the carbody structure model.

In most coach applications it is important to consider the carbody structural flexibility so that carbody vibrations and their negative effect on ride comfort are represented. Figure 12.3 exemplifies how carbody accelerations increase in amplitude and frequency when the carbody structural flexibility is modelled in the vehicle–track simulations. Frequencies about 10 Hz are usually prominent and for vertical accelerations this can cause significant discomfort as people are most sensitive to vertical accelerations of 8 to 10 Hz, cf. ISO 2631.<sup>2</sup> In fact, rms-values of comfort-weighted vertical accelerations may be doubled due to the carbody structural flexibility. Slender carbodies in vehicles running at fairly high speed on relatively poor track are most prone to this kind of acceleration amplification.

In railway vehicle dynamics simulations it is desirable to represent the carbody structural flexibility by only a limited number of degrees of freedom in addition to the six rigid body degrees of freedom (longitudinal, lateral, vertical, roll, pitch, and yaw motions). The most common way of finding such a representation is to use the body's eigenmodes found from eigenvalue analysis of the free body. Then, we get six rigid body modes and a number of structural modes, preferably those with the lowest eigenfrequencies. Figure 12.4 shows an example with the four lowest structural modes for a coach carbody. Note that the first structural mode has an eigenfrequency of about 10 Hz, cf. the discussion on ride discomfort above.



**FIGURE 12.3** Influence of carbody structural flexibility on carbody accelerations (example from Carlbom<sup>3</sup>). The accelerations are evaluated at the mid-carbody.



**FIGURE 12.4** Four eigenmodes and eigenfrequencies for a free carbody (example). Dashed lines indicate the undeformed carbody and finite element mesh. Modes 1 to 6 are rigid body modes, i.e.,  $f_1$  to  $f_6 = 0 \text{ Hz}$  (figure from Ref. 1).

The carbody vibrations are also promoted by a low damping of the carbody structure. However, in coaches insulation and other nonmetallic interior materials can increase the relative damping to approximately 2%. In fact, passengers can also provide the system with some additional damping, say up to 4%.<sup>4</sup> If possible the damping for the structural modes should be determined through laboratory tests. In carbody modelling, equipment and passengers may be modelled as separate bodies suspended to the carbody structure. Explicit introduction of seats and passengers into modelling also promotes evaluation of ride comfort on the seat cushion and not only on the floor.

When it comes to the bogie frame modelling we should first keep in mind that the most common bogie type worldwide, the three-piece bogie, consists of three bodies and not only of a single bogie frame body. Also, some bogies are equipped with a bolster beam to provide the bogie with significant yaw motion possibilities relative to the carbody. Such beams may be modelled as separate bodies.

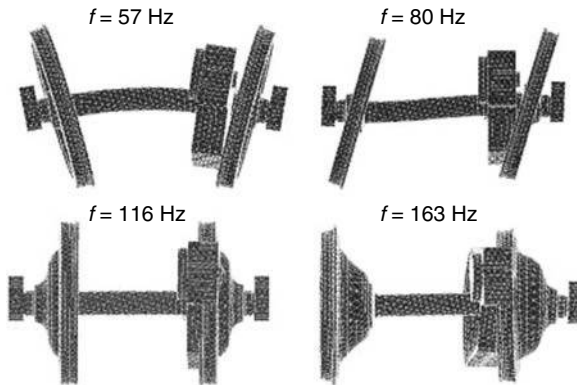
For a bogie frame, seen as a one-piece metal structure, the assumption of essentially rigid body behaviour is fair in most applications. Sometimes the torsional flexibility, about a longitudinal axis, is considered representing an improved negotiation of twisted tracks. Low-floor trams, without traditional wheel axles, need especially designed bogie frames and even if the speeds are low the frame flexibility probably needs to be considered in the bogie frame modelling for the dynamic simulations. The bogie frame flexibility can be determined through finite element analysis and may be represented in the MBD simulations by a limited set of eigenmodes of the free frame body. Such analysis can also give the mass and mass moments of inertia of the bogie frame.

Different braking equipment, for instance brake cylinders, is often attached to the bogie frame. Such equipment may be merged with the bogie frame in the modelling process. For powered bogies, the traction motors are usually mainly supported by the bogie frame. In certain applications the motors may be considered as separate bodies suspended in the bogie frame. The traction gear housing, etc. is also often partly supported by the bogie frame.

Last but not least among vehicle bodies, are the wheelsets. A plain wheelset normally consists of one solid unit with two wheels on a common wheel axle. The wheel diameter is often 0.7 to 1.0 m but both smaller and larger wheels exist. The axle diameter is usually 0.15 to 0.20 m. A typical mass for a plain wheelset is 1000 to 1500 kg. Additional mass can be introduced through brake discs, mounted on the axle or wheels, and traction gear (wheel axle gear wheel and part of the gear housing, etc.). The axle boxes, or journal-bearing boxes, also add some mass but they should not be considered in calculating the wheelset pitch moment of inertia. In fact, they are sometimes treated as separate bodies in the modelling.

Wheelsets are often modelled as rigid bodies, but wheel axle flexibility may affect the vehicle–track interaction. For instance, the axle torsion can cause ride instability and the axle bending can significantly alter the dynamic part of the wheel–rail forces.<sup>5</sup> Expressed in terms of lowest eigenfrequencies for a free wheelset, they may even be below 50 and 60 Hz in torsion and bending, respectively. More normally, these lowest frequencies are in the range of 60 to 80 Hz. For wheel diameters above, say, 1 m the wheels also contribute significantly to the wheelset structural flexibility, even for frequencies below 100 Hz.<sup>6</sup> An example of eigenmodes and eigenfrequencies for such a slender wheelset is shown in Figure 12.5. Here, the first and second bending modes in the vertical plane are shown along with two modes for which the wheels deform in an umbrella-like fashion.

For slender wheelsets the structural flexibility should be reflected in the MBD modelling. As for the bodies above, a limited number of the lowest eigenfrequencies and eigenmodes for a free



**FIGURE 12.5** Four eigenmodes and eigenfrequencies for a free powered wheelset (example). The wheel diameter is 1.3 m. The two lowest torsional modes are not shown but have the eigenfrequencies 48 and 362 Hz (figure from Chaar<sup>6</sup>).

wheelset is usually a fair representation of the wheelset flexibility. The relative damping of these modes is typically below 1%.

### C. VEHICLE MODELS — SUSPENSION COMPONENTS

Common suspension components in railway vehicles are coil springs, leaf springs, rubber springs, airsprings, friction dampers and hydraulic dampers (see Chapter 3). They play important roles in reducing bogie frame and carbody accelerations as well as dynamic wheel–rail forces. They also allow for proper curve negotiation but a too soft suspension causes problems with vehicle gauging. Bump stops and antiroll bars may therefore be introduced to mitigate such problems, the bars also reducing quasistatic lateral accelerations on the carbody floor. In the present context we may also denote traction rods as suspension components although their main task is to transfer longitudinal forces between bogie and carbody during acceleration or retardation. In addition, trailing arms and various linkages should be included here.

However, in this chapter, we restrict ourselves to passive suspension components; the reader should refer to Chapter 11 for details of active suspension. Thus, the suspension modelling here assumes that the suspension forces and moments are related to motions only at the interfaces with the connected bodies in question. However note that the static (vertical) force of an airspring can be altered through changing the air pressure without modifying the airspring height.

Static forces, or preloads, due to dead weight of bodies and payload are carried mainly by coil springs, airsprings, leaf springs, and rubber springs. For coil springs and airsprings this normally leads to compressive forces, whereas leaf springs and rubber springs are also subjected to shear forces and bending moments. In vehicle–track dynamic simulations the vehicle body motions are usually given as motions relative to the static equilibrium on tangent and horizontal track.

The static behaviour of the springs above can be determined through component measurements by slowly loading and unloading these components (for airsprings the air pressure is also changed accordingly). The vertical (axial) tangent stiffness of a single coil spring is virtually independent of the static load, whereas the airspring stiffness increases almost linearly with increasing preload. Rubber and leaf spring stiffnesses also often increase with increasing preload.

When unloading the coil spring the corresponding force–displacement graph will almost coincide with that of the loading phase; thus, the energy dissipation or hysteresis is very small. In contrast, the leaf spring undergoes significant hysteresis due to the sliding motions between the leaves. The airspring and rubber springs also experience some hysteresis due to internal friction-like mechanisms of the rubber parts. The hysteresis mentioned is due to friction rather than viscous effects since it will appear no matter how slowly the loading and unloading are realised.

For coil springs and airsprings the compressive preload gives rise to destabilising effects in the horizontal plane. An increasing preload will result in reduced horizontal (shear) stiffness for the coil springs and less increasing horizontal stiffness for the airsprings. To mitigate the destabilising effects of coil springs, two or three springs may be introduced side by side or inside each other. In the latter case the inner springs may not be activated at low preloads. This method is often used in freight wagons and produces a resulting progressive vertical stiffness.

During each simulation the suspension preloads are normally assumed constant. For the suspension springs proper static stiffnesses need to be defined, preferably based on static tests as indicated above or on appropriate calculations. Owing to curve negotiation and track irregularities, etc. the springs will deform and the suspension forces need to be changed in the simulations.

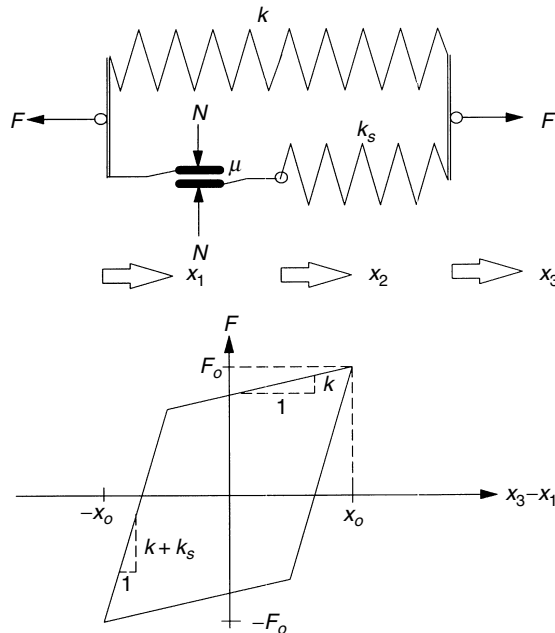
A reasonable start for suspension spring modelling is to assume models consisting of linear springs. The simplest three-dimensional model is one of three perpendicular linear springs, but the shear effect of coil springs and airsprings calls for models that consider the component height and compressive preload (see for instance [Refs. 7 and 8](#)). Nonlinear characteristics due to, for instance, clearances in coil spring sets or bump stops need special consideration.

Possible friction effects, which implies nonlinear models, should be dealt with. The stick–slip motions between leaves of leaf springs are probably the most obvious example of friction effects. Associated linkages of UIC running gear also experience combined rolling–sliding motions.<sup>9</sup> The most common friction model is that suggested by Coulomb. For a one-dimensional case such a model in series with a linear spring, which together has a linear spring in parallel, produces a resulting force–displacement as exemplified in Figure 12.6. In this way we get a parallelogram with an area of  $A = 4 \mu N(x_o - \mu N/k_s)$  corresponding to the energy dissipation per cycle. For increasing displacement amplitude  $x_o$ , this dissipation will increase whereas the stiffness  $S = F_o/x_o$  will decrease. Note that these two quantities and the graph itself are independent of the excitation frequency.

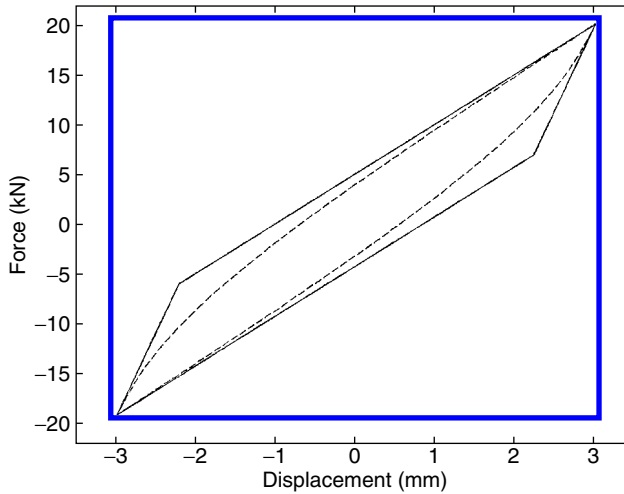
The model shown in Figure 12.6 might be sufficient in some situations but can only represent two distinct friction releases per cycle (at upper left and lower right corners). However, in particular, leaf springs have instead a successive friction release and the corners mentioned would be smoothed out. Also, rubber springs and airsprings show similar behaviour although with smaller hysteresis. A smooth friction model is therefore of interest. For instance, Refs 10 and 11 suggest such a model for leaf springs and rubber springs, respectively. Figure 12.7 shows a comparison of nonsmooth and smooth friction models. The latter is in better agreement with measurements and has only three parameters, similar to the nonsmooth model.

If we, for a given displacement amplitude, increase the frequency of excitation the force–displacement graphs of both leaf springs and coil springs show very little difference. This implies that models of the type above should be sufficient. However, for rubber springs and airsprings the viscous effects are significant thus suggesting frequency-dependent models.

The classic model of a linear spring in parallel with a linear viscous damper (dashpot) is sufficient here provided that the frequency range of interest is very limited. For larger frequency ranges this model provides a too strong frequency dependence, giving a very significant stiffness and damping at high frequencies. One common way to overcome this is to equip the model dashpot



**FIGURE 12.6** Simple friction model and corresponding force–displacement graph at harmonic displacement excitation with amplitude  $x_o$ ;  $x_o > \mu N/k_s$  (from Ref. 1).



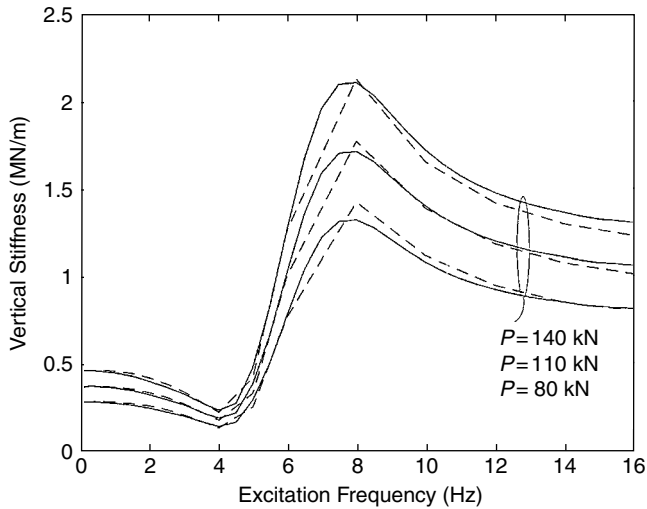
**FIGURE 12.7** Force–displacement graphs at harmonic displacement excitation for the model in Figure 12.6 and a model with smooth friction.

with a series spring, i.e., using a model as in Figure 12.6 but replacing the friction element with a dashpot. In this way the maximum stiffness will be  $k + k_s$ , as the excitation frequency  $\omega$  tends to infinity, and the maximum energy loss per cycle  $\pi k_s (x_o)^2 / 2$ , at  $\omega = k_s / c$  with  $c$  being the damping rate. However, for frequencies above  $4k_s / c$  the model's energy loss per cycle is already halved, a property not found from tests on rubber. To resolve this more parallel sets of dashpot with series springs may be added,<sup>12</sup> but then require a number of additional input parameters.

A proper choice of viscous model, cf. above, can describe reasonably well the frequency dependence of a rubber spring. However, the amplitude dependence, mainly due to friction, should also be represented as indicated above. In Ref. 11, a simple rubber spring model, although still representing both frequency and amplitude dependence, is suggested. In fact, comparing the stiffness of a large displacement and low frequency case with that of a small displacement and higher frequency case, reveals that the stiffness increase for the latter case can very well be greater owing to the smaller amplitude than to the higher frequency.

For the vertical dynamics of airsprings with auxiliary air volume, the models above are usually still not sufficient. The viscous effects are quadratic rather than linear with respect to velocity, introducing another nonlinearity into the airspring model. Moreover, an inertia effect will arise due to the very high accelerations of the air in the surge pipe connecting the two volumes. For small airspring displacements and no so-called orifice damping of the pipe, this inertia effect is very pronounced and thus needs to be modelled. Examples of such airspring models are Refs 8 and 13. Figure 12.8 shows an example of the very strong frequency dependence of an airspring system with the air bag subjected to a small displacement amplitude. This dependence is shown for different preloads and both simulation and measurement results are included. The figure first shows a decrease in stiffness followed by a significant stiffness increase at approximately 8 Hz. This effect can be devastating for vertical ride comfort.<sup>14</sup>

Hydraulic dampers may be modelled as linear or piecewise linear viscous dampers. For dampers with a high damping ratio (i.e., damping force changes very rapidly with damper velocity) it may be important to consider the inherent stiffness of the damper assembly as well, including rubber end bushings and internal structural and oil stiffnesses. This is particularly true for yaw dampers, which usually have a high damping ratio. Dampers normally have a force-limiting blow-off level, which also must be considered.



**FIGURE 12.8** Vertical airspring stiffness as a function of excitation frequency for the three preloads  $P$  and displacement amplitude 2 mm. Simulation (solid line) and measurement (dashed line) (from Berg<sup>8</sup>).

Flexibility in damper brackets, etc. must be considered as well, in particular, for dampers with high damping rates. It is as important as the internal oil stiffness. It is often desirable that brackets and attachment points are very stiff, in order not to reduce damper efficiency. Hydraulic dampers usually have some frictional effects as well.

Friction damping through plane surfaces sliding against each other during general two-dimensional motion is not simple to model, but one example is shown in Ref. 15. Moreover, the friction characteristics vary with the status of wear, humidity, and possible lubrication.

Except for the airspring vertical dynamics, the suspension models do not usually include inertia effects. However, the mass of each suspension component should be split 50/50 to the connected bodies. This is necessary in order not to underestimate the total vehicle mass in the modelling.

A final remark on the suspension modelling is the need for linearisation of nonlinear models when linear types of analyses are to be carried out.

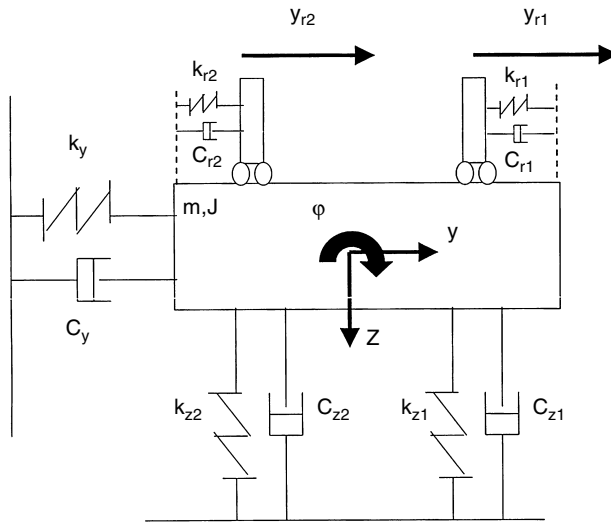
## D. TRACK MODELS

Track flexibility, nominal geometry (track layout) as well as track irregularities must be modelled and described.

Figure 12.9 shows a simple moving track model for lateral, vertical, and roll flexibilities of the track. Such a simple model may be sufficient for analysis of the vehicle's interaction with the track. However, this track model is, in most cases, too simple for analysing track behaviour in more detail. Stiffnesses and damping may have linear or nonlinear characteristics. Actual numerical data reflect the track construction, i.e., type of rails, fastenings, rail pads, sleepers, and ballast, as well as vehicle axle loads. The track flexibility will, above all, influence track forces, but under certain circumstances lateral dynamic stability will also be affected.

The nominal track geometry (layout) is defined by circular curve radii and lengths, lengths and types of transition curves, track cants, etc.

Track irregularities are normally given as lateral and vertical deviations of the track centreline (from nominal geometry), and by deviations in track cant and track gauge. It is important that track irregularities are representative for a longer section of track, and that they are also representative of



**FIGURE 12.9** Example of simple model for track flexibility — cross section. One piece of track is assumed to follow each wheelset (moving track) (from Claesson<sup>16</sup>).

the worst condition to be considered. A statistical analysis of track data is, therefore, very often needed.

## E. WHEEL–RAIL CONTACT MODELS

The characteristics of the wheel–rail contact is of crucial importance for the dynamic interaction between vehicle and track. Therefore, since the 1960s much research has been devoted to the fundamental issues of wheel–rail contact (see also Chapter 4 and Chapter 5). The wheel–rail contact is, as a rule, very nonlinear. This is true for the so-called contact functions but also for creep forces as functions of creepage and spin.

The stiff wheel–rail contact, including Hertz’s contact stiffness, but also the very high damping rates resulting from creep forces, provide high eigenvalues and therefore make it necessary to run time-domain simulations (by means of numerical integration) with short time steps (in the order of 0.1 to 1 msec). Thus, these simulations are quite computer intensive and time-consuming. Advanced wear simulations, cf. Ref. 17, are certainly computer intensive.

A wheel–rail model consists principally of a wheel–rail geometry module (including results from contact functions), a creep/spin calculation procedure and a creep force generator. The theories were described in Chapter 4.

The contact functions are usually calculated in a preprocessor program. Input parameters are wheel and rail geometry, rail inclination, track gauge, and wheelset flangeback spacing. Normally these quantities, apart from track gauge, are assumed to be constant over the whole simulated section; however, there may also be variations along the line, for instance, the possibility of *one-point contact* and *two-point contact* has to be considered. Simulations must automatically change from one state to another, depending on the actual wheel–rail conditions for each individual wheel.

Since wheel and rail parameters (profile shapes, gauge, rail inclination) change over time and over different track sections, many combinations and possibilities have to be included in the model and also have to be systematically investigated. A special case is negotiation of switches and crossings.

Quasistatic and dynamic performance are also dependent on wheel–rail friction, normally varying from approximately 0.6 (very dry rails) to 0.1 or 0.2 (wet or oily rails). This must also be



considered in the models. In particular, the worst case scenario is usually dynamic stability and wheel–rail wear under dry rail conditions. See also Chapter 5.

### III. SIMULATION METHODS

#### A. MULTIBODY SYSTEMS AND EQUATIONS OF MOTION

The first stage in setting up a computer model is to prepare a set of mathematical equations that represent the vehicle–track system. These are called the equations of motion and are usually second order differential equations that can be combined into a set of matrices. The equations of motion can be prepared automatically by the computer package, a user interface requiring the vehicle parameters to be described in graphical form or by entering sets of coordinates and other data describing all the important aspects of the bodies and suspension components.

The vehicle is represented by a network of bodies connected to each other by flexible, massless elements. This is called a MBS (see [Section II](#)) and the complexity of the system can be varied to suit the simulation and the results required. Each of the rigid bodies can be considered to have a maximum of six degrees of freedom, three translational and three rotational. Physical constraints may mean that not all of these movements are possible and the system can be simplified accordingly. Application of the constraint equations results in a set of equations of motion which are ordinary differential equations (ODE) or linear algebraic equations (LAE) and ODEs, depending on how the constraint equations are used.

Masses and moments of inertia for all bodies need to be specified. Points on the bodies, or nodes, are defined as connection locations, and dimensions are specified for these. Springs, dampers, links, joints, friction surfaces, or wheel–rail contact elements can be selected from a library and connected between any of the nodes. All of these interconnections may include nonlinearities such as occur with rubber or airsprung elements or as in damper blow-off valves or bump stop contact. Nonlinearities also occur at the wheel–rail contact point due to creep and flange contact. Owing to the presence of these nonlinearities the full equations of motion cannot normally be solved analytically. It is sometimes possible to linearise the equations of motion but otherwise a numerical method must be used to integrate the equations at small time intervals over the simulation period, the results at each point being used to predict the behaviour of the system at the next time step. These methods are discussed in the following sections.

The bodies in an MBS system are usually rigid but can be flexible with given modal stiffness and damping properties if required by the simulation. Modal properties can be measured or simulated using finite element (FE) tools. Some dynamic simulation packages include the possibility to model flexible bodies as beams or combinations of beams, and linking to or incorporation of FE models is also now becoming possible.

Inputs to the model are usually made at each wheelset. These can be idealised discrete events representing, for example, dipped joints or switches, or can be measured values from a real section of track taken from a track recording vehicle. Most railway administrations use track recording coaches running on the network and collecting track data at regular intervals. Additional forces, such as wind loading or powered actuators, may be specified (see Chapter 11). Depending on the purpose of the simulation a wide range of outputs, for example, displacements, accelerations, forces, at any point can be extracted.

#### B. SOLUTION METHODS

The method of solving the equations of motion will depend on the inputs to the model and the required output. Four of the main methods widely used in simulating vehicle dynamic behaviour are explained in the following sections and practical examples of these are given in Section V.

### C. EIGENVALUE ANALYSIS

All systems with mass and stiffness can vibrate and these vibrations occur most naturally at certain frequencies called modal frequencies and in certain patterns called mode shapes. If the equations of motion are linear (or can be linearised for certain equilibrium positions of the bodies or amplitudes of vibration) then an eigenvalue analysis can be carried out to determine the modal frequencies and mode shapes. This is also known as modal analysis and it may be useful for a vehicle designer or operator to have a knowledge of these modes to allow unwanted vibrations to be reduced.

Owing to the creep forces that are present at the wheel–rail interface the railway vehicle can be subject to self-exciting oscillations. These will occur as the vehicle is moving along the track, and their characteristics will depend on its forward velocity. Below a certain speed oscillations set up by a small disturbance will tend to die away or remain of very small amplitude, their energy being dissipated in the damping present. Above this speed, however, a similar disturbance will cause oscillations that grow until limited by the wheel flanges striking the rails.

This unstable behaviour is called hunting and can result in damage or derailment. The speed is called the critical speed for the vehicle. An eigenvalue analysis can be used to give information about the stability of vibrations at each mode and this is useful in establishing the critical speed of a vehicle above which hunting instability will occur.

In practice, the complex eigenvalues can easily be calculated for the particular equations of motion (for example using the Matlab “eig” routine) and the vehicle speed gradually increased until the real part is not negative but equal to zero. This corresponds to the critical speed of the vehicle.

Caution should be exercised in using this method for establishing the critical speed of a vehicle as it relies on the linearised equations of motion. In particular, the wheel–rail interface is highly nonlinear even over small displacements and the linearised conicity parameter, which must be used for a linear analysis, cannot fully represent the situation. The interested reader should refer to [Section V.E.2](#) and to [Refs 18, 19, 89 and 91](#).

An alternative method is to carry out a time-domain simulation using the full nonlinear equations of motion and observing the rate at which the oscillations of vehicle motion (especially wheelset lateral displacement) die down after a disturbance.

### D. STOCHASTIC ANALYSIS

The vehicle–track system is modelled using a MBS and linearised as previously described.

Inputs to the model can be made at each wheelset. These inputs correspond to vertical and lateral track irregularities and deviations in gauge and cross level. For a stochastic analysis these inputs usually take the form of a describing function, that is, a function of the spectral density of the amplitude of the particular parameter against frequency of the irregularity. The parameters of the describing function can be obtained from the measured track data.

The stochastic analysis method is useful for evaluating the general lateral or vertical behaviour of a vehicle to a particular type of track. Responses of various vehicles can then be compared or the effect of minor changes to the design of a vehicle evaluated. The equations of motion can only be linear and this type of solution method cannot be used to show response to discrete inputs such as bad track joints.

This is an ideal method to use when vehicle ride is of interest. The frequency spectra of the output can easily be analysed against available recommended levels. Caution must be exercised when using this technique for lateral motion and lateral ride as the nonlinearities in this case are severe.

### E. TIME-STEPPING INTEGRATION

The most powerful method available for simulating the dynamic behaviour of a vehicle is to solve the equations of motion fully at each of a series of very small time steps. All the nonlinearities of

the system can be considered and the equations updated accordingly at each time step. A wide range of numerical methods is available for this type of simulation, for example, the Runge Kutta techniques are widely used. The size of each step must be small enough to ensure that the solution does not become unstable but the penalty of using a smaller time step is, of course, a longer simulation time. Some solvers use a varying time step which is automatically adjusted to suit the current state of the simulation. The fastest simulators are now able to solve the equations of motion at faster than real time even for simulations involving complex suspensions or multiple vehicles. This type of simulation is sometimes called “dynamic curving” as it is most suitable when a vehicle negotiates a series of curves of differing radii or a curve of changing radius.

Each term of the matrix equation of motion is set out separately in a subroutine and at each time step every term is evaluated. This means that the stiffness or damping coefficient of each suspension element in the model can be calculated with reference to the relevant displacements or velocities. Typical suspension nonlinearities encountered are bump stops, or multistage suspension elements such as dampers.

At each time step the equations of motion are set up and all the suspension nonlinearities are evaluated. The creepages and creep forces between the wheels and rails are evaluated and the resulting accelerations at each body for each degree of freedom are calculated. The displacements and velocities are calculated through the integration routine and stored, the elapsed time is increased, and the complete calculation step repeated. The whole process keeps stepping until the preset maximum time or distance is reached.

This solution method is very powerful because of the ease with which it can accommodate nonlinearities in the equations of motion.

## F. QUASISTATIC SOLUTION METHOD

This is a special case in the overall study of the behaviour of a rail vehicle. When a vehicle is negotiating a curve of constant radius at a constant speed the wheelsets and bogies will take up a certain fixed attitude after a short period of transient motion (provided the vehicle is stable at this speed). The aim of this analysis is to predict the steady-state attitude and the resulting wheel–rail and suspension forces. The method is known as “steady-state curving.” No track irregularities or other varying inputs can be considered when using this technique.

## IV. COMPUTER SIMULATION

Using modern computer packages it is possible to carry out realistic simulation of the dynamic behaviour of railway vehicles. The theoretical basis of the mathematical modelling used is now mature and reliable, and programs often originally written by research institutes have been developed into powerful, validated, and user-friendly packages.

### A. HISTORICAL DEVELOPMENT

In analysis of the contact between a railway wheel and a rail the first step is to establish the location and the size and shape of the contact patch (or patches). As the cross-sectional profiles of the wheel and the rail can be quite complex shapes most computer simulation packages have a preprocessor, which puts the wheel and rail profiles together for a given wheelset and track and establishes where the contact will occur. A description of the cross-sectional profiles is prepared from the designs or measured using a device such as the widely used “Miniprof.” More details of these methods are given in Chapter 13.

The classical theory of contact was developed by Hertz<sup>20</sup> in 1882 when he was a 24-year-old research assistant at the University of Berlin. He demonstrated that the contact area between two nonconformal bodies of revolution would be elliptical and established a method for calculating

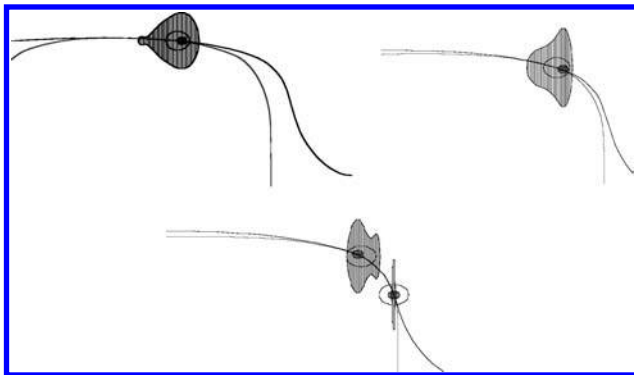
the semiaxes of the ellipse and the pressure distribution within the contact patch. The Hertz theory is strictly restricted to frictionless surfaces and perfectly elastic solids but it still provides a valuable starting point for most contact problems and is included in most computer programs which deal with wheel–rail contact.

Some software packages use Hertz theory to establish elliptical contact patches around the contact point. The normal load on the contact point is required and the calculation may be iterative to allow the correct load distribution between the contact points to be found. In tread contact the radii of curvature are only changing slowly with position and the contact patch is often close to elliptical in shape. However, if the radii are changing sharply or the contact is very conformal the contact patch may be quite nonelliptical and the Hertz method does not produce good results. For example, in 1985 Knothe and Hung<sup>21</sup> set out a numerical method for calculating the tangential stresses for nonelliptical contact.

Multi-Hertzian methods split the contact patch into strips with Hertz contact being calculated for each strip. Some iteration may be required to establish the correct normal load distribution across the whole contact patch if the contact is being treated as a constraint. Pascal and Sauvage<sup>22</sup> developed a method using an equivalent ellipse which first calculated the multi-Hertzian contact and then replaced it with a single ellipse which gives equivalent forces. In the methods developed by Kik and Piotrowski<sup>23</sup> an approximate one-step method is used and some results for an S1002 wheel and a UIC60 rail are shown in Figure 12.10. In the semi-Hertzian methods developed by Ayasse et al.<sup>24</sup> the contact is treated as Hertzian for the longitudinal curvatures (along the rail) and non-Hertzian for the lateral curvatures (across the rail).

In 1916 Carter<sup>25</sup> introduced the concept of creepage or microslip between the wheel and the rail and the corresponding creep force which was generated. A fuller treatment of creep forces was reported by Vermeulen and Johnson<sup>26</sup> in 1964, and in 1967 Kalker<sup>27</sup> provided a full solution for the general three-dimensional case with arbitrary creepage and spin. Haines and Ollerton divided the contact area into strips parallel to the direction of rolling and predicted areas of adhesion and slip in the contact patch.<sup>28</sup>

Heuristic methods for predicting creep force were developed initially by Vermeulen and Johnson<sup>26</sup> based on a cubic equation for creep force as it nears saturation. The method developed by Shen et al.<sup>29</sup> is widely used and is very fast but results are approximate and become less accurate when spin (relative rotation about the normal axis) is high. Polach<sup>30</sup> has developed a method which works well at high levels of creepage and spin and also includes the falling value of the coefficient of friction as slip velocity increases.



**FIGURE 12.10** Nonelliptical contact patches for profile combination S1002/UIC 60 (from Kik and Piotrowski<sup>23</sup>).

Kalker produced various computer algorithms for the calculation of creep forces according to his theory. The program CONTACT based on Kalker's "exact" theory, which includes non-Hertzian contact, is relatively slow and not practical for use at every time step in a numerical integration. Table interpolation routines are available, such as USETAB,<sup>31</sup> which interpolate between values of creep force precalculated by CONTACT. FASTSIM<sup>32</sup> is based on Kalker's "simplified theory," which assumes an elliptical contact patch with a flexible layer between the two rigid bodies.

Experimental measurements taken by Brickle<sup>33</sup> with a twin disk rig and Illingworth<sup>34</sup> with a roller rig validated Kalker's results provided that the surfaces in contact were free from contamination. Hobbs<sup>35</sup> proposed that the Kalker coefficients be factored by 0.6 to take normal levels of contamination into account.

When surface roughness effects are significant the creepage creep force relationship is affected and Bucher et al.<sup>36</sup> have proposed methods for dealing with this situation. Knothe, together with Gross-Thebing<sup>37</sup> has also looked at the effect of rapidly varying creepages which were not previously considered. When the contact characteristics have been calculated by non-Hertzian methods (see above) a modified version of FASTSIM has to be used as explained by Kik and Piotrowski.<sup>23</sup>

Once an understanding of the wheel-rail contact had been established the way was open for a full analysis of the dynamic behaviour of a railway vehicle. This was encouraged by a prize offered by the Office for Research and Experiments (ORE) of the Union of International Railways (UIC) in 1950 for the best analysis of the stability of a two-axle railway vehicle. The prize winners were Pospel, Beaufeoy, and Matsudaira.<sup>38</sup>

All of the prize winners had used a linear analysis of the problem but de Pater<sup>39</sup> formulated the hunting behaviour as a nonlinear problem. Van Bommel<sup>40</sup> later published nonlinear equations for a two-axle vehicle using wheel and rail profiles and a creep force-creepage law measured by Müller for the ORE committee.

With the advent of analogue and then digital computers it became possible for these equations to be solved for real problems and for nonlinearities to be included more easily.

Wickens<sup>41</sup> at British Rail led a group who improved the analysis to include an understanding of the wheelset as a feedback mechanism and applied first analogue and then digital computer methods to the problem. This resulted in a new high-speed two-axle freight vehicle with much improved stability and provided a basis for the work on the advanced passenger train and for the development of the software tools used today in the U.K.

Müller<sup>42</sup> carried out one of the first simulations with analogue computers of a bogie vehicle running into a curve. He<sup>43</sup> also recognised the importance of the inclusion of nonlinear wheel profiles and included tabulated geometric data which was measured for a combination of worn wheels and rails. This was taken further by Cooperrider et al.<sup>44</sup> who produced an algorithm for combining measured wheel and rail profiles to produce the nonlinear parameters such as rolling radius difference and contact angles as the wheelset moved laterally across the track.

The early programs tended to split up the types of behaviour to simplify the task of calculation. Programs for calculation of the vehicle attitude and the forces developed during steady-state curving were one example of this. An eigenvalue analysis of the linear or linearised equations of motion was used to give information about the natural frequencies and mode shapes of the oscillations and to predict limits to stable running (but note the discussion in Section III.C and Section V.E). Time-stepping integration could be carried out if the systems were nonlinear but it was usually necessary to separate the vertical behaviour (involving bounce and pitch of the bodies) and lateral behaviour (yaw, roll, sway). Longitudinal dynamics, which is more important when dealing with long freight trains, was also handled separately. As computing power developed it became less necessary to handle each aspect of the vehicle behaviour separately, and powerful numerical methods were applied in the time domain unless a frequency domain output was required.

True<sup>18</sup> applied the theory of nonlinear dynamics to the behaviour of a railway vehicle and showed that failure to consider the nonlinearity of the wheel–rail contact can lead to an inaccurate estimate of the critical speed of the vehicle. Recently, Schupp<sup>45</sup> has described a method using numerical bifurcation analysis to simulate the nonlinear behaviour of railway vehicles. These methods have resulted in the software PATH which has been used together with SIMPACK.

## B. MULTIBODY SIMULATION TOOLS

The early packages used text-based interfaces where vehicle parameters were listed in a particular order or using key words to provide the input to the simulation. User-friendly graphical interfaces were added and packages developed to allow engineers to test the effects of making changes to any part of the system and to animate the output; for example, ADAMS/Rail where the user works with a vehicle model through a graphical user interface which allows interaction with the model in the same way as a computer-aided design system (Figure 12.11 and Figure 12.12).

A large number of computer codes have been developed by railway organisations to assist in the design of suspensions and the optimisation of track and vehicles. Some of these have been combined into general purpose packages and some examples of those currently in widespread use are given here although this is not a comprehensive list and the aim is to illustrate the variety of programs that are in use today.

Multibody dynamics theory is used to develop the equations of motion for the system and these are processed by a solver which produces the results of interest. A review of the main multibody simulation packages and the methods that they used was carried out by Schielen.<sup>46</sup>

One of the early complete packages, MEDYNA<sup>47</sup> (Mehrkörper-Dynamik) was developed at the German Aerospace Research organisation DLR together with MAN and the Technical University of Berlin. MEDYNA was based on a MBS with small rigid body motions relative to a global reference frame which allowed large motions. The linearised kinematic equations of motion for each body are formulated with respect to the global reference frame. SIMPACK was developed later by the same team at DLR and as it was intended for road vehicles and other systems as well as rail vehicles it allowed nonlinear kinematics from the start. The equations of motion are formulated in terms of relative coordinates and can be generated symbolically and numerically in an implicit and explicit form. The kinematics of elastic bodies were developed to allow stress stiffening effects to be taken into account.

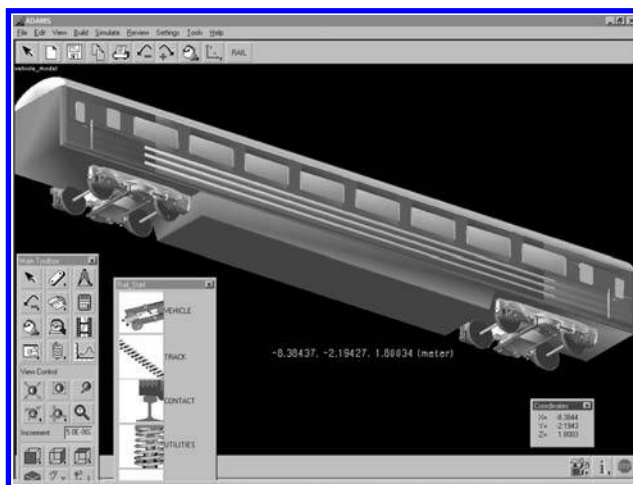
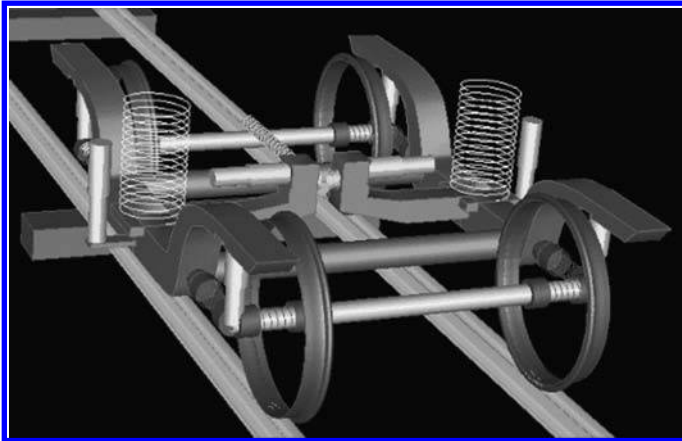


FIGURE 12.11 ADAMS/Rail graphical user interface with vehicle model.



**FIGURE 12.12** ADAMS/Rail parameterised bogie model.

ADAMS is one of the most popular dynamic simulation codes worldwide and in 1995 it entered the railway vehicle simulation market as ADAMS/Rail,<sup>48</sup> initially by including wheel–rail contact methods developed by NedTrain and later by licensing the wheel–rail contact elements from MEDYNA.

In the U.S.A. the Association of American Railroads (AAR) funded the development of a program to simulate the behaviour of a railway vehicle negotiating a curve. This was developed into the general purpose simulation package NUCARS<sup>49</sup> (New and Untried Car Analytic Regime Simulation). NUCARS has been used to improve the dynamic behaviour of the three-piece freight bogie.

The French National Transport Research Institute INRETS developed a multibody simulation code VOCO (Voiture en Courbe) in 1987 with a reference frame that permitted simulation on long curves. The inclusion of friction damping was possible from the beginning because the code was initially used to simulate the Y25 bogie. A commercial version of this named VOCOLIN<sup>50</sup> in 1991 allowed simulation of the wheel–rail contact with a multi-Hertzian approach. A second approach was made by Gimenez et al.<sup>51</sup> and this was incorporated into the code VOCODYM.

In the U.K. British Rail Research developed a number of computer programs to analyse different aspects of railway vehicle dynamic behaviour as has been mentioned above. These have been brought together into one coherent package VAMPIRE<sup>52</sup> which is now supported by AEA Technology Rail.

Real Time VAMPIRE<sup>55</sup> is an interesting development which has been made possible by advances in computer technology and the very high simulation speed possible with VAMPIRE. Track data can be fed into VAMPIRE directly from a track recording coach and the behaviour of vehicles on that section of the track predicted in real time. All the typical vehicle dynamics outputs can be generated for the specific vehicle model that has been loaded. This allows derailment risk, passenger comfort, track force, or any other normal output to be produced and these values are available immediately for track engineers to use, for example, in prioritising maintenance.

In Sweden, modelling of railway vehicles using computers started at ASEA in 1971. Initially, the analysis was carried out in the frequency domain with linear models and then, in 1973, a nonlinear, time-stepping integration program was developed. This program separated lateral and vertical modes and was used in the development of the X15 high-speed test train and the Rc4 locomotive in 1975.<sup>53</sup> In 1992 the development of a new three-dimensional calculation program started and software development was transferred to a new company called DESolver. This new three-dimensional, general computer code, together with all earlier pre- and postprograms became in 1993 the new railway vehicle analysis tool called GENSYs.<sup>54</sup>



### C. FLEXIBLE BODIES

Most of the multibody simulation packages allow the inclusion of flexible bodies, such as car bodies, where bending or torsion can be significant. Information about each mode for the flexible body must be included and this is often taken from a finite element analysis of the body carried out outside the multibody simulation package.

Track models used in multibody simulations are generally relatively simple one- or two-layer rigid body models and the track support conditions are usually constant along the rail or move with the vehicle. More details of track models are given in Chapter 6. Although this is likely to be adequate for many types of simulation it may not allow full representation of all the dynamic response modes of the rail. Knothe et al.<sup>56</sup> have reviewed developments in this area. More detailed track models have been developed, for example, by Corus Rail Technologies as part of their “Track System Model”,<sup>57</sup> shown in Figure 12.13.

An alternative to separate vehicle and track models is a finite element model of the track integrated into the vehicle dynamics software. This should enable the track response to be accurately captured and subsequently improve the accuracy of the vehicle response, which in turn should improve the accuracy of wheel–rail forces. This approach is currently being developed by some of the vehicle software packages. FE flexible models of a bridge structure have been successfully developed using ANSYS and incorporated into ADAMS/Rail. SIMPACK includes flexible bodies using an inbuilt flexible element called SIMBEAM, and also using an FE interface called FEMBS. FEMBS uses the standard input data (SID) file present in the majority of FE packages to create a reduced modal representation of the complete FEM. With the continuing increase in processing speed it may be that the MBS type program will be superseded by software based on finite element analysis which allows easy treatment of flexible bodies and locates stress and fatigue problems within bodies.

### D. BENCHMARKING

Owing to the high level of complexity of the software codes developed for simulation of railway vehicle dynamics there is a high level of interest in comparing the results of the different codes for certain test cases.

An early benchmark was proposed by ERRI<sup>58</sup> based on a passenger coach. In the exercise initiated at the Herbertov workshop on “MBS applications to problems in vehicle system dynamics”,<sup>59</sup> in 1990 and reported on by Kortüm and Sharp<sup>60</sup> the computer codes that were able to

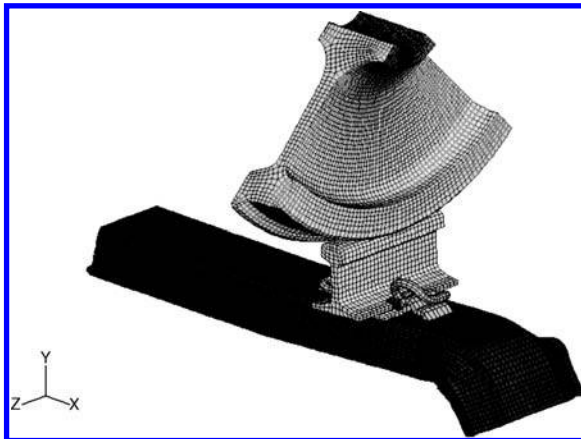
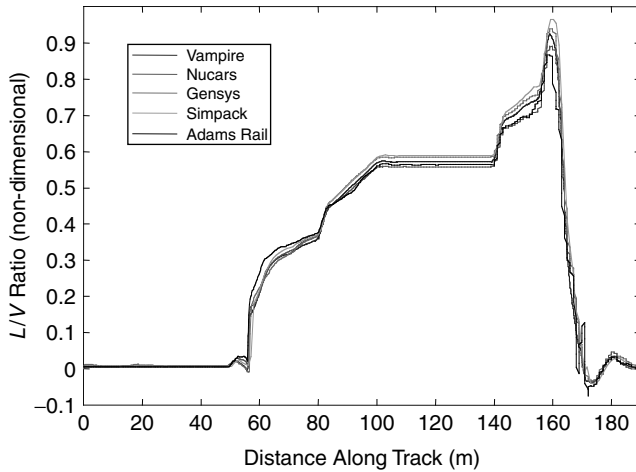


FIGURE 12.13 The CRT track system model.<sup>57</sup>





**FIGURE 12.14**  $L/V$  ratio from the Manchester benchmarks<sup>61</sup> (vehicle 1, track case 1, 4.4 m/sec).

handle wheel–rail contact were asked to simulate a single wheelset and a bogie. The wheelset benchmark was proposed by Pascal and participants were required to calculate the lateral deflection of the specified wheelset in response to a lateral force of 20 kN and to find the level of lateral force at which the wheelset would derail. In the bogie benchmark defined by Kik and Pascal, participants were required to predict the behaviour of the bogie in a vehicle running on straight and curved track at several speeds. Not all codes participated fully in the exercise but some interesting results were shown.

In the Manchester benchmarks published in 1999,<sup>61</sup> two simple vehicles and four matching track cases were defined to allow comparison of the capabilities of computer simulation packages to model the dynamic behaviour of railway vehicles. One of the aims of this benchmark was to try to encourage railway organisations to accept simulations carried out using any reliable computer simulation package and not to insist on one particular tool. Simulations were carried out with five of the major packages (VAMPIRE, GENSY, SIMPACK, ADAMS/Rail-MEDYNA, and NUCARS) and the results and statements of methods were presented.

A number of outputs were requested for each of the track cases, each of which was designed to test a particular potential vehicle problem. One of the most useful indicators of derailment potential is the lateral/vertical ( $L/V$  or  $Y/Q$ ) force ratio at each wheel. In a curve it is usually the outer wheel where derailment takes place and Figure 12.14 shows the  $L/V$  ratio for the outer wheel on the first wheelset for one of the benchmark vehicles. The peak value occurs at a dip designed to test vehicle suspension and shows that all five packages give good agreement on the nearness to derailment of this vehicle.

## V. DYNAMICS IN RAILWAY VEHICLE ENGINEERING

### A. INTRODUCTION

Computer analyses and simulations of vehicle dynamics constitute an integral part of engineering processes during the development and design of new and modified railway vehicles. Virtual prototyping computer tools have made considerable progress in recent years. The simulation of railway vehicle system dynamics can be coupled with other simulations from structural mechanics, aerodynamics, controls, electrotechnics, etc.<sup>62</sup> to a virtual development process. Modern simulation packages provide powerful and important analysis and design tools that are well-suited to the concurrent engineering process demands in the railway industry.<sup>63</sup>

In spite of the significant progress made in the virtual development process, very few publications exist that describe the dynamics methodology used in the railway industry (e.g., Refs 64 and 65). Although a wide range of methods is applied for dynamic analyses in industrial practice, some of the methods are typical to the development of railway vehicles. These typical vehicle dynamic methods applied during the engineering process are presented here. A description is supplied concerning the aim of the methods, important influencing parameters, typical input parameters, and output values. The methods are illustrated with examples from industrial application. In Section V.B the railway vehicle development engineering process and the role of vehicle dynamics in the same is described. Section V.C provides an overview concerning the aim and methods of dynamic calculations applied during the engineering process. Section V.D presents calculations of eigenbehaviour, and Section V.E shows examples of linearised and nonlinear stability analyses. Section V.F deals with the simulation of running behaviour on track with irregularities, and Section V.G with analysis of curving behaviour. In Section V.H, various special running conditions under influence of external loads are presented.

### B. RAILWAY VEHICLE ENGINEERING PROCESSES

The phases of the engineering process of railway vehicles as a function of time and knowledge of input parameters are illustrated in Figure 12.15. The engineering process of railway vehicle dynamics commences with a feasibility study and concept analysis. The major emphasis of engineering work is concentrated on the optimisation and verification design phases. The process continues with the realisation phase and test and qualification phases. Warranty and field support issues also benefit from vehicle dynamics analyses if required. Service experience provides feedback concerning the vehicle dynamics behaviour predicted during engineering.

Railway vehicle dynamics simulations are already applied at an early stage for concept investigation and feasibility study. The models applied here are normally based on other similar vehicles and most of the input parameters are estimated. Input parameter uncertainty is covered by parameter variation and sensitivity tests.

The main part of the dynamic calculations takes place during vehicle design. In the optimisation (preliminary) design phase, the topology of the vehicle is specified, but the input parameters are based on a first estimation only. The parameters of suspension and other coupling elements have to be optimised based on the target values given by the standards and vehicle specification. All limit and target values have to be checked and design changes introduced if necessary. Output values required for other engineering activities, such as movements and loads, are calculated to support the engineering process.

In the verification (detail) design phase, the vehicle topology is frozen. The vehicle is modelled in detail in order to acquire an exact vehicle behaviour. The parameters of elements are specified, and if possible, also verified by measurements on prototype elements. Parasitic effects such as

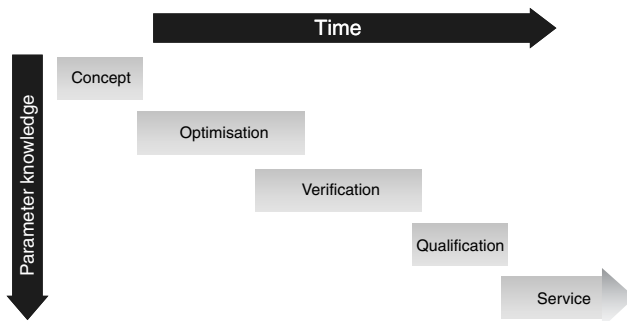


FIGURE 12.15 Engineering process of railway vehicle design.

vertical and lateral parasitic stiffness of traction rods, antiroll bars, or other elements have to be taken into consideration. The structural elasticity also has to be taken into account if an influence can be expected on the stiffness parameters, damper performance, or vehicle behaviour.

The aim of the verification phase is to prove conformity with the limiting values stated in the norms and regulations, as well as the vehicle specification. The verification calculation should provide proof of whether the vehicle will pass the acceptance test. The verification calculation report is often used to support the acceptance test, but can also be used to reduce the extent of type tests. Final movement and load collective calculations are applied to finalise the analyses of other specialists. A failure analysis and a sensitivity analysis can be provided as part of the verification calculations.

Later on, the vehicle acceptance and qualification tests are supported by dynamic calculations if necessary. During the test the predicted and measured vehicle performance can be compared and the modelling improved to gain experience for future projects.

Following the vehicle acceptance tests, the vehicles go to service. Although the engineering process is completed, a feedback from service promotes the improvement of modelling and simulation in future projects. Requirements for improvement arising during service can lead to warranty issues, which in turn, may necessitate vehicle dynamics simulation support.

### C. TASKS AND METHODS IN DYNAMIC ANALYSIS

Vehicle dynamics calculations utilised during railway vehicle engineering can be divided into the following categories depending on the activities supported by the analyses:

- Fulfillment of customer specification
- Vehicle acceptance tests
- Risk assessment
- Support of other specialists during the design process

Typical dynamic analyses worked out during vehicle engineering are listed in [Figure 12.16](#) together with the aim of the analysis according to categories mentioned above. A possible method to formulate each of the analyses is also provided.

Computer simulations that examine running safety, track loading, ride characteristics, and ride comfort of the vehicle are the primary items of investigation. These are the characteristics that are tested during the on-track test as part of the vehicle acceptance procedure. Fulfillment of the vehicle performance according to the customer specification and required standards<sup>66,67,68,69</sup> has to be proven and demonstrated (see also Chapter 13). Simulations to investigate and optimise these issues are also part of the risk assessment and risk management process.

In addition, other important factors such as stresses and cumulative load distribution (see e.g., [Refs 70–73](#)), spring displacements, vehicle gauging (see Chapter 7), influence of external loads acting on the vehicle or train configuration (see [Section V.H](#)) have to be taken into consideration. The simulations also fulfill an in-house requirement to support other specialists during the design process, and deliver the necessary input for analysis of structure mechanics, specification of components, and negotiations with suppliers.

Further issues concerning dynamic analyses related to risk assessment or problem solving could be the questions of wheel out-of-roundness,<sup>74</sup> rail corrugations,<sup>75</sup> rolling contact fatigue of wheels,<sup>76,77</sup> and rails<sup>78</sup>. The importance of these issues is increasing due to the tendency for higher speeds, increasing axle loads, and smaller wheel diameters increasing the risk of damage on wheels and rails.

There are no standards and recommendations available which specify how dynamic simulations of railway vehicles should be carried out. Specialists or companies have usually developed their own methods for vehicle modelling, analysis procedure, and result assessment. A guideline is

Task	Type of analysis	Calculation method				
		Eigenvalue analysis	Quasi-static analysis	Simulation		
				Straight track	Full curve	Curve transition
Internal need	Eigenbehaviour	X		X		
Risk assessment Customer's specification Vehicle acceptance Safety Support of other specialists	Carbody sway in curve		X		X	
	Safety against derailment		X		X	X
	Track shift force			X	X	
	Stability	X		X		
	Ride characteristics			X	X	
	Track loading		X		X	
	Ride comfort			X	X	X
	Wear		X		X	
	Gauging			X	X	X
	Influence of external loads			X	X	X
	Load collectives			X	X	X

FIGURE 12.16 Typical dynamic analyses and calculation methods applied in railway vehicle engineering.

provided by the standards and specifications for measurements and acceptance tests of vehicles (see Chapter 13). The simulations must prove whether the vehicle performance required by the specification can be achieved and that the limit values will be fulfilled. However, on the one hand, it is difficult to apply the procedure for measurements in simulations, and on the other hand, some additional possibilities in simulations are not available in measurements. Therefore, the simulations constitute a combination of the conditions specified in the standards for the measurements, together with a feasible and most efficient calculation method.

The simulation results can be basically structured in two possible ways:

- By the vehicle dynamics performance specified in the standard or specifications
- By the methods used in the calculations

The advantage of the first structure is a clear and easy comparison with requirements, while the second structure provides a simpler and better overview of the methods, input values, and other conditions used in simulations. In our overview of the typical methods used in the engineering process we follow the second structure based on calculation methodology.

## D. EIGENBEHAVIOUR

### 1. Eigenvalue Analysis

The eigenvalue analysis allows the vehicle–track model to be examined and first information to be obtained concerning suspension properties. This should be done at the commencement of the dynamic calculations as a first optimisation step. The application of linearised calculations for vehicles with strong nonlinearities (e.g., friction damping between carbody and bogie frame) is incorrect. A simulation of eigenbehaviour as described in Section V.D.2 would provide a suitable alternative.

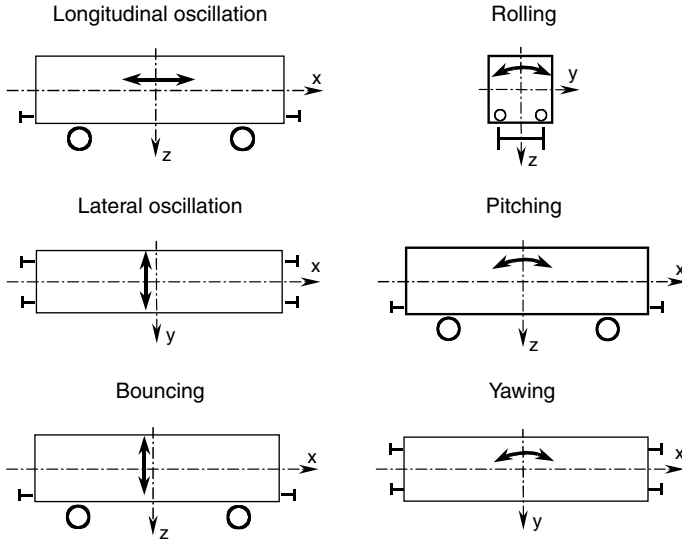


FIGURE 12.17 Basic carbody modes.

The eigenbehaviour should be investigated for tare (empty) as well as for full load. The calculation can be done for any speed, but as kinematic (speed-dependent) oscillations should be excluded, it is advisable to apply very low speed, e.g., 1 m/sec (zero speed is usually not allowed in simulation tools).

The eigenfrequencies, eigendamping, and eigenmodes can be used for testing the vehicle model. An asymmetry of the modes or implausible values can provide an indication concerning incorrect or missing parameters. All eigenvalues should possess sufficient damping (as a minimum 5% of critical damping is recommended). For correct modelling, all eigendamping properties of rubber elements, as well as other parasitic damping have to be considered in the model parameters.

The modes and nomenclature used for carbody eigenbehaviour are shown in Figure 12.17 and Figure 12.18. The sway mode, as combined lateral movement and rotation about longitudinal axis, is present in two forms with different heights of the rotation centre (see Figure 12.19): lower sway mode and upper sway mode.

Table 12.1 shows, as an example, the carbody eigenmodes of the passenger coach from the Manchester benchmark.<sup>61</sup>

The bouncing eigenfrequency of the carbody should be approximately 1 Hz. The frequency of lower sway mode should be higher than 0.5 Hz, otherwise there is a risk of motion sickness arising.

The eigendamping of the carbody modes should be between 15 and 30%. This target value is usually difficult to achieve for all carbody eigenmodes. For lower sway mode values slightly below 15% are also acceptable. The damping of the carbody yaw and upper sway modes is usually higher than 30%, but this is not critical from a running behaviour perspective.

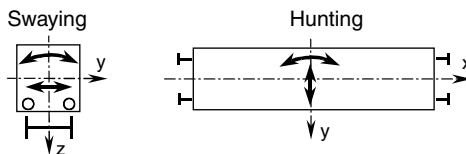


FIGURE 12.18 Combined carbody modes.

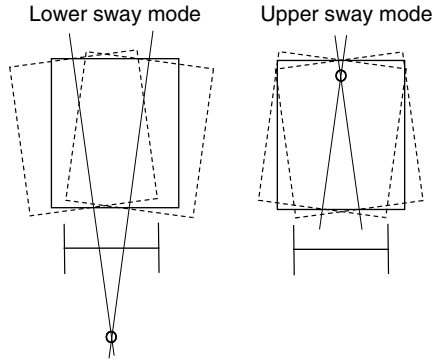


FIGURE 12.19 Two modes of carbody sway movement.

**TABLE 12.1**  
Eigenfrequency and Eigendamping of a Passenger Railway Coach from the Manchester Benchmark<sup>61</sup>

Mode	$f$ (Hz)	$D$ (%)
Bouncing	1.07	13.4
Pitching	1.28	16.0
Lower sway	0.58	21.1
Upper sway	1.10	45.2
Yawing	0.73	53.7

**2. Simulation of Eigenbehaviour**

The eigenbehaviour can also be assessed by using time integration. This is mainly suitable for vehicles with strong nonlinearities, where a linearisation would be an unacceptable simplification.

The investigation can be carried out in a similar manner to measurements, simulating a reaction of the vehicle on a single track excitation or on nonzero initial conditions. Figure 12.20 shows the simulated eigenbehaviour of a metro vehicle in comparison with measurement. The carbody sway oscillation was measured by a so-called wedge test. At the beginning of the test, wedges having

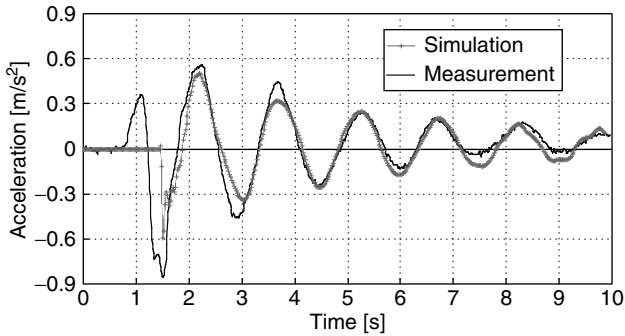


FIGURE 12.20 Comparison of lower sway mode oscillation observed during wedge test and detected by simulation.

a height of 15 to 25 mm are positioned on the top of one rail in front of each wheel at one side of the vehicle. The vehicle starts to move slowly, the wheels roll over the wedges and excite carbody swaying. The sum of the measured lateral accelerations on the ceiling and on the floor of the carbody provides the signal characterising the lower sway mode shown in [Figure 12.20](#). Pitch, bounce, and yaw modes can be measured and investigated by numerical simulations in a similar manner.

## E. STABILITY ANALYSIS

### 1. Introduction

The self-excited wheelset and bogie movement becomes unstable from a certain speed, as described in Section III.C. The frequency of the sinusoidal wheelset and bogie motion is related to the wheel–rail contact geometry. If only the wheelsets and bogies or running gears are involved in the unstable movement, we refer to bogie instability or bogie hunting. If the wheel–rail contact conditions lead to low frequency, the carbody sometimes starts to move together with the bogies. In this case we refer to carbody instability or carbody hunting.

Experimental stability investigations are exacting due to an important influence of nonlinear contact geometry of wheel and rail as well as general contact conditions and friction coefficient. In the vehicle type tests, vehicles are checked concerning the fulfillment of the stability up to the specified test speed, but the margin of the critical speed at which the vehicle becomes unstable is normally not investigated. It is therefore seldom possible to prove the margin of stability up to the critical speed and to compare the stability calculation with measurements during vehicle design.

Owing to the wide range of input conditions and limited experimental experience, the stability simulations provide the most diversified type of analysis. Methods such as nonlinear and linearised calculations can be applied in various versions. The methods can vary depending on whether they are based on the theory of mechanics or on the experience from measurements. The following sections present several feasible methods as they are or may be used for stability analysis in the industrial application. The description is divided into sections concerning linearised analysis using eigenvalue calculation and nonlinear time-domain simulations. In Section V.E.2 the linearisation of the contact between wheelset and track is described. Following this, bogie stability and carbody stability investigations are discussed individually. In Section V.E.3 the nonlinear stability analysis concentrates on the investigations of bogie stability. For nonlinear simulations regarding carbody stability, methods and limits as described in Section V.F can be used.

### 2. Linearised Stability Analysis

For linearised stability calculations the contact of wheelset and track has to be linearised in a different way to the other coupling elements. The parameters of a linear wheel–rail contact model depend on the lateral amplitude of wheelset movement used for linearisation. In parallel, they depend on the gauge and on the shape of the wheel and rail profiles. The characteristic parameter is the equivalent conicity (see Chapter 4).

The most widely used quasilinear wheel–rail contact model<sup>79,80</sup> uses three parameters:

1. Equivalent conicity  $\lambda$
2. Contact angle (contact slope) parameter  $\varepsilon$
3. Roll parameter  $\sigma$

The parameters of this linearised contact model can be determined by:

- Linearisation of nonlinear profiles for specified linearisation amplitude.
- Variation of equivalent conicity, setting the other parameters as its function.

Despite the fact that only the first method is exact, the second method is often used to vary the conicity in a wider range without the nonlinear profiles having to represent those conditions.

The contact angle parameter as a function of conicity was investigated by ERRI (European Railway Research Institute, formerly ORE) for common combinations of wheel and rail profiles. In the report ORE B 176<sup>81</sup> the following functions are specified.

For the combination of wheel profile S 1002 with rail UIC 60, inclination 1:40 (e.g., Germany, Austria, and Switzerland)

$$\varepsilon = 85\lambda \tag{12.1}$$

For the combination of wheel profile P8 with rail BS 113A, inclination 1:20 (e.g. Great Britain, France, Italy)

$$\varepsilon = 50(\lambda - 0.05) \tag{12.2}$$

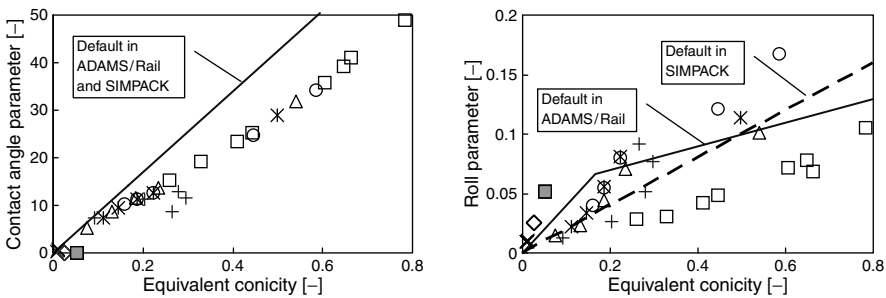
Equation 12.1 is the most widely used option for calculation of the contact angle parameter as a function of conicity. For roll parameter, the following function is used in ADAMS/Rail

$$\begin{aligned} \sigma &= 0.4\lambda & \text{for } \lambda \leq 0.166 \\ \sigma &= 0.05 + 0.1\lambda & \text{for } \lambda > 0.166 \end{aligned} \tag{12.3}$$

In SIMPACK it is the function:

$$\sigma = 0.2\lambda \tag{12.4}$$

Figure 12.21 shows the default values of the contact angle parameter  $\varepsilon$  and roll parameter  $\sigma$  used as default values in simulation tools (Equations 12.1, 12.3 and 12.4) compared with nonlinear profile combinations linearised for an amplitude of 3 mm. As can be seen from the comparison, the calculation of contact angle and roll parameters only provides very imprecise information, and can lead to results which are significantly different to the linearisation of nonlinear wheel–rail geometry.

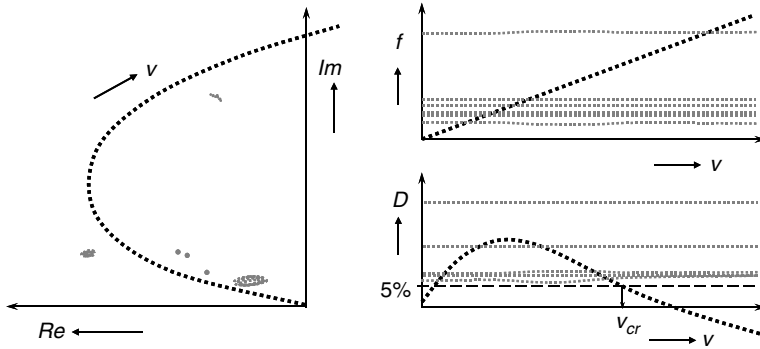


Wheel/rail combinations:

- |                          |                                  |
|--------------------------|----------------------------------|
| △ S1002 / UIC60 1:40     | × S1002 / UIC60 1:20             |
| × S1002 / UIC54 1:40     | ○ S1002 / UIC54E 1:40            |
| ◇ Cone 1:40 / UIC60 1:20 | ■ Cone1:20 / 115RE 1:40          |
| + P8 / UIC60 1:20        | □ S1002 / UIC54E 1:40 worn crown |

**FIGURE 12.21** Contact angle parameter and roll parameter as functions of equivalent conicity as applied for linear calculations in simulation tools and calculated applying the quasilinearisation for a lateral wheelset amplitude of 3 mm on different combinations of wheel and rail profiles and gauge values between 1430 and 1438 mm.



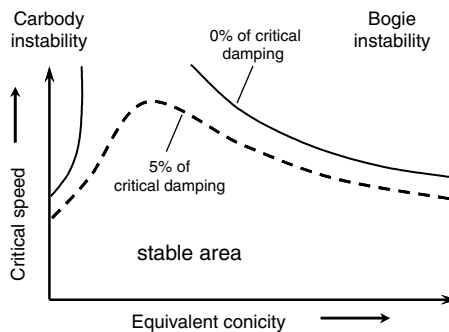


**FIGURE 12.22** Root locus curve in complex plane (left) and frequency  $f$  and relative damping  $D$  as function of speed (right).

A root locus curve represents a set of eigenvalue calculations with speed as the vehicle parameter. In this way the influence of speed on the eigenbehaviour of the vehicle can be observed (see Figure 12.22). The eigenmode is unstable if the real part of the eigenvalue is positive. For engineering applications, the speed at which the vehicle achieves 5% of critical damping of the lowest damped mode can be used for setting the critical speed.

A set of root locus calculations for a varying parameter is described as a stability map or stability diagram. A typical parameter used as an independent variable is the equivalent conicity (see Figure 12.23). In general, two areas with low critical speed exist. For high values of equivalent conicity, the limiting mode is the bogie hunting. In the low conicity range the limiting mode is a combined movement of carbody and bogies — carbody instability.<sup>73,80</sup> The vehicle is unstable for all speeds higher than the critical speed of bogies, whereas the carbody instability sometime disappears with increasing speed. Although the form of the stability diagram can differ within a wide range dependent on the input parameters, there is typically a sector limited by the bogie instability and, in most cases, also a sector limited by low damped carbody modes.

As the bogie stability decreases with increasing conicity, the stability should be mainly investigated for the upper range of conicity anticipated in operation. Since the worst case scenarios of wheel–rail contact are usually not known during the design process, the calculations are carried out for a parameter range based on experience. In UIC 518,<sup>66</sup> the conicity range as a function of speed is specified (Table 12.2), for which the vehicle should be verified by measurements to run without instability. However, it is well-known that the real situation varies for different railway companies and that conicity can reach higher values than specified in UIC 518 (see Refs 82 and 83).



**FIGURE 12.23** Example of typical stability map (stability diagram).

**TABLE 12.2**  
**Maximum Value of Equivalent Conicity Specified for Acceptance Tests**  
**According to UIC 518<sup>66</sup>**

Maximum Speed	Maximum Conicity
$V \leq 140$ km/h	0.50
$140 \text{ km/h} < V \leq 200$ km/h	0.40
$200 \text{ km/h} < V \leq 230$ km/h	0.35
$230 \text{ km/h} < V \leq 250$ km/h	0.30
$250 \text{ km/h} < V \leq 280$ km/h	0.25
$280 \text{ km/h} < V \leq 350$ km/h	0.15

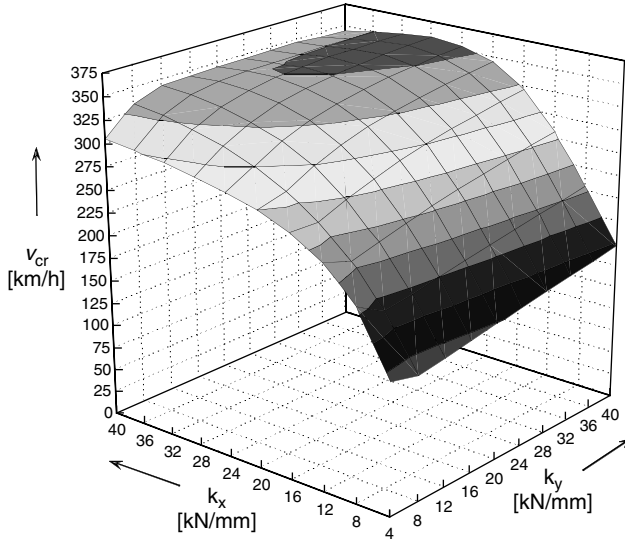
The typical conditions for linearised bogie stability analysis are summarised in Table 12.3. As the critical speed reaches the lowest value for dry wheel–rail contact, the linearised analysis should be realised for full creep coefficients of Kalker’s linear theory. For the lowest damped kinematic bogie mode a minimum 5% of critical damping is recommended. The stability criterion mentioned should be met for the maximum test speed which, according to UIC 518,<sup>66</sup> is the maximum vehicle speed increased by 10% or by 10 km/h for maximum speed below 100 km/h.

The linearised bogie stability computations are a very useful instrument during the concept investigation and design optimisation. The influence of several parameters on the critical speed can be investigated and an optimum range identified. However, for exact analysis, nonlinear calculations have to be used allowing simulation of real running conditions.

Figure 12.24 shows an example of parameter variation for longitudinal and lateral axle guidance stiffness of a four-axle locomotive. The levels with constant critical speed usually possess an approximate form of hyperbola. In the presented case the influence of longitudinal axle guidance stiffness plays a more important role because of large yaw damping between body and bogie. Better stability is achieved for large axle guidance stiffness. However, during design optimisation the trade-off between stability and curving requirements has to be solved (see Section V.G).

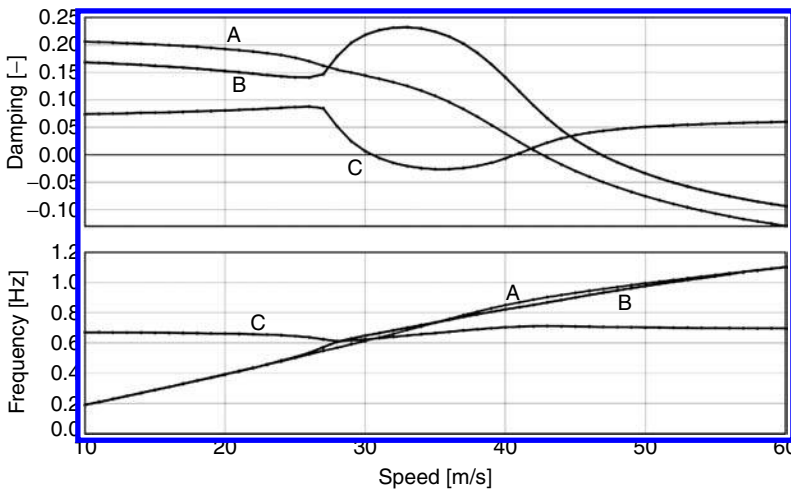
**TABLE 12.3**  
**Typical Conditions for Linearised Bogie Stability Analysis**

Input Parameter	Recommended Value or Conditions
Wheel–rail contact geometry	Variation of equivalent conicity up to the maximum value expected in the service, representing various combinations of worn wheel and rail profiles as well as gauge variation
Wheel–rail creep-force law	Full creep coefficients of Kalker’s linear theory (dry rail)
Vehicle state	1. Intact 2. Failure mode: failure or reduced effect of yaw dampers
Vehicle loading	1. Tare (empty) 2. Full (crush) load
Vehicle speed	Speed variation from low up to high speed, in minimum up to the maximum test speed (usually maximum service speed + 10%)



**FIGURE 12.24** Example of stability map of a four-axle locomotive: diagram of critical speed  $v_{cr}$  (speed at 5% of critical damping) in function of longitudinal and lateral axle guidance stiffness for equivalent conicity of 0.45.

At speeds for which the bogie sinusoidal frequency approaches the natural frequency of the vehicle carbody on the suspension, the possibility of considerable interaction may arise, leading to instabilities during which the amplitude of the carbody is large relative to that of the wheelsets. The bogie movement is coupled with carbody movement, usually yaw or lower sway carbody modes, sometimes a combination of both, or also in combination with carbody pitching. The kinetic energy is transferred from bogie to carbody so that the damping of the bogie eigenmode increases whereas the damping of the carbody decreases as explained, e.g., in Refs 73 and 84 (see Figure 12.25).



**FIGURE 12.25** Example of carbody instability (C) at very low conicity (from Mahr<sup>84</sup>). Damping of the sinusoidal movement of trailing bogie (B) increases whereas the damping of carbody decreases for the speed between 27 and 45 m/sec (A, B — sinusoidal movement of leading or trailing bogie, respectively; C — carbody sway movement).

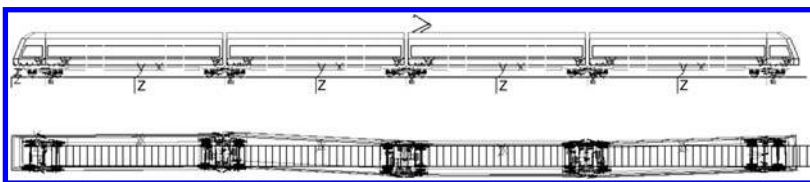
For two-axle vehicles or for freight wagons with laterally stiff coupling between the bogie and carbody, carbody instability with a frequency of approximately 2 Hz can lead to derailment risk as illustrated in Ref. 85. For vehicles with secondary suspension, carbody instability does not usually lead to the limits used for assessment of stability in measurements being exceeded, as the measurements are oriented on the bogie instability. The carbody oscillations cause deterioration of the lateral comfort behaviour. A significant comfort worsening may have already occurred for low damped eigenmode in spite of the fact that the carbody is still stable. For this reason it is recommended that the eigendamping of carbody modes should reach more than 10% of critical damping for the whole range of speeds expected in service. Owing to the influence on running comfort, the low damped carbody oscillations are often investigated from a comfort analysis standpoint.<sup>86</sup> On occasion, this coincidence of carbody and bogie eigenfrequency is referred to as resonance. However, as the motion is self-excited by wheelset movement, it constitutes low damping or instability of the carbody as also explained in Refs 73 and 80. The carbody instability is promoted by low values of conicity and creep coefficient. In service, the unstable or low damped carbody eigenbehaviour leads to increased oscillations for a certain speed range, mainly on smooth track whereas, on bad quality track, the eigenbehaviour is disturbed by track irregularity. The appearance of low damped or unstable carbody eigenmodes is dependent on all parameters of the vehicle, so that it is not easy to assess the risk of this phenomenon. The secondary suspension parameters are usually mentioned as being of influence. Should carbody instability be avoided, the lateral stiffness and lateral damping between the bogie and body may not exceed a certain limit.<sup>80</sup>

The axle guidance can also influence carbody instability. According to Refs 103 and 104, vehicles with forced steering are more sensitive to carbody instability than conventional vehicles. Carbody oscillations can also often be observed on articulated vehicles. In Figure 12.26 an example is given of an unstable eigenmode of a four-car articulated train. The cars are moving in a snaking mode with a frequency of approximately 1.2 Hz at a speed of 140 km/h and conicity of 0.05. A similar phenomenon is well-known from articulated LRV and trams.<sup>87</sup> On articulated vehicles the low damped carbody modes can be avoided by introducing damping between carbodies. As a design solution, two parallel longitudinal dampers between the carbodies are used. The vertical position of these intercar dampers can be at floor or roof level. The intercar dampers are also implemented on high-speed vehicles, e.g., on Shinkansen trains in Japan,<sup>88</sup> where the carbody oscillations are likely to be promoted by an aerodynamic phenomenon in tunnels.

It is of advantage to investigate the risk of unstable or low damped carbody oscillations with linearised calculations. Typical conditions and parameters used are shown in Table 12.4. In this manner, a set of speeds and parameter variations can be investigated very quickly and the coincidence of frequencies can be observed.

### 3. Nonlinear Stability Analysis

Two main groups of methods are utilised for nonlinear stability analysis in the railway vehicle engineering process. The first group of methods is based on the theory of nonlinear dynamics, the second on the measuring methods for bogie instability in the vehicle acceptance tests.

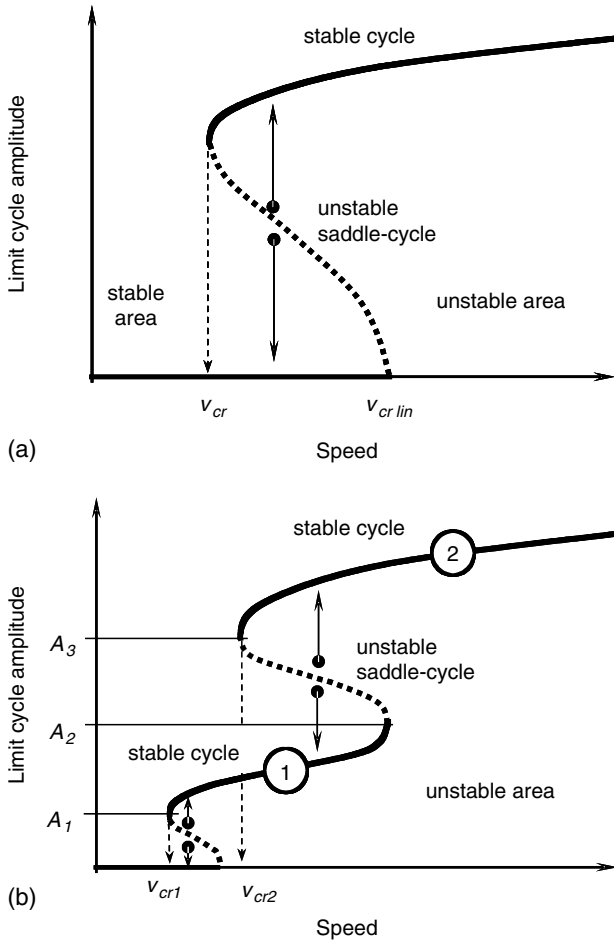


**FIGURE 12.26** Shape of an unstable carbody eigenmode of an articulated train at a speed of 140 km/h and conicity of 0.05.

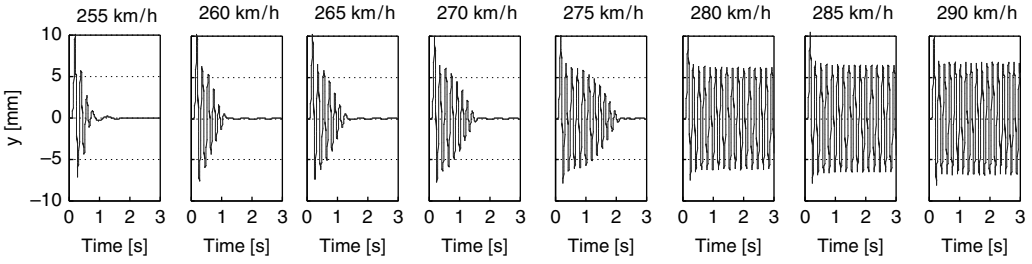
**TABLE 12.4**  
**Typical Conditions for Linearised Carbody Stability Analysis**

Input Parameter	Recommended Value or Conditions
Wheel–rail contact geometry	Low equivalent conicity (<0.1)
Wheel–rail creep-force law	1. Full creep coefficients of Kalker’s linear theory (dry rail) 2. Reduced creep coefficients of Kalker’s linear theory, reduction factor 0.2–0.6 (wet rail)
Vehicle state	Intact
Vehicle loading	Tare (empty)
Vehicle speed	Speed variation between very low speed and maximum service speed

Bogie instability is a nonlinear phenomenon. The amplitude of a limit cycle as a function of vehicle speed is presented in a bifurcation diagram<sup>89</sup> (see Figure 12.27). For certain speeds in the area of an unstable saddle cycle, the vehicle can either run stably or with a limit cycle depending on the initial conditions (initial disturbance). For speeds over the critical speed  $v_{cr}$ , large wheelset oscillations suddenly occur, if the initial disturbance of the wheelset reaches a certain limit



**FIGURE 12.27** Amplitude of limit cycles as function of speed in a bifurcation diagram. (a) large limit cycles only; (b) small and large limit cycles.



**FIGURE 12.28** Example of a multisimulation to calculate the bifurcation diagram. Limit cycles with large amplitude commence suddenly above 275 km/h.

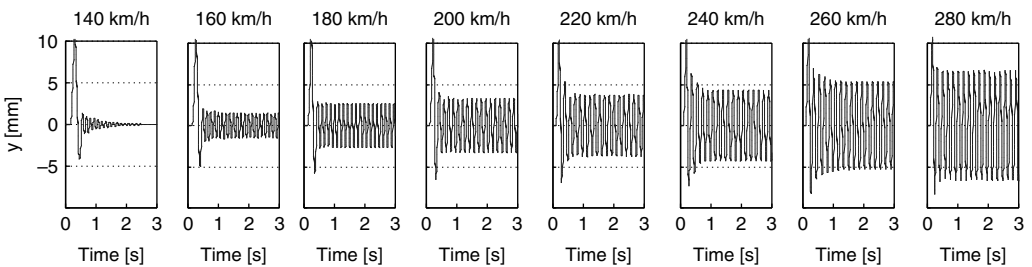
(Figure 12.27a). Sometimes, there are also limit cycles with a small amplitude present in the bifurcation diagram (Figure 12.27b). The limit cycle amplitude identified will be positioned on curve sequence 1 for any disturbances between  $A_1$  and  $A_2$ , or on curve sequence 2 for disturbances larger than  $A_3$ , respectively. For disturbances between  $A_2$  and  $A_3$ , the amplitude of the limit cycle will depend on the position of the initial disturbance against the unstable saddle cycle, as shown by arrows.

The critical speed of a bogie running on ideal track without disturbance is the linear critical speed  $v_{cr \text{ lin}}$ . The nonlinear critical speed is the lowest speed at which limit cycle occurs. In the case given in Figure 12.27b, from a theoretical point of view, the critical speed is  $v_{cr1}$ , but below speed  $v_{cr2}$  only a limit cycle with very small amplitude exists and no flange-to-flange wheelset movement is present. A limit cycle with such a small amplitude is often not safety-critical and will not necessarily lead to exceeding the stability limit in engineering applications.<sup>90</sup>

Direct computation of a complete bifurcation diagram is not yet state of the art in multibody simulation tools, although the method has been developed and tested.<sup>45</sup> The stable cycle line of the bifurcation diagram can also be calculated by a set of simulations as a reaction on initial disturbance. As an output value, lateral or yaw movement of the wheelset can be applied. Examples of multisimulations to calculate the bifurcation diagrams are given in Figure 12.28 and Figure 12.29.

Similar types of tests with excitation by initial disturbance are used for stability investigations on a roller rig. In this case a single lateral disturbance in the form of a half-cosine with wavelength of 10 m and excitation amplitude of 5 to 10 mm is used. This test can be achieved virtually by simulation. The critical speed based on a limit cycle with large amplitude can be determined, however, without information concerning the influence of excitation amplitude.

Owing to strong nonlinearity of the wheel–rail contact, the linear critical speed  $v_{cr \text{ lin}}$  can be higher than the nonlinear critical speed  $v_{cr}$ . In experiments on the roller rig in Munich it was



**FIGURE 12.29** Example of a multisimulation to calculate the bifurcation diagram. Limit cycles commence already for speeds above 140 km/h. At first the amplitude is small, then increases with speed and only reaches large flange to flange movement for speeds higher than 260 km/h.

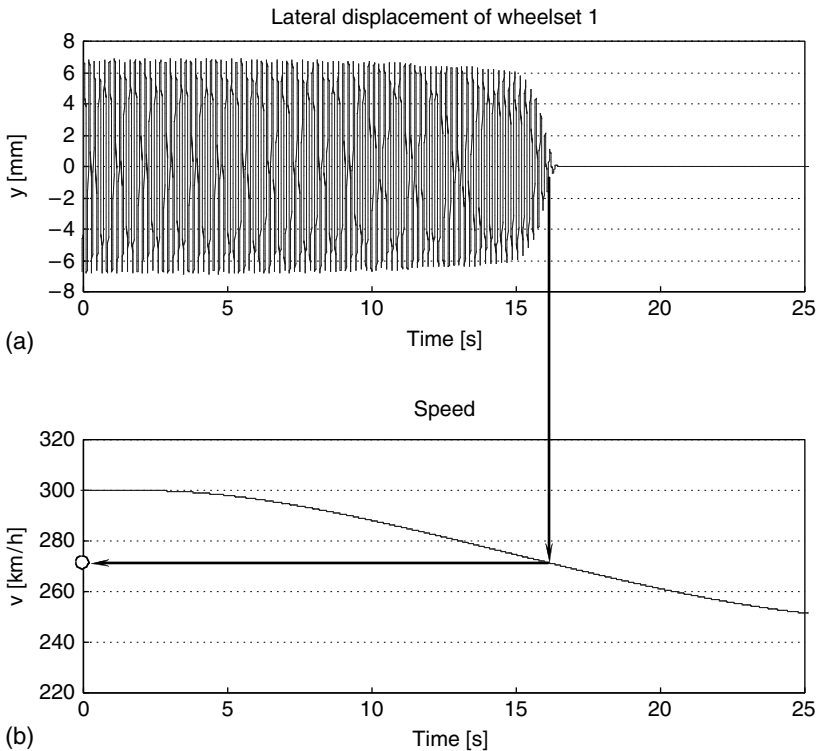
demonstrated that, during speed increase and decrease, the sinusoidal wheelset movement does not start and stop at the same speed. As explained in Refs 73 and 91, on a smooth track the instability commences at the linear critical speed and stops at the nonlinear critical speed. Some authors<sup>92</sup> recommend investigation of the nonlinear critical speed by simulating a run on ideal track (without excitation), starting at a high speed with unstable oscillations and reducing the speed continuously (see Figure 12.30). The critical speed is then the speed at which the wheelset oscillation stops. The result of this method is in good agreement with other methods, if the instability occurs suddenly with large limit cycles. Otherwise, if limit cycles with small amplitudes are also present (Figure 12.29), the critical speed would be significantly lower than the value of critical speed used in the engineering applications (see also Ref. 90).

To prove bogie stability during the engineering process, the methods specified for measurements and acceptance tests can also be applied. Vehicle running on straight track with irregularities under the worst wheel–rail contact conditions is simulated and criteria for measurements and vehicle acceptance are used for assessment.

According to UIC 518<sup>66</sup> and EN 14363,<sup>67</sup> the rms-value of the sum of guiding forces (track shifting force) is used in full on-track test. The limiting value is dependent on the static wheel load  $Q_0$  and is equal to a half of the limit according to Prud’homme:

$$(\sigma \Sigma Y)_{\text{lim}} = \frac{(\Sigma Y_{2m})_{\text{lim}}}{2} = \frac{1}{2} \left( 10 + \frac{2Q_0}{3} \right) \text{ [kN]} \tag{12.5}$$

The rms-value is a continuous average value over 100 m distance calculated with steps of 10 m.



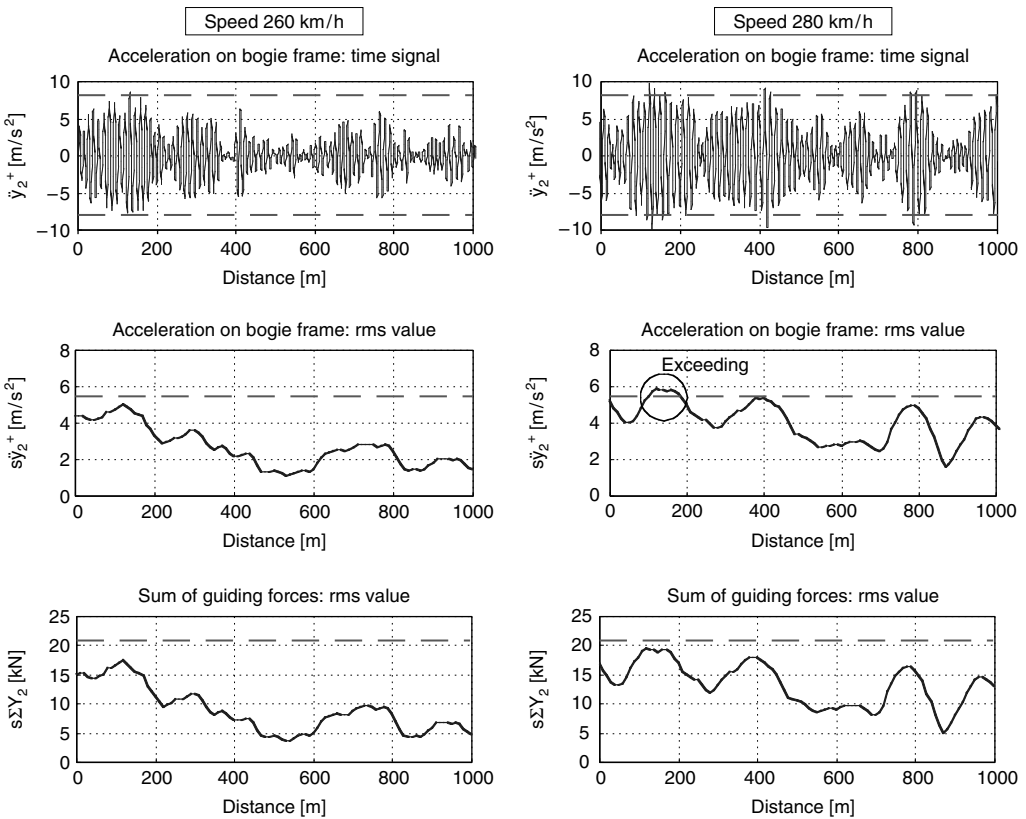
**FIGURE 12.30** Nonlinear stability analysis by simulation of run on ideal track with continuously decreasing speed. Vehicle and simulation parameters are identical to the example in Figure 12.28.

Another practice for proving stability in engineering applications is to analyse lateral accelerations on the bogie frame. The simplified method uses acceleration filtered with band-pass filter ( $f_0 \pm 2$ ) Hz, where  $f_0$  is the frequency of unstable bogie oscillations. The investigated signal, the rms-value over 100 m distance calculated with steps of 10 m, should be compared with the limit value specified in function of bogie mass  $m^+$  in tons as

$$(\dot{y}^+)_{lim} = \frac{1}{2} \left( 12 - \frac{m^+}{5} \right) [m/s^2] \tag{12.6}$$

Lateral acceleration on the bogie frame is traditionally used to check the bogie stability, also without the calculation of rms-value. The limiting value is then 8 m/sec<sup>2</sup>, however, the limit is not valid for exceeding once. To fulfill the criterion, the amplitude must, instead, not exceed the limit more than six times in succession.

Figure 12.31 shows an example of stability assessment using simulation results from a run on straight track with measured irregularities. All the signals mentioned can be used to prove the stability by simulation. However, the various limits mentioned above can lead to differing results.<sup>82,83</sup> In the example of simulation runs in Figure 12.31, one of the criteria is exceeded for the speed of 280 km/h, whereas the other criteria are still fulfilled (however, with only a small margin). Compared with stability analysis using a bifurcation diagram, the stability limits mentioned above will be achieved first when the limit cycles with large amplitudes occur.



**FIGURE 12.31** Comparison of criteria for stability analysis from simulation of running on track with measured irregularities. Vehicle and simulation parameters are identical to those in Figure 12.28 and Figure 12.30 (dashed lines = limit values).



**TABLE 12.5**  
**Typical Conditions for Nonlinear Bogie Stability Analysis**

Input Parameter	Recommended Value or Conditions
Track design	Straight track
Track irregularity	1. Ideal track with single lateral disturbance 2. Measured track irregularities
Wheel–rail contact geometry	Variation of wheel and rail profiles, including worn wheel and rail profiles and gauge variation, representing conditions expected in service
Wheel–rail creep-force law	Nonlinear theory, friction coefficient 0.4 to 0.5 (dry rail)
Vehicle state	1. Intact 2. Failure mode: failure or reduced effect of yaw dampers
Vehicle loading	1. Tare (empty) 2. Full (crush) load
Vehicle speed	1. Speed variation up to the critical speed 2. Maximum test speed (usually maximum service speed + 10%)

For the nonlinear stability analysis, the worst wheel–rail contact conditions have to be used. The sole application of nominal parameters and nominal wheel and rail profiles is definitely insufficient. The equivalent conicity increases for narrowed gauge. Wear of wheels usually leads to a reduced curve radius of wheel tread, top-of-rail wear on straight track leads to reduced curve radius of rail corner. Both wear phenomena lead to increased conicity.<sup>82</sup> To cover the worst case conditions, stability investigations should be carried out for worn shapes of wheel and rail profiles and possibly for narrowed gauge. As measured worn profiles are often not available, an artificial gauge narrowing is sometimes used to achieve increased conicity. By reducing the gauge the equivalent conicity increases, but the lateral movement of the wheelset is then limited and the vehicle behaves differently to the same equivalent conicity on a wider gauge. It is therefore better to use worn wheel and rail profiles with nominal or widened gauge to cover the anticipated range of conditions. Another important parameter is the friction coefficient between wheel and rail. As the critical speed decreases with increasing friction coefficient, dry conditions should be used for analysis. The friction coefficient applied is usually between 0.4 and 0.5. Typical conditions and parameters used for nonlinear bogie stability analysis are summarised in Table 12.5.

## F. RUN ON TRACK WITH IRREGULARITIES

### 1. Definition of Running Behaviour, Ride Characteristics and Comfort

Simulation of run on track with measured irregularities allows the prediction of the vehicle running behaviour that can be anticipated during the type test. According to EN 14363,<sup>67</sup> the running behaviour<sup>a</sup> constitutes the characteristics of the vehicle or running gear with regard to interaction

<sup>a</sup> In the UIC 518, the term “running behaviour” is used in the sense of “ride characteristics” in EN 14363, in parallel to the safety and track fatigue.

between the vehicle and track. Running behaviour is a general description covering the following specific terms: running safety, track loading, and ride characteristics.

For the running safety on straight track, the sum of the guiding forces provides a criterion. However, the limit value for this criterion is normally fulfilled if the bogie and vehicle are stable. This criterion, together with the simulations related to stability, was explained in Section V.E.

For the track loading, the term “track fatigue” is also sometimes used because the limit values are related to the fatigue of the track components. Investigation of track loading is important for assessment of the run on curved tracks and will be presented in Section V.G.

The ride characteristics provide assessment of the dynamic behaviour of vehicle analysing accelerations at the vehicle body, whereas the ride comfort assesses the influence of vehicle dynamic behaviour on the human body. Although both criteria use acceleration signals, the analysis and limit or target values differ. Simulations related to the analysis of ride characteristics and ride comfort are carried out first on straight track as explained in the following sections. However, curves and transitions generally also have to be considered in order to take the same conditions into account as during measurements.

## 2. Ride Characteristics

To simulate a run of a vehicle on track, measured irregularities are usually applied. An overview concerning definition and properties of measured track irregularities can be found in several references, e.g., in Ref. 93. Another possibility is the use of synthesised irregularities with specified spectral density. Track irregularity data often used in European countries are the spectral density “low level” and “high level” according to ORE B176 described in Ref. 81.

The typical conditions for simulations related to ride characteristics, as well as ride comfort, are described in Table 12.6. It is comprehensible that the irregularity has a significant influence on the ride characteristics. However, it is not easy to make a definite assessment of the track quality, as the same vehicle can demonstrate different ride characteristics on differing tracks. A change from one track to another does not always demonstrate a clear tendency to an overall improvement or deterioration of the vibration behaviour. In fact, the occurrence of opposing trends at various points in the vehicle is possible.

The ride characteristics are also influenced by the parameters of the wheel–rail combination. The best ride characteristics are usually achieved at medium equivalent conicities, i.e., 0.10 to 0.25. In the case of detailed investigations concerning ride comfort the wheel–rail contact may play an important role, as detailed in Section V.F.3.

Figure 12.32 shows an example of ride characteristics simulation for a four-axle locomotive with a maximum service speed of 140 km/h running on synthetic irregularities “low level” according to ORE B176<sup>81</sup> at a speed of 154 km/h. Several versions of nonlinear wheel–rail contact geometry were applied in the simulations. A wide range of equivalent conicity variation was achieved applying a new wheel profile together with either variation of gauge or variation of profile and inclination of rails, respectively. Vertical and lateral accelerations in the driver’s cab filtered with a band-pass filter 0.4 to 10 Hz are presented in Figure 12.32 as a function of equivalent conicity evaluated for the lateral wheelset amplitude of 3 mm.

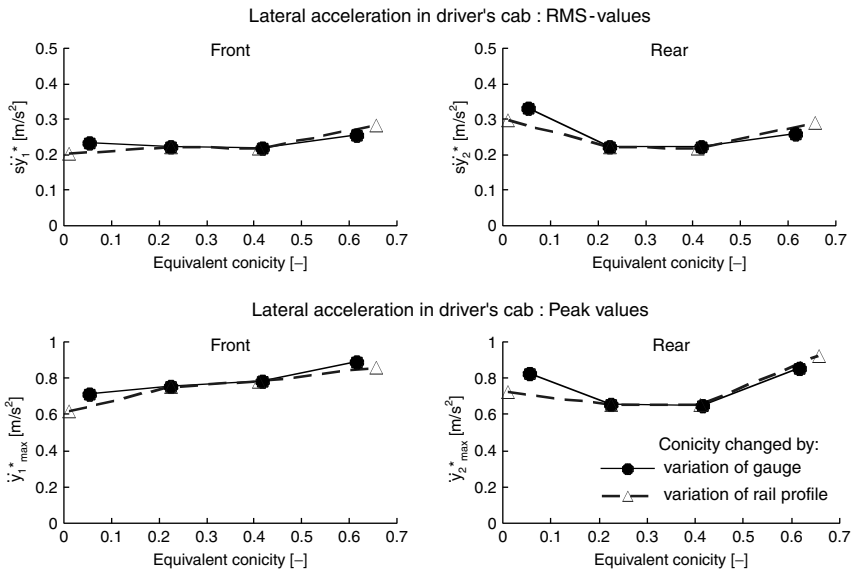
In the case of vehicles with airsprings, running on emergency suspension should also be simulated as required during acceptance tests, in addition to the intact conditions (see example in Figure 12.33).

Loading of the vehicle usually leads to lower natural frequencies, and in consequence, to an improvement in ride comfort in the vertical direction. The vehicle will therefore be mainly examined without loading (tare).

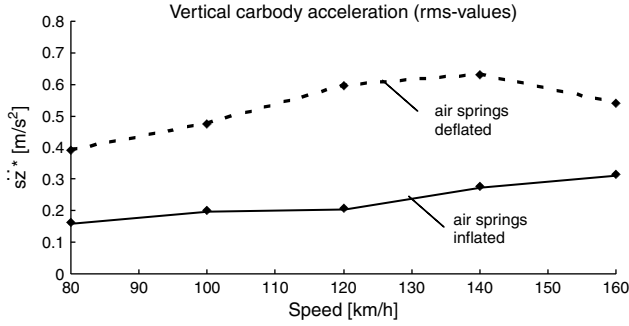
The vibrations cumulate relevant to increasing running speed, which in turn results in a deterioration of the ride characteristics. For this reason the limit values of the ride characteristics are examined during the acceptance test at the maximum test speed (usually maximum service speed

**TABLE 12.6**  
**Typical Conditions for Simulation of Ride Characteristics and Comfort**

Input Parameter	Recommended Value or Conditions
Track design	Straight track, curves with typical radius; optional comfort analysis in curve transitions (tilting trains)
Track irregularity	According to the specification and conditions on the railway network; measured track irregularity if possible
Wheel–rail contact geometry	Nominal profiles of wheel and rail, nominal gauge; sensitivity analysis of gauge narrowing and widening, analysis of influence of worn wheel and rail profiles
Wheel–rail creep-force law	Nonlinear theory, friction coefficient 0.4 (dry rail); analysis of influence of reduced friction coefficient 0.1 to 0.3 (wet rail)
Vehicle state	1. Intact 2. Failure mode: airsprings deflated (for ride characteristics only)
Vehicle loading	Tare (empty)
Vehicle speed for ride characteristics analysis	Maximum test speed (usually maximum service speed + 10%)
Vehicle speed for ride comfort analysis	1. Maximum service speed 2. Speed of carbody pitch and bounce resonance, see Section V.F.3



**FIGURE 12.32** Running characteristics of a four-axle locomotive as a result of simulations with nonlinear wheel–rail profiles and presented as a function of equivalent conicity.



**FIGURE 12.33** Influence of vehicle speed on vertical ride characteristics of a four-axle motor coach for intact and deflated airsprings.

+ 10%) and are therefore also investigated at this speed during the engineering process. A negative influence on the ride characteristics will be caused through a resonance by harmonic components of the excitation with natural frequencies of the vehicle, (see also Section V.F.3) so that the maximum test speed may not necessarily provide the worst ride characteristics values. Figure 12.33 shows a comparison of the ride characteristic of a rail car in intact condition (with airsprings inflated) and in an emergency run status (airsprings deflated). Whereas the acceleration values increase over the whole range concurrent to the speed in the case of the vehicle with airsprings inflated, during a run with airsprings deflated a maximum occurs in the resonance frequency range of the emergency suspension, followed by a slight reduction in the values.

### 3. Ride Comfort

Ride comfort characterises passenger well-being in relation to mechanical vibrations, thereby taking the physiological characteristics of the human body into consideration. The manner in which the human body experiences ride comfort differs depending on the frequency and amplitude of the vibrations. In order to take this influence into account frequency weighting filters are applied. The filters vary for the vertical and lateral direction and also differ depending on the standard applied.

The most widely used comfort analysis methods are the comfort index (*N*-value) analysis according to UIC 513,<sup>94</sup> specified also in ENV 12299,<sup>95</sup> rms-method specified in ISO 2631,<sup>2</sup> and Sperling’s method (comfort value *W<sub>z</sub>*). In a simplified method, accelerations in vertical and lateral (for *N*-value also in longitudinal) directions are evaluated on the floor of the carbody. Another option is a measurement on the seat and seat-back in conjunction with the human body. The signals are measured in 5 sec sequences and analysed for 60 blocks (see the standards or Ref. 96 for details). The comfort index *N* is calculated as a 95%-value from the statistical histogram of accelerations in all three directions, whereas the rms-value according to ISO 2631 or *W<sub>z</sub>*-value according to Sperling are evaluated separately in vertical and lateral direction. Owing to differing weighting filters and methodology, it is difficult to transform one comfort index to another one without a complete analysis. An attempt to identify such transformation formulae between *W<sub>z</sub>*-values and *N*-values has been made.<sup>96</sup>

The influence on ride comfort is similar to that on the running characteristics. The simulation conditions are therefore the same as for ride characteristic simulation, as can be seen in Table 12.6 (Section V.F.2). However, as the specifications regarding ride comfort are usually stricter than the standards for ride characteristics, the simulation investigations and prognoses of the ride comfort are more demanding than investigations of the ride characteristics.

In contrast to other calculations, a very important role is played in the comfort calculations with regard to the modelling of the carbody (see Section II.B and Figure 12.3). The elastic carbody structure, the connection of the carbody with the bogie, and the distribution and suspension

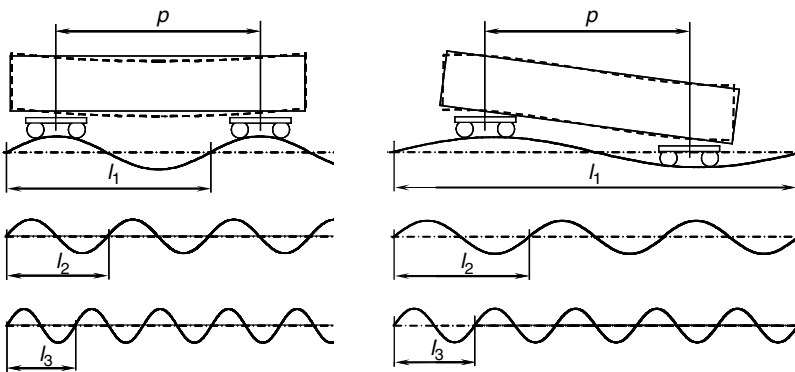
of the apparatus play decisive roles. While the influence of the flexible carbody structure above the bogies is relatively small, the vibrations at the vehicle centre and at the vehicle ends are strongly influenced by the flexibility of the structure. In a vehicle model with a rigid carbody the vibration behaviour in the vehicle centre is typically better than above the bogies. If the flexibility of the carbody structure is taken into consideration, the vibration behaviour at the vehicle centre achieves higher values and is usually higher than the values above the bogies, particularly in the case of light-weight vehicles possessing low carbody structural stiffness. In order to achieve realistic simulations it is therefore necessary to model the carbodies as elastic structures (see [Section II.B](#) and [Refs 97 and 98](#)).

Vibration comfort is unfortunately strongly influenced by the position in the vehicle. Above the bogies the low frequency vibration forms of the carbody (e.g., bouncing, pitching, and swaying) prevail. In the centre of the carbody a frequency of approximately 8 to 12 Hz typically dominates, which indicates the eigenmode of vertical carbody bending. Standards and specifications relevant for comfort tests usually indicate that the limit values or target values are to be fulfilled at all positions in the vehicle. This signifies that, by way of simulations, the most unfavourable position is to be located. Practically, the ride comfort is usually examined at nine locations on the carbody: longitudinally above both bogies, and in the centre of the vehicle, and transversally left, in the centre and to the right in the running direction.

Some unfavourable conditions that are particularly detrimental for comfort are:

- Low damping or even instability of the carbody mode initiated through the coupling of the self-excited sinusoidal bogie movement with carbody eigenmode as described in [Section V.E.2](#)
- Resonance from the eigenmode of the vehicle components with the periodic excited mode

To study the influence of vehicle speed, we have to consider the excitation by track unevenness. The track irregularity can be broken down into the harmonic components with the aid of a Fourier transformation. Of these, the wavelengths possess special significance for the ride comfort, which are in a particular relationship to the bogie pivot pin distance  $p$  (see also [Refs 86 and 97](#)). As can be seen in [Figure 12.34](#), at a particular speed bouncing, pitching, or bending of the carbody is excited by some wavelengths. These wavelengths can be described as being “critical wavelengths” for the speed in question. In the case of carbody bouncing and the first bending mode those critical



**FIGURE 12.34** Wavelengths exciting vertical carbody modes: bouncing and first bending mode (left), pitching and second bending mode (right).

wavelengths  $l_m$  are:

$$l_m = \frac{P}{m} \quad m = 1, 2, 3, \dots \tag{12.7}$$

The resonance occurs when, at particular speed  $v_m$ , the carbody is excited by the critical wavelength of a carbody eigenfrequency  $f_i$ .

$$\frac{v_m}{l_m} = f_i \quad m = 1, 2, 3, \dots \tag{12.8}$$

The critical resonance speeds for bouncing and for the first bending form of the carbody can be derived from Equation 12.7 and Equation 12.8 as:

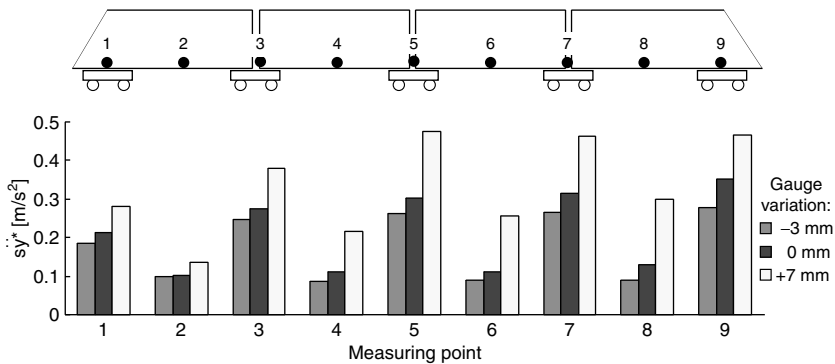
$$v_m = \frac{P}{m} f_i \quad m = 1, 2, 3, \dots \tag{12.9}$$

In a similar manner we can obtain the resonance speeds for pitching or the second bending form of the carbody (see Figure 12.34) as:

$$v_n = \frac{P}{n - \frac{1}{2}} f_i \quad n = 1, 2, 3, \dots \tag{12.10}$$

The resonance speeds for the bouncing and pitching movement of the carbody as a rigid body are low, which is why the influence on ride comfort is not very large. In particular, the first bending mode of the carbody structure is liable to strong excitation. While the first resonance speed for the bending of the carbody usually lies above the maximum speed, the second or third resonance speeds are often critical and should be examined with the flexible carbody. The resonance with carbody pitching movement leads to high accelerations during an emergency run on deflated airsprings in the speed range 80 to 140 km/h, which sometimes necessitates a speed limitation.

The parameters of the wheel–rail combination influence the comfort results, whereby these influences are strongly dependent on the vehicle design. At low conicity the self-excited low-damped eigenmode of the carbody can occur, as described in Section V.E. If the vehicle has a tendency to such vibrations in a specific speed range, the situation will worsen in the case of gauge widening (decrease of conicity), and improve in the case of gauge narrowing (increase of conicity); see the example of an articulated commuter train in Figure 12.35. Higher conicities and the allied



**FIGURE 12.35** Lateral accelerations on the floor of the carbodies of an articulated commuter train (rms-values filtered according to ISO 2631) at a speed of 160 km/h.

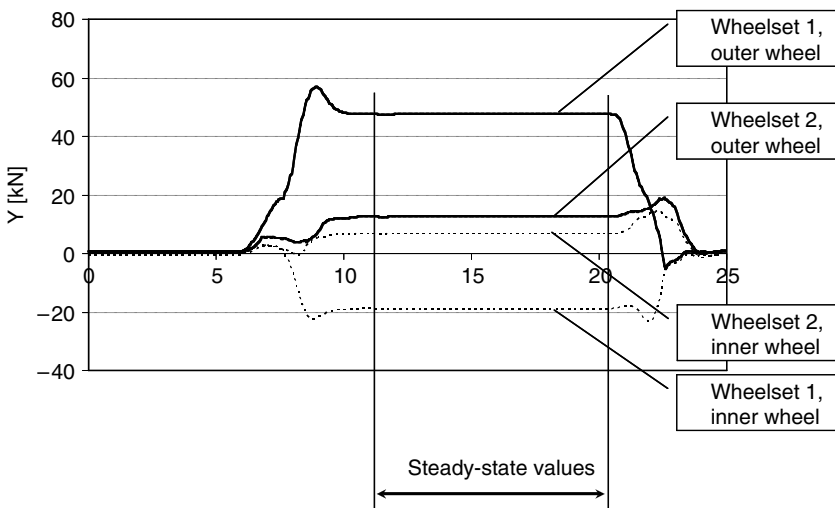
increasing tendency of the bogie to hunting, lead to moments on the carbody structure. A corresponding eigenmode of the carbody structure will be excited and may lead to poor ride comfort. In Ref. 86 an example is given whereby an increase of the equivalent conicity from 0.2 to 0.55 leads to a tripling of the calculated rms-values at the sides in the middle of the carbody, while the influence of the conicity in the central longitudinal axis remains slight.

## G. CURVING

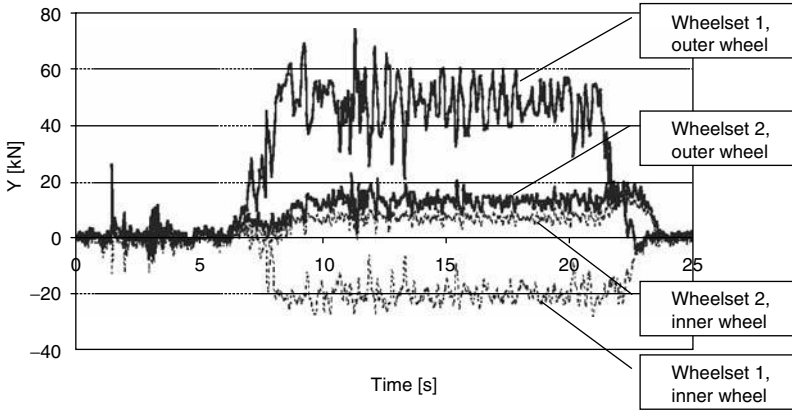
### 1. Assessment of Curving Properties

During curving the vehicle is guided by the forces between wheel and rail in the lateral direction. These guiding forces can reach very high values, particularly when heavy vehicles run through narrow curves. For example, Figure 12.36 shows lateral forces between wheel and rail at a locomotive bogie when running with uncompensated lateral acceleration of  $1.1 \text{ m/sec}^2$  through a curve with a radius of 300 m. The vehicle is guided into the curve direction mainly by the lateral force on the flange of the outer leading wheel on each bogie. An opposite, smaller lateral force exists on the inner wheel of the leading wheelset. The lateral forces of the trailing wheels are relatively small and can have the same or opposite direction to each other. This is a typical force distribution for a bogie with stiff wheelset guidance running in a curve with a small radius. Considering an ideal track curvature, the simulation results are in accordance with Figure 12.36. The steady-state curving can be solved in an efficient way by special tools or routines for quasistatic analysis, if such routines are available in the simulation tool. They can also be found by time integration, simulating a run on ideal track without irregularities from a straight through curve transition into a curve with constant radius. However, the steady-state values can also be calculated from a simulation with measured irregularities, calculating average values (50%-values) from the dynamic forces in the same way as they are evaluated in measurements. Taking track irregularities into account, the forces are superimposed by dynamic effects as can be seen in Figure 12.37.

The following presentation of the curving analyses is oriented on the criteria according to standards where the limits for both dynamics as well as steady-state values are specified. The typical conditions used for simulations of curving are shown in Table 12.7. Since the wheel–rail



**FIGURE 12.36** Curving simulation. Lateral forces on the leading bogie of a four-axle locomotive running on an ideal track through a curve with a radius of 300 m.



**FIGURE 12.37** Curving simulation. Lateral forces on the leading bogie of a four-axle locomotive running on a track with measured irregularities through a curve with a radius of 300 m.

forces increase with increasing friction coefficient, dry conditions of wheel–rail contact are used for simulations. Depending on the bogie and vehicle design, the results can be sensitive to the wheel–rail contact geometry. To test the worst case conditions with regard to wheel–rail contact, worn rail profiles from curves should be used. For investigations into track loading, a fully loaded vehicle should be taken into account, whereas for investigations of derailment safety an empty vehicle is the most critical case.

**TABLE 12.7**  
**Typical Conditions for Simulations of Curving**

Input Parameter	Recommended Value or Conditions
Track design	1. Typical curve radii including transitions 2. The smallest curve radius on the network (outside of depot area)
Track irregularity	According to the specification and conditions on the railway network; measured track irregularity if possible
Wheel–rail contact geometry	Nominal wheel and rail profiles, nominal gauge, gauge widening in tight curves according to the specification; influence analysis of worn wheel and rail profiles, mainly for self-steering wheelsets
Wheel–rail creep-force law	Nonlinear theory, friction coefficient 0.4 (dry rail)
Vehicle state	Intact
Vehicle loading	1. Full (crush) load 2. Tare (empty); relevant for derailment safety investigation
Vehicle speed	Speed variation in function of curve radius and cant deficiency



## 2. Running Safety

To assess running safety in curves, the following criteria are used:

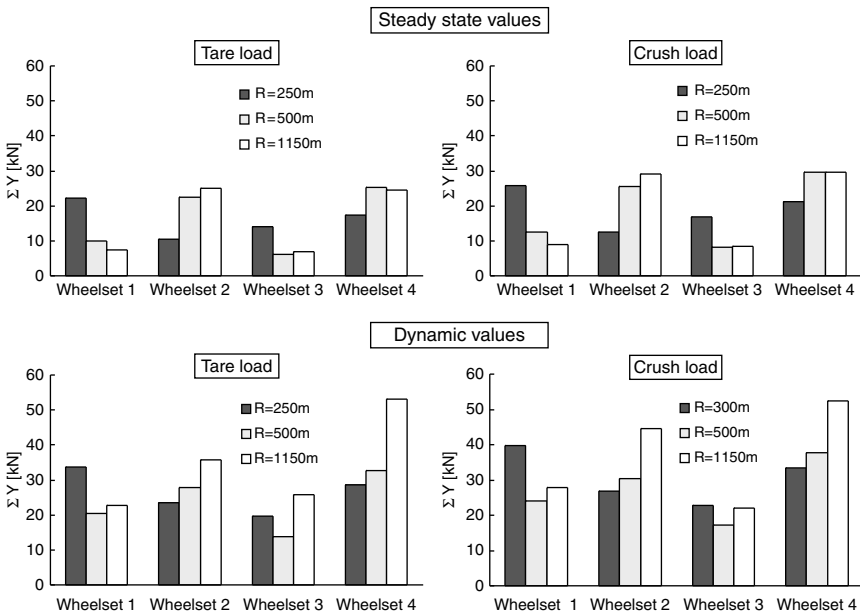
- Sum of guiding forces (track shifting force)
- Quotient of guiding force and vertical wheel force  $Y/Q$  ( $L/V$ ), often linked to Nadal’s criterion; this is covered in detail in Chapter 8.

For the sum of guiding forces, the limit used for an average value over 2 m distance is according to UIC 518<sup>66</sup> and EN 14363<sup>67</sup> the value according to Prud’homme:

$$\sum Y_{2m} = \alpha \left( 10 + \frac{2Q_0}{3} \right) [kN] \tag{12.11}$$

with  $\alpha$  taking account of greater variations in geometrical dimensions and of the state of vehicle maintenance ( $\alpha = 0.85$  for freight wagons,  $\alpha = 1.0$  for other vehicles). The criterion is valid on straight track as well as in curves. However, in curves the dynamic value excited by irregularities is superimposed by the steady-state value, which makes this limit more critical, mainly when running at high speed through curves with a large radius.

To reduce the sum of guiding forces, both steady-state as well as the dynamic value of the track shifting force has to be as low as possible. The sum of all steady-state track shifting forces acting on the vehicle is defined by the product of vehicle mass and lateral acceleration, which is a function of curve radius and cant deficiency. This criterion is challenging mainly for tilting trains running with large cant deficiency. An optimisation is limited to an equal distribution of the forces on wheelsets. Figure 12.38 gives an example of simulated track shifting forces of a regional vehicle running through a curve with a lateral acceleration of  $1.1 \text{ m/sec}^2$ . In tight curves the highest value occurs on the leading wheelset. In a large curve radius, the highest value is achieved on the trailing wheelset.



**FIGURE 12.38** Distribution of the sum of guiding forces of a four-axle motor coach running through a curve with uncompensated lateral acceleration of  $1.1 \text{ m/sec}^2$ .

Besides the vehicle design, curve radius and cant deficiency, the distribution of the track shifting forces is dependent on the friction coefficient and contact geometry between wheel and rail.

Similar to the steady-state values, the dynamic values of track shifting force are high for large cant deficiency. Furthermore, as can be observed in Figure 12.38, the values are high for a vehicle with maximum (crush) load and for curves with a large radius.

### 3. Track Loading and Wear

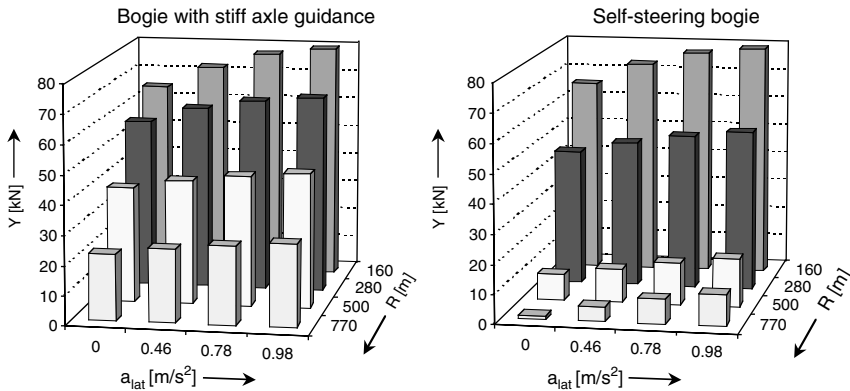
For track loading the individual forces between wheel and rail in a vertical and lateral direction are assessed: lateral guiding force  $Y$  and vertical force  $Q$ .

The limit values refer to both the steady-state as well as the dynamic values. With regard to the dynamic forces between wheel and rail, the limit value is only given for the vertical force in UIC 518<sup>66</sup> and EN 14363,<sup>67</sup> and only considering frequencies up to 20 Hz. For the steady-state forces according to UIC 518 and EN 14363 the following limit values apply: guiding force  $Y \leq 60$  kN and vertical wheel–rail force  $Q \leq 145$  kN. The given limit values refer to railway tracks with a maximum allowable load of 22.5 t per wheelset, and take account of rails with a minimum weight of 46 kg per metre and the minimum value of rail strength of 700 N/mm<sup>2</sup>. For tracks with higher axle loads and greater load capacity higher values may also be accepted.

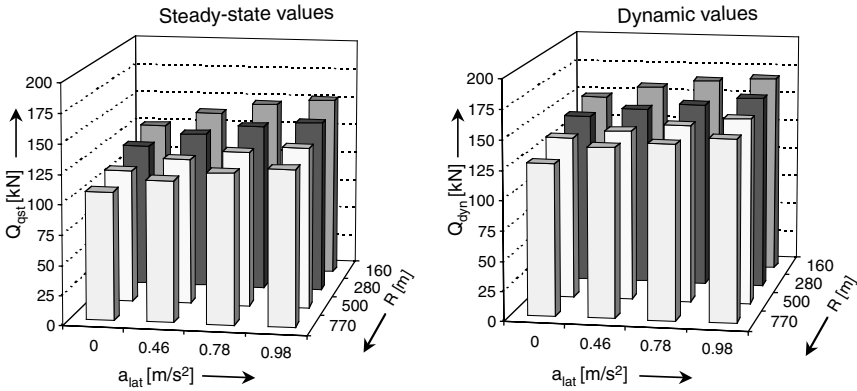
The guiding force  $Y$  achieves high values, particularly in narrow curve radii, during dry weather, and with heavy vehicles such as locomotives. Figure 12.39 gives an example, with a four-axle locomotive, of the steady-state guiding force  $Y$  as a function of curving radius and uncompensated lateral acceleration. The guiding force increases significantly concurrent with the decreasing curve radius. Increasing uncompensated lateral acceleration leads to a slight increase of the guiding force. In Figure 12.39 two bogie versions are compared:

- Conventional bogie design with longitudinal stiff axle guidance
- Self-steering bogie with longitudinal very soft axle guidance and cross-coupled wheelsets

The self-steering bogie demonstrates a good efficiency down to a small curve radius. The design presented here for the standard gauge locomotive significantly reduces the guiding force in large, medium, and small curve radii. In very small curve radii, approximately below 250 m, the creep forces cannot overcome the reaction of the axle guidance, and the guidance force is practically the same for both versions.



**FIGURE 12.39** Guiding force  $Y$  of the outer leading wheel in a curve as a function of lateral acceleration and curve radius with a four-axle locomotive as example. Comparison of stiff and self-steering bogie design.



**FIGURE 12.40** Steady-state and dynamic values of the vertical wheel–rail force  $Q$  of the outer leading wheel in a curve as a function of lateral acceleration and curve radius using a four-axle locomotive with stiff axle guidance as example.

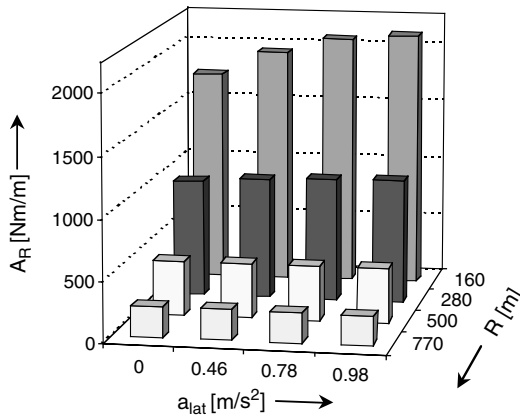
Figure 12.40 shows calculated steady-state and dynamic values of vertical wheel–rail forces for the same simulation runs as used for the investigation of guiding forces mentioned above. As a track irregularity, “low level” according to Ref. 81 was used. The vertical forces are high mainly for large uncompensated lateral acceleration and heavy vehicles such as locomotives.

Wheel and rail wear in curves represents an important aspect in respect to vehicle and track maintenance. In order to enable a simplified assessment of the wear, the friction work  $A_R$  effected (so-called wear index) is calculated as the sum of products from creep forces and creepages:

$$A_R = |F_{x,s_x}| + |F_{y,s_y}| + |\Theta\omega| \tag{12.12}$$

The last term mentioned, caused by spin  $\omega$  and moment  $\Theta$  around the normal to the contact surface, is very limited during wheel tread contact and can usually be neglected. Most subject to wearing is the wheel flange surface of the outer leading wheel. Wear is extremely dependent on the curve radius and increases significantly in very narrow curve radii (see Figure 12.41).

The life cycle of the wheels is not only influenced by the total friction work, but more significantly by the wear distribution over the lateral wheel profile. The alteration of the wheel



**FIGURE 12.41** Wear index of the outer leading wheel in a curve as a function of lateral acceleration and curve radius using a four-axle locomotive with stiff axle guidance as example.

profile can be simulated by determining the material removal in the individual profile sequences based on a wear theory. As the wear depends on the momentary form of the wheel and rail profile, a new simulation must be carried out with the newly determined worn wheel profile. In order to make a wear development prognosis, the simulations and analysis of profile form must be repeated several times in cycle, whereby a sequence of statistically representative running conditions such as vehicle loading, running speed, curve radius, track profile, wheel–rail friction coefficient, tractive effort, and braking forces must be taken into consideration. The procedure and examples of such extensive investigations can be seen in Refs 17, 99–101.

#### 4. Curving Optimisation Using Self-Steering and Interconnected Wheelsets

During the design of wheelset guidance, the contradictory objective arises of improving the curving performance, while achieving running stability at high speed. Independent of the form of the wheelset guidance and suspension design in the horizontal plane, the stability and curving performance can be described by two axle guidance stiffness parameters: the shear stiffness and the bending stiffness.<sup>92,102,103</sup> In order to achieve optimal curving properties, the bending stiffness should be low. However, the vehicle then lacks stability. To achieve the required stability the wheelsets have to be restrained by an increase in shear stiffness. For conventional bogies a limit exists to the shear stiffness that can be provided in relation to the bending stiffness. The trade-off between stability and curving in which the bending stiffness should be reduced is restricted.

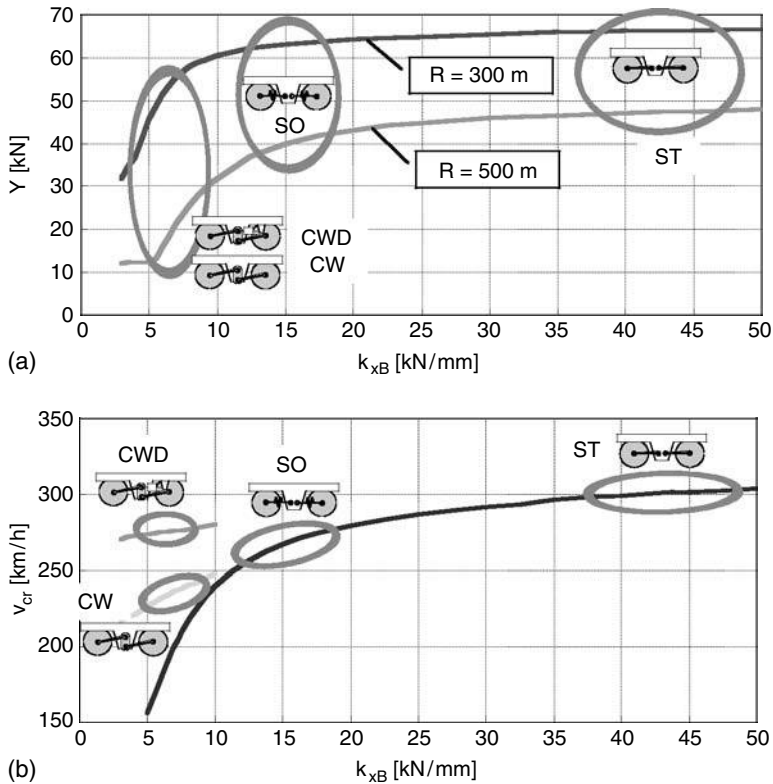
This limitation can be improved if the wheelsets are connected to each other directly or by a mechanism fitted on the bogie frame. In this way the trade-off between curving and stability can be improved, reducing the bending (steering) stiffness while retaining the range of achievable critical speed. An overview of design options and realised examples can be found in Refs 102 and 104. The application of a cross anchor on three-piece bogies and the service experience gained is described in Ref. 92. Another way of improving curving performance is the so-called forced steering of bogies or wheelsets by a carbody yaw movement or by angle between the carbodies.<sup>102,103</sup> A design with coupled single-axle running gears suitable for articulated vehicles is presented in Ref. 105. This construction combines the self-steering ability of soft axle guidance with forced steering by the carbody through secondary suspension.

An influence on the stability and curving performance of a locomotive with four different versions of modular axle guidance system was evaluated in Refs 106 and 107:

- Longitudinal stiff axle guidance (ST)
- Longitudinal soft axle guidance (SO)
- Longitudinal very soft axle guidance combined with wheelset coupling shaft (CW)
- Longitudinal very soft axle guidance combined with wheelset coupling shaft and dampers of coupling shaft (CWD)

The proposed versions were compared based on the equivalent longitudinal axle guidance stiffness  $k_{xB}$  for the bending (steering) mode of the wheelsets. The advantage of the soft axle guidance combined with cross-coupling of the wheelsets becomes apparent during curve negotiation as can be seen in Figure 12.42a. Both the guiding force of the leading wheel and the wear index decrease with reducing stiffness  $k_{xB}$ . The stability declines with lessening stiffness of the axle guidance, but in the case of soft axle guidance with interconnected wheelsets and damping of the coupling shaft, a stability comparable with the stiff axle guidance version can be achieved (see Figure 12.42b).

As the radial adjustment of the self-steering wheelsets in curves is achieved through creep forces in the contact between wheel and rail, the running characteristic is influenced by the conditions in the wheel–rail contact. The influence of wheel–rail contact geometry, tractive effort,



**FIGURE 12.42** Guiding force in curves (a) and critical speed (b) as a function of equivalent longitudinal axle guidance stiffness. Marked areas display recommended equivalent stiffness for the proposed versions of modular axle guidance (from Polach<sup>107</sup>).

and flange lubrication on the self-steering ability of a locomotive bogie has been analysed<sup>107</sup> and a comparison made between:

- Self-steering bogie with longitudinal very soft axle guidance and cross-coupling of the wheelsets by coupling shaft (CWD)
- Conventional bogie design with longitudinal stiff axle guidance (ST)

The results presented were calculated using a full nonlinear locomotive model in the simulation tool SIMPACK. Curve radii  $R = 300$  and  $500$  m and uncompensated lateral acceleration  $a_{lat} = 0.98$  m/s<sup>2</sup> were applied. In order to assess the curving performance, values on the outer wheel of the leading wheelset are presented for steady-state guiding force  $Y$  [kN] and wear index  $A_R$  [Nm/m].

The lateral rail profile and rail inclination has an influence on the self-steering ability of the wheelsets. Figure 12.43 demonstrates the guiding forces between the wheel profile S 1002 and the rail UIC 60. The guiding force is lower in the case of rail inclination 1:40 because the wheel profile S 1002 is optimised for this rail inclination. In order to estimate the influence of rail wear on self-steering, calculations were carried out on rail profiles which were identified from measurements as characteristic worn rail profiles in curves and used for calculation of wheel wear by Jendel<sup>101</sup> (see Figure 12.45). The heavily worn outer rail and small rolling radii difference lead to a reduction in the self-steering ability (see Figure 12.44). Even though the effectiveness of the self-steering is significantly lower, the self-steering bogie is still more favourable, particularly with regard to wear.

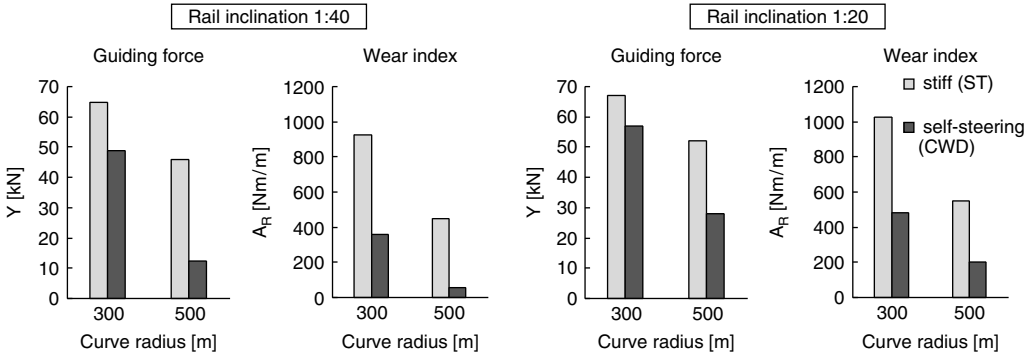


FIGURE 12.43 Influence of rail inclination on the guiding force and wear index.

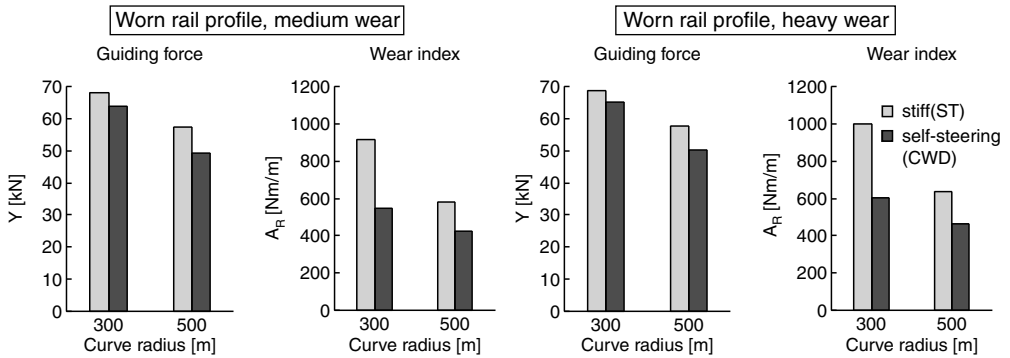


FIGURE 12.44 Influence of worn profile (Figure 12.45) on the guiding force and wear index. Results for new wheel profile — see Figure 12.43 (left diagram).

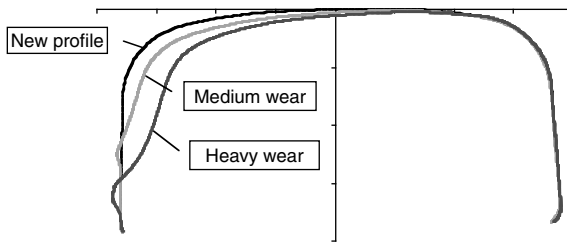


FIGURE 12.45 Worn rail profiles of high rail in curve as used in the sensitivity analysis.<sup>101</sup>

Figure 12.46 illustrates the influence of tractive effort on the examined values  $Y$  and  $A_R$ . With increasing tractive effort the creep forces between wheel and rail reach saturation point. The longitudinal creep forces incurred by the varying rolling radii difference reach lower values and self-steering is reduced. Wear at full tractive force is mainly caused by tractive creep and is therefore hardly influenced by self-steering.

The radial adjustment of the wheelsets will be slightly reduced by the influence of the wheel flange lubrication, so that the guiding force achieves a higher value than without lubrication. However, the wheel flange lubrication definitely has a positive effect on wear. As can be seen in Figure 12.47, the wear index demonstrates values which are approximately five times lower than

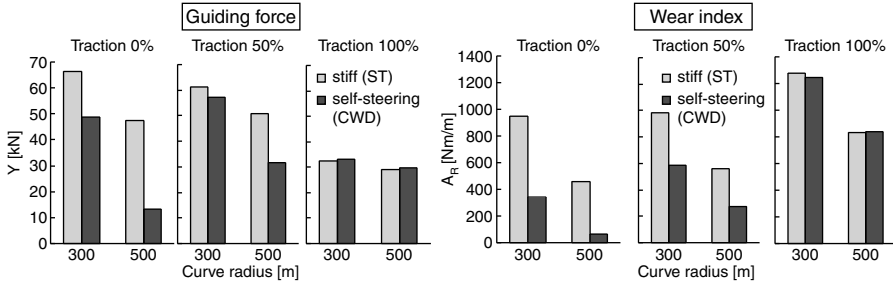


FIGURE 12.46 Influence of tractive effort on curving performance ( $a_{lat} = 1.1 \text{ m/s}^2$ ).

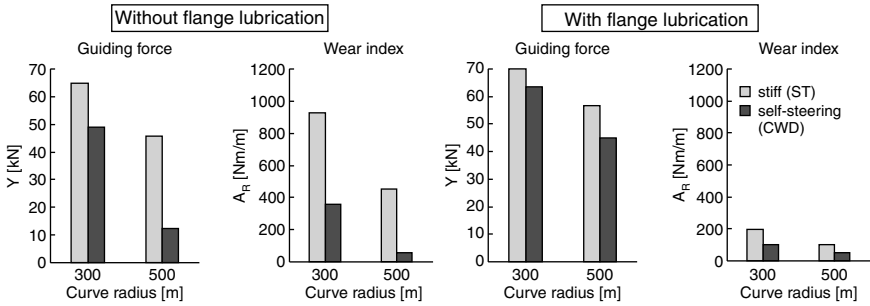


FIGURE 12.47 Influence of flange lubrication on curving performance (wheel–rail friction coefficient: tread 0.4, flange 0.1).

without lubrication. The utilisation of flange lubrication on self-steering bogies can lead to a further reduction in flange wear if a slight increase in the guiding force can be accepted.

The sensitivity analysis demonstrates the influence of operating conditions such as rail inclination, rail wear, tractive force, and wheel flange lubrication on the self-steering of the wheelsets. Despite the sensitivity to the effective operating conditions, the self-steering bogie generally achieves better characteristics and provides significant potential for savings in connection with maintenance of vehicle and infrastructure when compared with bogies with stiff axle guidance.

## H. RUNNING DYNAMICS UNDER THE INFLUENCE OF EXTERNAL LOADS

### 1. Influence of Crosswind

Crosswind safety of trains is a multidisciplinary subject that embraces the topics of aerodynamics and vehicle dynamics. In some cases a risk assessment is required that takes into account the combined effects of a train operating on a specific line. It is the most critical vehicle in a train that needs to be analysed. In a homogeneously composed train it is usually the leading car that is subjected to the largest aerodynamic loads.

The crosswind behaviour of a railway vehicle is determined by the following points, which however must be considered jointly and not individually:

- External vehicle design, characteristics of the running gears, track parameters such as radius, cant, and track quality
- Meteorological marginal conditions, in particular, the local occurrence of strong winds on the railway line

- Aggravating circumstances, such as running on high embankments and viaducts where winds are normally greater, and sudden gusts, for instance, when exiting a tunnel

A procedure to establish safe operation under the influence of strong crosswinds has been developed by, for example, the German Railways.<sup>108</sup> The procedure involves a comparison with reference vehicles and transportations, followed by assessment and derivation of measures in order to achieve equal or better performance than that of the reference standard. The following steps of the procedure are considered:

- Identification of the characteristic wind curves, which describe the maximum permissible wind speed as a function of the train speed and running conditions
- Analysis of the assigned tracks in order to determine the local permissible wind speed for running the vehicle at each point of the track, which results in the so-called operating curve
- Definition of wind occurrence frequency; meteorological investigations in order to evaluate the occurrence frequency of strong winds at each point on the track
- Cumulated occurrence frequency of events that exceeds the values of the wind characteristic curves; summation of the local occurrence of wind speed exceeding values of the characteristic wind curve over the total length of the track
- Comparison of the evaluated occurrence frequency of wind exceeding with the reference value identified using the identical method as applied for the reference case.

The assessment of the characteristic wind curves is carried out by means of multibody simulations, where the aerodynamic forces acting on the vehicle are either measured in a wind tunnel or calculated with computational fluid dynamics, respectively, cf. Ref. 109. The criteria that lead to the characteristic wind curves are based on the standard for vehicle acceptance test. However, some limit values are increased because of the exceptional situation.

Another probability analysis method for safety assessment in strong crosswinds applied on a high-speed line in Sweden is described in Ref. 110.

## 2. Influence of Coupler Forces

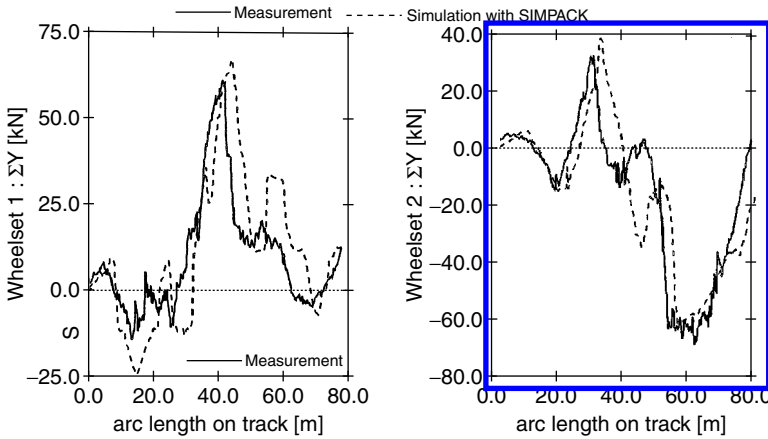
The longitudinal forces acting in long trains when pushing or during braking can lead to derailment risk in curves and in S-shaped curves, mainly in the case of long two-axle freight wagons. In order to determine the acceptable longitudinal force for the investigated wagon, a test in an S-shaped curve is specified in UIC 530-2.<sup>111</sup> The tested freight wagon, without load, should be placed between two so-called frame wagons which are loaded to attain an axle load of 20 t. A locomotive is coupled on each end of the composition in order to control the buffer forces and the speed.

The test track consists of an S-shaped curve with 150 m radius and 6 m intermediate straight track. The speed during the test should be 4 to 8 km/h. The criteria for identification of the acceptable longitudinal force are (in addition to the quotient  $Y/Q$ ):

- Sum of guiding forces (track shifting force)
- Wheel lift of nonguiding wheels over a distance longer than 2 m less or equal to 50 mm
- Minimum overlap of buffer plates of 25 mm
- Deformation of axle fixing plates

The tests described were simulated in Ref. 112 using the simulation tool SIMPACK and were compared with measurements. A model consisting of two frame wagons and tested freight wagon of type KIs442 was used. The simulation results and a comparison with measurement can be seen in [Figure 12.48](#) and [Figure 12.49](#). The speed is 2 m/s and the applied longitudinal force 195 kN. When



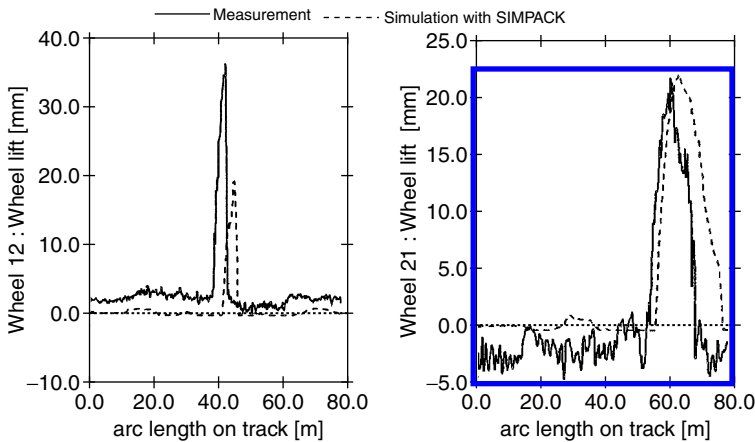


**FIGURE 12.48** Comparison simulation — measurement. Sum of lateral forces acting at wheelset 1 and wheelset 2 of the tested wagon at a longitudinal compression force of 195 kN (from Schupp and Jaschinski<sup>112</sup>).

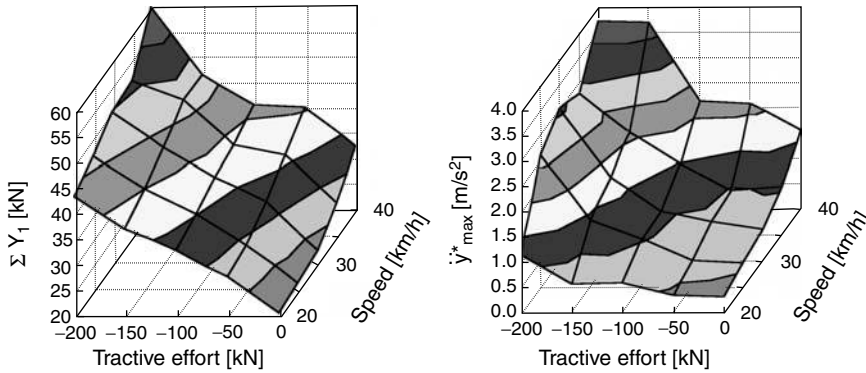
the formation passes over the experimental track, wheel lift occurs on both wheelsets. At first the left wheel of wheelset 1 loses contact after the train has traversed a distance of about 40 m. At this point wheel lift only takes place for a short period (see Figure 12.49). After covering a distance of about 55 m, the right wheel of wheelset 2 begins to lift-off and remains out of contact for nearly 15 m travel distance.

Figure 12.48 indicates a good agreement of the track shifting force between measured and simulated values. In Figure 12.49 a very good agreement can be seen between simulation and measurement of the wheel lift for the trailing wheelset but, in the case of the leading wheelset the agreement is not so positive. However, wheel lift is very sensitive to minor parameter changes and a large statistical variation is evident in the test results, even when the same longitudinal forces are applied. The investigation provides the possibility of simulating the influence of longitudinal load in long train compositions.

Pushing a train in a push–pull operation in commuter service causes a deterioration of the ride characteristics and the running safety in a series of turnouts and cross points (e.g., in a switch connection between parallel tracks). The consequences of pushing the train through an S-shaped



**FIGURE 12.49** Comparison simulation — measurement. Vertical wheel lift  $d_z$  at wheelset 1 left and wheelset 2 right of the tested wagon at longitudinal compression force of 195 kN (from Schupp and Jaschinski<sup>112</sup>).



**FIGURE 12.50** Track shifting force and the maximum lateral acceleration in the carbody during pushing the train through an S-shaped curve. Negative tractive effort = longitudinal compression in the train composition (from Polach<sup>113</sup>).

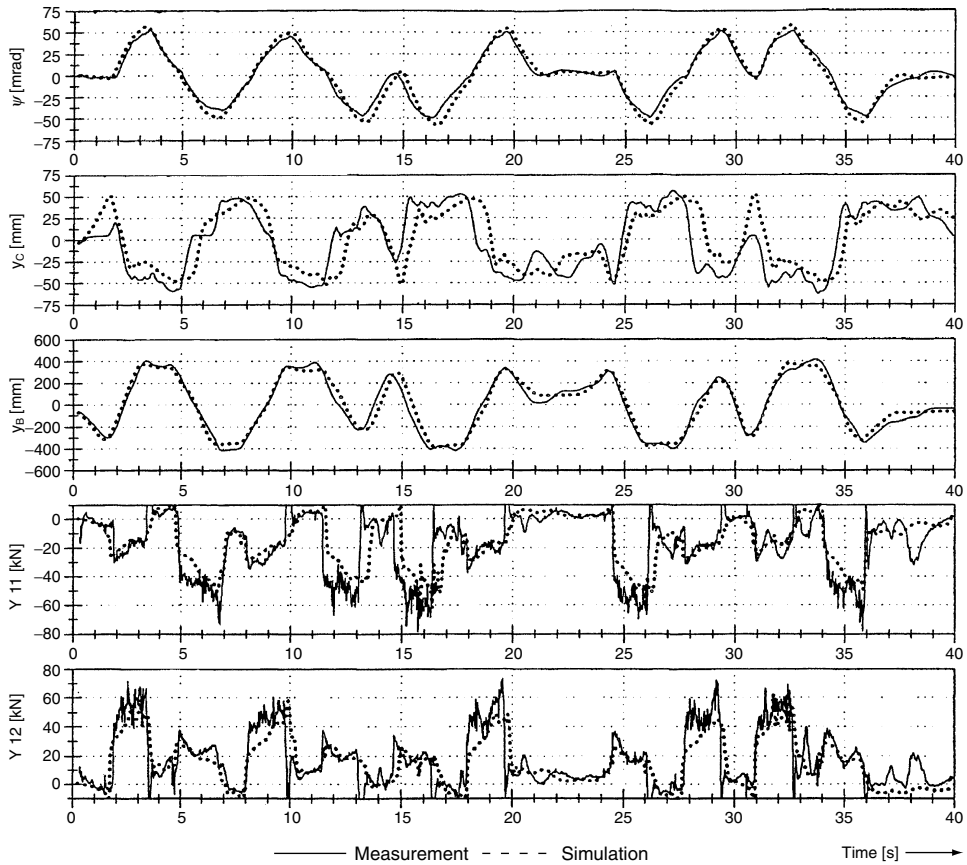
curve can be divided into steady-state and dynamic effects which, depending on the construction of the vehicle, can be of larger or lesser significance. The steady-state effect leads to an increase in the sum of the guiding forces on the leading wheelset of the leading bogie. The dynamic effect is associated with the movement of the carbody above the yawing bogie in a lateral direction from one side to the other. In the case of soft secondary lateral suspension, the carbody will be strongly accelerated and abuts on the other side with the lateral bump stop between the bogie and carbody, which in turn leads to peak of the lateral acceleration in the carbody.

Simulations and tests concerning this problem were carried out in the preparatory scope for commuter service of Intercity trains at SBB (Swiss Federal Railways). The simulations with the program VAMPIRE were realised with a model composed of four vehicles.<sup>113</sup> The push-pull forces were modelled in a similar manner to the running tests with preloading of the train composition with inversed longitudinal forces on the first and last vehicles. Figure 12.50 demonstrates the results of the peak value evaluation during a run through an S-shaped curve with radii of 150 m and length of the intermediate straight track of 6 m. The running characteristics deteriorate with increasing speed and increasing pushing force. The simulation results were proved by comparison with measurements. Figure 12.51 shows a comparison of the measured and calculated time diagrams for a test run in Zürich main station with ten turnouts in total, having radii of 160 to 300 m. The presented test run was carried out with a tractive effort of 150 kN and a speed of 40 km/h. Although the exact track irregularity was unknown and only an ideal track was assumed, the time diagrams demonstrate good agreement and the analysis was confirmed to be suitable for the engineering investigations.

### 3. Interaction between Vehicle and Traction Dynamics

The investigations of traction chain dynamics are usually limited to the rotational masses of traction chain, see e.g. Ref. 114. However, high adhesion utilisation and sophisticated vehicle dynamics design of modern traction vehicles demand complex simulations which, at the same time take into consideration the mechanical, electrotechnical, and control system fields. To carry out this kind of simulation, together with the complex modelling of systems mentioned, differing creep-force models used in traction dynamics and vehicle system dynamics have to be combined into one model. A suitable method has been described in Refs 30, 115 and 116.

In the following examples, a co-simulation of vehicle dynamics and traction control under adverse adhesion conditions are shown. The vehicle model is represented by the locomotive 12X of

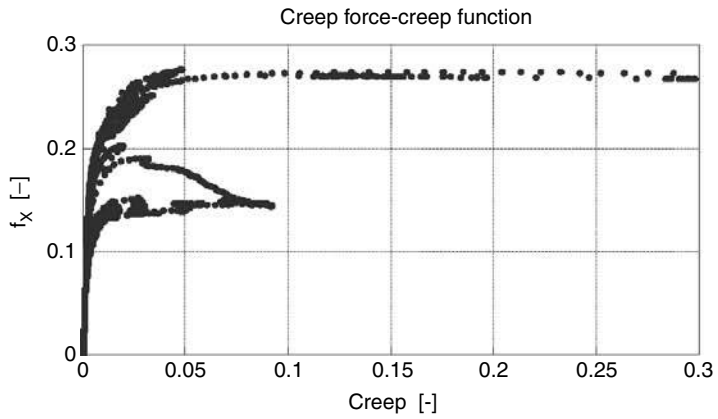
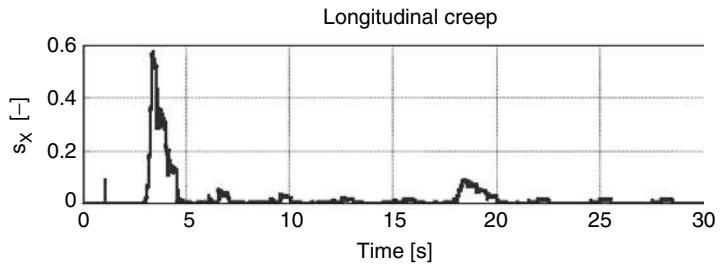
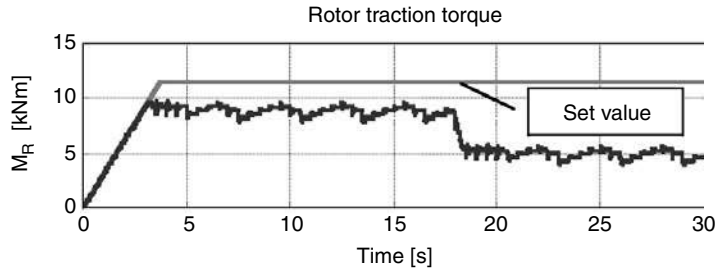
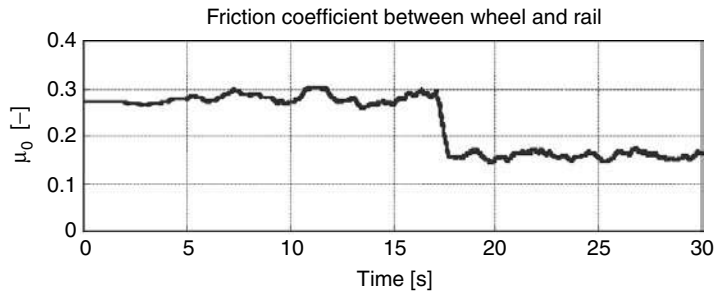


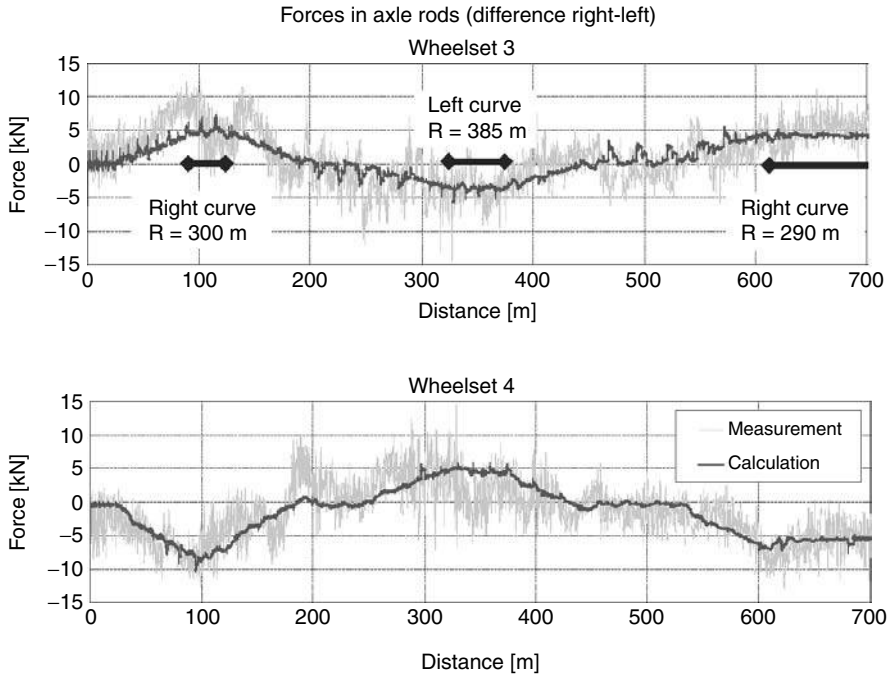
**FIGURE 12.51** Comparison of measurement and simulation of pushing the test train with effort of 150 kN through the turnouts in Zürich main station. The dimensions (top–down) are: yaw angle between carbody and bogie; lateral displacement between carbody and bogie; lateral buffer offset of neighbouring carbodies; guiding force on left wheel of the leading wheelset; guiding force on right wheel of the leading wheelset (from Polach<sup>113</sup>).

Bombardier modelled in the simulation tool SIMPACK. The MBS-model of the locomotive and one wagon (representing the hauled train) was extended to the full model of traction system. The controller was modelled in the computer code MATLAB-SIMULINK. During co-simulation, both programs are running in parallel while only exchanging a few channels.<sup>106</sup>

Figure 12.52 demonstrates a simulation of traction control reaction following a sudden worsening of adhesion conditions (see the input function of friction coefficient between wheel and rail in Figure 12.52). In addition to the sudden reduction of friction coefficient after approximately 30 m of track, small stochastic oscillations are superimposed. The torque on the rotor is unable to reach the set value due to adverse wheel–rail adhesion. The maximum achievable value is obtained by means of an adaptive traction controller. Following the sudden reduction of friction coefficient, the creep at first increases, but the controller stabilises the working point at a new adhesion optimum with low creepage after a short transition period.

Another co-simulation study shows starting and acceleration of a locomotive hauling a train on a curved sloping track. Figure 12.53 presents a comparison of measurement and simulation of longitudinal forces in wheelset linkages. The observed forces on the straight track and on the left and right curves are very close in measurement and simulation. The comparison validates the proposed method as suitable for computer simulation of traction vehicles running at adhesion limit.





**FIGURE 12.53** Comparison of calculated (dark line) and measured (nonfiltered) (light-coloured line) longitudinal forces in wheelset linkages during starting and acceleration of a locomotive hauling a train on a curved sloping track (from Polach<sup>30</sup>).

## VI. CONCLUSIONS

At present vehicle dynamics constitutes a significant and indispensable component of railway vehicle engineering. During the design period a prognosis of the running characteristics can be carried out with the aid of computer simulations. This enables a reduction in the extent of the test and shortening of the vehicle design period.

On the one hand, further developments to dynamic analyses in railway vehicle engineering evolve through new challenges, and on the other hand, by the broadening of objectives and possibilities. During the development of new railway vehicles the range of technical possibilities is being continuously extended and the technical limits exploited. From a dynamics standpoint, this tendency is accompanied by new challenges as well as new risks, for instance higher demands on material, out-of-roundness of wheels, rail corrugation, rolling contact fatigue, etc. Currently, new methods concerning reduction of the risks previously mentioned are the subject of intensive research.

The dynamic simulations also support a larger scope of development in other technical sectors. For instance, it is possible to take the whole dynamic behaviour of the vehicle or train composition into consideration for tasks such as investigation of loads acting on vehicle components or clearances between the components during the run on the track, crash behaviour, or aerodynamic analyses. Examples of such examinations have been tested and publicised and a broad industrial application will soon follow. The typical tasks of running dynamics will also evolve further, with the result that in future complex running tests can be virtually simulated and evaluated in accordance with measuring criteria.

The development of software tools for simulation of wheel–rail contact and the dynamic behaviour of railway vehicles on track has grown with the development of computing power. From the earliest analogue computers to the modern powerful digital processors, the equations that

govern the contact location, pressure distribution, and tangential creep forces have been developed and then coded into computer programs. These programs are now combined into a number of powerful and reliable computer simulation packages, and improvements have also been made in modelling the components of vehicle–track systems.

Dynamic analysis and simulations will continue to retain and further develop their well-established position in railway vehicle engineering.

## ACKNOWLEDGMENTS

The authors would like to thank their colleagues from the Rail Technology Unit at Manchester Metropolitan University, the Railway Group at KTH in Stockholm, and Bombardier Transportation in Winterthur for their help in preparing this chapter.

## NOMENCLATURE

$a_{\text{lat}}$ :	uncompensated lateral acceleration (cant deficiency)
$A_R$ :	wear index
$f_i$ :	carbody eigenfrequency, $i = 1, 2, 3 \dots$
$f_0$ :	frequency of unstable bogie oscillations
$F_x$ :	wheel–rail creep force in the longitudinal direction
$F_y$ :	wheel–rail creep force in the lateral direction
$k_x$ :	longitudinal axle guidance stiffness
$k_y$ :	lateral axle guidance stiffness
$k_{xB}$ :	equivalent longitudinal axle guidance stiffness for bending wheelset mode
$l_m$ :	wavelength of track irregularity critical for excitation of a carbody mode, $m = 1, 2, 3 \dots$
$m^+$ :	bogie mass
$p$ :	bogie pivot pin distance
$Q$ :	vertical wheel–rail contact force
$Q_0$ :	static vertical wheel–rail contact force
$R$ :	curve radius
$s_x$ :	longitudinal creep between wheel and rail
$s_y$ :	lateral creep between wheel and rail
$\dot{y}^+$ :	rms-value of lateral acceleration on the bogie frame
$\dot{y}^*$ :	rms-value of lateral acceleration in the carbody
$\dot{z}^*$ :	rms-value of vertical acceleration in the carbody
$s\Sigma Y$ :	rms-value of the sum of guiding forces (track shifting force)
$v$ :	vehicle speed
$v_{\text{cr}}$ :	nonlinear critical speed
$v_{\text{cr lin}}$ :	linear critical speed
$y$ :	lateral wheelset displacement
$\ddot{y}^+$ :	lateral acceleration on the bogie frame
$\ddot{y}^*$ :	lateral acceleration in the carbody
$Y$ :	guiding force (lateral wheel–rail contact force)
$\ddot{z}^*$ :	vertical acceleration in the carbody
$\varepsilon$ :	contact angle (contact slope) parameter
$\lambda$ :	equivalent conicity
$\sigma$ :	roll parameter
$\omega$ :	spin creep between wheel and rail
$\Theta$ :	moment around the normal to the wheel–rail contact
$\Sigma Y$ :	sum of guiding forces (track shifting force)

## REFERENCES

1. Andersson, E., Berg, M., and Stichel, S., *Rail Vehicle Dynamics*, Text Book, Division of Railway Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 2005.
2. ISO 2631-1: *Mechanical Vibration and Shock — Evaluation of Human Exposure to Whole-Body Vibration. Part 1: General requirements*, 2nd ed., 1997-05-01, corrected and reprinted 1997-07-15, ISO, Geneva, 1997.
3. Carlbom, P., *Structural Flexibility in a Rail Vehicle Car Body — Dynamic Simulations and Measurements*, TRITA-FKT Report 1998:37, Division of Railway Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 1998.
4. Carlbom, P., Passengers, seats and carbody in rail vehicle dynamics, *Vehicle Syst. Dyn.*, 37(Suppl.), 290–300, 2002, 17th IAVSD Symposium of Vehicles on Roads and Tracks, Copenhagen, Denmark, August 2001.
5. Chaar, N., *Structural Flexibility Models of Wheelsets for Rail Vehicle Dynamics Analysis — A Pilot Study*, TRITA-FKT Report 2002:23, Division of Railway Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 2002.
6. Chaar, N., Experimental and numerical modal analyses of a loco wheelset, *Vehicle Syst. Dyn.*, 41(Suppl.), 597–606, 2004, 18th IAVSD Symposium of Vehicles on Roads and Tracks, Atsugi, Kanagawa, Japan, August 2003.
7. Eickhoff, B. M., Evans, J. R., and Minnis, A. J., A review of modelling methods for railway vehicle suspension components, *Vehicle Syst. Dyn.*, 24(6–7), 469–496, 1995.
8. Berg, M., A three-dimensional airspring model with friction and orifice damping, *Vehicle Syst. Dyn.*, 33(Suppl.), 528–539, 2000, 16th IAVSD Symposium of Vehicles on Roads and Tracks, Pretoria, South Africa, August–September 1999.
9. Jönsson, P.-A. and Andersson, E., Influence of Link Suspension Characteristics on Freight Wagon Lateral Dynamics, *Sixth International Conference on Railway Bogies and Running Gear*, Budapest, Hungary, September 2004.
10. Fancher, P. S., Ervin, R. D., MacAdam, C. C., and Winkler, C. B., Measurement and Representation of the Mechanical Properties of Truck Leaf Springs, *SAE Paper No. 800905*, August 1980.
11. Berg, M., A non-linear rubber spring model for rail vehicle dynamics analysis, *Vehicle Syst. Dyn.*, 30, 197–212, 1998.
12. Nicolin, J. and Dellmann, T., Über die modellhafte Nachbildung der dynamischen Eigenschaften einer Gummifeder, *ZEV-Glas. Ann.*, 109(4)1985.
13. Evans, J. R., Rail vehicle dynamics simulations using VAMPIRE, In *The Manchester Benchmarks for Rail Vehicle Simulation*, Iwnicki, S., Ed., 1999, *Vehicle Syst. Dyn.* 31 (Suppl.).
14. Sundvall, P., *Comparisons between Predicted and Measured Ride Comfort in Trains — A Case Study on Modelling*, M.Sc. Thesis, TRITA-FKT Report 2001:19, Division of Railway Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 2001.
15. Xia, F., Modelling of wedge dampers in the presence of two-dimensional dry friction, *Vehicle Syst. Dyn.*, 37(Suppl.), 565–578, 2002, 17th IAVSD Symposium of Vehicles on Roads and Tracks, Copenhagen, Denmark, August 2001.
16. Claesson, S., *Modelling of Track Flexibility for Rail Vehicle Dynamics Simulation*, M.Sc. Thesis, Report TRITA AVE 2005:26, Division of Railway Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 2005.
17. Enblom, R. and Berg, M., Wheel Wear Modelling Including Disc Braking and Contact Environment — Simulation of 18 Months of Commuter Service in Stockholm, *14th Int. Wheelset Congress*, Orlando, Florida, USA, October 2004.
18. True, H., Does a Critical Speed for Railroad Vehicles exist?, RTD-Vol. 7, *Proceedings of the 1994 ASME/IEEE Joint Railroad Conference*, Chicago IL, American Society of Mechanical Engineers, pp. 125–131, March 22–24, 1994.
19. Goodall, R. M. and Iwnicki, S. D., Non-linear dynamic techniques v. equivalent conicity methods for rail vehicle stability assessment, *Vehicle Syst. Dyn.*, 41(Suppl.), 791–801, 2004.
20. Hertz, H., Über die Berührung zweier, fester elastischer Körper, *Journal für reine und angewandte Mathematik*, 92, 156–171, 1882.

21. Knothe, K. and Hung, L. T., Determination of the tangential stresses and the wear for the wheel–rail rolling contact problem, *Proceedings of the Ninth IAVSD Symposium*, Swets & Zeitlinger, Lisse, 1985, pp. 264–277, Linköping.
22. Pascal, J.-P. and Sauvage, G., New method for reducing the multicontact wheel/rail problem to one equivalent rigid contact patch, *Proceedings of the 12th IAVSD Symposium, Lyon, France*, Swets & Zeitlinger, Lisse, 1991.
23. Kik, W. and Piotrowski, J., A fast approximate method to calculate normal load at contact between wheel and rail and creep forces during rolling, *Proceedings of the Second Mini Conference, Contact Mechanics and Wear of Rail/Wheel systems*, Zobory, I., Ed., Budapest, TU, 1996.
24. Ayasse, J. B., Chollet, H., and Maupu, J. L., *Parametres Caracteristiques du contact roue-rail*, Rapport INRETS No. 225, Janvier 2000, ISBN 0768-9756.
25. Carter, F. W., The electric locomotive, *Proc. Inst. Civ. Engs.*, 221, 221–252, 1916.
26. Vermeulen, P. J. and Johnson, K. L., Contact of non-spherical elastic bodies transmitting tangential forces, *J. Appl. Mech. Trans. ASME*, 86, 338–340, 1964.
27. Kalker, J. J., *On the rolling of two elastic bodies in the presence of dry friction*, Doctoral Thesis, Delft University of Technology, 1967.
28. Haines, D. J. and Ollerton, E., Contact stress distributions on elliptical contact surfaces subjected to radial and tangential forces, *Proc. Inst. Mech. Engs.*, 177, 1963.
29. Shen, Z. Y., Hedrick, J. K., and Elkins, J. A., A comparison of alternative creep force models for rail vehicle dynamics analysis, *Proceedings of the Eighth IAVSD Symposium, Cambridge MA., August 1983*, Swets & Zeitlinger, Lisse, 2000, pp. 728–739.
30. Polach, O., Creep forces in simulations of traction vehicles running on adhesion limit, *Wear*, 258, 992–1000, 2005.
31. Kalker, J. J., *Book of tables for the Hertzian creep-force law*, Report 96–61, Delft University of Technology, Faculty of Technical Mathematics and Informatics, ISSN 0922-5641, 1996.
32. Kalker, J. J., A fast algorithm for the simplified theory of rolling contact, *Vehicle Syst. Dyn.*, 1–13, 1982, Swets & Zeitlinger B.V.-Lisse.
33. Brickle, B. V., *The steady state forces and moments on a railway wheelset including flange contact conditions*, Doctoral Thesis, Loughborough University of Technology, 1973.
34. Illingworth, R., *The mechanism of railway vehicle excitation by track irregularities*, Doctoral Thesis, University of Oxford, 1973.
35. Hobbs, A. E. W., *A Survey of Creep*, British Railways Technical Note DYN 52, 1967.
36. Bucher, F., Knothe, K., and Theiler, A., Normal and tangential contact problem of surfaces with measured roughness, *Wear*, 253, 204–218, 2002.
37. Knothe, K. and Gross-Thebing, A., Derivation of frequency dependant creep coefficients based on an elastic half-space model, *Vehicle Syst. Dyn.*, 15(3), 133–153, 1986.
38. de Possel, R., Beaufefoy, J., and Matsudaira, T., Papers awarded prizes in the competition sponsored by Office of Research and Experiment (ORE) of the International Union of Railways (UIC). ORE-Report RP2/SVA-C9, ORE, Utrecht, 1960.
39. de Pater, A. D., The approximate determination of the hunting movement of a railway vehicle by aid of the method of Krylov and Bogoljubov, *Appl. Sci. Res.*, 10, 205–228, 1961.
40. van Bommel, P., *Application de la theorie des vibrations nonlineaires sur le problem du mouvement de lacet d'un vehicule de chemin de fer*, Doctoral dissertation, Technische Hogeschool Delft, 1964.
41. Wickens, A. H., The dynamic stability of railway vehicle wheelsets and bogies having profiled wheels, *Int. J. Solids and Struct.*, 1, 319–341, 1965.
42. Müller, C. Th., Dynamics of railway vehicles on curved track, *Proc. Inst. Mech. Eng.*, 180(Pt3F), 45–57, 1965–1966.
43. Müller, C. Th., *Wear profiles of wheels and rails*, Office of Research and Experiment (ORE) of the International Union of Railways (UIC), ORE-Report C9/RP6, Utrecht, 1960.
44. Cooperrider, N. K., Hedrick, J. K., Law, E. H., Kadala, P. S., and Tuten, J. M., *Analytical and experimental determination of nonlinear wheel/rail geometric constraints*, Report FRA-O&RD 76-244, US Dept. of Transportation, Washington, 1975.
45. Schupp, G., Computational bifurcation analysis of mechanical systems with applications to railway vehicles, *Vehicle Syst. Dyn.*, 41(Suppl.), 458–467, 2004.
46. Schielen, W., Ed., *Multibody Systems Handbook*, Springer, Berlin, 1990.



47. Wallrapp, O. and Führer, C., MEDYNA — an interactive analysis program for geometrically linear and flexible multibody systems, In *Multibody Systems Handbook*, Schielen, W., Ed., Springer, Berlin, pp. 203–223, 1990.
48. Kik, W., Moelle, D., Bogo, C., and Ferrarotti, G., The Manchester benchmarks, ADAMS/Rail — MEDYNA statement of methods, *Vehicle Syst. Dyn.*, 31, 49–65, 1999.
49. Blader, F. B., and Klauser, P. E., *User's manual for NUCARS Version 1.0*, Report R-734, Association of American Railroads, Chicago, IL, September 1989.
50. Chollet, H., Ayasse, J. B., Fleuret, J. S., and Lévêque, E., A generalised conicity criterion for the wheel–rail contact, *Proceedings of the Eighth Mini Conference on Vehicle System Dynamics, Identification and Anomalies, TU Budapest 2002*.
51. Gimenez, J. G., Martin, L. M., Pascal, J. P., and Maupu, J. L., IAVSD Railway benchmark No.2 Sidive and Voco code solutions, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of the 12th IAVSD Symposium*, Sauvage, G., Ed., Swets and Zeitlinger, Lisse, pp. 551–565, 1992.
52. Evans, J. R., Rail vehicle dynamic simulations using VAMPIRE, *Vehicle Syst. Dyn.*, 31(Suppl.), 119–140, 1999, Swets & Zeitlinger, Lisse.
53. Andersson, E., Simulation von Spurkräften und Laufeigenschaften, *ZEV-Glasers Annalen*, 101(8/9), 339–347, 1977.
54. Persson, I., GENSYS in railway vehicle modelling. In Iwnicki, S., Ed., The Manchester Benchmarks for Rail Vehicle Simulation, *Vehicle Syst. Dyn.*, 31, Suppl., 1999.
55. Gilbert, B., Vampire — opportunities for fast, optimized railway simulations, *Proceedings of the 16th European ADAMS User Conference*, Berchtesgaden, 2001.
56. Knothe, K., Willie, R., and Zastra, B. W., Advanced Contact Mechanics — Road and Rail, *Vehicle Syst. Dyn.*, 53(4–5), 361–407, 2001.
57. Jaiswal, J., Blair, S., Stevens, A., Iwnicki, S., and Bezin, Y., *A Systems Approach to Evaluating Rail Life*, Railway Engineering 2002, Forde, M., Ed., Glasgow University, London, July 2002.
58. ERRI: B176/3 *Benchmark Problem — Results and Assessment*, B176/DT290, Utrecht, 1993.
59. Kortüm, W. and Sharp, R. S., The IAVSD Review Of Multibody Computer Codes For Vehicle System Dynamics, *Proceedings of the Third ASME Symposium on Transportation Systems*, ASME Winter Annual Meeting, Anaheim, CA, November 9–13, 1992.
60. Kortüm, W. and Sharp, R. S., Eds., *Multibody Computer Codes in Vehicle System Dynamics, Supplement to Vehicle System Dynamics*, Vol. 22, Swets & Zeitlinger, Lisse, 1993.
61. Iwnicki, S. D., Ed., Computer Simulation of Rail Vehicle Dynamics, *Vehicle System Dynamics*, Vol. 30, Numbers 3–4 September 1998.
62. Kortüm, W., Schielen, W. O., and Arnold, M., Software tools: from multibody system analysis to vehicle system dynamics, In *Mechanics for a New Millennium*, Aref, H. and Phillips, J., Eds., Kluwer Academic Publishers, Dordrecht, pp. 225–238, 2001.
63. Stribersky, A., Moser, F., Rulka, W., and Trautenberg, W., Advances in combined structural dynamics and system dynamics analysis of rail vehicles, *Vehicle Syst. Dyn.*, 37(Suppl.), 465–477, 2002.
64. Ofierzynski, M., Simulationstechnik bei der SIG Schweizerische Industrie-Gesellschaft, *ZEV + DET Glasers Annalen*, 118(2/3), 105–126, 1994.
65. Ofierzynski, M., Industrieller Einsatz der Simulationstechnik, Hochbruck, H., Knothe, K., Meinke, P., Ed., *Systemdynamik der Eisenbahn*, Fachtagung in Hennigsdorf am 13 und 14, Oktober 1994, Hestra-Verlag, pp. 171–180, 1994.
66. UIC Code 518: *Testing and Approval of Railway Vehicles from the Point of View of their Dynamic Behaviour–Safety–Track Fatigue–Ride Quality*, International Union of Railways, 3rd ed., Paris, October 2005.
67. EN 14363: *Railway Applications — Testing for the Acceptance of Running Characteristics of Railway Vehicles — Testing of Running Behaviour and Stationary Tests*, CEN, Brussels, June 2005.
68. ORE Frage B 55: *Entgleisungssicherheit von Güterwagen in Gleisverwindungen*, Bericht Nr. 8, Utrecht, April 1983.
69. 49CFR213, FRA Regulations, Title 49: *Transportation, Part 213 — Track safety standards*, Code of Federal Regulations (CFR), Federal Railroad Administration. Internet: <http://www.gpoaccess.gov/cfr>
70. Stichel, S. and Knothe, K., *Ein Konzept für die Einführung des Betriebsfestigkeitsnachweises bei Schienenfahrzeugen aufgrund von Simulationsrechnungen*, VDI-Forschritt Berichte, Reihe 12, No. 234, VDI-Verlag, Düsseldorf, Germany, 1995.

71. Bieker, G. and Flach, M., The use of simulations for calculating load spectra for railway vehicles, *RTR Railway Technical Review*, 10–14, No. 2, 1998.
72. Rottler, A. and Eschenauer, H. A., Optimal layout of bogie frames considering fatigue strength, In *Proceedings of DETC '98, ASME Design Engineering Technical Conference*, September 13–16, Atlanta, GA, 1998.
73. Knothe, K. and Stichel, S., *Schienefahrzeugdynamik*, Springer, Berlin, Heidelberg, New York, 2003.
74. Nielsen, J. C. O. and Johansson, A., Out-of round railway wheels — a literature survey, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 214(F2), 79–91, 2000.
75. Grassie, S. L. and Kalousek, J., Rail corrugations: characteristics, causes and treatments, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 207(F1), 57–68, 1993.
76. Ekberg, A. and Kabo, E. Rolling contact fatigue of railway wheels and rails — an overview, *Proceedings of the 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems (CM2003) in Gothenburg*, Sweden, Appendix II, pp. 11–24, June 10–13, 2003.
77. Ågren, J. and Enblom, R., Wheel damages on the Regina train and the need for further research — a summary, *Proceedings of the 6th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems (CM2003) in Gothenburg*, Sweden, pp. 105–112, June 10–13, 2003.
78. Evans, J. R. and Dembosky, M. A., Investigation of vehicle dynamic influence on rolling contact fatigue on UK railways, *Vehicle Syst. Dyn.*, 41(Suppl.), 527–536, 2004.
79. Mauer, L., The modular description of the wheel to rail contact within the linear multibody formalism, In *Advanced Railway Vehicle System Dynamics*, Kisilowski, J. and Knothe, K., Eds., Wydawnictwa Naukowo-Techniczne, Warsaw, pp. 205–244, 1991.
80. Wickens, A. H., *Fundamentals of Rail Vehicle Dynamics: Guidance and Stability*, Swets & Zeitlinger Publishers, Lisse, 2003.
81. ORE B176: Bogies with steered or steering wheelsets, Report No. 1: Specifications and preliminary studies, Vol. 2, Specification for a bogie with improved curving characteristics, ORE, Utrecht 1989.
82. Müller, R., Schienenkopfpflege aus den Gesichtspunkten des Zusammenwirkens Rad/Schiene, In *Proceedings of Fifth International Conference on Railway Bogies and Running Gears*, Zobory, I., Ed., Budapest, Hungary, pp. 199–215, 24–26 September, 2001.
83. Kolbe, Th., Schwingungstechnisches Verhalten von Schienenfahrzeugen unter dem Einfluss der Berührung Rad/Schiene, *Proceedings of Sixth International Railway Vehicle Conference*, Dresden, Germany, pp. 35–37, 8–10 October 2003.
84. Mahr, A., *Abstimmung der Rad-Schiene-Geometrie mit dem lateralen Fahrzeugverhalten: Methoden und Empfehlungen*, Thesis, Shaker Verlag GmbH, Aachen 2002.
85. Stichel, S., On freight wagon dynamics and track deterioration, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 213(F4), 243–254, 1999.
86. Ofierzynski, M. and Brundisch, V., Fahrkomfort von Schienenfahrzeugen — Die Zuverlässigkeit moderner Simulationstechnik, *ZEV + DET Glasers Annalen*, 124(2/3), 109–119, 2000.
87. Krettek, O. and Miluczky, A., Zum Einfluss der Dachdämpfer auf den Lauf von Niederflur-Multigelenkstadtbahnwagen, In *Proceedings of Fifth International Conference on Railway Bogies and Running Gears*, Zobory, I., Ed., Budapest, Hungary, pp. 279–291, 24–26 September, 2001.
88. Fujimoto, H. and Miyamoto, M., Measures to reduce the lateral vibration of the tail car in a high speed train, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 210(F2), 87–93, 1996.
89. True, H., On the theory of nonlinear dynamics and its application in vehicle systems dynamics, *Vehicle Syst. Dyn.*, 31, 393–421, 1999.
90. Polach, O. and Vetter, A., Methods for running stability prediction and their sensitivity to wheel/rail contact geometry, In *Extended abstracts of the Sixth International Conference on Railway Bogies and Running Gears*, Budapest, pp. 62–64, 13–16 September 2004.
91. Knothe, K. and Böhm, F., History of stability of railway and road vehicles, *Vehicle Syst. Dyn.*, 31, 283–323, 1999.
92. Orlova, A., Boronenko, Y., Scheffel, H., Fröhling, R., and Kik, W., Tuning von Güterwagendrehgestellen durch Radsatzkopplungen, *ZEVrail Glasers Annalen* 126, Tagungsband SFT Graz 2002, pp. 200–212, 2002.
93. Frederich, F., Die Gleislage — aus fahrzeugtechnischer Sicht, *ZEV Glasers Annalen*, 108(12), 355–362, 1984.

94. UIC Code 513: *Guidelines for Evaluating Passenger Comfort in Relation to Vibration in Railway Vehicles*, 1st ed., 1.7. 1994, International Union of Railways, Paris 1995.
95. ENV 12299: *Railway applications — Ride comfort for passengers — Measurement and evaluation*, European prestandard, CEN, Brussels, February 1999.
96. Kim, Y. G., Kwon, H. B., Kim, S. W., Park, C. K., and Park, T. W., Correlation of ride comfort evaluation method for railway vehicles, *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit*, 217, 73–88, 2003.
97. Diana, G., Cheli, F., Collina, A., Corradi, R., and Melzi, S., The development of a numerical model for railway vehicles comfort assessment through comparison with experimental measurements, *Vehicle Syst. Dyn.*, 38, 165–183, 2002.
98. Götsch, M. and Sayir, M., Simulation of riding comfort of railway vehicles, *Vehicle Syst. Dyn.*, 37(Suppl.), 630–640, 2002.
99. Pearce, T. G. and Sherratt, N. D., Prediction of wheel profile wear, *Wear*, 144, 343–351, 1991.
100. Zobory, I., Prediction of wheel/rail profile wear, *Vehicle Syst. Dyn.*, 28, 221–259, 1997.
101. Jendel, T., Prediction of wheel profile wear — comparisons with field measurements, *Wear*, 253, 89–99, 2002.
102. Wickens, A. H., Steering and stability of the bogie: vehicle dynamics and suspension design, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 205, 109–122, 1991.
103. Bell, C. E. and Hedrick, J. K., Forced steering of rail vehicles: stability and curving mechanics, *Vehicle Syst. Dyn.*, 10, 357–386, 1981.
104. Scheffel, H., Unconventional bogie designs — their practical basis and historical background, *Vehicle Syst. Dyn.*, 24, 497–524, 1995.
105. Polach, O., Coupled single-axle running gears — a new radial steering design, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 216(F3), 197–206, 2002.
106. Polach, O., Optimierung moderner Lok-Drehgestelle durch fahrzeugdynamische Systemanalyse, *El Eisenbahningenieur*, 53(7), 50–57, 2002.
107. Polach, O., Curving and stability optimisation of locomotive bogies using interconnected wheelsets, *Vehicle Syst. Dyn.*, 41(Suppl.), 53–62, 2004.
108. Matschke, G., Grab, M., and Bergander, B., Nachweis der Sicherheit im Schienenverkehr bei extremem Seitenwind, *ETR Eisenbautechnische Rundschau*, 51(4), 200–206, 2002.
109. Diedrichs, B., On computational fluid dynamics modelling of crosswind effects for high-speed rolling stock, *Proc. Inst. Mech. Eng., Part F, J. Rail Rapid Transit*, 217, F3 pp. 203–226, 2003.
110. Andersson, E., Häggström, J., Sima, M., and Stichel, S., Assessing cross-wind safety of trains, *Proceedings of the International Symposium on Speed-up and Service Technology for Railway and Maglev Systems 2003 (STECH'03)*, Tokyo, pp. 293–300, August 19–22 2003.
111. UIC-Code 530-2: *Wagons: Running safety*, 4th ed., 1.7.1985, Reprint 1.7.1997, International Union of Railways, Paris, 1999.
112. Schupp, G. and Jaschinski, A., Investigation of running behaviour and derailment criteria for freight cars by dynamic simulation, In *COMPRAIL 96, Computers in Railways V*, Vol. 2, Computational Mechanics Publications, Southampton, UK, pp. 13–22, 1996.
113. Polach, O., Laufverhalten beim Schieben des Zuges — Simulation und Versuch, *Proceedings of the 12th International Conference Current Problems in Rail Vehicles*, Žilina, pp. 141–152, 1995.
114. Gold, P.W., Schelenz, R., Holzapfel, M., and Bühren, M., Möglichkeiten zur Optimierung mechanisch-elektrischer Antriebssysteme durch Simulation, VDI Bericht 1346, *Mechanisch-Elektrische Antriebstechnik 1997*, VDI-Verlag, Germany, 1997.
115. Polach, O., Rad-Schiene-Modelle in der Simulation der Fahrzeug- und Antriebsdynamik, *Elektrische Bahnen*, 99(5), 219–230, 2001.
116. Polach, O., Influence of locomotive tractive effort on the forces between wheel and rail, In *ICTAM, Selected papers from the 20th International Congress of Theoretical and Applied Mechanics held in Chicago*, 28 August–1 September 2000, *Vehicle Syst. Dyn.*, 35(Suppl.), 7–22, 2001.

---

# 13 Field Testing and Instrumentation of Railway Vehicles

*Julian Stow and Evert Andersson*

## CONTENTS

I.	Introduction .....	424
A.	Reasons for Testing.....	424
II.	Common Transducers .....	425
A.	Displacement Transducers .....	425
B.	Accelerometers.....	426
1.	Piezoelectric Accelerometers .....	427
2.	Capacitive Accelerometers.....	429
C.	Strain Gauges .....	430
1.	Bridge Circuits .....	432
D.	Force-Measuring Wheelsets.....	433
1.	Measuring Lateral Forces between Wheelset and Axle Box .....	435
2.	Measuring Lateral and Vertical Wheel–Rail Forces — The Axle Method....	435
3.	Measuring Lateral and Vertical Wheel–Rail Forces — Wheel Methods .....	436
4.	Compensation for Undesired Parasitic Effects .....	437
E.	Vehicle Speed and Position Measurement .....	438
1.	The AC Tachogenerator.....	438
2.	Hall Effect Probes .....	439
3.	Ground Speed Radar .....	439
4.	Determining Vehicle Position.....	440
III.	Test Equipment Configuration and Environment.....	441
A.	Transducer Positions on Vehicles.....	442
IV.	Data Acquisition.....	443
V.	Measurement of Wheel and Rail Profiles.....	445
VI.	Track Geometry Recording .....	447
A.	Manual Survey .....	448
B.	Track Geometry Trolley .....	449
C.	Track Recording Vehicles .....	449
D.	Chord Offset Measuring Systems .....	451
VII.	Examples of Vehicle Laboratory and Field Tests.....	452
A.	Static/Quasi-Static Tests .....	452
1.	Wheel Unloading Test.....	452
2.	Bogie Rotational Resistance Test .....	452
3.	Sway Test .....	454

4. Body Modes Tests.....	455
B. Dynamic Tests.....	455
References.....	456

## I. INTRODUCTION

An understanding of testing and instrumentation methods is essential to the accurate construction and validation of railway vehicle dynamic models. The dynamics engineer may need to produce specifications for test work, understand the applicability and limitations of data produced, and manipulate test results to provide comparisons with modelling work. This chapter provides an overview of the situations in which the engineer may require test data, together with an introduction to common techniques and equipment used, both in the laboratory and for conducting vehicle testing on-track. The examples given largely relate to vehicle testing which forms the most relevant body of work for the vehicle dynamics engineer. However, dynamic simulation is increasingly used in rail/track related investigations where many of the same techniques may be applied.

### A. REASONS FOR TESTING

During the development of a new vehicle or modification of an existing one, requirements for test work may arise for a number of different reasons:

*Component characterisation* — tests may be required to establish the properties of the various components that make up the suspension in order to allow the initial construction of a model. Such tests are normally carried out in the laboratory using small to medium sized test machines or dedicated test rigs.

*Determination of parasitic or secondary effects* — once assembled vehicles (particularly modern passenger vehicles) can exhibit behaviour that is difficult to predetermine from the individual suspension components. These parasitic effects typically arise from the summation of a number of small stiffness contributions from components such as anti-roll bars, traction centres, and lateral and yaw dampers in directions other than those in which they are mainly designed to operate. Other effects, which may need to be quantified arise from flexibilities in mounting brackets or similar, as well as internal flexibility in dampers. Such tests are normally carried out statically or quasi-statically on a complete vehicle in the laboratory.

*Structural testing* — the testing of vehicle body structures and bogie frames for strength, fatigue life, and crashworthiness is a complete subject in itself and beyond the scope of this chapter. However, the dynamicist may need to obtain parameters to enhance models, particularly with respect to simulation of ride and passenger comfort. Typical examples include the vertical and lateral bending modes of vehicle bodies and the torsional stiffness of bogie frames. Tests are most often carried out in the lab using bare body shells or bogie frames mounted in dedicated structural test rigs.

*Validation testing* — it is generally necessary to increase confidence in the correct operation of models by comparing the results with those from a series of tests. Such tests may be on bogies or complete vehicles. At a basic level, these may be carried out quasi-statically in the laboratory, but any extensive validation is likely to require on-track tests under a range of conditions to fully understand the dynamic behaviour of the vehicle. The level of validation required will ultimately depend upon the intended use of the models.

*Acceptance tests* — all railway administrations require new or modified vehicles to undergo a series of tests to demonstrate safe operation for various conditions. Such tests may be specific to an individual company or country or, as in the case of European Standards,<sup>1,2</sup> may allow a vehicle to operate across a number of countries. The exact requirements for these can vary widely but will usually comprise of a mixture of lab and field tests. Many administrations now allow some of these requirements to be met by simulation of the test procedure using a suitably validated vehicle

dynamics model. In any case, simulation of these tests forms a common part of vehicle development to ensure that proposed designs meet the required standards. As such tests will be carried out on all vehicles accepted for service, they may also provide a useful source of information to validate models of existing vehicles.

In addition, the dynamics engineers may also be involved with testing to assess performance against specified criteria such as passenger comfort or to investigate problems with existing rolling stock.

*Reproducing track geometry* — many simulation tasks will require the use of “real” track geometry measured by a high-speed recording vehicle or hand operated trolleys. Although such data is generally presented as “ready to use,” experience has shown that an appreciation of the measuring systems and instrumentation used is vital to ensure that an accurate reconstruction of the track geometry can be obtained.

*Measuring wheel and rail profiles* — accurate representations of worn wheel and rail profiles are vital to understanding vehicle (and track) behaviour. A number of proprietary devices are available to measure profiles, however, as with track geometry, accurate results will be aided by an understanding of the principles behind their operation.

## II. COMMON TRANSDUCERS

This section provides a brief overview of the range of transducers commonly encountered to measure displacement, acceleration, and force.

### A. DISPLACEMENT TRANSDUCERS

These are used for measuring linear or rotational displacements. The most common type of transducer is the linear variable differential transformer (LVDT). This comprises a transformer with a single primary coil and two secondary coils wound onto a hollow cylindrical tube as shown in Figure 13.1. Within this tube, a ferromagnetic core can move up and down. The primary coil at the centre of the tube is excited with an AC signal and this induces a voltage in the secondary coils. The secondary coils are normally connected as shown in Figure 13.2. This arrangement, known as “series opposition,”<sup>3</sup> has the effect of producing zero output voltage with the core in its central or zero position. As the core is moved, the coupling between the primary and one of the secondary coils increases whilst the coupling with the other secondary coil decreases in direct proportion. With correct arrangement of the coils and core, the resulting output voltage will be linear over the majority of the stroke. It should be noted that as the core moves past the zero position (central on the primary coil), the output voltage undergoes a 180° phase shift.

In practice, a transducer that requires AC input and produces AC output is inconvenient, so a signal processing module is used in conjunction with the LVDT. This senses the zero-passing phase shift described above and uses this to distinguish between AC signals of equal amplitude either side of the zero position. The resulting conditioned output is therefore a positive or negative DC voltage either side of the zero position. The signal conditioning module usually also converts a DC supply



**FIGURE 13.1** LVDT (courtesy of RDP Electronics). *Source:* From RDP Electronics Ltd. — Catalogue. With permission.

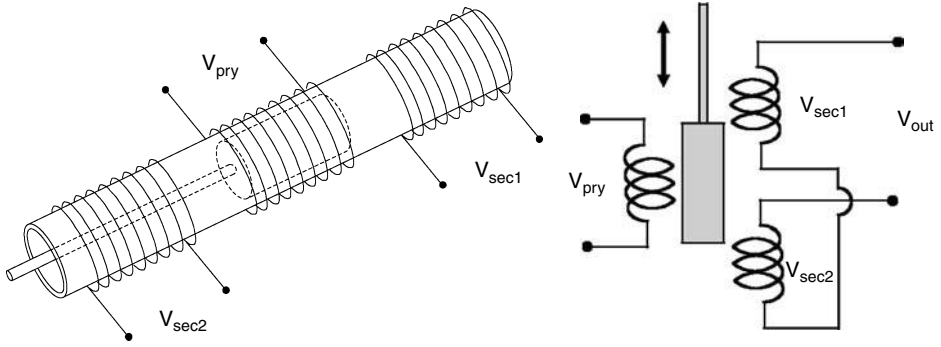


FIGURE 13.2 Schematic of primary and secondary coils in an LVDT.

voltage into the required AC excitation for the primary coil. The signal conditioning may be in a separate module but is often incorporated within the transducer casing itself.

LVDTs have the advantage of being inherently non-contact devices and therefore have no wearing parts. They typically achieve better than  $\pm 1\%$  linearity over their specified range and are available commercially in measuring ranges from a few millimetres up to approximately 0.5 m. LVDTs generally operate on input voltages up to 24 V DC and may be obtained with floating cores or with a sprung loaded plunger.

### B. ACCELEROMETERS

Accelerometers are electromechanical transducers that convert vibration into an electrical signal. Unlike displacement and velocity, acceleration can be measured as an absolute, rather than relative, quantity. This factor combined with the accuracy, robustness, and good frequency response/sensitivity of modern accelerometers makes them ideal for use in vehicle dynamics test applications.

Figure 13.3 shows a simplified accelerometer. A mass (the seismic mass) is mounted within a rigid casing on a spring and damper. Accelerometers are designed such that the natural frequency of the seismic mass is high compared to the desired measuring frequency range. In such an

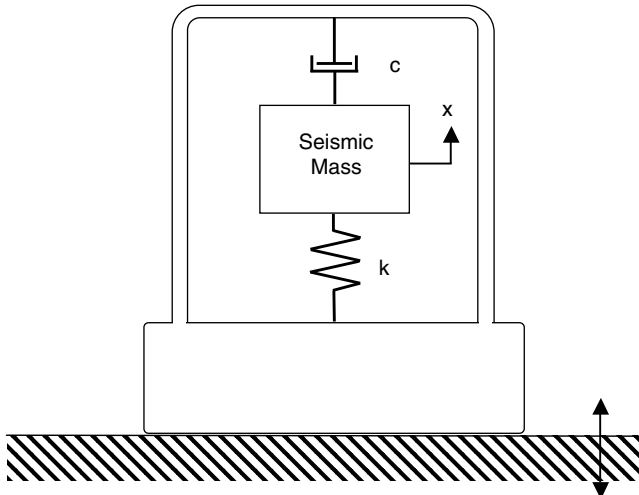


FIGURE 13.3 Simplified accelerometer.

arrangement, the amplitude of displacement of the seismic mass will be directly proportional acceleration exciting the transducer. It follows, therefore, that accelerometers work by sensing the relative displacement of the seismic mass with respect to the transducer casing. It can be shown, mathematically,<sup>3</sup> that the maximum useful frequency range of an accelerometer, around 20–30% of the transducers natural frequency, is achieved with a damping ratio of 0.7. This damping ratio also provides almost zero phase distortion.

### 1. Piezoelectric Accelerometers

The most commonly used type of accelerometer is the piezoelectric accelerometer. The sensing element in such devices is a slice or disc of piezoelectric material. Such materials develop an electrical charge when they are subjected to mechanical stress. A number of naturally occurring materials exhibit this effect (e.g., quartz), but transducers typically employed man-made materials of a family known as “ferroelectric ceramics.”<sup>4</sup>

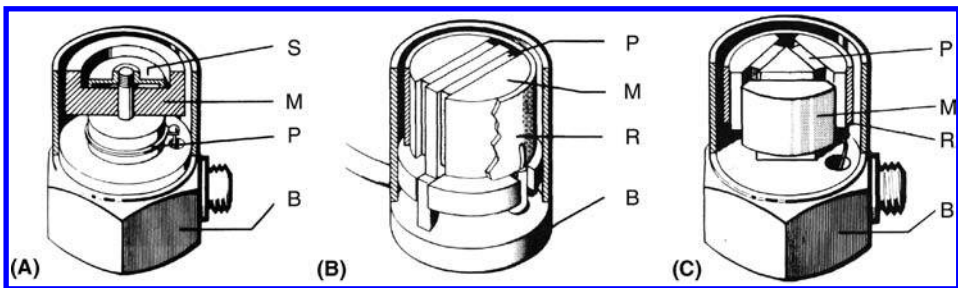
Practical accelerometer designs typically employ a seismic mass resting upon, or suspended from, a number of slices of the piezoelectric material. The vibration of the seismic mass within the accelerometer exerts a force on the piezoelectric material and a charge is developed that is proportional to the force exerted. Three common designs of accelerometer are illustrated in Figure 13.4 below.

The centre mounted compression design is a relatively simple arrangement where the mass is mounted on a centre pillar with a spring to provide the preloading. The mass acts in compression on the piezoelectric element. These designs have the advantage of good useable bandwidth. However, as the base and centre pillar act as a stiffness in parallel with the piezoelectric element, any bending of the base or thermal expansion can cause erroneous readings.

Planar shear designs feature two slices of piezoelectric material either side of the centre post, each having a seismic mass attached to it. The masses are held in place by a clamping ring which preloads the piezoelectric elements and results in a high degree of linearity. The charge induced by the shear forces acting on the piezoelectric elements is collected between the housing and the clamping ring. In this design, the sensing elements are effectively isolated from the base and these designs therefore have good resistance to base strains and temperature variations.

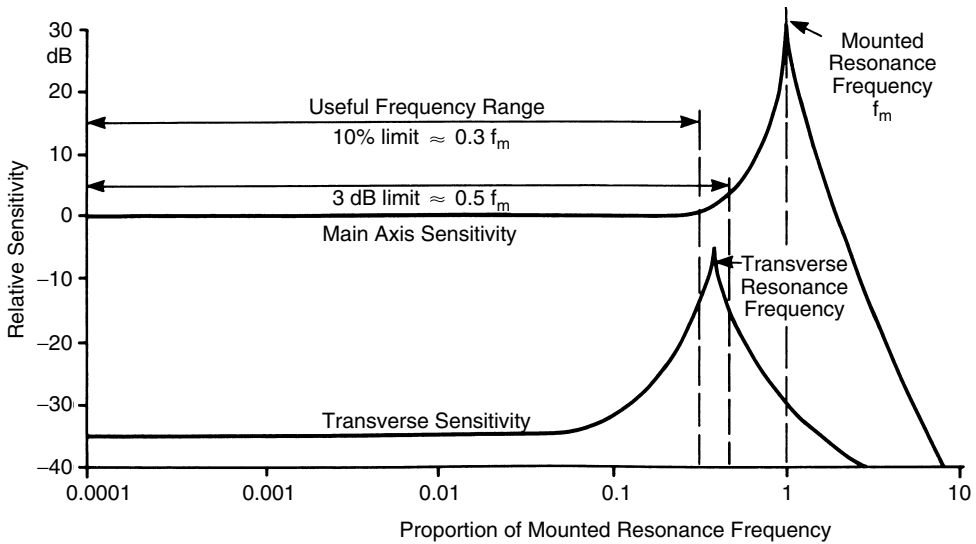
The delta shear design is similar to the planar shear version described above. In this case, three masses and piezoelectric slices are mounted radially to the centre pillar at 120° to each other. Once again, a clamping ring preloads the elements. In addition to good resistance to base strain and temperature changes, these designs also have high resonant frequency and sensitivity.

An understanding of the useable bandwidth of a device is vital when selecting the correct accelerometer for a vehicle test application. The typical frequency response of a piezoelectric accelerometer is shown in Figure 13.5. As described above, the upper frequency limit for the device



**FIGURE 13.4** Common accelerometer designs, (a) centre mounted compression, (b) planar shear and (c) delta shear® (courtesy of Bruel and Kjaer). *Source:* From Serridge, M. and Licht, T. R., *Piezoelectric Accelerometers and Vibration Preamplifier Handbook Bruel and Kjaer*, Revised November 1987. With permission.

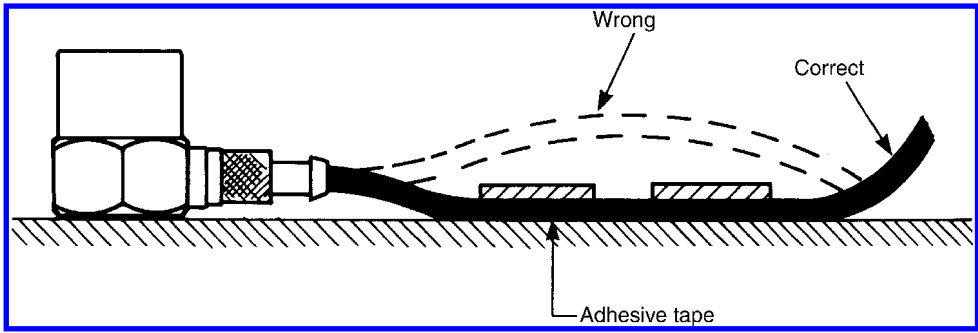




**FIGURE 13.5** Frequency response of an accelerometer (courtesy of Bruel and Kjaer). *Source:* From Serridge, M. and Licht, T. R., *Piezoelectric Accelerometers and Vibration Preamplifier Handbook Bruel and Kjaer*, Revised November 1987. With permission.

will be dictated by the resonant frequency of the accelerometer. A commonly used rule of thumb is that the upper frequency limit should be no more than one third of the resonant frequency. For general purpose piezoelectric accelerometers, the resonant frequency may be of the order of 20 kHz, putting their upper useable limit way above anything likely to be required for a vehicle dynamics application. Piezoelectric accelerometers have one very important limitation with regard to vehicle dynamics test applications. The lower frequency limit is determined not by the accelerometer itself but by the RC time constant of the charge amplifier used to condition the signal from the transducer. Whilst it is possible to sense very low frequencies using preamplifiers with very high impedance, general purpose equipment may limit the lower useable frequency limit to 1–3 Hz. Clearly, this has important implications for the vehicle dynamics engineer. A general purpose piezoelectric accelerometer may be quite acceptable for mounting on unsprung masses such as the axle box, but may be operating near its lower limit when mounted on a bogie frame. Vehicle body modes may occur at 0.5 Hz or less and therefore would be below the lower limit of a general purpose piezoelectric accelerometer. Piezoelectric accelerometers are generally not capable of measuring the quasi-static accelerations due to curving. In the case of bogie and body measurements, the capacitive accelerometer, described in Section II.B.2 below, will provide a solution to this problem.

An ideal accelerometer design would only respond to vibrations applied to the main sensing axis. However, in practice, most accelerometers will exhibit some sensitivity to excitations at  $90^\circ$  to the main axis, known as the transverse sensitivity. These are caused by small irregularities in the piezoelectric material causing the axis of maximum sensitivity to be slightly misaligned with the operating axis of the accelerometer. It can be imagined, therefore, that the transverse sensitivity will not be constant and subsequently there will be directions of maximum and minimum sensitivity at  $90^\circ$  to each other. Some accelerometer designs will indicate the direction of minimum transverse sensitivity on the accelerometer body to aid correct mounting of the device. It is generally found that the transverse resonant frequency is lower than the main resonant frequency and therefore falls within the useable bandwidth of the accelerometer. However, at the relatively low frequencies of interest to the vehicle dynamics engineer, the maximum transverse sensitivity is usually less than 4% of the main axis sensitivity.<sup>4</sup>



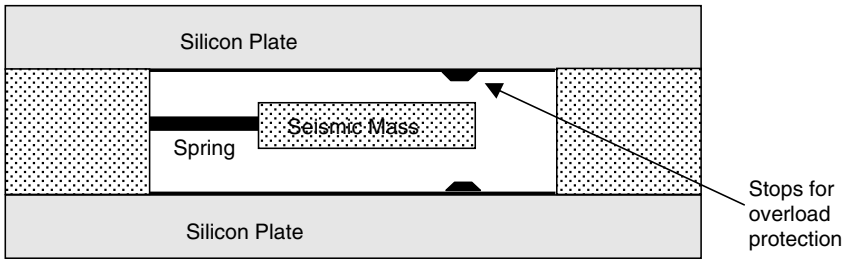
**FIGURE 13.6** Accelerometer mountings and cabling (courtesy of Bruel and Kjaer). *Source:* From Serridge, M. and Licht, T. R., *Piezoelectric Accelerometers and Vibration Preamplifier Handbook Bruel and Kjaer*, Revised November 1987. With permission.

To achieve reliable results, care should be taken when mounting and cabling piezoelectric accelerometers. The device may be mounted using either a stud screwing into a hole tapped directly into the test component or by gluing it in place onto a clean flat surface. However, in many railway vehicle applications the accelerometer will be mounted on its stud to a bracket which will in turn be bolted or clamped to the vehicle. In this case, the mounting bracket and clamping arrangement should be as rigid as possible to ensure that the measured data is not degraded by vibrations or deflection of the mounting itself. A thin mica washer is often placed between the base of the accelerometer and the mounting which, when used in conjunction with an insulated stud, increases the electrical isolation of the accelerometer from the test piece. It should be noted that dropping an accelerometer onto a hard surface, such as a workshop floor, can cause a shock load that exceeds the maximum design limit and damages the device permanently.

As described above, the piezoelectric sensor generates a small charge when subjected to mechanical stress. Any noise generated between the accelerometer and the signal conditioning/charge amplifier module can therefore adversely affect the accuracy of the results obtained. Flexing of the accelerometer cables can induce a charge as a result of the separation of the layers within the co-axial cable, known as the triboelectric effect. These charges can be sufficiently large to induce significant “noise” when measuring low levels of vibration. It therefore follows that accelerometer cables should be securely clamped or taped in position to prevent flexing of the cables that induce such charges (see Figure 13.6). In addition, cable runs between accelerometers and charge amplifiers should be as short as possible as the signal conditioning unit will generally output a strong DC voltage which will be less sensitive to noise than the incoming signal from the transducer. An alternative (which may be preferable in many vehicle test applications) is to use an accelerometer with an in-built preamplifier that performs some or all of the required signal conditioning. For piezo-electric accelerometers, this is usually in the form of a miniature “charge amplifier” which produces an output voltage proportional to the charge generated by the accelerometer. As with all test equipment, cabling runs should avoid sharp bends and be routed away from sources of electrical and magnetic interference, such as traction equipment and current collectors (third rail shoes/pantographs).

## 2. Capacitive Accelerometers

Although piezoelectric accelerometers have very high frequency upper useable limits, they can be limited to around 1 Hz at the lower end of their frequency range (dependent on the charge amplifier employed), as described in Section II.B.1. In contrast, capacitive accelerometers have no lower limit on their useable frequency range as they are capable of giving a DC or static response. They also have a number of other attributes that make them attractive for railway vehicle test



**FIGURE 13.7** Schematic arrangement of a typical capacitive accelerometer.

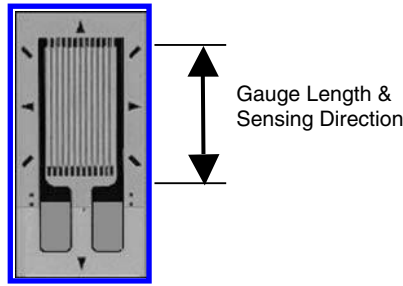
applications. They generally exhibit no phase shift at low frequencies, are insensitive to thermal effects and electro-magnetic interference, have a high signal to noise ratio and a low transverse sensitivity, typically around 1%.

The sensor comprises a tiny seismic mass etched onto a slice of silicon which is interposed between two further silicon plates which act as electrodes. The plates are arranged as a capacitive half-bridge, and the small space between the plates is filled with a gas which provides the necessary damping to the seismic mass.<sup>5</sup> The arrangement is shown in schematic form in Figure 13.7. A useful feature of the design is that the plates provide a mechanical stop for the seismic mass, preventing damage by shock loadings. When the accelerometer is stationary, the mass is central between the plates. Applying an acceleration causes the mass to move towards one of the plates and unbalances the capacitive half-bridge. This results in a charge which is proportional to the applied acceleration. Devices of this type normally include the signal processing within the accelerometer package. Therefore, although the half-bridge is excited by a high frequency AC voltage, the accelerometer requires only a low current (few milliamps) DC input and provides a DC output that can be fed directly to a data logger.

Due to the modest upper frequency limit, mounting methods are less critical than for piezoelectric accelerometers, with the device being glued or bolted to the test components. Once again, however, it is essential to avoid any additional vibrations that may result from insufficiently stiff mounting brackets. Typical capacitive accelerometers suitable for body or bogie mounting may have accelerations of  $\pm 2$  g with a frequency response of 0–300 Hz, or  $\pm 10$  g with a frequency response of 0–180 Hz. It should be noted that, in general, increasing the acceleration range is achieved at the expense of lower sensitivity. As static devices, capacitive accelerometers also measure acceleration due to gravity. The output from the accelerometer will therefore be the sum of the vibration being measured and the component of the acceleration due to gravity acting on the main sensing axis.

### C. STRAIN GAUGES

The science of force and strain measurement is a complex one and it is not possible to provide more than an introduction to the subject in this context. Strain gauges operate on the principle of measuring the change of resistance of a conductor when it is subjected to a strain. This change of resistance is generally measured using a bridge circuit as described below. The most common form of strain gauge is the foil gauge in which the required pattern is etched onto a thin metal foil, a simple example of which is shown in Figure 13.8 below. A good strain gauge will have two apparently conflicting requirements. It must have a short “gauge length” in order to provide a point measurement of strain on the test specimen, whilst having the longest possible conductor to give the maximum change in resistance per unit strain. It is for this reason that most foil gauges use a folded or “concertina” pattern, as illustrated in Figure 13.8.



**FIGURE 13.8** Foil strain gauge (courtesy of Micro Measurements Inc.). *Source:* From Micro Measurements Catalogue. With permission.

The change in resistance of the gauge is related to the strain (i.e., the change in length of the gauge) by a constant known as the gauge factor.

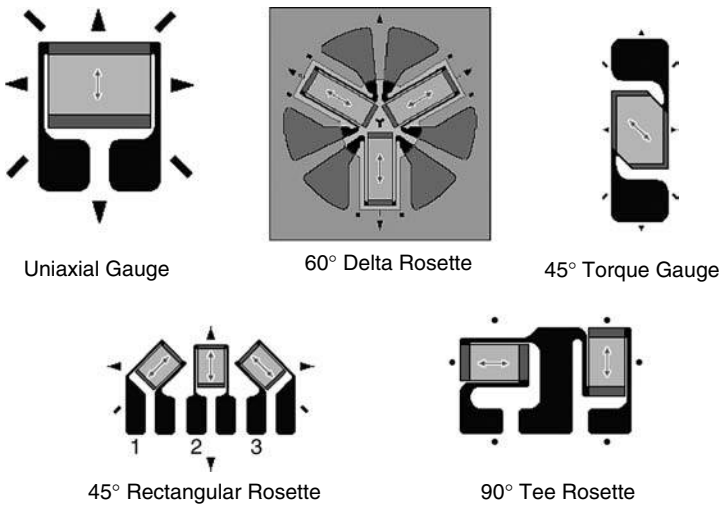
$$k = \frac{\Delta R/R}{\Delta L/L}$$

since strain is defined as  $\epsilon = \Delta L/L$

$$k = \frac{\Delta R/R}{\epsilon}$$

where  $k$  = gauge factor;  $\Delta R$  = change of resistance;  $R$  = unstrained resistance;  $\Delta L$  = change in gauge length;  $L$  = unstrained gauge length; and  $\epsilon$  = strain (normally quoted in terms of micro-strain).

The higher the gauge factor, the higher the sensitivity of the gauge. Good linearity is also a key requirement for accurate measurement; foil strain gauges typically have gauge factors around  $k = 2$  and linearity varying from  $\pm 0.1\%$  at  $4000 \mu\epsilon$  to  $\pm 1\%$  at  $10,000 \mu\epsilon$ . Many configurations of foil strain gauges are available for a variety of strain measuring applications, a selection of which are shown in Figure 13.9. Less commonly used are thick-film and semi-conductor strain gauges. These have considerably higher sensitivity than foil gauges with  $k = 10$  to  $20$  and  $k \approx 50$ , respectively.



**FIGURE 13.9** Examples of various strain gauge configurations (courtesy of Micro Measurements Inc.). *Source:* From Micro Measurements Catalogue. With permission.

Foil gauges are particularly delicate items, and considerable care is required when mounting them on a test component if accurate and reliable measurements are to be achieved. Mounting surfaces must be polished to a good surface finish and then cleaned with specialist solvents and cleaning agents to remove any oil contamination and oxides and ensure that the surface is at the optimum pH for bonding. An adhesive will then be applied to the surface of the part and the gauge, complete with its backing tape, will be pressed onto the surface. Finally, the backing tape is peeled away leaving the gauge bonded to the test specimen. Gauges may be supplied with leads already attached to them, or the leads may be soldered in place once the gauge is mounted. In either case it is good practice to bond a terminal in place adjacent to the gauges from which the main leads can be led away to the bridge circuit. Once the installation is complete, the gauges should be tested before being finally encapsulated in a protective coating to prevent ingress of water and other contaminants. Strain gauges are now available with various mounting systems including weldable gauges that are fixed to the component using a small spot welder.

Strain gauges have a number of applications in the fields of rail vehicle testing. They are commonly the basis of force-measuring devices such as load cells and force-measuring wheels. Strain gauges may also be attached directly to components that are being tested under laboratory conditions. In both these instances the load cell or component can be mounted in a test machine and the resulting strain can be calibrated against a known force input. Strain gauges are also widely used in structural test applications such as determining the strain regimes present in bogie frames or vehicle bodies for testing performed either in the laboratory or on track. The data generated may be used to provide validation for finite element models or as the basis of fatigue calculations. However, as gauges measure strain at singular point(s), considerable care and skill is required to ensure that the critical elements of a structures behaviour are captured.

## 1. Bridge Circuits

The change of resistance generated by strain gauges is very small, typically of the order of a few hundredths of an ohm. The most convenient means of measuring such changes is with a Wheatstone bridge. This comprises of four resistances connected to a DC power supply as shown in [Figure 13.10](#). If  $R_1 = R_2 = R_3 = R_4$ , the bridge is said to be balanced and a voltmeter connected across the bridge as shown will read 0 V. It can be shown<sup>3</sup> that:

$$R_1/R_3 = R_2/R_4$$

This equation highlights two important factors about the Wheatstone bridge. If more than one strain gauge is connected in the measuring circuit, the sensitivity of the bridge can be increased. It is also apparent that changes in resistance on one half of the bridge may “balance” by changing the resistance of the other half. As described below, this provides a useful method of compensating for the temperature sensitivity of strain gauges.

In order to measure strains, the resistors shown in [Figure 13.10](#) are replaced by one or more strain gauges (which are of course variable resistors whose resistance changes with applied strain). At the start of the test, the balancing potentiometer is used to balance the bridge, giving 0 V at the voltmeter. Applying the test load will then cause the resistance of the strain gauge to change and unbalance the bridge again, producing a voltage output that is proportional to the applied strain. The actual strain can then be calculated. It is not uncommon to find situations where it is not possible to calibrate the strain measurement system using a known test load, for example, when gauging a large structure such as a vehicle body. In these cases it is possible to undertake an electrical calibration by placing a high resistance in parallel to the active arm(s) of the bridge. This method is known as shunt calibration and assumes that the surface strain in the test component is fully transmitted to the strain gauge in which it produces a linear response. The fact that the active gauge itself is not a part of the calibration is clearly a drawback, however, providing sufficient care

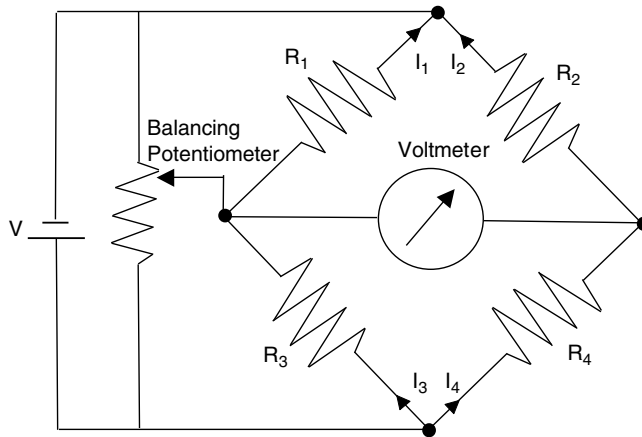


FIGURE 13.10 Wheatstone bridge (with balancing potentiometer).

is exercised, results from such set-ups may be used with a reasonable degree of confidence. Where long cable runs are used, it is essential that the shunt resistance is applied across the gauge with the cable in place to ensure that the effects of cable resistance are accounted for in the calibration.

A bridge may contain one, two, or four strain gauges in the arrangements shown in Figure 13.11, and these are known as a  $\frac{1}{4}$ -bridge,  $\frac{1}{2}$ -bridge and full bridge, respectively. A  $\frac{1}{4}$ -bridge will have the lowest sensitivity of these three arrangements. If no precautions are taken it may also produce errors if the gauges used are sensitive to thermal effects. In order to prevent this, one of the resistors adjacent to the active gauge may be replaced with a “dummy” gauge. This will be an identical strain gauge to the active one that is subject to the same environmental conditions but is not subject to loading, achieved, for example, by mounting it on an unstressed part of the test component. Both gauges will be exposed to any temperature changes and the effect is to cancel out any resulting unbalance on the bridge due to thermal effects. A  $\frac{1}{2}$ -bridge will have a higher sensitivity than a  $\frac{1}{4}$ -bridge as the additional strain gauge will produce a larger unbalanced voltage across the Wheatstone bridge. The presence of two active gauges will also cancel out any thermal effects as described above. A full bridge will have the greatest sensitivity of the three arrangements and will similarly be self-compensating for temperature changes (Figure 13.11).

Strain gauges are available that are self-compensating for temperature changes and the need for dummy gauges is therefore eliminated. However, changes in temperature can also affect the resistance of the lead wires and connectors and, if no dummy is present, such changes may unbalance the bridge resulting in errors in the strain measurement. Such errors can be minimised by the use of a “three-wire” arrangement such as that described in.<sup>6</sup>

It should be noted that the output voltage changes from strain bridges are usually very small and therefore should be amplified as close to the bridge as possible. Once again, cabling should be fully screened and carefully installed to prevent unwanted noise from interfering with the test data. It may be advisable to include dummy gauges in the system that are subject to the same environmental conditions, wiring and connection arrangements as the active gauges but are not subjected to strain. These can be used to assist in determining the level of noise present.

#### D. FORCE-MEASURING WHEELSETS

Recently proposed European standards<sup>1,2</sup> call for the assessment of wheel–rail forces — track forces — in newly developed or essentially modified main line rail vehicles, particularly for those operating at higher speeds. National standards and practices often call for track force assessment

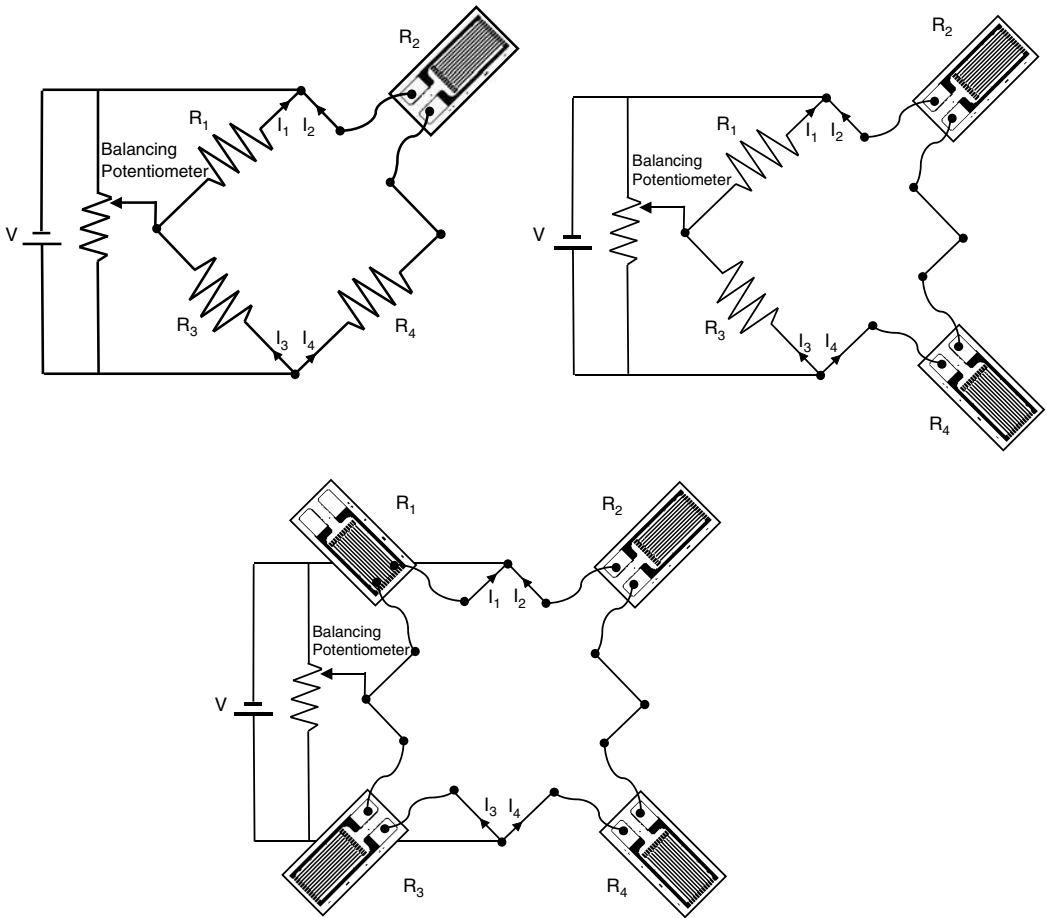


FIGURE 13.11  $\frac{1}{4}$ ,  $\frac{1}{2}$ , and full bridge arrangements.

for acceptance testing. The data generated by force-measuring wheelsets is, obviously, also useful for direct validation of vehicle dynamics simulations. With force-measuring wheelsets, run over appropriate sections of track, it is possible to evaluate the vehicle–track interaction continuously under different operating conditions. Force-measuring wheelsets are often called instrumented wheelsets.

With the more advanced wheelsets it is possible to measure lateral and vertical, and sometimes even longitudinal wheel–rail forces, continuously during “on-track” tests. With correct wheelset design and instrumentation, measuring precision can be good (of the order of 5–10%). However, the instrumentation processes, as well as the procedures during the tests, are sometimes tedious and time consuming.

The sort of measuring device to be used is partly regulated in the current European standard proposals. For low-speed vehicles with conventional running gear, modest axle load and modest cant deficiency it is not mandatory to measure  $Y$  and  $Q$  forces. In such cases a simplified method with instrumented wheelsets measuring just the lateral forces between wheelsets and the axle boxes are considered as sufficient. In some cases (for example, freight wagons with standard running gear at ordinary speed) instrumented wheelsets are not required. However, above a certain speed (usually  $> 160$  km/h or  $> 120$  km/h for freight wagons),  $Y$  and  $Q$  forces must be assessed. This is also the case for higher axle load or cant deficiency. In the European standard assessment of  $Y$  and  $Q$

forces this is referred to as the “normal measuring method.” This requires a more advanced technique for the instrumented wheelsets than for the simplified method described above.

It should be pointed out that there is no known technique available for measuring the forces at the wheel–rail interface directly. Instead, reactions such as strains or accelerations must be measured in structures affected by the wheel–rail forces, i.e., in wheels, axles, or axle boxes. A brief overview of each of the techniques available is given below, whilst a detailed description of the different methods available can be found in.<sup>7</sup> Historically, strains in the track structure have also been measured, but such techniques are outside the scope of this chapter.

### 1. Measuring Lateral Forces between Wheelset and Axle Box

The simplest form of a force-measuring wheelset is to install strain–force-measuring devices between axle journals and the axle boxes. The lateral force can then be estimated by the calibrated strain–force relationships. The operating principle of the device is shown in Figure 13.12. In this case, the lateral axle force is referred to as the  $H$ -force. It is similar, but not identical, to the track shifting force  $S$ . The difference is due to the wheelset mass force.

This simple  $H$ -force method can be further developed by attaching an accelerometer to the wheelset, measuring the lateral wheelset acceleration. This makes it possible to calculate and compensate for the lateral mass force of the wheelset. With this technique it is possible to achieve a fairly good idea of the total lateral track shift force  $S$  between the wheelset and the track. However, with this method it is only possible to measure the total lateral force on the wheelset or on the track; it is not possible to separate the lateral force between the two wheels, i.e., the  $Y$ -forces, or to measure vertical  $Q$ -forces.

### 2. Measuring Lateral and Vertical Wheel–Rail Forces — The Axle Method

Through the measurement of bending moments in the axle, on four cross sections, it is possible to estimate approximate vertical and lateral forces on the wheels, if mass forces generated by the wheelset are neglected. By additionally measuring two torques, approximate longitudinal forces can be calculated. Thus, with six measured moments and torques it is possible to determine six forces (two longitudinal, two lateral, and two vertical) “on-line.” The principle is shown in Figure 13.13. Moments and torques are measured by strain gauge bridges. Signals are transmitted to and from the axle through slip-ring devices, inserted at one of the axle journals, or by radio transmission.

This principle of measuring axle moments and torques seems, at first sight, to be fairly simple, efficient, and accurate. A further advantage is that wheels can be changed on the instrumented axle. However, this method has two major disadvantages:

- Forces on the wheels may be applied at various positions. For example, the lateral position of the contact area may change as much as  $\pm 35$  mm over the wheel tread, thus

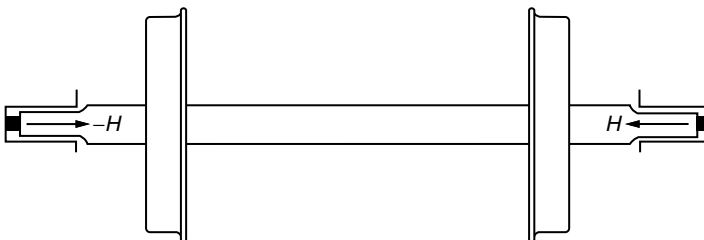
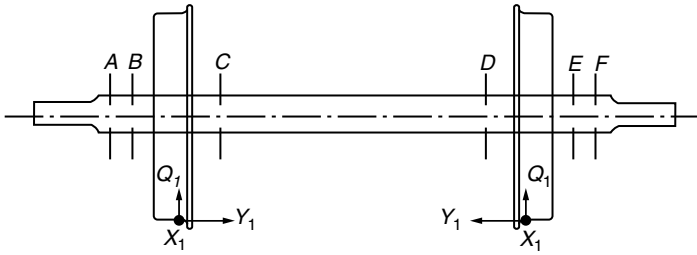


FIGURE 13.12 Lateral force  $H$ , measured between axle journal and axle box.





**FIGURE 13.13** By measuring six bending moments and torques, approximate measures of forces  $x$ ,  $y$ ,  $q$  on the wheels can be determined.

the position of the vertical force application will also change. The changing positions will also change the moments measured in the axle, thus introducing errors that cannot be compensated because the actual position at which the force is applied is not known.

- Moments in the axle are dependent, to a small degree, on the vertical mass forces, due to the unsprung mass of the axle and other unsuspended parts of the wheelset. Thus, it is not possible to fully assess the effects of the unsprung mass on the vertical dynamic forces.

Due to the deficiencies described above, this method has been further developed. By applying strain gauges on the wheel webs, the effects of varying positions are reported to be compensated. However, this makes the method more complicated and approaches the “wheel methods” described in the next section.

### 3. Measuring Lateral and Vertical Wheel–Rail Forces — Wheel Methods

“Wheel methods” can be divided into two different techniques, either measuring strains in the spokes of spoked wheels, or measuring strains in the wheel web of ordinary railway wheels, i.e., in the web between the axle and the outer wheel rim.

The “spoked wheel method” is not frequently used nowadays, mainly due to the need to design and manufacture special wheels. In addition, the calibration procedures are tedious and time consuming. However, with properly designed and calibrated instrumented spoked wheels, this method is reported to produce a good accuracy. The mass forces of the unsprung mass are, to a large extent, included in the measured quantities.

The most frequently used method today — besides the simplified “axle box method” described in Section II.D.1 also is the wheel web method. Within this method a number of different technologies are used. The basic principle is that strains are measured at various locations on the wheel web as a result of the applied forces on the wheel as shown in Figure 13.14. A number of strain gauges are applied on the same web, usually in the radial direction on the inside as well as on the outside of the wheel. However, these locations must be carefully selected.

Figure 13.15 shows an example of measured strains in single strain gauges of one wheel web as the wheel rotates and the wheel is loaded by lateral forces  $Y$  or vertical forces  $Q$ . In order to achieve signals proportional to the applied load, the strain gauges must be combined in Wheatstone bridges in an intelligent and precise way. Separate bridges are required for the lateral  $Y$  forces and the vertical  $Q$  forces. Sometimes, two bridges are used for the same force on the same wheel, installed at different wheel angles. In this case, additional data processing is needed to combine the two bridge signals. In a few cases, forces are measured in all the three directions: longitudinal, lateral, and vertical. Signals are usually transferred to and from the wheels via slip rings, although radio transmission may also be used.

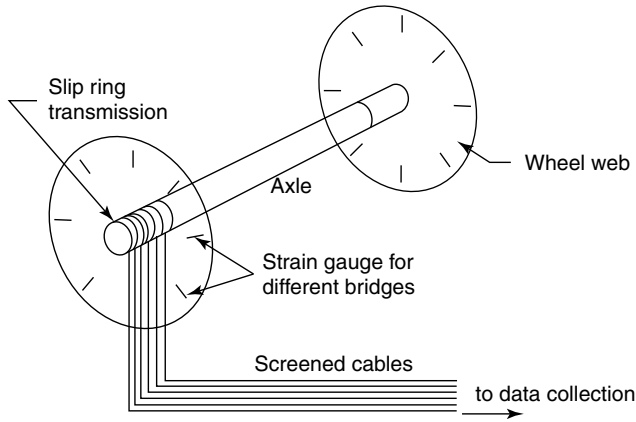


FIGURE 13.14 Schematic arrangement of force-measuring wheelset using the wheel web method.

#### 4. Compensation for Undesired Parasitic Effects

An important issue for force-measuring wheelsets is how to compensate for “parasitic” effects and possible cross talk between forces in the longitudinal, lateral, and vertical directions. The parasitic effects include the influence of wheel rotation, temperature, and temperature distribution and, finally, the location of the forces on the wheels. As described in the previous section, the lateral position of the vertical force application will change by as much as  $\pm 35$  mm over the wheel tread, which may generate errors in the output signal. Also, electro-magnetic noise must be carefully considered as very strong currents (1000–2000 A) will sometimes pass just some 50–100 mm away from the wheels and the cabling. Effects of water and humidity, temperature and mechanical impact must also be considered.

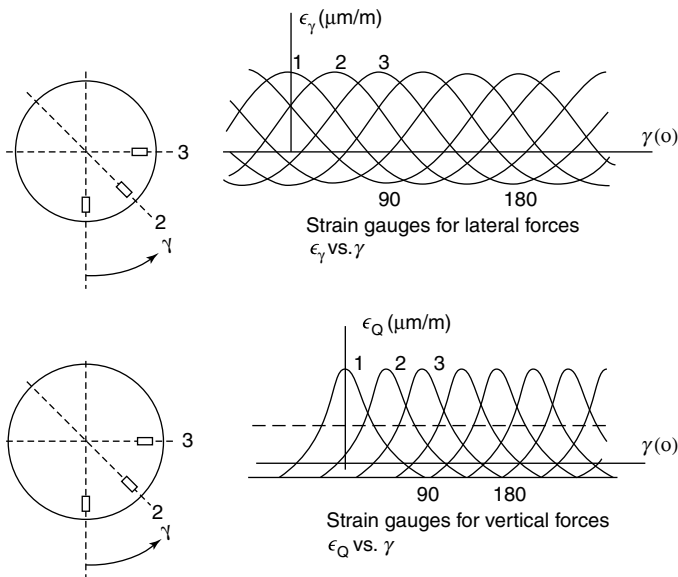


FIGURE 13.15 Example of measured strains in single strain gauges of one wheel web as the wheel rotates —  $y$  and  $q$  applied at the lowest part of the wheel.

The overall goal is to achieve output signals proportional to the applied loads with a minimum of cross talk and parasitic effects. Several techniques are used for this purpose, for example, in France, Germany, Sweden, U.S.A., and China. The selection of locations for the strain gauges, their connection in bridges and the additional data processing, may vary considerably from one laboratory to another. The design, calibration, and operation of this type of equipment is a highly specialised subject and a detailed description is beyond the scope of this chapter.

The principle advantage of the “wheel web method” is that it is possible to measure the  $Y$  and  $Q$  forces continuously quite close to the wheel–rail interface and hence most of the dynamic effect from the unsprung mass is included in the measured data. It is possible to measure quite high-frequency forces (up to at least 100 Hz). The measuring accuracy may be good or at least acceptable (within 5–10% under normal conditions) if wheels and the whole system are properly designed and calibrated. The major drawback is the volume of work required for system design and calibration, requiring specialised knowledge which generally makes the technique very expensive. A further drawback is that the instrumented wheelsets must very often be specifically designed for the type of vehicle to be tested.

## E. VEHICLE SPEED AND POSITION MEASUREMENT

A prerequisite of almost all on-track testing applications will be the ability to determine vehicle speed and position. This section discusses the most commonly used approaches to this problem.

### 1. The AC Tachogenerator

The most commonly encountered means of generating a vehicle speed signal is the AC tachogenerator. This is, effectively, a two-phase induction motor, comprising a rotating magnet with a pair of stator coils arranged at  $90^\circ$  to each other and to the axis of rotation as shown in Figure 13.16. One coil is excited with a constant frequency AC signal. The resulting eddy currents in the core induce an AC voltage in the sensing coil which is proportional to the rotational velocity of the core. The direction of rotation can also be determined from the device as the output voltage phase will change by  $180^\circ$  when the rotation direction is reversed. AC tachometers are generally used as they are less susceptible to noise and “ripple” of the signal than their DC equivalents. They can also be fairly robust, an important consideration as axle box mounted equipment may be exposed to very high accelerations.

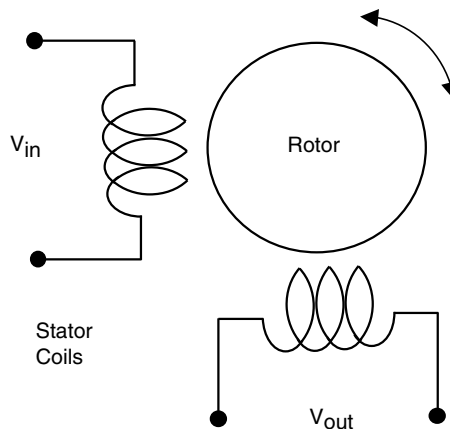


FIGURE 13.16 AC tachogenerator.

In order to determine actual vehicle speed and distance travelled from devices of this sort, the vehicle's wheel diameter at the nominal rolling position must be measured. Inevitably, the accuracy is limited by the lateral movement of the wheel from the nominal rolling position. If measurements are carried out over extended periods, the wheel diameter must be remeasured to compensate for the effects of wheel wear.

## 2. Hall Effect Probes

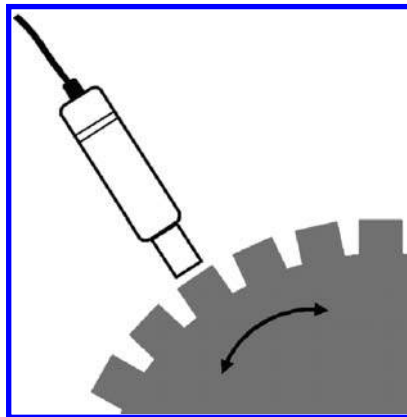
An alternative means of measuring wheel rotational speed is with a Hall Effect probe. If a conductor with a current flowing through it is placed in a magnetic field whose direction is normal to the direction of the current, a voltage (the Hall voltage) will be induced across the width of the conductor. This is due to the magnetic field causing the electrons to take a curved path through the conductor. The effect is found most strongly in semiconductors and these are therefore the basis of commercially available Hall Effect probes.<sup>8</sup> Rotational speed measurement is achieved by combining the probe with a ferrous toothed wheel, as shown in Figure 13.17. As the ferrous tooth passes the probe, it causes the reluctance of the probes internal magnetic circuit to change and this change produces a varying Hall voltage where the frequency is proportional to the rotation speed. Once again, if the resulting speed signal is used to estimate distance travelled, the wheel diameter must be known and appropriate corrections made for wear.

## 3. Ground Speed Radar

This is a noncontact device that relies upon the Doppler effect to measure the vehicle speed. The Doppler effect is based upon the frequency shift that occurs when energy waves radiate from, or are reflected off, a moving object. A familiar example is the change in pitch in the noise from a train passing at speed. Due to the Doppler effect, the pitch increases as the train approaches and then lowers as it departs. For a ground speed radar device, a high, known, frequency signal is transmitted from the radar, aimed at a point on the track beneath the vehicle. The reflected signal will be detected by the sensor and the phase shift from the original transmitted signal will be calculated allowing the velocity of the vehicle relative to the stationary target (the track) to be determined. The Doppler frequency shift is given by:

$$F_d = 2V(F_0/c) \times \cos \theta$$

where  $F_d$  = Doppler shift Hz;  $V$  = velocity;  $F_0$  = transmitter frequency Hz (typically 25–35 GHz);  $c$  = speed light; and  $\theta$  = offset angle.



**FIGURE 13.17** Hall Effect rotational speed sensor.

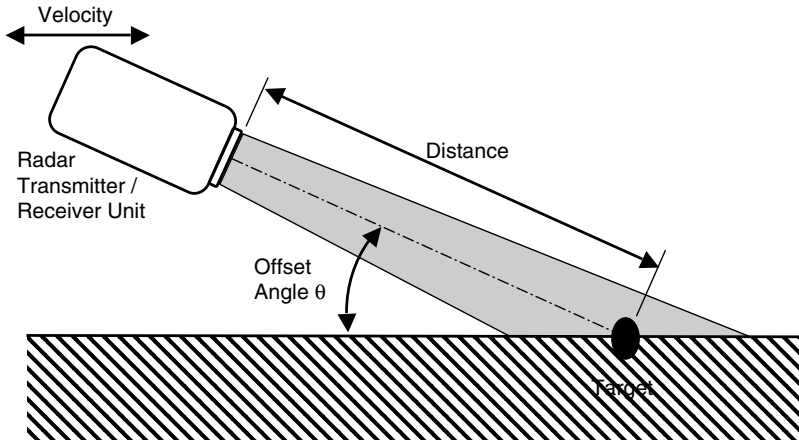


FIGURE 13.18 Simple Doppler effect speed sensor.

A simplified arrangement of a Doppler effect speed sensor is shown in Figure 13.18. It should be noted that the sensor is usually mounted at an angle as shown. This represents a compromise between the strength of the return signal, which is greatest with the sensor vertical and reducing the sensitivity to vertical motion due to the vehicle/bogie bouncing and pitching on its suspension. Larger offset angles can introduce a “cosine error”<sup>9</sup> as targets at the edge of the conical beam will be at a slightly different angle to those in the centre of the beam. Commercially available ground speed radar devices may have two sensors, one of which faces forwards while the other faces backwards. This arrangement can be used to automatically correct mounting or vehicle pitch errors. Devices of this type are typically mounted between 0.3 and 1.2 m from the ground and have an accuracy for speed measurement of better than 1% above 50 km/h. Below this speed the accuracy reduces somewhat (say within 0.5 km/h below 50 km/h) and as such these devices are less suitable when accurate low speed measurements are required.

#### 4. Determining Vehicle Position

Before the advent of global positioning systems (GPS), a vehicle’s position was recorded in the test data by means of marking events such as mile/km posts in a separate data channel on the logger. This could be carried out manually by observations from the test coach, or alternatively a known position at the start of the route could be synchronised with a data file containing a list of features and their locations for the chosen test route (known in the U.K. as a “route setting file”). These would then be written to the test data based on the distance calculated from the vehicle speed. When such a method is used, it is normally necessary to resynchronise against an observed position to remove the effects of measurement errors from the speed/distance measurement device.

Modern GPS systems make establishing the location of a vehicle relatively straightforward, provided an asset register is available to relate the logged GPS position to the location of stations and other infrastructure features. GPS is based upon a network of 24 satellites orbiting around 12,000 miles above the Earth. They are arranged so that a GPS receiver should be able to see signals from four of these satellites at any given time. Each satellite transmits low power radio signals which can be detected by a GPS receiver on Earth. The signals contain information which allow the GPS receiver to determine the location of each satellite it is tracking and how far it is from the receiver. Knowing this information for three satellites allows the receiver to calculate a two-dimensional position (latitude and longitude) whilst adding the position of a fourth satellite produces a 3-D position (latitude, longitude, altitude). A more detailed description of the system may be found in Ref.10. Typical accuracies for GPS systems are in the order of 6–12 m. A system

known as Differential GPS (DGPS) uses Earth based “reference stations” at known locations to determine corrections to the satellites transmitted positions. Using DGPS, accuracies of 1–5 m or better may be achieved.

For railway test applications, the GPS antenna is mounted on the roof of the vehicle to ensure that the maximum number of satellites is visible. However, since GPS is a “line-of-sight” system, deep cuttings, tunnels, high buildings, and other obstructions will prevent the system from working. In addition, it may be considered advisable to confirm the logged location by “marking” the logged data, either by a manually activated signal against known locations (mileposts, etc.) or automatically by recording signals from trackside balises or signalling devices (e.g., AWS/TPWS loops in the U.K.).

### III. TEST EQUIPMENT CONFIGURATION AND ENVIRONMENT

The elements of a typical test arrangement are shown in Figure 13.19. The example shown might be appropriate for gathering data on the suspension behaviour or ride comfort of a vehicle, but the principles apply equally to many on-vehicle or lab test tasks.

Data from the selected transducers is passed through the appropriate conditioning electronics and transmitted to the analogue side of a data logger. Such signals are usually in the form of DC voltages although some devices utilise AC voltage or varying current with a steady DC voltage. The signals are passed through an analogue to digital converter and stored in digital format in the data logger memory. Many systems allow real-time display of the incoming data and this is very useful for checking that the measurement system is performing correctly and that sensitivity settings for different channels are correctly configured. It is essential when vehicle acceptance or safety tests are being undertaken.

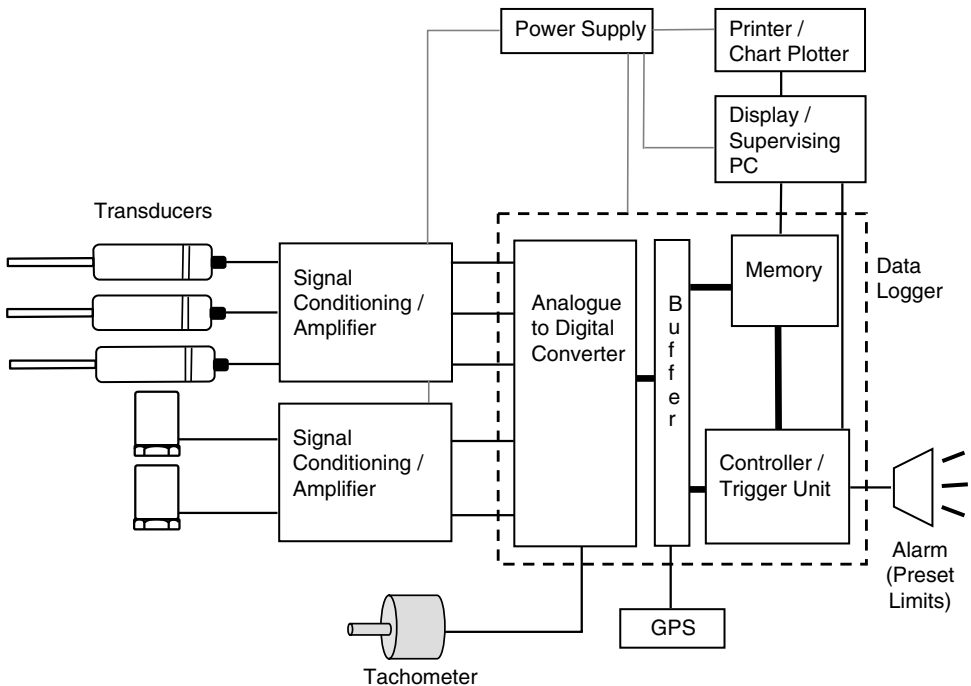


FIGURE 13.19 Typical test equipment configuration.

The railway represents an aggressive environment in which to perform test measurements, and careful attention to detail is required if reliable test set-ups on vehicles or track are to be implemented. Transducers and cabling may be subject to high levels of transient vibrations, particularly if mounted on unsprung components (e.g., axle boxes or rails). Acceleration levels of the order of 30–50 g are not uncommon on axle boxes, and peaks of up to 100 g have been recorded. Extremes of temperature and weather must be catered for and all externally mounted components should resist the ingress of moisture and dirt. The test data must not be affected by electrical noise from a wide variety of source including high voltage AC or DC traction systems. Conversely, the test set-up must preclude the possibility of generating electro-magnetic interference that could adversely affect train control or signalling systems. Fortunately, most commonly used transducers are low voltage, low power devices and hence the problem does not arise.

The following list highlights typical requirements for a reliable installation:

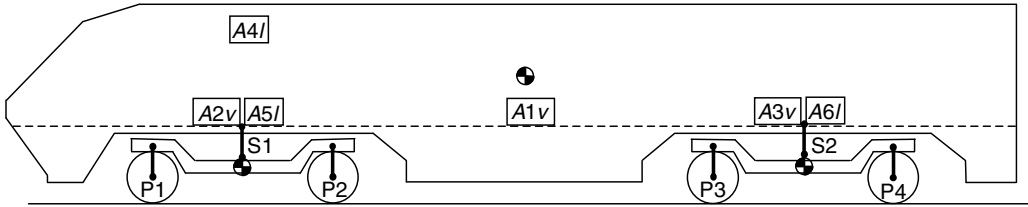
- All brackets should be rigid and robust and should prevent unwanted backlash between the bracket and transducer.
- Cabling should be secured to prevent (insofar as possible) movement. This is particularly important in the case of connections to transducers or joints between cables.
- Cables should be routed to avoid sources of electrical noise (traction motors, pantographs/collector shoes, wiring looms, generators etc.).
- Where possible, signal conditioning should be carried out within the transducers. Where this is not possible (for example, with some types of accelerometer requiring a separate charge amplifier), the distance between the device and the conditioning unit should be as short as possible as very low power signals may be rendered useless by even modest amounts of electrical interference.
- Cables and connectors should be shielded to the highest available standards.
- All transducers, cables, and enclosures should be sealed to a recognised standard such as IP66/IP67.
- Transducers should be selected to prevent the likely peak vibration transients exceeding the “shock loading” specification for the device.
- Cabling lengths should be minimised and be routed inside vehicles at the earliest opportunity. Cables should be arranged tidily and long coils of spare cable should be avoided at the end of cable runs.
- Care should be taken to ensure that expensive transducers and data loggers cannot be exposed to damaging voltage “spikes.”

Sources of power are an important consideration when conducting field testing. It is generally recommended that when using power other than from the mains (for example, from generators or vehicle sources), suitable voltage stabilisation and surge protection devices are used. In many cases, field power supplies cannot be guaranteed and an alternative battery backup should be arranged to guard against the loss of important test data.

## A. TRANSDUCER POSITIONS ON VEHICLES

One of the most common instrumentation applications that the vehicle dynamics engineer will encounter is the fitting of accelerometers and displacement transducers to the body, bogies and suspension, for ride test, passenger comfort, or dynamic response track tests. When choosing the locations for instrumentation, it is important to clearly understand the vibration modes that each transducer will “see” to ensure that the correct data is gathered and that the desired information can be derived from it.

It is evident from [Figure 13.20](#) that a vertical accelerometer mounted on the vehicle floor in line with the vehicle centre of gravity ( $A1v$ ) will largely sense to the bounce mode responses to the track



**FIGURE 13.20** Simple transducer layout. Key:  $Axv$ –vertical accelerometers,  $Axl$ –lateral accelerometers,  $Px$ –primary suspension LVDTs,  $Sx$ –secondary suspension LVDTs.

input. As it is not practical to actually mount the accelerometer at the centre of gravity, some accelerations due to the change in floor height as the vehicle rolls will also be detected. Vertical accelerometers mounted on the vehicle floor above the bogie centre pivots ( $A2v$  and  $A3v$ ) will sense both the bounce and pitch mode responses. However, whilst accelerations due to body bounce will be in phase at both transducers, the accelerations due to pitching will be  $180^\circ$  out of phase. Assuming that the carbody is perfectly stiff, i.e., that no flexible modes occur, the pitch and bounce components of the signal at accelerometer  $A2v$  can therefore be separated thus:

$$A2_{\text{bounce}} = (A2v + A3v)/2$$

$$A2_{\text{pitch}} = (A2v - A3v)/2$$

Similarly, accelerometers  $A5l$  and  $A6l$  will sense body lateral and yaw responses in and out of phase, respectively.  $A4l$  will likewise sense a combination of body yaw, lateral, and roll modes. Providing that sufficient transducers have been provided, and their locations chosen carefully, it should be possible to reliably establish the natural frequencies of the various modes of vibration of the vehicle body. The sensed accelerations on the body will also include the effects of flexible body modes. These may also be of interest to the dynamics engineer (for example, when considering the effect of the first body bending mode on passenger comfort). Accelerometers will also detect inputs from body mounted mechanical equipment such as internal combustion engines and compressors. However, these will often occur at constant frequencies, somewhat higher than the frequencies of interest to the dynamics engineer and may, if desired, be easily removed by filtering.

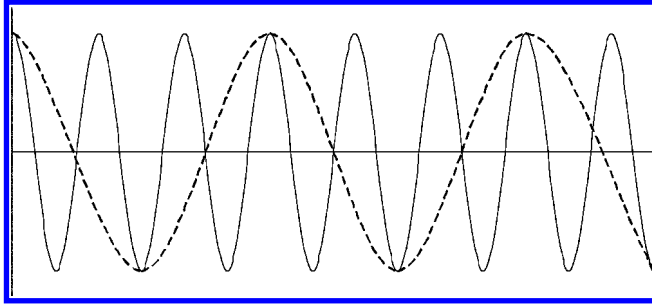
Similar considerations to those described above will also apply to accelerometers mounted on the bogie frame or displacement transducers fitted across the primary or secondary suspension. Accurate records should be made of the mounting positions of all transducers on the vehicle to allow for later correction of geometric effects.

#### IV. DATA ACQUISITION

Test data must be collected and store in a suitable form for later analysis. In the past this was often carried out using magnetic media such as tape recorders to allow large volumes of data to be stored. However, modern computer systems are able to directly store such large volumes of data, making tape storage largely redundant. Modern data loggers are usually either in the form of a PC with suitable additional hardware cards and software, or a standalone device with a PC compatible up-link. In either case the analogue signals from the test devices must be converted to digital form to allow the data logger to store them. In addition to logging varying voltage or current signals from transducers, loggers may also have additional hardware inputs for digital signals and serial data (for example an RS232 connection to a GPS).

As computers/data loggers store information in digital format, and most transducers provide an analogue signal, an analogue-to-digital converter (ADC) is required to convert the signal.





**FIGURE 13.21** Aliasing due to insufficient sampling rate.

The process of sampling and converting the signal leads to the possibility of two forms of inaccuracy in the digitised signal, known as aliasing and quantisation errors.

A key decision when setting up a data logging system is the sampling rate. This is the time interval at which the logger takes a “snapshot” (sample) of the incoming analogue signal. If the sampling rate is too low, a high frequency signal may appear incorrectly as a lower one, as illustrated in Figure 13.21. Although the aliased frequency shown here is one third the actual signal frequency, a whole series of aliases are possible depending upon the sampling frequency chosen. Theoretically, the chosen sampling rate should be at least twice the highest frequency (of interest) in the sampled signal. However, in practice, it is normal to sample at between five and ten times the highest frequency required in order to ensure good representation of amplitude as well as frequency.

An additional aliasing problem can arise if the sampled signal is degraded by an unknown, high frequency, noise component. A sampling rate of, say, ten times the highest frequency of interest, could cause this noise to be aliased into the measurement frequency range giving incorrect results. The solution in this case is to low pass filter the signal using an analogue filter prior to sampling to remove the unwanted component. This precaution is known as anti-alias filtering.

Quantisation errors are introduced by the ADC forcing a continuous analogue signal into a limited number of discrete levels (binary digits or “bits”). The error will be present in all digitised signals and has the effect of restricting the dynamic range of the signal. The magnitude of the error will be proportional to the resolution of the ADC. Quantisation errors are not generally a serious restriction for modern data loggers. However, care should be taken to ensure that the full range of the ADC is used. For example, consider an ADC using an 8-bit conversion, which allows 256 discrete states of the converted signal. If the ADC range is set to  $\pm 1$  V and a  $\pm 1$  V signal is converted, the digitised signal will consist of 256 discrete values and will give a good representation of the original analogue signal. However, if the same signal is passed through with the ADC range set to  $\pm 10$  V, the  $\pm 1$  V signal may only have 256/10 discrete values and a significant quantisation error results. It also follows, therefore, that the potential for quantisation errors are larger when a signal has a wide dynamic range.

Once the signal has been converted to digital format, further operations can be performed easily, such as filtering, bias or offset removal, amplification, etc. Detailed discussion of digital signal processing techniques is beyond the scope of this chapter, but numerous texts exist to guide the interested reader.

Important considerations when selecting a data logger include:

- The number of transducers to be used in the test, and hence the number of channels required by the data logger.
- The rate at which each transducer is to be sampled. Modern data loggers are generally capable of sampling at very high frequencies (of the order of kHz), many times higher than required for most railway vehicle dynamics applications. Many loggers allow

different sample rates to be set for different channels, though these must generally be divisible by the highest sample rate chosen.

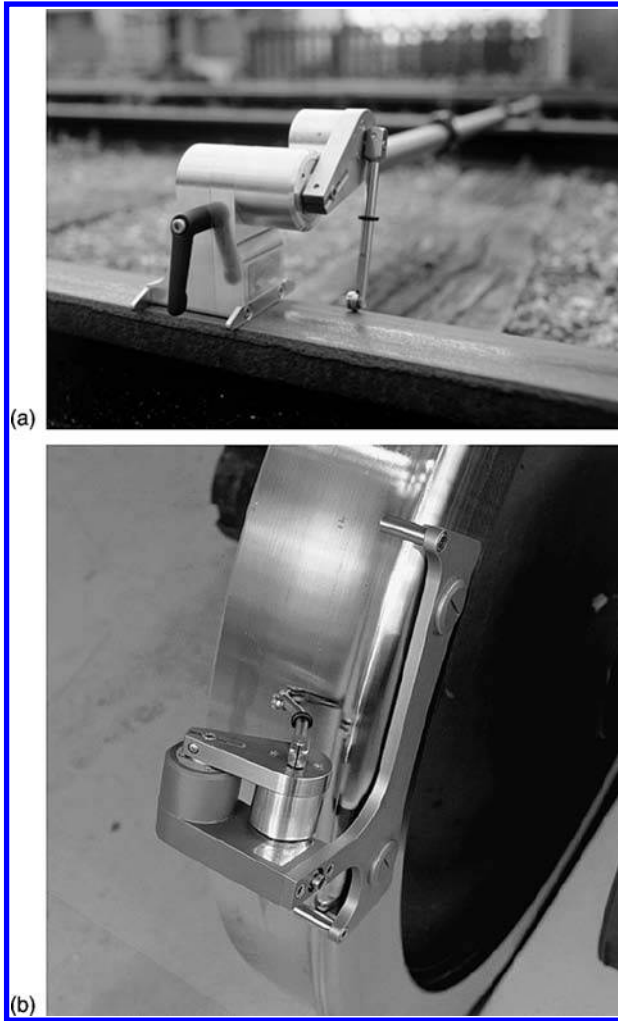
- The duration of the test and hence the volume of data to be stored. Assuming each sample for one transducer represents one byte, this can be easily calculated by multiplying the sample rate, the number of channels sampled and the required test duration to give the storage volume needed.
- The resolution of the analogue-to-digital converter. This must be high enough to allow the transducer data to be collected to the required precision.
- The construction of the data logger. A PC may be appropriate for lab test work, but on-train applications may require a more rugged construction. This will depend partly upon whether the logger is to work in a test vehicle on a temporary basis or to remain in a service vehicle over extended periods. It should be noted that traditional computer hard disk drives are not reliable when subjected to substantial or prolonged vibration and, for on-train systems, consideration should be given to using more rugged components (such as flash memory).
- The method of sampling employed. Most high-end data loggers sample all the required channels simultaneously ensuring that the data is perfectly synchronised. However, modern logging devices are available very cheaply that, whilst only sampling each channel sequentially, can do so at high frequencies. Where a small number of channels is only required to be sampled at low frequency, these loggers may be adequate.
- The interface required during the tests. A logger may be required to record data unattended, in which case no interface is required. Alternatively, a real-time display of all the data being logged may be needed allowing, for example, alarm thresholds set to give warnings if predetermined levels are exceeded on any channel. Logging tasks which also require extensive real-time calculations to be performed on the logged data may require the development of dedicated software.

It is always advisable to have the ability to view measured data on-line when conducting on-track testing work where the ability to repeat tests is limited. This is particularly important on the first day of an extended test programme in order to ensure that all data is as expected and to ensure that errors do not occur in the complete measuring system due to faulty scaling factors, electrical noise, etc.

## V. MEASUREMENT OF WHEEL AND RAIL PROFILES

The accurate measurement of wheel and rail profiles is critical to many vehicle dynamics simulation tasks. Early measuring devices for such tasks relied on moving a stylus over the profile, the shape being transferred via a linkage with a pen attached to a piece of tracing paper. An alternative mechanical measuring method used an indexing plate to allow a dial-test-indicator to be moved to a number of known positions around the wheel or rail, a reading being taken at each. Such devices tended to be labourious to use and required the data to be manually entered into a computer for use in dynamics simulations. Several proprietary devices are available for making such measurements. These devices either use a mechanical linkage or a laser beam to electronically record the cross sectional rail or wheel profile in terms of Cartesian coordinates. Vehicle dynamicists will usually require a high degree of accuracy to be maintained when taking profile measurements as wheel/rail forces may be significantly affected by variations in shape of a few tenths of a millimetre. In general, measuring systems to provide this level of accuracy rely on readings being carried out manually during track walks or depot visits. Caution should be exercised if using data from automated in-track or on-train inspection systems as the need to carry out measurements at speed may limit the accuracy of the recorded profile.

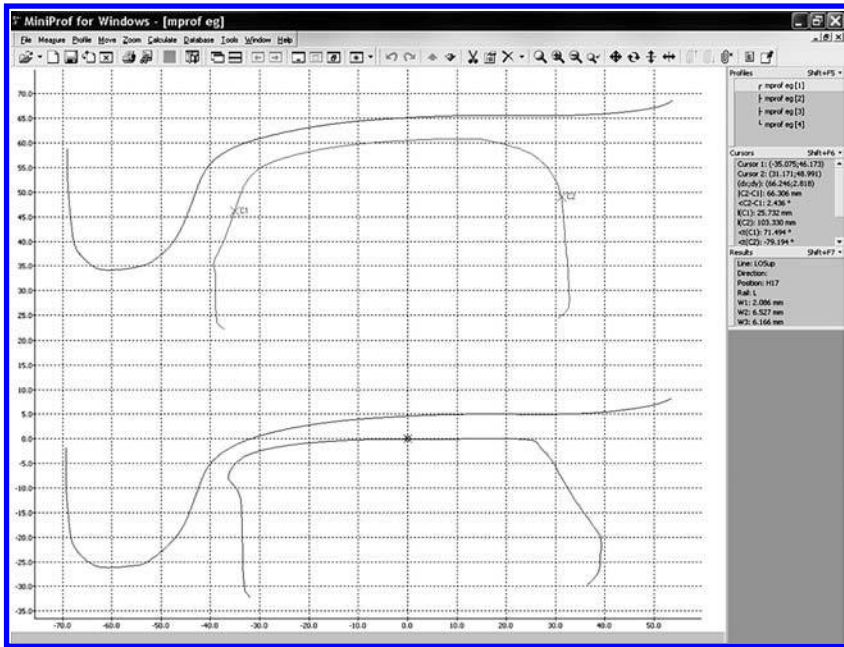
The device must be capable of measuring not only the profile, but also its correct orientation in the lateral-vertical plane. Small errors causing rotations of a wheel or rail cross sectional profile will



**FIGURE 13.22** MiniProf rail and MiniProf wheel measuring instruments (courtesy of Greenwood Engineering). *Source:* Catalogue — Courtesy of Greenwood Engineering.

appreciably alter the resultant contact conditions and the resulting equivalent conicity and must be avoided for realistic simulations. Several devices use the flangeback as the datum position for wheel profile measurements. This simple approach may generate significant errors when measuring profiles where the flangeback is worn, for example, by contact with check rails. Experience shows that this simple approach therefore has limitations. For measurement of both wheel and rail profiles, the device should have an outrigger which bears on the opposite wheel or rail to ensure that the measuring head remains in plane of the track/wheelset. Such equipment is not generally available for measuring wheels. In all cases, however, a system using two measurement heads capable of recording a pair of wheel or rail profiles and their relative orientation simultaneously will provide superior results.

As the requirement is generally to gather profiles which are representative of a given location or situation, the exact position at which profiles are measured should be chosen with some care to avoid localised pits or defects. Profiles should always be cleaned prior to taking the measurement as contamination on the wheel or rails (e.g., dirt, wear debris, or grease) may prevent accurate recording of the profile shape. Similarly, it is particularly critical that the measuring head of contacting devices is also kept clean.



**FIGURE 13.23** Viewing measured profiles using MiniProf software (courtesy of Greenwood Engineering).  
*Source:* Catalogue — Courtesy of Greenwood Engineering.

An example of a widely used device is the MiniProf (Figure 13.22). The measuring head in this case (which is common to both wheel and rail devices) consists of two arms and two rotary optical encoders. These are used to determine the position of a magnetic measuring wheel. As the operator moves the measuring wheel around the profile, the signals from the encoders are logged by a laptop computer. The profile of the wheel or rail being measured is then calculated from the position of the two arms as obtained from the encoders, with suitable correction being applied to allow for the varying contact position on the measuring wheel itself. The use of a wheel rather than a stylus has a filtering effect upon the measured profile, as the wheel cannot follow very small surface irregularities. However, providing that the radius of the measuring wheel remains significantly smaller than the smallest wheel or rail profile radius, this effect is of little importance. Software supplied with the MiniProf allows profiles to be viewed and provides the facility to undertake various geometric wear calculations against “reference” profiles (Figure 13.23).

Although less widely used, other profile measuring devices exist, only some of which are suitable for use in vehicle dynamics simulations or for a detailed study of wheel–rail contact conditions. The general requirement is for a high degree of measuring accuracy, probably better than 0.02 mm. Some devices using scanning lasers are suitable for this application, although other displacement or rotation measuring transducers are also likely to be suitable. In the past, systems designed to give “real time” measurements of wheel profiles, either track based automated systems in depots or train based systems, have not achieved sufficiently high accuracy levels. This is because some of the accuracy is sacrificed for speed of measurement. However, with advances in scanning laser technology, this may well change in the future.

## VI. TRACK GEOMETRY RECORDING

The vehicle dynamics engineer frequently requires data describing real, representative track geometry as a basis for simulating vehicle behaviour. It is convenient, for both maintenance engineers and vehicle dynamicists, to separate the long wavelength features that represent the

design layout of the track from the short wavelength features that form the variation from the design (i.e., the track irregularities). This usually results in a description based on the following five geometrical terms:

*Curvature* — the lateral design layout of the track radii (long wavelength). Curvature is defined as the inverse of the curve radius in units of rad/km. However curvature may also be quoted as a “versine” measurement in mm, this being the distance from the centre of a chord of known length to the rail.

*Cant* — the vertical difference in height of the left and right rails, (long wavelength).

*Lateral alignment* — the short wavelength lateral track irregularities.

*Vertical alignment* — the short wavelength vertical track irregularities.

*Gauge* — the distance between the rails measured at a specified distance below the crown of the rail. This is typically 14 mm in the U.K. and Europe and 5/8 in. in North America. Gauge may be given as an absolute value or a variation from a nominal gauge (e.g., 1435 mm European standard gauge).

The geometry of railway track may be recorded using one of several techniques.

### A. MANUAL SURVEY

The surveyor establishes a datum (or several datums) position on the site to be surveyed, usually by placing a marker in the ground. A theodolite is then used to record the position of the left and right rail with reference to the datum position. Considerable care is required to produce accurate results from these techniques. Good results have been achieved by using a high accuracy “autotracking” theodolite measuring to a target placed on the fixed end of a cant and gauge stick above the rail gauge corner. The theodolite is then used to measure the position of one rail, and the position of the adjacent rail is determined from the cant and gauge measurements displayed on the stick (see Figure 13.24).

In the absence of other methods, useful results may be obtained for track design curvature by conducting a versine survey. In this case, a chord (wire) of fixed length is stretched along the high rail of the curve and the distance between the centre of the chord and the gauge corner is measured. Chord lengths of 10, 20, or 30 m are common depending on the curve radii to be measured. The chord length is normally chosen so that the measured versine does not exceed 150 mm on the tightest curve to be surveyed. Successive versine measurements are taken at frequent intervals, with the maximum recommended interval being half the chord length. Increasing the measurement frequency will increase the detail contained in the survey results. The radius of curvature at any mid-chord position can then be calculated as follows:

$$R = \frac{C^2}{8V}$$

where:  $R$  = curve radius (m);  $C$  = chord length (m); and  $V$  = versine (mm).



**FIGURE 13.24** Cant and gauge stick (courtesy of Abtus Ltd.). *Source:* Abtus Ltd. — Catalogue. With permission.

However, the ability of such techniques to “see” short wavelength lateral irregularities is inherently limited as the position of the datum (the wire itself) depends upon the track irregularities at either end of the wire. This survey method is simple and cheap to carry out and requires limited equipment, but has the disadvantage of needing three people to undertake the survey.

Manual surveys are generally fairly slow to carry out and are therefore limited to short sections of track.

## B. TRACK GEOMETRY TROLLEY

A number of proprietary recording trolleys are available. These carry instrumentation and a data logging system to allow track irregularities to be measured. They commonly measure vertical irregularities cross-level and gauge and (less commonly) lateral irregularities and curvature. An electronic record of the geometry is stored in the on-board data logger for later retrieval, and many trolleys also provide a paper-trace of the stored data. Some feature on-line calculations allowing alarms to be set for exceedances in, for example, track twist over a given distance. Trolleys are generally pushed at walking pace by an operator, though some are self-propelled at low speed. The length of line that can be surveyed by this method is generally greater than for manual surveys, but is limited by the slow recording speed and the capacity of the on-board data logger and power supplies (Figure 13.25).

## C. TRACK RECORDING VEHICLES

Most railway administrations operate dedicated track geometry recording vehicles. These vehicles are equipped with measuring systems, often based upon the inertial principle described below, which allow data to be gathered at high speeds (typically up to 185 km/h). Such specialist vehicles have extensive data storage and analysis capabilities which allow regular surveys of entire routes to be undertaken at normal running speeds. These vehicles provide the most commonly used source of data for vehicle dynamics engineers.

The following description of the inertial measurement is based upon the track recording systems used in the U.K. and described in Refs 11–13. The general principles are, however, common to all systems of this type. The signal from an accelerometer mounted on the vehicle body is double integrated to provide a displacement measurement. This is then low-pass filtered to



**FIGURE 13.25** Track geometry recording trolley (courtesy of Abtus Ltd.). *Source:* Abtus Ltd. — Catalogue. With permission.



remove the long wavelength design information, effectively creating a moving average datum for the measurement of the shorter wavelength features.

Vertical (track top) measurements are made using one wheelset on the vehicle as the sensor. Displacement transducers are fitted across the primary and secondary suspension as shown in Figure 13.26, with an accelerometer mounted directly above them. Subtracting the suspension displacements from the body displacement (double integrated from the acceleration) gives the track top profile. As the suspension movements are removed from the final answer, the system is effective regardless of suspension type.

Noncontact measurement of the track gauge at high speed presents a considerable challenge. Early systems projected a narrow beam of light onto the railhead and used cameras to measure the intensity of the reflected light and thus determine the location of the gauge face. Recent developments include using a laser beam is guided onto a point 14 mm below the crown of the rail, the normal gauge measuring point, by a mirror. This mirror is “steered” by a galvanometer, responding to suspension displacement measurements. Alternatively, a fanned array of laser beams may be projected directly onto the railhead. In either case the reflected laser light is then used to measure the position of left and right rails, and these are combined to obtain the gauge (Figure 13.27).

The lateral irregularity of the track is obtained by subtracting the rail position, measured by the laser displacement sensors described above from the inertial datum produced by a body mounted lateral accelerometer.

The cross-level is determined by subtracting the difference in the vertical suspension displacements from the body roll angle obtained from an on-board gyroscope. The gyroscope also provides the plan view rate of turn of the body and this, together with the vehicle velocity, allows the curvature to be calculated.

The foregoing description is, of necessity, a somewhat simplified version of what is a sophisticated and complex measuring system. It is worth noting that the lateral irregularity and curvature channels are effectively short and long wavelength parts of the same signal. Track geometry data is normally supplied at 0.2–0.25 m intervals in the U.K. and will not therefore adequately capture very short wavelength features, less than 1.5 m, such as dipped joints. The system also does not capture wavelength greater than 70 m, and being inertially based will only provide data above 30 km/h. Other systems are now becoming available, capable of measuring

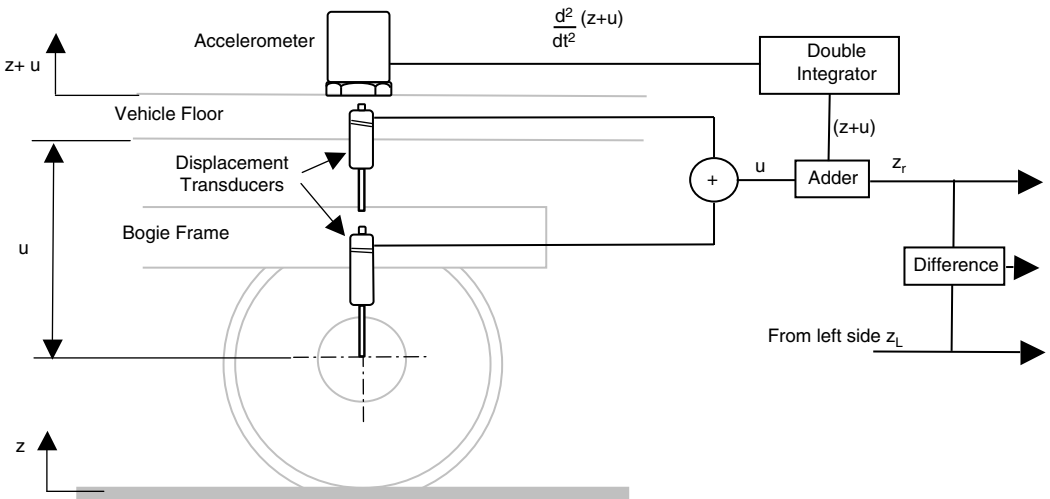
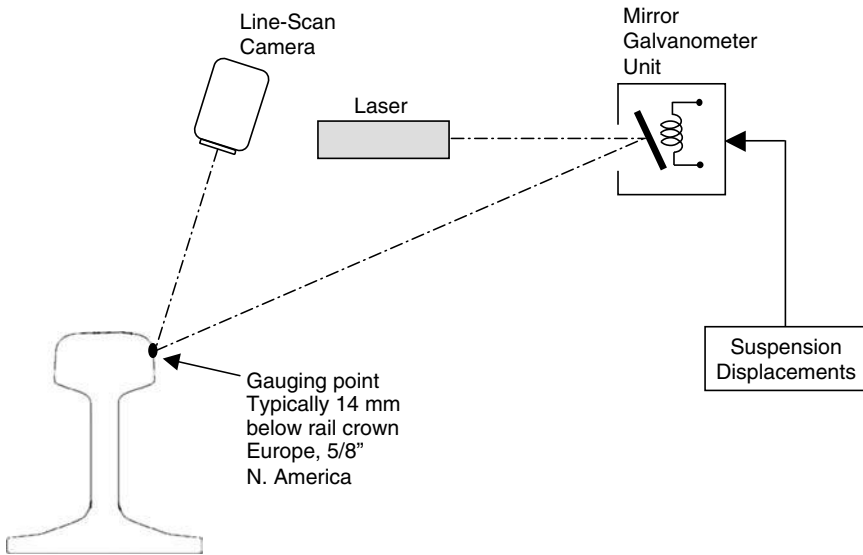


FIGURE 13.26 Vertical measuring system schematic.



**FIGURE 13.27** Noncontact measurement of rail position.

longer wavelengths up to at least 100 m, for example the Swedish STRIX system, developed by Banverket Production.

It can be seen that all the data required by the dynamics engineer to reconstruct the track geometry is available from this type of track geometry vehicle. However, several operations must be carried out to ensure it is suitable:

- As transducers are mounted on different parts of the vehicle, there will be an offset or lag in the some raw data channels depending upon the position of the transducers used to obtain them. This offset may also vary depending upon the direction in which the vehicle is running. These offsets are normally removed from the data at source, but this should be confirmed.
- Filtering data introduces both phase and amplitude distortion in the filtered signal compared with the original. As the cross-level and gauge channels are not filtered, the filtering will also introduce an offset between channels. Clearly, this offset is not a realistic representation of the real track geometry and, as it can significantly affect simulation results, it must be corrected. This is done by re-passing (backfiltering) the data through a filter of the same design as that originally used. This restores both the phase and amplitude distortions caused by filtering.

Track recording vehicles may also record a range of other parameters such gradient and cant deficiency and may derive other measures, for example, track twist or cyclic top from the raw data.

#### **D. CHORD OFFSET MEASURING SYSTEMS**

Vehicle based chord offset measuring systems rely on the same three-point measurement of versines described in Section VI.A above. In this case however, the chord is the vehicle wheelbase and a third wheelset or bogie is placed in the centre of the vehicle to provide the measurement at the 1/2-chord position. As discussed above, versine measurements have the distinct disadvantage that they have an inherent geometric filtering effect that prevents them from “seeing” particular wavelength ranges. Attempts have been made to improve this problem including fitting a number



of measuring wheels so that a better reference line can be used. However, this offsets the essential benefit of this method, namely its simplicity.

Chord offset based measuring vehicles are no longer in common usage and data measured by such systems should be treated with some caution by the dynamics engineer for the reasons discussed.

## VII. EXAMPLES OF VEHICLE LABORATORY AND FIELD TESTS

This section provide examples of laboratory and field tests that may be commonly encountered by, or provide useful information for, the vehicle dynamics engineer. Whilst they are based specifically on U.K. practice, similar tests are employed by many railway administrations worldwide.

### A. STATIC/QUASI-STATIC TESTS

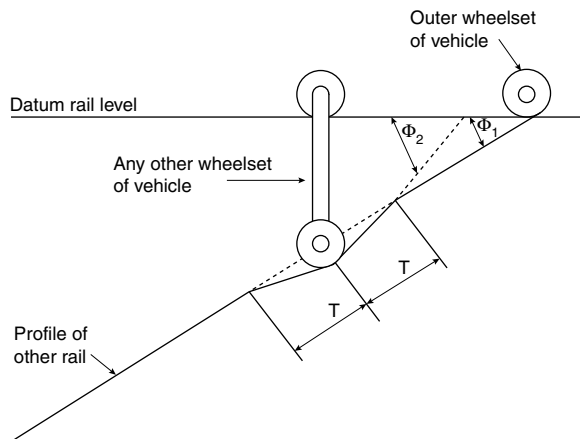
These tests are normally carried out on a vehicle in a specialist laboratory. The results may be used to gain confidence in the general behaviour of a vehicle model and also to estimate the additional (parasitic) stiffness present in the completed vehicle. However, as the dynamic behaviour can vary considerably from the static behaviour, some comparisons against dynamics tests (such as ride tests) are required to enable a vehicle model to be fully validated.

#### 1. Wheel Unloading Test

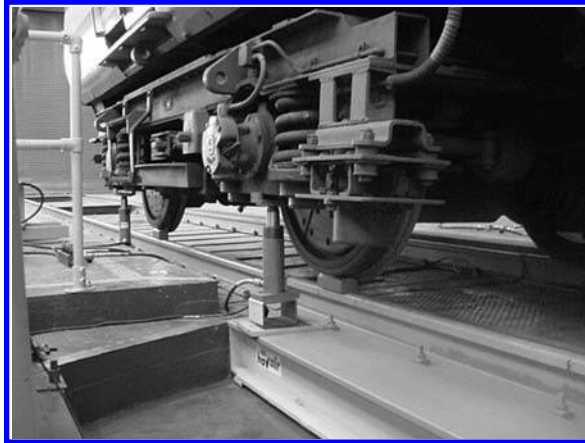
This test is detailed in Appendix A of Ref. 14. Packings are placed under the wheels on one side of the vehicle to reproduce the track twist feature shown in Figure 13.28. The resulting wheel loads are measured using a load cell and expressed in terms of the change from the static load  $\Delta Q/Q$ . The test is repeated to place each corner of the vehicle in turn at the bottom of the dip. The limiting  $\Delta Q/Q$  value is normally specified as 0.6 (Figure 13.29).

#### 2. Bogie Rotational Resistance Test

This test is detailed in Appendix B of Ref. 14. One bogie is placed on a turntable and rotated both clockwise and anticlockwise to an angle that represents the minimum operating curve radius for the



**FIGURE 13.28** GM/RT2141 Appendix A, track twist geometry for  $\Delta Q/Q$  tests. *Source:* RSSB — GM/RT 2141 Railway Group Standard.



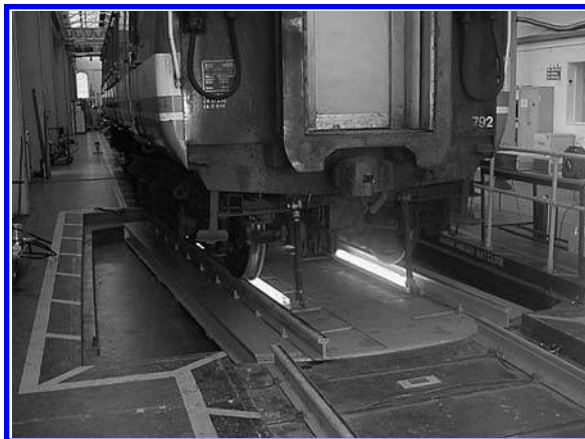
**FIGURE 13.29** Wheel unloading test.

vehicle. The test is usually performed at 0.2 and 1°/sec rotation speeds and the torque required to rotate the bogie is measured (Figure 13.30). Where yaw dampers are fitted, they may or may not be disconnected during the test depending on the test requirements. If yaw dampers with positional control (i.e., designed to blow-off at a particular bogie rotation angle) are included, the results from the test will reflect both the velocity and displacement dependent nature of such an arrangement. It follows that the vehicle model must be simulated in the same condition if comparable results are to be achieved. The resulting X-factor is calculated as follows:

$$X = \frac{\text{Body to bogie yaw torque}}{\text{Wheelbase} \times \text{axleload}}$$

The limiting value is 0.1 except for freight vehicles < 8 tonnes axle load.

The measured bogie rotation torques are particularly useful when confirming the behaviour of a vehicle model with friction sidebearers or airspring secondary suspensions. Typical examples of results from such tests for various vehicles are shown in Figure 13.31.



**FIGURE 13.30** Bogie rotation test.

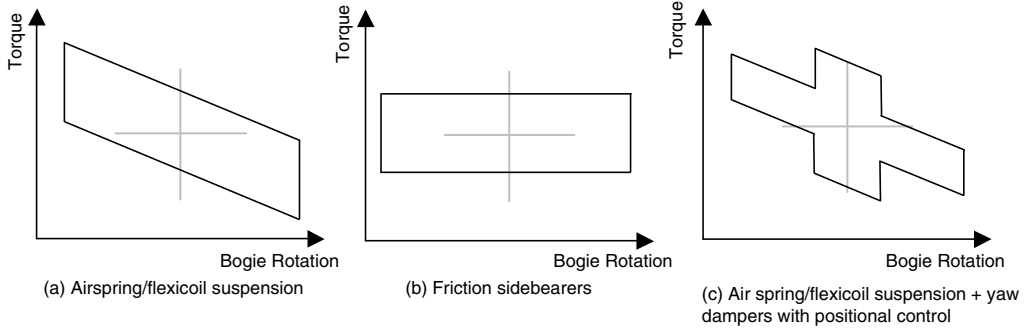


FIGURE 13.31 Typical bogie rotation test results for common suspension types.

Before a bogie rotation test is carried out, the bogie is normally moved slowly to its maximum rotation position to check clearances between all body and bogie mounted equipment.

### 3. Sway Test

Such tests are normally carried out to generate the input data for the kinematic gauging process or to verify that the vehicle will remain within a predetermined static envelope. Their usefulness to the dynamics engineer is in enabling a reasonable estimate of the parasitic stiffnesses (particularly in roll and lateral directions) to be made. A number of targets are fixed to the end of the vehicle at cantrail, waistrail, and solebar level. One side of the vehicle is raised in stages to approximately 10° of cant. A theodolite placed some distance from the end of the vehicle is used to measure the displacement of the targets as the vehicle is raised. Additional measurements of vertical and lateral suspension movement may prove useful when comparing test and model results (Figure 13.32).



FIGURE 13.32 Sway test. Source: Photo — courtesy of Serco Railtest Ltd.

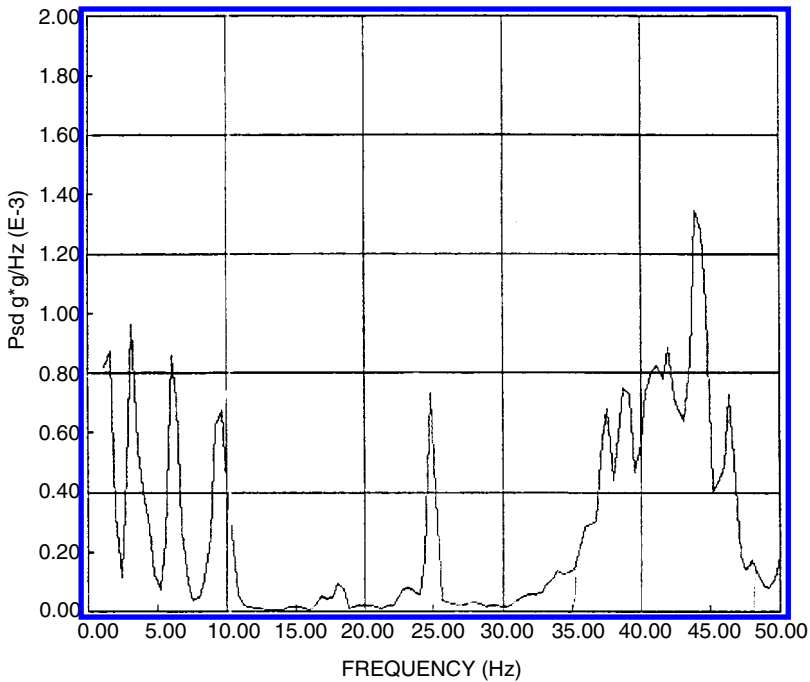


FIGURE 13.33 PSD of body vertical acceleration, mainline diesel locomotive.

#### 4. Body Modes Tests

Such tests are generally commissioned to provide the dynamics engineer with confirmation of the fundamental vibration modes of various parts of the vehicle. They are useful for comparison with the eigenvalue analysis used at the model checking stage as they will include the effects of all the parasitic stiffnesses. Tests are usually performed by disconnecting the vehicle dampers and then applying a sine sweep displacement signal in the direction of interest via an electric or hydraulic actuator. The response of the vehicle body is measured using both vertical and lateral accelerometers mounted at various positions along the length of the vehicle. These tests may also be used to confirm the first body bending mode.

#### B. DYNAMIC TESTS

Vehicle ride tests provide the means to validate the dynamic behaviour of a model. The vehicle will typically be instrumented with accelerometers at various positions on the body, bogie frame, and axle boxes, and this may be complemented with displacement transducers to measure primary and secondary suspension displacements. To be of maximum value for model validation purposes, the data should also include the vehicle speed and location and be accompanied by recent track recording coach data for the test route and measured wheel profiles from the vehicle at the time of the test.

Raw acceleration data from ride tests is normally expressed in terms of the power spectral density (p.s.d.) of the signal vs. its frequency range, as shown in the example in Figure 13.33. The test data p.s.d. plots are likely to include peaks representing the rigid body modes, flexible modes, and rail length passing frequencies. They may also include peaks associated with the traction unit, originating from vibration of the engine or driveline. The results may be further sorted into speed banded p.s.d. to assist in identifying speed dependent effects such as excitation of various

body/bogie modes and response to cyclic track geometry. Additionally, the test data may be analysed in sections according to track type to determine the vehicle's response to changes in track construction (welded, jointed etc.) and track quality. Care should be taken to ensure that a full understanding of the vehicle behaviour is reached under a range of conditions including any evidence of hunting or of body modes being driven by other modes such as bogie pitch.

The measured wheel profiles, track geometry, and speed data can be used to set up a time stepping integration (transient analysis) for the model in question, simulating the test run and outputs. The outputs from this simulation will be accelerations (and displacements) from specified positions on the vehicle such that they replicate the signals seen by the real instrumentation. Test data and model results can then be compared both in terms of time histories and power spectral densities. It is important to note that if the test data is filtered, the characteristics of the filter must be known to allow the simulated data to be treated in the same way.

Wheel–rail forces themselves can also be validated if test data is available from force–measuring wheelsets. However, their use may be restricted due to their high cost and the difficulties of acceptance.

## REFERENCES

1. CEN. *Railway Applications — Testing for the Acceptance of Running Characteristics of Railway Vehicles — Testing of Running Behaviour and Stationary Tests*. prEN14363, December 2004.
2. UIC: *Testing and Approval of Railway Vehicles from the Point of View of their Running Behaviour — Safety — Track Fatigue — Ride Quality*. Code 518 OR, Draft September 2003.
3. Turner, J. and Hill, M., *Instrumentation for Scientists and Engineers*, Oxford University Press, Oxford, 1999.
4. Serridge, M. and Licht, T. R., *Piezoelectric Accelerometers and Vibration Preamplifier Handbook*. Bruel and Kjaer, Revised November 1987.
5. Berther, T., Gautschi, G. H., and Kubler, J., Capacitive accelerometers for static and low-frequency measurements, *Sound Vib.*, June, 1996.
6. Window, A. L. and Holister, G. S., Eds., *Strain Gauge Technology*, Applied Science Publishers, Barking, 1982.
7. Gabriellsson, O., Analysis of a New Method for Measuring Wheel–Rail Forces, Report 1996, 24, KTH Railway Technology, 1006.
8. Sinclair, I., *Sensors and Transducers*, 3rd ed., Butterworth Heinemann, London, 2001.
9. Anon: Fundamentals of Non-Contact Speed Measurement Using Doppler Radar. *Application Note 1000 Rev 1.1*, GMH Engineering, 2003. (Available from <http://www.gmheng.com>).
10. Anon: GPS Guide for Beginners. *Garmin Corporation Part no 190-00224-00* Revision A, 2000. (Available from <http://www.garmin.com>).
11. Lewis, R. B., Track recording techniques used on British rail, *IEE Proc.*, 131(3), Part B, May 1984, .
12. Lewis, R. B. and Richards, A. N., A compensated accelerometer for the measurement of railway track crosslevel, *IEEE Trans. Veh. Technol.*, 37(3), August 1988.
13. Track Recording Handbook. *Railtrack Line Code of Practice RT/CE/C/038*, Issue 1, Rev B, February 1996.
14. Resistance of Railway Vehicles to Derailment and Rollover. *Railway Group Standard GM/RT2141*, Issue 2, October 2000.

---

# 14 Roller Rigs

*Weihua Zhang, Huanyun Dai, Zhiyun Shen, and Jing Zeng*

## CONTENTS

I.	Introduction .....	458
II.	The History of Roller Rigs .....	459
III.	The Test Technique and Classification of Roller Rigs .....	460
IV.	Examples of Roller Rigs .....	462
A.	The Chengdu Roller Rig .....	462
1.	Structure of the Chengdu Roller Rig .....	462
a.	Degrees of Freedom of the Rollers .....	462
b.	Components of the Roller Rig .....	464
i.	Test Unit .....	464
ii.	Driving System .....	465
iii.	Hydraulic System .....	466
iv.	Monitoring System .....	466
v.	Data Acquisition and Processing System .....	466
vi.	Auxiliary System .....	466
c.	Function of the Roller Unit .....	467
d.	Curve Simulation .....	467
2.	Characteristics of the Chengdu Roller Rig .....	469
3.	Scope of Test Function .....	469
B.	Naples Roller Rig .....	470
1.	Structure .....	470
a.	Power Supply System .....	470
b.	Overhead Line Simulator .....	471
c.	Roller Test Bench .....	471
d.	Drive and Brake System .....	472
e.	Control System .....	472
f.	Data Acquisition System .....	473
g.	Lifting Equipment .....	473
2.	Function .....	473
C.	The Tokyo Roller Rig (RTRI) .....	474
1.	Roller Unit .....	474
2.	Flywheel Assembly .....	476
3.	Driving Device .....	476
4.	Hydraulic Power Source .....	476
5.	Bogie Holding Device .....	476
6.	Building and Foundations .....	476
D.	The Pueblo Roller Rig (TTC) .....	476
V.	Operation and Results .....	477
A.	Test Methods .....	477

1.	Status of the Test Vehicle .....	478
2.	Status of the Roller Rig .....	478
3.	Stability Test .....	478
4.	Dynamic Simulation Test .....	479
5.	Curve Simulation Test .....	480
6.	Power Test .....	481
7.	Modal Analysis Test .....	481
8.	Storage Security Test .....	481
B.	Differences between Roller and Track .....	481
1.	Differences of Geometry Relationship .....	481
a.	Calculation Method of Wheel–Roller Geometry Relationship .....	481
b.	Geometry Difference of Wheel–Rail Contact and Wheel–Roller Contact .....	484
2.	Difference in Creep Coefficient .....	487
3.	Differences in Stability .....	489
4.	Difference in Vibration Response .....	491
5.	Difference in Curve Simulation .....	493
C.	Influence of Setup Errors on Roller Rig Vehicle Stability .....	494
1.	Diameter of the Rollers .....	494
2.	Gauge of the Rollers .....	494
3.	Cant of the Rollers .....	494
4.	Coefficient of Friction on the Contact Surface .....	496
5.	Vehicle Position on the Roller Rig .....	496
6.	Yaw Angle between Wheelset and Roller Axle .....	497
D.	Examples of Comparison Results with Theoretical Analysis .....	497
1.	Stability Test .....	498
a.	Stability Tests and Theoretical Analysis of a Freight Car .....	498
b.	Stability Test and Theoretical Analysis of a Passenger Car .....	500
2.	Ride Comfort Test and Theoretical Analysis of a Passenger Car .....	502
VI.	Conclusions .....	504
	References .....	504

## I. INTRODUCTION

Experience teaches us that the complete development cycle (design, prototype, railway line tests, industrialisation, production start up, in service adjustment, and corrective action) for a new vehicle featuring a significant level of systems and technological innovation is very time-consuming. The duration and efficiency of prototype experimentation activities is a key element in the development of new rolling stock, and is instrumental in terms of technical and economic success. An awareness of this technological and competitive development has led the railway industry or research institutes to commit extensive financial and technical resources to the creation of test facilities. The main objectives are to reduce the time (and therefore the cost) of testing new vehicles, to make as wide a range of tests available as possible, in order to achieve maximum levels of performance, reliability, and availability in the shortest possible time.

A roller rig is a type of railway vehicle testing plant. First, it is a system capable of testing a vehicle in a running condition without field tests, and second, it allows the study of interaction between a railway wheel and the rail.

The application of roller rigs to the study of vehicle system dynamics and the development of high-speed trains and other railway vehicles has become more widespread in recent decades. Roller rigs are used by researchers and railway organisations around the world to assist in understanding

the behaviour of railway vehicles and developing faster, safer and more efficient railways. Roller rigs have contributed to many current designs of railway vehicles.

Roller rigs have been proved useful for both basic research and development of innovations in suspensions and vehicle components. Full-scale roller rigs offer the advantages that the experiments are independent of weather conditions, individual phenomena can be investigated, and the experiments and the constraints as well as the particular conditions are reproducible.

Full-scale roller rigs have been proven as powerful tools, not only for the demonstration of vehicle dynamics for students, but also for the validation of theoretical work and the test of new concepts of innovative vehicle designs.

## II. THE HISTORY OF ROLLER RIGS

Roller rigs were originally used for the investigation of the performance of steam locomotives over 100 years ago. One of the earliest such plants was built at the Swindon works of the Great Western Railway in 1904 (see [Figure 14.1](#)).<sup>1-3</sup> The rollers of this rig were moveable and could be adjusted so that the centre of each driving wheel was exactly about the centre of each roller. High speeds could be attained while the engine remained stationary, and a braking arrangement on the rollers measured the traction power of the locomotive at various speeds.

In 1957, a full-scale roller rig with two axles was used at the Railway Technical Research Institute, Japan, which used an eccentric roller to create a sinusoidal excitation. In about 1960, tests of bogies commenced on the newly built full-scale roller rig. This roller rig was put into use for about 30 years and played a very important role in studies related to protection against freight car derailment, regenerative braking, Shinkansen bolsterless bogies, etc. In order to meet the demand for high-speed vehicle tests, a new four-axle, full-scale roller rig with the facility for roller lateral and vertical excitations began construction in 1987.

A roller rig was built in Vitry, France in 1964 by CAFL Company, which allows lateral and vertical motions of the roller on each axle simultaneously using simple hydraulic control methods. By using the roller rig, the vertical, lateral, and yaw vibration frequencies, amplitudes and resonance can be measured. In particular, the influence of the change of the vertical and lateral excitation forces due to the impart force on the vehicle running performance can be studied and the running safety and ride performance can also be investigated.

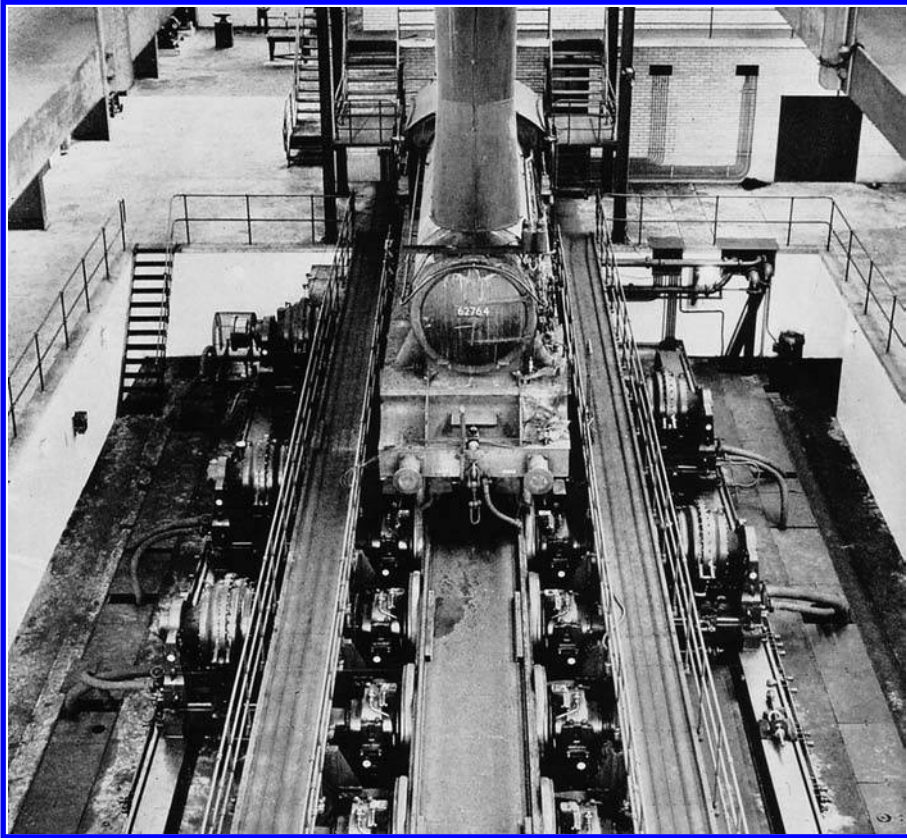
The roller rig in Berlin, Germany, was built in 1967. This roller rig allows evaluation of traction equipment, acceptance tests for vehicle springs, and assessment of braking systems.

The construction of a roller rig at the BR Research Centre in Derby, began in 1959 and was completed in 1971. This roller rig had the capacity to assess braking power, resonant vibration, and vehicle stability. Latterly, the roller rig has been modified to a modal analysis test stand mainly used for the vibration analysis of vehicle suspension systems.

In 1977, a full-scale roller rig was built in Munich, Germany at Deutsche Bahn AG. The rollers have four degrees of freedom including vertical, lateral, inclination, and rotation. The servohydraulic excitation control system was adopted for the roller rig and can accurately simulate track conditions for the dynamic simulation of a vehicle operating on tracks. The rig is mainly utilised for the measurement of the dynamic performance of vehicles and determination of the effects of vehicle modifications on the system performance. The Munich roller rig has played a very important part in the development of ICE high-speed trains.

In 1978, a roller rig, called the roll dynamics unit, with vibrations applied through the wheels to simulate track conditions, began operation at Pueblo, Colorado, U.S.A. The rig consisted of two separate test stands, one for roller-based testing, and another used as a vibration stand. The rolling stand can be used for hunting stability and traction power simulation tests and the vibration stand used for studies of suspension system features, vehicle system natural frequencies,





**FIGURE 14.1** One of the earliest roller rigs for steam locomotives at the Swindon works of the Great Western Railway.

fatigue strength, and freight load reliability, etc. The rig was designed for speeds of up to 480 km/h.

A roller rig called the curved track simulator was set up at the National Research Council in Ottawa, Canada. It consisted of two pairs of rollers in a flexible frame that permitted the yawing motion of the roller axle to simulate curving. The frame floated on hydrostatic bearings. Unfortunately, the roller rig has since been dismantled.

In 1995, a four-axle roller rig was built at the State Key Laboratory of Traction Power (Southwest Jiaotong University) at Chengdu, China. This roller rig was built for the optimum design and testing of railway vehicles. Each roller can vibrate in lateral and vertical directions. In 2002, two new sets of rollers were added, to form six axles, allowing locomotives to be tested on the rig. In China there are four roller rigs for whole vehicle tests and two rigs for bogie tests. Scale roller rigs have also been used in many research laboratories and this will be discussed in Chapter 15.

### III. THE TEST TECHNIQUE AND CLASSIFICATION OF ROLLER RIGS

The main aim of building roller rigs is to provide controlled conditions for investigation and optimisation of railway vehicle performance. The following situations for simulating vehicle operation can be carried out entirely or partly by a roller rig:

- To measure the stability of railway vehicles
- To study wheel–rail interactions

- To simulate the vibration of vehicles running on track with different irregularity conditions
- To simulate the process of train acceleration or braking

A roller rig acts as a track simulator; the rollers with rail profiles form an endless track. It cannot only simulate the running of vehicles on straight track by the rotation of the rollers, but can also simulate track irregularities by the excitation of rollers in a number of axes. When applying rotational resistance to the rollers, the roller rig can provide the traction force to simulate traction and braking effort of a vehicle. Therefore, railway vehicle test facilities can be classified as:

- RTU — a pure rolling roller rig. RTU has the basic function of simulating railway vehicles running on a straight line, with or without traction forces.
- RVTU — a rolling and vibrating roller rig. RVTU not only simulates the vehicles on straight track, but can also simulate track irregularities by the excitation of the rollers, again, with or without traction forces.
- VTU — a test rig using a short vibrating rail under each vehicle wheel, to reproduce track irregularities. The VTU cannot simulate wheel–rail contact and traction power as there is no wheel rolling motion.

Table 14.1 classifies roller rigs from around in the world based on the above groupings. Most are of type RTU. Owing to their combined features, RVTU rigs are inherently more useful in the development of railway vehicles.

The track irregularity shown in Figure 14.2 can be considered to have four components: gauge, cross level, lateral alignment, and vertical profile. When the vertical and lateral disturbances of the left and right rails are indicated as  $z_L, y_L, z_R, y_R$ , the four type of track irregularities can be described as:

1. Gauge =  $(y_L - y_R)/2$
2. Lateral alignment =  $(y_L + y_R)/2$
3. Cross level =  $(z_L - z_R)/2$
4. Vertical profile =  $(z_L + z_R)/2$

Table 14.2 describes how to use the rollers to simulate track. Ultimately, a roller rig should be capable of simulating track irregularities and also curve negotiation, but to date, there is no rig which is able to simulate both. Only the roller rigs in Munich and in Chengdu possess part curving functions. The roller rig in Munich has now been decommissioned and the roller rig in Chengdu cannot simulate the rail movement along the tangent direction when gauge and lateral alignment irregularities exist. To achieve this function, the rollers must be able to yaw about their vertical axis.

**TABLE 14.1**  
**Classification of Roller Rigs**

Type	China Chengdu	Germany Munich	America Pueblo	Japan Tokyo	Germany Berlin	Italy Naples	France Vitry	China Qingdao	China Dalian
RTU			✓		✓			✓	✓
RVTU	✓	✓		✓			✓		
VTU			✓					✓	

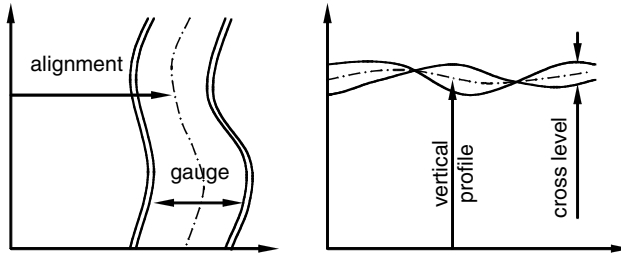


FIGURE 14.2 Track irregularity inputs.

## IV. EXAMPLES OF ROLLER RIGS

Roller rigs have been established for many years and the following section details five representative roller rigs that have played an important role in the development of railway vehicles.

### A. THE CHENGDU ROLLER RIG<sup>4,5</sup>

China has more than 70,000 km of railway line. The service speed of passenger trains on main lines has risen from approximately 50 to 70 km/h in the 1980s to approximately 140 to 160 km/h in the 1990s. This rise in vehicle speed has been mainly attributed to the use of test facilities, especially roller rigs. The successful application of these rigs within China has resulted in a growth in their use. There are now six roller rigs in service, but only the roller rig in Chengdu has function of rotation and vibration combined together.

The roller rig in Chengdu was developed by the State Key Laboratory of Traction Power (Southwest Jiaotong University). Each roller can move in vertical and lateral directions independently under servo control. Design of the roller rig began in 1989 and it came into service in 1995. From 1995 to 2005 more than 50 railway vehicles were tested. The original roller rig had four roller sets (allowed testing of up to four-axle vehicles) with the two rollers of each set constrained to have the same rotational speed. These constraints meant the roller rig could only simulate a four-axle vehicle running on straight track, with a maximum gauge variation of between 1435 and 1676 mm. As the rig was heavily utilised in the development of new railway vehicles, it was extended to six roller sets and the structure improved during 2002. Four roller sets of the new modified rig have the ability of gauge variation of between 1000 and 1676 mm and the two rollers of each roller set can be run at different rotational speeds. This roller rig is shown in Figure 14.3.

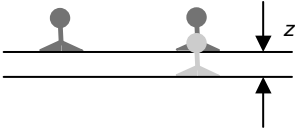
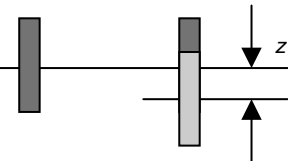
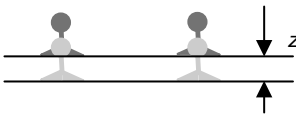
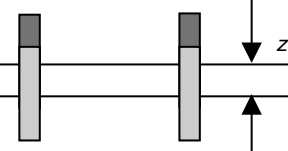
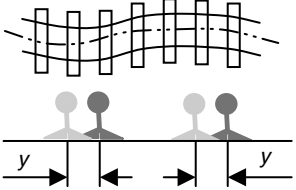
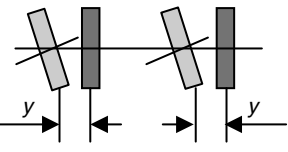
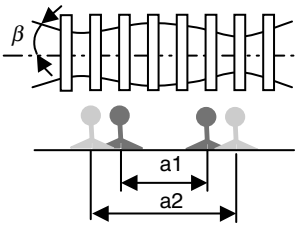
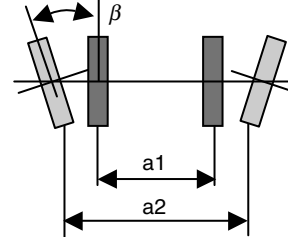
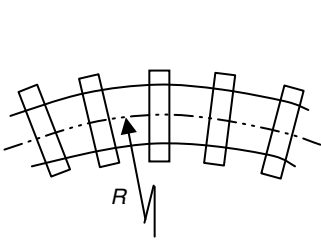
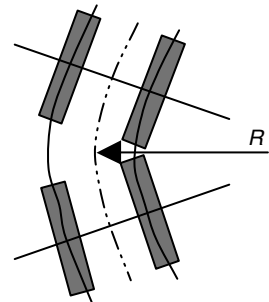
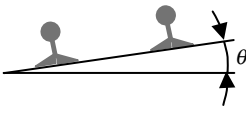
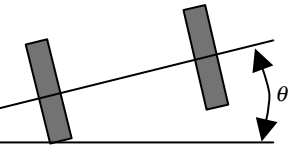
#### 1. Structure of the Chengdu Roller Rig

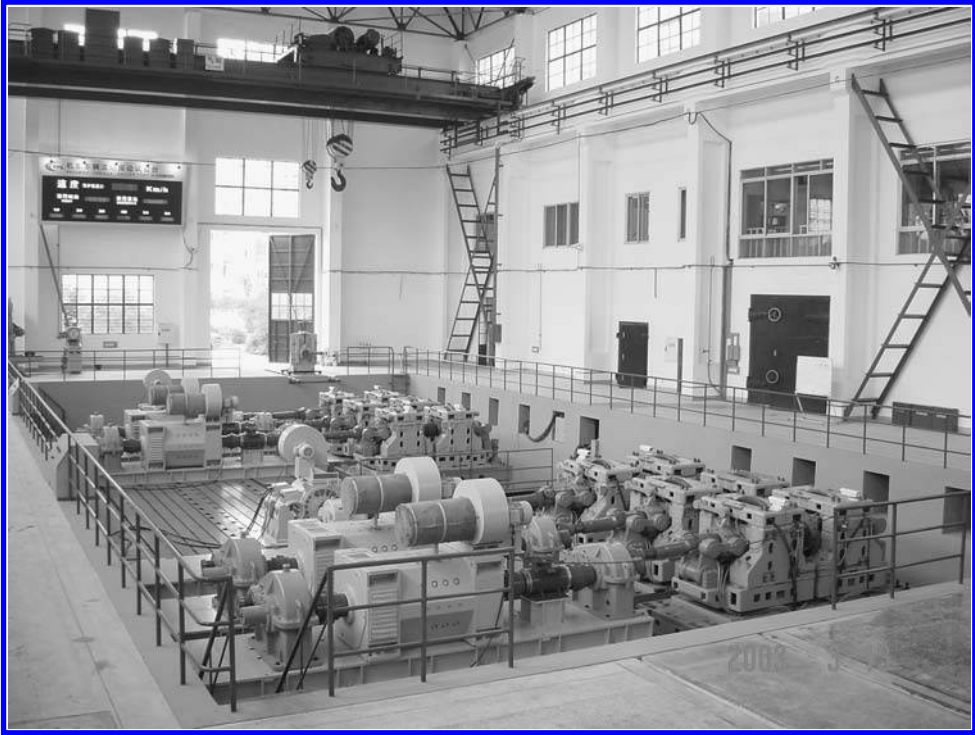
##### a. Degrees of Freedom of the Rollers

An ideal roller rig should have the degrees of freedom as described in Table 14.2. In reality, considering design, manufacturing, and financial constraints, the most useful of the described degrees of freedom were chosen. The degrees of freedom of the Chengdu roller rig are shown in Figure 14.4:

- Movement of the two rollers independently in the  $Y$  direction simulates the track irregularities of gauge and lateral alignment.
- Movement of the two rollers independently in the  $Z$  direction simulates the track irregularities of cross level and vertical profile.
- Turning of the two rollers about the  $X$  axis is to simulate the cant angle in curving.

**TABLE 14.2**  
**Relationship between Status of Rails and Rollers**

No.	Type of Irregularity	Status of Rails	Status of Rollers	Remark
1	Cross level			Left roller and right roller move relatively vertically
2	Vertical profile			Left roller and right roller move synchronously vertically
3	Alignment			Left roller and right roller move in lateral and yaw directions
4	Gauge			Left roller and right roller move in lateral relatively and corresponding yaw motion
5	Curve			Left roller and right roller are set in curved position and rotate at different speeds
6	Cant in curve			Left roller and right roller tilt synchronously



**FIGURE 14.3** The roller rig in Chengdu (without flywheel).

- Rotation of the two rollers at the same speed about the  $Y$  axis is to simulate the forward speed of vehicle on straight track, and with different speeds is to simulate curving.
- Turning of the two rollers about the  $Z$  axis allows simulation of the track tangent in a curve.

The linear motions of the two rollers in the  $Y$  and  $Z$  directions and also the rotation of the two rollers about the  $Y$  axis are controlled during the roller rig operation. The rotation about the  $X$  and  $Z$  axes is applied only for curve simulation which is preset before the test. That is, this roller rig can simulate straight track and circular curved track with track irregularities.

#### *b. Components of the Roller Rig*

Hydraulic actuators provide the movements of the two rollers in  $Y$  and  $Z$  directions, and the rotation about the  $Y$  axis is driven by the motor. The roller rig is composed of several subsystems including the test unit, driving system, hydraulic system, monitoring system, and data acquisition and processing system. The whole test system of the roller rig is shown in [Figure 14.5](#). The main power supply for the rig comes from the railway power supply with 25 kV and 50 Hz or a low-level power supply of 380 V at 50 Hz.

##### *i. Test Unit*

The roller rig has six test units. They are independent and can be moved according to different vehicle configurations. Each test unit consists of a roller unit and a driving unit, as shown in [Figure 14.6](#). The driving unit can provide different rotational speeds and torque to the rollers, via a double-articulated universal joint, according to the task required. Within the driving unit, there is a DC motor, two

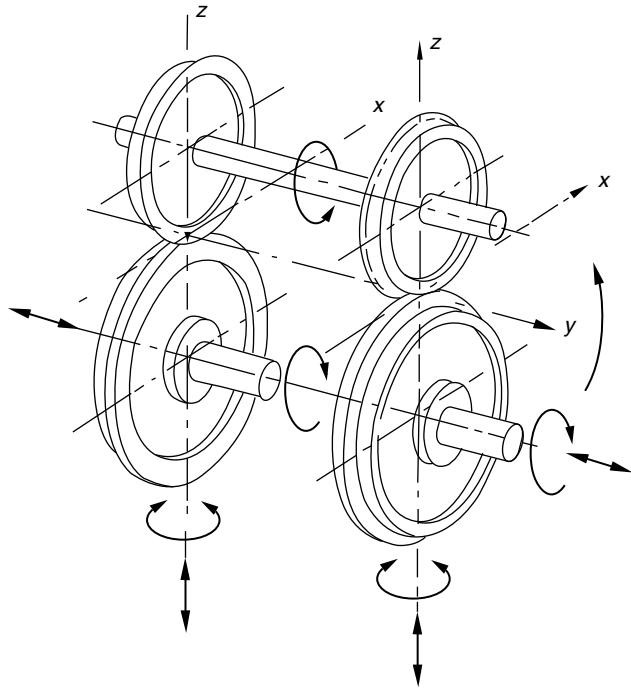


FIGURE 14.4 Freedom of the rollers.

gearboxes, a flywheel, and a torsion meter, all fixed to a welded frame. The motor can work as a driving motor or a generator according to the requirement of driving or braking. The flywheel is used to maintain the running stability of the roller rig and to simulate the inertia of the vehicle. Gearbox II is used to accelerate the flywheel, while gearbox I is used to apply different rotational speeds and torques by setting the transmission ratio as 1:1, 2:1 (for high torque) or 1:2 (for high speed).

ii. *Driving System*

The driving system consists of a remote control computer, digital controller, converter, motor excitation, resistance, and motor. Using a feedback control technique, the operator can control the motor operation according to the defined running speed or operating torque through the remote

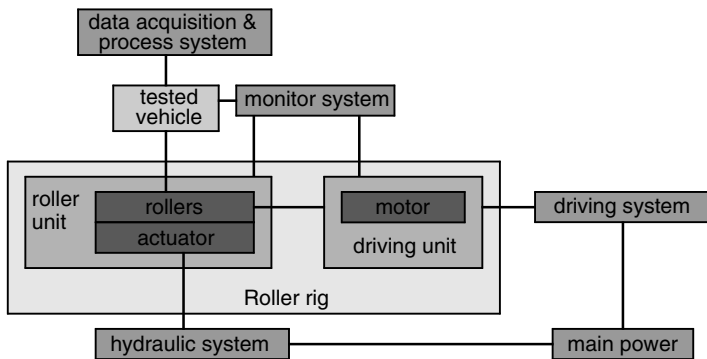


FIGURE 14.5 System of roller rig.

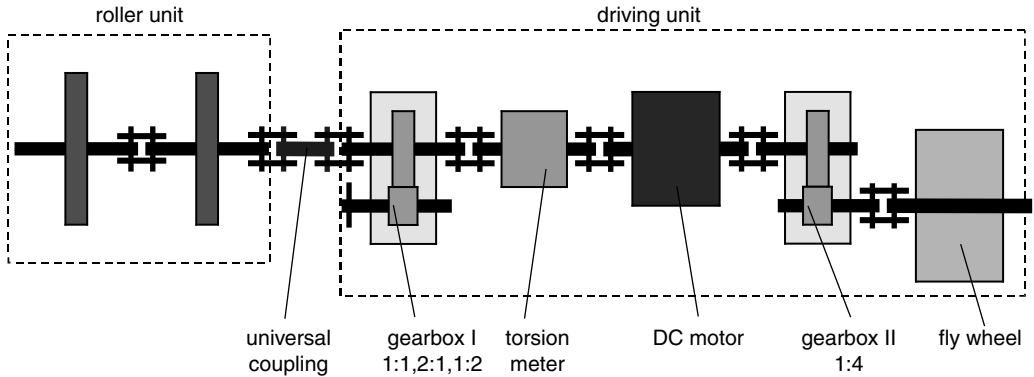


FIGURE 14.6 Test unit.

control computer. The maximum difference between the rotation speeds of the six test units can be controlled within 0.5%. Therefore, the driving system can ensure that the six roller sets are rotating almost synchronously without any mechanical connection.

### iii. Hydraulic System

The movements of the rollers in the vertical and lateral directions are provided by the lateral and vertical hydraulic actuators; the roller rig is a complicated system with a total of 24 actuators for the 12 rollers. By using the digital controller, motion of the actuators is controlled by displacement-based PID feedback control.

### iv. Monitoring System

The operation of the roller rig is surveyed by a monitoring system. The monitoring system can display the roller rig running speed and torque, temperature of bearings, and lubricating oil. Through the 12-channel video system, the status of the roller rig and the tested vehicle, the contact condition of roller and wheel can be monitored. The system also has the function of overload protection and safety interlocking.

### v. Data Acquisition and Processing System

According to the railway vehicle test evaluation standards, the responses of the tested vehicle should be recorded during the test. The data acquisition and processing system can measure the signals of displacement, velocity, acceleration, strain, pressure, temperature, voltage, current, etc. All signals can be measured, conditioned, and sent to the acquisition computer via a network link. Up to 200 channels of data can be acquired.

The terminal operating computers for the driving system, hydraulic system, monitor system, and data acquisition and processing system are arranged on a desk in the control room, as shown in Figure 14.7.

### vi. Auxiliary System

The following are auxiliary facilities of the roller rig:

- Test shed — the test shed with a length of 72 m and width of 24 m is divided into two sections, one section is for the roller rig, the other for test preparation and locating facilities for component tests.
- Component test stands — vehicle suspension parameter measurement, fatigue test, etc.
- Power supply — there are two power supply systems. The civil power system with three-phase AC 10 kV is used for the driving motors of the roller rig. The hydraulic system and





**FIGURE 14.7** The control room.

other systems use 380 V. The railway power system of AC 25 kV is to power locomotives under test.

- Crane — there are two gantry cranes in the test shed with capacity of 50 t each.

#### *c. Function of the Roller Unit*

The roller unit is the key part of the roller rig. A schematic of the roller unit of the Chengdu rig is shown in [Figure 14.8](#). The roller is supported by a U-shaped frame via a pair of roller bearings. The rollers can rotate within the U-frame via a universal coupling. The rollers, together with the U-frame, can move in a vertical direction using the hydraulic actuators. The shaft of the roller is fixed within a bearing housing, which can be pushed or pulled through a lever arm by a hydraulic actuator. The upper part of the roller unit is installed on a subframe, which can rotate via a ring bearing, that allows rotation through a predefined angle to suit a simulated curve radius. The left and right rollers can be moved independently; a gear coupling allows the transfer of rotational torque and variation in drive speed of the left and right rollers. The subframe also allows the upper part of the roller unit to tilt using a lift actuator to simulate the cant angle of a curve. The whole roller unit is seated on a base frame, which allows longitudinal movement to accommodate variation in vehicle wheelbase.

#### *d. Curve Simulation*

A special characteristic of the Chengdu roller rig is that the left and right roller can rotate at different speeds, which means the roller rig can simulate a wheelset in a curve. This function is realised by a complicated differential driving system. The principle of the differential driving system for one test unit is shown in [Figure 14.9](#). It can be seen that there is no direct coupling between the left and right roller. The motor drive is divided into two by the cone gear in gearbox I.



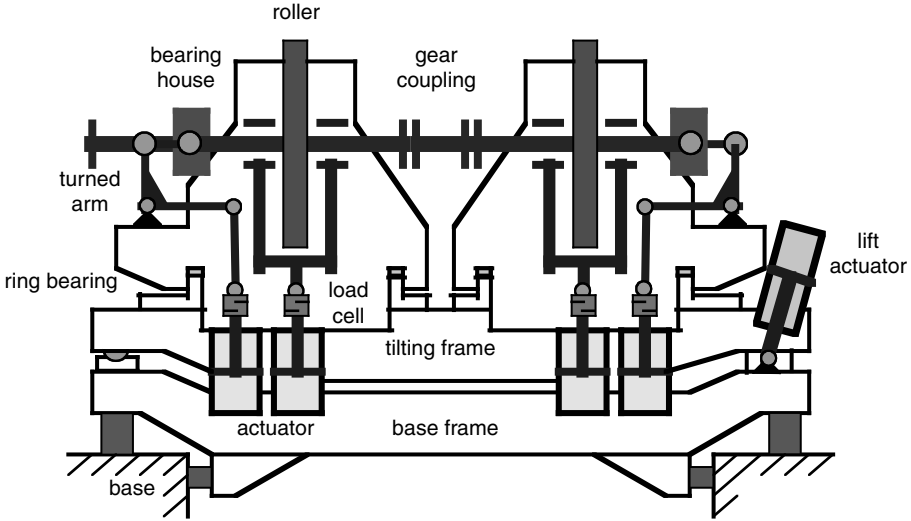


FIGURE 14.8 Sketch of Chengdu roller unit.

One gear takes drive directly to the right roller through a shaft and universal coupling. Another drive is transferred to the cone gearbox II and then to the differential gearbox. When the drive passes through the differential gearbox, the speed of input shaft can be different from the speed of the output shaft according to the speed of the differential speed motor. The modified drive speed is then transferred to the left roller through cone gearboxes III and VI. This allows the left roller to run with the same rotational direction as the right roller but at a different speed to simulate the speed differential between left and right wheels during curving. Figure 14.10 shows a picture of the roller rig components for simulating curve negotiation.

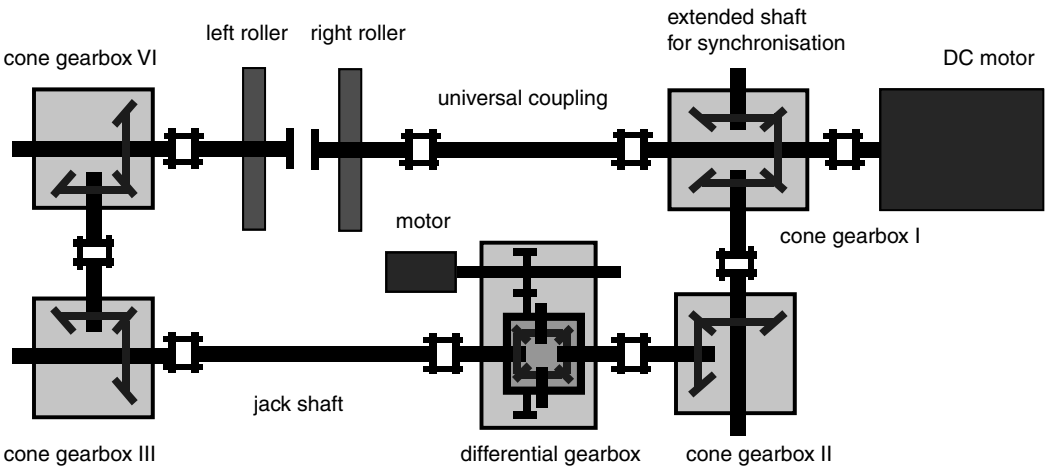


FIGURE 14.9 Sketch of differential driving system.

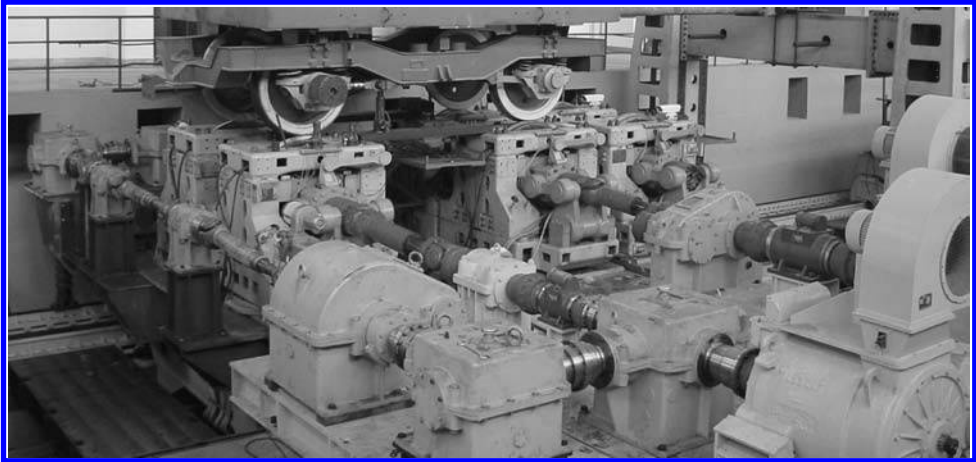


FIGURE 14.10 The roller rig with differential driving system.

**TABLE 14.3**  
**Main Characteristic Parameters of Chengdu Roller Rig**

Exciting in Vertical		Exciting in Lateral	
Maximum Frequency $f_{vmax}$	30 Hz	Maximum Frequency $f_{vmax}$	30 Hz
Maximum Amplitude $A_{vmax}$	$\pm 10$ mm	Maximum Amplitude $A_{hmax}$	$\pm 10$ mm
Maximum Acceleration $a_{vmax}$	$\pm 4$ g	Maximum Acceleration $a_{hmax}$	$\pm 5$ g
Maximum traction force per axle $F_e$	10 T	Maximum axle load $Mw$	25 T
Maximum motor power/brake $W$	1200/1500 kW	Maximum speed $V$	450 km/h
Maximum cant angle $\phi_{max}$	$7^\circ$	Distance between bogies $L$	4 ~ 22 m
Bogie wheelbase $l$	1600 ~ 3500 mm	Range of gauge $A_0$	1000 ~ 1676 mm
Minimum curve radius $R$	200 m	Maximum wheelset numbers $Nz$	6

## 2. Characteristics of the Chengdu Roller Rig

The main parameters of the roller rig are shown in Table 14.3. The rig can test conventional four-axle and up to six-axle railway vehicles. The roller rig can run in either active or passive mode for vehicle testing depending on whether testing a locomotive or unpowered trailer car.

## 3. Scope of Test Function

The Chengdu roller rig is not only used to verify the performance of railway vehicles, but also for basic research duties, such as wheel–rail contact mechanics, wear, noise, etc. Its main functions are listed below:

1. Basic research on wheel–rail creep theory
2. Study of derailment mechanisms
3. Hunting stability
4. Dynamic response and ride comfort
5. Curve simulation and operating safety
6. Vibration, mode shape, dynamic stresses of railway vehicles and their components
7. Goods load safety

8. Driving or braking power tests and optimisation of train operation
9. Wheel–rail wear, adhesion and control
10. Wheel–rail interaction forces and the forces on components of vehicle systems
11. Wheel–rail noise and noise reduction
12. Static and dynamic parameter measurement of railway vehicle systems

## B. NAPLES ROLLER RIG<sup>7–9</sup>

The roller rig at the Ansaldo Transport Research Centre in Naples, Italy, was completed at the end of 1992. Initially, the roller rig has been configured with four rollers (axle), but is designed to allow future upgrading to test six-axle vehicles. The rig is mainly used to test railway locomotives for traction simulation. Its rollers can only rotate about the X-axis to simulate running on straight track. Figure 14.11 shows an overview of the Naples roller rig. The rig allows the testing of vehicles with the following characteristics:

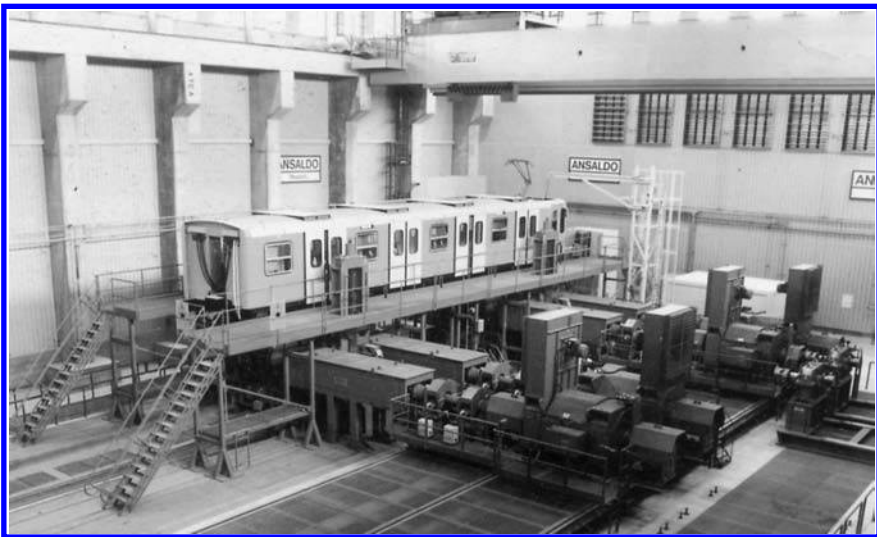
- Total weight (max.) of a four-axle vehicle (25 t/axle, 250 km/h) 100,000 kg
- Wheel diameter 500–1400 mm
- Maximum traction effort per axle 100 kN
- Maximum speed (22 t/axle with 1500-mm diameter rollers) 300 km/h
- Range of gauges available 600–1700 mm
- Bogie wheelbase 1400–3500 mm
- Distance between bogies (four-axle vehicle) 5200–22,000 mm
- Maximum continuous power at the axle 1500 kW

### 1. Structure

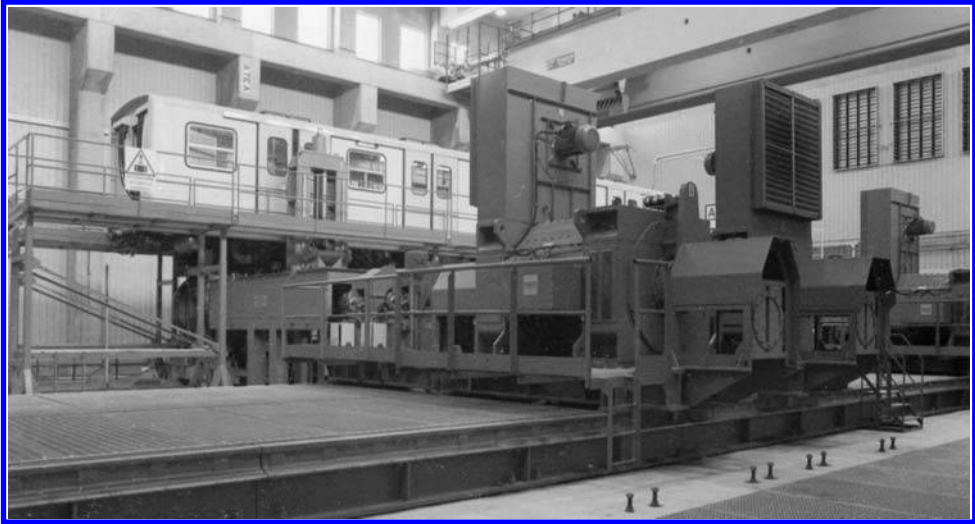
The Naples roller rig is basically composed as follows.

#### a. Power Supply System

Since the power supply of railway systems in Europe varies, the difference not only being voltage but also current (AC or DC), the power supply system of the Naples roller rig is flexible.



**FIGURE 14.11** Overview of the Naples roller rig.



**FIGURE 14.12** Overview of drive and brake systems of the Naples roller rig.

The primary power is ENEL (Italian National Power Utility) power. Using a 60/9 kV transformer, the ENEL power transforms to 9 kV using a three-phase bus bar in the power room. The power used for locomotives is 25 kV, 50 Hz or 15 kV,  $16\frac{2}{3}$  Hz; they are drawn from the 9 kV bus bar (Figure 14.12).

#### *b. Overhead Line Simulator*

Power is fed to the test locomotive through its own pantograph. The real field power supply is characterised by:

- A no-load voltage at the current terminal point which varies on the basis of absorption by other vehicles
- Power supply system equivalent resistance and inductance which vary as a function of vehicle position
- Limited maximum power during recovery

The overhead line power supply control system makes it possible to reproduce the above conditions in the test facility (programmable). Depending on the tests to be performed, the following types of overhead line power feeding system controls are possible:

- Power supply in stable condition
- Power supply in variable condition
- Power supply with voltage jumps
- Power supply with line section simulation (programmed as a function of the distance covered by the vehicle during simulation)

#### *c. Roller Test Bench*

This is the main part of the railway vehicle test facility. The rollers are fixed on an axle. Normally, the gauge of two rollers on one axle is 1435 mm. To test vehicles with different gauge, rollers must be exchanged with other rollers from a “roller park.” Each of these rollers is mechanically

connected to an electrical machine regulated to deliver a resistant torque with an amplitude and sign which varies on the basis of the dynamic train behaviour.

#### d. Drive and Brake System

The drive and brake system is of a DC type and suited to the typical operation of each traction motor. Each roller is connected to a generator/motor. The roller drive control functions are as follows:

- Speed control
- Torque control
- Synchronous roller speed control

Speed and torque control are available for synchronous roller speeds. The driving system guarantees precision of the order of 0.1%.

When the motor works as a generator, the power will be recovered and directed back to the 9 kV bus bar, a bidirectional conversion section is used. The converters are very powerful; the main characteristics of one converter are 1600 kW, 750 V, 0/230/1065 rpm. This roller rig motor connects directly to the roller axle without a gearbox or braking device.

#### e. Control System

System automation is based on the distributed intelligence structure used for industrial process control. The control system, shown in Figure 14.13, comprises two levels. Level I, shown in Figure 14.14a, is the installation's basic automation and comprises:

- One test facility control desk (Figure 14.14b) — all functions indicated in OIS paragraph, signalling, and emergency control
- One locomotive control desk (Figure 14.14c) — vehicle initialisation, manual command entry, locomotive status display, and emergency control
- One mobile supervision desk — test facility operating status display, reference display for manual operation
- Four programmable logic controllers (PLC) — Data acquisition from the plant and vehicle/drives under test, control logic implementation, protection logic implementation

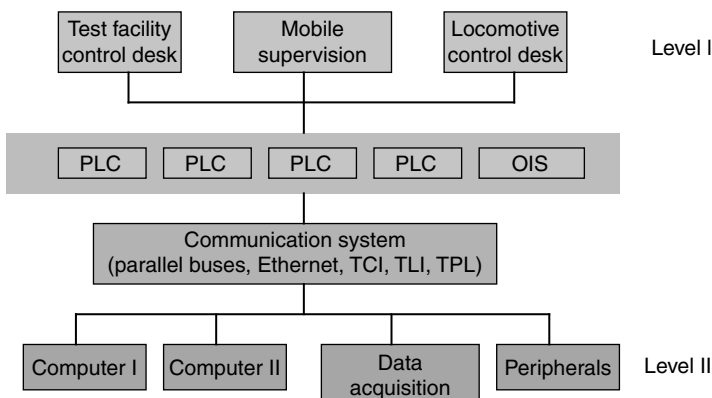


FIGURE 14.13 Control system.



**FIGURE 14.14** Control system: (a) level I of control system; (b) test facility control desk; (c) locomotive control desk.

- One operator interface system — configuration of plant video pages, configuration of basic plant data, selection of remote initialisation mode for test facility, remote control and status display of circuit breakers and disconnecters, manual command entry, test mode display, monitoring of data from test facility and test item, alarm display and printout

Level II is used to perform general plant control and supervision. Two computers with relative peripherals perform the following function:

- Computer 1 — control and supervision
- Computer 2 — development and backup

#### *f. Data Acquisition System*

This system acquires signals from the test vehicle during testing and can receive up to 400 signals, including high frequency and high voltage signals (1 MHz, > 1 kV), medium frequency signals (< 50 kHz), low frequency signals (< 1 kHz), and temperature signals. This structure includes:

- A computer with relative peripherals — data storage, data processing, data display
- Acquisition cards — A/D conversion
- Interface equipment — signal conditioning

#### *g. Lifting Equipment*

In order to position the vehicle on the rollers, two 160 tonne twin-bridge cranes operating in tandem are available. Four electromechanical jacks are also provided to lift the test vehicle and ensure perfect alignment between the wheel axes and the corresponding roller axes.

## **2. Function**

The Naples roller rig is a successful vehicle test facility for power testing. The test type can be short duration tests, long duration, or series of tests. A few examples of tests performed are described as follows:

1. Standard start-up test — the aim is to study the behaviour of vehicles in various start-up condition covered by specification and to verify the global vehicle traction effort, traction effort during transients, and speed as a function of time.
2. Standard simulated path test — the aim is to simulate the motor car (locomotive) running from station to station. The characteristics of the path, such as height profile of the line, radii of curvature, speed limits, tunnels, stops, no load voltage, are simulated.

3. Standard braking test — the objectives are verification of the electrical braking effort of the vehicle as a function, electrical and pneumatic braking effort, and speed.
4. Standard line voltage step variation test — the objectives are verification of vehicle operating continuity during traction and braking, verification of traction and braking effort continuity, and verification of performance with undue intervention of the vehicle protection devices.

### C. THE TOKYO ROLLER RIG (RTRI)<sup>10,11</sup>

The roller rig at the Railway Technical Research Institute (RTRI) of Japan was built in 1957 and has contributed greatly to research and development including the high-speed bogies used for the Shinkansen trains, countermeasures against derailment of freight cars, regenerative braking, and bolsterless bogies, and has also played an important role in technical cooperation with domestic and overseas railways. However, the old plant was becoming outdated and its functions limited, as 30 years have passed since its completion. It had become difficult to operate from the viewpoint of noise and vibration to the surrounding area. After 30 years of operation the plant was renewed with respect to the requirements of future railway vehicle development. The new facility was brought into operation at the end of 1989 (see Figure 14.15).

The general arrangement of the roller rig is shown in Figure 14.16. The maximum testing speed of the new roller rig is 500 km/h. One of the major additions to the rig was the facility for inducing vibrations using actuators.

The principal devices in the plant are outlined below:

#### 1. Roller Unit

Gauge	1,000 ~ 1676 mm(variable)
Minimum wheelbase	1600 mm
Maximum test speed	500 km/h
Maximum axle load	200 kN
Diameter of roller	1500 mm
Lateral displacement	0 ~ 1 Hz Max. $\pm 30$ mm 3 Hz Max. $\pm 10$ mm 10 Hz Max. $\pm 2$ mm Maximum acceleration 10 m/sec <sup>2</sup>
Vertical displacement	0 ~ 1.8 Hz Max. $\pm 12$ mm 25 Hz Max. $\pm 0.4$ mm Maximum acceleration 10 m/sec <sup>2</sup>
Rolling displacement	0 ~ 2 Hz Max. $\pm 0.011$ rad 15 Hz Max. $\pm 0.0006$ rad Maximum acceleration 5 rad/sec <sup>2</sup>

The principal functions of this plant consist of rollers on which the bogie is tested, hydrostatic bearings, hydrostatic couplings, actuators for lateral vibration, movable beds, etc.

A set of rollers and shaft assemblies is supported by hydrostatically controlled bearings and connected with a hydraulically controlled actuator for lateral vibration through hydrostatic couplings.

Three rollers are fitted on each shaft, and usually they are fixed at standard gauge (1435 mm) and narrow gauge (1067 mm). When any other gauge is necessary, it can be obtained by adjusting the position of a roller. Thus, gauge-changing work can be carried out easily.

A supporting frame supports a set of rollers-and-shaft assembly, and is installed on the movable bed. The required wheelbase can be obtained by means of moving these beds longitudinally.





The frame can turn on a movable bed (by up to 10 mrad) and the relative position between the rollers and the test bogie can be finely adjusted by means of a mechanism, “balancing on a ball.” A side shaft is provided to avoid a difference in rotations between rollers, thereby achieving synchronous rotation.

## 2. Flywheel Assembly

The purpose of the flywheel is to absorb or discharge energy corresponding to the inertia of the vehicle being tested. The equivalent mass per axle of the rolling stock mass is given by a flywheel assembly for each roller.

Equivalent mass:

Max. 20 t

Min. 6 t

## 3. Driving Device

A DC motor of 500 kW is used for each set of rollers-and-shaft assembly, to give the driving device sufficient capacity to absorb the load of the rolling stock. The motor is controlled either in speed or torque mode. The driving shaft is designed to allow the rollers to move laterally.

Driving motor rating: voltage 440 V, current 1250 A, capacity 500 kW

## 4. Hydraulic Power Source

The hydraulic power source is used for power of the lateral vibration actuators, hydrostatic bearings, and hydrostatic coupling, and consists of pump unit, flow control unit, drain collection unit, oil cooling unit, etc. Lubrication of each gear set and bearing (excepting hydrostatic bearing) is by forced lubrication.

## 5. Bogie Holding Device

The bogie to be tested is positioned on the rollers by means of the load frame which is connected to the bogie holding device. The load frame is mounted with a mass equivalent to the car body and is designed to withstand the force due to acceleration and deceleration.

## 6. Building and Foundations

Besides the above-mentioned devices and mechanisms, electric power converter, controlling device, operation and monitoring board, and measurement controlling equipment are installed in the building, which is lined with soundproofing countermeasures.

The foundations, with a mass of 9000 t, is provided under the building so that vibration originating from the mechanical system under test does not transfer to the neighbouring area.

## D. THE PUEBLO ROLLER RIG (TTC)<sup>1,13</sup>

The roller rig, called the roll dynamics unit (RDU, see [Figure 14.17](#)), is a test machine for imparting rotational excitation to the wheels of a rail vehicle. The Rail Dynamics Laboratory, sponsored by the Federal Railroad Administration of the United States Department of Transportation, has been built to provide a powerful test tool for investigating a wide range of rail dynamics problems.

Through a system of drive motors, flywheels, and rollers, the RDU is capable of simulation of both nonpowered vehicles, such as boxcars and passenger cars, and for absorbing power produced



**FIGURE 14.17** Roll dynamics unit of TTC.

by self-propelled vehicles, including locomotives and transit cars. The maximum speed is 230 km/h. The unit is composed of modular elements, which can be positioned to match various bogie spacing, axle spacing, and rail equipment gauges.

Each test vehicle wheel rests on and is driven by a supporting roller. Each pair of rollers, mounted on a common shaft, is attached to a drive train, which provides inertia. This interface between the vehicle wheelsets and the roller pair simulates the vehicle travelling over track. The roller allows simulation of vehicles on tangent track having no lateral or vertical irregularities, and also allows simulation of flat curve geometry. Through its flywheels, the RDU is able to simulate resistive forces associated with accelerating or braking of a vehicle.

The RDU enables support and drive of the wheelsets of a four-axle rail vehicle or locomotive bogie. Six- or eight-axle locomotives and cars can be tested with the use of auxiliary support stands. Additionally, the roller rig is equipped with a reaction frame, providing a mounting base for two hydraulic actuators. These actuators are used to apply lateral forces to the side frame of a bogie. The forces can be either steady or vibratory, and can be either push–push or push–pull against the bogie’s side frame. This setup provides a very precise, versatile test tool when complemented with high resolution displacement instrumentation. Test conditions are monitored and controlled by the RDU control systems and a pair of computers. The test vehicle responses to the excitation are sensed, processed, and recorded by an acquisition system.

## V. OPERATION AND RESULTS

### A. TEST METHODS

The test methods for railway vehicles running on roller rigs differ between testing institutions. The following are test methods used for the Chengdu roller rig and have been formed over many years’ testing experience.

## 1. Status of the Test Vehicle

- In order to ensure the test vehicle is in good running condition, it should run on the roller rig for about 10 h or 500 km before the test starts. During the trial, the running speed should cover the design speed of the vehicle.
- If using a dummy car body instead of a real car body, the mass, moments of inertia, and centre of gravity of the dummy car body should be controlled within an error range of 15% compared with the real car body.

## 2. Status of the Roller Rig

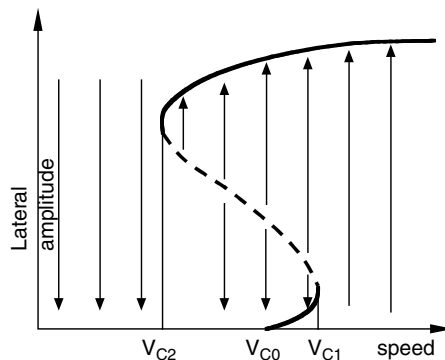
In order to ensure precision of the test, the following are checked:

- The wear of the rail profile should be less 0.2 mm.
- The roller diameter difference should be less than 0.5 mm for the same roller unit, 1 mm for the bogie, and 2 mm for the vehicle.
- The displacement error of the actuator used on the railway test facility should be less than 5%. The phase error relative to the command signal at 40 Hz should be less than 60°.
- Two longitudinal fixation bars for the middle of the car are positioned in place of the couplers with ball joints, and the length of the bar should be greater than 1 m. One longitudinal fixation bar for the locomotive is positioned in the place of the couplers with ball joints and the length of the bar should be greater than 1.5 m.
- The errors of roller altitude for the same bogie should be less than 2 mm, and for the entire vehicle should be less than 4 mm.

## 3. Stability Test

Motion stability of a railway vehicle is one of the most important factors of vehicle dynamic behaviour. The main objective of performing a stability test is to identify the vehicle hunting critical speed (see also Chapters 2 and 12). Before introducing the stability test method, we first briefly review the concept of vehicle stability.

A typical limit cycle diagram of wheelset motion is shown in Figure 14.18. In this figure, the solid line indicates a stable limit cycle and the dashed line indicates an unstable limit cycle. When the vehicle speed is lower than  $V_{C2}$ , the vehicle system is always stable under any track disturbances. When the vehicle speed is between  $V_{C2}$  and  $V_{C0}$ , the system equilibrium position is



**FIGURE 14.18** Limit cycle diagram of wheelset motion:  $V_{C0}$ , linear critical speed;  $V_{C1}$ , nonlinear critical speed;  $V_{C2}$ , nonlinear critical speed.

stable at small track disturbances and unstable at large track disturbances where a limit cycle oscillation will occur. When the vehicle speed is between  $V_{C0}$  and  $V_{C1}$ , the system equilibrium position is unstable, and small limit cycle oscillations emerge with small track disturbances and large limit cycle oscillations with large track disturbances. Finally, when the vehicle speed is greater than  $V_{C1}$ , the system jumps to large limit cycle oscillation at any track disturbance and flange contact may occur. Therefore,  $V_{C0}$  which is the Hopf bifurcation point can be defined as the linear critical speed.  $V_{C1}$  and  $V_{C2}$  can be defined as the nonlinear critical speeds. The nonlinear critical speed  $V_{C2}$ , where the first stable limit cycle appears, is normally lower than linear critical speed  $V_{C0}$ , thus it should be taken as the speed limit for the vehicle running on tracks.

The object of the stability test is to find out the three speed points  $V_{C0}$ ,  $V_{C1}$ , and  $V_{C2}$ . The method of stability testing on the roller rig (for RTU and RVTU) is as follows:

1. Gradually increasing the roller rig speed under pure rolling condition, when a small limit cycle oscillation appears, then the speed  $V_{C0}$  is found. Even when the roller rig is in a pure rolling condition, the roller rig always shows small disturbances due to roller (wheel) surface roughness or driving disturbance.
2. Increasing the roller rig speed continuously, when the hunting motion of vehicle system jumps to large amplitude oscillation or even flange contact, then the speed  $V_{C1}$  is found.
3. Reducing the speed of the roller rig slowly, when the severe hunting motion reduces to an equilibrium position, then the speed  $V_{C2}$  is found.

Since the RVTU is capable of both roller rolling and induced vibrations, it can be used to search the actual critical speed of the vehicle system using actual track irregularity inputs. The actual critical speed is between  $V_{C2}$  and  $V_{C1}$ , which is related to the track conditions.

In the field line test, the vehicle stability is estimated by the bogie acceleration, which is filtered with a band pass of approximately 0.5 to 10 Hz. If the peak values of acceleration have exceeded  $10 \text{ m/sec}^2$  six times consecutively, the vehicle is said to be unstable.

For the stability test, not only should the critical speed be measured, but also the mode shapes of hunting need to be determined. This is achieved by measuring the lateral displacements of car body, bogie frames, and wheelsets.

#### 4. Dynamic Simulation Test

Using the RVTU can simulate the running of the vehicle with track irregularity inputs. Normally, the responses of vehicles, such as accelerations or displacements on car body or bogie, are measured. According to test standards, for instance UIC518, the ride performance is calculated according to the acceleration response in the car body.

Let  $z_L$ ,  $y_L$ ,  $z_R$ ,  $y_R$  indicate the irregularity inputs of left and right rails in the time domain for the first roller unit, then the inputs of other roller units are delayed by certain time intervals. The time delays, considering the six-axle vehicle in [Figure 14.19](#) as an example, can be calculated as follows:

$$t1 = l/v \quad (14a)$$

$$t2 = 2l/v \quad (14b)$$

$$t3 = L/v \quad (14c)$$

$$t4 = (L + l)/v \quad (14d)$$

$$t5 = (L + 2l)/v \quad (14e)$$

where  $l$  is the bogie wheelbase,  $L$  is the distance between bogie centres, and  $v$  is the running speed.

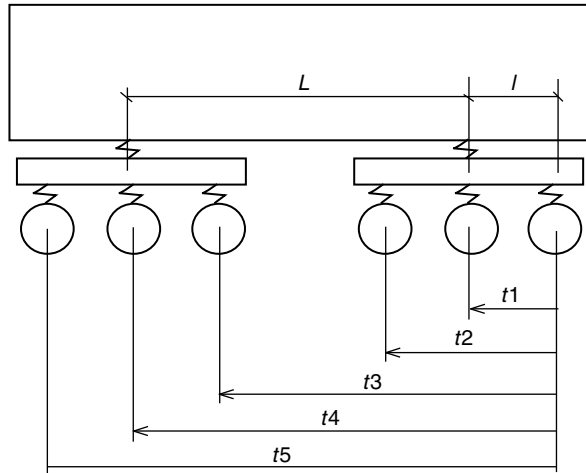


FIGURE 14.19 Time delay of input signals.

In the dynamic test, the measuring items normally include: the accelerations and displacements of car body, bogie frame, and wheelsets (axle box); the relative displacements of primary and secondary suspensions; the relative displacements of motors; the wheel/rail forces; stresses of key parts; temperature of bearing or gear case; etc.

## 5. Curve Simulation Test

In order to simulate the vehicle curving performance, the roller rig should have the ability to:

- Simulate the curve — set the rollers in a radial position
- Simulate the speeds of inner wheel and outer wheel — to attain different speeds of inner roller and outer roller by using a differential driving system
- Simulate the superelevation and centrifugal force — to set the cant angle of roller unit to simulate the unbalanced centrifugal force caused by superelevation and centrifugal force.

The steps to carrying out a curve simulation test are:

- Set the roller set in a radial position of curve.
- Widen the gauge of roller set according to the curve radius.
- Set up the cant angle of the roller unit according to the unbalanced centrifugal force.
- Lift the test vehicle onto the roller rig and locate it by fixation bars in longitudinal direction. The fixation bars are set at an angle with respect to the car body centreline according to the simulated curve radius.
- The roller rig runs at the prescribed speed.
- Adjust the speed difference of inner and outer rollers according to the curve radius and running speed.
- For steady-state simulation, the roller rig is in a rolling only condition and the wheel–rail interaction forces can be measured.
- For dynamic simulation, the track irregularity inputs are considered. The wheel–rail interaction forces and the responses such as acceleration, displacement of car body, bogie frame, and wheelset can be measured.
- Then the derailment ratio  $Q/P$ , lateral force  $H$ , wheel load reduction rate  $\Delta P/P$ , and ride index  $W$  can be obtained.

## 6. Power Test

The aim of the power test is to evaluate the driving or braking behaviour of the motor car. The power test is performed mainly at the Naples roller rig. A few examples of power tests performed in the Naples roller rig are described in Section IV.C.2.

## 7. Modal Analysis Test

When the car body, bogie frame, and wheelset are considered as rigid bodies and are suspended by primary and secondary suspensions, the vehicle is a typical multibody system. Thus the natural vibrations of the vehicle system will appear at some frequencies. By using the RVTU or VTU, the vibration modes (self-vibration frequency and modal shape) can be determined.

To perform the modal analysis test, the lateral, vertical, roll, pitch and yaw motions are normally considered. The rollers (or movable short rails) should be excited in separate modes. A swept sine wave or white noise is used as the roller rig inputs. Through analysing the responses of car body and bogie frames, the resonance points under each mode will be determined and the modal frequencies can be obtained.

## 8. Storage Security Test

Loading methods are very important for some special goods, such as columned goods (pipes, cans, wood), destructible goods (glass, apparatus) and explosive goods (nuclear material, detonators). The use of RVTU or VTU can validate the goods behaviour, such as:

- The security of loading under vibration, impact, and lateral force (superelevation, centrifugal force)
- The stability of goods after long-distance travel
- The behaviour of a new loading method
- The dynamic environment of goods during transportation, such as vibration acceleration, temperature, pressure, and force.

A roller rig has great flexibility and can perform diverse tasks such as estimation of creep forces, wear of wheel and rail profiles, adhesion between wheel and rail, natural frequency and mode, response control, etc.<sup>12–23</sup>

## B. DIFFERENCES BETWEEN ROLLER AND TRACK

### 1. Differences of Geometry Relationship

#### a. Calculation Method of Wheel–Roller Geometry Relationship

In theoretical analysis of railway vehicle dynamics, the first step is the determination of the wheel–rail geometry relationship.<sup>24,25</sup> The main difference between a vehicle running on track and on a roller rig is the wheel–rail contact and wheel–roller contact. In order to understand the behaviour of a vehicle running on a roller rig, a numerical method for wheel–rail geometry calculation called the “line tracing method” will be introduced below.<sup>26</sup>

Owing to the symmetry of the wheelset structure a half wheelset model is considered. The analysis model is shown in Figure 14.20.  $OXYZ$  is the absolute coordinate axis for the roller,  $G''wX''Y''Z''$  is the wheelset coordinates, and  $G''wX'Y'Z'$  is the wheelset fixed coordinate.  $\varphi$  and  $\theta$  are the yaw angle and roll angle with respect to  $G''wX''Y''Z''$ . Term  $Y''_{Gw}$  is the lateral displacement of wheelset. According to directional cosine principle, the axis line  $G''wY'$  in wheelset coordinate

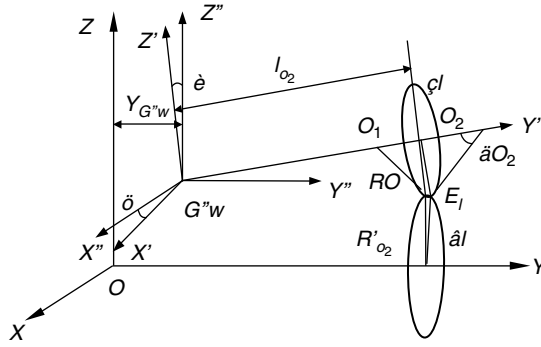


FIGURE 14.20 Model of wheel–roller geometry relationship.

$G''wX''Y''Z''$  can be described as:

$$L_x = -\cos \theta \sin \varphi; \quad L_y = \cos \theta \cos \varphi; \quad L_z = \sin \theta$$

In the figure,  $E_1$  is the contact point,  $\delta_{O_2}$  is the contact angle,  $\eta_l$  is the contact circle on wheel,  $O_2$  is the centre of the contact circle,  $R_{O_2}$  is the radius of the contact circle,  $l_{O_2}$  is the distance from the contact circle to the centre of the wheelset. The normal line from contact point  $E_1$  intersects the wheelset axis line  $G''wY'$  at point  $O_1$ . Then, the distance from  $O_1$  to  $O_2$  is:

$$O_1O_2 = H = R_{O_2} \operatorname{tg} \delta_{O_2} \tag{14.1}$$

The distance from  $O_1$  to  $E_1$  is

$$O_1E_1 = \frac{R_{O_2}}{\cos \delta_{O_2}} \tag{14.2}$$

The coordinates of  $O_1$  and  $O_2$  in the axis system  $G''wX''Y''Z''$  are

$$\begin{aligned} X_{O_2} &= l_{O_2} L_x \\ Y_{O_2} &= l_{O_2} L_y \end{aligned} \tag{14.3}$$

$$Z_{O_2} = l_{O_2} L_z$$

$$\begin{aligned} X_{O_1} &= (l_{O_2} - H) L_x \\ Y_{O_1} &= (l_{O_2} - H) L_y \end{aligned} \tag{14.4}$$

$$Z_{O_1} = (l_{O_2} - H) L_z$$

The equation of the contact circle in  $G''wX''Y''Z''$  is

$$\begin{aligned} L_x(X - X_{O_2}) + L_y(Y - Y_{O_2}) + L_z(Z - Z_{O_2}) &= 0 \\ (X - X_{O_2})^2 + (Y - Y_{O_2})^2 + (Z - Z_{O_2})^2 &= R_{O_2}^2 \end{aligned} \tag{14.5}$$

The equation of the plane, which is perpendicular to  $G''wX''Z''$ , passes  $O_1$ , and has an angle  $\beta_l$  with respect to axis  $Z''$ , is

$$Z = -ctg\beta_l X + C \quad (C \text{ is a constant}) \tag{14.6}$$

Considering the coordinates of point  $O_1$ , Equation 14.6 becomes

$$X = -tg\beta_l Z + (l_{O_2} - H)(L_z tg\beta_l + L_x) = -tg\beta_l Z + K \tag{14.7}$$

where  $K = (l_{O_2} - H)(L_z tg\beta_l + L_x)$ .

Inserting Equation 14.7 into the first formula of Equation 14.5, we get

$$Y = \frac{L_x tg\beta_l - L_z}{L_y} Z + \frac{L_x(X_{O_2} - K) + L_y Y_{O_2} + L_z Z_{O_2}}{L_y}$$

if

$$K_1 = \frac{L_x tg\beta_l - L_z}{L_y} \quad \text{and} \quad K_2 = \frac{L_x(X_{O_2} - K) + L_y Y_{O_2} + L_z Z_{O_2}}{L_y}$$

we have

$$Y = K_1 Z + K_2 \tag{14.8}$$

Inserting Equation 14.8 into the second formula of Equation 14.5, we get

$$AZ^2 + BZ + C = 0 \tag{14.9}$$

where  $A = tg\beta_l + K_1^2 + 1$

$$B = 2(-tg\beta_l K + X_{O_2} tg\beta_l - Y_{O_2} K_1 - Z_{O_2} + K_1 K_2)$$

$$C = (K - X_{O_2})^2 + (K_{O_2} - Y_{O_2})^2 + Z_{O_2}^2 - R_{O_2}^2$$

From Equation 14.9,  $Z$  can be derived as

$$Z = \frac{-B - \sqrt{B^2 - 4AC}}{2A} \tag{14.10}$$

Here, the location of contact point  $E_l$  is determined according to Equation 14.7, Equation 14.8, and Equation 14.10. However, in those equations, there is an unknown variable  $\beta_l$  (for wheel–rail case,  $\beta_l = 0$ ). It is necessary to add another equation. Generally, the wheelset centre is assumed to be in the same vertical plane of roller. Thus, there is a relation:

$$X + R'_{O_2} \sin \beta_l = 0 \tag{14.11}$$

When  $l_{O_2}$  changes, the other values such as  $R_{O_2}$ ,  $R'_{O_2}$ , and  $\delta_{O_2}$  will also be changed, and the coordinates of possible contact points  $E_l$ ,  $X_{E_l}$ ,  $Y_{E_l}$ ,  $Z_{E_l}$ , and  $\beta_l$  can be obtained. A series of possible contact points form a line called the “tracing line.” The contact point  $E_l$  must be on the tracing line. Considering the right side of the wheelset and adjusting the roll angle  $\theta$ , the contact points  $E_l$  and  $E_r$  can be determined under the condition of ensuring that the left and right wheel keep contact with the rollers. Furthermore, the radius of contact circle  $R_{l,r}$ , contact angle  $\delta_{l,r}$  and the distance of contact circle to the centre of wheelset  $l_{l,r}$  can be decided. The difference of radius of contact circle  $\Delta R$ ,



difference of contact angle  $\Delta\delta$ , equivalent conicity  $\lambda_e$ , contact angle parameter  $\varepsilon_e$ , wheelset gravitational stiffness  $K_{sy}$ , and gravitational angular stiffness  $K_{s\varphi}$  can be calculated as:

$$\Delta R = R_l - R_r; \quad \lambda_e = \frac{\Delta R}{2Y_{G''w}}$$

$$\Delta\delta = \delta_l - \delta_r; \quad \varepsilon_e = \frac{\Delta\delta}{2Y_{G''w}}(l_l - l_r)$$

$$K_{sy} = \frac{-\sin \delta_l(Z_{E_r} \sin \delta_r - Y_{E_r} \cos \beta_r \cos \delta_r) + \sin \delta_r(Z_{E_l} \sin \delta_l + Y_{E_l} \cos \beta_l \cos \delta_l)}{Y_{G''w}D}$$

$$K_{s\varphi} = \frac{(-X_{E_r} \sin \delta_l + Y_{E_l} \sin \beta_l \cos \delta_l)(Z_{E_r} \sin \delta_r - Y_{E_r} \cos \beta_r \cos \delta_r)}{\varphi D} + \frac{(X_{E_l} \sin \delta_r - Y_{E_r} \sin \beta_r \cos \delta_r)(Z_{E_l} \sin \delta_l + Y_{E_l} \cos \beta_l \cos \delta_l)}{\varphi D}$$

where

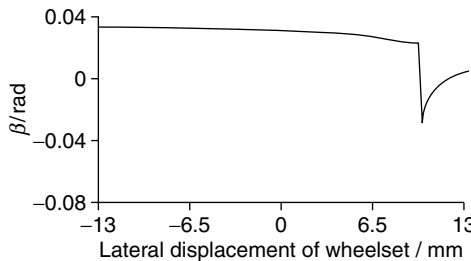
$$D = \cos \beta_l \cos \delta_l (Z_{E_r} \sin \delta_r + Y_{E_l} \cos \beta_r \cos \delta_r) + \cos \beta_r \cos \delta_r (Z_{E_l} \sin \delta_l - Y_{E_r} \cos \beta_l \cos \delta_l)$$

*b. Geometry Difference of Wheel–Rail Contact and Wheel–Roller Contact*

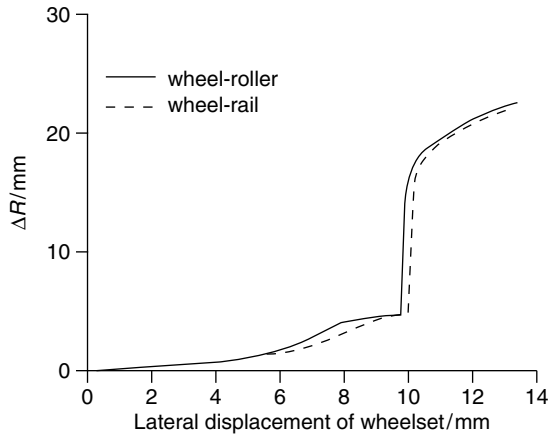
A roller rig uses a limited radius roller instead of straight track. When the yaw angle  $\varphi$  of the wheelset is zero, the geometry relationship of wheel–roller contact is the same as wheel–rail contact. If the yaw angle is not zero, a difference of geometry exists. The following parameters for wheel, rail, and roller are considered:

Track gauge is 1435 mm, wheel diameter is 915 mm, roller diameter is 1370 mm, rail cant is 1/40, type of wheel profile is LM (a Chinese worn-type wheel profile), type of rail profile (roller) is 60 kg (similar to UIC S1002 profile).

For analysing the difference between wheel–rail contact geometry and wheel–roller contact geometry, yaw angle  $\varphi$  is chosen as  $3^\circ$ . First, for wheel–rail contact, the normal line of contact point is always in the vertical plane, which means the angle  $\beta$  is zero. However, for wheel–roller contact, there is an angle between the normal line and vertical plane. Figure 14.21 shows the angle of left wheel contact point  $\beta_l$  at different lateral displacements of the wheelset. When the yaw angle is  $3^\circ$ , the angle  $\beta$  is about 0.038 rad ( $\approx 2.2^\circ$ ). When the lateral displacement of the wheelset reaches 10 mm, a change of the angle  $\beta_l$  occurs and the value of  $\beta_l$  becomes negative, resulting in flange contact.



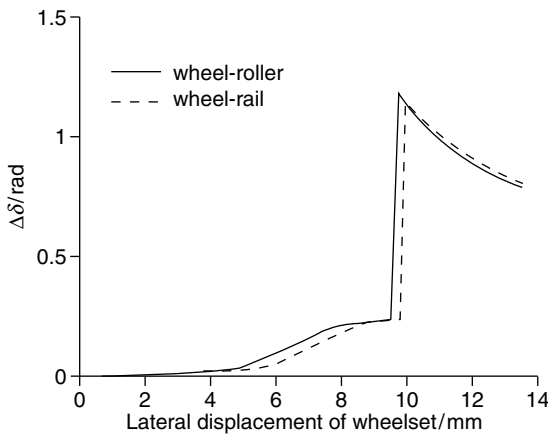
**FIGURE 14.21** Angle  $\beta_l$ .



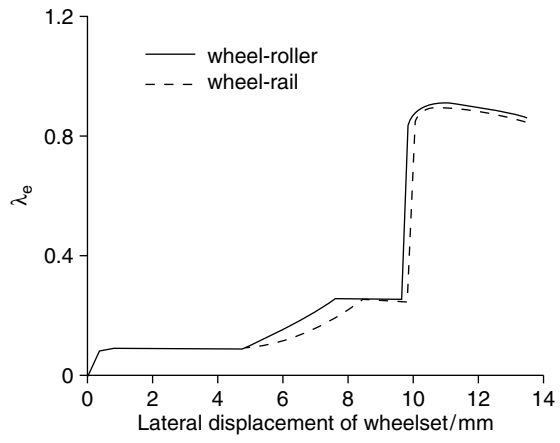
**FIGURE 14.22** Comparison of  $\Delta R$ .

Figure 14.22 illustrates the comparison of the difference in radius of the contact circle  $\Delta R$ . When the lateral displacement is less than 5 mm, there is almost no difference in  $\Delta R$ . Once the lateral displacement is larger than 5 mm, the difference in radius of contact circle  $\Delta R$  for wheel–roller contact is larger than wheel–rail contact. At approximately 10 mm flange contact occurs and the value of  $\Delta R$  increases rapidly. The flange contact for wheel–roller contact appears earlier than wheel–rail contact. The situation for the difference in contact angle  $\Delta\delta$  is similar to the case of  $\Delta R$ . After flange contact, the difference in contact angle  $\Delta\delta$  of wheel–roller contact is smaller than wheel–rail contact, which is shown in Figure 14.23. Since the equivalent conicity  $\lambda_c$  is derived from  $\Delta R$ , results are similar, as shown in Figure 14.24. Figure 14.25 shows the comparison of roll angle  $\theta$ . It can be seen that the roll angle  $\theta$  in wheel–roller is larger than in the case of wheel–rail.

Comparing the wheelset gravitational stiffness  $K_{sy}$  and gravitational angular stiffness  $K_{s\varphi}$ , it is found that the stiffness  $K_{sy}$  is nearly the same for wheel–rail contact and wheel–roller contact, but there is a big difference for angular stiffness  $K_{s\varphi}$ . At  $3^\circ$  of wheelset yaw angle,  $K_{s\varphi}$  for wheel–rail contact is nearly zero before the flange contact, but for wheel–roller contact,  $K_{s\varphi}$  is negative before flange contact (see Figure 14.26). Therefore, the vehicle running on the roller rig is less stable due to the negative value of  $K_{s\varphi}$ .



**FIGURE 14.23** Comparison of  $\Delta\delta$ .



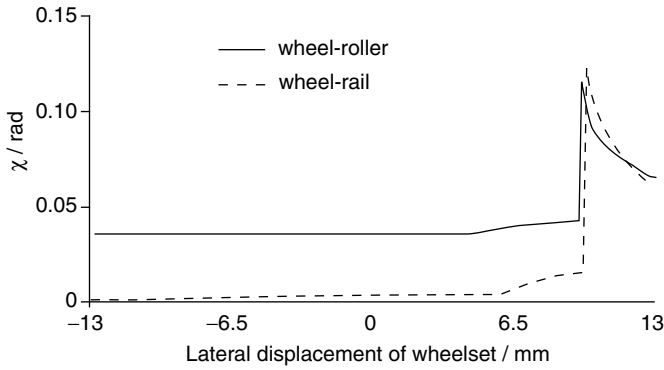


FIGURE 14.27 Comparison of  $\chi$ .

The contact point is not at the bottom of the contact cycle, and there is a leading angle  $\chi$  shown in Figure 14.27. It is apparent that the leading angle  $\chi$  for wheel–roller contact is larger than the value in the case of wheel–rail contact. This is the reason why the absolute value of gravitational angular stiffness  $K_{s\varphi}$  for wheel–roller contact is larger than for wheel–rail contact.

## 2. Difference in Creep Coefficient

The creep forces are calculated by the following formula

$$T_1 = -f_{11}\gamma_1, \quad T_2 = -f_{22}\gamma_2 - f_{23}\omega_3, \quad M_3 = f_{23}\gamma_2 - f_{33}\omega_3 \quad (14.12)$$

where  $T_1$ ,  $T_2$ , and  $M_3$  are longitudinal creep force, lateral creep force, and spin creep moment.  $\gamma_1$ ,  $\gamma_2$ , and  $\omega_3$  are the longitudinal, lateral, and spin creepages, which are related to the movement of wheel and rail (roller).  $f_{11}$ ,  $f_{22}$ ,  $f_{23}$ , and  $f_{33}$  are the creep coefficients, which are defined as:

$$f_{11} = E(ab)c_{11}; \quad f_{22} = E(ab)c_{22}; \quad f_{23} = -f_{32} = E(ab)^{3/2}c_{23}; \quad f_{33} = E(ab)^2c_{22} \quad (14.13)$$

where  $E$  is the modulus of rigidity.  $a$  and  $b$  are the length of the semiaxis of the contact ellipse in the rolling and lateral directions.  $C_{ij}$  are the Kalker’s creepage and spin coefficients, which depend only on Poisson’s ratio  $\sigma$  and the ratio of the semiaxes of the contact ellipse. Thus, the creep coefficients are significantly affected by the contact ellipse dimensions (see also Chapter 4, Section III E).

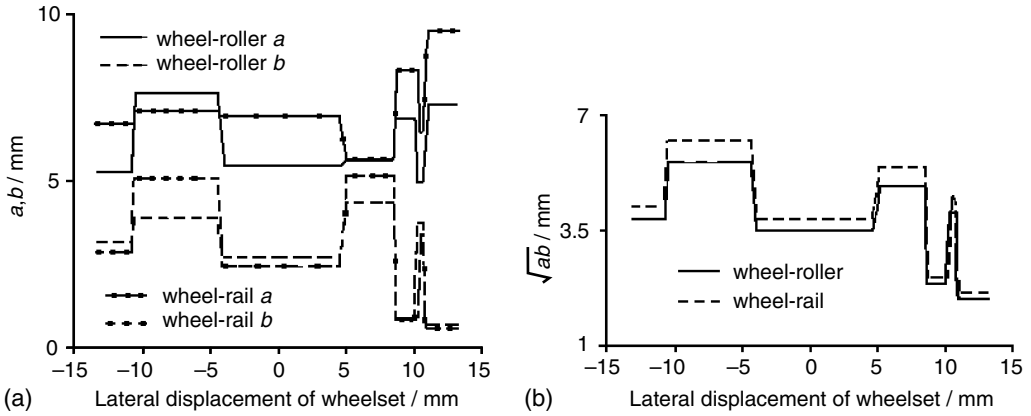
According to Hertz’s static solution, the semiaxis of the contact ellipse  $a$  and  $b$  can be expressed as:

$$\left(\frac{a}{m}\right)^3 = \left(\frac{b}{n}\right)^3 = \frac{3N(1 - \sigma)^2\rho}{2\pi E} \quad (14.14)$$

where  $N$  is the normal force.  $m$  and  $n$  are coefficients which can be found by using Hertz’s table.  $\rho$  is the characteristic length which can be described as:

$$\frac{1}{\rho} = \frac{1}{R_1} + \frac{1}{R'_1} + \frac{1}{R_2} + \frac{1}{R'_2} \quad (14.15)$$

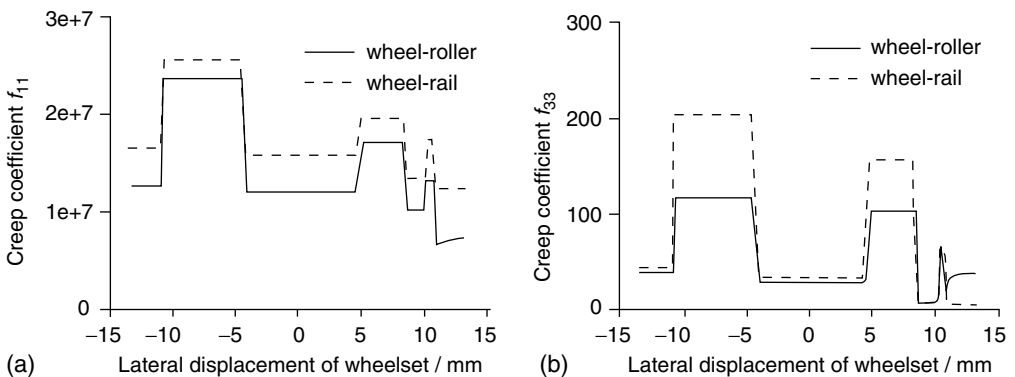
where  $R_1$  and  $R'_1$  are the principal radii of wheel profile.  $R_2$  and  $R'_2$  are the principal radii of rail or roller profile. For wheel–rail contact,  $R_2 = \infty$ . Thus it is obvious that characteristic length  $\rho$  for wheel–rail contact is larger than the case for wheel–roller contact. For example, principal radii of



**FIGURE 14.28** Comparison of contact ellipses: (a) semiaxes of the contact ellipse; (b) area of the contact ellipse.

wheel profile  $R_1 = 475.5$  mm and  $R'_1 = 500$  mm, the principal radii of roller profile  $R_2 = 685$  mm ( $\infty$  for rail) and  $R'_2 = 300$  mm, the characteristic length  $\rho$  for wheel–rail contact is 134.0 mm, and for the wheel–roller contact is 112.4 mm; the difference is about 20%. Under the condition of an axle load of 11,400 kg, the semiaxes of the contact ellipse  $a$  and  $b$  for wheel–rail and wheel–roller contact is shown in Figure 14.28a. Since the profiles of wheel and rail are composed of different curves and lines, the calculated  $a$  and  $b$  are not constants.

When the wheelset lateral displacement is zero, the semiaxis  $a$  for wheel–roller contact is shorter than the value for wheel–rail contact. The reason is that the radius of the roller is less than the radius of the rail, while the semiaxis  $b$  for wheel–roller contact is little larger than the value for wheel–rail contact. This situation is easily explained, since the axle load for the two cases is the same, when the semiaxis  $a$  becomes smaller, the semiaxis  $b$  will become larger in order to support the same wheel load. When the wheelset is at its central position, the area of contact ellipse  $\sqrt{ab}$  for wheel–roller contact is smaller than for wheel–rail contact (see Figure 14.28b). According to Equation 14.13, the area of the contact ellipse directly affects the creep coefficients, so it can be estimated that the creep coefficients for wheel–roller contact should be smaller than the values for wheel–rail contact. Figure 14.29 shows the comparison of the creep coefficients  $f_{11}$  and  $f_{33}$ .



**FIGURE 14.29** Comparison of creep coefficients: (a) creep coefficient  $f_{11}$ ; (b) creep coefficient  $f_{33}$ .

The value of  $f_{11}$  for wheel–roller contact is 20% less than for wheel–rail contact. For  $f_{33}$ , the maximum difference is about 40%.

### 3. Differences in Stability

There are two main reasons for the difference between wheel–rail contact and wheel–roller contact. The first is that the creep coefficients for wheel–roller contact are smaller than the values for wheel–rail contact, and the second is that the wheelset gravitational angular stiffness  $K_{s\phi}$  is negative for wheel–roller contact when the yaw angle  $\phi$  is not zero. In order to discuss the difference of vehicle stability on a roller rig, the equations of motion of a single wheelset with primary suspension are considered,<sup>27,28</sup>

$$\begin{aligned}
 M_w \ddot{y}_w + \frac{2f_{22}}{v} \dot{y}_w + \left( K_{py} - \frac{2f_{23}\varepsilon}{R_0 b} + \frac{W\varepsilon}{b} \right) y_w + \frac{2f_{23}}{v} \dot{\phi}_w - 2f_{22} \phi_w &= 0 \\
 I_{wz} \ddot{\phi}_w - \frac{2f_{23}}{v} \dot{y}_w + \left( \frac{2f_{11}\lambda_e b}{R_0} - 2f_{33} \frac{\varepsilon}{R_0 b} \right) y_w + \left( \frac{2f_{11}b^2}{v} + 2f_{33} \right) \dot{\phi}_w \\
 + (K_{px}b_1^2 + K_{s\phi} + 2f_{23})\phi_w &= 0
 \end{aligned}
 \tag{14.16}$$

where  $y_w$  and  $\phi_w$  are the lateral displacement and yaw angle of wheelset.  $M_w$  and  $I_{wz}$  are the mass and moment of inertia around axis  $z$ .  $K_{px}$  and  $K_{py}$  are the suspension stiffness in the rolling direction and the lateral direction.  $R_0$  is the wheel contact radius.  $b$  is the half distance between contact points.  $b_1$  is the half distance between primary suspension.  $W$  is the axle load.  $v$  is the forward speed.

Through variable transformation, Equation 14.16 can be presented as

$$\dot{Y} = AY \tag{14.17}$$

The condition of nonzero solution is that the following form is satisfied

$$|\lambda I - A| = 0 \tag{14.18}$$

Expanding Equation 14.18, it can be written as

$$a_0\lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4 = 0 \tag{14.19}$$

According to the Hurwitz law of stability, the following condition should be satisfied if the wheelset is stable

$$\begin{aligned}
 \Delta_1 &= a_1 > 0 \\
 \Delta_2 &= \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0 \\
 \Delta_3 &= \begin{vmatrix} a_1 & a_3 & a_0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} > 0 \\
 \Delta_4 &= a_4\Delta_3 > 0
 \end{aligned}$$

Since Equation 14.16 can ensure that the coefficient  $a_i$  in Equation 14.19 is positive, then  $\Delta_1$  and  $\Delta_2$  are larger than zero, so only  $\Delta_3$  needs to be considered. When  $\Delta_3 = 0$ , the critical

speed  $V_c$  can be presented as:

$$V_c^2 = \frac{\left[ 2f_{11}b^2 \left( K_{py} + \frac{W\varepsilon}{b} \right) + 2f_{22}(K_{px}b^2 + K_{s\varphi}) \right] (2f_{22}I_{wz} + 2f_{11}b^2M_w)}{\frac{\lambda_e}{R_0b} (2f_{22}I_{wz} + 2f_{11}b^2M_w)^2 - \left[ M_w(K_{px}b_1^2 + K_{s\varphi}) - I_{wz} \left( K_{py} + \frac{W\varepsilon}{b} \right) \right]^2} \quad (14.20)$$

As discussed, considering a wheelset on a roller rig, the gravitational angular stiffness  $K_{s\varphi}$  has a negative value and this will result in the numerator decreasing and the denominator increasing. Thus, the critical speed  $V_c$  reduces. This result has been described in other papers.<sup>29–33</sup>

Figure 14.30 shows a Chinese high-speed passenger car with TB wheel profile (a conical profile). The hunting stability of the vehicle on roller rig and on rail are compared. Figure 14.31 is the limit cycle of the vehicle on a roller rig and on rail. It can be seen that the critical speed of the vehicle on the roller rig is less than the critical speed on rail. The speed difference is approximately 40 to 60 km/h. As would be expected, the speed difference between a vehicle on a roller rig and on rail is not the same for different vehicles or different parameters. Figure 14.32 shows the influence of the longitudinal stiffness primary suspension (it is 24 MN/m in Figure 14.31). It is seen that the speed difference between a vehicle on a roller rig and on rail becomes larger with an increase in the longitudinal stiffness. Therefore, it is necessary to modify the critical speed of the vehicle being tested on the roller rig by theoretical analysis or by a test method. Using a test method to modify the critical speed has been adopted on the Munich roller rig<sup>6</sup> by changing the states of the rollers, such as widening the gauge, or tilting the rollers, according to the requirement of equivalent conicity.



FIGURE 14.30 Passenger car on the roller rig.

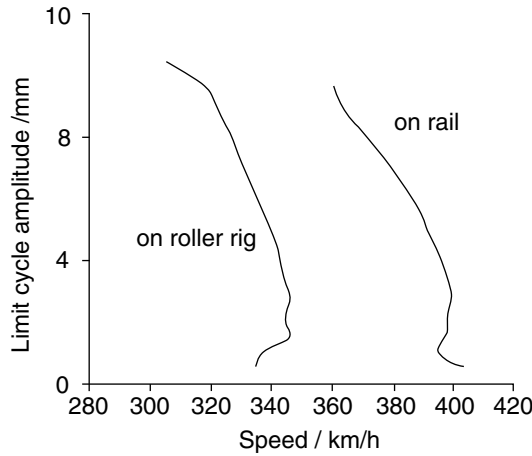


FIGURE 14.31 Comparison of critical speed.

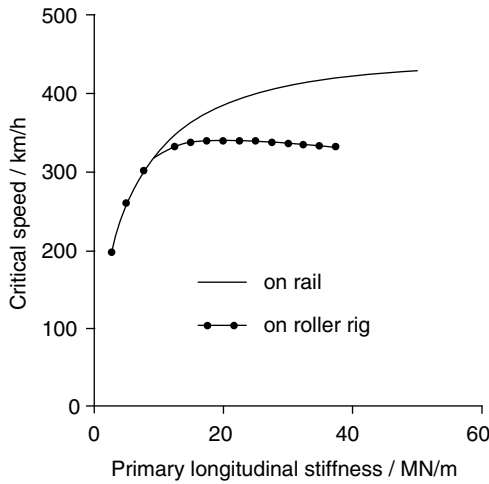


FIGURE 14.32 Influence of stiffness.

**4. Difference in Vibration Response**

Dynamic tests include not only stability tests but also the dynamic response test, which is performed to investigate the vibration response of a car body subject to track irregularity inputs. Rig results can be compared with line tests, and the difference in vibration response for the vehicle tested on the RVTU should be clear.

Inclusion of track irregularity inputs in the equations of motion of the wheelset gives:

$$M_w \ddot{y}_w + \frac{2f_{22}}{v} \dot{y}_w + \left( K_{py} - \frac{2f_{23}\epsilon}{R_0 b} + \frac{W\epsilon}{b} \right) y_w + \frac{2f_{23}}{v} \phi_w - 2f_{22} \phi_w = f_{22} \left( \frac{\dot{y}_{tl}}{v} + \frac{\dot{y}_{tr}}{v} \right) \quad (14.21)$$



$$\begin{aligned}
 I_{wz}\ddot{\phi}_w - \frac{2f_{23}}{v}\dot{y}_w + \left( \frac{2f_{11}\lambda_e b}{R_0} - 2f_{33}\frac{\varepsilon}{R_0 b} \right) y_w + \left( \frac{2f_{11}b^2}{v} + 2f_{33} \right) \dot{\phi}_w + (K_{px}b_1^2 + K_{s\phi} + 2f_{23})\phi_w \\
 = \frac{f_{33}}{v}(\dot{\phi}_{tl} + \dot{\phi}_{tr})
 \end{aligned}$$

where  $\dot{y}_{tl}$  and  $\dot{y}_{tr}$  are the lateral velocities of the left and right tracks.  $\dot{\phi}_{tl}$  and  $\dot{\phi}_{tr}$  are the yaw angular velocities of the left and right rails. Owing to the reduction in creep coefficients  $f_{ii}$  and the gravitational angular stiffness  $K_{s\phi}$ , the damping and stiffness values for the variable  $\phi_w$  are reduced, and therefore the vibration response of a vehicle on a roller rig will be larger than the case of running on rail.

By taking a Chinese high-speed passenger car as an example, shown in Figure 14.30, the comparison for the responses on roller rig and on rail is made. Figure 14.33 shows the acceleration response in the frequency domain. When the frequency is larger than 6 Hz, the acceleration response is nearly the same between roller rig and rail but at a low frequency the acceleration response of vehicle on roller is larger than the case on rail.

The example has been also analysed by using the code MEDYNA.<sup>4</sup> The following power spectral density of track irregularity is used:

$$S_{yy}(\omega) = S_{zz}(\omega) = \frac{0.1078 \times 105}{1 + 0.6804 \times \omega^2 + 0.2886 \times 10^{-3} \times \omega^4} \tag{14.22}$$

The sampling point is at floor height above the bogie, with 1 m offset from the centre. The ISO2631 standard is used to evaluate the difference in acceleration response. Comparison of the results is shown in Table 14.4. Data above the line in the row is for lateral and data below is for

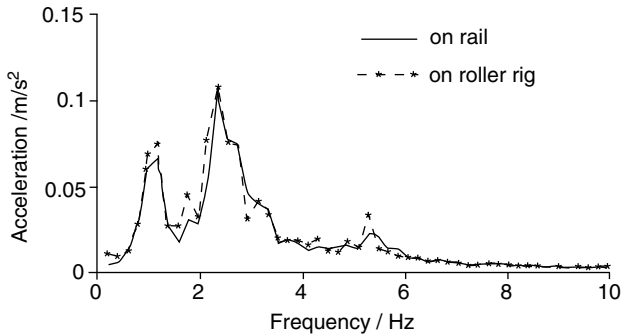


FIGURE 14.33 Comparison of acceleration response.

TABLE 14.4 Comparison of Acceleration Response

Speed/km/h	120	160	220	320	400
On rail	0.3045	0.4162	0.4769	0.5615	0.7027
	0.1504	0.2586	0.3646	0.5317	0.6221
On roller rig	0.3053	0.4230	0.4910	0.5970	0.6896
	0.1540	0.2650	0.3835	0.6160	0.7788

vertical. It is obvious that the results for a vehicle on a roller rig are larger than the results on rail, whether for lateral or vertical.

### 5. Difference in Curve Simulation

Curve simulation is difficult to imitate on a roller rig, especially for a vehicle passing through a transition curve. Owing to this, curve simulations are limited to circular curves. The cant of the roller unit is used to simulate the unbalanced centrifugal force caused by superelevation and centrifugal force. The cant angle of the roller unit can be expressed as:

$$\theta_d = \frac{v^2}{gR} - \frac{h}{2d} \tag{14.23}$$

where  $v$  is speed,  $R$  is the radius of curvature,  $h$  is the superelevation, and  $2d$  is the gauge.

In order to show the differences in curve simulation, an example with a curve radius of  $R = 300$  m and superelevation  $h = 90$  mm is presented in Figure 14.34. A passenger vehicle modeled using the SIMPACK vehicle dynamics code was used in this example. The results shown in Figure 14.34 indicate that the lateral displacement of wheelset, lateral wheel-rail force  $Q$ , derailment factor  $Q/P$ , wheel load reduction ratio  $\Delta P/P$  on a roller rig are a little smaller than the values on rail. The difference is less than 10%, showing that curve simulation on a roller rig can basically show the characteristics of vehicles passing through a circular curve.

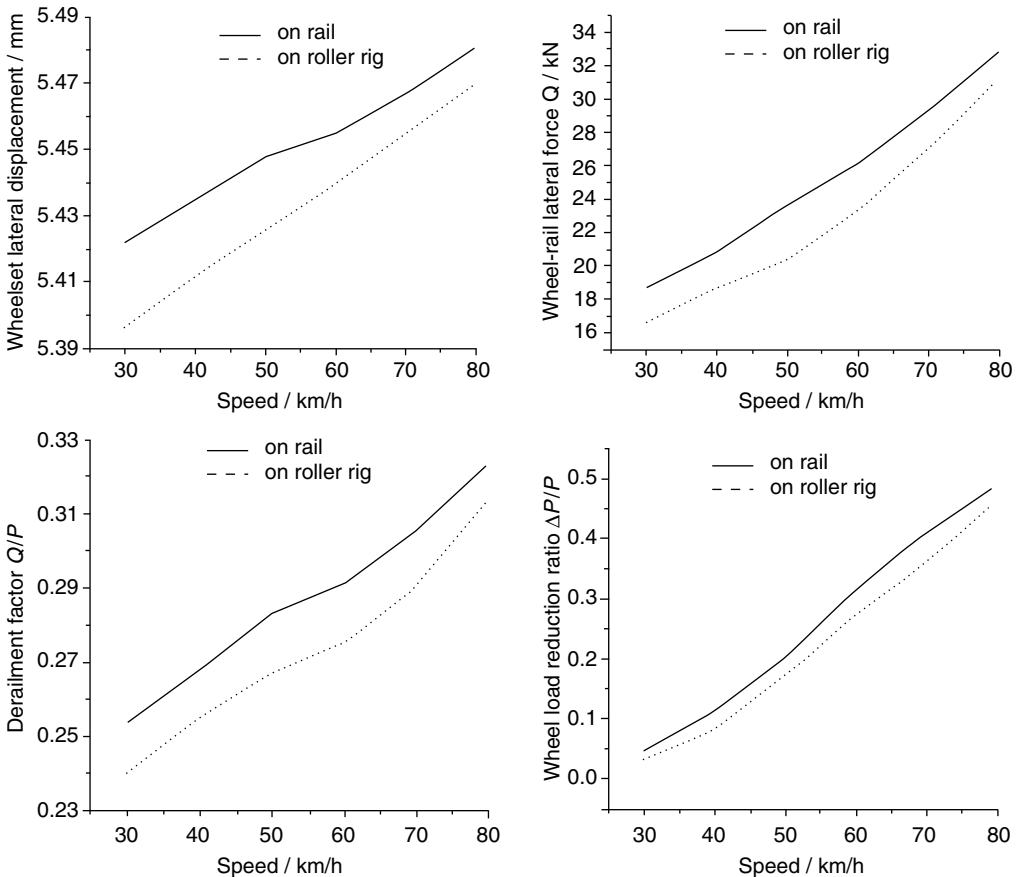


FIGURE 14.34 Comparison of curve simulation.

### C. INFLUENCE OF SETUP ERRORS ON ROLLER RIG VEHICLE STABILITY

The roller rig should have the ability to evaluate accurately the dynamic behaviour of the vehicle system but various setup errors always exist on a roller rig due to wear, incorrect installation, and different environmental (friction) conditions. The following shows the influence of various setup errors of a roller rig on vehicle stability using the example of a Chinese high-speed passenger car (shown in Figure 14.30) with a TB wheel profile.<sup>4,34</sup>

#### 1. Diameter of the Rollers

Wear and machining errors on the roller surface are practically unavoidable. The nominal diameter of the rollers in Chengdu is 1370 mm, with an allowable minimum diameter of 1300 mm. A smaller roller radius will cause larger negative gravitational angular stiffness  $K_{s\phi}$ , which may cause hunting critical speed to become lower. Figure 14.35 shows the limit cycle of the passenger car on the roller with diameters of 1370 and 1300 mm. It is seen that the smaller the roller diameter, the lower the critical speed.

#### 2. Gauge of the Rollers

For an RTU, the gauge of the rollers is fixed during the test but after adjusting the rollers to meet a different gauge, an error in the gauge may occur. For an RVTU, the gauge can be changed arbitrarily. Theoretical results indicate that the influence of gauge is small when using a TB wheel profile but when using the LM worn-type wheel profile the influence of gauge on critical speed is greater, as seen in the results of Figure 14.36. When the gauge is widened, the critical speed increases, which is one of the methodologies used on the Munich roller rig to match test results to line measurements.

#### 3. Cant of the Rollers

In the early era of railways the cant of the rail was zero, later it became 1/20 and currently the normal cant is 1/40. The influence of different cant is shown in Figure 14.37. It is evident that when the roller is without cant, the critical speed is at its lowest, with increase in cant, the critical speed increases, but within the range of 1/40 to 1/10 the influence is relatively small.

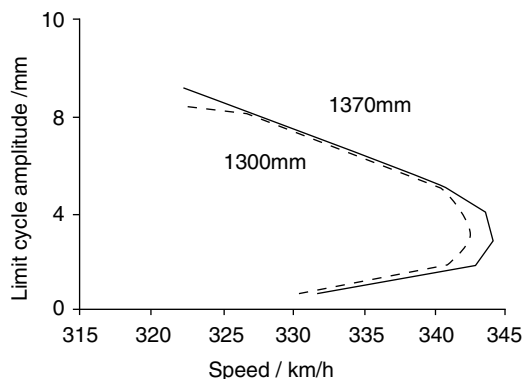
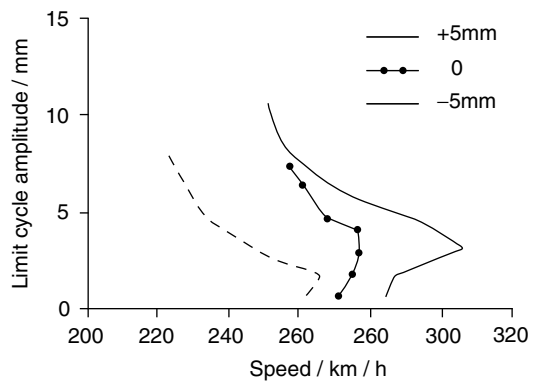


FIGURE 14.35 Influence of diameter of roller.



### 4. Coefficient of Friction on the Contact Surface

A wheel rolling on rollers makes the contact surface smooth and contamination makes the friction characteristic change. The influence of friction coefficient on stability is shown in Figure 14.38. It is seen that the lower the friction coefficient, the higher the critical speed.

### 5. Vehicle Position on the Roller Rig

During the test the vehicle sits on top of the rollers and is fixed in the longitudinal direction. Since errors exist in the wheelbase, bogie centre distance, and position of roller unit, it is difficult to position accurately the wheels on top of the rollers (see Figure 14.39). Such errors will influence the critical speed, as shown in Figure 14.40. When all the wheels of the vehicle are forward of the centreline of the rollers by 10 mm, the critical speed reduces, contrarily when all wheels are rearward by 10 mm, the critical speed is increased. Such interesting results are validated by the stability tests.

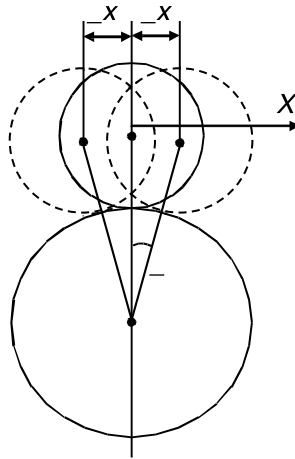


FIGURE 14.39 Wheel position.

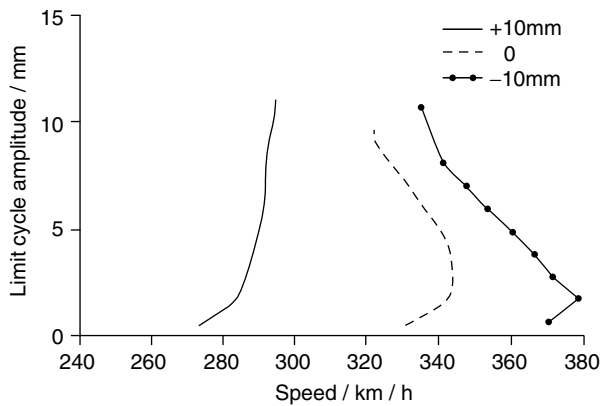


FIGURE 14.40 Influence of friction coefficient.

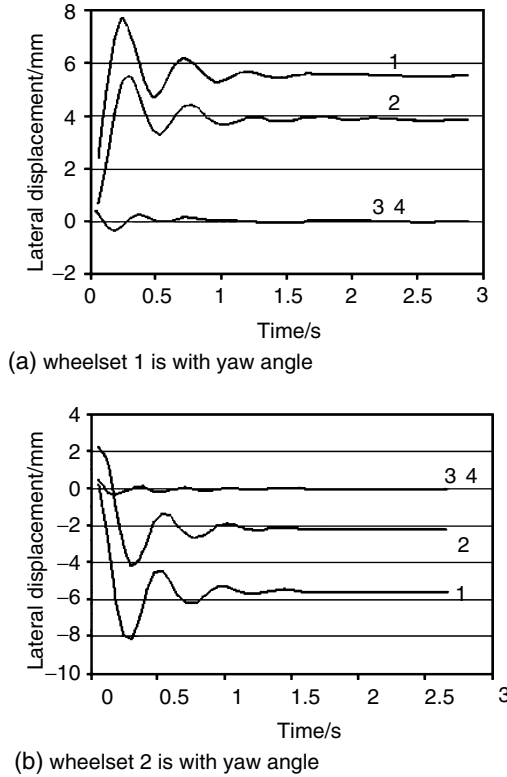


FIGURE 14.41 Influence of wheelset yaw angle.

### 6. Yaw Angle between Wheelset and Roller Axle

The wheels of the test vehicle should stay on top of the rollers and be parallel with the roller axle. However, the wheelsets within a bogie are difficult to keep parallel with each other. There is a yaw angle between the wheelset and roller axes which results in the wheelset moving to one side of roller unit or even into flange contact. Figure 14.41 shows the simulation results of vehicle motion due to a wheelset yaw angle. It is known that when wheelset one is set a yaw angle, it moves from the central position, which in turn causes wheelset two to move from the central position. When wheelset two is set to a yaw angle, again, both wheelsets one and two leave the central position. Thus if one wheelset has a yaw angle with respect to roller axle, it will cause the bogie to go to one side. Normally, this yaw misalignment of the wheelsets cannot be adjusted during the test. In order to allow the wheelset to align centrally on the roller unit, it is necessary to adjust the yaw angle of the roller axle to match the wheelset yaw angle.

It should be stated that the above results are based on a Chinese high-speed passenger car. The influence of the setup error of the roller rig on stability may be inconsistent with numerical values for other vehicles. However the general trend should be the same.

### D. EXAMPLES OF COMPARISON RESULTS WITH THEORETICAL ANALYSIS

The following are examples of vehicle dynamic tests carried out on the Chengdu roller rig. The test results are compared with theoretical analyses.



FIGURE 14.42 Freight car on the roller rig.

## 1. Stability Test

### a. Stability Tests and Theoretical Analysis of a Freight Car<sup>4</sup>

The first example of a stability test is a heavy haul freight car, designed and manufactured by Qiqihaer Rolling Stock Works. The main purpose of the test was to evaluate the hunting stability and ride performance of the car. The test was carried out in 1995. The freight car on the roller rig is shown in Figure 14.42.

The heavy haul freight car is designed for an axle load of up to 25 t and a maximum speed of 120 km/h, using a three-piece bogie. In order to improve the stability of three-piece bogies, cross-bracing between two side-frames, elastic side bearers, and axle box rubber springs are incorporated in the design. The car body is available in two types, the normal body has a bogie separation of 9.2 m and the other a bogie separation of 7.8 m.

The roller rig speed acceleration or deceleration rate is chosen as 1.25 km/h per second. It has been found through tests that the critical speed for the car with a long car body is approximately 90 to 98 km/h and is approximately 78 to 92 km/h for the car with a short body. The distribution of test results is due to nonlinearities such as friction, clearance, and bump stops. The hunting motion of the wheelset is shown in Figure 14.43. The hunting motions for the front and rear bogie are in opposite directions. When the hunting motion occurs, the flange contact appears, which means the critical speed is  $V_{C1}$  (see Figure 14.18). Then, gradually lowering the roller rig speed, the freight car becomes stable at approximately 40 to 50 km/h (see Figure 14.43b), which means the critical speed  $V_{C2}$  is very low and cannot meet the design speed. After investigating the suspension parameter, the reasons for low stability are:

- There is longitudinal clearance in the side bearer (of about 1 mm). The side bearer should work longitudinally as a friction damper, but due to the clearance it acts as a spring and cannot provide the intended yaw friction torque.
- The bogie warp stiffness is only 1 MN/rad (design value is 6 MN/rad).
- The longitudinal stiffness of primary suspension is strongly nonlinear due to the structure of the suspension.
- The stiffness value reduces with increasing displacement, which is shown in Figure 14.44.

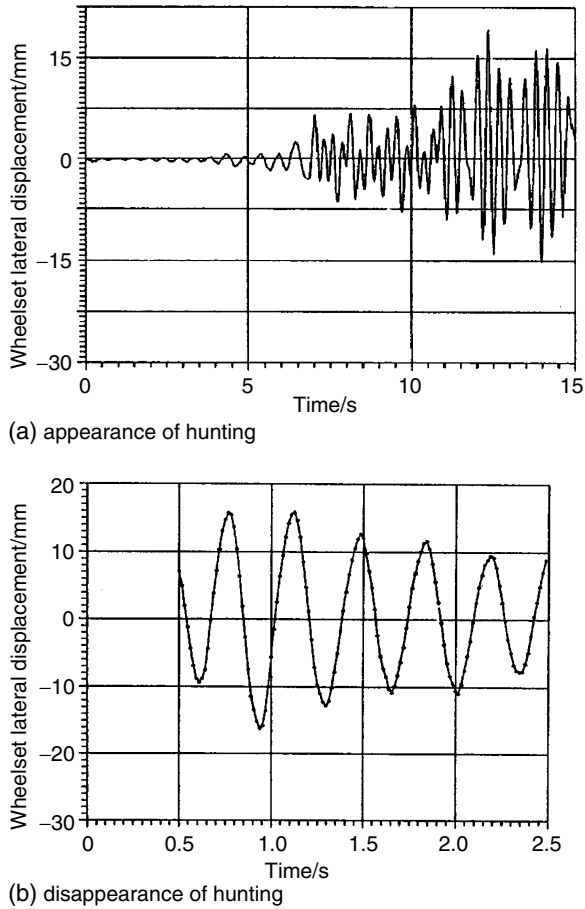


FIGURE 14.43 Record signal of hunting.

The stability of the freight car has been analysed using a simulation method with measured parameters. The degrees of freedom of the freight car are listed in Table 14.5. A total of 44 degrees of freedom are taken into consideration. A nonlinear wheel–roller geometry relationship is used and the object of the simulation is to recreate the stability test process. The start speed is

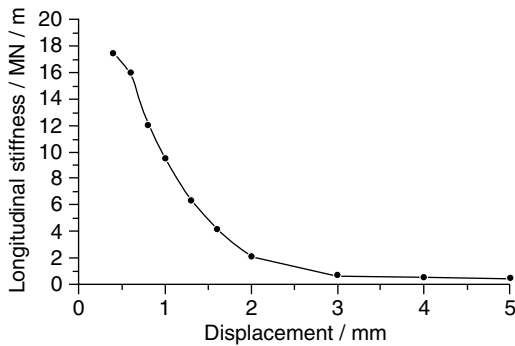


FIGURE 14.44 Longitudinal stiffness.



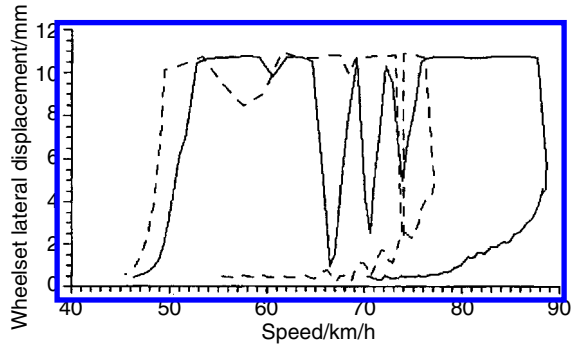


FIGURE 14.45 Wheelset lateral displacement amplitude.

TABLE 14.5  
Degrees of Freedom of Freight Car

Coordinate Rigid Body	X Longitudinal	Y Lateral	Z Vertical	$\theta$ Roll	$\varphi$ Pitch	$\psi$ Yaw
Wheelset ( $i = 1 \sim 4$ )	$x_w$	$y_w$			$\varphi_w$	$\psi_w$
Side frame ( $i = 1,2$ ) ( $r, l$ )	$x_f$	$y_f$	$z_f$		$\varphi_f$	$\psi_f$
Bolster <sup>a</sup> ( $i = 1,2$ )						$\psi_H$
Car body	$x_c$	$y_c$	$z_c$	$\theta_c$	$\varphi_c$	$\psi_c$

<sup>a</sup> Other motions of bolster are considered in car body motions.

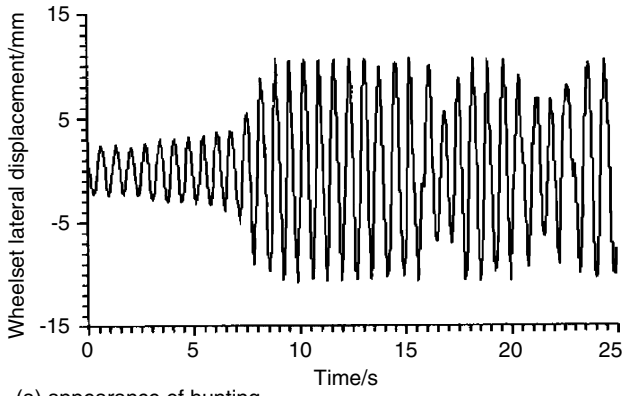
70 km/h for the long car-body car and 55 km/h for short car-body car. Figure 14.45 shows the simulation results for the wheelset lateral motion, the solid line is for the long car and the dashed for the short car. It can be seen that the hunting motion at small amplitudes occur at 78 to 88 km/h for the long car and at 70 to 77 km/h for the short car. When the speed is greater than 88 km/h for the long car and 77 km/h for the short car, the hunting motion increases and flange contact occurs. Then, gradually lowering the speed, the system becomes stable at speeds of 53 and 49 km/h for the long car and short car, respectively. Figure 14.46 depicts the hunting motion of wheelset in amplitude and phase. It can be seen that the calculated results correlate well with the test results.

Field testing of the vehicles have shown that the long car becomes unstable at approximately 100 km/h, and after increasing the warp stiffness and improving the side bearers, the freight car can run in service at 120 km/h.

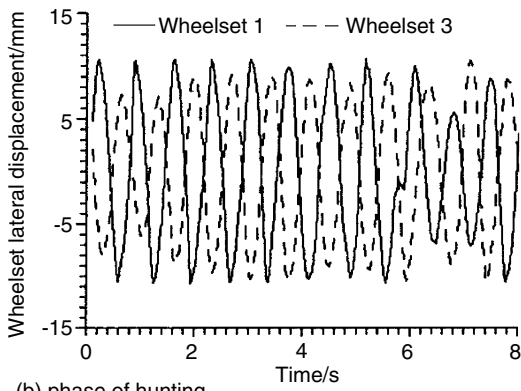
*b. Stability Test and Theoretical Analysis of a Passenger Car<sup>31</sup>*

A stability test on a passenger car was carried out on the Chengdu roller rig in 1998. The passenger car was manufactured by Hanjin Industry of Korea and was in service on the Heng-guang line in China. The passenger car uses a stainless steel car body made in Korea and the bogies are made in China. Its design speed is 200 km/h. The vehicle is shown on the roller rig in Figure 14.47.

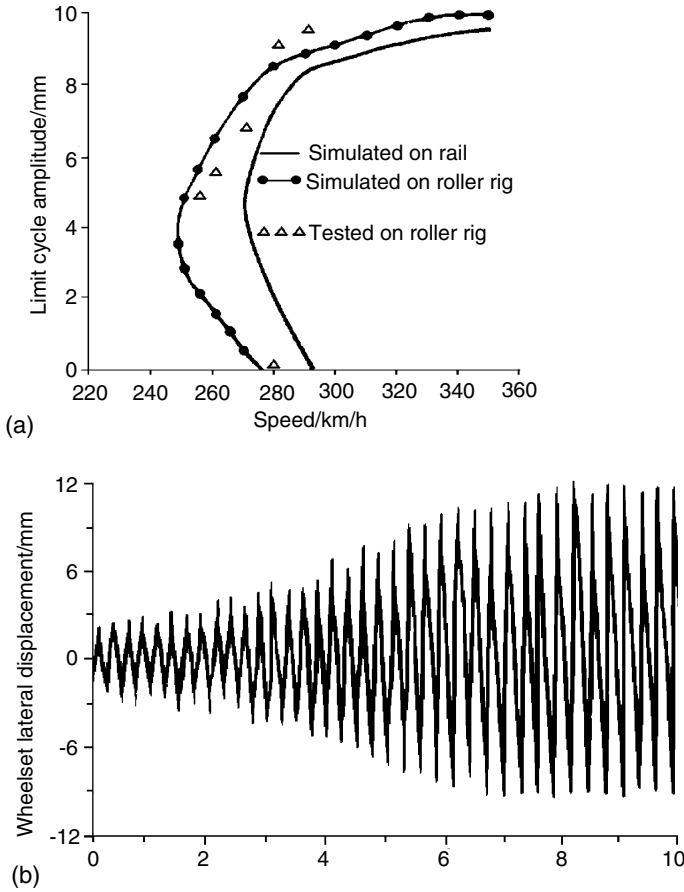
Before the test was undertaken, the stability of the passenger car on rail and on the roller rig was theoretically analysed. The results are shown in Figure 14.48a; the solid line shows the wheelset limit cycle amplitude on rail, and the solid line with a black circle are the results running on the roller rig. It can be seen that the critical speed on the roller rig is lower than on rail, the difference



(a) appearance of hunting



(b) phase of hunting



**FIGURE 14.48** Wheelset hunting motion: (a) limit cycle amplitude; (b) appearance of hunting motion.

being about 20 km/h. The triangle symbols shown in Figure 14.48a are the test results measured on the roller rig. The measured linear and nonlinear critical speeds are approximately 280 and 255 km/h which correlate quite well with the theoretical values. Figure 14.48b shows that the wheelset hunting motion is very sinusoidal in form. The test results demonstrated the tested passenger car can run at a design speed of 200 km/h.

**2. Ride Comfort Test and Theoretical Analysis of a Passenger Car<sup>35,36</sup>**

A dynamic performance test on a wide-gauge (1676 mm) coil spring passenger bogie manufactured by the Changchun Car Company was carried out on the Chengdu roller rig in 2002. The test vehicle on the rig is shown in Figure 14.49. The main purpose of the test was to evaluate the hunting stability and ride comfort of the passenger car for use in Pakistan.

The dynamic performance of the vehicle system is also simulated. For simplicity in modelling the complex system, the following assumptions were made:

- The wheelset, bogie frame, and car body are considered as rigid bodies.
- The adjacent vehicles are not considered and only one vehicle is used in the simulation.

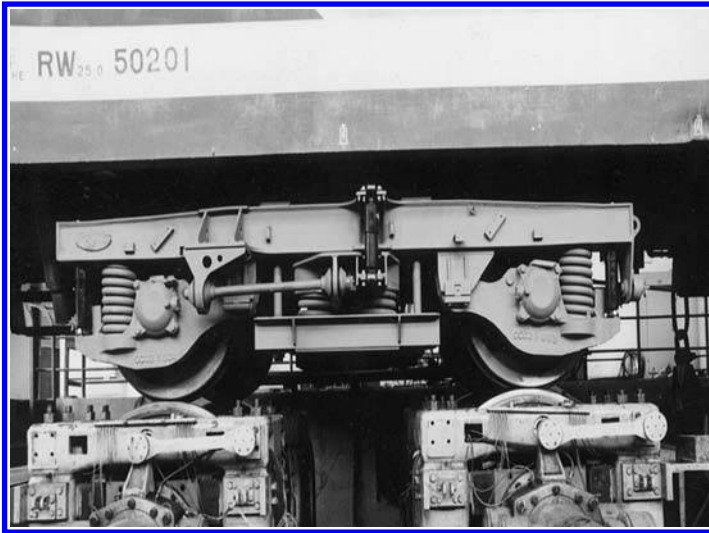


FIGURE 14.49 Wide-gauge passenger car bogie on roller rig.

**TABLE 14.6**  
Degrees of Freedom of the Vehicle System

DOF Body	Lateral	Vertical	Roll	Pitch	Yaw
Car body	$Y_c$	$Z_c$	$\Phi_c$	$\theta_c$	$\psi_c$
Bogie frame	$Y_{b1\sim 2}$	$Z_{b1\sim 2}$	$\Phi_{b1\sim 2}$	$\theta_{b1\sim 2}$	$\psi_{b1\sim 2}$
Wheelset	$Y_{w1\sim 4}$	$Z_{w1\sim 4}$	$\Phi_{w1\sim 4}$	—	$\psi_{w1\sim 4}$

The total number of degrees of freedom of the vehicle system, which is listed in Table 14.6, is 31.

The dynamic simulation program TPLDYNA, developed by the State Key Laboratory of Traction Power, was used for the dynamic performance computations of the vehicle on rail. The

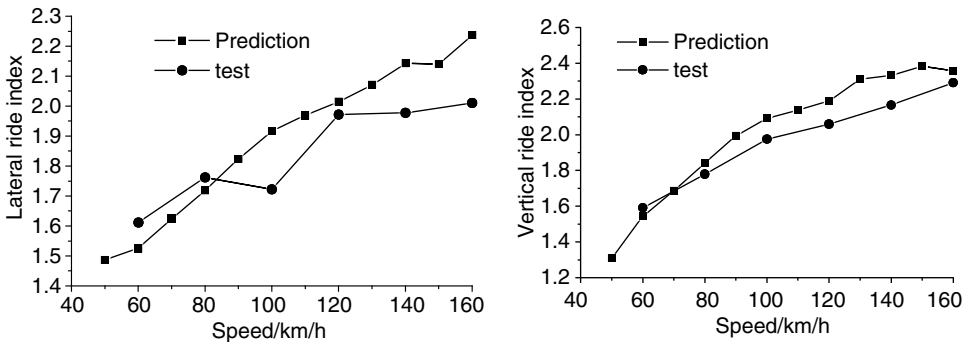


FIGURE 14.50 Comparison of test results and theoretical predictions.

stochastic irregularities of the U.S.A. grade 5 track irregularity spectrum was used in the dynamic simulation and also for ride comfort testing on the roller rig.

Through the simulation and test, the vibration accelerations on the car body were determined. The ride index calculation method is in accordance with Chinese national standard GB5599-85 (Railway Vehicles Specification for evaluation of the dynamic performance and accreditation test).

The lateral and vertical ride indices of the test and theoretical predictions, at different speeds, are shown in [Figure 14.50](#). It can be seen that the test results are a little lower than the simulation results when at higher running speeds.

## VI. CONCLUSIONS

From a brief historical review of roller test rigs for railway vehicles, it can be seen that roller rigs have played an important role in railway history, especially in the development of high-speed trains. Roller rigs are very useful in the testing of railway vehicles for the purpose of examining dynamic performance, estimating and optimising parameters, checking capabilities in extreme situations (e.g., high speed, violent impact), etc. Although the roller rig is becoming increasingly complex, in order to simulate vehicle running accurately, the inherent errors in wheel–rail contact must always be considered and these lead to test results for hunting speed, vibration response (ride index) and wheel–rail force under curve simulation which can differ from field tests. For this reason, roller rigs cannot replace field testing. Now some famous roller rigs (such as the roller rigs in Derby, Pueblo, and Munich) have been decommissioned, it could be said that the golden age for roller rigs has passed. The reason is probably that roller rigs have already taught us many things and knowledge has increased with respect to vehicle design, structures, and performance. Additionally, as simulation techniques become more advanced the need for costly testing plants diminishes. Since the roller rig is useful in developing high-speed trains, they will still be used in developing countries such as China.

## REFERENCES

1. Jaschinski, A., Chollet, H., Iwnicki, S., Wickens, A., and Von Würzen, J., The application of roller rigs to railway vehicle dynamics, *Vehicle Syst. Dyn.*, 31, 345–392, 1999.
2. Anon, Locomotive testing plant at Sweden, *The Engineer*, 100, 621–622, 1905.
3. Carling, D. R., Locomotive testing stations, parts 1 and 2, *Trans. Newcomen Soc.*, 45, 105–182, 1972.
4. Weihua, Z., *Dynamic Simulation Study of Railway Vehicle*, Ph.D. Dissertation, Southwest Jiaotong University, P.R. of China, 1996, June.
5. Ma, S. J., Zhang, W. H., Chen, G. X., and Zeng, J., Full scale roller rig simulation for railway vehicles, 13th IAVSD Symposium of Vehicles on Roads and Tracks, Chengdu, P.R. of China, August 1993, *Vehicle Syst. Dyn.*, 23(Suppl.), 346–357, 1994.
6. Deutsche Bahn, A. G., Welte Druck GmbH, Central Division for Research and Testing Research and Test Centre 3 *The Roller Rig*, Munich-Freimann, 1994, December.
7. Rizzi, C. and De Luca, S., La sala sperimentazione veicoli ferroviari del centro ricerche ansaldo trasporti. *Ingegneria Ferroviaria*, 48(7), 459–464, 1993, in Italian.
8. Vitrano, F. and Ghislanzoni, F., Descrizione della sala sperimentazione veicoli ferroviari. *Ingegneria Ferroviaria*, 48(7), 465–479, 1993, in Italian.
9. Vitrano, F., Pasten, L., and Flego, U., Le prove su veicoli ferroviari mediante banco a rulli. *Ingegneria Ferroviaria*, 48(7), 480–492, 1993, in Italian.
10. Hoshiya, S., Rolling stock testing plant capable of testing at a speed of 450 km/h, *Jpn. Rail. Eng.*, 102, 12–15, 1987.
11. Miyamoto, M., JR rolling stock testing plant capable of testing at a speed of 500 km/h, *Q. RTRI*, 32(4), 211–212, 1991.

12. Iwnicki, S. D. and Shen Z. Y., Collaborative railway roller rig project, *Proceedings of the SEFI World Conference on Engineering Education*, Portsmouth, p. 255, September 1992.
13. Cooperrider, N. K., Law, E. H., Fries, R. H., and Haque, I., *State-of-the-Art-Car (SOAC), creep forces and dynamic response on the Roll Dynamics Unit*, Arizona State University, Department of Mechanical and Aerospace Engineering, Report CR-R 82037, 1982 September.
14. Fries, R. H., *Estimation of transit rail vehicle parameter from roller rig tests*, Ph.D. Dissertation, Arizona State University, 1983.
15. Fries, R. H., Dynamic response of a transit rail vehicle on a roller rig, The Dynamics of Vehicle on Road and Tracks, Eighth IAVSD Symposium of Vehicles on Roads and Tracks, Cambridge, MA, August 15–19, 1983, *Vehicle Syst. Dyn.*, 13, 1984.
16. Fries, R. H., Cooperrider, N. K., and Law, E. H., Experimental investigation of freight car lateral dynamics, *J. Dyn. Syst., Meas., Control*, 103, 201–210, 1981.
17. Matsumoto, A., Sato, Y., Nakata, M., Tanimoto, M., and Qi, K., Wheel–rail contact mechanics at full scale on the test stand, *Wear*, 191, 101–106, 1996.
18. Dreher, G., Jobst, F., Götz, G., Mauer, L., Meinke, P., Mielcarek, A., Nestmeier, J., and Örley, H., Active hunting control of a wheel/rail vehicle demonstrated up to 530 km/h on the German roller test rig, The Dynamics of Vehicle on Road and Tracks. Eighth IAVSD Symposium of Vehicles on Roads and Tracks, Cambridge, MA, August 15–19, 1983, *Vehicle Syst. Dyn.*, 13, 1984.
19. Elkins, J. A. and Wilson, N. G., Train resistance measurements using a roller rig, The Dynamics of Vehicle on Road and Tracks, Ninth IAVSD Symposium of Vehicles on Roads and Tracks, Linköping, Sweden, June, 1985, *Vehicle Syst. Dyn.*, 15, pp. 86–99, 1986.
20. Dai, H. Y., Robust Performance Analysis of Active Suspension with model Uncertainty Using Structured Singular Value,  $\mu$  approach, 13th IAVSD Symposium of Vehicles on Roads and Tracks, Budapest, Hungary, August 1997, *Vehicle Syst. Dyn.*, 29(Suppl.), pp. 635–647, 1998.
21. Haque, I. and Law, E. H., Steady state techniques in creep force estimation: results from full scale tests of the SOAC vehicle on a roller rig, *J. Dyn. Syst., Meas., Control*, 111, 61–68, 1989.
22. Francesco, B., Stefano, B., and Ferruccio, R., Wear of railway wheel profiles: a comparison between experimental results and a mathematical model, 17th IAVSD Symposium of Vehicles on Roads and Tracks, Lyngby, Denmark, August 2001, *Vehicle Syst. Dyn.*, 37(Suppl.), pp. 478–489, 2003.
23. Zhang, W. H., Chen, J. Z., Wu, X. J., and Jin, X. S., Wheel/rail adhesion and analysis by using full scale roller rig, *Wear*, 22253(1–2), pp. 82–88, 2003.
24. Eisaku, S. and Masayuki, M., Geometrical contact between wheel and roller rig, *Nippon Kikai Gakkai Ronbunshu, C Hen/Transactions of the Japan Society of Mechanical Engineers, Part C*, 59(562), 1686–1693, 1993.
25. Keiji, Y., Norio, F., and Tatsuro, S., Fundamental consideration of geometrical contact between wheel and rail, *Nippon Kikai Gakkai Ronbunshu, C Hen/Transactions of the Japan Society of Mechanical Engineers, Part C*, 56(528), 2132–2138, 1990.
26. Zhan, W. H. and Chen, L. Q., A study of wheel–roller contact relations on rail vehicle roller rig system, *J. Southwest Jiaotong Univ.*, 30(1), 76–91, 1995, in Chinese.
27. Zhan, F. S., *Locomotive Dynamics*, Issue: Publishing Company of Chinese Railway, 1990, (in Chinese).
28. Zhang, W. H. and Chen, L. Q., Analysis of difference of rail vehicle stability on track and on roller rig, *J. China Rail. Soc.*, 19(1), 32–36, 1997.
29. De Pater, A.D., The motion of a railway wheelset supported by a pair of rollers as compared with the motion of such a wheelset along a tangent track, Delft University of Technology, Laboratory for Engineering Mechanics, Report 1012, 1993.
30. Dukkipati, R. V., Analysis of the lateral stability of a Track on NRC curved track simulator, *J. Southwest Jiaotong Univ.*, 1994.
31. Zeng, J., Zhang, W. H., Dai, H. Y., Wu, X. J., and Shen, Z. Y., Hunting instability analysis and  $H_{\infty}$  controlled stabilizer design for high speed railway passenger car, 13th IAVSD Symposium of Vehicles on Roads and Tracks, Budapest, Hungary, August 1997, *Vehicle Syst. Dyn.*, 29(Suppl.), pp. 655–668, 1998.
32. Yan, J. M., Wang, K. W., and Dai, H. Y., A Study of the high speed passenger bogies utilizing the full scale roller testing rig, *Vehicle Syst. Dyn.*, 23(Suppl.), 554–565, 1993.

33. Dukkipati, R.V., Parametric study of the lateral stability of a rail bogie on a roller rig, Proceedings of the Institution of Mechanical Engineers, Part F, *J. Rail Rapid Transit*, 213(1), pp. 39–47, 1999.
34. Zhang, W. H., Influence of roller rig state errors to rail vehicles stability testing results, *Chinese J. Mech. Eng.*, 33(3), 49–53, 1997.
35. Wu, P. B., *Dynamic Performance Computations for the Pakistan Wide-Gauge Passenger Car*, Research Report, No. TPL2001-09, State Key Laboratory of Traction Power, Southwest Jiaotong University, P.R. of China, September 2001.
36. Zeng, J. and Dai, H. Y., *Dynamic Performance Test for the Pakistan Wide-Gauge Coil Spring Bogie of Passenger Car Using Roller Testing Rig*, Research Report, No. TPL2002-09, State Key Laboratory of Traction Power, Southwest Jiaotong University, P.R. Of China, May 2002.

---

# 15 Scale Testing

*P. D. Allen*

## CONTENTS

I.	Introduction .....	507
II.	A Brief History of Scaled Roller Rigs.....	508
III.	Survey of Current Scaled Roller Rigs .....	508
	A. The Scaled Rig of DLR .....	508
	1. Design Overview.....	509
	B. The Scaled Rig of MMU .....	509
	1. Design Overview.....	509
	C. The Scaled Rig of INRETS .....	510
	1. Design Overview.....	512
IV.	Roller Rigs: The Scaling Problem .....	512
	A. The Scaling Strategy of MMU .....	515
	1. Principles.....	515
	2. Materials.....	515
	3. Equations of Motion .....	516
	4. Scaling and Wheel–Rail/Roller Forces.....	516
	B. The Scaling Strategy of DLR.....	519
	C. The Scaling Strategy of INRETS .....	521
	D. Tabular Comparison of Scaling Strategies.....	523
V.	Scaling Errors .....	523
VI.	Conclusions .....	525
	Acknowledgments .....	525
	References.....	526

## I. INTRODUCTION

The use of roller rigs for the investigation of railway vehicle dynamics has been discussed in Chapter 14. Their operation, application, and the changes in vehicle response introduced due to geometric and kinematic differences between running a rail vehicle on rollers, as opposed to track, were described in detail. Full scale rigs are a useful tool when assessing the dynamic performance of a prototype vehicle, especially in the days when numerical simulation of vehicle dynamics was not as well developed as it is today, but their frequency of use for prototype design is in decline, as computer techniques become more popular. Their high costs act against their widespread use.

Where the area of research is rather broader, and the behaviour of a particular prototype vehicle is not the primary area of interest, then the use of a scaled roller rig can offer a number of advantages. The most obvious of these advantages is, of course, the space occupied by the rig, this is coupled to a large cost saving and an ease of operability. A scaled rig is much easier to maintain and the mechanical handling of the test vehicle is more manageable. It is also far easier to change



a large number of vehicle parameters without great effort. However, these advantages must be offset by a number of negative factors. These are primarily concerned with the effect of scaling down the vehicle dimensions. From a scientific viewpoint, it is not acceptable to reduce the dimensions of the vehicle without giving due consideration to the effect of these changes. It is of great importance, if reliable scaled results are to be obtained, to adopt a scientifically based scaling strategy. The outcome of this strategy will dictate how well the rig will relate to the full scale, whether this is in terms of vehicle dynamics, wheel–rail forces or even wear.

This chapter describes a number of scaled roller rigs, used as research tools, and how each of the institutions involved have handled the issues related to scaling. Examples are given of the errors which can be introduced by different types of scaling strategies.

The fundamental ideas of similarity, that is, maintaining correlation between a scale model and the full scale, can be traced back to the work of Reynolds,<sup>1,2</sup> or even earlier. Analogous to Reynolds approach, similarity of mechanical systems with respect to dynamic behaviour and elastic deformation can be defined.

Small-scale testing of railway vehicles on roller rigs has been carried out for different purposes, including the verification and validation of simulation models, the investigation of fundamental railway vehicle running behaviour (nonlinear response, limit cycles, etc.), for the development and testing of prototype bogie designs with novel suspensions, in order to support field tests and computer simulations and last but not least for teaching and demonstration of railway vehicle behaviour. Small-scale tests at various institutions have proven that under laboratory conditions, influences of parameters can be revealed which often cannot be separated from stochastically affected measurements of field tests, which is of course also true for full scale rigs.

## II. A BRIEF HISTORY OF SCALED ROLLER RIGS

Investigations using scaled models of railway vehicles on scaled tracks were performed by Sweet et al.<sup>3,4</sup> at Princeton University in 1979 and 1982. Experiments were concerned with the mechanics of derailment of dynamically scaled 1/5 model of a typical three-piece freight truck design widely used in North America. Careful attention was given to the scaling of clearances. Forces were scaled according to similarity laws, including the effects of inertia, gravitation, spring stiffness, creep, and dry friction. These methods have been adopted and are used in many of the currently adopted scaling strategies.

One of the first investigations in Germany on a scaled roller rig was performed by the RWTH Aachen.<sup>5</sup> Other designs of scaled roller rigs followed, in 1984 at DLR Oberpfaffenhofen,<sup>6–8</sup> in 1985 at the Institut National de Recherche sur les Transports et leur Securite (INRETS) in Arcueil,<sup>9</sup> and in 1992 at the Rail Technology Unit of the Manchester Metropolitan University (MMU).<sup>10–12</sup>

## III. SURVEY OF CURRENT SCALED ROLLER RIGS

There are a number of scaled roller rigs which are used for research and demonstration purposes. The design and operation of some of the scaled roller rigs in use today are described in the following section.

### A. THE SCALED RIG OF DLR

The Institute for Robotics and System Dynamics of DLR has been involved in the development of simulation software for railway vehicle dynamics based upon multibody modelling techniques since the early 1970s.

The institute was interested in the nonlinear running behaviour of railway passenger vehicles and experiments became very important for the validation of modelling work which was carried out to predict the dynamic response of the vehicle. In wheel–rail dynamics, the nonlinear forces

involved play a dominant role in the onset of vehicle hunting, a phenomenon which is caused by a bifurcation of the system's equations of motion into a periodic solution or limit cycle as it is commonly referred to. For this reason, DLR developed a scaled roller rig, with a single bogie vehicle running on the rollers. The primary functions of the rig were to perform the above validation but also to assist in the verification of parts of DLR's dynamic simulation software, SIMPACK.

The bogie was a scaled-down version of the MAN bogie.<sup>6,7</sup> The emphasis of the first series of tests were the fundamentals of modelling and experimental methods in wheel–rail dynamics. Once this first stage of work was completed, including investigations of limit cycle behaviour of the bogie,<sup>8</sup> DLR adapted the rig to concentrate on the development of unconventional wheelset concepts, and contributed to fundamental research in this field.<sup>13</sup>

## 1. Design Overview

The roller rig at DLR is a 1/5 scale rig consisting of two rollers, each of which is composed of a hollow cylinder, with a wall thickness of about 20 mm. At each end of the cylinder a disc is attached, which has formed around its circumference a 1/5 scale UIC60 rail profile. The diameter of this part of the roller is 360 mm and the separation of the disc is 287 mm, which is 1/5 of the standard track gauge of 1435 mm. The advantage of this type of roller construction is that the design provides a very high torsional stiffness, which is important in maintaining a true creepage relationship between the wheel and roller. This stiffness coupled with a large rotational inertia, which makes the rollers insensitive to disturbances of their rotational velocity, makes the arrangement well suited to simulating tangent track behaviour. A plan view drawing of the rig is shown in [Figure 15.1](#).

The distance between the rollers can be varied to accommodate different bogies with different wheelbase. A feature of this arrangement is the inclusion of a “Schmidt-Coupling,” this device is a parallel crank mechanism and allows the change of wheelbase without disruption to the drive arrangement. As can be seen in the sectional view, the rollers are mounted on cones, this allows easy removal of the rollers for changes to rail profile or gauge.

The rollers are interconnected using a toothed belt with a specified longitudinal stiffness to maintain the synchronisation of the roller speeds at all times. The roller speed can be varied from 0 up to 168 km/h, depending on the rolling resistance of the vehicle being modelled. The general arrangement, showing a MAN bogie being tested on the rig is shown in [Figure 15.2](#).

## B. THE SCALED RIG OF MMU

A 1/5 scale roller rig was set up at Manchester Metropolitan University (MMU) in 1992 for use in the investigation of railway vehicle dynamic behaviour and to assist in research, consultancy, and teaching activities. Research activities then focused on the evaluation of a novel design of differentially rotating wheelset and the quantification of errors inherent in roller rig testing.<sup>12</sup> The roller rig is currently being used to investigate the behaviour of independently driven wheelsets for light rail applications.

### 1. Design Overview

The roller rig at Manchester Metropolitan University is of 1/5 scale, and consists of four rollers supported in yoke plates incorporating the rollers supporting bearings, with the interconnection between the roller pairs being provided by the use of splined and hook jointed shafts. While these shafts do not offer the degree of torsional stiffness given by the DLR arrangement, they do allow the simulation of lateral track irregularities by enabling rotational movement of the rollers about a vertical axis (yaw), coupled with a lateral movement of the rollers as a pair. The roller motion is provided by servo hydraulic actuators which are connected directly to the rollers supporting yoke plates, these actuators being controlled by a digital controller which allows the inputs to follow

defined waveforms or measured track data. The longitudinal and lateral position of the rollers can be adjusted by means of a system of linear bearings for changing the wheelbase and the gauging of the rollers. Drive is supplied to the rollers via a belt, with pulleys on each roller drive shaft, allowing the rig to operate at scaled speeds of up to 400 km/h. The bogie is modelled on the BR Mk IV passenger bogie,<sup>14</sup> but it can be easily modified. The purpose of the rig was to demonstrate the behaviour of a bogie vehicle under various running conditions and acquire nominal data from the vehicle responses. A plan view drawing of the roller rig is shown in Figure 15.3.

The bogie vehicle parameters were selected to represent those of a typical high speed passenger coach (the BR Mk4 passenger coach). The wheel profiles are machined scaled versions of BR P8 profile and the rollers have a scale BS110 rail profile with no rail inclination. The bogie running on the rig can be seen in Figure 15.4.

### C. THE SCALED RIG OF INRETS

INRETS is the French national research institute and within this is a group specialising in wheel-rail interaction, with particular interest in the novel variations of freight bogies. The test facility was originally commissioned in 1984 and was used intensively until 1992.

The first railway vehicle to be tested on the rig was the Y25 type, a UIC bogie,<sup>15</sup> which is very common in Europe. The bogie was selected as it is particularly difficult to model with conventional

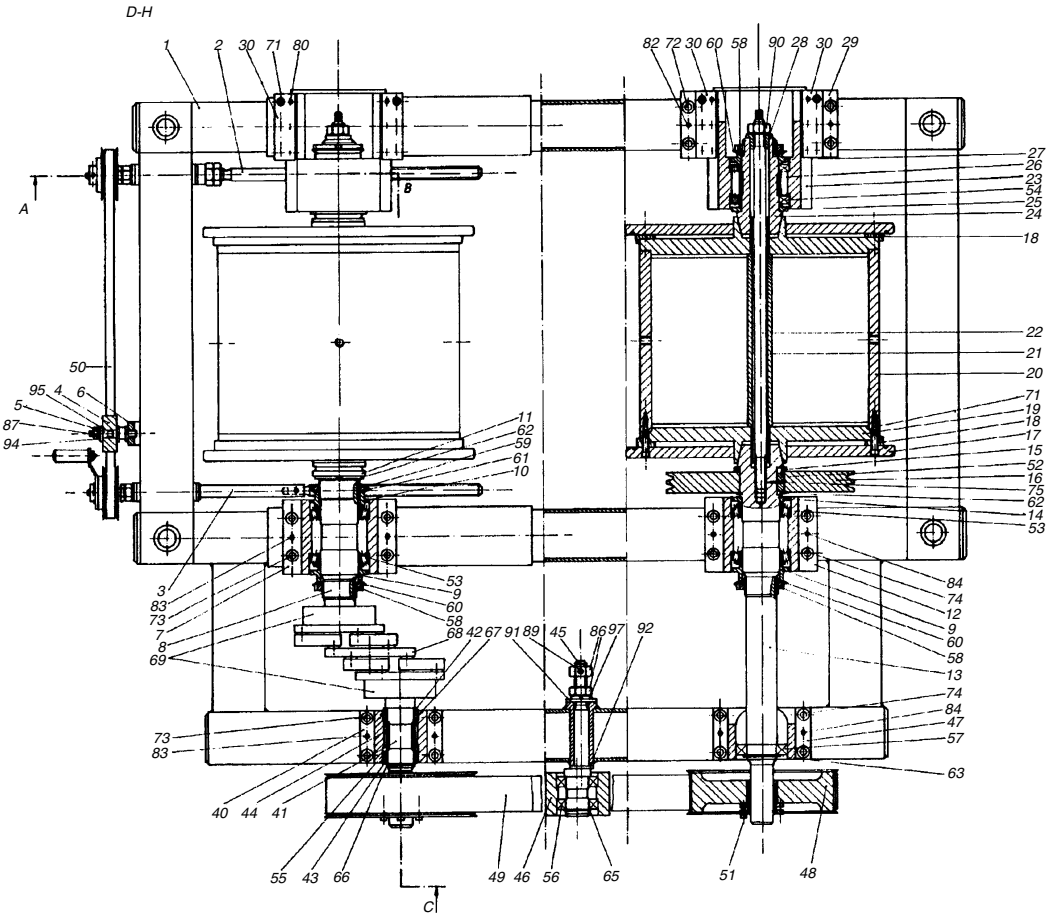
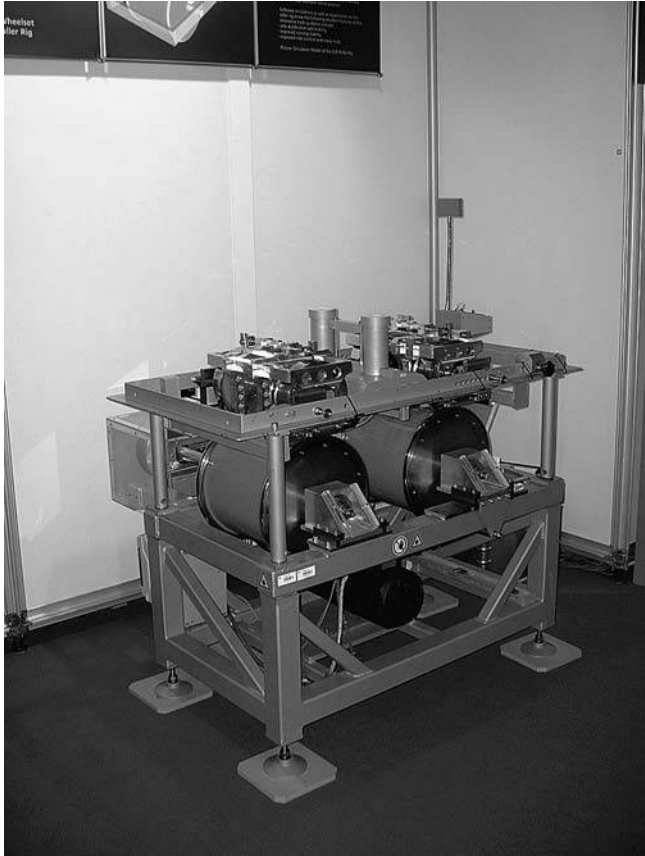
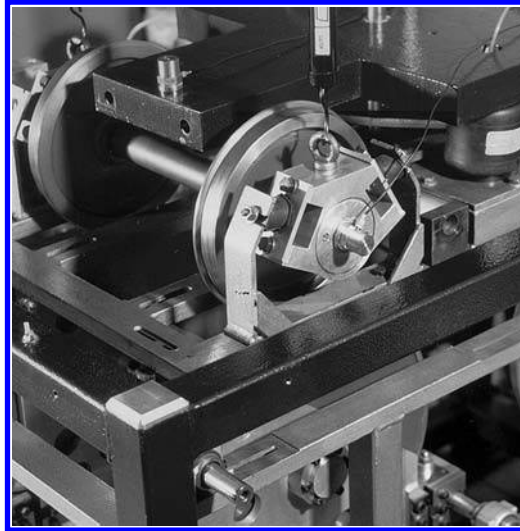


FIGURE 15.1 Plan view drawing of the 1/5 scale DLR roller rig.





**FIGURE 15.4** Bogie vehicle on the MMU roller rig platform.

computer software as there are several dry friction dampers within the suspension. Studies focused on optimising the stability of the bogie under varying vertical loads and differing suspension parameters.

Latterly the rig has been used for the quasistatic measurement of the Kalker coefficients and currently experiments are being carried out dealing with squeal noise and braking performance.

### 1. Design Overview

The test rig of the Institut National de Recherche sur les Transports et leur Sécurité (INRETS) was originally designed as a large flywheel of 13 m diameter to test linear motors for the Bertin AeroTRAIN transport vehicle and this is the reason the wheel is of such large diameter. Weighing 40 t, the wheel is driven by a linear 2 MW motor, which can power the wheel to a periphery speed of 250 km/h.

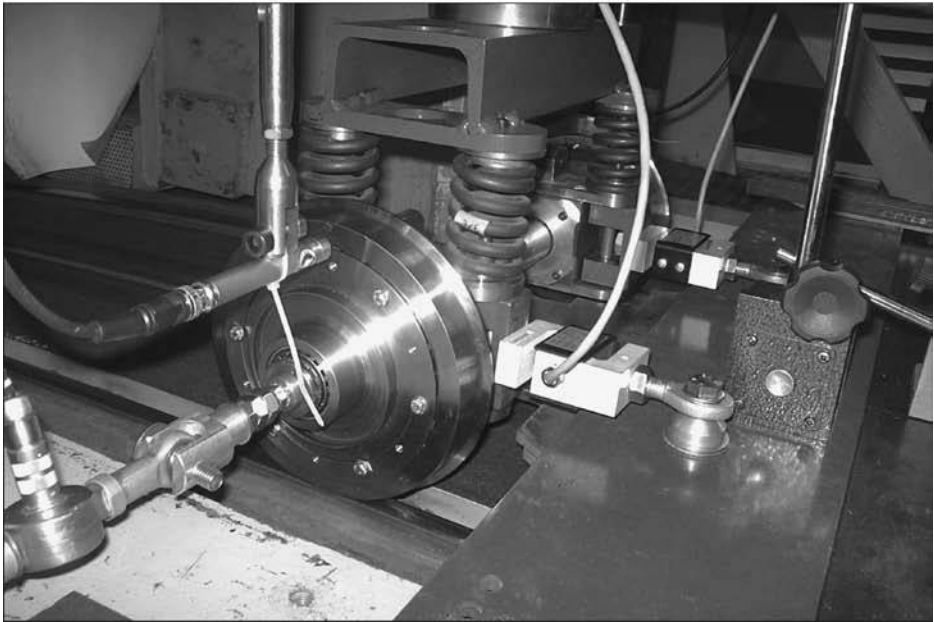
The flywheel was not designed to support very high vertical loads and therefore a scaling factor of 1/4 was chosen for the rig. INRETS were already familiar with similarity laws used in scale models, and developed a specific strategy for dynamic similarity, with respect to the preservation of the elasticity of the bodies, especially for the wheel–rail contact area. The rig is illustrated in [Figure 15.5](#).

The large diameter of the test wheel made the rig at INRETS particularly suitable for investigating the contact between wheel and rail as the radius of curvature was far closer to approaching that of conventional track, when compared to any other roller rig in existence, resulting in the size and shape of the contact patch being closer to reality.

The wheel can only rotate about the horizontal axis, and therefore an angle of attack can only be generated by yawing the vehicle wheelset relative to the track. A hydraulic ram is fitted at the top of the wheel to allow the variation of the vertical load on the tested bogie.

## IV. ROLLER RIGS: THE SCALING PROBLEM

Similarity laws and the correlated problem of scaling are of interest for the transformation of experimental results from a scaled model to the full scale design. There are various possible



approaches to scaling, including using the methods of dimensional analysis to establish several dimensionless groups from which the scaling factors can be derived, workers include, Jaschinski,<sup>8</sup> Illingworth,<sup>21</sup> and Chollet.<sup>15</sup> Other methods include first deriving the equations of motion and then calculating the scaling factors required for each term to maintain similarity.

This latter method is known as inspectional analysis and requires a sound understanding of the equations of motion, which is achievable in this particular field.

Choice of material properties is also a factor in the scaling method used, particularly if the simulation work requires the preservation of the levels of strain at the contact point. British Rail used aluminum wheels and rollers,<sup>18</sup> while Matsudaira et al.,<sup>19</sup> used steel, and Sweet et al.<sup>3</sup> used plastic.

The starting point for defining a system of similarity is the definition of the length scaling factor and defining the general terms. These general terms are outlined below, it is from these that the scaling strategies of the three research institutions are developed.

$$\varphi_l = \frac{l_1}{l_0} \quad (15.1)$$

where  $l_1$  is a characteristic length of the full scale and  $l_0$  that of the scaled model. In the same way, a time scaling factor can be derived.

$$\varphi_t = \frac{t_1}{t_0} \quad (15.2)$$

With these definitions, scaling factors for cross-section,  $\varphi_A$ , volumina,  $\varphi_V$ , velocity,  $\varphi_v$ , and acceleration,  $\varphi_a$ , follow:

$$\varphi_A = \varphi_l^2 \quad (15.3)$$

$$\varphi_V = \varphi_l^3 \quad (15.4)$$

$$\varphi_v = \frac{\varphi_l}{\varphi_t} \quad (15.5)$$

$$\varphi_a = \frac{\varphi_l}{\varphi_t^2} \quad (15.6)$$

When the density scaling  $\varphi_\rho$  is

$$\varphi_\rho = \frac{\rho_1}{\rho_0} \quad (15.7)$$

then the scaling factors for mass,  $\varphi_m$ , moment of inertia,  $\varphi_I$ , and inertial force,  $\varphi_F$ , can be derived:

$$\varphi_m = \varphi_\rho \varphi_l^3 \quad (15.8)$$

$$\varphi_I = \varphi_m \varphi_l^2 \quad (15.9)$$

$$\varphi_F = \frac{m_1 a_1}{m_0 a_0} = \varphi_m \varphi_a = \frac{\varphi_\rho \varphi_l^4}{\varphi_t^2} \quad (15.10)$$

Once these general definitions have been developed, the scaling strategies of each institution can be used to derive the following quantities used in studies of wheel–rail interaction:  $\varphi_T$ , scaling factor for creep forces;  $\varphi_{ab}$ , scaling factor for the elliptical size of the contact patch;  $\varphi_E$ , scaling

factor for Young's modulus;  $\varphi_\nu$ , scaling factor for Poisson's ratio;  $\varphi_\varepsilon$ , scaling factor for strain;  $\varphi_\sigma$ , scaling factor for stress;  $\varphi_\mu$ , scaling factor for the active coefficient of friction;  $\varphi_c$ , scaling factor for stiffness;  $\varphi_d$ , scaling factor for damping; and  $\varphi_f$ , scaling factor for frequency.

## A. THE SCALING STRATEGY OF MMU

### 1. Principles

The important aspects of the behaviour that are being studied in a dynamic analysis are the displacements, velocities, and acceleration of the various bodies and the forces between these bodies and at the wheel–rail/roller interface.

As the most common measurements in dynamic studies are made in the form of time histories or frequency spectra the scaling factor for time and therefore frequency should be unity.

$$\varphi_t = 1 \quad (15.11)$$

The roller rig has been built to 1/5 full size to give suitable dimensions for construction and laboratory installation:

$$\varphi_l = 5 \quad (15.12)$$

Following the Equation 15.1 to Equation 15.6, gives rise to the following expressions:

$$\text{for displacement} \quad \varphi_l = 5 \quad (15.13)$$

$$\text{for velocity} \quad \varphi_v = 5 \quad (15.14)$$

$$\text{for acceleration} \quad \varphi_a = 5 \quad (15.15)$$

and

$$\text{for frequency} \quad \varphi_f = \frac{1}{\varphi_t} = 1 \quad (15.16)$$

which is convenient for comparison of these values.

### 2. Materials

Various options were available for the material used in the construction of the roller rig, but for ease of construction and to allow a reasonably practical wear life of the wheels and rollers it was convenient to use steel for these bodies. This is not a great disadvantage as the roller rig is not used to perform wear investigations, which would require correlation with the full scale case. The material properties are then similar on scale and full size:  $\varphi_\rho = 1$  for density;  $\varphi_E = 1$  for Young's modulus;  $\varphi_\nu = 1$  for Poisson's ratio;  $\varphi_\mu = 1$  for coefficient of friction.

Therefore the scaling factor for mass, considering Equation 3.7.8 is:

$$\varphi_m = 5^3 \quad (15.17)$$

and for rotational inertia, due to Equation 15.9:

$$\varphi_I = 5^5 \quad (15.18)$$



### 3. Equations of Motion

The equations of motion for a dynamic system govern the relationship between force and acceleration (and therefore velocity and displacement). In general terms, the basic equation is expressed in the form of a force balance and all force terms in the equation, for similarity, should equate to the force scaling term,  $\varphi_F$ .

$$m\ddot{x} + c\dot{x} + kx = F \tag{15.19}$$

and in the angular form:

$$I\ddot{\theta} + c_T\dot{\theta} + k_T\theta = T \tag{15.20}$$

where  $m$  is the mass;  $I$  is the moment of inertia;  $c, c_T$  are the damping coefficients;  $k, k_T$  are the stiffnesses;  $F$  is the applied force; and  $T$  is the applied torque.

Therefore, for the scale model, Equation 3.7.11 and Equation 3.7.12 become:

$$m\ddot{x}\left(\frac{\varphi_m\varphi_l}{\varphi_f^2}\right) + c\dot{x}\left(\frac{\varphi_c\varphi_l}{\varphi_f}\right) + kx(\varphi_k\varphi_l) = F(\varphi_F) \tag{15.21}$$

$$I\ddot{\theta}\left(\frac{\varphi_l}{\varphi_f^2}\right) + c_T\dot{\theta}\left(\frac{\varphi_{c_T}}{\varphi_f}\right) + k_T\theta\left(\frac{\varphi_{k_T}}{\varphi_f}\right) = T(\varphi_T) \tag{15.22}$$

For the translational case, from Equation 3.7.13 and for similarity:

$$\left(\frac{\varphi_m\varphi_l}{\varphi_f^2}\right) = \left(\frac{\varphi_c\varphi_l}{\varphi_f}\right) = (\varphi_k\varphi_l) = (\varphi_F) \tag{15.23}$$

therefore using the previously derived scaling factors for,  $\varphi_l, \varphi_m$ , and  $\varphi_f$ :

$$\varphi_l^4 = \varphi_c\varphi_l = \varphi_k\varphi_l = \varphi_F \tag{15.24}$$

giving  $\varphi_c = 5^3$  for the translational damping coefficient;  $\varphi_k = 5^3$  for the translational stiffness constant; and  $\varphi_F = 5^4$  for the applied force.

For the rotational case, from Equation 3.7.14 and for similarity:

$$\left(\frac{\varphi_l}{\varphi_f^2}\right) = \left(\frac{\varphi_{c_T}}{\varphi_f}\right) = \left(\frac{\varphi_{k_T}}{\varphi_f}\right) = (\varphi_T) \tag{15.25}$$

therefore using the previously derived scaling factors for,  $\varphi_l$  and  $\varphi_f$ :

$$\varphi_l = \varphi_{c_T} = \varphi_{k_T} \tag{15.26}$$

giving  $\varphi_{c_T} = 5^5$  for the rotational damping coefficient;  $\varphi_{k_T} = 5^5$  for the rotational stiffness constant; and  $\varphi_T = 5^5$  for the applied torque.

The above terms of power  $x^5$  are validated by considering that a translational spring of stiffness  $k$  will give a torsional stiffness of  $kl^2$ , hence giving rise to the power raise of two. Therefore similarity is maintained in all equations with forces scaling at  $5^4$  and torques scaling at  $5^5$ .

### 4. Scaling and Wheel–Rail/Roller Forces

A complete study of the scaling methodology must also include the effect of scaling on the equations governing the wheel–rail/roller interaction. A complete derivation of the equations of motion for a railway vehicle is not required for this study as the wheel–rail forces act through the

wheelset alone. Therefore the equations of motion for a single bogie vehicle, which includes some simple suspension forces is sufficient. The creep forces are derived from Kalker's linear theory.

The lateral equation of motion for a simple linear vehicle model can be represented by the following expressions:

$$m\ddot{y}_w + 2f_{22}\left(\frac{\dot{y}_w}{v} - \psi_w\right) + 2f_{23}\left(\frac{\dot{\psi}_w}{v} - \frac{\varepsilon_0}{l_0 r_0}\right) + \frac{w\varepsilon_0 y_w}{l_0} + d_y(\dot{y}_w - \dot{y}_b - a\dot{\psi}_b + h\dot{\theta}_b) + c_y(y_w - y_b - a\psi_b + h\theta_b) \quad (15.27)$$

and the terms influencing the yaw of the wheelset:

$$I_z \ddot{\psi}_w + 2f_{11}\left(\frac{l_0^2 \dot{\psi}_w}{v} + \frac{l_0 \lambda y_w}{r_0}\right) - 2f_{23}\left(\frac{\dot{y}_w}{v} - \psi_w\right) + 2f_{33}\left(\frac{\dot{\psi}_w}{v}\right) + c_\psi(\psi_w - \psi_b) \quad (15.28)$$

where  $m$  is the wheelset mass;  $y_w$  is the wheelset lateral displacement;  $y_b$  is the bogie lateral displacement;  $\psi_w$  is the wheelset yaw angle;  $\psi_b$  is the bogie yaw angle;  $\theta_b$  is the bogie roll angle;  $d_y$  is the wheelset–bogie lateral damping (per wheelset);  $c_y$  is the wheelset–bogie lateral stiffness (per wheelset);  $c_\psi$  is the wheelset–bogie yaw stiffness (per wheelset);  $w$  is the axle load;  $\lambda$  is the effective conicity;  $l_0$  is the semi gauge;  $a$  is half the bogie wheelbase;  $h$  is the height of the bogie centre of gravity above the wheelset axis;  $v$  is the forward speed of the vehicle;  $\varepsilon_0$  is the rate of change of contact angle with  $y_w$ ;  $r_0$  is the rolling radius with the wheelset central;  $f_{11}, f_{22}, f_{23}, f_{33}$  is Kalker's linear creep coefficients.

The equations governing the linear creep coefficients are:

$$\begin{aligned} f_{11} &= (ab)GC_{11} & f_{23} &= (ab)^{3/2}GC_{23} \\ f_{22} &= (ab)GC_{22} & f_{33} &= (ab)^2GC_{33} \end{aligned} \quad (15.29a)$$

where  $C_{ii}$  are Kalker's tabulated creep coefficients,  $G$  is the modulus of rigidity and  $a$  and  $b$  are the contact patch semi-axes.

Hertz theory governs the size of the contact patch and the relevant equations are also quoted as follows:

$$ab = mn[3\pi N(k_1 + k_2)/4k_3]^{2/3} \quad (15.29b)$$

where

$$k_1 = \frac{1 - \nu_R^2}{E_W} \quad k_2 = \frac{1 - \nu_W^2}{E_R} \quad (15.29c)$$

and

$$k_3 = \frac{1}{2} \left[ \frac{1}{r_1} + \frac{1}{r_1'} + \frac{1}{r_2} + \frac{1}{r_2'} \right] \quad (15.29d)$$

$m$  and  $n$  are the elliptical contact constants,  $N$  is the normal force and the other parameters are as previously quoted.

The scaling factors can then be calculated:

$$\varphi_{k_1} = \varphi_{k_2} = \frac{1}{\varphi_E} = 1 \quad (15.30)$$

$$\varphi_{k_3} = \frac{1}{\varphi_l} = 5^{-1} \quad (15.31)$$

If the scaling factor for the normal force,  $\varphi_N$  is  $5^4$ , as with all other forces then the scaling factor for the contact patch area,  $\varphi_{(ab)}$  will be:

$$\varphi_{(ab)} = \left( \varphi_N \frac{\varphi_{k_1}}{\varphi_{k_3}} \right)^{\frac{2}{3}} = \left( \frac{\varphi_F}{\varphi_{k_3}} \right)^{\frac{2}{3}} = 5^{3.33} \quad (15.32)$$

From Equation 15.32 and Equation 15.29a, we can evaluate the scaling for the linear creep coefficients.

Therefore:

$$\varphi_{f_{11}} = \varphi_{f_{22}} = \varphi_E \varphi_{(ab)} = 5^{3.33} \quad (15.33)$$

$$\varphi_{f_{23}} = \varphi_E (\varphi_{(ab)})^{\frac{3}{2}} = 5^5 \quad (15.34)$$

$$\varphi_{f_{33}} = \varphi_E (\varphi_{(ab)})^2 = 5^{6.66} \quad (15.35)$$

With a normal force scaling factor of  $5^4$ , we have a conflict with the vehicle weight scaling factor due to its mass multiplied by the acceleration due to gravity:

$$\varphi_w = \varphi_m \varphi_g = 5^3 \quad (15.36)$$

conflicting with

$$\varphi_N = 5^4$$

This conflict can be resolved by the use of support wires, with incorporated spring balances connected to each axle box, to remove the required amount of weight.

Considering the describing equations of motion given in Equation 15.27 and Equation 15.28, each term can be evaluated, including the scaling factors derived above for the linear creep coefficients, to check for the required scaling factor. A factor of  $5^4$  when considering a force term and  $5^5$  for a torque scaling factor. All terms agree with the scaling strategy and give perfect scaling apart from those listed below:

*Force terms* (required  $\varphi_F = 5^4$ )

$$2f_{22} \frac{\dot{y}_w}{v} \psi_w \quad \text{gives a force scaling, } \varphi_F = 5^{3.33} \quad (15.37)$$

$$\frac{w\varepsilon_0 y_w}{l_0} \quad \text{gravitational stiffness term gives, } \varphi_F = 5^3 \quad (15.38)$$

*Torque terms* (required  $\varphi_T = 5^5$ )

$$2f_{11} \left( \frac{l_0^2 \dot{\psi}_w}{v} + \frac{l_0 \lambda y_w}{r_0} \right) \text{ gives a torque scaling, } \varphi_T = 5^{4.33} \quad (15.39)$$

$$2f_{33} \left( \frac{\dot{\psi}_w}{v} \right) \quad \text{gives a torque scaling, } \varphi_T = 5^{5.66} \quad (15.40)$$

In practice, the value of  $f_{33}$  is much smaller than  $f_{11}$  and  $f_{22}$ , the gravitational stiffness term in Equation 15.38 is of the same order as the  $f_{33}$  terms during normal tread running of the wheel, therefore the major error sources with respect to scaling the simulated, scaled forces, with those of a full size vehicle are Equation 15.37 and Equation 15.39.

## B. THE SCALING STRATEGY OF DLR

AS DLR was involved in the development of simulation software for railway vehicle dynamics, and in particular the nonlinear lateral dynamics which leads to the instability known as hunting. This instability is caused by a bifurcation in describing differential equations into a limit cycle and the strategy for the scaling of the roller rig was developed with respect to this.

Therefore, much like the MMU group, the starting point for the DLR scaling strategy focused on the nonlinear lateral behaviour of a single wheelset, suspended to an inertially moving body. An example equation is described in Ref. 17, for a wheelset with conical treads. The first component of this system of two coupled equations of motion is shown below:

$$\frac{m}{\chi} \ddot{y}_w = \frac{I_y \Gamma v}{\chi r_0} \dot{\psi}_w \frac{m g b_0}{\chi} \frac{c_y}{\chi} y_w + T_y + T_x \psi_w \quad (15.41)$$

The symbols used above denote the same quantities as described in the previous section, with the addition or replacement of the wheelset's rotational moment of inertia,  $I_y$ ; the longitudinal creep force,  $T_x$ ; the lateral creep force;  $T_y$ ;  $\Gamma = \delta_0/l_0 - r_0\delta_0$ ; the cone angle,  $\delta_0$ ;  $\chi = \Gamma l_0/\delta_0$ ;  $b_0 = 2\Gamma + \Gamma^2(R_R + r_0)$ ; the transverse radius of the rail head,  $R_R$ .

Multiplying the scaleable parameters and variables in Equation 15.41 with the previously defined scaling factors and re-arranging:

$$\frac{m}{\chi} \ddot{y}_w = \frac{I_y \Gamma v}{\chi r_0} \dot{\psi}_w \frac{m g b_0}{\chi} y_w \frac{\varphi_l^2}{\varphi_l} \frac{c_y}{\chi} y_w \frac{\varphi_c \varphi_l^2}{\varphi_m} + (T_y + T_x \psi_w) \frac{\varphi_T \varphi_l^2}{\varphi_m \varphi_l} \quad (15.42)$$

Dynamically, the scale wheelset behaves similarly to the full scale, if Equation 15.41 and Equation 15.42 coincide. This requires that the following conditions hold:

$$\begin{aligned} \frac{\varphi_l^2}{\varphi_l} = 1 &\Rightarrow \varphi_v = \sqrt{\varphi_l} && \text{velocity scaling} \\ \frac{\varphi_c \varphi_l^2}{\varphi_m} = 1 &\Rightarrow \varphi_c = \varphi_\rho \varphi_l^2 && \text{stiffness scaling} \\ \frac{\varphi_T \varphi_l^2}{\varphi_m \varphi_l} = 1 &\Rightarrow \varphi_T = \varphi_\rho \varphi_l^3 && \text{creep force scaling} \end{aligned} \quad (15.43)$$

It can be seen that for similarity to be maintained the scaling factors for the Equation 15.43 above cannot be freely chosen and are a function of the principle scaling terms derived in Section IV.A. This result is identical to that found by Matsudaira et al.,<sup>19</sup> from investigations carried out in 1968 at the RTRI of the Japanese railways. From the constraint equations (a relationship between the normal forces, gyroscopic, gravitational, applied, and creep forces, see Ref. 8), together with the scaling method described above, the scaling factors for the constraint forces, the mass and creep

forces can be derived:

$$\varphi_N = \varphi_m = \varphi_T = \varphi_P \varphi_l^3 \tag{15.44}$$

This results in a scale factor for friction coefficient  $\mu$ :

$$\varphi_\mu = 1$$

Assuming that Kalker’s nonlinear theory is used for calculation of the contact forces, then similarity is required for the dimensions of the contact ellipse, if the calculated Kalker creep coefficients, and hence creep forces are to be correct. This requires that:

$$\varphi_E = \varphi_v = 1$$

If this condition is adhered to, then the scaling factor for density is derived as follows:

Kalker’s theory requires that  $\varphi_T = \varphi_{ab}$  and  $\varphi_{ab} = (\varphi_N \varphi_l)^{2/3}$  and  $\sqrt{\varphi_{ab}} = \varphi_e$  (the contact ellipse mean radius), then the scale of this radius becomes:

$$\varphi_e^3 = \varphi_N \varphi_l = \varphi_\rho \varphi_l^4 \tag{15.45}$$

Assuming geometric similarity for the contact ellipse,  $\varphi_e = \varphi_l$ , Equation 15.45, results in the definition of the density scaling factor:

$$\varphi_\rho = \frac{1}{\varphi_l}$$

This scaling factor, which would result in perfect scaling for the contact ellipse and Kalker coefficients, when considering the length scale factor  $\varphi_l = 5$ , requires a density which is very difficult to achieve. It was considered by DLR that exact scaling of the contact patch was only necessary at low levels of creepage and not so important during the analysis of limit cycles, where saturation of the creep forces occurs and the exact shape and size of the contact patch does not influence the creep forces (gross sliding within the contact region). Considering the above practical limitations, the density scaling factor was chosen as:

$$\varphi_\rho = \frac{1}{2}$$

which can be easily achieved and has proven through testing to give good experimental results. With

$$\varphi_l = 5$$

and considering the above-mentioned limitation with respect to density, the other scaling factors can be determined as follows:

$$\varphi_v = \sqrt{\varphi_l} = \sqrt{5} \quad \text{velocity}$$

$$\varphi_t = \frac{\varphi_l}{\varphi_v} = \sqrt{5} \quad \text{time}$$

$$\varphi_a = \frac{\varphi_l}{\varphi_t^2} = 1 \quad \text{acceleration}$$

$$\varphi_m = \varphi_T = \varphi_N = \varphi_F = \varphi_\rho \varphi_l^3 = 62.5 \quad \text{mass and force}$$

$$\begin{aligned}\varphi_I &= \varphi_\rho \varphi_l^5 = 1562.5 && \text{moment of inertia} \\ \varphi_c &= \varphi_\rho \varphi_l^2 = 12.5 && \text{spring stiffness} \\ \varphi_d &= \frac{\varphi_\rho \varphi_l^3}{\varphi_v} = \varphi_\rho \varphi_l^{5/2} = 27.95 && \text{viscous damping} \\ \varphi_f &= \frac{1}{\varphi_t} = \frac{1}{\sqrt{5}} && \text{frequency} \\ \varphi_\mu &= \frac{\varphi_T}{\varphi_N} = 1 && \text{coefficient of friction} \\ \varphi_e &= (\varphi_N \varphi_l)^{1/3} = 6.79 && \text{contact ellipse}\end{aligned}$$

Table 15.1 below shows some typical parameters for a generic test vehicle using the DLR scaling strategy.

## C. THE SCALING STRATEGY OF INRETS

Within the INRETS institution the main area of research focus was the experimental validation of Kalker's creep coefficients. The vehicle scale at INRETS is large compared to other rigs at 1:4, coupled with the very large roller diameter means the rig is suitable for the analysis of wheel/rail

**TABLE 15.1**  
**Generic Test Vehicle Parameters**

Parameter	Full-Size	1/5 Scale
<b>Bogie</b>		
Bogie frame mass	487.50 kg	7.8 kg
Wheel mass	281.25 kg	4.5 kg
Axle mass	275.00 kg	4.4 kg
Bogie roll inertia	218.75 kg/m <sup>2</sup>	0.14 kg/m <sup>2</sup>
Bogie pitch inertia	103.13 kg/m <sup>2</sup>	0.066 kg/m <sup>2</sup>
Bogie yaw inertia	192.19 kg/m <sup>2</sup>	0.123 kg/m <sup>2</sup>
Wheel rotational inertial	51.56 kg/m <sup>2</sup>	0.033 kg/m <sup>2</sup>
Axle rotational inertia	3.13 kg/m <sup>2</sup>	0.002 kg/m <sup>2</sup>
<b>Vehicle Body</b>		
Body mass	2037.50 kg	32.6 kg
Body roll inertia	1403.13 kg/m <sup>2</sup>	0.898 kg/m <sup>2</sup>
Body pitch inertia	1339.06 kg/m <sup>2</sup>	0.857 kg/m <sup>2</sup>
Body yaw inertia	2342.19 kg/m <sup>2</sup>	1.499 kg/m <sup>2</sup>
<b>Wheel Dimensions</b>		
Wheel diameter	1.0 m	0.2 m
Gauge	1.435 m	0.287 m
<b>Primary Suspension</b>		
Longitudinal stiffness	8.30 × 10 <sup>5</sup> N/m	6.64 × 10 <sup>4</sup> N/m
Lateral stiffness	8.30 × 10 <sup>5</sup> N/m	6.64 × 10 <sup>4</sup> N/m
Vertical stiffness	5.90 × 10 <sup>7</sup> N/m	4.73 × 10 <sup>6</sup> N/m
Normal force	11,496 N	183.94 N

contact. The validation of Kalker's theory requires exact representation of the contact patch and its elasticity, to allow accurate measurement of the quasistatic creepage and creep force relationships. Therefore the basis of the scaling strategy was obtained by adopting a stress scaling factor of

$$\varphi_{\sigma} = \frac{\varphi_F}{\varphi_l^2} = 1.$$

This means that the stresses in the scale and full scale test vehicle are the same. In addition to the advantages in investigating Kalker's theory, this stress scale factor results in a spring scaling factor which is proportional to the length factor. This helps in the design of suspension components as size and internal stresses are the same as the full scale:

$$\varphi_c = \frac{\varphi_F}{\varphi_l} = \varphi_l = 4 \quad (15.46)$$

When similarity of elastic forces, together with similarity of gravitational forces is required, then the following is true:

$$\varphi_c \varphi_l = \varphi_m \varphi_g \quad (15.47)$$

where  $\varphi_g$  is the scaling factor for gravity. Equation 15.46 shows that for the requirement of a valid frequency scaling factor, then the frequency of a mass spring system should be the same as that of an equivalent gravitational oscillator, such as a pendulum, this condition yields that:

$$\varphi_w^2 = \frac{\varphi_c}{\varphi_m} = \frac{\varphi_g}{\varphi_l} \quad (15.47a)$$

Assuming the density scaling factor  $\varphi_{\rho} = 1$ , Equation 15.47 leads to a gravity scaling factor of

$$\varphi_g = \frac{1}{\varphi_l} = \frac{g_1}{g_0} \quad (15.48)$$

The above equation essentially results in a different scaling factor for forces generated through gravitation and those generated from inertia and in a similar way to the MMU strategy, this can be achieved by application of external forces, which adds to the effective weight, without increasing mass. Considering Equation 15.42, this results in a scaling factor for weight of

$$\varphi_w = \frac{m_1 g_1}{m_0 g_0 \varphi_l} \Rightarrow \varphi_w = \varphi_l^2 = 16 \quad (15.49)$$

whereas the inertial force scaling factor with  $\varphi_g = 1$

$$\varphi_m \varphi_g = \varphi_{\rho} \varphi_l^3 \Rightarrow \varphi_w = \varphi_l^3 = 64 \quad (15.50)$$

Using this strategy, increasing the weight through an external force which does not change the mass of the body allows the derivation of scaling factors for velocity, time, and acceleration to be formed from the frequency Equation 15.5 and Equation 15.6.

$$\left( \frac{\varphi_v}{\varphi_l} \right) = \frac{\varphi_c}{\varphi_m} \Rightarrow \varphi_v = 1$$

$$\varphi_t = \varphi_l$$

$$\varphi_a = \varphi_l \quad (15.51)$$

**TABLE 15.2**  
**Comparison of Scaling Strategies**

Scaling	MMU	DLR	INRETS
<b>Geometry</b>			
Length	5	5	4
Cross-section	25	25	16
Volume	125	125	64
<b>Material</b>			
Density	1	0.5	1
Mass	125	62.5	64
Inertia	3125	1562.5	1024
Elasticity	G,E,cij	approximate	G,E,cij
<b>Parameters</b>			
Time	1	$\sqrt{5}$	4
Frequency	1	$1/\sqrt{5}$	1/4
Velocity	5	$\sqrt{5}$	1
Acceleration	5	1	4
Stress	25	5	1
Strain	5	5	1
Stiffness	125	12.5	4
<b>Forces</b>			
Inertial forces	625	62.5	16
Gravitational forces	Reduced by 1/5	62.5	Multiply by 4
Spring forces	Modified	62.5	Scaled
Viscous damping forces	Modified	62.5	Not considered

The above scaling strategy results in similarity of vertical dynamics, together with elastic contact, normal and tangential stresses, which in turn allows the lateral dynamics to be accurately represented.

**D. TABULAR COMPARISON OF SCALING STRATEGIES**

To summarise the strategies discussed in the above section, the scaling parameters have been listed in Table 15.2.

**V. SCALING ERRORS**

As discussed, a scaling strategy is selected based on the type of analysis work to be carried out on the rig, this type specific selection of the strategy has to be performed as perfect scaling cannot be achieved. The example below, using the scaling strategy of the Manchester Metropolitan University, illustrates the level of error which can be encountered.

The errors caused by the scaling of a vehicle can be expressed by the following equations, which have been reproduced from Section IV.A.4 for convenience:

*Force terms* (required  $\varphi_F = 5^4$ )

$$2f_{22} \left( \frac{\dot{y}_w}{v} - \psi_w \right) \quad \text{gives a force scaling, } \varphi_F = 5^{3.33} \quad (15.52)$$



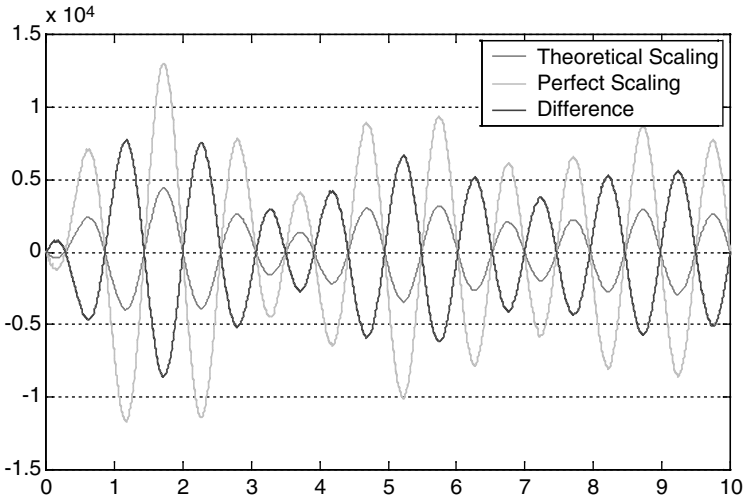


FIGURE 15.6 Lateral force scaling error.

Torque terms (required  $\varphi_T = 5^5$ )

$$2f_{11} \left( \frac{l_0^2 \dot{\psi}_w}{v} + \frac{l_0 \lambda y_w}{r_0} \right) \quad \text{gives a torque scaling, } \varphi_T = 5^{4.33} \quad (15.53)$$

$$2f_{33} \left( \frac{\dot{\psi}_w}{v} \right) \quad \text{gives a torque scaling, } \varphi_T = 5^{5.66} \quad (15.54)$$

The expected level of error from the terms highlighted above can be quantified by performing an analysis of a two degree of freedom wheelset model. The results of the theoretical scaling strategy are plotted against the same model but simulated with a perfect scaling strategy.

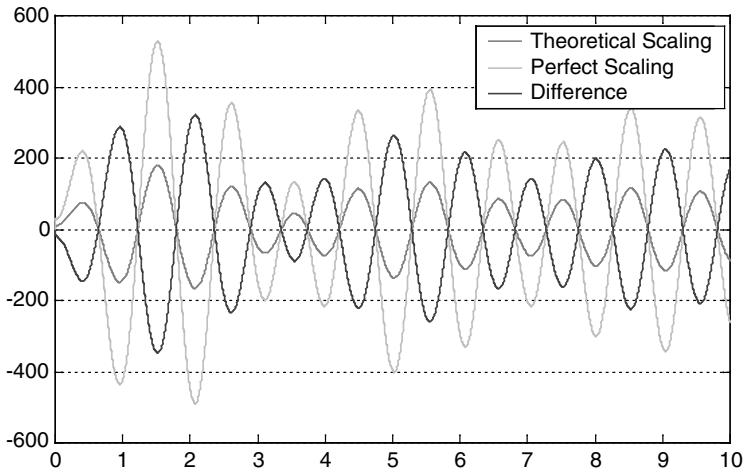


FIGURE 15.7 Wheelset torque scaling error.

The plots shown in Figure 15.6 and Figure 15.7 quantify the errors due to the scaling factors derived in Equation 15.52 to Equation 15.54. The plots were produced with a two degree of freedom model, excited with a sinusoidal disturbance, with an amplitude of 0.5 mm and a frequency of  $2\pi$  radians per second. The forward speed of the vehicle was a 1/5 scale speed of 2 or 10 m/sec at full scale. The results have been scaled from the 1/5 scale values of the roller rig, to the full scale, all other terms achieve perfect scaling.

Although perfect scaling has not been achieved for these terms, experimental testing has shown that the adopted scaling strategy gives good agreement with the full scale, particularly when considering stability. For the purpose of relative studies between vehicles on the roller rig, the error in creep forces illustrated in Figure 15.6 and Figure 15.7 are not of great significance but modifications to the scaling method, to reduce this error, may be required if absolute values between scaled and full-size creep forces, were required.

## VI. CONCLUSIONS

This chapter has discussed the history and application of scaled roller rigs, their uses and has described the construction and operation of three scaled roller rigs from Manchester Metropolitan University, INRETS, and DLR. The three scaling strategies of the above institutions have been described in detail and the differences tabulated.

It is important when designing a new roller rig to first consider the primary use of the rig as this will help form the basis of the scaling strategy. The scaling strategy is the most important aspect of the rig development as it will ensure that the measured parameters are correctly related and obey laws of similarity.

In summary, the scaling strategy of the rig at MMU, Manchester, was developed using a comparison of the linearised differential equations for the scale and full size, the main purpose of the rig being the study of vehicle stability and general dynamic behaviour. Frequency was preserved at 1:1 for this type of analysis. The scaling method for the rig at DLR, Oberpfaffenhofen, was derived from a study of the full set of nonlinear equations of motion to give precise results for study of limit cycle behaviour and early validation of the dynamic multibody simulation software, SIMPACK. The large single wheel rig at INRETS in Grenoble, allows suspension parameters to be evaluated and the almost exact treatment of the contact conditions, allowed by the very large radius of roller. The rig has been used extensively for the validation of Kalker's theory and development of in-house contact mechanics software.

As has been detailed in Chapter 14, there are errors inherent in roller rig testing and these of course apply to scaled rigs. It must be realised that scaled rigs also have additional errors introduced by the scaling strategy, as perfect scaling for all parameters cannot be achieved. An example of possible scaling errors is given and the errors analysed using a typical two degree of freedom wheelset model. This analysis illustrates the importance of selecting a scaling strategy which suits the desired use of the rig.

It is sensible with a scaled roller to use the largest possible roller diameter irrespective of the scale defined for the bogie, as this will preserve the contact conditions with respect to running on conventional track. Results have been presented in the previous chapter as to the influence of roller diameter on various parameters, and these should be considered at the design stage of a scaled rig.

## ACKNOWLEDGMENTS

The author wishes to acknowledge the contribution of Alfred Jaschinski and Hugues Chollet, whose work has been presented, in part, under information submitted regarding the DLR and INRETS scaled roller rigs, respectively.<sup>20</sup>

## REFERENCES

1. Reynolds, O., An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinous, and the law of resistance in parallel channels, *Phil. Trans. R. Soc.*, 174, 935–953, 1883.
2. Reynolds, O., On the dynamical theory of incompressible viscous fluids and the determination of the criterion, *Phil. Trans. R. Soc.*, 186, 123–164, 1895.
3. Sweet, L. M., Sivak, J. A., and Putman, W. F., Non-linear wheelset forces in flange contact, part I: steady state analysis and numerical results, part 2 measurement using dynamically scaled models, *J. Dyn. Syst., Meas. Control*, 101, 238–255, 1979.
4. Sweet, L. M., Karmel, A., and Fairley, S. R., Derailment mechanics and safety criteria for complete railway vehicle trucks, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Seventh IAVSD Symposium, Cambridge, UK, September 1981*, Wickens, A., Ed., Swets & Zeitlinger, Lisse, 1982.
5. Cox, M. and Nicolin, H., Untersuchung des Schwingungsverhaltens von Schienenfahrzeugen mit Hilfe des Modellprüfstands am Institut für FSrdertechnik und Schienenfahrzeuge der RWTH Aachen, *Leichtbau der Verkehrsfahrzeuge*, 23(4), 91–95, 1979.
6. Jochim, M., Konstruktion eines Versuchsdrehgestells. Term study, Lehrstuhl B für Mechanik, TU-München und DLR, Oberpfaffenhofen, 1984.
7. Jochim, M., Analyse der Dynamik eines Schienenfahrzeuges. Diploma thesis, Lehrstuhl B für Mechanik, TU-München und DLR, Oberpfaffenhofen, 1987.
8. Jaschinski, A., On the application of similarity laws to a scaled railway bogie model. Dissertation, TU-Delft, 1990 and DLR-FB 90-06, Oberpfaffenhofen, 1990.
9. Heliot, C., Small-scale test method for railway dynamics, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of Ninth IAVSD Symposium, Linköping, Sweden, June 1985*, Nordstrom, C., Ed., Swets & Zeitlinger, Lisse, 1986.
10. Iwnicki, S. D. and Shen, Z. Y., *Collaborative Railway Roller Rig Project, Proceedings of SEFI World Conference on Engineering Education*, Portsmouth, September, 1992.
11. Iwnicki, S. D. and Wickens, A. H., Validation of a MATLAB railway vehicle simulation using a scale roller rig, *Veh. Syst. Dyn.*, 30(3), 257–270, 1998.
12. Allen, P. D., Error Quantification of a Scaled Roller Rig, Doctoral thesis, Manchester Metropolitan University, 2001.
13. Jaschinski, A. and Netter, H., Non-linear dynamical investigations by using simplified wheelset models, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 12th IAVSD Symposium, Lyon, France, August 26–30, 1991*, Sauvage, G., Ed., Swets & Zeitlinger, Amsterdam/Lisse, 1992.
14. Meinke, P. and Mauer, L., Koppelrahmen-Laufdrehgestell für ICE-Mittelwagen, VDI-Berichte, No. 634, pp. 203–219, 1987.
15. Chollet, H., Essais en similitude à l'échelle 1/4 de bogies de wagons de la famille Y25 INRETS-Report No. 78, 1988.
16. Mauer, L. and Meinke, P., Requirements of future high-speed running gears. RTR Special Railway Technical Review, September 1993.
17. Jaschinski, A., Grupp, F., and Netter, H., Parameter identification and experimental investigations of unconventional railway wheelset designs on a scaled roller rig, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 14th IAVSD Symposium, Ann Arbor, Michigan, 1995*, Segel, L., Ed., Swets & Zeitlinger, Lisse, 1996.
18. Wickens, A. H., The dynamics of railway vehicles on straight track: fundamental considerations of lateral stability, *Proc. I. Mech. Eng.*, 180, 29–44, 1965.
19. Matsudaira, T., Matsui, N., Arai, S., and Yokose, K., Problems on hunting of railway vehicle on test stand, *Trans. A.S.M.E.J. Eng. Ind.*, 91(3), 879–885, 1969, ser. B.
20. Jaschinski, A., Chollet, H., Iwnicki, S. D., Wickens, A. H., and Von Würzen, J., The application of roller rigs to railway vehicle dynamics, *Veh. Syst. Dyn.*, 31, 345–392, 1999.
21. Illingworth, R., Railway wheelset lateral excitation by track irregularities, In *The Dynamics of Vehicles on Roads and Tracks, Proceedings of 5th VSD-2nd IUTAM Symposium, Vienna, Austria, September 19–23, 1977*, Slibar, A., Ed., Swets & Zeitlinger, pp. 450–458, 1978.